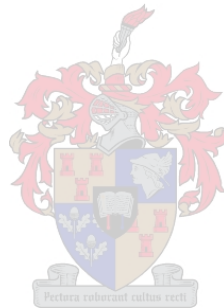


**DISCRIMINATING BETWEEN FOREST PLANTATION GENERA  
USING REMOTE SENSING AND MACHINE LEARNING  
ALGORITHMS**

By CALEY HIGGS

*A thesis presented in partial fulfilment of the requirements for the degree of Master  
of Science in the Faculty of Science at Stellenbosch University.*



Supervisor: Prof A van Niekerk

December 2021



## SUMMARY

Forest inventories are constructed on a compartmental level and contain information such as forest age, species/genus, location, and extent. An up-to-date forest inventory is critical for monitoring harvests, assessing the production of timber, planning, maximising production, assessing water use, and assessing timber quality. On a national scale, forest inventories are used for monitoring the impact forests have on the climate and stream flow, assessing the contribution forests have on alleviating poverty, monitoring forest trends, and supporting policy and trade decisions. Conventional methods for obtaining forest inventory information, such as plantation genus/species, is done in-field, which is time-consuming and costly. Remote sensing is a more efficient way to capture forest genus information. Very high-resolution, hyperspectral, and unmanned aerial vehicle (UAV) imagery have been shown to contain suitable spectral and spatial information for machine learning algorithms to differentiate between forest species. However, such data requires extensive processing and is expensive to acquire, making it unsuitable for mapping over larger areas. High-resolution imagery, such as Sentinel-2, combined with textural measures and vegetation indices as features in machine learning algorithms, have shown potential to differentiate between spectrally similar classes. However, it is not known what impact training sample configuration and size have on classification accuracies when classifying acacia, eucalyptus, and pinus (pine) genera. It is also not known whether signature extension is a viable method for reducing the time and effort spent on obtaining in situ training data when mapping forest plantations over a large and complex area.

This research set out two main experiments. The first experiment evaluated the impact of using an even, uneven, or an area-proportionate training sample configuration and size in a random forest machine learning model for classifying acacia, eucalyptus, and pine compartments. It was found that the study area that contained an uneven area planted with acacia, eucalyptus, and pine trees was classified more accurately using a balanced training sample configuration, compared to using an unbalanced and area-proportionate training sample configuration. It was also found that a saturation point exists where adding more training samples adds little value to the overall accuracy (OA). The saturation point was found to be  $\sim 57n$ , where  $n$  is the number of features used in the classification.

The second set of experiments was set out to test the viability of training data signature extension for constructing random forest machine learning models to differentiate between acacia, eucalyptus, and pine trees using Sentinel-2 imagery as input. The study area was split into 19 Sentinel-2 tiles spanning the Mpumalanga, KwaZulu-Natal, Eastern Cape, and Western Cape

provinces. Three separate random forest models were built using training data collected in one tile located in Mpumalanga, one tile located in KwaZulu-Natal, and one tile located in the Eastern Cape. A fourth model was built using training data from all three source tiles. The four models were applied to all 19 Sentinel-2 tiles to map forest plantation genera. The results show that a ~70% OA can be achieved if the training data is collected in areas with similar climates (rainfall seasonality) to the areas that are being mapped. In addition, it was found that signature extension distance (i.e. distance between the training data and the area being classified) should not exceed 500 km.

## **KEYWORDS**

Remote sensing, forest mapping, training sample configuration, training sample size, signature extension, random forest

## OPSOMMING

Bosplantasie inventarisering word op 'n kompartementele vlak saamgestel en bevat inligting soos die ouderdom, spesie/genus, ligging en omvang van die plantasie. 'n Bygewerkte bosplantasie inventaris is van kritieke belang vir die monitering van oeste, die assessering van houtproduksie, beplanning, maksimalisering van produksie en die assessering van watergebruik en houtgehalte. Op nasionale skaal word bosplantasie inventarisering gebruik om die impak wat bosbou op die klimaat en stroomvloei het te monitor, die bydraes wat bosbou maak om armoede te verlig te assesser, die tendense in bosbou te monitor en beleids- en handelsbesluite te ondersteun. Konvensionele metodes om bosinventarisering, soos plantasie-genus/spesie, te bekom word in die veld gedoen, wat tydrowend en duur is. Afstandswaarneming is 'n doeltreffender manier om boom-genus-inligting vas te lê. Daar is getoon dat baie-hoë-resolusie- en hiperspektrale beelde, asook beelde geneem uit onbemande lugvoertuie, geskikte spektrale en ruimtelike inligting bevat om masjienleer-algoritmes in staat te stel om tussen boomspesies te onderskei. Sodanige data verg egter omvattende verwerking en is duur om te bekom, wat dit ongeskik maak om groot gebiede te karteer. Hoë-resolusiebeelde, soos Sentinel-2, gekombineer met tekstuurmaatstawwe en plantegroei-indekse as veranderlikes in masjienleer-algoritmes, toon potensiaal om tussen klasse met soortgelyke spektrale eienskappe te kan onderskei. Dit is egter nie bekend hoe opleidingsdata konfigurasie en grootte die akkuraatheid van akasia, bloekom en pinus (denne) genera klassifikasies sal beïnvloed nie. Dit is ook nie bekend of klassifiseerder-uitbreiding 'n lewensvatbare metode is om die tyd en moeite benodig om opleidingsdata in situ te bekom, te verminder wanneer bosplantasies oor 'n groot gebied gekarteer word nie.

Hierdie navorsing het twee hoofeksperimente uiteengesit. Die eerste eksperiment het die impak van die gebruik van 'n gelyke, ongelyke of area-proporsionele opleidingmonsteropstelling en -grootte in 'n ewekansige-woud-masjienleermodel vir die klassifikasie van akasia-, bloekom- en denneplantasies geëvalueer. Meer akkurate resultate is vir die studiegebied wat 'n ongelyke area met akasia, bloekom en dennebome bevat behaal wanneer 'n gebalanseerde opleidingmonsteropstelling gebruik is. Daar is ook gevind dat 'n versadigingspunt bestaan waar die toevoeging van meer opleidingmonsters min waarde tot die algehele akkuraatheid (AA) toevoeg. Die versadigingspunt is  $\sim 57n$ , waar  $n$  die aantal veranderlikes wat in die klassifikasie gebruik word verteenwoordig.

Die tweede stel eksperimente is uitgevoer om die lewensvatbaarheid van klassifikasie-uitbreiding te toets. Ewekansige-woud-masjienleer is aangewend om tussen akasia, bloekom en dennebome, met Sentinel-2-beelde as toevoer, te onderskei. Die studiegebied is verdeel in 19 Sentinel-2-teëls

wat oor die Mpumalanga, KwaZulu-Natal, Oos-Kaap en Wes-Kaap provinsies strek. Drie afsonderlike ewekansige-woud-modelle is met behulp van opleidingsdata, wat onderskeidelik in een teël in Mpumalanga, een teël in KwaZulu-Natal en een teël in die Oos-Kaap ingesamel is, gebou. 'n Vierde model is met behulp van opleidingsdata van al drie bronteëls gebou. Die vier modelle is op al 19 Sentinel-2-teëls toegepas om plantasie genera te karteer. Die resultate toon dat 'n ~ 70% AA behaal kan word indien die opleidingsdata in gebiede met soortgelyke klimaat (reënval seisoenaliteit) as die areas wat gekarteer word, ingewin word. Daarbenewens is gevind dat die afstand van klassifiseerder-uitbreiding (d.w.s. afstand tussen die opleidingsdata en die area wat geklassifiseer word) nie 500 km moet oorskry nie.

### **SLEUTELWOORDE**

Afstandswaarneming, bosplantasie kartering, opleidingmonsterskema, klassifiseerder-uitbreiding, ewekansige woud.

## ACKNOWLEDGEMENTS

I sincerely thank:

- Prof van Niekerk for the constant guidance, advice, motivation, and always making time in his busy schedule for me;
- Dr Munch for always having her door open for any questions I had and any advice that I needed;
- Liezl Vermeulen for helping me with any Google Earth Engine issues that I had;
- Helene van Niekerk for all her language editing and for translating my summary into Afrikaans;
- The Department of Geography & Environmental Studies, particularly the Geoinformatics staff, for their valuable feedback at the scheduled contact sessions organised by Deirdre;
- The Water Research Commission for funding this project. Without them none of this would have been possible.
- My family for their emotional and financial support; and finally
- My friends for their emotional support.

This work forms part of a larger project titled “The application of national scale remotely sensed evapotranspiration (ET) estimates to quantify water use and differences between plantations in commercial forestry regions of South Africa” which was initiated and funded by the Water Research Commission (WRC) of South Africa (contract number K5/2966//4). More information about this project is available at [www.wrc.org.za](http://www.wrc.org.za).

## CONTENTS

<b>DECLARATION .....</b>	<b>ii</b>
<b>SUMMARY .....</b>	<b>iii</b>
<b>OPSOMMING .....</b>	<b>v</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>vii</b>
<b>CONTENTS .....</b>	<b>viii</b>
<b>TABLES .....</b>	<b>xi</b>
<b>FIGURES .....</b>	<b>xii</b>
<b>APPENDICES .....</b>	<b>xiii</b>
<b>ACRONYMS AND ABBREVIATIONS.....</b>	<b>xiv</b>
<b>CHAPTER 1: INTRODUCTION .....</b>	<b>1</b>
<b>1.1 PLANTATION FOREST CHARACTERISATION .....</b>	<b>2</b>
<b>1.2 REMOTE SENSING OF FORESTRY .....</b>	<b>2</b>
<b>1.2.1 Common remote sensing application in forestry.....</b>	<b>2</b>
<b>1.2.2 Regional forest mapping.....</b>	<b>3</b>
<b>1.2.3 In situ data for classifying forest plantations.....</b>	<b>4</b>
<b>1.3 PROBLEM STATEMENT.....</b>	<b>6</b>
<b>1.4 AIM AND OBJECTIVES.....</b>	<b>6</b>
<b>1.5 RESEARCH METHODOLOGY.....</b>	<b>7</b>
<b>CHAPTER 2: LITERATURE REVIEW .....</b>	<b>9</b>
<b>2.1 EARTH OBSERVATION .....</b>	<b>9</b>
<b>2.1.1 Principles of remote sensing.....</b>	<b>9</b>
<b>2.1.2 Image analysis paradigms .....</b>	<b>14</b>
<b>2.1.3 Classification.....</b>	<b>15</b>
2.1.3.1 Unsupervised.....	15
2.1.3.2 Supervised .....	16
2.1.3.3 Training samples and dimensionality.....	18
2.1.3.4 Knowledge-based .....	20
<b>2.1.4 Classification accuracy assessment.....</b>	<b>21</b>
<b>2.2 REMOTE SENSING IN FORESTRY .....</b>	<b>22</b>
<b>2.2.1 Mapping forests and other land covers .....</b>	<b>22</b>
<b>2.2.2 Forest type classification.....</b>	<b>25</b>



2.2.3	Genera/species classification .....	27
2.3	SUMMARY AND EVALUATION.....	31
<b>CHAPTER 3: IMPACT OF TRAINING SET CONFIGURATIONS FOR DIFFERENTIATING BETWEEN PLANTATION FOREST GENERA WITH SENTINEL-2 IMAGERY AND MACHINE LEARNING.....</b>		
3.1	ABSTRACT .....	34
3.2	INTRODUCTION.....	34
3.3	METHODS AND MATERIALS.....	38
3.3.1	Study areas.....	38
3.3.2	Data collection and preparation.....	39
3.3.2.1	Imagery.....	39
3.3.2.2	In situ data .....	40
3.3.3	Experimental design.....	40
3.3.4	Accuracy assessment .....	42
3.3.5	Spectral analysis .....	43
3.4	RESULTS.....	43
3.4.1	Genus spectral profiles.....	43
3.4.2	Classification.....	44
3.5	DISCUSSION .....	49
3.6	CONCLUSION.....	52
<b>CHAPTER 4: SIGNATURE EXTENSION AS A MACHINE LEARNING STRATEGY FOR MAPPING PLANTATION FOREST GENERA WITH SENTINEL-2 IMAGERY.....</b>		
4.1	ABSTRACT .....	53
4.2	INTRODUCTION.....	53
4.3	METHODS AND MATERIALS.....	57
4.3.1	Study area .....	57
4.3.2	Data collection and preparation.....	59
4.3.2.1	Imagery.....	59
4.3.2.2	In situ data .....	60
4.3.3	Experimental design.....	60
4.3.4	Accuracy assessment .....	61
4.3.5	Spectral analysis .....	62
4.3.6	Variable drivers.....	62

<b>4.4</b>	<b>RESULTS.....</b>	<b>63</b>
<b>4.4.1</b>	<b>Spectral profiles.....</b>	<b>63</b>
<b>4.4.2</b>	<b>Classification.....</b>	<b>64</b>
<b>4.5</b>	<b>Factors influencing signature extension.....</b>	<b>66</b>
<b>4.6</b>	<b>DISCUSSION .....</b>	<b>67</b>
<b>4.7</b>	<b>CONCLUSION.....</b>	<b>70</b>
<b>CHAPTER 5: DISCUSSION AND CONCLUSION.....</b>		<b>71</b>
<b>5.1</b>	<b>REVISITING THE AIMS AND OBJECTIVES.....</b>	<b>71</b>
<b>5.2</b>	<b>SYNTHESIS .....</b>	<b>72</b>
<b>5.3</b>	<b>VALUE OF RESEARCH.....</b>	<b>73</b>
<b>5.4</b>	<b>STUDY LIMITATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH.....</b>	<b>74</b>
<b>5.5</b>	<b>CONCLUSION.....</b>	<b>75</b>
<b>REFERENCES .....</b>		<b>78</b>
<b>APPENDICES .....</b>		<b>93</b>

## TABLES

Table 2-1 Selection of popular EO satellites.....	11
Table 2-2 Summary of forest and land cover-type mapping studies.....	24
Table 2-3 Summary of forest type mapping studies .....	26
Table 2-4 Summary forest genera and species mapping studies.....	29
Table 3-1 Features (bands, indices and textural measures) used as input to the classifications generated from a composite of individual Sentinel-2 images taken from 2019-06-30 to 2020-06-30 .....	39
Table 3-2 Summary of in situ data (forest compartment information) collated, including tree genus, age (mean and standard deviation) and planted area per study area .....	40
Table 3-3 Summary of Experiments A to H .....	41
Table 3-4 A summary table showing the overall accuracy, the standard deviation of the overall accuracy, kappa statistic, the standard deviation of the kappa statistic, consumer's and user's accuracy and the maximum OA and KS of the 100 iterations per sample size of experiments A to G conducted on Study Area 1 (WC) and Study Area 2 (KZN) .....	45
Table 4-1 Description of the Köppen-Geiger zones in South Africa.....	59
Table 4-2 Summary of in situ data (forest compartment information) collated, including tree genus and planted area per province .....	60
Table 4-3 Summary of in situ data (forest compartment information) collated, including tree genus and planted area per block .....	61
Table 4-4 The overall accuracy, kappa statistic, user's accuracy (UA), and producer's accuracy (PA) of each block for all experiments .....	65
Table 4-5 Intra-class spectral variability among tiles, as quantified by Jefferies-Matusita distance .....	67

## FIGURES

Figure 1-1 Research design for forest plantation mapping on a genus level and determining forest plantation ages using satellite imagery .....	8
Figure 2-1 EMS and its wavelengths .....	10
Figure 2-2 Spectral signatures for soil, water, and green vegetation .....	12
Figure 2-3 Confusion matrix showing commission and omission errors.....	21
Figure 3-1 Distribution of South African forest plantations (DAFF 2008) .....	37
Figure 3-2 Locations of Study Area 1 (a), along the southern coast of the Western Cape (WC) province, and Study Area 2 (b) in the KwaZulu-Natal (KZN) midlands.....	38
Figure 3-3 Genus distribution of in situ data in Study Area 1 (a) and Study Area 2 (b) .....	40
Figure 3-4 Average spectral signatures of the training samples, the standard deviation of the spectral signatures (shown as error bars), and the J-M separability score for all the training samples for (a) Study Area 1 (WC) and (b) Study Area 2 (KZN) .....	43
Figure 3-5 Trends of the overall accuracies, standard deviation of the overall accuracies for the 100 iterations, kappa statistics, standard deviation of the kappa statistics for the 100 iterations, user's accuracies, and producer's accuracies of experiments A to D for Study Area 1 and Study Area 2 .....	46
Figure 3-6 Trends of the overall accuracies, standard deviation of the overall accuracies for the 100 iterations, kappa statistics, standard deviation of the kappa statistics for the 100 iterations, user's accuracies, and producer's accuracies of experiments E to H on Study Area 1 and Study Area 2 .....	47
Figure 4-1 Location of compartments used in the study .....	57
Figure 4-2 Rainfall seasonality in South Africa.....	58
Figure 4-3 Köppen-Geiger climate zones of South Africa .....	58
Figure 4-4 Average spectral signatures for the training samples and J-M separability score for Tile 4 (a), Tile 10 (b), Tile 17 (c), and all source tiles in combination (d).....	64
Figure 4-5 Overall accuracy vs distance from the source tile for exp 1 (Tile 4) (a). exp 2 (Tile 10) (b), and exp 3 (Tile 17) (c), overall accuracy vs rainfall seasonality index for exp 1 (d), exp 2 (e), and exp 3 (f), and overall accuracy vs temperature index for exp 1(g), exp 2 (h), exp 3 (i) .....	66

## APPENDICES

- Appendix A: Indices of disagreement for Chapter 3 experiments at the initial iteration and then at every tenth iteration sample increase
- Appendix B: Binary classification (eucalyptus and pine) accuracies using signature extension with training data obtained from Tile 17

## ACRONYMS AND ABBREVIATIONS

ANN	Artificial neural network
AVHRR	Advanced Very High-Resolution Radiometer
CART	Classification and regression tree
CASI	Compact Airborne Spectrographic Imager
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station Data
CNN	Convolutional neural networks
DT	Decision tree
EMR	Electromagnetic radiation
EMS	Electromagnetic spectrum
EO	Earth observation
EVI	Enhanced vegetation index
GEDI	Global Ecosystem Dynamics Investigation
GEE	Google Earth Engine
GIAS	Geoscience Laser Altimeter System
GLCM	Grey level co-occurrence matrix
ha	Hectare
ISODATA	Iterative self-organising data analysis technique
J-M	Jefferies-Matusita
KNN	K-nearest neighbour
KS	Kappa statistic
KZN	KwaZulu-Natal
LCLU	Land cover and land use
LDA	Linear discriminative analysis
LiDAR	Light Detecting and Ranging
MDS	Multidimensional scaling
MERIS	Medium Resolution Imaging Spectrometer
MLC	Maximum likelihood classifier
MNF	Minimum noise fraction
MODIS	Medium Resolution Imaging Spectroradiometer
NDVI	Normalised difference vegetation index
NFI	National forestry inventory
NIR	Near-infrared

NN	Neural networks
OA	Overall accuracy
OBIA	Object-based image analysis
OOB	Out of bag
PA	Producer's accuracy
PALSAR	Phased Array type L-band Synthetic Aperture Radar
PCA	Principle component analysis
RF	Random forest
RGB	Red-green-blue
RMSE	Root mean square error
RS	Remote sensing
RT	Regression tree
RVI	Regular vegetation index
SAM	Spectral angle mapper
SAR	Synthetic aperture radio detecting and ranging
SAVI	Soil adjusted vegetation index
SfM	Structure for Motion
SPOT	Satellite pour l'observation de la terre, lit
SVM	Support vector machine
TM	Thematic mapper
UA	User's accuracy
UAV	Unmanned aerial vehicle
USA	United States of America
VHR	Very high-resolution
VI	Vegetation index
WC	Western Cape

## CHAPTER 1: INTRODUCTION

The forestry sector contributes 1% to South Africa's gross domestic product and offers employment opportunities to many rural communities, with about 165 900 workers being employed within those communities (Tibane & Vermeulen 2014).

The materials produced by plantations have different uses, such as timber and paper products (Mandy & Steve 2015), sustainable energy (Carle, Del Lungo & Varmola 2003) and tannin extracts for leather tanning (Mandy & Steve 2015). Many forest products are exported (Food and Agriculture Organization 2015), and the demand for these goods increases as uses of forestry products and populations grow.

The South African forestry sector is experiencing afforestation constraints as the licensing process has become cumbersome as a result of the National Water Act of 1998. The Act was put in place to ensure a sustainable distribution of water to all parties (Gush et al. 2002). The implementation of the Act led to a decrease of 80 000 hectares (ha) of forest plantations (Food and Agriculture Organization 2015) as only a certain amount of forestry is permitted per catchment.

National forest inventories have many uses, for example, to assess the quality and production of timber (Brown & Ball 2000), aid in planning, sustainably manage land use, maximise production, assess water use (Food and Agriculture Organization of the United Nations 2015), monitor timber harvests and rotations, and assist in decision-making. Information for silviculture, fixing rotation age, site management, and timber harvests are derived from inventories and are used in decision-making (Mati & Dawaki 2015). Inventories include data such as location, genus, age, species, yield, and water use. However, the existing South African forest inventories are incomplete and outdated (DAFF 2008). Previously, the collection of inventory data was dependant on funding and policy demand. Currently, forests are monitored every three years according to national indicators and criteria such as; the development and maintenance of forest resources, biological diversity in forests, the health and vitality of forests, the productive functions of forests, the protective and environmental functions of forests, and the social functions of forests (Government of South Africa 1998). Inventory data are collected through questionnaires that assess the use of forest resources on communal land and through licences. Private companies own 31% of commercial forests in South Africa, making them an important contributor to national forest inventories; however, the data comes at a cost (DWAF 2008). The current methods used to collect inventory data do not provide a comprehensive and up-to-date overview of forestry activities in South Africa. It is therefore important to develop a methodology whereby forest plantation genera can be mapped at



a national scale to frequently update the South African national forest inventory database (DWAF 2008).

## **1.1 PLANTATION FOREST CHARACTERISATION**

In South Africa, commercial forest species are evergreen trees (Mead 2013) categorised into three genera (Mandy & Steve 2015), of which 57% are pine trees, 35% are eucalyptus trees, and 8% are acacia (wattle) trees (Dye & Versfeld 2007).

The *Pinus* genus (henceforth referred to as pines) contains many different hybrids and species, of which several are grown in South Africa (Mandy & Steve 2015) for wood and paper products. Pines have needle-like leaves and grow for 25 to 30 years, reaching a height of about 30 m, before they are harvested (Mead 2013).

Eucalyptus trees are fast-growing (Albaugh, Dye & King 2013), which allows harvesting from when they are seven to ten years of age (Pillay 2012). The harvested trees are used for solid wood products, pulp, and paper (Mandy & Steve 2015).

Acacia trees, also known as wattle trees, are grown for their bark, which is used for leather tanning (Mandy & Steve 2015). Acacia trees reach maturity after one to five years (Wilson et al. 2011) and a height of 15 m (De Beer 1986).

## **1.2 REMOTE SENSING OF FORESTRY**

Remote sensing (RS) is a technology that uses a sensor to record reflected or emitted electromagnetic energy at a distance from the region of interest (Campbell & Wynne 2013). The recorded radiance is used to derive information about the earth's surface.

RS has been used in forestry to derive information about forest structure, crown closure estimates, health statuses, age estimation, as well as genus and species mapping.

### **1.2.1 Common remote sensing application in forestry**

The structural characteristics of forests such as height, volume, and basal area (Tang & Shao 2015) have been successfully captured using RS technologies such as photogrammetry (Campbell & Wynne 2013), light detecting and ranging (LiDAR) (Holmgren & Thuresson 1998), and structure from motion (SfM) (Tang & Shao 2015). It has been shown that using geoinformation systems for capturing structure data is faster than using conventional methods (Budei et al. 2018).

RS data such as aerial imagery (Campbell & Wynne 2013) or very high-resolution (VHR) satellite imagery (Tang & Shao 2015) has been used to accurately estimate the crown closure of plantations.

Data captured through RS can also be transformed into vegetation indices (VIs) to determine the health status of trees. This is possible as the chlorophyll in plants absorbs red light and strongly reflects radiation in the near-infrared (NIR) region of the electromagnetic spectrum (EMS). Researchers have successfully identified trees that are unhealthy using this method, and the methodology has directed foresters to apply fertilisers to specific problem trees (Tang & Shao 2015).

### **1.2.2 Regional forest mapping**

Forest mapping has been conducted at different scales. The scale of a forest mapping exercise is dependent on how big the study area is and its heterogeneity. The extent of a study can be global, continental, national, regional, or local (Herod 2016). The extent of regional studies varies, but, in general, smaller areas with high heterogeneity can be considered as regional (Muller et al. 2020).

A large volume of work has been done on mapping forests at global and continental scales. McRoberts et al. (2002) classified four states in America into forest and non-forest areas using stratified national land cover data for training, while Hagner & Reese (2007) and Tomppo et al. (2008) classified forest types in Sweden and Finland. A land cover classification was conducted on a global level using Medium Resolution Imaging Spectrometer (MERIS) fine resolution (300 m) data (Leroy et al. 2007). DeFries et al. (2000), Hansen et al. (2003) and Hansen et al. (2005) estimated the global percentage of tree cover using different machine learning algorithms and Medium Resolution Imaging Spectroradiometer (MODIS) imagery. However, the authors noted that it is difficult to quantify the accuracy of classifications at a global scale due to the lack of reference data. This is also a problem when it comes to training machine learning algorithms. Therefore, studies often have to rely on national inventory data, which is often incomplete, inaccurate, and/or out of date.

Research conducted in smaller areas have successfully discriminated between plantation and natural vegetation (Nery et al. 2019; Lück 2018), conifers, deciduous and mixed forests (Nangendo, Skidmore & Van Oosten 2007), and other land cover types (Baatuuwie & Leeuwen 2011) using medium resolution (15 m – 30 m) multispectral imagery. Classification approaches that use medium resolution multispectral imagery are generally unable to differentiate between spectrally similar features, such as plantation genera and species, but have been shown to successfully differentiate between plantation types, such as commercial forest plantations and natural forests, as the spectral properties of such classes are more dissimilar than plantation genera. The use of VHR multispectral imagery in classifications is useful in differentiating between forest species (Cho, Malahlela & Ramoelo 2015; Franklin & Ahmed 2018; Franklin, Ahmed & Williams 2017; Immitzer, Atzberger & Koukal 2012; Ke, Quackenbush & Im 2010; Pu & Landry 2012;

Wagner et al. 2019; Xie et al. 2019). VHR hyperspectral data have been used to successfully classify exotic forests (Peerbhay, Mutanga & Ismail 2013), coniferous forest species (Buddenbaum, Schlerf & Hill 2005), eucalyptus and pine plantations (Van Aardt & Norris-Rogers 2008), swamp tree species (Adam et al. 2012), and forest plantation species (Fagan et al. 2015; Voss & Sugumaran 2008), but—in addition to such data being very expensive (even more so than VHR multispectral imagery)—it requires extensive processing as it has high data dimensionality. Consequently, the use of VHR multispectral and hyperspectral data is not a viable solution for regional applications.

Active RS data have been successfully used to map forest plantations. SAR data are unaffected by atmospheric conditions and can differentiate between trees and other land cover types in classification algorithms. Specifically, synthetic aperture RADAR (SAR) derived metrics have been used in conjunction with multispectral derived metrics in classification algorithms to map forest plantations. SAR derived metrics are often used to differentiate between forest and non-forest areas (Chen et al. 2016; Dong et al. 2012; Dong et al. 2013).

Airborne LiDAR has been used in classification algorithms to map tree species (Budei et al. 2018; Heinzl & Koch 2011; Li, Hu & Noland 2013; Martinuzzi et al. 2013) and deciduous and coniferous trees (Yao, Krzystek & Heurich 2012). LiDAR data are expensive and therefore only commonly used in small areas. Spaceborne LiDAR sensors such as the Geoscience Laser Altimeter System (GLAS) (Lefsky 2010; Simard et al. 2011), IceSat (Xing et al. 2010), and Global Ecosystem Dynamics Investigation (GEDI) (Qi et al. 2019) have been used to capture tree heights over large areas. However, the GLAS satellite only operated discontinuously between 2003 and 2009 (Michez et al. 2016), and GEDI will only be operational for two years, making it unsuitable for operational forest monitoring solutions.

### **1.2.3 In situ data for classifying forest plantations**

Supervised classification requires a sufficient amount of in situ data used for training the model. The accuracy of supervised classification is influenced by the number, distribution, and quality of the training samples (Lu & Weng 2007). Factors such as the spectral variability within and among classes (Lu & Weng 2007; Mather 2004), the number of features (bands) used in the classification (Mather 2004), and the number of classes being classified (Campbell & Wynne 2013) should be considered when collecting training data for a supervised classification.

Forest plantations genera are spectrally similar making it difficult to differentiate between them. Literature has shown that the addition of textural feature and vegetation indices as bands increases the ability of classification algorithms to differentiate between forest plantation genera/species

(Vaglio Laurin et al. 2016). However, increasing the number of features (bands) to the classification often leads to the  $s \ll n$  problem, known as the Hughes effect, where  $s$  is the number of samples and  $n$  is the number of features. The Hughes effect states that as the number features increases the accuracy of the classification improves, but at some point the accuracies will decrease as the training data become sparser compared to the increased feature space (Ma et al. 2013).

The recommended number of training samples required for classification varies among literature. Campbell & Wynne (2013) recommend training area sizes exceeding 100 pixels. Whereas, Mather (2004) recommends obtaining more than  $30n$  samples per category, where  $n$  is the number of input features (bands). Similarly, Belgiu & Dragut (2016) propose collecting enough samples so that  $s > n$ . Congalton & Green (2019) suggests that, when the area being classified is smaller than 1 million hectares, the number of training samples should be about four times larger than the dimensionality of the dataset being classified (i.e.  $s \sim 4n$ ). Thanh Noi & Kappas (2017) showed that the overall accuracy of a classification generally increases as the training set size increases, but that a saturation point exists where the overall accuracy is unaffected with an increase in training set size. Adding training data beyond the saturation point is unnecessary (Foody 2009). Instead, the training data should be selected to maximise the spectral separability between classes by collecting information about the spectral variability between classes (Foody et al. 2006).

There are different views on whether an equal number of training samples are required per class. Congalton & Green (2019) and Colditz (2015) suggests that classes that occupy a larger area require more training samples as they are often more complex and have a large spectral variation. Mellor et al. (2015) showed that complex classes, like forests, can be mapped more accurately by using an unbalanced training data set. In contrast, Dalponte et al. (2013) and Millard & Richardson (2015) showed that when an unbalanced training data set is used, the class with the most training data dominates the classification result. Belgiu & Dragut (2016) found that the user's and producer's accuracies of the scarce classes increase when using an area-proportionate (unbalanced) sample set.

A disadvantage of supervised classifications is that the collection of training data can be time-consuming and costly, especially when mapping at a regional or national scales (Pax-Lenney et al. 2001). Signature extension or generalisation has been suggested to reduce the expense of training data collection. Signature extension is the process whereby a model is trained on one image and applied to other images or scenes (Laborte, Maunahan & Hijmans 2010). Signature extension can be applied across time and distance (Wang, Azzari & Lobell 2019).

### 1.3 PROBLEM STATEMENT

Incomplete and outdated forest plantation inventories are impeding effective forest plantation management. Tree genus is a characteristic of plantations that are critical for national forest planning and inventorying. It is therefore important to evaluate—and, if necessary, adapt—methodologies used to acquire the genus of forest plantations on a national level.

Although some successes have been reported on genera/species classifications using remotely sensed imagery (Budei et al. 2018; Hagner & Reese 2007; Lück 2018; Nery et al. 2019; Peerbhay, Mutanga & Ismail 2014), previous studies have mapped forest plantations on a local scale, and very few attempts have been made to map forest and other land cover types on a continental scale (DeFries et al. 2000; Hansen et al. 2003; Hansen et al. 2005), and forest types on regional and national scales (Franco-Lopez, Ek & Bauer 2001; Hagner & Reese 2007; McRoberts et al. 2002; McRoberts et al. 2007). It is difficult to map forest plantations over large areas as they often occur in widely distributed patches (Geldenhuys & Mucina 2006) and are spectrally similar. VHR multispectral imagery, hyperspectral imagery and LiDAR data have been used to successfully classify forest plantation species; however, such data are expensive, making it unsuitable for mapping over large areas. SAR data have been used to map forests and non-forests (Chen et al. 2016; Dong et al. 2012; Dong et al. 2013), but have not successfully been used to differentiate the forested areas into genera or species. High spatial resolution imagery has been used in classifiers to discriminate between forest types, genera, and species, but no studies have specifically discriminated between pine, acacia, and eucalyptus. Therefore, there is a research gap about the efficiency of RS methods to map and characterise forest plantations on regional or national scales. It is not clear which RS methods and data sources will be most effective for differentiating between the three main plantation forest genera in South Africa (pine, eucalyptus, and acacia) on a national scale. Also, factors such as climatic variations, tree age and densities, as well as different phenological characteristics of species, will likely have a negative effect on classification accuracies (Hansen et al. 2005). Although non-parametric machine learning algorithms have been shown to be effective for complex classification tasks, it is not clear how much training data and what training set configuration would be required to adequately represent such variations, and to what extent signature extension will lead to acceptable classification accuracies.

### 1.4 AIM AND OBJECTIVES

This study aims to evaluate RS and machine learning methodologies whereby forest plantation genera can be mapped at a national scale. Specifically, it aims to assess the impact of different sampling strategies on machine learning accuracies and to investigate the potential of signature extension for reducing training sample collection costs.

The following objectives have been set to achieve the research aim:

1. Carry out a literature review to identify the most appropriate techniques for classifying forest plantation genera and estimating forest plantation age;
2. Collect suitable data on forest genera that can be used to build and assess models;
3. Carry out RS experiments to identify the most effective training sample strategy for discriminating between forest plantation genera over large and diverse areas;
4. Evaluate signature extension (model transfer) as a means to reduce reliance on large sets of in situ genera data for training machine learning algorithms; and
5. Evaluate the findings and make recommendations for implementing RS techniques for carrying out forest inventories on a national scale.

## **1.5 RESEARCH METHODOLOGY**

This research is experimental in nature and will follow a deductive approach. Existing approaches to classify forest plantations will be adapted to characterise forest plantations in South Africa. The study will experiment with different sampling schemes and test the transferability of machine learning models. Primary data in the form of shapefiles containing genus data at compartmental level, obtained from forestry companies, will be used for training and validating the models. The results will be evaluated both quantitatively and qualitatively.

Figure 1-1 shows the research agenda. Chapter 1 outlines the research aim and objectives. Chapter 2 will review the literature and thoroughly evaluate previous studies relevant to this research. It will also provide a background of RS techniques that are frequently applied in forestry.

Chapter 3 will detail the methods and results for mapping acacia, eucalyptus and pine forest plantations using a balanced, unbalanced, and area-proportionate training sample configurations at different sample sizes. This will be followed by Chapter 4, in which the methods and results for evaluating the extent to which signature extension is viable for mapping forest plantation genera will be evaluated. Finally, Chapter 5 will synthesise the findings of the study, highlight the study limitations, and make recommendations future research.

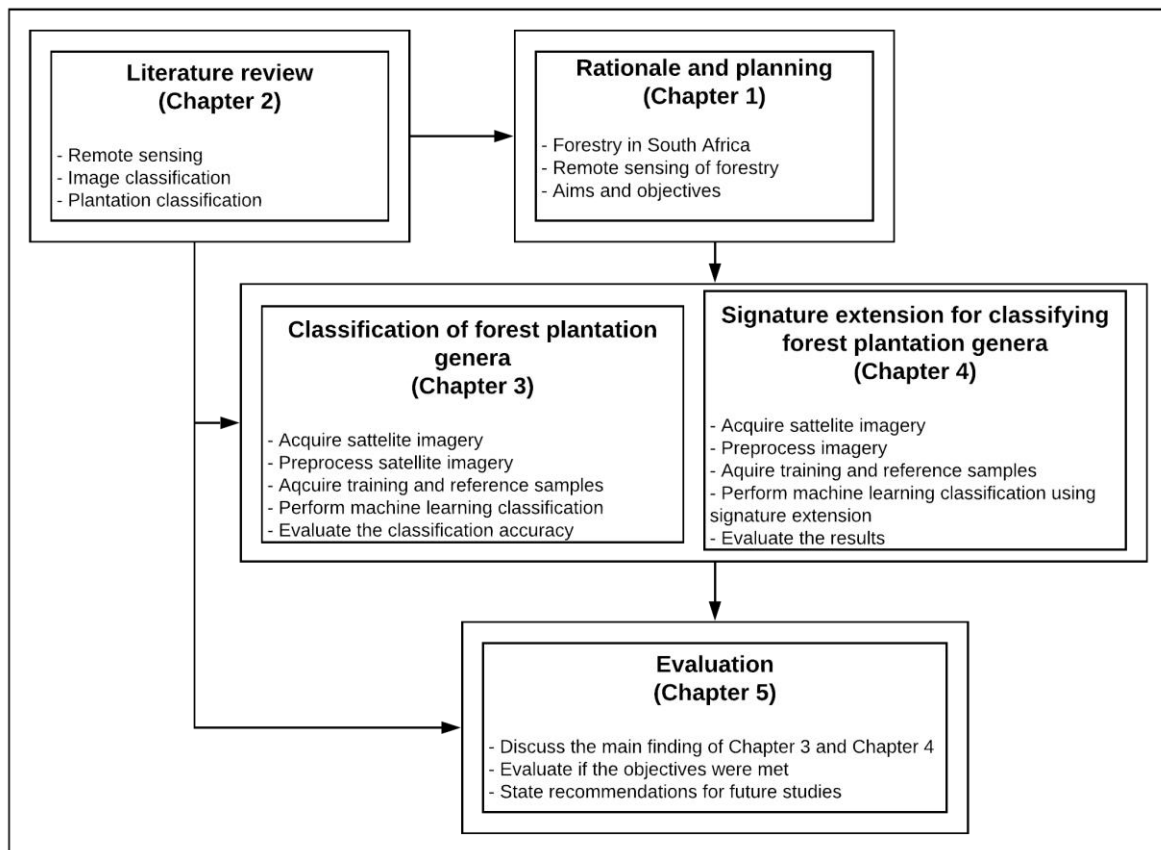


Figure 1-1 Research design for forest plantation mapping on a genus level and determining forest plantation ages using satellite imagery

## **CHAPTER 2: LITERATURE REVIEW**

Discriminating between forest plantation genera using RS can contribute to accurate and up-to-date forest inventory databases. It is therefore important to develop a methodology that accurately and automatically maps forest plantation genera at regional or national scales.

This chapter overviews the principles of RS and image classification to better understand existing RS methodologies used in forestry. Research on forest and other land cover mapping, forest type classification, and genera/species classification will be reviewed to elaborate on the research gaps identified in Chapter 1 and to identify suitable methods to achieve the aim of this study.

### **2.1 EARTH OBSERVATION**

Earth observation (EO) uses RS technologies to gather information about processes that occur on the earth's surface. The principles of RS, image classification paradigms and accuracy assessment will be discussed in the following subsections.

#### **2.1.1 Principles of remote sensing**

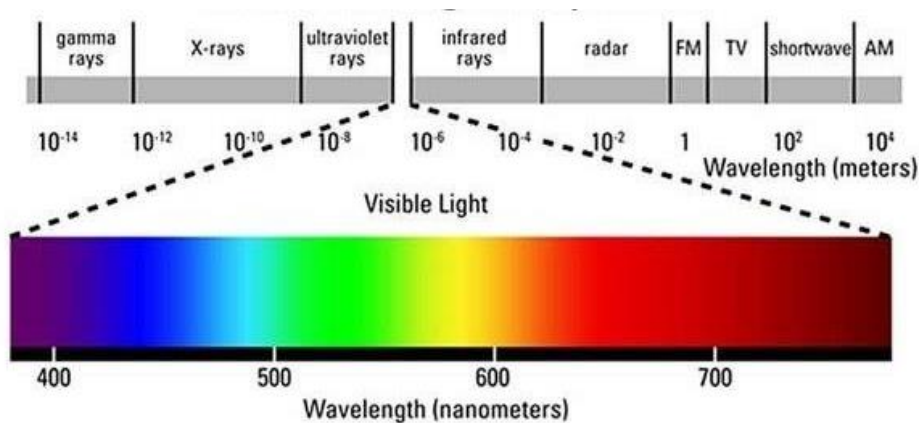
RS is the science of observing the earth's surface through emitted or reflected electromagnetic energy (Lillesand, Kiefer & Chipman 2019) that is captured at a distance by a sensor (Roughgarden, Running & Matson 2010). Spaceborne or airborne vehicles are most commonly used to carry RS instruments, although hand-held and tower-based instruments have also been used (Roughgarden, Running & Matson 2010). The radiation or signal that is reflected/emitted from the objects on the earth's surface is captured by the sensor and used to study those objects (Campbell & Wynne 2013).

RS sensors are either passive or active. Passive RS sensors use the electromagnetic radiation (EMR) generated by the sun (Bangira 2019), which is comprised of the ultraviolet to the infrared portion of the EMS and records the radiation that is emitted or reflected from the earth's surface. In contrast, active RS sensors broadcast a signal and record the strength and the time it takes the signal to return to generate a point cloud of the earth's surface (Bangira 2019). Examples of active RS sensors are SAR and LiDAR.

When EMR interacts with the earth's surface it is either reflected, transmitted or absorbed. The proportions of the EMR that is reflected, transmitted or absorbed varies with wavelength and the object it interacts with (Jackson & Huete 1991). The EMS can be categorised into seven wavelength regions (Figure 2-1) (Murthy 2004). RS sensors often capture different portions



(bands) of the EMR, providing information about the various features on the earth's surface (Shaw & Burke 2003).



Source: Silva et al. (2019)

Figure 2-1 EMS and its wavelengths

Three types of optical sensors exist: panchromatic, multispectral, and hyperspectral (Shaw & Burke 2003). Panchromatic sensors collect wavelengths in one band, while multispectral sensors collect data at a variety of different wavelength ranges in multiple bands (Shaw & Burke 2003). Hyperspectral sensors generate imagery that contains hundreds of bands consisting of data ranging from the visible to thermal infrared regions of the EMS (Shaw & Burke 2003).

Thousands of RS satellites exist and some have been operational for more than 40 years (Zhu et al. 2017). Table 2-1 summarises the characteristics of selected EO satellites.

The resolution of a remotely sensed image determines whether it can accommodate fine spectral, spatial and radiometric characteristics (Lillesand, Kiefer & Chipman 2019). The spectral resolution is the sensor's ability to capture radiation in small wavelength intervals. An image with a high spectral resolution contains many narrow bands from which the spectral responses can be used to better discriminate between spectrally similar objects (Lillesand, Kiefer & Chipman 2019). Temporal resolution refers to the time a satellite takes to revisit and capture the same scene. A satellite with a high temporal resolution improves the quantity and quality of images as the chance of capturing a cloud-free image increases (Lu & Weng 2007). The spatial resolution of an image determines the smallest discernible feature on the map. The higher the spatial resolution, the smaller the pixel, and the more detailed the image becomes (Lillesand, Kiefer & Chipman 2019).

Table 2-1 Selection of popular EO satellites

		Satellite (commission period)	Sensors	Spectral bands	Spatial resolution	Revisit Time	Availability
Low resolution	MODIS	Terra (1999 - ) Aqua (2002 - )	Moderate resolution imaging spectro-radiometer (MODIS)	36	Red, NIR (250 m) Blue, Green, IR (500 m) Thermal (1 km)	2 times, daily	Freely available
	AVHRR	NOAA (1978 - 2009) multiple	Advanced very high resolution radiometer (AVHRR)	6	VIS, NIR, Thermal (1 km)	2 times, daily	Freely available
Medium resolution	Landsat	Landsat - 4 (1992 - 2001)	Multispectral scanner (MSS); Thematic mapper <sup>TM</sup>	4	VIS, NIR, Thermal (68 m by 83 m )	18 days	Freely available
		Landsat - 5 (1985 - 2011 <sup>TM</sup> /2013 (M88))	MSS	4	VIS, NIR, Thermal (68 m by 83 m )	18 days	Freely available
			TM	7	VIR, NIR, Mid IR (30 m), Thermal (120 m)	16 days	Freely available
		Landsat - 7 (1999 - )	Enhanced thematic mapper plus (ETM+)	8	Panchromatic (15 m), VIR, NIR, Mid-IR, SWIR (30 m), Thermal (60m)	14 days	Freely available
		Landsat - 8 (2013 - )	Operational land imager (OLI)	9	Panchromatic (15 m), VIR, NIR, Mid-IR, SWIR (30 m)	15 days	Freely available
	Thermal infrared sensor (TIRS)		2	Thermal (100 m)	16 days	Freely available	
	ASTER	Terra (1999 - )	Advanced space-borne thermal emission and reflection radiometer (ASTER)	14	VIS, NIR (15 m), SWIR (30 m), Thermal (90 m)	16 days	Commercially, Research
	CBERS	China-Brazil Earth Resources satellite CBERS-4 (2014 - )	Multispectral camera (MUXCam); Panchromatic and multispectral camera (PanMUX); IRMSS-2 (Infrared multispectral scanner-2); WFI (Wide-field imager)	4, 4, 4, 4	VIS (20m); Panchromatic (5m), VIS, NIR (10 m); NIR, SWIR (40m), TIR (80m); VIS, NIR (64m)	26, 52, 26, 5 days	Freely available
	IRS	Indian Remote Sensing Satellite IRS-1A (1988 - 1996)	Linear imaging self-scanning sensor (LISS) - I; LISS - II	4,4	VIR, NIR, (72.5 m); VIR, NIR (36.25 m)	22 days	Commercially, Research
		IRS-1B (1991 - 2003)	LISS - I; LISS - II	4	VIR, NIR, (72.5 m); VIR, NIR (36.25 m)	22 days	Commercially, Research
	IRS-1C (1996 - 2007)	LISS - III	4	Panchromatic (5.8 m), VIR, NIR (23 m), SWIR (70 m)	24 days	Commercially, Research	
	IRS-1D (1997 – 2010)	LISS - III	5	Panchromatic (5.8 m), VIR, NIR (23 m), SWIR (70 m)	24 days	Commercially, Research	
High resolution	IRS	Resourcesat - 1 (2003 - 2013)	LISS - IV	4	Panchromatic, VIR, NIR (5.8 m)	5 days	Commercially, Research
		Resourcesat - 2/2A (2011 - )	LISS - IV	4	Panchromatic, VIR, NIR (5.8 m)	5 days	Commercially, Research
	Sentinel	Sentinel - 2 (A & B) (2015 & 2016 - )	Multispectral instrument (MSI)	13	VIS, NIR (10 m), SWIR (20 m), other (60 m)	5 days	Freely available
	SPOT	SPOT-1 (1986 - 1990)	Visible high resolution sensor (HRV)	4	Panchromatic (10 m), VIS, NIR (20 m)	26 days	Commercially, Research
		SPOT-2 (1990 - 2009)	Visible high resolution sensor (HRV)	4	Panchromatic (10 m), VIS, NIR (20 m)	27 days	Commercially, Research
		SPOT-3 (1993 - 1997)	Visible high resolution sensor (HRV)	4	Panchromatic (10 m), VIS, NIR (20 m)	28 days	Commercially, Research
		SPOT-4 (1998 - 2013)	Visible and infrared high-resolution sensor (HRVIR)	5	Mono-spectral (10 m), VIS, NIR, SWIR (20 m)	26 days	Commercially, Research
		SPOT-5 (2002 - 2015)	High resolution geometric sensor (HRG)	5	Panchromatic (2.5 m), VIS, NIR (10m), SWIR (20 m)	26 days	Commercially, Research
	SPOT 6 & 7 (2012 & 2014 - )	New AstroSat optical modular instrument (NAOMI)		Panchromatic (1.5 m), VIS, NIR (6 m)	5 days	Commercially, Research	
Very high resolution	IKONOS	IKONOS (2000 - 2015)	Optical sensor assembly (OSA)	5	Panchromatic (0.82 m), VIS, NIR (3.2 m)	Approx. 3 days	Commercially
	QuickBird	QuickBird (2001 - 2015)	Ball's global imaging system (BGIS 2000) sensor	5	Panchromatic (0.65 m), VIS,NIR (2.6 m)	3.5 days	Commercially
	RapidEye	RapidEye (2008 - ) multiple	Rapideye earth imaging system (REIS)	5	VIS, NIR (5 m)	5.5 days	Commercially
	GeoEye -1	GeoEye-1 (2008 - )	Geoeeye imaging system (GIS)	5	Panchromatic (0.41-0.46 m), VIS, NIR (1.65-1.84 m)	8 to 10 days	Commercially
	WorldView	WorldView-1 (2007 - )	Worldview-60 camera	1	Panchromatic (0.5 m)	2 to 6 days	Commercially
		WorldView -2 (2009 - )	Worldview -110 camera	9	Panchromatic (0.5 m), VIS, NIR (2 m)	1 to 3 days	Commercially
		WorldView - 3 (2014 - )	WV-3 imager, CAVIS	17	Panchromatic (0.3 m), VIS, NIR (1.2 m), SWIR (3.7m)	1 to 5 days	Commercially

Source: Dzikiti et al. (2019)

Different spectral bands can be combined and compared to observe subtle characteristics of features on the earth's surface. VIs are linear combinations or ratios using two or more spectral bands (Huete et al. 2002). VIs are used to enhance the vegetation properties (Figure 2-2), which enables interpreters to better study variations in vegetation vigour and biomass (Lukas et al. 2016). They are computed without knowledge of climate conditions or land cover and consequently provide an unbiased representation of vegetation (Jackson & Huete 1991).

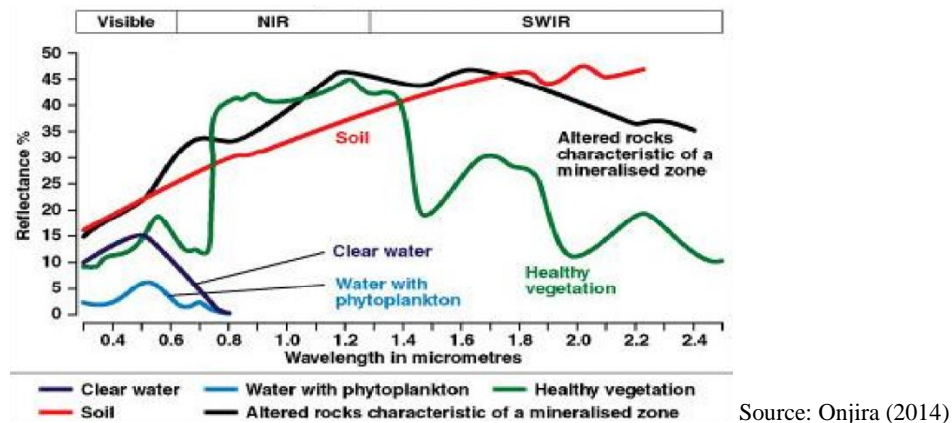


Figure 2-2 Spectral signatures for soil, water, and green vegetation

There are many different VIs, each intended for different applications. VIs aim to maximise the vegetation signal while minimising background signals (Jackson & Huete 1991). The most common VIs include the regular vegetation index (RVI), normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), and soil adjusted vegetation index (SAVI), all of which makes use of the visible to the NIR region of the EMS as vegetation strongly reflects NIR radiation and absorbs a large proportion of radiation in the visible region (Xue & Su 2017).

RVI is used to monitor vegetation variations during the peak growth period. The RVI is useful if the red band is measured with high accuracies (Jackson & Huete 1991) and is calculated by:

$$RVI = \frac{\rho R}{\rho NIR} \quad \text{Equation 2-1}$$

where  $\rho R$  is the red image band; and  
 $\rho NIR$  is the near-infrared image band.

RVI is sensitive to chlorophyll as it strongly reflects the NIR portion of the EMS (Huete et al. 2002). RVI can compare the growth of vegetation at a seasonal and inter-annual level. NDVI is a normalised ratio of RVI and is calculated by:

$$NDVI = \frac{\rho NIR - \rho R}{\rho NIR + \rho R} \quad \text{Equation 2-2}$$

where  $\rho_{NIR}$  is the near-infrared image band; and  
 $\rho_R$  is the red image band.

Although NDVI and RVI are functionally equivalent, RVI is unable to assess vegetation quantities at low densities, whereas NDVI can identify sparse vegetation as the normalised values enhance the visualisation of low RVI values (Jackson & Huete 1991). However, due to its nonlinearity, NDVI often saturates when there is high biomass as leaf pigments strongly absorb the red band and it is sensitive to canopy background noise (Huete et al. 2002). Soils and aerosols contribute to the signal received in the NIR and red bands causing higher and lower NDVI values respectively. Soils contribute more to the signal when vegetation is sparse, while aerosols contribute more when vegetation is dense (Liu & Huete 1995).

EVI is used to measure structural variations of vegetation (Huete et al. 2002). EVI overcomes the limitation of NDVI in high biomass regions by identifying the background signal and separating it from the signal emitted from the vegetation canopy. The equation uses the aerosol resistance term coefficients, which applies the blue band to correct for aerosol reflectance in the red band. The canopy adjustment term is used to compensate for canopy background noise. Therefore, EVI is suitable for measuring structural variations of vegetation such as leaf area index, plant physiognomy, and canopy type (Huete et al. 2002). EVI is defined by:

$$EVI = G \times \frac{(\rho_{NIR} - \rho_R)}{(\rho_{NIR} + C1 + \rho_R - C2 \times \rho_B + L)} \quad \text{Equation 2-3}$$

where  $G$  is the green band;  
 $\rho_{NIR}$  is the near-infrared band ;  
 $\rho_R$  is the red band;  
 $C1, C2$  is the coefficient of aerosol resistance term;  
 $\rho_B$  is the blue band; and  
 $L$  is the canopy background adjustment.

The main disadvantage of EVI is that it is sensitive to topographical effects. Therefore, it is important to remove topographical noise by applying bidirectional reflectance distribution function models before generating EVI (Matsushita et al. 2007). Similar to EVI, SAVI was developed to adjust for soil background noises that exist in the NDVI by incorporating a soil conditioning index into the equation (Xue & Su 2017). SAVI is calculated by:

$$SAVI = \frac{(\rho NIR - \rho R)(1 + L)}{(\rho NIR + \rho R + L)} \quad \text{Equation 2-4}$$

where  $\rho NIR$  is the near-infrared image band;  
 $\rho R$  is the red image band; and  
 $L$  is the soil conditioning index.

SAVI is equivalent to NDVI when  $L$  is set to 0, but SAVI is generally more accurate than NDVI (Xue & Su 2017).

In summary, VI's are useful for monitoring the growth, health status, and structural variations of different vegetation types and conditions.

### 2.1.2 Image analysis paradigms

RS has increased in popularity owing to improvements in digital data storage capacity and computing power, as well as the development of advanced software systems. Furthermore, many satellite data sources have become freely available and easily accessible, while the spatial, spectral, and temporal resolution of satellite images have also improved. This has led to the development and application of RS to Earth sciences, hydrological sciences, land use and land cover analysis, and plant sciences, etc. (Cracknell 2018).

RS is used in Earth science to map drainage patterns, detect lineaments, identify rocks and minerals, map soils and landscapes, and in photoclinometry. The tone of the imagery is used to identify drainage patterns as light tones are used to characterise porous soils. Similarly, brightness values are used to assess the orientation of the surface (Peng et al. 2003). Filtering and segmentation techniques are applied to remotely sensed images to detect lines or edges indicating an area of lineaments (Marghany & Hashim 2010). Mineral, soil, and landscape mapping use spectral signatures in classification algorithms to classify an image into mineral, soil and landscape groups (Mujabar & Dajkumar 2019).

Hydrospheric sciences apply active and passive RS techniques to locate the extent of water bodies by assessing a normalised difference water index derived from optical imagery (Kaplan & Avdan 2017) and/or the backscatter received by SAR sensors (Prasad, Garg & Thakur 2018). Additionally, SAR can detect surface water roughness, and the radiation detected by optical sensors can determine ocean surface water temperature. Hydrological models are better understood by assessing land cover and land use (LCLU) information derived from classifying remotely sensed images (Gao et al. 2018). Evapotranspiration rates are estimated by combining elevation

and optical imagery with data collected by meteorological stations and satellites (Neale et al. 2011).

Scenes of RS data are often classified into LCLU to better understand and manage land. Multi-temporal imagery is used to monitor the change of LCLU over time, driven by economic, social and environmental factors (Halefom et al. 2018).

Plant sciences rely heavily on RS, particularly in the agriculture and forestry sectors. In forestry, RS is often used to estimate crown closure, timber volumes, species, and forest plantation ages. This information can be collected periodically to generate up-to-date forestry inventory databases (Tomppo et al. 2008). Section 2.2 will elaborate on the applications of RS in forestry.

### **2.1.3 Classification**

A classified image is compiled by assigning every pixel in an image to a category. Each pixel is composed of values for each spectral band and is assigned to a class/category based on the similarity of these values to those of other pixels with known categories (Anand 2018).

Image classification is conducted using either a per-pixel or object-based image analysis (OBIA) approach. Per-pixel classification is the conventional method (Bhaskaran, Paramananda & Ramnarayan 2010) of evaluating and classifying individual pixels. A disadvantage of per-pixel classification is speckled results (Liu & Xia 2010). Speckle, or the so-called salt-and-pepper effect, arises from neighbouring pixels having high spectral heterogeneity (Kelly et al. 2011).

OBIA alleviates the effect of speckle by segmenting the image into homogenous and spatially continuous objects (Kelly et al. 2011). The results of image segmentation are influenced by the algorithm and its parameterisation. Object-based image classification assigns categories to the objects generated during the segmentation process based on the spectral, shape and spatial characteristics of each object (Blaschke 2010).

Classification algorithms can be categorised as unsupervised, supervised, and knowledge-based. The following subsections expand on each of these image classification types.

#### **2.1.3.1 Unsupervised**

Unsupervised classification follows a clustering approach whereby pixels with similar spectral values are statistically grouped into clusters of pixels. Pixels can either be individual pixels or multiple pixels grouped into objects through image segmentation, but for sake of brevity the term “pixels” will henceforth be used to refer to both scenarios. The analyst then assigns (or labels) informational classes (categories of interest) to the spectral groups (Naghdy et al. 2014).

The advantage of unsupervised classification is that the analyst does not need to have implicit knowledge of the area of interest before classification, there is less risk of human error, and unique spectral classes that may relate to informational classes are automatically identified (Naghdy et al. 2014). However, spectral classes do not always correspond to informational classes and may change over time, which reduces their value for automated classification methods (Naghdy et al. 2014). Examples of unsupervised classification algorithms are k-means, iterative self-organising data analytics technique (ISODATA), and modified k-means.

### 2.1.3.2 Supervised

Supervised classification categorises unlabelled pixels into informational classes using a sample of labelled pixels representing each informational class. The labelled pixels are thus used as training samples. Unlabelled pixels are assigned a label according to how similar their spectral characteristics are to those that are labelled (Basha et al. 2018).

The advantage of supervised classification is that informational classes are determined by the analyst (Basha et al. 2018). The disadvantage of supervised classification is that the classes selected by the analyst may be spectrally ambiguous, the training samples may not be representative as conditions may vary across the image influencing the pixel value, and collecting training samples is time-consuming (Campbell & Wynne 2013).

The maximum likelihood classifier (MLC) is a popular supervised classification algorithm. It determines the probability of each pixel belonging to a class by calculating the statistical distance between clusters using the variance, mean value and brightness of each cluster. MLC assumes that the spectral data within each informational class is normally distributed (Baatuwue & Leeuwen 2011).

Supervised classification methods such as k-nearest neighbour (KNN), classification and regression tree (CART), support vector machine (SVM), random forest (RF), and artificial neural networks (ANN) (Keuchel et al. 2003) are non-parametric as they do not make assumptions about data distribution and do not rely on statistical measures (Lillesand, Kiefer & Chipman 2019).

KNN classifies unlabelled objects based on the k-nearest known labels in feature space (the pixel values of each band). The analyst can set the value for k, which determines the number of closest neighbours to consider when classifying an unlabelled pixel. The final classification is based on a majority vote. The higher the k-value the more training data are required (Xie et al. 2019).

SVM attempts to find a hyperplane that best separates two classes in feature space. The feature space is transformed until a hyperplane can be used as a separator in the feature space. Mathematical functions such as radial basic function kernel, polynomial kernel, and linear kernels

are used to transform the feature space (Mather 2004). A disadvantage of SVM is that, if the number of features ( $n$ ) is much greater than the number of samples ( $s$ ) (i.e. when  $n \gg s$ ), the accuracy of the classification decreases.

Decision trees (DT), also called CART, evaluates the training samples and creates a set of thresholds that are used to recursively split the samples based on an attribute value test. The process is repeated until splitting the subsets adds no value to the informational classes. The main disadvantage of DT is that a large training dataset is needed and it is prone to overfitting (Baatuuwie & Leeuwen 2011).

RF is an ensemble classifier as it combines the results of multiple DT models. RF uses  $2/3$  of the sample data, known as subset R1, for training the model. The remaining  $1/3$  of the samples, known as the out of bag (OOB) sample subset, is used to estimate the variable importance and the classification error. At each node within each DT, the algorithm selects a random subset of input features (bands), known as F1. It then uses subset R1 to identify the feature in F1 that best splits the data for a class and creates new nodes. The process is repeated at each node. Once all nodes are split for each tree, the majority vote among trees decides the final output for the map (Budei et al. 2018).

ANNs are more complex than traditional statistical classifiers as they can model non-linear relationships. They contain three elements: an input layer, hidden layers and an output layer. The input layer contains the source data (imagery), hidden layers represent weights of association between classes and pixel values, and there can be many hidden layers. The output layer represents the classes for the desired output, which is defined by the training data during model building. The input data are passed through the network and weights are adjusted until the expected classification (defined by the training data) is achieved. Once the neural network (NN) is established, the input data can be replaced with other data. The disadvantages of ANNs are that they are complex and prone to overfitting (Han, Liu & Fan 2018).

Deep learning occurs when a multi-layered NN is formed, creating a deeper network than conventional NNs (Devi Mahalakshmi & Geethanjali 2019). A convolution NN (CNN) contains convolutional layers, max-pooling layers, and fully connected layers. Filters are applied to the convolutional layers, the dimensionality of the data is reduced in the max-pooling layers, and the fully connected layers ensure that all of the input data in one layer are connected to all of the units of the next layer (Devi Mahalakshmi & Geethanjali 2019). There are two main types of CNNs: LeNet and AlexNet. LeNet is a shallow NN containing two convolutional layers, two hidden layers, two subsampling layers, and an output layer (Devi Mahalakshmi & Geethanjali 2019). AlexNet is a deep CNN designed to classify an image into thousands of classes. AlexNet contains



five convolutional layers, three fully connected layers, three subsampling layers (Devi Mahalakshmi & Geethanjali 2019), and millions of parameters (Han, Liu & Fan 2018). NNs are advantageous as they can accept various numerical data even if the data does not have a statistical distribution, allowing them to process ancillary data to remotely sensed data (Mather 2004). A major disadvantage of NNs is that a large amount of training data and computing power is required (Han, Liu & Fan 2018).

### 2.1.3.3 Training samples and dimensionality

Training samples that are used in supervised classification methods are defined as pixels of known identity and are often collected from aerial photographs, satellite imagery (typically VHR), and/or maps, or using fieldwork (Lu & Weng 2007).

The number and distribution of training samples influence the accuracy of supervised classifications (Lu & Weng 2007). Different classification algorithms require different statistical characteristics of training data (Mather 2004). Campbell & Wynne (2013) suggest the number of categories, source of reference data, and the diversity of the categories should determine the number of training samples required in a classification. Areas of high spectral variability typically require more training samples (Lu & Weng 2007; Mather 2004). In addition, Mather (2004) suggests that the sample size should relate to the number of features (bands) from which the classification will occur, their statistical properties, and the diversity of spectral information within and among classes, referred to as intra- and inter-class spectral variability respectively.

The addition of bands to increase the separability between spectrally similar classes often leads to the  $s \ll n$  problem, where  $s$  is the number of samples and  $n$  is the number of features. This phenomenon, known as the Hughes effect, states that the classification accuracies increase gradually as features are added to the classifier, but at some point, the accuracies will decrease as the training data become sparser compared to the increased feature space (Ma et al. 2013). The Hughes effect is mitigated by increasing the number of samples or reducing the dimensionality of the data (Alonso, Malpica & De Agirre 2011).

Dimensionality reduction is achieved by using feature selection or feature extraction techniques. Feature selection finds the features that best represent the data, while feature extraction finds linear combinations of the original data to produce a smaller, optimised feature set (Zebari et al. 2020). A disadvantage of feature extraction is that some information is lost, while a disadvantage of feature selection is that excluded features may contain critical information for particular classes (Khalid, Khalil & Nasreen 2014). Pearson correlation, forward selection, backward elimination, and recursive elimination are examples of popular feature selection techniques. Feature extraction

techniques include principal component analysis (PCA), multidimensional scaling (MDS), and linear discriminative analysis (LDA). When the number of features cannot be reduced, the selection and number of training samples can be manipulated to mitigate the Hughes effect (Ma et al. 2013).

Training samples should represent distinct features, homogeneity and uniformity, and they must be placed throughout the image to capture the diversity in the image (Park & Lu 2015). Selecting homogenous samples becomes difficult for complex scenes, especially if medium- or low-resolution imagery is used in the classification. It is therefore important to consider the spatial resolution of the data being used in the classification and the complexity of the scene being mapped (Lu & Weng 2007).

Training samples are often collected using polygons (Lu & Weng 2007). The required number and size of training polygons vary among literature sources. Campbell & Wynne (2013) recommend training polygons exceeding 100 pixels, whereas Mather (2004) recommends obtaining more than  $30n$  samples per category, where  $n$  is the number of input features (bands). Similarly, Belgiu & Dragut (2016) propose collecting enough samples so that  $s > n$ . Congalton & Green (2019) suggest that when the area being classified is smaller than 1 million ha, the number of training samples should be about four times larger than the dimensionality of the dataset being classified (i.e.  $s \sim 4n$ ).

Thanh Noi & Kappas (2017) showed that the overall accuracy (OA) of a classification generally increases as the training set size increases, but that a saturation point exists where the OA is unaffected with an increase in training set size. Adding training data beyond the saturation point is unnecessary (Foody 2009). Instead, the training data should be selected to maximise the spectral separability between classes by collecting information about the spectral variability between classes (Foody et al. 2006).

There are different views on whether an equal number of training samples are required per class. Congalton & Green (2019) and Colditz (2015) suggest that classes that occupy a larger area require proportionately more training samples as they are often more complex and have a large spectral variation. Mellor et al. (2015) showed that complex classes, like forests, can be mapped more accurately by using an unbalanced training dataset. In contrast, Dalponte et al. (2013) and Millard & Richardson (2015) showed that when an unbalanced training dataset is used, the class with the most training data dominates the classification result. Belgiu & Dragut (2016) found that the user's and producer's accuracies of the scarce classes increase when using an area-proportionate (unbalanced) sample set.

A disadvantage of supervised classifications is that the collection of training data can be time-consuming and costly, especially when mapping at regional or national scales (Pax-Lenney et al. 2001). Signature extension or generalisation has been suggested to reduce the expense of training data collection. Signature extension is the process whereby a model is trained on one image and applied to other images or scenes (Laborte, Maunahan & Hijmans 2010). Signature extension can be applied across time and space (Wang, Azzari & Lobell 2019).

Zhang et al. (2019) used the concept of signature extension to create annual land cover maps. They created a spectral library containing the spectra of each land cover type to train an RF model that was subsequently applied to annual Landsat composites with OAs ranging from 71.3% to 80.7%. Similarly, Dannenberg, Hakkenberg & Song (2016) and Gray & Song (2013) collected training data from stable sites to create spectral signatures for training MLC and RF classifiers to map land cover classes across time. OAs ranging from 60% to 73.3% were achieved. Wang, Azzari & Lobell (2019) used signature extension in an RF model for crop type mapping across time and geographical distance. They found that an 80% OA is possible if the model is trained with data that contains similar growing degree days to the target image.

Phalke & Özdoğan (2018), Woodcock et al. (2001) and Li et al. (2020) classified land cover types over large areas by using signature extension with machine learning algorithms. They found that the OAs are affected by climatic, topological, and ecological differences. However, Olthof, Butson & Fraser (2005) and Verhulp & Van Niekerk (2016) found that as the distance increases away from the training scene, the classification accuracies deteriorate. Instead of signature extension, Knorn et al. (2009) used a chain approach to classify forest and non-forest areas over six Landsat scenes to reduce the impact of extension distance. Training data from one scene was collected and used to classify the scene. The portion of the classified scene that overlaps the neighbouring scene is used for training a second model, which is then applied to the neighbouring scene. This chain effect is repeated until all images are classified. However, as with conventional signature extension, the OA decreased as the distance from the first scene increased.

#### 2.1.3.4 Knowledge-based

Knowledge-based (expert systems) classification is another alternative to overcome the expense of collecting training data (Stephenson 2010). It applies a set of rules defined by an expert to classify remotely sensed images (Peled & Gilichinsky 2010). Expert systems consist of a knowledge base containing a set of rules (e.g. if-then statements), an inference engine that stores information about how to apply the rules, and a database that contains the transformed or raw datasets. The structure of a knowledge-based classification is similar to a DT, however, the splits are based on rules defined by an expert instead of algorithms (Watkins 2019). In comparison to a

supervised DT classifier, the rules defined in a knowledge-based classifier are based on the expert knowledge of the user, ancillary data and spectral information, mimicking how humans differentiate between classes on the earth's surface (Cohen & Shoshany 2002). Apart from not requiring training data, knowledge-based classification is advantageous in that the rules are transparent (open for scrutiny) and can be easily updated/modified and applied to other areas and/or data (Peled & Gilichinsky 2010). A disadvantage of a knowledge-based classification is that it relies on the knowledge of the user, resulting in an increased risk of human error (Baierle et al. 2019)

#### 2.1.4 Classification accuracy assessment

The classification scheme, the sampling method used to collect reference data, the amount of reference data, what reference data needs to be collected, and spatial autocorrelation are factors to consider when assessing the accuracy of a map (Lu & Weng 2007).

The reference data must preferably be collected at the same time the image was acquired and they need to be geographically co-registered to the imagery being classified. When collecting the reference data, the sampling scheme needs to have a random element to minimise bias. A simple random sampling method is appropriate for large sample sizes and a stratified random sampling scheme is suitable for a small sample size. Less than 50 samples per class are needed if the area being mapped is smaller than 1 million ha and the number of classes is less than 12 (Congalton & Green 2019).

Classified maps are evaluated in terms of site-specific and non-site-specific accuracy. Non-site-specific accuracy compares only the overall areas of each category of the map being classified to a reference map. Site-specific accuracy considers specific locations to assess the similarity between the reference data and the classified map (Congalton & Green 2019).

A confusion matrix (Figure 2-3) is constructed to calculate site-specific accuracies and documents the OA, error of commission, error of omission, and the kappa statistic (Lu & Weng 2007).

		Actual Change Status	
		Changed	No Change
Modeled Change Status	Changed	Correctly Classified	Commission Errors
	No Change	Omission Errors	Correctly Classified

Source: Pierce (2015)

Figure 2-3 Confusion matrix showing commission and omission errors

A confusion matrix is only suitable for classifications that have mutually exclusive and exhaustive categories and where every location on the map belongs to only one category (Lu & Weng 2007). The matrix compares the reference data and classified map at each reference sample to compute the different metrics. The OA is calculated by dividing the sum of the correctly classified pixel by the total number of pixels. The error of omission records when a feature is left out of the category being evaluated and errors of commission records when a feature is incorrectly included in the category being evaluated (Rwanga & Ndambuki 2017). A kappa statistic is used in conjunction with the OA. It considers the agreement of two maps by chance and is calculated by:

$$k = \frac{\text{observed} - \text{expected}}{1 - \text{expected}} \quad \text{Equation 2-5}$$

Pontius & Millones (2011) stated that kappa is unreliable for three reasons: firstly, kappa is a ratio making interpretation difficult as it is unknown whether the denominator or numerator is the dominating factor. Also, if the denominator is zero, the results will be undefined. Secondly, kappa does not indicate where the classification needs improving, and lastly, kappa compares the observed accuracy relative to accuracy due to chance, which is irrelevant. Quantity disagreement and allocation disagreement are more suitable for quantifying accuracy. Quantity disagreement is the similarity in the proportions of the categories, which is used to quantify the difference between the reference map and a comparison map. Allocation disagreement is the spatial allocation of the categories that are used to analyse the similarity in the proportion of the categories to quantify the difference between the reference map and a comparison map (Pontius & Millones 2011).

## **2.2 REMOTE SENSING IN FORESTRY**

RS has been used in forestry applications to identify trees, estimate crown closure and timber volumes, delineate compartments, and estimate tree ages. This section considers the following forestry applications of RS: mapping forested and other land cover types, classification of forest types, and genera/species classification. Specifically, RS methods/data that can be applied at regional/national scales are reviewed.

### **2.2.1 Mapping forests and other land covers**

Leroy et al. (2007) classified 250 m (bands 1-2), 500 m (bands 3-7), and 1 000 m (bands 8-36) MERIS imagery into land cover types on a global scale (Table 2-2). Temporal and spectral information was used in a clustering algorithm to capture phenological characteristics of land cover types. The classes were first labelled according to predefined classes and then an expert-based labelling system was applied to improve the labelling.

DeFries et al. (2000) generated a prototype global map depicting percentage tree cover and associated portions of trees with different longevity and leaf type using Advanced Very High-Resolution Radiometer (AVHRR) scenes. Data derived from a global network of Landsat scenes were used to determine weightings and endmembers for training. A mixture model was developed to estimate proportional coverage of woody vegetation, herbaceous vegetation and bare ground. A separate mixture model was developed for each continent to determine mixtures of broadleaf evergreen, broadleaf deciduous, and woody vegetation, depending on which types are present in each continent. Overall, the percentage of tree cover estimates appeared to depict the broad spatial patterns of forest distribution observed in Landsat thematic mapper (TM) scenes. Hansen et al. (2003) estimated the percentage tree cover globally using 500 m resolution MODIS imagery and a regression tree (RT) algorithm. Landsat images were used for training data collection. Hansen et al. (2005) tested 40-day composites and multi-temporal annual metrics derived from MODIS imagery for mapping percentage tree cover globally and continentally using an RT algorithm. MODIS global percentage tree cover data was used for training. Globally, the multi-temporal images produced better results than the composite images. The use of multi-temporal images produced better results over the African continent due to the different seasonal phenology experienced by common cover types in northern Africa compared to southern Africa. Continents that experience the same season at a given time, like North America, achieved similar results using the composite images and multi-temporal annual metrics in a RT.

Carrão et al. (2010) classified MERIS imagery of Portugal into land cover types using a linear discriminant classifier. Existing land cover data were used to assess the accuracy and an 80% OA was achieved. McRoberts et al. (2010) estimated the area of forested and non-forested plots on Landsat 5 scenes in four states of the United States of America (USA). They obtained an average location error of 7.9 m using supervised classification with stratified national land cover data as training. Tomppo et al. (2008) estimated forests and non-forested areas in Sweden and Finland using a KNN classification on Landsat scenes using national forestry inventory data for training. The root mean square error (RMSE) at a subsampling scale ranged from 5% to 16% for pine volume estimates.

Table 2-2 Summary of forest and land cover-type mapping studies

	Study	Methods	Results	Source
National Scale	Estimated forest/non-forest areas for Indiana, Iowa, Minnesota, and Missouri in the United States	Used a supervised classification with stratified national land cover data as training	An average error of 7.9 m was obtained	McRoberts et al. (2002)
	Estimates forests in Sweden and Finland	Used a KNN classification algorithm with different variables to classify Landsat images with NFI data for training	RMSE at a subsampling scale ranges from 5% to 16%	Tomppo et al. (2008)
	Estimate land cover types in Portugal	Used a linear discriminant classifier on MERIS imagery and tested with existing land cover data	Overall accuracy of 80% was achieved	Carrão et al. (2010)
Continental/Global Scale	Classifies land cover types on a global scale	Used a clustering algorithm with spectral and temporal information to capture phenological characteristics on MERIS imagery. The classes were then labelled into previously defined classes. An expert-based labelling system was then applied to improve the labels	Intensive validation campaign led by a network of experts from several parts of the world	Leroy et al. (2007)
	Generated a prototype global map depicting percentage tree cover and associated proportions of trees with different leaf longevity (evergreen and deciduous) and leaf types (broadleaf and needle-leaf) on a global scale	Used RT on AVHRR scenes with training data derived from a global network of Landsat scenes was used to determine weightings and end members	Overall the percentage tree cover estimates appear to depict the broad spatial patterns of forest distribution observed in the Landsat TM scenes	DeFries et al. (2000)
	Estimated percentage tree cover per 500m pixel globally	Used a supervised regression tree to classify MODIS imagery with Landsat grids as training data	Shows that MODIS data can discriminate percentage tree cover better than AVHRR	Hansen et al. (2003)
	40-day composites and multi-temporal annual metrics are tested in mapping percentage tree cover globally and continental	Used a regression tree with MODIS global percentage tree cover data for training	Globally - the metrics as input outperformed the composite images slightly. Africa - similar results as global due to the different seasonal phenology for common cover types. N America - the metrics and composites had similar results due to the common seasonal phenology change	Hansen et al. (2005)

### 2.2.2 Forest type classification

Baatuuwie & Leeuwen (2011) classified an ASTER image of a forest reserve in Ghana into built-up/bare, plantation forests, new plantations, natural forests, and mixed/degraded forest by exploiting an MLC, spectral angle mapper (SAM), and a DT classifier. The MLC was the most accurate achieving an 88.5% OA and a 0.85 kappa statistic. Nangendo, Skidmore & Van Oosten (2007) compared the ability of an MLC, SAM, MLC with an expert system, and SAM with an expert system to classify a Landsat scene of a forest reserve in Uganda into forest, woodland, and savannah types. The combination of SAM with an expert system was the most accurate, achieving a 96% OA. Lück (2018) used Landsat scenes covering three sites in South Africa to test the ability of a spectral rule-based classifier and two machine learning classifiers (RF and SVM) to discriminate between indigenous and forest plantations. The machine learning classifiers outperformed the spectral rule-based classifier obtaining between 93% and 98% OAs. Nery et al. (2019) utilised an OBIA with SVM, RF, and CART to classify a Landsat scene of an Australian catchment into indigenous and forest plantations. Texture metrics were calculated and added as features to the image, after which a principal component analysis (PCA) was applied to reduce the dimensionality of the data. A segmentation was conducted, training and testing data were identified, and the machine learning algorithms were implemented and compared. SVM achieved better results (92% OA) than RF (90%) and CART (88%).

Hagner & Reese (2007) applied an MLC to Landsat scenes of Sweden's forests to discriminate coniferous, deciduous and mixed-forest classes. Compared to Baatuuwie & Leeuwen (2011) (88.5%) a lower OA (74%) was achieved. However, Hagner & Reese (2007) classified a larger (and likely more complex) area than Baatuuwie & Leeuwen (2011) which may explain the lower accuracy.

Wagner et al. (2019) used OBIA with a CNN algorithm to classify a VHR image (WorldView2, WV2) of a forest biome in Brazil into eucalyptus and natural forest groups. A 95% OA was achieved. The combination of OBIA with CNN on VHR data (Wagner et al. 2019) outperformed the combination of OBIA with machine learning algorithms on medium resolution imagery (Nery et al. 2019).

Compared to the accuracies achieved in the studies that made use of VHR (Wagner et al. 2019) and medium resolution multispectral imagery (Baatuuwie & Leeuwen 2011; Hagner & Reese 2007; Lück 2018; Nangendo, Skidmore & van Oosten 2007 & Nery et al. 2019), it seems that hyperspectral data does not necessarily result in better accuracies when classifying forest plantation types/genera.



Table 2-3 Summary of forest type mapping studies

	<b>Study</b>	<b>Methods</b>	<b>Results</b>	<b>Source</b>
<b>Medium resolution imagery</b>	Classified built-up/bare, plantation forest, new plantations, natural forests, & mixed/degraded forests in a forest reserve	Tested ML, SAM, & DT using ASTER imagery	ML algorithm was the most accurate achieving an accuracy of 88.5%	Baatuuwie & Leeuwen (2011)
	Classified forest, woodland, and savannah types in a forest reserve	Used ML classifiers, SAM, ML + expert system, SAM + expert system to classify Landsat imagery	SAM + expert systems were the most accurate with a 96% accuracy	Nangendo, Skidmore & van Oosten (2007)
	Classified indigenous & plantation forests on a study area in SA	Tested rule-based and machine learning classifiers (RF & SVM) to classify Landsat imagery	Machine learning classifiers achieved 93% to 98% accuracy	Lück (2018)
	Classified indigenous & plantation forests over a catchment	Use OBIA. Used texture to segment and PCA, SVM, RF & CART to classify Landsat scenes	SVM was the most accurate (92%)	Nery et al. (2019)
	Classified coniferous, deciduous & mixed forests in a Sweden	Used an ML classifier with Landsat imagery	ML classifier achieved a 74% overall accuracy	Hagner & Reese (2007)
<b>Very high-resolution &amp; hyperspectral imagery</b>	Classified eucalyptus & natural forests in a forest Biome in Brazil	Used a segmentation technique followed by a CNN classification on WV2 imagery	Achieved an overall accuracy of 95%	Wagner et al. (2019)
<b>SAR imagery</b>	Classified deciduous and evergreen forests in Hainen Island	Used KNN classification on PALSAR imagery and then used MODIS imagery to identify rubber plantations	Achieved an 89% accuracy	Dong et al. (2012)
	Classified tropical forests and rubber plantations in Hainen Island	Used structural information derived from PALSAR & Vis derived from Landsat	Achieved a 86% accuracy	Chen et al. (2016)
	Classified rubber plantations in Danshou region close to Hainen Island using	Used PALSAR to delineate forest and Vis derived from Landsat to classify forest plantations	Achieved a 96% accuracy	Dong et al. (2013)

Dong et al. (2012) used an NN algorithm to classify a Phased Array type L-band Synthetic Aperture Radar (PALSAR) scene of Hainan Island into forested and non-forested areas, from which forest types were identified using a KNN algorithm to classify a MODIS image of the same scene. An OA of 89% was achieved for the PALSAR dataset and an 85% OA was achieved for the MODIS imagery. Similarly, Chen et al. (2016) utilised structural information derived from PALSAR imagery with VIs derived from Landsat imagery to map tropical forests and rubber plantations on Hainan Island, obtaining an 86% accuracy for the tropical forests and a 96% accuracy for the rubber plantations. Similar results were reported by Dong et al. (2013) who combined Landsat-derived metrics with PALSAR metrics to classify rubber plantations in the Danzhou region. From these studies, one can conclude that classification algorithms are capable of classifying SAR imagery into forest types, but are of greater value when differentiating between forest and non-forest land cover types.

### **2.2.3 Genera/species classification**

Xie et al. (2019) compared the ability of MLC, KNN, DT, RF, ANN, and SVM to classify seven forest tree species of a scene captured by bi-temporal ZiYuan-3 multispectral and stereo imagery. Spectral responses, textures, height, slope, and elevation were used to train the classification algorithms. MLC achieved the highest accuracy (89.4%), but RF (88.8%) and SVM (88.2%) obtained similar OAs. Vaglio Laurin et al. (2016) used VIs and texture metrics derived from Sentinel-2 data to distinguish between different forest species in a conservation area in Ghana. Jeffries-Matusita (J-M) separability, which shows how well the training data (regions of interests) are statistically different, was above 1.8 when texture metrics were included. Stabach et al. (2009) used an MLC to discriminate between forest species of Landsat 7 and Satellite Pour l'Observation de la Terre, lit (SPOT) 5 scenes. Low accuracies (69% to 72%) were achieved for both data sources due to the complexity and heterogeneity of the study area and the low spatial and spectral resolution of the imagery. Similarly, Franco-Lopez, Ek & Bauer (2001) used KNN to classify forest species using Landsat images of St Louis County and achieved low (47% to 52%) OAs.

Francois & Leckie (2006) used an MLC to classify individual trees in a research forest into seven species. The analysis was done using a VHR IKONOS image. A rule-based algorithm was applied to delineate individual tree crowns from which spectral signatures were generated for each of the seven species classes; a 67% OA was achieved. Ke, Quackenbush & Im (2010) obtained better results by applying a segmentation technique followed by DT classification on integrated LiDAR data and VHR Quickbird imagery to classify forest species in the Heidelberg Memorial Forest. Adding LiDAR data to the Quickbird imagery in the segmentation and classification processes increased the OA of the classification from 89% to 94%. Pu & Landry (2012) compared WV2 and

IKONOS data for mapping seven tree species using LDA and CART for classification in an object-based approach. Tree species were better differentiated using WV2 imagery compared to IKONOS imagery.

Immitzer, Atzberger & Koukal (2012) compared an object-based and per-pixel approach on VHR WV2 imagery by applying an RF algorithm to classify ten tree species. The object-based approach was more accurate (82%) than the per-pixel approach (73%). Similarly, Cho, Malahlela & Ramoelo (2015) used WV2 imagery and compared the ability of an object-based SVM and a per-pixel SVM classifier to map three forest plantation species in KwaZulu-Natal. The results showed that the object-based SVM is more accurate (89%) than the per-pixel classification (85%). Van Aardt & Norris-Rogers (2008) used CART to classify Compact Airborne Spectrographic Imager (CASI) hyperspectral data into eucalyptus and pine plantations. The dimensionality of the data was reduced by applying a minimum noise fraction (MNF) algorithm. The classification performed on the reduced dataset improved the outcome to an 85% OA and the computational time decreased. Furthermore, the classification improved to 97% OA when classifying plantations of the same age.

Franklin, Ahmed & Williams (2017) made use of OBIA with RF to classify an image captured by an unmanned aerial vehicle (UAV). Combining NIR with red-green-blue (RGB) bands and using crown shape, tone, texture and crown size showed an increase from 72% to 87% in OAs compared to just using NIR and RGB bands. Similarly, Franklin & Ahmed (2018) employed an object-based RF classification to differentiate four hardwood species on an UAV VHR image and an 78% OA was obtained.

Peerbhay, Mutanga & Ismail (2013) compared a classification of six forest species using a partial least square analysis on all the bands of an AISA Eagle hyperspectral image and selected bands of the same image from a variable importance algorithm. The OA improved from 80.61% to 88.78% using only the selected bands in the classification compared to using all of the bands in the classification. Buddenbaum, Schlerf & Hill (2005) used SAM and MLC to classify a HyMap hyperspectral image into coniferous forest species. Experiments were carried out using four sets of data, namely spectral information only, spectral information and stem density, spectral and textural information, and all the data together. The OA improved from 66% when using only the spectral information to 77.8% when texture metrics were added.

Table 2-4 Summary forest genera and species mapping studies

	Study	Methods	Results	Source
Medium resolution imagery	Classified 7 forest tree species	Used spectral responses, texture, height, slope, & elevation derived from Bi-temporal Ziyuan-3 multispectral and stereo imagery to train the algorithm	Achieved an 84.5% accuracy	Xie et al. (2019)
	Classified forest species in a conservation reserve in Ghana	Used vegetation indices & textures in SVM & ML	SVM was more accurate than ML. The addition of texture is very important	Vaglio Laurin et al. (2016)
	Classified forest species in a field research area	Used KNN vegetation plots to guide class assignment on Landsat & SPOT imagery	Very low accuracies were achieved	Stabach et al. (2009)
	Classified forest species over St. Louis Country	Used KNN with different hyper-parameters to classify Landsat images	Very low accuracies of 47% to 52% were achieved	Franco-Lopez, Ek & Bauer (2001)
Very high-resolution imagery	Classified individual trees into seven species in a research forest	Used a rule-based algorithm to delineate individual trees on IKONOS imagery. Applied an ML algorithm based on spectral signatures	Achieved a 67% overall accuracy	Francois & Leckie (2006)
	Classified forest species in Heidelberg Forest	Used LiDAR-derived metrics with Quickbird data to segment and classify the forest	Achieved a 91.6% kappa statistic	Ke, Quackenbush & Im (2010)
	Classified 7 tree species	Compared WV2 & IKONOS by using an OBIA with LDA & CART for classification	WV2 imagery was more accurate	Pu & Landry (2012)
	Classified 10 tree species	Compared OBIA to per-pixel classification using an RF classifier to classify WV2 imagery	OBIA was more accurate with an 82% overall accuracy	Immitzer, Atzberger & Koukal (2012)
	Classified 3 forest plantation species in a small area in KZN	Compared OBIA to per-pixel classification using SVM to classify WV2 imagery	OBIA Was more accurate with an 89% accuracy	Cho, Malahlela & Ramoelo (2015)
	Classified eucalyptus & pine plantations	Used MNF to reduce the dimensionality of CASI hyperspectral data & applied a CART classification	Achieved an 85% overall accuracy of mixed aged trees. Improved to 97% accuracy when classifying trees of the same age	Van Aardt & Norris-Rogers (2008)
UAV imagery	Classified conifer forests into species	Used an object-based classification using NIR, RGB, crown shape, tone, texture & crown size on UAV imagery	The addition of crown shape, tone, texture & crown size increase accuracy to 87%	Franklin, Ahmed & Williams (2017)

Continued overleaf

Table 2-4 continued

	Study	Methods	Results	Source
Hyperspectral imagery	Classified four hardwood species	Used object-based classification with an RF algorithm to classify UAV imagery	Achieved a 78% accuracy	Franklin & Ahmed (2018)
	Classified coniferous forest species using HyMap hyperspectral data	Compared spectral information, stem density and textural information using SAM and ML classifiers	Statistical textural information proved to be the most accurate achieving a 0.74 kappa statistic	Buddenbaum, Schlerf & Hill (2005)
	Classified swamp tree species using HyMap data	Used simple forward selection for dimensionality reduction & applied an RF classification	Achieved a 90.5% overall accuracy	Adam et al. (2012)
	Classified forest plantation species using AISA hyperspectral and LiDAR data	Used LiDAR data to segment the image and then used hyperspectral to classify the image using a rule-based classification an MNF for dimensionality reduction	Achieved a 57% accuracy	Voss & Sugumaran (2008)
LiDAR imagery	Classified tree species using multi-temporal and HyMap hyperspectral data	Used LDA to reduce the data. Multi-temporal data was combined with hyperspectral data in an RF classification	Obtained an 88.5% accuracy	Fagan et al. (2015)
	Classified tree species using full-waveform LiDAR	Used LDA to identify the important variables and applied a classification	Obtained a 91% accuracy when classifying conifers and broadleaves	Heinzel & Koch (2011)
	Classified 4 tree species using LiDAR-derived metrics	Used a genetic algorithm to select important variables to use in an LDA for classification	Obtained a 77% accuracy	Li, Hu & Noland (2013)
	Classified two study sites into nine species using multispectral LiDAR	Used an RF classification	Achieved a 75% accuracy	Budei et al. (2018)
	Classified forest species using LiDAR-derived metrics	Used a DT classify with LiDAR-derived measures	Obtained a 0.9 kappa statistic	Martinuzzi et al. (2013)

Adam et al. (2012) used RF and RT to classify swamp tree species using a HyMap image as input. A simple forward variable selection algorithm was used to reduce the dimensionality of the data. The OA of 90.5% was achieved using RF compared to 84.5% using RT. The classification improved from 88.4% to 90.5% and 80.5% to 84.5% using a reduced data set compared to the full dataset in RF and RT respectively.

Voss & Sugumaran (2008) classified a study area in northern Iowa into forest plantation species using OBIA. LiDAR data was used to segment the hyperspectral image to compensate for the shadow effect present in the hyperspectral image. A rule-based classification was applied using NDVI to separate vegetation and non-vegetation, LiDAR-derived height metrics to separate high vegetation and low vegetation, and LiDAR intensity to differentiate between coniferous and deciduous trees. An MLC was used to classify species within the coniferous and deciduous parent groups. The classification accuracies improved from 48% to 57% when including LiDAR data in the classification. Fagan et al. (2015) achieved superior results compared to Voss & Sugumaran (2008) by combining hyperspectral data with multi-temporal Landsat imagery as classification input data to differentiate between tree species in the lowlands of Costa Rica. The hyperspectral image was used in an RF algorithm to differentiate between tree species. The dimensionality of the hyperspectral data was reduced to the most important bands, which were then combined with multi-temporal data, resulting in an 88.5% OA.

### **2.3 SUMMARY AND EVALUATION**

Based on the literature reviewed, it seems that using low-resolution (>250 m) imagery as input to classifiers for forest type/species/genera mapping is not effective, likely due to spectral mixing that reduces the ability of the imagery to adequately represent small spectral differences in and among forests. Better results were achieved using LiDAR-derived metrics and hyperspectral data as input to classifiers. But LiDAR and hyperspectral data are processing-intensive and costly to acquire, which reduces their viability for forestry applications over large areas. Although VHR multispectral imagery has successfully been combined with machine learning to distinguish between forest genera in small areas, such data are also expensive to acquire over large areas. The backscatter recorded by SAR sensors has been successfully used to estimate vegetation structure. This data has potential for regional applications given that some data sources are freely available (e.g. Sentinel-1). High-resolution (10-30 m) multispectral imagery is another potential source of data as some sources (e.g. Sentinel-2 and Landsat-8) are freely available, but it is not known whether such data would be suitable for differentiating genera in regions with different environmental characteristics, such as those that support forestry in South Africa.

OBIA seems to work better than per-pixel analysis to distinguish between forest genera because the crown structure and transmissivity of forest plantation is better represented over multiple pixels rather than individual pixels, particularly when VHR imagery is used as input (Immitzer, Atzberger & Koukal 2012). However, object delineation and segmentation parameterisation is an ill-structured problem (Louw & Van Niekerk 2019). Given that most of the literature uses a per-pixel approach and that there is a level of uncertainty with image segmentation, this study will use a per-pixel approach. Supervised classification techniques generally perform better than unsupervised classification to distinguish between forest plantation genera. A likely explanation is that plantation forests are spectrally similar, which causes unsupervised classification techniques to group multiple genera in the same spectral class. Supervised classification seems to be more suitable because it uses training samples of predefined classes to label pixels/objects and can thus be manipulated to distinguish between forest plantations. Furthermore, multi-temporal images are more suited than single date images as phenological variations among forest plantation genera can be considered.

Forest plantation genera information is important for forest planning and inventorying. However, the current methods used to collect such information is time-consuming, costly, and only occurs every three years in South Africa. An operational solution that uses RS and machine learning to map genera at national scale will consequently be of great value. Satellite imagery with very high spatial and a medium to high spectral resolution is likely the most viable source of RS data to represent variations in the crown structure, texture, and spectral information of genera and to produce forest plantation maps over large and complex areas. Landsat-8, Sentinel-2, and Sentinel-1 are the most viable sources of data given that they have high spatial resolution, are freely available and have a global coverage. According to the literature, Sentinel-2 seems to show the most potential, given that it has a 10 m spatial resolution which is suitable for generating texture measures to capture crown structures. To date, Sentinel-2 has not yet been used to map forest plantation genera over large and complex areas such as South Africa.

A large number of training samples are needed for supervised classification to be an operational solution. The samples have to be collected in-field, but only need to be updated sporadically, given that forest plantations grow at a slow rate. The cost to collect samples over South Africa will be very costly, which reduces the value of a RS approach to map forest plantation genera. Therefore, it is vital to find a sampling strategy that will maximise classification accuracies while keeping costs as low as possible. This study is a first step towards the establishment of a sampling scheme for operational monitoring of forests at a national scale as it intends to provide a greater

understanding of how sampling factors (e.g. size, configuration/balance, and placement) influence genera classification accuracies.

Chapter 3 (objective 3) will investigate which training sampling configuration and size will result in the best accuracies for classifying forest plantation genera. The main purpose of the chapter is to identify the best training sample configuration and size so that the costs and time associated with training data collection can be minimised. Chapter 4 (objective 4) evaluates the efficiency of signature extension for classifying forest plantations over large and diverse areas. In essence, it evaluates how signature extension can be used to minimise training data collection costs and time. Chapters 3 and 4 were written as research articles and are intended for publication in scientific journals. Each chapter was consequently written as independent documents, which means that there is some duplication between the two chapters. Chapter 5 (objective 5) is a synthesis of the research and the research limitations and recommendations for operationalising the methods are also discussed.



## **CHAPTER 3: IMPACT OF TRAINING SET CONFIGURATIONS FOR DIFFERENTIATING BETWEEN PLANTATION FOREST GENERA WITH SENTINEL-2 IMAGERY AND MACHINE LEARNING**

### **3.1 ABSTRACT**

The genus of planted trees is a fundamental forest inventory variable and is generally collected using costly and time-consuming in situ observations. Although remotely sensed data and machine learning show potential for mapping genera at regional scales, such approaches require training data and there is uncertainty about which sampling configuration, sample size, and sample distribution will be most effective over large and complex areas. This study aimed to evaluate the effect of different sampling strategies (e.g. even, uneven, and area-proportionate) for training the random forest (RF) machine learning classifier to differentiate among acacia, eucalyptus, and pine trees using Sentinel-2 imagery as input. Sample size ( $s$ ) was related to the number of input features ( $n$ ) to better understand the potential impact of sample sparseness. The results show that an even sample with a maximum size (100%,  $s \sim 81n$ ) produced the highest overall accuracy (OA) (76.3%). A 6% difference in OA was recorded when the training samples were enlarged from  $\sim 20\%$  to 100% of the total samples. Although larger training set sizes ( $s > n$ ) resulted in higher OAs, a saturation point was reached at  $s \sim 57n$ . The eucalyptus class was the most accurately classified, followed by pine and then acacia. We concluded that RF yields the most accurate plantation genera maps when a balanced sampling scheme is used, but noted that accuracies are influenced by genus mix and spectral separability of classes. More work is needed to assess model transferability, particularly to areas where the genus proportions are severely skewed.

### **3.2 INTRODUCTION**

Forest inventories include information about compartment location, extent (area), planting date, tree genus/species/clone, yield, and water use (Mati & Dawaki 2015). The information derived from inventories is used for planning, for land management, to maximise production by assessing the mean annual increments of growth rates, to assess water use, to monitor timber harvests and rotations, and to assist in forest management (Mati & Dawaki 2015). Genus information is specifically used in allometric equations for carbon stock estimates and changes (Maniatis et al. 2011) (to estimate the biomass and carbon stock levels) (Basuki et al. 2009) and in streamflow reduction models (Gush et al. 2002). Although commercial forestry companies keep records of plantings, such data are not in the public domain and will likely never be made known as they can be used by competitors to gain a commercial advantage. In addition, many plantations are owned by small growers who are unlikely to keep plantation records (Forestry South Africa 2019). The

inventory data that do exist are often inaccurate, outdated, and inconsistent and there is no way of validating the data.

Inventory data are traditionally collected in-field, which is time-consuming and costly. A methodology whereby forest plantation genera can be cost-effectively mapped and frequently updated at regional/national scales would be beneficial for keeping forest inventories up-to-date (Food and Agriculture Organization 2015). Earth observation (EO) and remote sensing (RS) offers a potential solution as satellite images are captured over vast areas (global coverage) at regular (up to 15-minute) intervals, e.g. EUMESAT (Schulz et al. 2009). Although RS is often used for regional land cover mapping (Jokar Arsanjani, Tayyebi & Vaz 2016; Thompson 2019), land cover maps tend to only differentiate forests from other land covers (e.g. water, built-up areas, crops). Some land cover maps differentiate between indigenous and plantation forests (DEA 2019; Lück 2018) but do not disaggregate these land covers into types or genera.

Medium resolution (30 m – 250 m) imagery has been combined with machine learning algorithms to map forest species with varying success. For example, Landsat (30 m) and SPOT (20 m) imagery were used by Franco-Lopez, Ek & Bauer (2001) to generate a three-class species map with an accuracy of 47%, and by Stabach et al. (2009) to generate a 13-class species map with a 63% accuracy. Higher accuracies have been achieved using very high-resolution (VHR) satellite imagery (Cho, Malahlela & Ramoelo 2015; Francois & Leckie 2006; Immitzer, Atzberger & Koukal 2012; Ke, Quackenbush & Im 2010; Pu & Landry 2012), while unmanned aerial vehicle (UAV) data (Franklin, Ahmed & Williams 2017; Franklin & Ahmed 2018) and/or hyperspectral imagery (Buddenbaum, Schlerf & Hill 2005; Bujang & Baharum 2017; Fagan et al. 2015; Peerbhay, Mutanga & Ismail 2013; Voss & Sugumaran 2008) have also been used for this purpose. However, VHR, UAV and hyperspectral imagery are often too costly to acquire over large areas and require substantial computing power to process, making them less suitable for regional and/or national applications. The employment of high-resolution (10 m - 60 m) imagery—such as those generated by the Sentinel-2 constellation as part of the Copernicus Programme (ESA 2015)—is more viable, as it is freely available, frequently updated and can be cost-effectively processed on cloud computing platforms such as Google Earth Engine (GEE).

Nomura & Mitchard (2018) used Sentinel-2 data with supervised classification to map seven commercial forest plantation species/genera in Myanmar. They achieved a 95% overall accuracy (OA) using an unbalanced training dataset. Mngadi et al. (2019) classified seven species in the Clan forest plantation (located in South Africa) using Sentinel-2 bands in a linear discriminative analysis (LDA) and achieved an 84% OA. When adding Sentinel-1 vertical-vertical (VV) and vertical-horizontal (VH) features, the accuracies increased to 87%.

Vegetation indices (VIs) and textural measures, derived from the original bands, are frequently used to improve forest plantation species/genera classification accuracies (Nomura & Mitchard 2018; Vaglio Laurin et al. 2016). However, such image transformations increase the dimensionality of input data, which is problematic for many supervised classification algorithms, especially parametric classifiers such as the maximum likelihood classifier (MLC) (Loggenberg et al. 2018). High dimensionality can lead to the so-called  $s \ll n$  problem, where  $s$  is the number of samples per class and  $n$  is the number of features. This phenomenon, known as the Hughes effect, states that at some point, the increased number of features may result in lower accuracies unless the number of training samples is increased proportionally (Ma et al. 2013).

Belgiu & Dragut (2016) stated that training samples need to be statistically independent and representative of the class being mapped, while the number of training samples per class must be balanced and large enough so that  $s > n$ . In contrast, Congalton & Green (2019) recommend that  $s$  should be equal to 50 when  $n = 12$  ( $s \sim 4n$ ) and when the area being mapped is smaller than 1 million hectares (ha), while Mather (2004) suggests using  $10n$  to  $30n$  training samples. Thanh Noi & Kappas (2017) found that classification accuracy increases as sample size increases, but Foody (2009) showed that a saturation point exists, where adding more samples does not significantly increase classifier performance and the addition of samples beyond this point is a waste of resources. To minimise training sample collection cost, Foody et al. (2006) suggest collecting training data that maximise inter-class separability, while minimising intra-class spectral variability. However, such an approach requires prior knowledge of the spectral variability within and among target classes. Congalton & Green (2019) suggest collecting more samples for larger areas as they are often more complex. They also suggest collecting samples for each class relative to the importance of those classes for the mapping objective.

Dalponte et al. (2013) investigated the impact of using unbalanced training sets and found that the class with the most samples was favoured in the classification. This is in agreement with Millard & Richardson (2015), who found that the dominant class in the training set resulted in that class being the most accurate. In contrast, Mellor et al. (2015) showed that an imbalance in the training data can improve the accuracies of complex classes, such as open-canopy woodlands and forests, which occur over diverse ecosystems and topography. Generally, classes that occupy a larger area require more samples due to a large spectral variation (Colditz 2015), and when area-proportional samples ( $s/A$ ) are used, the producer's and user's accuracies of the smallest class being mapped increases (Belgiu & Dragut 2016).

Forest plantations are generally grown in areas with mean annual precipitations of 750 mm or more. When forest plantations are grown in semi-arid areas, such as South Africa, they often occur

in localised areas where the rainfall is sufficient, resulting in plantations being sparsely distributed in the landscape. Forest plantations in South Africa occupy only 1.2% of the land surface and cover approximately 1.2 million ha (Figure 3-1). It is roughly estimated that about 50% of plantations are planted with pine, 43% with eucalyptus, and 7% with acacia trees (Van der Zel 1995). In terms of the geographic make-up of these plantations, 40% are likely situated in Mpumalanga (MP), 40% in KwaZulu-Natal (KZN), 11% in the Eastern Cape (EC) and the remaining 9% in equal proportions in the Western Cape (WC) and Limpopo (Van der Zel 1995). The sparseness and proportion of each genus pose a unique machine learning challenge, particularly in terms of collecting representative training samples over such a vast area (Xulu et al. 2018). Although some successes have been achieved at local scales (Mngadi et al. 2019), no attempts have been made to map South Africa's main commercial forest genera (acacia, eucalyptus, and pine) at a regional/national scale. As with many EO machine learning applications, the main obstacle for producing such an inventory is the absence of suitable in situ data to train and build a model. Collecting in situ data over such a vast area would be very costly and there is uncertainty about which sampling configuration (total sample size, samples per class, and spatial distribution of samples) will be most effective for training a machine learning classifier to differentiate among genera at national scale.

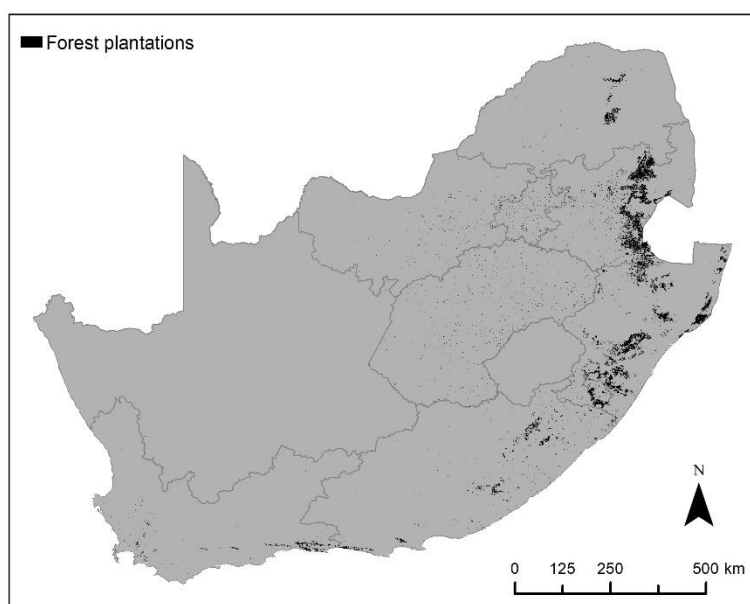


Figure 3-1 Distribution of South African forest plantations (DAFF 2008)

This study aims to investigate the impact of employing different training data set sampling configurations and sizes on the performance of a machine learning classifier (RF) for forest plantation genus mapping. Using Sentinel-2 imagery as input, different sampling strategies are used to train the classifier and the resulting accuracies are compared. Experiments are carried out in two very diverse study sites to assess the effect of enlarging training set size under balanced

(where  $s/n$  is constant for all classes), imbalanced ( $s < 12n$  for some classes), and area-proportional ( $s/A$  is constant) scenarios. The results are interpreted within the context of finding the most effective approach to training sample collection, particularly for regional/national forest genus mapping in semi-arid areas where plantations are sparsely distributed and where genus mix varies dramatically from one area to another. In contrast to previous studies, the focus of the present study is on the relative accuracies achieved among sampling strategies using a consistent feature space (a common set of imagery) and classification algorithm (RF), rather than a general assessment of the viability/suitability of using different machine learning algorithms and data sets for genus classification.

### 3.3 METHODS AND MATERIALS

#### 3.3.1 Study areas

Two study sites (Figure 3-2) were selected for carrying out the experiments. Study Area 1 stretches from the settlements Knysna to Kranshoek in the WC province of South Africa, while Study Area 2 spans settlements New Hanover to Osborn in KZN. These sites were chosen owing to the diversity of the genera planted and the availability of reference data (in situ forest plantation extents and genera). The selected sites are also climatically representative, as they are respectively located in the winter and summer rainfall regions of South Africa.

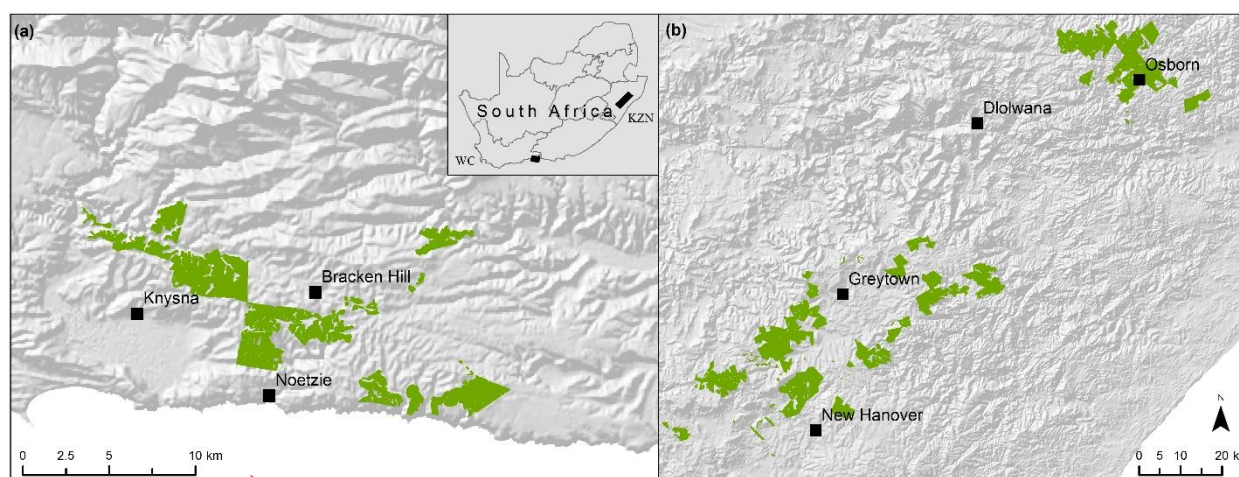


Figure 3-2 Locations of Study Area 1 (a), along the southern coast of the Western Cape (WC) province, and Study Area 2 (b) in the KwaZulu-Natal (KZN) midlands

Study Area 1 receives ~ 875 mm of rainfall per annum, with peak rainfall during July. The mean temperature in the area is 16.9°C (Kraaij et al. 2018). The annual rainfall in Study Area 2 ranges from 1 400 mm in the southwest to 1 570 mm in the northeast (peak rainfall during January) and has a mean temperature of 17.5°C (Fuller & Perrin 2001).

Study Area 1 is dominated by pine, followed by eucalyptus and then acacia, whereas Study Area 2 is planted equally by acacia, eucalyptus and pine.

### 3.3.2 Data collection and preparation

#### 3.3.2.1 Imagery

The GEE platform was used to access Sentinel-2 level-2A surface reflectance satellite imagery. Sentinel-2 imagery has a five-day temporal resolution and contains 13 spectral bands with spatial resolutions of 10 m (B2, B3, B4, B8), 20 m (B5, B6, B7, B8A, B11, B12) and 60 m (B1, B9, B10). A median composite image was generated from images dated 2019-06-30 to 2020-06-30. The reasoning behind using a composite image was to remove cloud contamination and to compensate for seasonal variations among the two study areas. Although it is well known that phenological variations and multi-temporal EO approaches can aid in genus classifications (Mngadi et al. 2019), we purposefully excluded such variations as it would introduce uncertainty to our sampling strategy results. For instance, some training sample sets may benefit more from seasonal variations than others, which would add complexity and potentially skew the findings.

Machine learning algorithms can better distinguish between forest plantation genera when adding textural measures and vegetation indices as features to multispectral imagery (Vaglio Laurin et al. 2016 & Xie et al. 2019). Therefore, the normalised differential vegetation index (NDVI), enhanced vegetation index (EVI), entropy, and grey level co-occurrence matrix (GLCM) measures were derived and added as bands to the composite image, which resulted in a total of 37 features (Table 3-1).

Table 3-1 Features (bands, indices and textural measures) used as input to the classifications generated from a composite of individual Sentinel-2 images taken from 2019-06-30 to 2020-06-30

Bands	Textural Measures	
B1 – coastal aerosol	angular second moment	cluster prominence
B2 – blue	contrast	sum average
B3 – green	correlation	sum entropy
B4 – red	difference entropy	cluster shade
B5, B6, B7, B8A – vegetation red edge	dissimilarity	sum variance
B8 – NIR	difference variance	variance
B9 – water vapour	entropy	
B10 – SWIR - cirrus	inverse difference moment	
B11, B12 – SWIR	information measure of correlation 1	
	information measure of correlation 2	
Vegetation Indices	inertia	
NDVI	maximum correlation coefficient	
EVI		

### 3.3.2.2 In situ data

In situ (ground-truthed) data at plantation compartment level were collated from several South African commercial forestry companies as polygons in a shapefile format. This data included records of the species, genus, and planting date of the trees in each compartment. Table 3-2 provides an overview of the collected in situ sample data per genus within each study area.

Table 3-2 Summary of in situ data (forest compartment information) collated, including tree genus, age (mean and standard deviation) and planted area per study area

Genus	Study Area 1 (WC)					Study Area 2 (KZN)				
	Age (mean)	Age (stddev)	Area (ha)	Area (%)	#	Age (mean)	Age (stddev)	Area (ha)	Area (%)	#
Acacia	35.94	12.61	118.14	3.56	40	8.75	7.10	3869.7	33.33	406
Eucalyptus	43.18	20.49	145.79	4.40	50	6.52	4.57	3869.4	33.33	718
Pine	9.97	11.08	3052.81	92.04	940	8.78	6.50	3869.3	33.33	478
<b>Total</b>			<b>3316.74</b>		<b>1030</b>			<b>11608.4</b>		<b>1602</b>

Key: Standard deviation (stddev); Number of compartments (#)

Figure 3-3 shows the geographic distribution of in situ data within each study area. Compared to Study Area 1, the in situ data in Study Area 2 spans a larger area and are more widely distributed. The areas covered by each genus in Study Area 1 is unbalanced (dominated by pine), while in Study Area 2 the area covered by each genus is more or less equal.

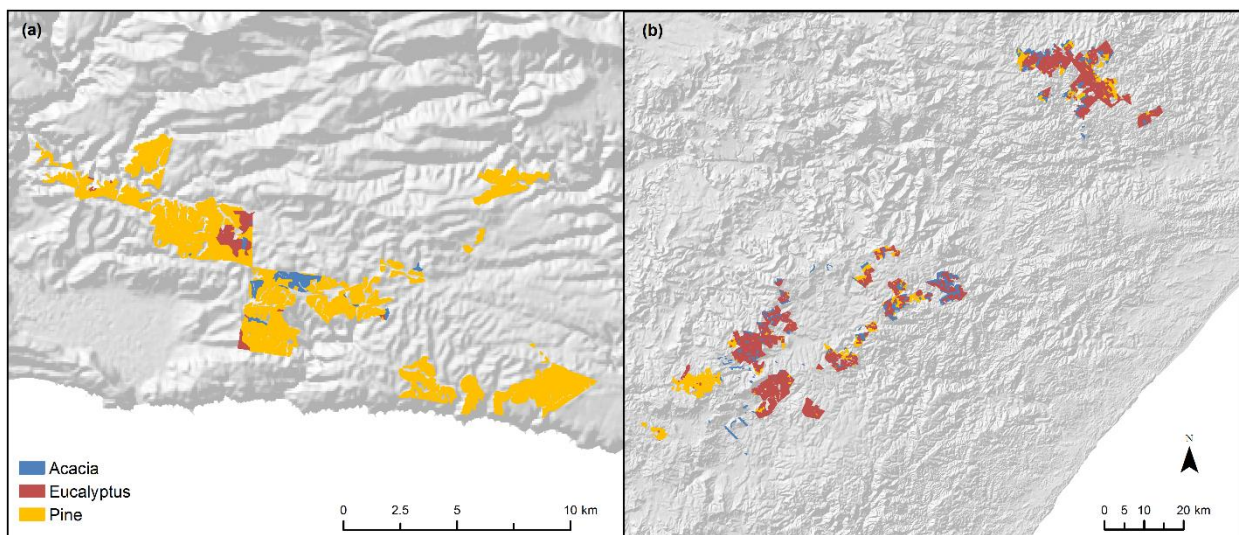


Figure 3-3 Genus distribution of in situ data in Study Area 1 (a) and Study Area 2 (b)

### 3.3.3 Experimental design

The in situ data were consolidated into a shapefile and the compartments were dissolved to form a single polygon per genus. Random points were generated in each polygon to obtain 3 000 samples per genus (i.e. 9 000 in total) in both study areas. This sample set was denoted *SS\_0* and

represents 2.7% and 0.76% of the total area for which in situ data were available for all three genera in both study areas. The Sentinel-2 band values were then extracted at each point in *SS\_0* using GEE. The training data was collected with a minimum distance of 10 m between samples within each genus polygon. The polygons from which the training data were collected contain many compartments of different tree ages and some with different species. The variety of age and species per genera and the distance between the random samples minimises autocorrelation in the training data. Each experiment made use of the 9 000 possible point samples for training the model. The experiments (Table 3-3) were designed to investigate the combined effect of training set size and imbalance.

Table 3-3 Summary of Experiments A to H

Exp. A	Exp. B	Exp. C	Exp. D	Exp. E	Exp. F	Exp. G	Exp. H
Select 50 samples per genera. Iteratively add 50 samples per genera until 3 000 samples per genera are used.	Select 50 samples per genera. Iteratively add 50 eucalyptus and pine samples until 3 000 eucalyptus and pine samples are used.	Select 50 samples per genera. Iteratively add 50 acacia and 50 pine samples until 3 000 acacia and pine samples are used.	Select 50 samples per genera. Iteratively add 50 acacia and 50 eucalyptus samples until 3 000 acacia and eucalyptus samples are used.	Select 50 samples per genera. Iteratively add 50 pine samples until 3 000 pine samples are used.	Select 50 samples per genera. Iteratively add 50 eucalyptus samples until 3 000 eucalyptus samples are used.	Select 50 samples per genera. Iteratively add 50 acacia samples until 3 000 acacia samples are used.	Select 3 acacia, 4 eucalyptus and 50 pine samples. Iteratively add 3 acacia, 4 eucalyptus and 50 pine samples until 3 000 pine samples are used.

*Experiment A* aimed to quantify the effect of a balanced training sample set in different sample size scenarios. The experiment starts with randomly selecting 50 (~ 1.6%) samples per genus from *SS\_0*. When compared to the number of features ( $n$ ), this initial subset's size ( $s$ ) can be expressed as  $s \sim 2n$  (i.e. about two samples per feature/dimension, per genus). The training sample was then enlarged by iteratively adding 50 (randomly selected) samples from *SS\_0* until all (100%,  $s \sim 81n$ ) of the samples in *SS\_0* were utilised for training the model.

*Experiments B, C, and D* are similar to *Experiment A*, except that the size of the initial subset ( $s = 50$ ) was not enlarged for one of the genera. For instance, in *Experiment B*, the acacia subset was kept at 50 samples, while the subsets for the other two genera (eucalyptus and pine) were enlarged by 50 samples per iteration until all available samples were utilised ( $s = 3\ 000$ ). Similarly, the training subsets for eucalyptus and pine were kept at 50 samples in *Experiments C* and *D* respectively. These experiments were conducted to assess the effect when one class (genus) is severely under-sampled.

*Experiments E, F, and G* are similar to *Experiments B, C, and D*, except that the size of the initial subset ( $s = 50$ ) was kept constant for two of the genera. For instance, in *Experiment E*, the acacia



and eucalyptus subsets were kept at 50 samples each, while the subset of pine was enlarged with 50 samples per iteration until all the available samples were utilised ( $s = 3000$ ). Similarly, the training subsets for acacia and pine, and then eucalyptus and pine, were kept at 50 samples each in *Experiments F* and *G* respectively. These experiments assessed the effect when two classes (genera) are severely under-sampled.

The purpose of *Experiment H* was to assess the effect of area-proportional training datasets on classification accuracies. As with the other experiments, *Experiment H* starts with 50 randomly selected points (sampled from  $SS_0$ ). In Study Area 1 (WC), the sample size was iteratively enlarged by 0.02% for acacia, 0.03% for eucalyptus, and 0.56% for pine samples until 4% ( $s \sim 3n$ ), 6% ( $s \sim 5n$ ), 100% ( $s \sim 81n$ ) of the acacia, eucalyptus, and pine samples were selected. In Study Area 2, *Experiment H* equates to *Experiment A* given that the planted areas are more or less equal.

### 3.3.4 Accuracy assessment

An independent test sample set of 100 points per genus were used to calculate the OA, kappa statistic (KS), user's accuracy (UA), producer's accuracy (PA), indices of disagreement (Pontius & Millones 2011), and a McNemar's test. The OA measures the percentage of pixels that are correctly classified and the KS measures the chance agreement between the reference and classified maps. The PA and UA are used to quantify the performance of each class. The PA shows the occurrence of features on the ground that are correctly shown on the classified map. The error of omission can be calculated by  $1 - PA$ . The UA shows the occurrence of the class on the map that will be present on the ground. The error of commission can be calculated by  $1 - UA$  (Foody 2002).

The RF algorithm was run 100 times per iteration to remove any random error due to the stochastic nature of the algorithm, and the mean OA, KS, UA, and PA were calculated from confusion matrices. The standard deviations of the 100 iterations of the OA and KS were also calculated to assess to what extent the calculated means are representative of the 100 iterations (Rodriguez-Galiano et al. 2012).

Indices of disagreement and McNemar's statistical test were calculated for seven of the sample increase iterations, namely for the initial iteration (50 samples), and then at every tenth iteration. Indices of disagreement were analysed to validate the statistics derived from the confusion matrix, and McNemar's test was used to analyse whether the differences between the OAs were statistically significant.

### 3.3.5 Spectral analysis

The reflectance values of the samples ( $SS_0$ ) were used to develop a spectral profile of each genus within each study area to assist with the interpretation of the results. A pair-wise Jeffries-Matusita (J-M) distance separability analysis was carried out in R-studio using the J-M built in package to better understand inter-class variations. The J-M distance quantifies the average distance between two classes in feature space based on a density function (probability distribution) of each class (Mahdianpari et al. 2019). Both mean and variance are considered in the distance calculations. The J-M distances range from 0 to 2, where 0 represents a low separability between classes and 2 represents a high separability between classes.

## 3.4 RESULTS

### 3.4.1 Genus spectral profiles

The spectral properties of the genera, as extracted from the S2 imagery and all  $SS_0$  samples, are shown in Figure 3-4. The spectral properties of trees are dependent on many factors, including age, species composition and location. The spectral signatures of the acacia and pine classes are very similar, especially in Study Area 2 (KZN). Although it seems that there are some differences in the mean values in bands B4, B6 to B8, and B11 to B12, there is a large spectral overlap between the classes (indicated by the error bars), which suggests that the genera are spectrally similar. This is confirmed by the J-M distance scores (shown as bars in Figure 3-4), where eucalyptus and pine are the most separable, followed by acacia and eucalyptus, and to a lesser degree acacia and pine. The highest separability (0.8) in Study Area 1 (WC) is between eucalyptus and pine in B12 (Figure 3-4a). These two genera are consistently the most separable in this region, which corresponds well with the spectral profiles (mean reflectance values of eucalyptus and pine are most distinct in almost all bands).

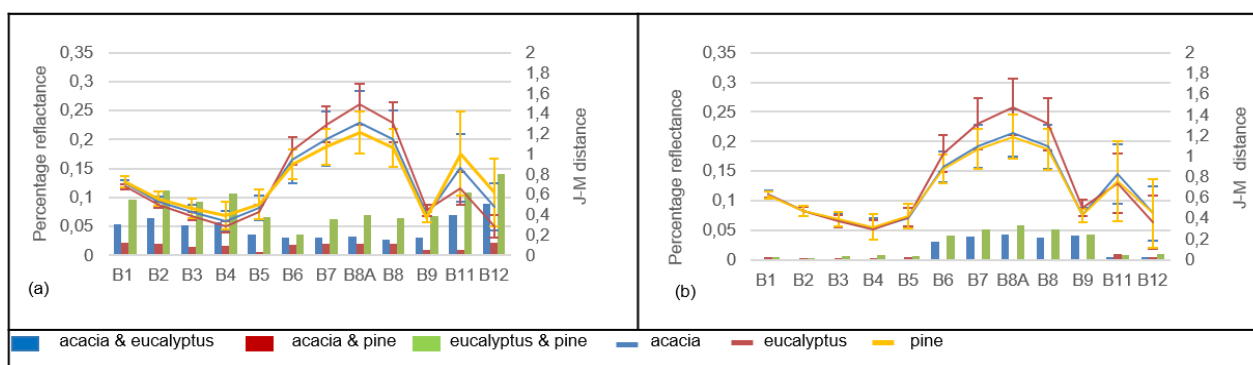


Figure 3-4 Average spectral signatures of the training samples, the standard deviation of the spectral signatures (shown as error bars), and the J-M separability score for all the training samples for (a) Study Area 1 (WC) and (b) Study Area 2 (KZN)

In Study Area 1 (WC), the visible (B2, B3, and B4) and shortwave infrared (SWIR) (B11 and B12) region of the electromagnetic spectrum (EMS) provides higher separability scores (0.4-0.7) than the red edge region (0-0.3) of the EMS (Figure 3-4a).

In contrast, all three genera in Study Area 2 (KZN) have low separability in the visible and SWIR region of the EMS. The strongest separability scores (0-0.33) were recorded in the red edge (B6 and B7), NIR (B8 and B8A) and water vapour (B9) bands (Figure 3-4b). Eucalyptus and pine trees are most separable, followed by acacia and eucalyptus.

When the spectral profiles are compared, it is clear that pine and acacia trees have very similar spectral responses, while those of eucalyptus trees are more distinct. It is also clear that the separability scores of the genera vary dramatically between the two study areas.


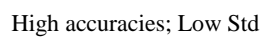
### 3.4.2 Classification

Table 3-4 summarises the classification results using the traditional confusion matrix. The results showing the indices of disagreement are shown in Appendix A and are not discussed as they coincide with the OA and KS metrics (Table 3-4). The mean OA of all experiments in Study Area 1 (79.8%) was significantly ( $p > 0.001$ ) higher than that of Study Area 2 (76.7%). From *Experiment A* it is evident that larger training sample set sizes ( $s > n$ ) resulted in higher OAs. In Study Area 1 (WC), the highest mean OA (76.3%) and best individual classification result (Max OA = 81.3%) were achieved when all (100%) of the available samples ( $SS_0$ ) were used for training. However, the improvement in accuracy was not linear (Figure 3-5: A1), with the OA increasing by only 2% when the sample size was enlarged from 67% ( $s \sim 54n$ ) to 100% ( $s \sim 81n$ ). The same trend can be observed in Study Area 2 (KZN) (Figure 3-5: A4), although fewer (50%) training samples were required for achieving maximum OA (70.7%).

When the PAs and UAs are considered in Study Area 1 (WC) (Figure 3-5: A2 and A3), it is evident that the eucalyptus class was the most accurately classified, as it has the lowest errors of omission and commission, followed by pine and then acacia. In Study Area 2 (KZN), the PAs and UAs of the three genera are relatively similar.

Table 3-4 A summary table showing the overall accuracy, the standard deviation of the overall accuracy, kappa statistic, the standard deviation of the kappa statistic, consumer's and user's accuracy and the maximum OA and KS of the 100 iterations per sample size of experiments A to G conducted on Study Area 1 (WC) and Study Area 2 (KZN)

Experiment	Iteration	Study area 1 (WC)												Study Area 2 (KZN)													
		Consumers Accuracy						Producers Accuracy						Max OA	Max KS	Consumers Accuracy						Producers Accuracy					
		OA	OA Std	KS	KS Std	Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine	Acacia	Eucalyptus			Pine	Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine	Max OA	Max KS			
	1	0.586	0.038	0.380	0.056	0.538	0.628	0.583	0.371	0.767	0.621	0.703	0.555	0.459	0.046	0.189	0.069	0.458	0.487	0.440	0.449	0.493	0.435	0.553	0.330		
Exp. A (Even samples)	20	0.715	0.020	0.573	0.031	0.665	0.772	0.705	0.614	0.802	0.729	0.763	0.645	0.626	0.025	0.439	0.037	0.622	0.638	0.619	0.611	0.652	0.617	0.666	0.499		
	40	0.744	0.020	0.616	0.031	0.697	0.803	0.731	0.655	0.828	0.748	0.800	0.700	0.656	0.015	0.483	0.022	0.660	0.664	0.644	0.645	0.674	0.649	0.683	* 0.525		
	60	* 0.763	0.023	0.644	0.034	0.723	0.818	0.747	0.684	0.840	0.764	* 0.813	0.720	* 0.674	0.011	0.511	0.017	0.675	0.681	0.667	0.669	0.685	0.669	0.680	0.520		
	Mean	* 0.725	0.022	0.588	0.033	0.676	* 0.780	0.718	0.625	* 0.816	0.734	0.770	0.655	* 0.618	0.020	0.427	0.030	0.622	* 0.629	0.606	0.615	* 0.624	0.616	0.670	0.505		
Exp. B (50 acacia)	20	0.600	0.015	0.400	0.022	0.700	0.661	0.545	0.061	0.880	0.858	0.637	0.455	0.498	0.025	0.247	0.037	0.638	0.526	0.467	0.069	0.699	0.727	0.560	0.340		
	40	0.610	0.012	0.415	0.018	0.716	0.677	0.550	0.043	0.907	0.879	0.650	0.475	0.506	0.020	0.259	0.030	0.641	0.542	0.470	0.039	0.733	0.745	0.557	0.335		
	60	0.608	0.013	0.413	0.019	0.740	0.677	0.548	0.036	0.910	0.880	0.630	0.445	0.510	0.018	0.265	0.027	0.572	0.545	0.478	0.026	0.746	0.756	0.563	0.345		
	Mean	0.606	0.017	0.408	0.025	0.683	0.661	0.556	0.073	0.895	0.849	0.639	0.458	0.499	0.024	0.249	0.036	0.585	0.531	0.468	0.076	0.699	0.723	0.560	0.340		
Exp. C (50 eucalyptus)	20	0.586	0.037	0.379	0.055	0.523	0.832	0.585	0.702	0.296	0.759	0.657	0.485	0.525	0.026	0.287	0.039	0.536	0.709	0.493	0.723	0.131	0.720	0.573	0.360		
	40	0.589	0.029	0.384	0.044	0.547	0.893	0.576	0.740	0.233	0.795	0.653	0.480	0.537	0.020	0.306	0.030	0.538	0.787	0.519	0.775	0.085	0.751	0.577	0.365		
	60	0.583	0.030	0.374	0.045	0.553	0.908	0.563	0.757	0.197	0.793	0.663	0.495	0.534	0.018	0.301	0.027	0.531	0.810	0.521	0.779	0.068	0.754	0.607	0.410		
	Mean	0.585	0.033	0.377	0.049	0.533	0.843	0.581	0.704	0.285	0.765	0.658	0.487	0.530	0.024	0.296	0.036	0.533	0.756	0.508	0.736	0.131	0.724	0.586	0.378		
Exp. D (50 pine)	20	0.638	0.025	0.457	0.037	0.544	0.682	0.900	0.783	0.838	0.293	0.693	0.540	0.527	0.025	0.291	0.038	0.521	0.520	0.651	0.727	0.745	0.111	0.577	0.365		
	40	0.651	0.021	0.476	0.031	0.557	0.701	0.946	0.822	0.877	0.253	0.703	0.555	0.531	0.023	0.297	0.034	0.531	0.523	0.672	0.765	0.759	0.070	0.590	0.385		
	60	0.650	0.024	0.475	0.035	0.552	0.717	0.956	0.838	0.894	0.218	0.700	0.550	0.539	0.021	0.309	0.032	0.540	0.528	0.733	0.780	0.776	0.062	0.587	0.380		
	Mean	0.639	0.025	0.459	0.037	0.550	0.686	0.890	0.781	0.851	0.287	0.699	0.548	0.526	0.024	0.290	0.036	0.521	0.521	0.655	0.731	0.734	0.113	0.584	0.377		
Exp. E (50 acacia & 50 eucalyptus)	20	0.514	0.041	0.271	0.062	0.547	0.787	0.432	0.100	0.504	0.938	0.573	0.360	0.415	0.041	0.123	0.062	0.516	0.692	0.371	0.130	0.194	0.921	0.527	0.290		
	40	0.487	0.036	0.231	0.054	0.529	0.815	0.413	0.067	0.424	0.970	0.550	0.325	0.402	0.031	0.103	0.046	0.543	0.751	0.364	0.102	0.151	0.953	0.493	0.240		
	60	0.450	0.034	0.176	0.051	0.507	0.811	0.390	0.051	0.328	0.972	0.513	0.270	0.390	0.022	0.085	0.033	0.549	0.778	0.358	0.068	0.129	0.972	0.470	0.205		
	Mean	0.506	0.035	0.260	0.052	0.527	0.781	0.436	0.110	0.481	0.928	0.546	0.318	0.414	0.034	0.122	0.050	0.538	0.702	0.374	0.143	0.189	0.911	0.497	0.245		
Exp. F (50 acacia & 50 pine)	20	0.542	0.026	0.313	0.040	0.609	0.478	0.709	0.200	0.955	0.470	0.620	0.430	0.434	0.032	0.151	0.047	0.560	0.391	0.611	0.159	0.917	0.225	0.507	0.260		
	40	0.513	0.023	0.269	0.035	0.584	0.444	0.768	0.144	0.976	0.419	0.570	0.355	0.411	0.029	0.116	0.043	0.579	0.374	0.635	0.107	0.957	0.168	0.490	0.235		
	60	0.507	0.026	0.260	0.039	0.574	0.437	0.778	0.116	0.983	0.422	0.560	0.340	0.389	0.027	0.083	0.040	0.560	0.364	0.595	0.088	0.968	0.111	0.450	0.175		
	Mean	0.531	0.028	0.297	0.042	0.596	0.471	0.726	0.177	0.956	0.460	0.583	0.375	0.423	0.030	0.135	0.045	0.558	0.388	0.614	0.154	0.912	0.205	0.482	0.223		
Exp. G (50 eucalyptus & 50 pine)	20	0.572	0.036	0.358	0.054	0.450	0.829	0.832	0.913	0.442	0.360	0.650	0.475	0.459	0.031	0.188	0.046	0.415	0.588	0.582	0.915	0.219	0.243	0.553	0.330		
	40	0.535	0.037	0.302	0.056	0.425	0.883	0.858	0.958	0.356	0.290	0.607	0.410	0.445	0.028	0.168	0.042	0.397	0.601	0.700	0.959	0.159	0.218	0.517	0.275		
	60	0.529	0.034	0.293	0.051	0.420	0.895	0.900	0.969	0.333	0.285	0.623	0.435	0.442	0.029	0.163	0.043	0.394	0.622	0.727	0.976	0.147	0.204	0.550	0.325		
	Mean	0.553	0.035	0.329	0.052	0.442	0.832	0.828	0.906	0.411	0.341	0.627	0.440	0.459	0.031	0.189	0.046	0.417	0.582	0.630	0.914	0.216	0.247	0.540	0.310		
Exp. H (area-proportional samples)	20	0.531	0.034	0.297	0.052	0.599	0.761	0.448	0.124	0.545	0.925	0.572	0.345	0.626	0.025	0.439	0.037	0.622	0.638	0.619	0.611	0.652	0.617	0.666	0.499		
	40	0.549	0.028	0.323	0.042	0.621	0.789	0.457	0.134	0.574	0.937	0.594	0.376	0.656	0.015	0.483	0.022	0.660	0.664	0.644	0.645	0.674	0.649	0.683	0.525		
	60	0.560	0.025	0.340	0.037	0.624	0.807	0.464	0.137	0.601	0.940	0.632	0.410	0.674	0.011	0.511	0.017	0.675	0.681	0.667	0.669	0.685	0.669	0.680	0.520		
	Mean	* 0.530	0.028	0.295	0.043	0.573	0.789	0.443	0.108	0.545	0.938	0.587	0.365	0.618	0.020	0.427	0.030	0.622	0.629	0.606	0.615	0.624	0.616	0.670	0.505		

Key: Low accuracies  High accuracies; Low Std  High Std; Overall accuracy (OA); Standard deviation (Std); Kappa statistic (KS); Maximum (Max); Important results (\* Bold text)

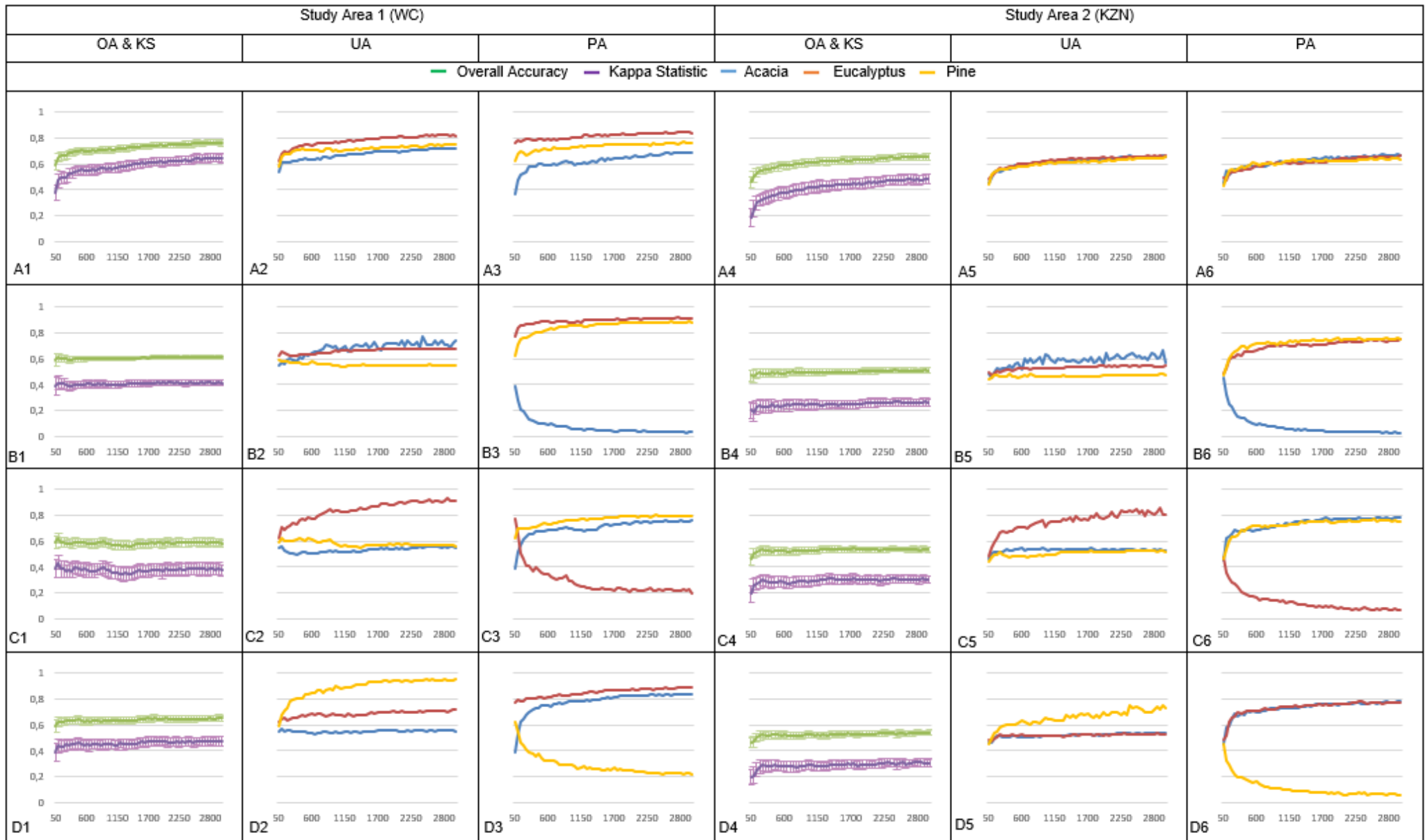


Figure 3-5 Trends of the overall accuracies, standard deviation of the overall accuracies for the 100 iterations, kappa statistics, standard deviation of the kappa statistics for the 100 iterations, user's accuracies, and producer's accuracies of experiments A to D for Study Area 1 and Study Area 2

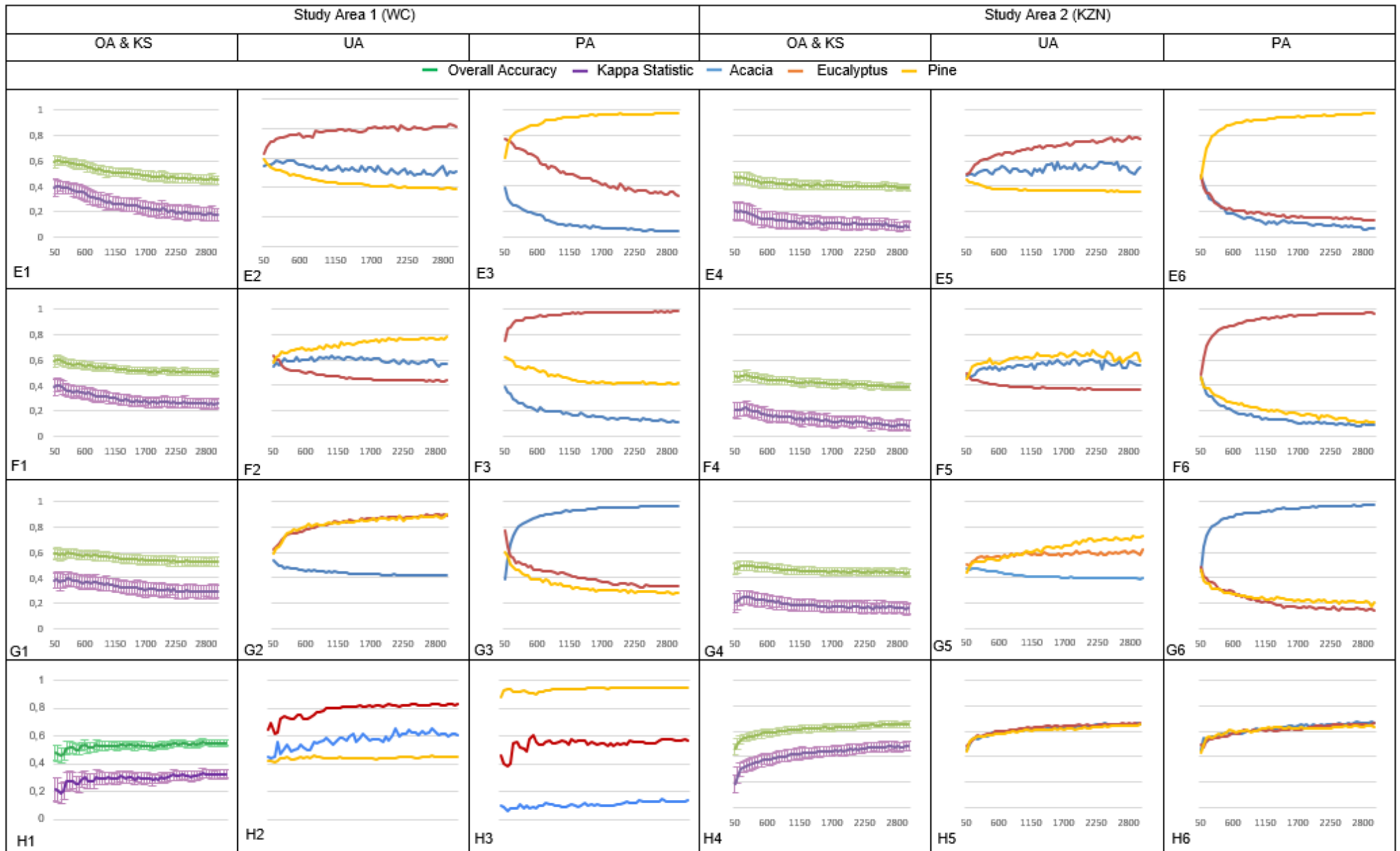


Figure 3-6 Trends of the overall accuracies, standard deviation of the overall accuracies for the 100 iterations, kappa statistics, standard deviation of the kappa statistics for the 100 iterations, user's accuracies, and producer's accuracies of experiments E to H on Study Area 1 and Study Area 2

The experiments dealing with imbalanced training sets (*Experiments B to D*) showed that the OAs and KSs remain relatively constant as sample imbalance increases. The UAs of the class for which the training subset was not enlarged (i.e. is kept at 50 samples throughout all iterations) are generally higher (Figure 3-5: B2, B5, C2, C5, D2, D5) than those of the two classes for which the training subsets were systematically enlarged, which indicates that the errors of commission (i.e. false positives) are relatively low for this class. Conversely, the PAs (Figure 3-5: B3, B6, C3, C6, D3, D6) of the genus for which the training set was kept constant is consistently lower compared to those of the other two genera.

In *Experiments E to G* (where two genera were kept at 50 samples throughout all iterations), the OA and KS decrease as the sample imbalance increases. The UA of the class for which the sample size was systematically enlarged is generally lower (Figure 3-6: E2, E5, F2, F5, G2, G5) than the two classes for which the training subset was kept constant (i.e. was kept at 50 samples throughout all iterations). Conversely, the PA (Figure 3-6: E3, E6, F3, F6, G3, G6) of the genus for which the training set was enlarged is consistently higher compared to the other two genera for which the training sets were kept constant.

In *Experiment H*, Study Area 2 (KZN) has consistently higher OAs and KSs than Study Area 1 (WC). The OAs and KSs for Study Area 1 (Figure 3-6: H1) are relatively low and stable when the sample size was enlarged in proportion to the area being mapped. Eucalyptus has the highest UA, followed by acacia and then pine (Figure 3-6: H2). The PA of pine is the highest, followed by eucalyptus and then acacia (Figure 3-6: H3). Conversely, the increases in OAs and KSs for Study Area 2 (Figure 3-5: H4) is initially dramatic, but then moderates as the sample sizes were enlarged. Additionally, the UAs and PAs among classes are relatively similar in this study area (Figure 3-6: H5, H6).

In summary, the results show that the best accuracies were achieved when using an equal number of training samples per class (mean OA of 72.2% for *Experiment A*). OAs generally increased as the training sample sets were enlarged, but only when training datasets were balanced. Comparatively, low classification accuracies were achieved when unbalanced training samples were used. A trend of low error of commission (i.e. false positives) and high errors of omission (false negatives) was noted for under-represented classes (those for which sample sizes are not systematically enlarged). Classification accuracies were generally higher (mean OA of 61.8%) in Study Area 2, where the area covered by the sampled genera were similar in proportions. However, accuracies were generally low (mean OA of 53%) when an area-proportional sampling strategy was used to differentiate the genera in Study Area 1, i.e. where the targeted area of interest (i.e. population from which sampling was done) was dominated by one class (pine).

### 3.5 DISCUSSION

Satellite RS, combined with machine learning, is effective for monitoring plantation forests (Lück 2018) and for differentiating forest plantation genera/species (Mngadi et al. 2019). However, machine learning requires labelled (in situ) data for model training and validation. The cost of collecting large quantities of such data can impede regional implementations. Given that classification accuracies are strongly related to the size of training datasets (Foody et al. 2006), operational solutions should ideally strike a balance between classification accuracies and training data collection efforts. However, very little is known about how training set size will impact forest plantation genera classification accuracies. It is also not clear whether the number of samples per class should be proportional to the targeted populations, as suggested by Colditz (2015), Millard & Richardson (2015), and Shetty (2019), or if an equal number of samples per class will be most effective, as suggested by Mellor et al. (2015) and Thanh Noi & Kappas (2017). A further complication is that the distributions and proportions of genera vary greatly from region to region – often with one or two dominating genera – which means that sampling strategies might have to be region-dependent.

In this study, we investigated the effect of different sampling schemes on the performance of the RF machine learning classifier for differentiating between three genera in two diverse regions – using 37 features extracted from a cloud-free Sentinel-2 composite image as input. Our findings agree with those of Myburgh & Van Niekerk (2014), who showed that higher (~90%) classification accuracies were achieved when the number of samples per class was increased from  $s \sim 0.5n$  to  $s \sim 2n$ , but that the relationship between the number of samples per class and accuracy is non-linear. This is attributed to the reduction of sparsity as samples are added until the training samples sufficiently represent inter- and intra-class spectral variations. Similarly, Heydari & Mountrakis (2018) found that classification accuracy increased by 2% to 6% when the training sample size was enlarged from  $s \sim 37n$  to  $s \sim 370n$  respectively, but accuracies tended to flatten out at  $s = 2218$  ( $s \sim 370n$ ). Enlarging the training set beyond this saturation point (Foody 2009) did not add any value to the classification. Based on our data (*Experiment A*), a saturation point was reached at  $s \sim 2100$  ( $s \sim 57n$ ) when the samples per class were balanced. This suggests that collecting more than  $57n$  training samples per class may be superfluous for our particular application.

Our results show that the RF classifier did not handle the unbalanced training datasets well. Generally, low errors of commission (i.e. false positives) and high errors of omission (false negatives) were noted for minority (under-represented) classes (i.e. those for which sample sizes were not systematically enlarged). Pixels belonging to the minority class were often erroneously allocated to the majority class. This agrees with Mellor et al. (2015) who found that majority



classes performed well at the expense of minority classes when the training sets were unbalanced. The RF classifier is an ensemble of tree classifiers, where each tree makes its decision on a set of rules derived from labelled features (training data) (Pal 2005) and a majority vote is used to determine the consensus class (Gislason, Benediktsson & Sveinsson 2006). Classes with larger training data sets tend to have fewer false negatives (omission errors) as a larger set of rules (splitters) can be constructed. In contrast, classes with limited training data are likely to have fewer false positives (error of commission) as insufficient data are available for constructing splitting rules.

It is clear from our results that spectral separability among classes also played a role in classification accuracies, with low separability classes generally requiring a larger proportion of samples compared to those with high separability (Mellor et al. 2015). In our application, the spectral properties of eucalyptus trees are the most distinct, which resulted in larger differences in UAs among classes when the sample size of eucalyptus was contained (C2 compared to B2 and D2 in Figure 3-5). The SWIR wavelengths are known to be absorbed by vegetation with high moisture content (Manna & Raychaudhuri 2020). This suggests that the water content varies between the plantation genera in Study Area 1 as water availability is scarce. In contrast, the genera in Study Area 2 have similar SWIR values, suggesting that the genera have similar water content as a result of more rainfall. Trees with similar water content will absorb and reflect similar amount of radiation in the SWIR region of the EMS.

Colditz (2015) and Shetty (2019) evaluated the effect of different sampling designs on ML classifications for land cover mapping and found that area-proportional samples resulted in the best accuracies. This agrees with Millard & Richardson (2015) who recommended randomly selected, area-proportional training sets. In contrast to these findings, an area-proportional sampling scheme performed poorly in our Study Area 1 (*Experiment H*). Study Area 1 is dominated by pine (ratios of acacia, eucalyptus and pine are about 1:1:26), which resulted in highly skewed training sets in *Experiment H*. According to Figure 3-5 H2 and H3, the pine class was over-classified (i.e. low UA, high error of commission), while the acacia and eucalyptus classes were under-classified (i.e. low PA, high error of omission). For example, the error of omission ( $EO=1-PA$ ) of the acacia class exceeded 86% throughout the experiment, even when  $s \sim 3n$  for this class. Contrastingly, in *Experiment A* (balanced samples) the error of omission for acacia was 37% when  $s \sim 2n$  (initialisation of *Experiment A*). This suggests that an area-proportional sampling approach is not effective when the populations (total number of pixels) of some classes are severely skewed (i.e. where some classes dominate, while others are under-represented). Based on our results, a balanced training set (*Experiment A*) produced the most consistent (stable) per class (i.e.

PAs and UAs) results for both the skewed (Study Area 1) and equivalent (Study Area 2) populations. This agrees with the findings of Foody & Mathur (2004), Mellor et al. (2015), and Thanh Noi & Kappas (2017).

Our results show that OAs were often not a good reflection of overall class performance, particularly when training data were unbalanced. For instance, it is noticeable that the mean OAs of *Experiment D* in Study Area 1 is 64%, while it is clear that pine has a low mean PA (29%) and acacia has a low UA (55%). This demonstrates that OA is an unreliable measure of accuracy, particularly when the training data are skew. In contrast, the mean KS of *Experiment D* in Study Area 1 is low (0.025), which better depicts the overall performance of the experiment. Similarly, the OAs of *Experiment B – G* are relatively high, while they resulted in varying UAs and PAs. The relatively low KSs of these experiments are consequently better indicators of the overall classification performance and they corresponded to indices of disagreement (see Appendix A for a summary showing the indices of disagreement). This agrees with Viera & Garrett (2005) and Thanh Noi & Kappas (2017), who observed that OAs were often deceiving when training datasets were unbalanced.

Although the purpose of this study was not to assess the efficacy of machine learning for genus classification, our findings show that accuracies exceeding 80% can be achieved when a composite Sentinel-2 image (and its derivatives) is used as input. Better results are likely when the methodology is extended to incorporate multi-temporal imagery, data fusion (e.g. the combination of SAR data), feature selection/extraction, and alternative machine learning techniques. Operationalising such approaches will enable more frequent updates of national forest inventories at a reduced cost compared to traditional (field survey) methods. However, this study showed that training sample design significantly affects classification results and that data collection efforts must be carried out accordingly.

Although this study demonstrated potential for classifying the main forest plantation genera in two very diverse study areas (WC and KZN), more work is needed to evaluate the transferability of machine learning models to other regions. Our study showed that the sample set should be ~ 57n within each area being mapped, but we did not assess whether samples can be extended to classify genera in other regions where no samples are available. Signature extension and model transferability should be investigated in future work as they could reduce the costs of in situ training data collection.

### 3.6 CONCLUSION

Although it is known that the accuracy of machine learning approaches for classifying remotely sensed imagery is affected by training data size and configuration, little is known about how different sampling strategies may affect the differentiation of forest plantation genera. This study evaluated the effect of different training dataset configurations on RF's ability to differentiate between acacia, eucalyptus, and pine trees in two diverse regions within South Africa. The RF algorithm was implemented using Sentinel-2 bands, as well as several spectral and textural indices, as input features ( $n = 37$ ). In situ forest plantation data were used to generate various sample sets, which were used to train and assess the ability of the RF classifier to differentiate among genera.

Our findings show that although spectral separability of classes affected the ability of RF to accurately differentiate among forest plantation genera, classification accuracies were mainly influenced by training sample size and imbalance. Balanced training datasets produced the most accurate and consistent results, while unbalanced training data did not work well for differentiating genera using RF. It is clear that sampling scheme design is critical for regional forest genera mapping implementations and that a balanced approach is needed for in situ (labelled) data collection efforts. Although the purpose of the study was not to assess the value of Sentinel-2 imagery for genus classifications, our results suggest that such imagery, combined with machine learning, holds much potential for this purpose. Future research should consider a multi-temporal approach to exploit the phenological differences among genera. Model transferability and/or sample extension should also be investigated to minimise sample collection efforts. Answering these questions may lead to operational solutions for mapping forest plantation genera at regional/national scales and for regularly updating forest inventories. Accurate and up-to-date inventories will allow for improved land and forest management, water use assessments, carbon stock estimates and streamflow reduction modelling.

## **CHAPTER 4: SIGNATURE EXTENSION AS A MACHINE LEARNING STRATEGY FOR MAPPING PLANTATION FOREST GENERA WITH SENTINEL-2 IMAGERY**

### **4.1 ABSTRACT**

Plantation forest inventory data such as tree genus is important for supporting forestry decisions and policymaking, but such data are traditionally collected in-field, which is time-consuming and costly. Although machine learning and remote sensing technologies have shown great potential to reduce the time and effort for mapping forest plantation genera at regional scales, they rely on training (labelled) data which are also costly to collect over large areas. One approach to reducing the effort of training data collection is to make use of signature extension, whereby training data collected in one area is used to train and apply a machine learning model in a different area. This study aimed to evaluate the viability of training data signature extension for constructing random forest (RF) machine learning models to differentiate between acacia, eucalyptus and pine trees using Sentinel-2 imagery as input. The study was carried out over a large area (about 4 920 km<sup>2</sup>) in South Africa. The study area was divided into 19 tiles of 100 x 100 km (each tile coincides with the footprint of a Sentinel-2 tile) from which three were chosen for sourcing (collecting) training data. Four experiments were conducted. In the first experiment, a fixed number (3 000 per genera) of training samples were collected in the first source tile and used to build an RF model. The resulting model was then applied and assessed in all 19 tiles. This protocol was repeated in the second and third experiments using the training data collected in the other two source tiles respectively. In the final experiment, training data from all three source tiles were combined and applied to all 19 tiles. The mean overall classification accuracy of each classified tile was compared to the extension distance (i.e. the distance between the target and source tile), differences in rainfall seasonality, and variation in the mean annual temperature among tiles to gain an understanding of how signature extension efficiency is influenced by distance and environmental conditions. The results show that signature extension is viable (~ 70% overall accuracies) over distances of up to 500 km, but only if the source and target tiles represent areas with similar rainfall regimes.

### **4.2 INTRODUCTION**

Forest plantation inventory data includes the location, extent, planting date, tree species/genus/clone, water use, and yield at individual compartmental levels (Mati & Dawaki 2015). Such data are fundamental for forest management, which involves planning, land management, analysing growth rates to maximise production, assessing water use, monitoring

rotations and harvests. Forest inventories are useful to monitor the status of forests, identify trends in the forestry industry, model climate change, carry out hydrological modelling, quantify forestry's contribution to alleviating poverty, and inform trade and policy decisions at a national scale (Food and Agriculture Organization 2007). It is therefore vital to maintain and continuously update national forest inventories. Although commercial forestry companies maintain detailed inventories of the forests under their management, such information is often not publicly available as competitors may use it to gain a commercial advantage. Many small-scale timber farmers do not keep detailed inventory information (FSA 2019).

Inventory data, such as the genus of trees planted, are usually collected in-field, which requires manpower and time (Food and Agriculture Organization 2015). RS and image classification offers a more efficient way for collecting such information, especially over large regions.

Machine learning algorithms and medium resolution (30 m – 250 m) satellite imagery have been used to map forest species with varying success. For instance, Franco-Lopez, Ek & Bauer (2001) used SPOT (20 m) and Landsat (30 m) imagery with machine learning algorithms to map three forest species with an accuracy of 47%, while Stabach et al. (2009) classified 13 species with a 63% overall accuracy (OA). Higher accuracies have been achieved using very high-resolution (VHR) imagery (Cho, Malahlela & Ramoelo 2015; Francois & Leckie 2006; Immitzer, Atzberger & Koukal 2012; Ke, Quackenbush & Im 2010; Pu & Landry 2012), hyperspectral imagery (Buddenbaum, Schlerf & Hill 2005; Bujang & Baharum 2017; Fagan et al. 2015; Peerbhay, Mutanga & Ismail 2013; Voss & Sugumaran 2008), and/or unmanned aerial vehicle (UAV) data (Franklin, Ahmed & Williams 2017; Franklin & Ahmed 2018). However, UAV, VHR and hyperspectral imagery are not suitable for mapping at a national or regional scale due to the imagery being costly and the need for sufficient computing power for processing the very large volumes of data. High-resolution imagery (10-60 m) such as Sentinel-2, is freely available, can be processed on the cloud through platforms like Google Earth Engine (GEE), and has a high (5-day interval) temporal resolution, making it ideal for mapping over large areas.

Nomura & Mitchard (2018) mapped seven forest plantation species using supervised classification with Sentinel-2 data in Myanmar, achieving a 95% OA. Mngadi et al. (2019) used Sentinel-2 bands in a linear discriminative analysis (LDA) to classify seven species in the Clan forest plantation (located in South Africa) and obtained an 84% OA. They found that adding Sentinel 1 VV and VH features increased accuracies by 3%.

Image classification is generally done by collecting training data in one image and applying the classifier to that same image scene (Knorn et al. 2009). This is problematic when classifying large areas that are covered by multiple image scenes, as the training data collection becomes time-

consuming and costly. In machine learning modelling, signature extension or generalisation reduces the effort and cost of training data collection (Pax-Lenney et al. 2001). Signature extension is the process of training a model on one image scene and applying it to other image scenes (Laborte, Maunahan & Hijmans 2010).

Pax-Lenney et al. (2001) classified Landsat images into conifer forests across time and space using artificial neural networks (ANN). The ANN was trained using in situ data obtained from one image and applied to other images. There was a mean decline of 8 to 13% in OA when the model was trained with a raw image and applied to images that were atmospherically corrected. It was concluded that the atmospheric correction algorithm used influences the accuracy of a model when signature extension is applied. Olthof, Butson & Fraser (2005) classified ten Landsat scenes into land cover classes using a k-means classifier to initially create sample cluster signatures in one reference Landsat scene. The sampled cluster signatures were used to classify all other Landsat scenes by using a minimum distance classifier to produce similar clusters to those in the reference scene. This was done using two radiometric correction algorithms on the Landsat scenes to evaluate whether they affect the OA when using signature extension to map land cover. However, both sets of images produced an OA of ~71% and it was concluded that the accuracy is influenced more by extension distance (distance between training data and the area being classified) than the radiometric correction algorithm used. Knorn et al. (2009) also found that an increased extension distance decreases the OA of the classification. Instead of creating sample cluster signatures, they created a chain classifier whereby a support vector machine (SVM) algorithm was trained and applied on the same Landsat scene. The part of the classified image that overlaps with a second Landsat scene formed the training data for the second scene. This process is repeated to form a “chain” classifier until all the Landsat scenes in the chain are classified. An average OA loss of 1.9% was reported per neighbouring scene. Verhulp & Van Niekerk (2017) tested the transferability of CART for land cover classification. The model was trained on four mosaicked coastal Landsat scenes and applied to two inland scenes. The OA of the four coastal scenes was 80.6%, while the two inland scenes had OAs of 61.4% and 83.6% respectively. Similar to Olthof, Butson & Fraser (2005) and Knorn et al. (2009), they concluded that accuracies decrease as the distance increases from the training scene (extension distance).

Woodcock et al. (2001) mapped forest change using signature extension by training an ANN with data from the Rocky Mountains and applied it to Cascades of Oregon and visa-versa, achieving a 15.3% OA. They concluded that ecological/topographic/climatic differences affect the effectiveness of signature extension more than atmospheric correction techniques and extension distance. Similarly, Phalke & Özdoğan (2018) found that agro-ecological conditions affect the

generalisation (transferability) of machine learning models when mapping croplands vs non-crop lands. Wang, Azzari & Lobell (2019) trained an RF classifier using data in one state and applied the model to eight other states. The results showed that the accuracies consistently exceeded 80% when the model is trained and applied in areas with similar growing degree days.

It is clear from the literature that signature extension is a viable approach for reducing the costs associated with collecting labelled data to train machine learning algorithms. However, many factors influence its effectiveness and it is not known whether signature extension is a viable strategy for mapping forest plantation genera over large regions. There is often a trade-off between extension distance and classification accuracy, while climatic variations can also negatively influence signature extension efficiency. These factors may outweigh the benefits of signature extension in complex regions such as South Africa, where forest plantations are sparsely distributed and where climate gradients are dramatic (range from subtropical summer rainfall to semi-arid winter rainfall). A further complicating factor is that genera are planted in unequal proportions throughout South Africa, with pine trees being favoured in the Western and Eastern Cape provinces, while most plantations in KwaZulu-Natal and Mpumalanga are planted with eucalyptus trees.

This study aims to quantify and map the relationship between genus classification accuracy and training sample cluster distribution. The RF classifier is used to classify forest plantations covered by a set of 19 Sentinel-2 tiles of 100 x 100 km in size—distributed throughout the forestry regions of South Africa and representing eight Köppen-Geiger climatic zones—into three target genera (pine, eucalyptus, and acacia). Several experiments are carried out. In the first experiment, all 19 tiles are classified using a model trained with samples collected in a tile with a mid to late summer precipitation pattern and a mean summer temperature of 18°C (Köppen-Geiger climatic zone Cwb). In the second experiment, training data from a tile with an early, mid, and late summer precipitation pattern and a mean summer temperature of 20°C (Köppen-Geiger climatic zones Cwb, Cwa, Cfb, Cfa) are used, while the third experiment uses training data from a tile that experiences precipitation throughout the year, but with a wetter winter and a mean summer temperature of 19 °C (Köppen-Geiger climatic zones Cfb, CSk, BWk). In the final experiment, training data from all three source tiles are used to classify all 19 tiles. An independent set of in situ data is used to quantify the classification accuracy in each of the 19 tiles. The relationship between classification accuracy, extension distance and rainfall seasonality is quantified and interpreted within the context of optimising in situ data collection efforts for operational mapping of plantation genera at a national scale.

### 4.3 METHODS AND MATERIALS

#### 4.3.1 Study area

The study area covers an area of ~ 4 920 km<sup>2</sup> and includes plantations located in Mpumalanga, KwaZulu-Natal (KZN), EC, and Western Cape (WC) provinces of South Africa (Figure 4-1). These sites were chosen due to the genera planted and the availability of in situ data. These sites are also representative of the commercial forest plantations in South Africa.

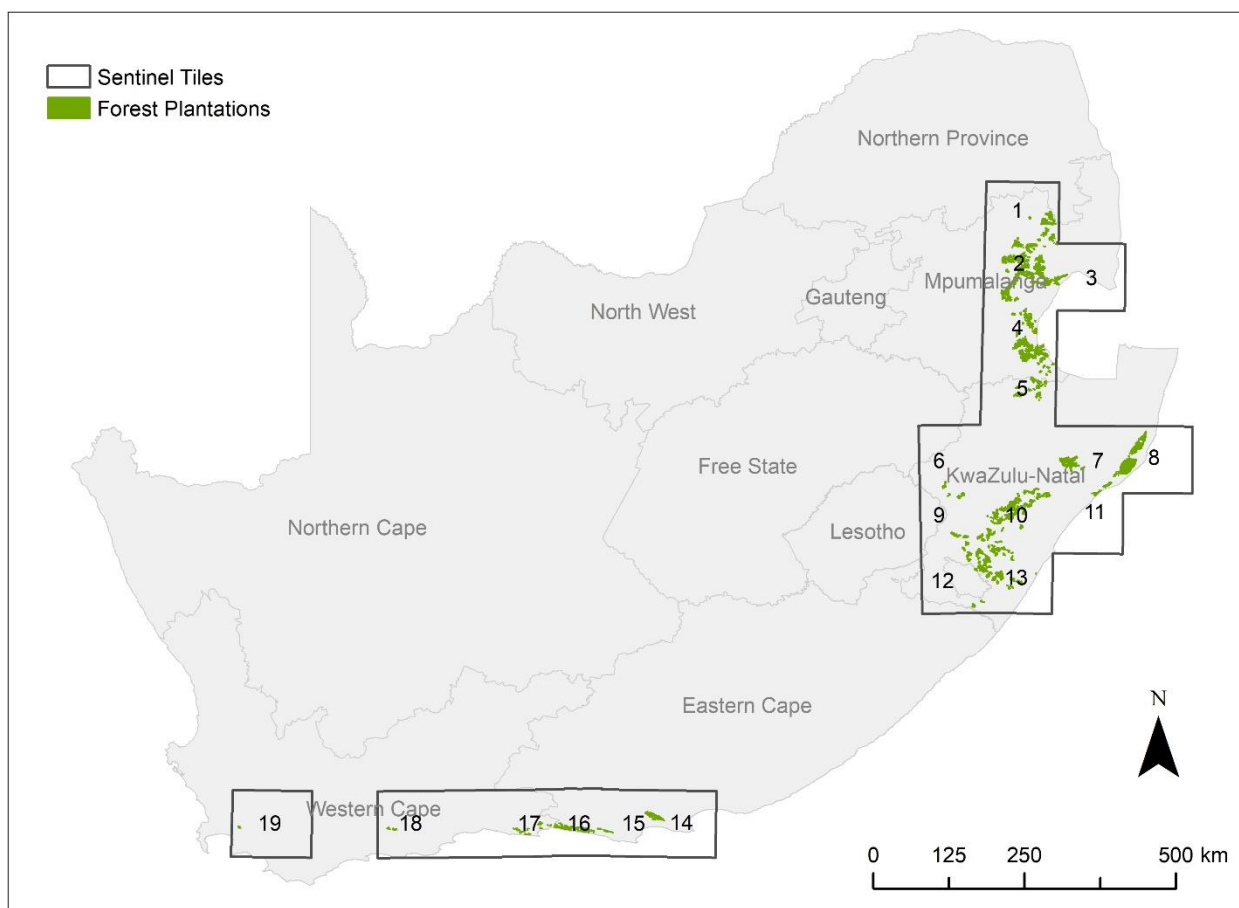


Figure 4-1 Location of compartments used in the study

The South African rainfall patterns vary from winter rainfall in the southwest to early summer rainfall in the northeast. Regions with mid-summer rainfall, late summer rainfall, very late summer rainfall, and a bimodal to summer rainfall are also present (Figure 4-2). The WC experiences winter rainfall, while the rest of the country generally experiences summer rainfall (Botai, Botai & Adeola 2018).



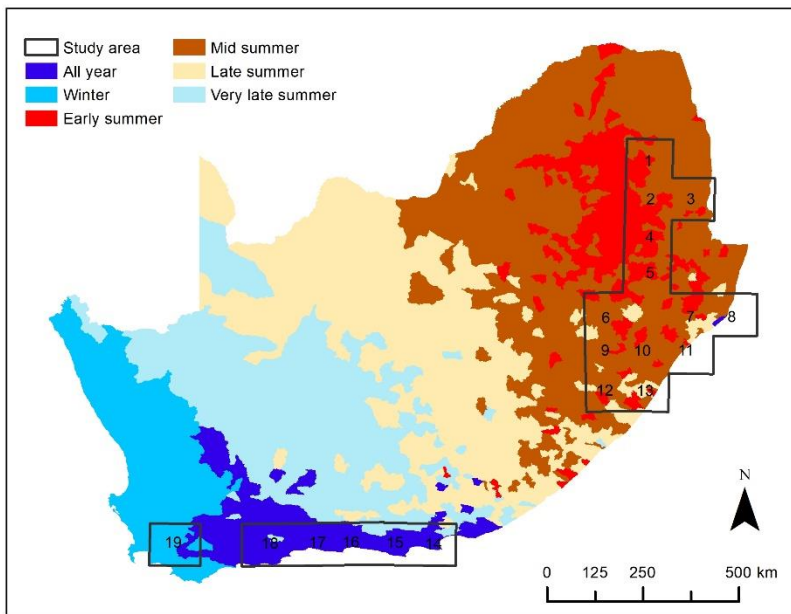


Figure 4-2 Rainfall seasonality in South Africa

Most (80%) of the plantations are located in Mpumalanga and KZN, with each being home to about 40% of the total plantations, while the EC represents about 11% of the total commercial plantations. The remainder are located in the WC, particularly along the south coast. The variation in rainfall and temperatures across South Africa are represented by no less than 13 Köppen-Geiger climate zones (Schulze 1947) (Figure 4-3), while the targeted forest plantations are located the nine listed in Table 4-1.

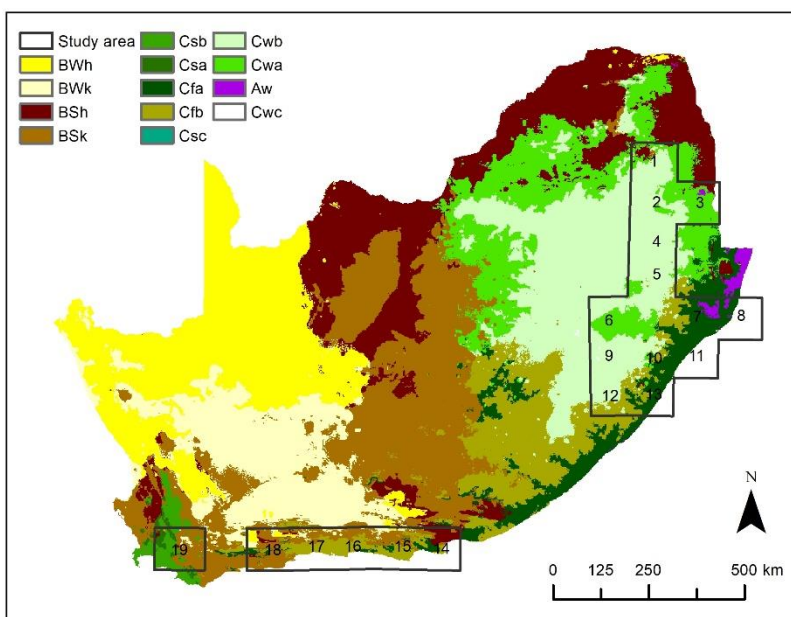


Figure 4-3 Köppen-Geiger climate zones of South Africa

Table 4-1 Description of the Köppen-Geiger zones in South Africa

Köppen-Geiger zone	Conditions
BSh	Semi-arid, hot and dry
BSk	Semi-arid, cool and dry
Csb	Summers long, dry and cool
Csa	Summers long, dry and hot
Cfa	Wet all seasons, summers long and hot
Cfb	Wet all seasons, summers long and cool
Cwb	Winters long, dry and cool
Cwa	Winters long, dry and hot
Aw	Tropical wet, dry winter season

From this regional overview, it should be clear that the classification of South Africa's forest plantation genera poses a unique challenge as plantations are sparsely distributed over a large region. Collecting in situ genera data (for training machine learning models) over such a large region is prohibitively expensive. Although signature extension is a viable option, its effective implementation in a region with dramatic climatic variations will likely depend on the quantity and distribution of the training samples used for model building.

### 4.3.2 Data collection and preparation

#### 4.3.2.1 Imagery

Sentinel-2 level-2A satellite imagery was accessed through the GEE platform. Sentinel-2 imagery contains 13 spectral bands and has a five-day temporal resolution. Four bands (B2, B3, B4, B8) have a spatial resolution of 10 m, while six bands (B5, B6, B7, B8A, B11, B12) and three bands (B1, B9, B10) have a spatial resolution of 20 m and 60 m respectively. Median pixel values of the images dated 2019-06-30 to 2020-06-30 were calculated to form a composite image. The composite image was produced to remove cloud contamination and to compensate for seasonal variations among the plantations. Although it is known that multi-temporal earth observation approaches and phenological variations can aid in genus classifications (Mngadi et al. 2019), such variations were purposefully excluded from our experiments as seasonal variations may benefit some training sets more than others, which in turn will add complexity to the findings; i.e. using a single composite image allows for direct comparisons. The assumption is that a multi-temporal approach will improve classification accuracies, but testing this assumption is outside the scope of this study.

The normalised differential vegetation index (NDVI), enhanced vegetation index (EVI), entropy, angular second moment, contrast, dissimilarity, and difference variance were derived and added as bands to the composite image, resulting in 18 features.

#### 4.3.2.2 In situ data

In situ (ground-truthed) data at a plantation compartment level were collated from several South African commercial forestry companies. This data included records of the species, genus, and planting date of the trees in each compartment. Table 4-2 provides an overview of the in situ sample data per genus within each province.

Table 4-2 Summary of in situ data (forest compartment information) collated, including tree genus and planted area per province

	Acacia (ha)	Acacia (%)	Eucalyptus (ha)	Eucalyptus (%)	Pine (ha)	Pine (%)	Total (ha)
MP	1429	0.65	117543	53.83	99369	45.51	218341
KZN	4832	2.02	199035	83.23	35277	14.75	239144
EC	102	0.39	220	0.85	25650	98.76	25972
WC	118	1.44	181	2.21	7902	96.35	8201
Total	6482	1.32	316980	64.47	168198	34.21	491659
Sample size	30	0.46	30	0.01	30	0.02	90

Key: Hectare (ha); Mpumalanga (MP); KwaZulu-Natal (KZN); Eastern Cape (EC); Western Cape (WC)

KZN contains the most sampled acacia and eucalyptus plantations, while most pine compartments are located in Mpumalanga. The sampled compartments in the EC and WC are sparser.

#### 4.3.3 Experimental design

The in situ genus data were consolidated into a shapefile and geographically split by Sentinel-2 tiles. The resulting groups of genus polygons were labelled 1 to 19 (Figure 4-1).

The total area covered per genus per tile was calculated to identify the tiles with the most evenly distributed genera (Table 4-3). Tiles 4 and 10 were selected as source tiles (for model training) as they contained a relatively large percentage of acacia compartments, as well as a sufficient number of eucalyptus and pine compartments. Tile 17 was selected as the third source tile, as it contained the most acacia in the southwestern part of the country.

Table 4-3 Summary of in situ data (forest compartment information) collated, including tree genus and planted area per block

Block ID	Acacia		Euc		Pine		Total Area (ha)	Area (m2)
	Area (ha)	#	Area (ha)	#	Area (ha)	#		
1	0.00000	0	24085.22280	2613	11252.55382	844	35337.77661	353377766.1
2	1.59898	1	40744.05417	2519	71008.31219	3422	111753.9653	1117539653
3	0.52017	1	1145.12945	112	6466.84901	524	7612.498628	76124986.28
4	1411.15116	117	53764.61090	3990	21706.04727	1374	76881.80934	768818093.4
5	249.56682	34	17628.11371	1435	2634.10616	270	20511.78668	205117866.8
6	0.00000	0	2066.87156	155	0.00000	0	2066.871556	20668715.56
7	1249.37819	101	26784.86371	1787	3710.84429	262	31745.08619	317450861.9
8	6.66984	1	54811.62499	3199	253.17720	60	55071.47203	550714720.3
9	4.75796	1	18903.90321	1580	11091.76549	916	30000.42666	300004266.6
10	2647.13672	316	54675.70224	4825	22487.42953	1965	79810.26848	798102684.8
11	312.62138	42	3609.25956	282	0.00000	0	3921.88094	39218809.4
12	406.73914	35	35880.90225	2202	6606.75598	521	42894.39737	428943973.7
13	720.11692	65	46153.98210	2808	4445.49181	326	51319.59083	513195908.3
14	1.14891	1	49.11812	19	9919.45114	1748	9969.718167	99697181.67
15	92.52951	24	140.05582	62	13139.63649	3137	13372.22182	133722218.2
16	111.73580	38	195.01783	81	17297.28591	4431	17604.03954	176040395.4
17	118.13468	40	145.44645	50	3044.90083	942	3308.481957	33084819.57
18	0.00000	0	12.03381	5	1894.22266	327	1906.256476	19062564.76
19	0.00000	0	1.86178	2	727.20226	246	729.064035	7290640.35

Key: Hectares (ha); number of compartment polygons (#)

Four experiments were set out to evaluate the viability of sample extension for classifying forest plantation genera. All experiments used 9 000 training samples (3 000 per genera). A stratified (per genus) random sampling scheme was used to generate samples, with the minimum distance between samples set to 10 m. The genus polygons from which the training samples were generated contain compartments of different tree ages and species. This variation, combined with the minimum distance between samples, reduced autocorrelation in the training data.

The Sentinel-2 band values were extracted at each training point (pixel) and four RF models were built and applied to each tile. The first experiment (Experiment 1) used training data from Tile 4, while the model in the second experiment (Experiment 2) was trained using samples from Tile 10. Experiment 3 was trained with samples from Tile 17. For comparison purposes, Experiment 4 used training data from all three tiles.

#### 4.3.4 Accuracy assessment

An independent test data set of 100 samples per genus in each tile was used to conduct the accuracy assessments. The OA, kappa statistic (KS), consumer's accuracy (CA), and producer's accuracy (PA) were deduced from confusion matrices. OA measures the percentage of pixels that are correctly classified and the KS measures the chance agreement between the reference and classified maps. The PA and UA are used to quantify the performance of each class. The PA shows the occurrence of features on the ground that are correctly shown on the classified map. The error of omission can be calculated by 1-PA. The UA shows the occurrence of the class on the map that will be present on the ground. The error of commission can be calculated by 1-UA (Foody 2002).

### 4.3.5 Spectral analysis

The reflectance values of the training samples were used to develop a spectral profile of each genus within each tile to assist with the interpretation of the results. A pair-wise Jeffries-Matusita (J-M) distance separability analysis was carried out to better understand inter-class variations. The J-M distance quantifies the average distance between two classes in feature space based on a density function (probability distribution) of each class (Mahdianpari et al 2019). Both mean and variance are considered in the distance calculations. The J-M distances range from 0 to 2, where 0 represents a low separability and 2 a high separability between classes.

### 4.3.6 Variable drivers

Several factors can influence the efficacy of signature extension. The geographical distance between the training samples and the scene being classified (i.e. extension distance) has been shown to influence accuracies (Olthof, Butson & Fraser 2005; Pax-Lenney et al. 2001), while Li et al. (2020), Phalke & Özdoğan (2018), Verhulp & Van Niekerk (2017), Wang et al. (2016) and Woodcock et al. (2001) showed that climatic variations can also affect the success of signature extension implementations. Given that the study area spans eight climatic regions, each with unique rainfall seasonality and temperature profiles, rainfall seasonality and annual mean temperature were examined along with extension distance to evaluate to what extent these factors influence the efficiency of signature extension.

The distance from the centroid of the source tile to the centroid of all other tiles was calculated. The long-term mean monthly and annual precipitation was obtained from Schulze (2007). A seasonality index was calculated according to Walsh & Lawler (1981):

$$SI = \frac{1}{Ri} \sum_{n=1}^{n=12} \left| X_{in} - \frac{Ri}{12} \right| \quad \text{Equation 4-1}$$

where  $Ri$  is the total annual long-term mean precipitation; and  
 $X_{in}$  is the long-term monthly mean precipitation for month  $n$ .

Similarly, the long-term mean monthly and annual temperatures were obtained from Schulze (2007) for each tile to calculate a temperature seasonality index:

$$TI = \frac{1}{Ri} \sum_{n=1}^{n=12} \left| X_{in} - \frac{Ri}{12} \right| \quad \text{Equation 4-2}$$

where  $Ri$  is the long-term mean annual temperature; and  
 $X_{in}$  is the long-term monthly mean temperature for month  $n$ .

The OAs (dependent variable) for Experiments 1, 2, and 3 were plotted against the rainfall seasonality, temperature seasonality and extension distance (independent variables) to establish if there was a relationship between the OAs and independent variables for each experiment.

In addition, the intra-class J-M distance score was calculated for each experiment using the training data and testing data for each class in every tile (i.e. the training data for acacia in the source tile vs the testing data for acacia in the classified tiles, the training data for eucalyptus in the source tile vs the testing data for eucalyptus in the classified tiles, and the training data for pine in the source tile vs the testing data for pine in the classified tiles, for every experiment). This was done to compare the spectral variations of genera among tiles and to assist in the interpretation of the results.

## **4.4 RESULTS**

### **4.4.1 Spectral profiles**

The spectral signatures extracted from each source tile (i.e. Tiles 4, 10 and 17) and all three source tiles in combination are shown in Figure 4-4. The spectral signatures of eucalyptus are consistently more distinct, compared to those of pine and acacia. The greatest dissimilarities among classes are noticeable in B6 to B8, although the standard deviation bars indicate large overlaps. The dissimilarity of eucalyptus to pine and acacia are confirmed by the J-M scores (shown as bar graphs in Figure 4-4). The acacia and pine classes consistently produced the lowest separability scores. The classes seem to be the most separable in Tile 17 (Figure 4-4c), followed by the classes in Experiment 1 (Figure 4-4a), then the classes in Experiment 4 (Figure 4-4d), and to a lesser degree the classes in Experiment 2 (Figure 4-4b).

Generally, the separability between the classes in the red edge region of the electromagnetic spectrum (EMS) is higher than in the visible and SWIR region of the EMS. In Tile 17, the separability in the visible and SWIR region of the EMS is higher compared to the other two source tiles. The SWIR region of the EMS is known to be absorbed by vegetation with high water content. Therefore, the reflectance in the SWIR region of the EMS will vary for trees located in water scarce areas – i.e. Tile 17 – as they retain different amounts of moisture.

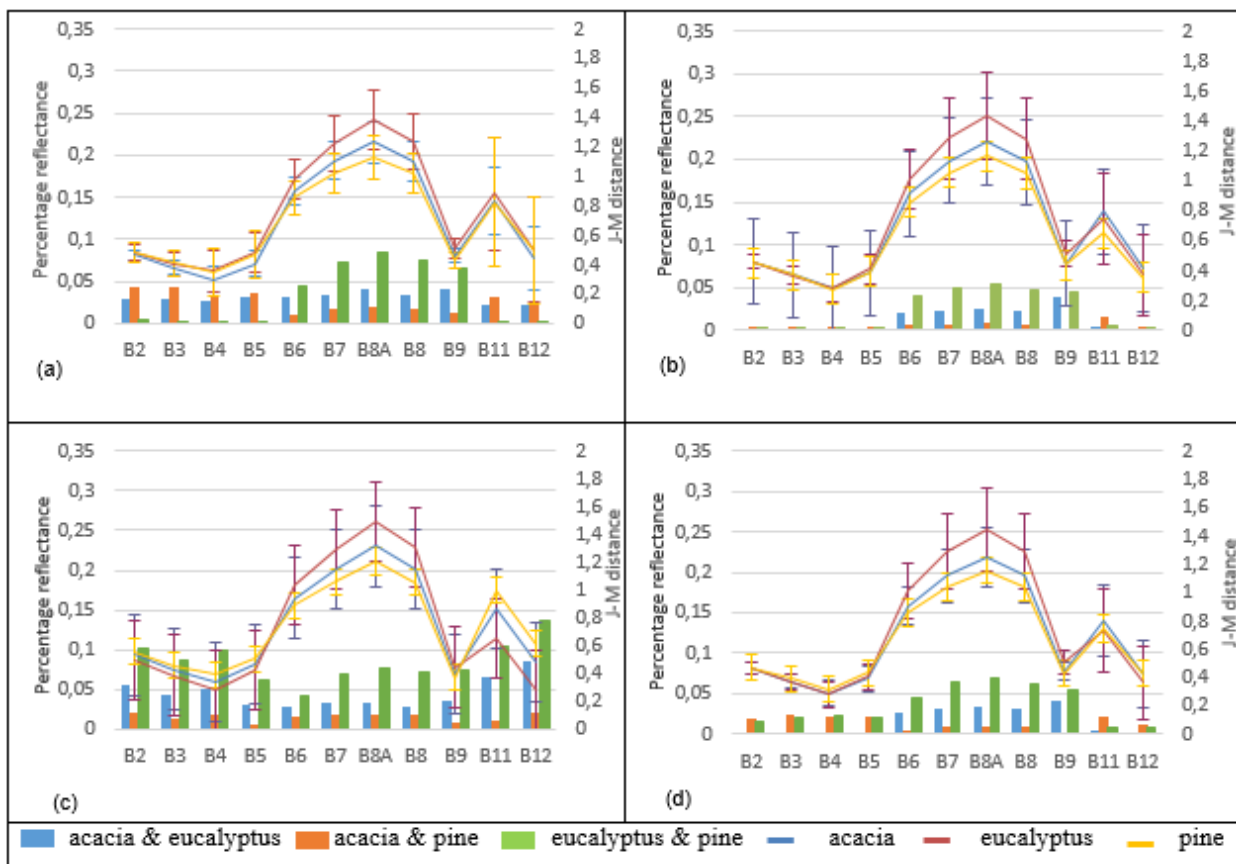


Figure 4-4 Average spectral signatures for the training samples and J-M separability score for Tile 4 (a), Tile 10 (b), Tile 17 (c), and all source tiles in combination (d)


#### 4.4.2 Classification

Table 4-4 summarises the classification results for each experiment. The mean OA was the highest in Experiment 4 (70%), followed by Experiment 1 (68%), Experiment 2 (62%), and Experiment 3 (52%). From Experiment 4, it is evident that collecting training data in a large number of climatic regions resulted in higher OAs. In Experiments 1 and 2, the mean OAs for Tiles 1 to 13, located in the summer rainfall regions, are higher than the mean OAs in Tiles 14 to 19, located in the winter and bimodal to summer rainfall regions. Similarly, in Experiment 3, the mean OAs for Tiles 14 to 19, located in the winter and bimodal to summer rainfall regions, are higher than the mean OAs in Tiles 1 to 13, located in the summer rainfall regions.

In general, the pine class was classified with higher PAs than the eucalyptus and acacia classes. However, when considering the UAs, the pine and eucalyptus classes returned similar results. The acacia class consistently produced the lowest mean PA and UA in all experiments.

Table 4-4 The overall accuracy, kappa statistic, user’s accuracy (UA), and producer’s accuracy (PA) of each block for all experiments

Tile	Exp 1. Train on block 4						Exp 2. Train on block 10						Exp 3. Train on block 17						Exp 4. Train on block 4, 10 and 17													
	OA	KS	Producers Accuracy			Users Accuracy			OA	KS	Producers Accuracy			Users Accuracy			OA	KS	Producers Accuracy			Users Accuracy										
			Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine			Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine			Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine					
1	0,835	0,686		0,900	0,770		0,849	0,917	0,875	0,767		0,870	0,880		0,978	0,917	0,745	0,528		0,660	0,830		0,880	0,762	0,845	0,712		0,860	0,830		0,945	0,883
2	0,547	0,320	0,000	0,870	0,770	0,000	0,424	0,846	0,577	0,365	0,030	0,780	0,920	0,429	0,431	0,821	0,490	0,235	0,060	0,510	0,900	0,429	0,331	0,682	0,563	0,345	0,010	0,770	0,910	0,333	0,425	0,784
3	0,553	0,330	0,020	0,870	0,770	0,167	0,446	0,828	0,593	0,390	0,020	0,850	0,910	0,143	0,924	0,469	0,453	0,180	0,020	0,450	0,890	0,067	0,302	0,736	0,567	0,350	0,020	0,810	0,870	0,105	0,910	0,453
4	* 0,933	0,900	0,970	0,930	0,900	0,942	0,894	0,968	0,840	0,760	0,880	0,790	0,850	0,936	0,823	0,773	0,510	0,265	0,010	0,530	0,990	0,200	0,930	0,416	* 0,843	0,765	0,930	0,660	0,940	0,894	0,904	0,764
5	0,907	0,860	0,990	0,980	0,750	1,000	0,797	0,962	0,900	0,850	1,000	0,970	0,730	0,917	0,836	0,973	0,543	0,315	0,030	0,740	0,860	0,300	0,771	0,443	0,917	0,875	0,990	0,870	0,890	0,952	0,926	0,873
6	0,670	0,000		0,670			1,000		0,510	0,000		0,510			1,000		0,360	0,000		0,360			1,000		0,440	0,000		0,440			1,000	
7	0,820	0,730	0,740	0,960	0,760	0,902	0,722	0,894	0,793	0,690	0,820	0,900	0,660	0,739	0,796	0,868	0,517	0,275	0,100	0,590	0,860	0,370	0,922	0,411	0,887	0,830	0,890	0,900	0,870	0,840	0,978	0,853
8	0,480	0,220	0,000	1,000	0,440		0,629	0,312	0,527	0,290	0,000	0,990	0,590	0,000	0,429	0,983	0,353	0,030	0,020	0,630	0,410	0,069	0,643	0,237	0,603	0,405	0,030	0,910	0,870	0,300	0,919	0,455
9	0,870	0,805	0,980	0,820	0,810	0,899	0,820	0,890	0,910	0,865	0,970	0,810	0,950	0,933	0,964	0,848	0,457	0,185	0,000	0,490	0,880	0,000	0,817	0,374	0,883	0,825	0,970	0,800	0,880	0,915	0,879	0,854
10	0,820	0,730	0,850	0,780	0,830	0,850	0,813	0,798	* 0,907	0,860	0,970	0,820	0,930	0,858	0,911	0,959	0,480	0,220	0,100	0,410	0,930	0,556	0,745	0,410	* 0,877	0,815	0,940	0,790	0,900	0,847	0,849	0,938
11	0,945	0,894	0,940	0,950		0,990	0,979		0,960	0,920	0,980	0,940		0,952	0,979		0,260	0,103	0,060	0,460		0,462	0,807		0,915	0,837	0,980	0,850		0,933	0,977	
12	0,710	0,565	0,630	0,810	0,690	0,829	0,600	0,775	0,807	0,710	0,710	0,810	0,800	0,877	0,705	0,889	0,400	0,100	0,080	0,290	0,830	0,400	0,426	0,392	0,793	0,690	0,710	0,900	0,770	0,866	0,682	0,895
13	0,810	0,715	0,720	0,930	0,780	0,889	0,705	0,897	0,870	0,805	0,760	0,970	0,880	0,905	0,789	0,946	0,470	0,205	0,040	0,470	0,900	0,500	0,588	0,425	0,843	0,765	0,760	0,970	0,800	0,854	0,752	0,976
14	0,507	0,260	0,010	0,620	0,890	1,000	0,827	0,397	0,330	-0,005	0,030	0,620	0,340	0,042	0,756	0,231	0,647	0,470	0,950	0,050	0,940	0,922	0,500	0,503	0,843	0,765	1,000	0,610	0,920	0,893	1,000	0,724
15	0,427	0,140	0,000	0,520	0,760	0,000	0,403	0,463	0,290	-0,065	0,000	0,330	0,540	0,000	0,452	0,302	0,713	0,570	0,920	0,350	0,870	0,719	0,875	0,659	0,413	0,120	0,010	0,290	0,940	0,071	0,509	0,410
16	0,457	0,185	0,000	0,540	0,830	0,000	0,425	0,503	0,337	0,005	0,000	0,390	0,620	0,000	0,506	0,344	0,723	0,585	0,870	0,440	0,860	0,757	0,772	0,672	0,450	0,175	0,070	0,380	0,900	0,318	0,535	0,435
17	0,553	0,330	0,070	0,860	0,730	0,500	0,562	0,549	0,507	0,260	0,420	0,830	0,270	0,506	0,525	0,458	* 0,933	0,900	0,990	0,950	0,860	0,892	0,931	0,989	* 0,753	0,630	0,680	0,790	0,790	0,850	0,790	0,658
18	0,410	-0,013		0,010	0,810		0,091	0,519	0,185	0,027		0,020	0,350		0,069	0,972	0,245	-0,203	0,000	0,490	0,000	0,329	0,000	0,455	0,145		0,000	0,910		0,000	0,645	
19	0,625	0,254		0,350	0,900		0,796	0,581	0,140	-0,376		0,020	0,260		0,077	0,210	0,640	0,301		0,330	0,950		1,000	0,590	0,475	-0,010		0,020	0,930		1,000	0,490
Mean	0,678	0,469	0,461	0,756	* 0,776	0,641	0,673	* 0,712	0,624	0,427	0,506	* 0,701	0,675	0,549	0,682	* 0,704	0,523	0,277	0,283	0,484	* 0,809	0,443	* 0,714	0,512	* 0,704	0,528	0,599	0,664	* 0,878	0,665	* 0,788	0,711

Key: Low accuracies  High accuracies; Overall accuracy (OA); Kappa statistic (KS)



In Experiment 3, the PAs of acacia in Tiles 14 to 17 exceed 92% but are less than 10% in Tiles 1 to 13. This suggests that for the acacia class the model built in the winter rainfall region is not transferable to the summer rainfall region, and visa-versa. Furthermore, the OAs and KSs of the tiles containing no or little acacia are low, which coincide with the producer's and user's accuracies of acacia.

#### 4.5 Factors influencing signature extension

Figure 4-5 compares classification accuracies achieved by signature extension to 1) extension distance; 2) rainfall seasonality; and 3) temperature seasonality. It is clear that the OAs decrease as extension distance increases, with a general decline of 3%, 6% and 2% per 100 km for Experiment 1, 2 and 3 respectively. The OAs are frequently below 50% when the extension distance exceeded 500 km (Figure 4-5a, b, and c). The statistical relationship between OA and extension distance was strong ( $R^2=0.723$ ) in Experiment 2, but very weak ( $R^2=0.18$ ) in Experiment 3.

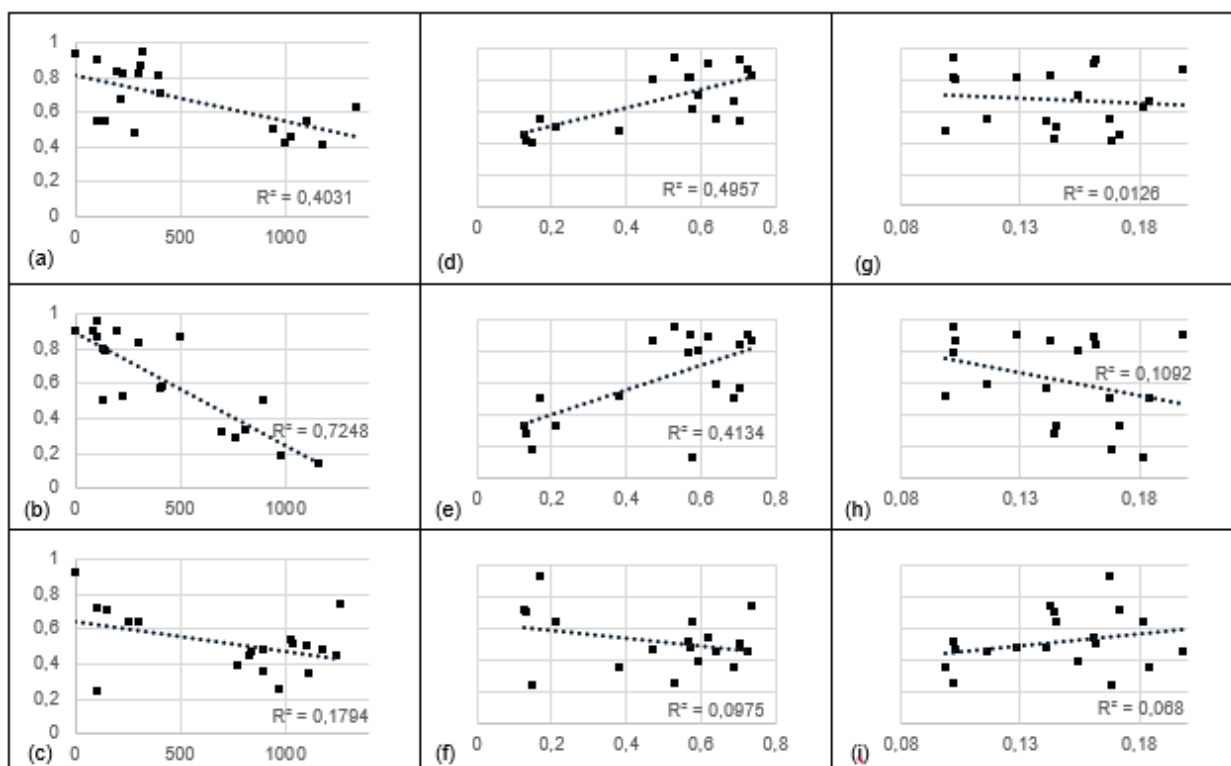


Figure 4-5 Overall accuracy vs distance from the source tile for exp 1 (Tile 4) (a), exp 2 (Tile 10) (b), and exp 3 (Tile 17) (c), overall accuracy vs rainfall seasonality index for exp 1 (d), exp 2 (e), and exp 3 (f), and overall accuracy vs temperature index for exp 1(g), exp 2 (h), exp 3 (i)

In general, higher accuracies were obtained for tiles with a similar rainfall seasonality to the source tile. In Experiment 1, the statistical relationship between OA and rainfall seasonality was stronger ( $R^2=0.5$ ) than extension distance ( $R^2=0.403$ ), but weaker ( $R^2=0.413$ ) in Experiment 2. As with

extension distance, the relationship with rainfall seasonality was weak ( $R^2=0.098$ ) in Experiment 3. The relationship between OA and temperature is weak ( $R^2<0.11$ ) in all three experiments.

In summary the OAs have the strongest relationship with rainfall seasonality in Experiment 1, whereas extension distance was the strongest driver of OAs in Experiments 2 and 3. Temperature has the weakest effect on the OA for all experiments.

When comparing the intra-class spectral variability (as quantified by the J-M distances calculated among tiles) between the training data from Tile 4, Tile 10, Tile 17 and the testing data in each tile (Table 4-5), it is clear that the intra-class spectral variability of the pine class is lower compared to the eucalyptus and acacia classes. As can be expected, the intra-class J-M scores are lowest in the source tiles (i.e. Tile 4 for Experiment 1, Tile 10 for Experiment 2, and Tile 17 for Experiment 3), highlighted in light red. However, for Experiment 3 the intra-class J-M distance for eucalyptus in the source tile (Tile 17) is high compared to acacia and pine. This suggests a high level of spectral variability within the eucalyptus class in Tile 17.

Table 4-5 Intra-class spectral variability among tiles, as quantified by Jefferies-Matusita distance

Tile	Acacia				Eucalyptus				Pine			
	Exp. 1	Exp. 2	Exp. 3	Exp.4	Exp. 1	Exp. 2	Exp. 3	Exp.4	Exp. 1	Exp. 2	Exp. 3	Exp.4
1	*	*	*	*	0,036021	0,071421	0,178974	0,150773	0,085509	0,066405	0,154094	0,071503
2	0,846965	0,659844	0,557646	0,718093	0,100123	0,079800	0,163766	0,117189	0,079160	0,138964	0,099115	0,165579
3	0,520246	0,540368	0,617000	0,461913	0,016117	0,082967	0,189245	0,197526	0,014865	0,243143	0,095059	0,153343
4	0,036810	0,145830	0,271102	0,097380	0,024328	0,048830	0,098924	0,116788	0,192562	0,280406	0,184719	0,120275
5	0,156926	0,311246	0,462086	0,237308	0,040567	0,096734	0,198991	0,184942	0,096257	0,107595	0,060767	0,105117
6	*	*	*	*	0,266487	0,157846	0,131336	0,119271	*	*	*	*
7	0,127041	0,155233	0,196951	0,111042	0,525696	0,495462	0,463688	0,491236	0,419992	0,309221	0,319825	0,286444
8	0,533910	0,454983	0,318231	0,436594	0,050214	0,025781	0,130784	0,074949	0,202394	0,059103	0,251540	0,049949
9	0,301538	0,243916	0,226401	0,193789	0,080136	0,018051	0,115917	0,046059	0,080646	0,048857	0,139161	0,099897
10	0,169355	0,016581	0,112053	0,055150	0,236011	0,166263	0,161181	0,169721	0,065686	0,080619	0,126359	0,144395
11	0,140155	0,091358	0,183300	0,064712	0,110311	0,024296	0,127919	0,030825	*	*	*	*
12	0,426103	0,282815	0,403646	0,342484	0,152781	0,059707	0,126695	0,025788	0,342422	0,127299	0,403581	0,161340
13	0,295138	0,186072	0,298535	0,245586	0,229244	0,159164	0,196297	0,198789	0,206503	0,018089	0,257450	0,120270
14	0,900773	0,826481	0,896230	0,799896	0,188707	0,182043	0,210317	0,290078	0,165443	0,145128	0,143137	0,061376
15	0,371235	0,219516	0,283162	0,241469	0,183408	0,177378	0,208672	0,290862	0,214660	0,175862	0,168029	0,046284
16	0,419815	0,261304	0,320818	0,301634	0,269122	0,174927	0,090264	0,159347	0,139972	0,100693	0,109091	0,107741
17	0,344436	0,124653	0,020820	0,187822	0,526282	0,547075	0,629712	0,706655	0,145590	0,167028	0,065255	0,122088
18	*	*	*	*	0,854256	0,718710	0,722674	0,726749	0,054784	0,232744	0,094758	0,198638
19	*	*	*	*	0,147371	0,106154	0,164236	0,123071	0,134490	0,211483	0,066986	0,158806
Mean	0,372696	0,301347	0,344532	0,299658	0,212483	0,178558	0,226821	0,222138	0,155349	0,147802	0,161113	0,127826

Key: Low J-M scores  High J-M scores, \* class not present

## 4.6 DISCUSSION

RS technologies have been effectively used for mapping commercial forest plantations (Mngadi et al. 2019). One limitation of machine learning is that sufficient in situ data are required for training and testing the models generated. In situ data collection is costly and time-consuming, especially when collecting large quantities over large areas. Ideally, a balance must be struck between training data collection efforts and classification accuracies targeted, particularly when operational solutions are sought. Very little is known about whether signature extension is a viable solution for reducing training sample collection costs when mapping forest plantation genera at a national scale. Olthof, Butson & Fraser (2005), Knorn et al. (2009), and Verhulp & Van Niekerk (2016) showed that with signature extension there is a trade-off between extension distance and

OA, while Woodcock et al. (2001), Phalke & Özdoğan (2018), and Wang, Azzari & Lobell (2019) suggested that seasonal variabilities can also negatively affect signature extension implementations. However, little is known about how these factors will impact a signature extension approach for mapping forest plantation genera at regional scales, particularly in areas where plantations are sparsely distributed and where climatic variations are dramatic.

In this study, we evaluated the extent to which signature extension can be used, along with RF machine learning, to map forest plantation genera across a large region in South Africa, using 18 features derived from a composite Sentinel-2 image as input. We found that when the model was trained with in situ data obtained from an area (Sentinel-2 tile) in the northeast (with summer rainfall) and applied to the southwestern parts (with bimodal to summer and winter rainfall) of the study area, the accuracies were generally low. In comparison, much better results were obtained when the same model was applied to areas with similar climatic profiles to where the training data were collected. This is attributed to larger inter-class spectral variability as extension distance increases and climatic conditions change, shown in Figure 4-5.

This is in agreement with Olthof, Butson & Fraser (2005), who generated spectral signatures in a source scene and applied them to other scenes. The OAs were strongly affected by extension distance. Similar conclusions were drawn by Knorn et al. (2009), who found a 1.9% average decrease in OA as extension distance increased by one Landsat scene (i.e. ~ 1% per 100 km). Whereas, we observed a mean decrease of ~ 4% in OA per 100 km increase in extension distance.

Our findings are also in agreement with Woodcock et al. (2001), who found that high accuracies can be achieved when the source and classified tiles are in the same climatic regions, but when the model is applied across different climatic regions the accuracies decrease. They concluded that the accuracies are influenced by extension distance and climatic variation. Our findings show that lower accuracies are achieved when models are trained with a data collected in areas with different rainfall seasonality to the area being mapped.

The inter-class spectral analysis showed that there was high separability in the SWIR region of the EMS when the source tiles were located in the winter and bimodal to summer rainfall regions, suggesting that moisture content may have been a discriminating factor (Manna & Raychaudhuri 2020). The SWIR wavelengths are known to be absorbed by vegetation with high moisture content (Manna & Raychaudhuri 2020). The source tiles located in the summer rainfall regions had a low separability in the SWIR region of the EMS suggesting that moisture content did not have an influence on the separability of the genera, because the annual rainfall is higher resulting in the trees having less water stress (Manna & Raychaudhuri 2020).

Our results show that pine compartments were generally classified with high accuracy, with mean PAs ranging from 70% to 88%. This is attributed to pine having a more consistent spectral profile—relative to the other two genus classes—throughout the study region (supported by the low intra-class J-M scores in each block shown in Table 4-5). Acacia had the highest intra-class J-M scores (Table 4-5), which is directly related to the low PAs of acacia, which ranged from 28% to 60%.

Furthermore, Tile 6 contained eucalyptus, Tiles 1, 2, 18, and 19 contained eucalyptus and pine, and Tile 11 contained acacia and eucalyptus. A disadvantage of machine learning is that if the model has been trained with three classes, all unknown pixels will be categorised into three classes regardless of whether all three classes are present in the area being classified. As a result, Tiles 2, 6, 18, and 19 consistently had low OAs and KSs as some classes were absent from those tiles.

Our findings show that an average OA of ~70% can be achieved using signature extension for mapping acacia, eucalyptus and pine trees if the training samples are collected from tiles that adequately represent variation in rainfall seasonality. Furthermore, our results suggest that an extension distance of 500 km or less is recommended. The source tiles from where the training samples are collected should also contain all of the target classes. Temperature seasonality did not have a significant effect on OAs and can be disregarded when mapping forest plantation genera over large and complex areas.

Our findings demonstrate the potential of signature extension for forest plantation genera mapping at a regional scale. However, more work is needed to evaluate whether multi-temporal variables will improve the differentiation between acacia, eucalyptus, and pine. It is expected that such imagery will represent at least some of the phenological differences among genera and will increase classification accuracies. Furthermore, acacia plantations are spectrally similar to pine plantations and suffered from low accuracies because they are in the minority (in the study area). More work is needed to evaluate whether sequential binary classifications (Dong et al. 2013) will improve the OAs. For instance, by initially classifying eucalyptus and pine plantations, and then further extracting acacia plantations from the pine class. In addition, more work is needed to evaluate how the reduction of the number and distribution of training samples for classes with low intra-class separability will affect overall classification accuracies. The intra-class spectral variability of the pine class was generally low and relatively high accuracies were achieved in most tiles, even when the training data was collected in one source tile (Experiments 1, 2 and 3). It would be worth investigating differential signature extension configurations, whereby extension distance is increased for some classes (e.g. pine in our case), while shorter distances are used for

classes that are more difficult to differentiate (e.g. acacia and eucalyptus). Such an approach could potentially further reduce in situ data collection efforts.

#### 4.7 CONCLUSION

The training data used in machine learning influence the OA of the classified map. Training data are traditionally collected in the image scene being mapped, which can be prohibitively expensive for operational applications (such as genera mapping) at regional scales. This study evaluated the viability of machine learning and signature extension for mapping forest plantation genera at a regional scale, using Sentinel-2 bands and derivatives (e.g. NDVI, EVI, and textural measures) as input ( $n = 18$ ). Three RF models were built using training data collected in three different Sentinel-2 tiles. A fourth model was produced by combining the training data from all three tiles. Each model was applied to 19 Sentinel-2 tiles and the resulting OAs were plotted against signature extension distance, as well as rainfall and temperature seasonality.

Our findings show that extension distance and seasonal variations negatively influence overall classification accuracy. The location and clustering of the training data are critical for the implementation of signature extension. Our results show that OAs of ~70% can be achieved for mapping plantation forest genera as long as the extension distance does not exceed 500 km and if the machine learning models are trained and applied in areas of similar rainfall seasonality. Specifically in South Africa, three source tiles, located in Mpumalanga, KwaZulu-Natal, and the Eastern Cape, are highly suitable for collecting training data. Further efforts for operational mapping of forest genera in South Africa should consider the impact of differential signature extension, whereby extension distance for some classes is more than others. Other approaches, such as multi-temporal image analysis and sequential binary classification may also improve results. Answering these questions may further reduce in situ data collection efforts and potentially increase the OAs of genera maps. Operationalising these methods may contribute to more frequent updating of forest inventories, and at a reduced cost. Up-to-date and accurate inventory information will help monitor the state and trends of forestry activities, provide valuable input to climate change modelling, support policymaking and trade decisions, and monitor the socio-economic impact of forests at regional and national scales.

## **CHAPTER 5: DISCUSSION AND CONCLUSION**

### **5.1 REVISITING THE AIMS AND OBJECTIVES**

Forest inventories are used to monitor water levels, assess climate change, monitor forest trends, evaluate the contribution of forestry to the economy, and support decisions regarding policymaking and trade. On a local scale, inventory data are used for planning, sustainable land management, assessing water use, monitoring timber harvests and rotations, and site management. General data are important for carbon stock estimations and changes, and used as an input for allometric equations for predicting biomass and carbon stock levels, and are used in streamflow reduction models. Such information is recorded by commercial forestry companies but it is not in the public domain due to competitors potentially gaining a commercial advantage. In addition, small plantation owners are unlikely to record genus information. Currently, genus information is collected in-field, which is costly and time-consuming. Therefore, this research was conducted to find an operational solution for mapping forest plantation genera to aid forest inventorying over large areas.

RS technologies have shown to be a viable alternative to collecting inventory data. Supervised classification methods are more suitable than unsupervised methods for mapping forest plantation species/genera, as unsupervised methods are unable to group pixels into the correct genus/species group. However, supervised methods require sufficient training data to produce accurate genera maps. It is unknown which sampling configuration and size will be most effective for the machine learning algorithms to map forest plantation genera. Furthermore, image classification is usually done per image scene, which is problematic when mapping over large areas, as the training data collection becomes cumbersome. It is not known whether signature extension—training a model on one image scene and applying it on other image scenes—will produce accurate genera maps.

The literature review (Objective 1) identified many different sources of imagery and machine learning algorithms used to map forest plantation genera. High-resolution imagery, such as Sentinel-2, has been used with machine learning algorithms to map land cover types that are spectrally similar. However, the literature review revealed conflicting recommendations on which training sampling configuration and size is best for mapping forest plantation genera. Therefore, in situ plantation data were acquired from commercial forest plantation companies (Objective 2) to develop experiments to investigate the factors that influence accuracies. In this study, Sentinel-2 imagery was obtained through the GEE data catalogue to create composite images spanning a year to compensate for seasonal variations so that the experiments can be directly compared to each other.

Chapter 3 (Objective 3) was experimental, as different sampling configurations and sample sizes were tested to identify an optimal sampling design for obtaining training data for machine learning algorithms to map three forest plantation genera (acacia, eucalyptus, pine). Furthermore, training data collection becomes cumbersome when mapping over large areas as image classification is usually done per image scene. A potential solution is to use signature extension. In Chapter 4, signature extension using RF was evaluated (Objective 4) for mapping genera across four provinces in South Africa. The main aim was to gain a better understanding of how extension distance and environmental factors influence the accuracies when mapping forest plantation genera. The findings of these experiments are synthesised and discussed in the next section. This is followed by a critical review of the study, as well as some recommendations for further research (Objective 5).

## 5.2 SYNTHESIS

The aim of Chapter 3 was to investigate the impact of using different sampling strategies for collecting training data on the performance of a machine learning classifier for differentiating between forest plantation genera.

An even number of 3 000 samples per genera was selected from two diverse study areas; one in KwaZulu-Natal and one in the Western Cape. An RF machine learning classifier was used with 37 features extracted from a Sentinel-2 composite image representing one year of images (31-06-2019 to 31-06-2020). Eight experiments were carried out using different sample sizes to quantify the effect of a balanced, unbalanced, and area-proportionate training data set on the accuracies of the classified maps.

The results showed that the RF model did not handle unbalanced training datasets well. The minority classes generally showed high omission and low commission errors, while the majority classes showed low omission and high commission errors. The spectral separability of the classes was also evaluated, showing that the eucalyptus class was more separable from the acacia and pine classes. As a result, the differences between the UAs of the classes were larger when the number of training samples for the eucalyptus class was limited. An area-proportionate sampling scheme did not work well in the Western Cape study area as it was dominated by the pine class (the ratio of acacia, eucalyptus, and pine was 1:1:26), resulting in a highly skewed training data set. The pine class was generally over-classified, while the eucalyptus and acacia classes were under-classified. A balanced training data set produced the most consistent map accuracies in both study areas.

Chapter 4 aimed to investigate the potential of signature extension by quantify the relationship between training sample cluster distribution and genus classification accuracy. The study area

spanned Mpumalanga, KwaZulu-Natal, the Eastern Cape and the Western Cape, representing nine Köppen-Geiger climatic zones and six rainfall seasonality regions. The study area was divided into 19 Sentinel-2 tiles to test the viability of signature extension for mapping forest plantation genera. Acacia, eucalyptus, and pine, were mapped using an RF model with signature extension. Four experiments were carried out. For the first experiment, training data from one Sentinel-2 tile (located in Mpumalanga) was used to build an RF model, which was then used to classify all 19 Sentinel-2 tiles. Similarly, in Experiment 2 and 3, samples from one Sentinel-2 tile (located in KwaZulu-Natal and the Eastern Cape respectively) were used for building two RF models, which was then applied to all tiles. In the final experiment, training samples from three tiles (one in each of the aforementioned provinces) were used for model building, which was then applied to all tiles. The experiments revealed that accuracies decreased as the distance increased from the training tile. In addition, it was found that rainfall seasonality had a strong impact on accuracies. Specifically, accuracies dropped when the models were applied to areas with different rainfall seasonality (e.g. summer vs. winter rainfall). A separability analysis showed that the pine class is very stable across the study area, while the spectral characteristics of the eucalyptus class are more variable. The acacia class showed the most spectral variability of the three classes. This suggests that the collection of pine samples can be clustered, while the eucalyptus and acacia classes require a sampling configuration that is stratified according to seasonal variations (e.g. per Köppen-Geiger climatic zone). In general, samples should be within 500 km of the areas being mapped.

The findings of this research demonstrate the importance of considering different sampling configurations and sizes to maximise the capability of machine learning algorithms to differentiate between spectrally similar classes, specifically when using Sentinel-2 data and its derivatives as input features.

### **5.3 VALUE OF RESEARCH**

This research shows that RS technologies and machine learning can contribute to updating forest inventories which will, in turn, benefit forest management, predictive modelling, forest monitoring, climate change monitoring, and decision-making regarding trade and policies. Although VHR, hyperspectral, and LiDAR data have been used in machine learning algorithms to successfully map forest plantation genera/species, this study focused on finding a regional/national solution. Freely available Sentinel-2 imagery was identified as the most viable source of data. The study area spans Mpumalanga, KwaZulu-Natal, the Eastern Cape, and the Western Cape provinces of South Africa and represents nine climatic regions and six rainfall seasonality regions.



The results of Chapter 3 showed that accuracies of up to 76.3% can be achieved for mapping forest plantation genera on a local scale using a balanced sample configuration with a sample size of  $\sim 57n$ . These findings can be used to reduce the time taken to obtain training data.

Furthermore, the results of Chapter 4 showed that the training data obtained from regions with similar climatic conditions and within 500 km of the area being mapped are most effective for mapping plantation forest genera at regional scales using the signature extension in an RF model. This finding can be used to reduce in situ data collection efforts, by clustering samples in areas with similar rainfall seasonality and within 500 km of the mapping scene. In addition, a 70% OA can be achieved for mapping acacia, eucalyptus and pine in South Africa if the training data are collected from three Sentinel 2 tiles—one in Mpumalanga, one in KwaZulu-Natal, and one in the Eastern Cape.

#### **5.4 STUDY LIMITATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH**

One limitation of this study is that the classification was conducted on delineated forest compartment boundaries. This means that the forest compartment boundaries need to be delineated prior to analysis if the methods employed in this study are to be replicated. An alternative is to make use of existing land cover maps (DEA 2019) to mask known forest plantations, followed by image segmentation to produce objects (polygons) that can be used instead of delineated forest compartments. One can also make use of multi-temporal Sentinel-2 images for delineating compartment boundaries using Canny edge detection and watershed segmentation, as suggested by Watkins & Van Niekerk (2019).

Another limitation of the study was that RF was the only classification algorithm used. Although previous research has shown that RF is suitable for plantation classifications (Adam et al. 2012; Franklin & Ahmed 2018; Immitzer, Atzberger & Koukal 2012; Lück 2018 & Nery et al. 2019), other machine learning algorithms have shown much potential to differentiate between spectrally similar classes (Nery et al. 2019; Vaglio Laurin et al. 2016). It would be worthwhile comparing the accuracies of forest plantation mapping using MLC, SVM, NN and RF.

This study showed that Sentinel-2 bands, combined with textural measures and VIs, are suitable for mapping forest plantation genera. However, Mngadi et al. (2019) showed that fusing optical imagery with active satellite imagery such as SAR can improve the differentiation between forest plantation genera/species. Future studies should investigate whether SAR imagery fused with Sentinel-2 can improve the classification of forest plantation genera using RF.

Chapter 3 showed that there is a relationship between the number of samples and the OAs achieved by machine learning algorithms. All the Sentinel-2 bands (13 bands), NDVI, EVI, and all GLCM texture measures available in GEE were used as input to the classifications, totalling 37 features. According to the Hughes phenomenon, as the number of features increases, so do the accuracies, but at some point, the accuracies start to decrease unless the number of samples is also increased (Ma et al. 2013). It would be interesting to determine which of these features are most important for differentiating between forest plantation genera. Future studies should consider applying feature selection methods to reduce the number of features. This, in theory, should further reduce the number of samples needed to differentiate between forest plantation genera.

This research aimed at classifying the three most common plantation genera in South Africa. However, compared to eucalyptus and pine, acacia compartments are scarce. This resulted in low producer and user's accuracies for acacia, which in turn decreased the overall accuracies of the classifications. An experiment (shown in Appendix B) that excluded acacia in the Eastern and Western Cape showed a significant increase in the OA from 65% to 74%. Furthermore, pine is more distinct from eucalyptus than it is from acacia, resulting in better producer's and user's accuracies when acacia was left out of the classification.

This study purposefully used a composite image spanning a year to compare the results of each experiment. However, it is known that multi-temporal variables can improve the separability between genera as phenological characteristics can be represented (Fagan et al. 2015). It is recommended to include a multi-temporal variable to initially classify the eucalyptus and pine classes (leaving out acacia due to its scarcity). It is then suggested that the pine class be further subdivided (classified) into pine and acacia classes using spectral and temporal information.

Although per-pixel classification was used in this study, as it was the most common method in literature for classifying forest plantations, some research suggests that OBIA produces better accuracies than per-pixel approaches (Immitzer, Atzberger & Koukal 2012). It is recommended that future studies compare OBIA and per-pixel approaches for classifying acacia, eucalyptus, and pine trees. However, given that object delineation and segmentation parameterisation is an ill-structured problem (Louw & Van Niekerk 2019), image segmentation will add a level of uncertainty to the mapping workflow which will have to be taken into account in OBIA experimental designs.

## 5.5 CONCLUSION

Forest inventories are important for forest management, monitoring timber harvest, predictive modelling, and assessing the contribution of forestry to the economy. Traditional methods for

forest inventory data collection are costly and time-consuming. This research set out to evaluate whether freely available Sentinel-2 imagery and its derivatives could be used as input to machine learning to differentiate between the three main forest plantation genera in South Africa. The aim was to develop an RS methodology whereby forest plantation genera can be mapped at a national scale. The aim was achieved by investigating the impact of training sample size and configuration when mapping forest plantation genera, and by investigating the viability of signature extension as a strategy for mapping forest plantation genera at a national scale.

The use of Sentinel-2 bands, NDVI, EVI, and textural measures as input to machine learning algorithms produces genera maps with ~76% OA using a balanced training sample size of 57n. An important finding of the research was that the training sample configuration and size used influences the machine learning algorithms' ability to differentiate between forest plantation genera. The most accurate results were produced using a balanced training sample configuration. Additionally, the number of samples used also influences the classification accuracies. This research showed that as the number of samples increases, so do the OAs, but when the number of samples reaches 57n, the increase in accuracies become marginal.

Signature extension can be used to map forest plantation genera at a national scale. However, the samples must be collected in areas with similar seasonal characteristics to the areas being mapped and the distance away from the training sample site should be < 500 km. To map forest plantation genera at national scale in South Africa, it is recommended that in situ data be collected from one Sentinel-2 tile in Mpumalanga, one in KwaZulu-Natal, and one on the border of the Eastern and Western Cape.

The guidelines developed in this research can contribute towards regularly mapping forest plantation genera at regional scales and with minimal costs. The South African forestry sector contributes 1% to its GDP, employing ~165 900 workers from many rural communities (Tibane & Vermeulen 2014). Since South Africa implemented the National Water Act of 1998, the licensing process for afforestation has become cumbersome. The Act states that there should only be a certain amount of forestry activity per catchment to ensure all parties have access to a sustainable amount of water (Gush et al. 2002). As a result, there has been a decrease of 80 000 ha of forestry activity. It is therefore important to make use of forestry inventories to maximise the production of forestry.

Currently, forest inventory in South Africa is collected every three years, making use of questionnaires and information gained from the licensing process (DWAF 2008). Regularly updated forest inventories will greatly improve forest management and can be used to evaluate the contribution that forestry makes to the economy and ensure sustainable practices so that future

generations can benefit from the important services and products that forests provide. Updated forest inventories will also improve plantation land management, planning, production by regularly assessing the annual growth rates, the assessment of water use, and forest management. Furthermore, updated genera data will specifically improve the accuracy of allometric equations used to estimate carbon stock and biomass levels, and the accuracy of streamflow models.

## REFERENCES

- Adam EM, Mutanga O, Rugege D & Ismail R 2012. Discriminating the papyrus vegetation (Cyperus papyrus L.) and its co-existent species using random forest and Hyperspectral data resampled to HYMAP. *International Journal of Remote Sensing* 33, 2: 552–569.
- Albaugh JM, Dye PJ & King JS 2013. Eucalyptus and water use in South Africa. *International Journal of Forestry Research* 1: 1–11.
- Alonso MC, Malpica JA & De Agirre AM 2011. Consequences of the Hughes phenomenon on some classification techniques. *American Society for Photogrammetry and Remote Sensing Annual Conference 2011*, May: 32–40.
- Anand A 2018. Unit 13 Image Classification. *Processing and Classification of Remotely Sensed Images*, 41–58.
- Baatuwue NB & Van Leeuwen IL 2011. Evaluation of three classifiers in mapping forest stand types using medium resolution imagery: a case study in the Offinso Forest District, Ghana. *African Journal of Environmental Science and Technology* 5, January: 25–36.
- Bangira T 2019. Mapping Surface Water in Complex and Heterogeneous Environments Using Remote Sensing. Doctoral dissertation. Stellenbosch: Stellenbosch University, Department of Geography and Environmental Studies.
- Baierle IC, Sellitto MA, Frozza R, Schaefer JL & Habekost AF 2019. An artificial intelligence and knowledge-based system to support the decision-making process in sales. *South African Journal of Industrial Engineering* 30, 2: 17–25.
- Basha SM, Rajput DS, Poluru RK, Bharath Bhushan S & Basha SAK 2018. Evaluating the performance of supervised classification models: Decision tree and Naïve Bayes using KNIME. *International Journal of Engineering and Technology(UAE)* 7, 4: 248–253.
- Basuki TM, van Laake PE, Skidmore AK & Hussin YA 2009. Allometric equations for estimating the above-ground biomass in tropical lowland Dipterocarp forests. *Forest Ecology and Management* 257, 8: 1684–1694.
- Belgiu M & Dragut L 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing Random forest in remote sensing* 114: 24–31.
- Bhaskaran S, Paramananda S & Ramnarayan M 2010. Per-pixel and object-oriented classification methods for mapping urban features using Ikonos satellite data. *Applied*

*Geography* 30, 4: 650–665.

- Blaschke T 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 1: 2–16.
- Botai CM, Botai JO & Adeola AM 2018. Spatial distribution of temporal precipitation contrasts in South Africa. *South African Journal of Science* 114, 7–8: 1–9.
- Brown C & Ball J 2000. *World View of Plantation Grown Wood<sup>1</sup>*. Forestry Department, Food and Agriculture Organization of the United Nations.
- Buddenbaum H, Schlerf M & Hill J 2005. Classification of coniferous tree species and age classes using hyperspectral data and geostatistical methods. *International Journal of Remote Sensing* 26, 24: 5453–5465.
- Budei BC, St-Onge B, Hopkinson C & Audet FA 2018. Identifying the genus or species of individual trees using a three-wavelength airborne lidar system. *Remote Sensing of Environment* 204, October 2017: 632–647.
- Bujang MA & Baharum N 2017. Guidelines of the minimum sample size requirements for Cohen 's Kappa. *Epidemiology Biostatistics and Public Health* 14, 2.
- Campbell J & Wynne R 2013. Introduction to Remote Sensing. *Remote Sensing* 5, 1: 337–375
- Carle J, Del Lungo A & Varmola M 2003. *The need for improved forest plantation data*. Rome: Forestry Department, Food and Agriculture Organization.
- Carrão H, Araújo A, Gonçalves P & Caetano M 2010. Multitemporal MERIS images for land-cover mapping at a national scale: A case study of Portugal. *International Journal of Remote Sensing* 31, 8: 2063–2082.
- Chen B, Li X, Xiao X, Zhao B, Dong J, Kou W, Qin Y, Yang C, Wu Z, Sun R, Lan G & Xie G 2016. Mapping tropical forests and deciduous rubber plantations in Hainan Island, China by integrating PALSAR 25-m and multi-temporal Landsat images. *International Journal of Applied Earth Observation and Geoinformation* 50: 117–130.
- Cho MA, Malahlela O & Ramoelo A 2015. Assessing the utility WorldView-2 imagery for tree species mapping in South African subtropical humid forest and the conservation implications: Dukuduku forest patch as case study. *International Journal of Applied Earth Observation and Geoinformation* 38: 349–357.
- Cohen Y & Shoshany M 2002. A national knowledge-based crop recognition in Mediterranean environment. *International Journal of Applied Earth Observation and Geoinformation* 4: 75-87.

- Colditz RR 2015. An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms. *Remote Sensing* 7, 8: 9655–9681.
- Congalton RG & Green K 2019. *Assessing the Accuracy of Remotely Sensed Data*. 3<sup>rd</sup> ed. Taylor & Francis: CRC Press.
- Cracknell AP 2018. The development of remote sensing in the last 40 years. *International Journal of Remote Sensing* 39, 23: 8387–8427.
- DAFF 2008. Report on commercial timber resources and primary round wood processing in South Africa. Department of Agriculture, Forestry and Fisheries.
- Dalponte M, Ørka HO, Gobakken T, Gianelle D & Næsset E 2013. Tree species classification in boreal forests with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* 51, 5: 2632–2645.
- Dannenberg MP, Hakkenberg CR & Song C 2016. Consistent classification of landsat time series with an improved automatic adaptive signature generalization algorithm. *Remote Sensing* 8, 8.
- De Beer H 1986. *Black wattle*. Pretoria: Department of Agriculture and Water Supply.
- DeFries RS, Hansen MC, Townshend JRG, Janetos AC & Loveland TR 2000. A new global 1-km dataset of percentage tree cover derived from remote sensing. *Global Change Biology* 6, 2: 247–254.
- DEA 2019. South African National Land-Cover 2018 Report & Accuracy Assessment .4: 1–39. Department of Environmental Affairs, Pretoria, South Africa.
- DWAF 2008. Expression of interest to revise, update and create the national forestry inventory and plantation database for KwaZulu-Natal. Department of Water Affairs and Forestry, Pretoria, South Africa.
- Devi Mahalakshmi S & Geethanjali V 2019. Plant classification using deep learning. *Journal of International Pharmaceutical Research* 46, 3: 745–749.
- Dong J, Xiao X, Chen B, Torbick N, Jin C, Zhang G & Biradar C 2013. Mapping deciduous rubber plantations through integration of PALSAR and multi-temporal Landsat imagery. *Remote Sensing of Environment* 134: 392–402.
- Dong J, Xiao X, Sheldon S, Biradar C & Xie G 2012. Mapping tropical forests and rubber plantations in complex landscapes by integrating PALSAR and MODIS imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 74: 20–33.

- Dye P & Versfeld D 2007. Managing the hydrological impacts of South African plantation forests: An overview. *Forest Ecology and Management* 251, 1–2: 121–128.
- Dzikiti BS, Jarman C, Jovanovic N, Le Maitre D, Mashimbye E, Stephenson G, Van Niekerk A & Vermuelen D 2019. The application of national scale remotely sensed evapotranspiration (ET) estimates to quantify water use and differences between plantations in commercial forestry regions of South Africa. Report to the Water Research Commission, Stellenbosch.
- ESA 2015. *ESA's Optical High-Resolution Mission for GMES Operational Services*. European Space Agency.
- Fagan ME, DeFries RS, Sesnie SE, Arroyo-Mora JP, Soto C, Singh A, Townsend PA & Chazdon RL 2015. Mapping species composition of forests and tree plantations in northeastern Costa Rica with an integration of hyperspectral and multitemporal landsat imagery. *Remote Sensing* 7, 5: 5660–5696.
- Food and Agriculture Organization 2007. *Brief on National Forest Inventory NFI*. September. Available from: <https://www.forestresearch.gov.uk/tools-and-resources/national-forest-inventory/>.
- Food and Agriculture Organization 2015. *Southern Africa's Forests and People - Investing in a Sustainable Future*. Available from: <http://www.fao.org/3/i4894e/i4894e.pdf>.
- Foody GM 2009. Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing* 30, 20: 5273–5291.
- Foody GM 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80, 1: 185–201.
- Foody GM & Mathur A 2004. Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification. *Remote Sensing of Environment* 93, 1–2: 107–117.
- Foody GM, Mathur A, Sanchez-Hernandez C & Boyd DS 2006. Training set size requirements for the classification of a specific class. *Remote Sensing of Environment* 104, 1: 1–14.
- FSA 2019. Timber plantation ownership. Forestry South Africa.
- Franco-Lopez H, Ek AR & Bauer ME 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote Sensing of Environment* 77, 3: 251–274.
- Francois A & Leckie DG 2006. The individual tree crown approach to Ikonos images. *Photogrammetric Engineering & Remote Sensing* 72, 11: 1287–1297.



- Franklin SE & Ahmed OS 2018. Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *International Journal of Remote Sensing* 39, 15–16: 5236–5245.
- Franklin SE, Ahmed OS & Williams G 2017. Northern conifer forest species classification using multispectral data acquired from an unmanned aerial vehicle. *Photogrammetry Engineering & Remote Sensing* 83, 7: 501–507.
- Fuller JA & Perrin MR 2001. Habitat assessment of small mammals in the Umvoti Vlei conservancy, KwaZulu-Natal, South Africa. *African Journal of Wildlife Research* 31, 1–2: 1–12.
- Gao H, Sabo JL, Chen X, Liu Z, Yang Z, Ren Z & Liu M 2018. Landscape heterogeneity and hydrological processes: a review of landscape-based hydrological models. *Landscape Ecology* 33, 9: 1461–1480.
- Geldenhuys CJ, Mucina L 2006. Towards a new national forest classification for South Africa. In: S.A. Ghazanfar & H.J Beentje (eds), *Taxonomy and ecology of Africa plants, their conservation and sustainable use*, pp. 111-129. Royal Botanic Gardens, Kew.
- Gislason PO, Benediktsson JA & Sveinsson JR 2006. Random forests for land cover classification. *Pattern Recognition Letters* 27, 4: 294–300.
- Gray J & Song C 2013. Consistent classification of image time series with automatic adaptive signature generalization. *Remote Sensing of Environment* 134: 333–341.
- Gush MB, Scott DF, Jewitt GPW, Schulze RE, Hallowes LA & Görgens AHM 2002. A new approach to modelling streamflow reductions resulting from commercial afforestation in south africa. *Southern African Forestry Journal* 196, 1: 27–36.
- Hagner O & Reese H 2007. A method for calibrated maximum likelihood classification of forest types. *Remote Sensing of Environment* 110, 4: 438–444.
- Halefom A, Teshome A, Sisay E & Ahmad I 2018. Dynamics of land use and land cover change using remote sensing and GIS: A case study of Debre Tabor Town, South Gondar, Ethiopia. *Journal of Geographic Information System* 10, 02: 165–174.
- Han D, Liu Q & Fan W 2018. A new image classification method using CNN transfer learning and web data augmentation. *Expert Systems with Applications* 95: 43–56.
- Hansen MC, DeFries RS, Townshend JRG, Carroll M, Dimiceli C & Sohlberg RA 2003. Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS Vegetation Continuous Fields Algorithm. *Earth Interactions* 7, 10: 1–15.

- Hansen MC, Townshend JRG, DeFries RS & Carroll M 2005. Estimation of tree cover using MODIS data at global, continental and regional/local scales. *International Journal of Remote Sensing* 26, 19: 4359–4380.
- Heinzel J & Koch B 2011. Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation* 13, 1: 152–160.
- Herod A 2016. Scale: The local and the global. May: 217–235.
- Heydari SS & Mountrakis G 2018. Effect of classifier selection, reference sample size, reference class distribution and scene heterogeneity in per-pixel classification accuracy using 26 Landsat sites. *Remote Sensing of Environment* 204, February 2017: 648–658.
- Holmgren P & Thuresson T 1998. Satellite remote sensing for forestry planning—A review. *Scandinavian Journal of Forest Research* 13, 1–4: 90–110.
- Huete A, Didan K, Miura T, Rodriguez EP, Gao X & Ferreira LG 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83: 195–213.
- Immitzer M, Atzberger C & Koukal T 2012. Tree species classification with Random forest using very high spatial resolution 8-band worldView-2 satellite data. *Remote Sensing* 4, 9: 2661–2693.
- Jackson RR & Huete AR 1991. Interpreting veget indices. *Preventive Veterinary Medicine* 11: 185–200.
- Jokar Arsanjani J, Tayyebi A & Vaz E 2016. GlobeLand30 as an alternative fine-scale global land cover map: Challenges, possibilities, and implications for developing countries. *Habitat International* 55: 25–31.
- Kaplan G & Avdan U 2017. Object-based water body extraction model using Sentinel-2 satellite imagery. *European Journal of Remote Sensing* 50, 1: 137–143.
- Ke Y, Quackenbush LJ & Im J 2010. Remote sensing of environment synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment* 114, 6: 1141–1154.
- Kelly M, Blanchard SD, Kersten E & Koy K 2011. Terrestrial remotely sensed imagery in support of public health: New avenues of research using object-based image analysis. *Remote Sensing* 3, 11: 2321–2345.
- Keuchel J, Naumann S, Heiler M & Siegmund A 2003. Automatic land cover analysis for

Tenerife by supervised classification using remotely sensed data. *Remote Sensing of Environment* 86, 4: 530–541.

Khalid S, Khalil T & Nasreen S 2014. A Survey of feature selection and feature extraction techniques in machine learning. Paper delivered at the Science Information Conference, London.

Knorn J, Rabe A, Radeloff VC, Kuemmerle T, Kozak J & Hostert P 2009. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sensing of Environment* 113, 5: 957–964.

Kraaij T, Baard JA, Arndt J, Vhengani L & van Wilgen BW 2018. An assessment of climate, weather, and fuel factors influencing a large, destructive wildfire in the Knysna region, South Africa. *Fire Ecology* 14, 2: 1–12.

Laborte AG, Maunahan AA & Hijmans RJ 2010. Spectral signature generalization and expansion can improve the accuracy of satellite image classification. *PLoS ONE* 5, 5.

Lefsky MA 2010. A global forest canopy height map from the moderate resolution imaging spectroradiometer and the geoscience laser altimeter system. *Geophysical Research Letters* 37, 15: 1–5.

Leroy M, Bicheron P, Latham J, Gregorio A Di, Witt R, Herold M, Sambale J, Achard F, Durieux L, Plummer S & Weber J 2007. GlobCover : European space agency service for global land Cover from MERIS. *IEEE International Geoscience and Remote Sensing*: 2412–2415.

Li J, Hu B & Noland TL 2013. Classification of tree species based on structural features derived from high density LiDAR data. *Agricultural and Forest Meteorology* 171–172: 104–114.

Li Q, Qiu C, Ma L, Schmitt M & Zhu XX 2020. Mapping the land cover of africa at 10 m resolution from multi-source remote sensing data with google earth engine. *Remote Sensing* 12, 4: 1–22.

Lillesand T, Kiefer R & Chipman J 2019. *Remote Sensing and Image Interpretation* 1, 7: 59-84.

Liu D & Xia F 2010. Assessing object-based classification: Advantages and limitations. *Remote Sensing Letters* 1, 4: 187–194.

Liu HQ & Huete A 1995. Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing* 33, 2: 457–465.

Loggenberg K, Strever A, Greyling B & Poona N 2018. Modelling Water Stress in a Shiraz

- Vineyard Using Hyperspectral Imaging and Machine Learning. *Remote Sensing* 10, 2: 202–216.
- Lu D & Weng Q 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing* 28, 5: 823–870.
- Lück W 2018. Generating Automated forestry geoinformation products from remotely sensed imagery. Master's thesis. Stellenbosch: Stellenbosch University, Department of Geography and Environmental Studies.
- Lukas V, Novák J, Neudert L, Svobodova I, Rodriguez-Moreno F, Edrees M & Kren J 2016. The combination of UAV survey and Landsat imagery for monitoring of crop vigor in precision agriculture. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 41, July: 953–957.
- Ma W, Gong C, Hu Y, Meng P & Xu F 2013. The Hughes phenomenon in hyperspectral classification based on the ground spectrum of grasslands in the region around Qinghai Lake. *International Symposium on Photoelectronic Detection and Imaging 2013: Imaging Spectrometer Technologies and Applications* 8910: 89101G.
- Mahdianpari M, Salehi B, Mohammadimanesh F, Homayouni S & Gill E 2019. The first wetland inventory map of Newfoundland at a spatial resolution of 10 m using Sentinel-1 and Sentinel-2 data on the Google Earth Engine cloud computing platform. *Remote Sensing* 11, 1: 43.
- Mandy L & Steve H 2015. Tree farming guidelines for private growers. *Science of The Total Environment* 2: 10–15.
- Maniatis D, Malhi Y, Saint André L, Mollicone D, Barbier N, Saatchi S, Henry M, Tellier L, Schwartzenberg M & White L 2011. Evaluating the potential of commercial forest inventory data to report on forest carbon stock and forest carbon stock changes for REDD+ under the UNFCCC. *International Journal of Forestry Research* 2011: 1–13.
- Manna S & Raychaudhuri B 2020. Mapping distribution of Sundarban mangroves using Sentinel-2 data and new spectral metric for detecting their health condition. *Geocarto International* 35, 4: 434–452.
- Marghany M & Hashim M 2010. Lineament mapping using multispectral remote sensing satellite data. *International Journal of Physical Sciences* 5, 10: 1501–1507.
- Martinuzzi S, Gould WA, Vierling LA, Hudak AT, Nelson RF & Evans JS 2013. Quantifying tropical dry forest type and succession: Substantial improvement with lidar. *Biotropica* 45,

2: 135–146.

- Mather PM 2004. *Computer Processing of Remotely-Sensed Images*. Third edit. England: John Wiley & Sons Ltd.
- Mati A & Dawaki SA 2015. Role of forest inventory in sustainable forest management : A review. *Intrnational journal of Forestry and Horticulture* 1, 2: 33–40.
- Matsushita B, Yang W, Chen J, Onda Y & Qiu G 2007. Sensitivity of the enhanced vegetation index (evi) and normalized difference vegetation index (ndvi) to topographic effects: A case study in high-density Cypress Forest. *Sensors* 7: 2636–2651.
- McRoberts RE, Cohen WB, Erik N, Stehman S V. & Tomppo EO 2010. Using remotely sensed data to construct and assess forest attribute maps and related spatial products. *Scandinavian Journal of Forest Research* 25, 4: 340–367.
- McRoberts RE, Tomppo EO, Finley AO & Heikkinen J 2007. Estimating areal means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery. *Remote Sensing of Environment* 111, 4: 466–480.
- McRoberts RE, Wendt DG, Nelson MD & Hansen MH 2002. Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sensing of Environment* 81, 1: 36–44.
- Mead DJ 2013. *Sustainable management of Pinus radiata plantations*. Rome: Food and Agriculture Organization.
- Mellor A, Boukir S, Haywood A & Jones S 2015. Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS Journal of Photogrammetry and Remote Sensing* 105: 155–168.
- Mendoza F & Lu R 2015. *Hyperspectral Imaging Technology in Food and Agriculture*: 9–56.
- Michez A, Bauwens S, Bonnet S & Lejeune P 2016. Characterization of forests with LiDAR technology. *Land Surface Remote Sensing in Agriculture and Forest*: 331–362.
- Millard K & Richardson M 2015. On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote Sensing* 7, 7: 8489–8515.
- Mngadi M, Odindi J, Peerbhay K & Mutanga O 2019. Examining the effectiveness of Sentinel-1 and 2 imagery for commercial forest species mapping. *Geocarto International* 0, 0: 1–12.

- Mujabar PS & Dajkumar S 2019. Mapping of bauxite mineral deposits in the northern region of Saudi Arabia by using Advanced Spaceborne Thermal Emission and Reflection Radiometer satellite data. *Geo-Spatial Information Science* 22, 1: 35–44.
- Muller SJ, Sithole P, Singels A & Van Niekerk A 2020. Assessing the fidelity of Landsat-based fAPAR models in two diverse sugarcane growing regions. *Computers and Electronics in Agriculture* 170, March.
- Murthy VRK 2004. Satellite remote sensing and GIS Applications in agricultural meteorology. *Satellite remote sensing and GIS applications in agricultural meteorology*: 235.
- Myburgh G & Van Niekerk A 2014. Impact of training set size on object-based land cover classification: A comparison of three classifiers. *International Journal of Applied Geospatial Research* 5, 3: 49–67.
- Naghdy GA, Todd C, Olaode A & Naghdy G 2014. Unsupervised Classification of Images: A Review. *International Journal of Image Processing (IJIP)* 8, 5: 325.
- Nangendo G, Skidmore AK & van Oosten H 2007. Mapping East African tropical forests and woodlands - A comparison of classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing* 61, 6: 393–404.
- Neale CMU, Geli H, Taghvaeian S, Masih A, Pack RT, Simms RD, Baker M, Milliken JA, O'Meara S & Witherall AJ 2011. Estimating evapotranspiration of riparian vegetation using high resolution multispectral, thermal infrared and lidar data. *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIII* 8174: 81740P.
- Nery T, Sadler R, Solis Aulestia M, White B & Polyakov M 2019. Discriminating native and plantation forests in a Landsat time-series for land use policy design. *International Journal of Remote Sensing* 40, 11: 4059–4082.
- Nomura K & Mitchard ETA 2018. More than meets the eye: Using Sentinel-2 to map small plantations in complex forest landscapes. *Remote Sensing* 10, 1693.
- Olthof I, Butson C & Fraser R 2005. Signature extension through space for northern landcover classification: A comparison of radiometric correction methods. *Remote Sensing of Environment* 95, 3: 290–302.
- Onjira P 2014. Application of remote sensing and rainfall-run-off inundation modeling to near-real time flood monitoring in Kenya. Master's thesis. Tokyo: National Graduate Institute for Policy Studies.
- Pal M 2005. Random forest classifier for remote sensing classification. *International Journal of*

*Remote Sensing* 26, 1: 217–222.

- Park B, Lu R with Mendoza F 2015. *Hyperspectral Imaging Technology in Food and Agriculture*. New York: Springer Science+Business Media LLC New York
- Pax-Lenney M, Woodcock CE, Macomber SA, Gopal S & Song C 2001. Forest mapping with a generalized classifier and Landsat TM data. *Remote Sensing of Environment* 77, 3: 241–250.
- Peerbhay KY, Mutanga O & Ismail R 2013. Commercial tree species discrimination using airborne AISA Eagle hyperspectral imagery and partial least squares discriminant analysis (PLS-DA) in KwaZulu-Natal, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing* 79: 19–28.
- Peerbhay KY, Mutanga O & Ismail R 2014. Investigating the capability of few strategically placed worldview-2 multispectral bands to discriminate forest species in KwaZulu-Natal, South Africa. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7, 1: 307–316.
- Peled A & Gilichinsky M 2010. knowledge-based classification of land cover for the quality assesment of GIS database. Paper delivered at ISPRS joint workshop on: Core spatial databases-updating, maintenance and services – from theory to practice conference, Haifa.
- Peng W, Wheeler DB, Bell JC & Krusemark MG 2003. Delineating patterns of soil drainage class on bare soils using remote sensing analyses. *Geoderma* 115, 3–4: 261–279.
- Phalke AR & Özdoğan M 2018. Large area cropland extent mapping with Landsat data and a generalized classifier. *Remote Sensing of Environment* 219, October 2017: 180–195.
- Pierce KB 2015. Accuracy optimization for high resolution object-based change detection: An example mapping regional urbanization with 1 m aerial imagery. *Remote Sensing* 7, 10: 12654–12679.
- Pillay M 2012. Classical genetics and traditional breeding in Musa. *Genetics, Genomics, and Breeding of Eucalypts*: 34–55.
- Pontius RG & Millones M 2011. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* 32, 15: 4407–4429.
- Prasad NR, Garg V & Thakur PK 2018. Role of sar data in water body mapping and reservoir sedimentation assesment. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4, 5: 151–158.

- Pu R & Landry S 2012. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment* 124: 516–533. [online]. Available from: <http://dx.doi.org/10.1016/j.rse.2012.06.011>
- Qi W, Lee S, Hancock S, Luthcke S, Tang H, Armston J & Dubayah R 2019. Improved forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data. *Remote Sensing of Environment* 221: 621–634.
- Rodriguez-Galiano VF, Chica-Olmo M, Abarca-Hernandez F, Atkinson PM & Jeganathan C 2012. Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment* 121: 93–107.
- Roughgarden J, Running S. & Matson PA 2010. What does remote sensing do for ecology? *Ecological Society of America* 72, 6: 1917–1922.
- Rwanga SS & Ndambuki JM 2017. accuracy assessment of land use/land cover classification using remote sensing and GIS. *International Journal of Geosciences* 08, 04: 611–622.
- Schulz J, Albert P, Behr HD, Caprion D, Deneke H, Dewitte S, Dürr B, Fuchs P, Gratzki A, Hechler P, Hollmann R, Johnston S, Karlsson KG, Manninen T, Müller R, Reuter M, Riihelä A, Roebeling R, Selbach N, Tetzlaff A, Thomas W, Werscheck M, Wolters E & Zelenka A 2009. Operational climate monitoring from space: The EUMETSAT satellite application facility on climate monitoring (CM-SAF). *Atmospheric Chemistry and Physics* 9, 5: 1687–1709.
- Schulze BR 1947. The climates of South Africa according to the classifications of Köppen and Thornthwaite. *South African Geographical Journal* 29, 1: 32–42.
- Shaw G & Burke H 2003. Spectral imaging for remote sensing. *Lincoln Laboratory* 14, 1: 121–126.
- Shetty S 2019. Analysis of machine learning classifiers for LULC classification on Google Earth Engine. Master's Thesis. Twente: University of Twente, Faculty of Geo-Information Science and Earth Observation.
- Silva JR, Rodrigues WP, Ruas KF, Paixão JS, de Lima RSN, Filho JAM, Garcia JAC, Schaffer B, Gonzalez JC & Campostrini E 2019. Light, photosynthetic capacity and growth of papaya (*Carica papaya* L.): A short review. *Australian Journal of Crop Science* 13, 3: 480–485.
- Simard M, Pinto N, Fisher JB & Baccini A 2011. Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences* 116, 4: 1–12.



- South Africa 1998. National forests act, Act 30 of 1998. *Government Gazette of South Africa* 400, 30.10.1998.
- Stabach JA, Dabek L, Jensen R & Wang YQ 2009. Discrimination of dominant forest types for Matschie's tree kangaroo conservation in Papua New Guinea using high-resolution remote sensing data. *International Journal of Remote Sensing* 30, 2: 405–422.
- Stephenson G 2010. A Comparison of Supervised and Rule-Based Object-Oriented Classification for Forest Mapping. Master's thesis. Stellenbosch: Stellenbosch University, Department of Geography and Environmental Studies.
- Tang L & Shao G 2015. Drone remote sensing for forestry research and practices. *Journal of Forestry Research* 26, 4: 791–797.
- Thanh Noi P & Kappas M 2017. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors* 18, 18.
- Thompson M 2019. *South African National Land-Cover 2018 Report & Accuracy Assessment*. Pretoria: Department of Environmental Affairs and Tourism.
- Tibane E & Vermeulen A 2014. South Africa yearbook 2013/2014. *Science And Technology* 21, December: 33–59.
- Tomppo E, Olsson H, Ståhl G, Nilsson M, Hagner O & Katila M 2008. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment* 112, 5: 1982–1999.
- Vaglio Laurin G, Puletti N, Hawthorne W, Liesenberg V, Corona P, Papale D, Chen Q & Valentini R 2016. Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data. *Remote Sensing of Environment* 176: 163–176.
- Van Aardt J & Norris-Rogers M 2008. Spectral-age interactions in managed, even-aged Eucalyptus plantations: Application of discriminant analysis and classification and regression trees approaches to hyperspectral data. *International Journal of Remote Sensing* 29, 6: 1841–1845.
- Van Der Zel DW 1995. Accomplishments and Dynamics of the South African Afforestation Permit System. Pretoria: Department of Water Affairs and Forestry.
- Verhulp J & Van Niekerk A 2016. Effect of inter-image spectral variation on land cover separability in heterogeneous areas. *International Journal of Remote Sensing* 37, 7: 1639–

1657.

- Verhulp J & Van Niekerk A 2017. Transferability of decision trees for land cover classification in a heterogeneous area. *South African Journal of Geomatics* 6, 1: 30–46.
- Viera AJ & Garrett JM 2005. Understanding interobserver agreement: the kappa statistic. *Family Medicine* 37, 5: 360–363.
- Voss M & Sugumaran R 2008. Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach. *Sensors* 8, 5: 3020–3036.
- Wagner FH, Sanchez A, Tarabalka Y, Lotte RG, Ferreira MP, Aidar MPM, Gloor E, Phillips OL & Aragão LEOC 2019. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *Remote Sensing in Ecology and Conservation* 5, 4: 360–375.
- Walsh RPD & Lawler DM 1981. Rainfall Seasonality: Description, Spatial Patterns and Change Through Time. *Weather* 36, 7: 201–208.
- Wang S, Azzari G & Lobell DB 2019. Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. *Remote Sensing of Environment* 222, November 2018: 303–317.
- Watkins B 2019. Agricultural field boundary delineation using earth observation methods and multi-temporal sentinel-2 imagery. Master's thesis. Stellenbosch: Stellenbosch University: Department of Geography and Environmental Studies.
- Wilson JRU, Gairifo C, Gibson MR, Arianoutsou M, Bakar BB, Baret S, Celesti-Grapow L, Ditomaso JM, Dufour-Dror JM, Kueffer C, Kull CA, Hoffmann JH, Impson FAC, Loope LL, Marchante E, Marchante H, Moore JL, Murphy DJ, Tassin J, Witt A, Zenni RD & Richardson DM 2011. Risk assessment, eradication, and biological control: Global efforts to limit Australian acacia invasions. *Diversity and Distributions* 17, 5: 1030–1046.
- Woodcock CE, Macomber SA, Pax-Lenney M & Cohen WB 2001. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment* 78, 1–2: 194–203.
- Xie Z, Chen Y, Lu D, Li G & Chen E 2019. Classification of land cover, forest, and tree species classes with Ziyuan-3 multispectral and stereo data. *Remote Sensing* 11, 2: 1–27.
- Xue J & Su B 2017. Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors* 2017: 1–17.

- Xulu S, Peerbhay KY, Forests S & Gebreslasie M 2018. Remote sensing of forest health and vitality: A South African perspective. *Southern Forests: a Journal of Forest Science* 2018: 1–16
- Yao W, Krzystek P & Heurich M 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment* 123: 368–380.
- Zebari RR, Abdulazeez AM, Zeebaree DQ, Zebari DA & Saeed JN 2020. A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *Journal of Applied Science and Technology Trends* 01, 02: 56–70.
- Zhang X, Liu L, Chen X, Xie S & Gao Y 2019. Fine land-cover mapping in China using Landsat datacube and an operational SPECLib-based approach. *Remote Sensing* 11, 9.
- Zhu L, Suomalainen J, Liu J, Hyypä J, Kaartinen H & Haggren H 2017. A Review: Remote Sensing Sensors. In *Multi -purposeful Applicatioin of Geospatial Data*, 20–39.

## APPENDICES

Appendix A	Indices of disagreement for Chapter 3 experiments at the initial iteration and then at every tenth iteration sample increase	94
Appendix B	Binary classification (eucalyptus and pine) accuracies using signature extension with training data obtained from Tile 17	95

## APPENDIX A

Indices of disagreement for Chapter 3 experiments at the initial iteration and then at every tenth iteration sample increase

		Study Area 1 (WC)										Study Area 2 (KZN)									
		Figure of Merit			Omission			Commission			OA	Figure of Merit			Omission			Commission			OA
		Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine		Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine	
Exp. A (Even samples)	1	49	68	56	15	4	8	9	8	10	73	33	38	34	14	19	15	18	14	17	52
	10	66	77	70	9	5	3	5	4	8	83	47	53	48	11	12	11	12	9	12	66
	20	60	75	68	10	6	4	8	4	8	80	45	55	49	17	10	8	11	10	13	66
	30	64	76	72	8	5	5	7	4	6	83	52	58	51	11	10	9	10	8	12	70
	40	70	90	74	7	2	4	5	2	6	87	51	59	58	11	9	8	11	8	9	72
	50	72	85	75	7	2	3	4	3	6	87	56	61	50	11	10	7	8	7	13	72
60	69	86	74	7	3	3	5	2	6	86	55	63	52	12	8	8	8	8	12	72	
Exp. B (50 acacia)	10	6	65	58	0	16	19	31	1	3	65	6	50	37	1	20	28	31	7	11	51
	20	3	67	56	0	15	22	32	1	3	64	8	51	40	1	27	18	31	3	13	54
	30	2	71	53	0	11	25	33	2	2	63	1	51	42	0	24	22	33	4	10	53
	40	2	71	56	0	12	23	33	1	2	64	0	50	45	0	25	21	33	4	9	53
	50	1	74	53	0	11	25	33	1	3	64	0	55	46	0	23	20	33	2	9	56
	60	0	73	53	0	12	24	33	0	3	64	0	50	42	0	27	20	33	3	11	53
Exp. C (50 eucalyptus)	10	51	47	59	21	1	10	6	17	8	69	42	8	45	26	0	21	8	31	9	52
	20	53	38	65	24	0	7	3	21	7	69	53	38	65	24	0	7	3	21	7	69
	30	53	30	68	25	0	7	2	23	6	68	53	30	68	25	0	7	2	23	6	68
	40	52	25	61	25	0	10	3	25	7	65	52	25	61	25	0	10	3	25	7	65
	50	56	25	64	23	0	10	2	25	6	67	56	25	64	23	0	10	2	25	6	67
	60	56	25	62	22	0	11	2	25	6	67	56	25	62	22	0	11	2	25	6	67
Exp. D (50 pine)	10	55	73	34	23	6	0	2	5	22	71	47	47	4	24	22	1	7	7	32	54
	20	50	75	29	25	7	0	4	3	24	69	47	45	4	27	19	0	5	10	32	54
	30	55	79	23	24	6	0	2	3	26	70	48	47	0	25	21	0	5	8	33	54
	40	52	79	19	26	6	0	2	2	27	68	51	52	2	22	21	0	5	5	33	57
	50	53	84	19	26	5	0	2	1	27	70	48	50	0	25	20	0	5	6	33	55
	60	55	82	26	25	4	0	1	3	25	71	49	48	1	22	24	0	6	6	33	55
Exp. E (50 acacia & 50 eucalyptus)	10	12	63	48	4	4	33	29	10	1	60	17	22	36	1	2	52	27	26	3	44
	20	5	57	43	1	1	44	32	14	0	54	8	19	36	1	3	55	31	27	1	41
	30	2	54	42	0	1	46	33	15	0	52	5	11	35	0	0	61	32	30	0	38
	40	2	44	40	0	1	50	33	18	0	49	1	6	34	0	0	64	33	31	0	35
	50	0	36	38	0	0	54	33	21	0	45	8	4	34	0	0	62	31	32	0	37
	60	0	20	36	0	0	60	33	27	0	40	5	6	34	0	0	63	32	31	0	37
Exp. F (50 acacia & 50 pine)	10	25	52	52	4	30	4	24	1	14	62	9	38	13	3	54	2	30	0	29	41
	20	24	48	44	3	36	3	25	0	17	58	6	35	4	1	61	1	31	0	32	37
	30	16	45	46	2	40	2	28	0	17	55	7	35	8	0	61	0	31	0	31	38
	40	11	43	42	2	44	2	29	0	19	52	5	34	1	1	64	0	32	0	33	35
	50	8	41	41	2	47	1	31	0	19	50	0	34	3	0	65	0	33	0	32	34
	60	5	42	44	2	47	1	32	0	18	50	5	34	1	0	64	0	32	0	33	35
Exp. G (50 eucalyptus & 50 pine)	10	49	54	40	34	1	1	1	15	20	65	39	18	18	49	3	3	1	27	27	45
	20	45	42	36	41	0	0	0	19	21	59	37	10	22	53	1	3	2	30	25	43
	30	43	37	30	44	0	0	0	21	23	56	37	13	13	56	0	2	0	29	29	42
	40	43	38	27	45	0	0	0	21	24	55	37	11	16	56	1	1	0	30	28	42
	50	41	32	24	48	0	0	0	23	25	52	37	9	13	57	1	2	0	30	29	41
	60	41	37	22	47	0	0	0	21	26	53	36	4	15	59	1	1	0	32	28	40

Figure A.1 A summary table showing the overall accuracy, the figure of merit, omission errors and commission errors for Experiments A to G conducted Study Area 1 (WC) and Study Area 2 (KZN).

## APPENDIX B

Binary classification (eucalyptus and pine) accuracies using signature extension with training data obtained from Tile 17

Table B.1 The overall accuracy, kappa statistic, user's accuracy, and producer's accuracy of Tile 14 to 19 for Experiment 5, which used tile 17 as the source tile to train the RF model to classify eucalyptus and pine

Tile	Exp 5. Train on block 17 (Eucalyptus & Pine)					
	OA	KS	Producers Accuracy		Users Accuracy	
			Eucalyptus	Pine	Eucalyptus	Pine
14	0,50	0,00	0,00	1,00	0,00	1,00
15	0,72	0,43	0,45	0,98	0,96	0,64
16	0,75	0,49	0,51	0,98	0,96	0,67
17	0,98	0,95	1,00	0,95	0,95	1,00
18	0,50	0,00	1,00	0,00	0,00	0,50
19	0,94	0,87	0,92	0,95	0,95	0,92
Mean	0,73	0,46	0,65	0,81	0,64	0,79