

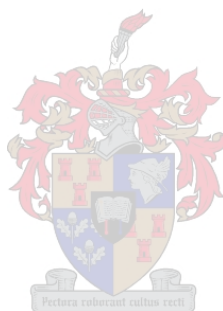
---

REGIONAL MAPPING OF SPEKBOOM CANOPY COVER USING VERY  
HIGH RESOLUTION AERIAL IMAGERY

By DUGAL HARRIS

---

Dissertation presented for the degree of Doctor of Philosophy in the Faculty of  
Science at Stellenbosch University



Promotor: Prof A van Niekerk

December 2019

## DECLARATION

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

With regard to Chapters 3, 4 and 5 the nature and scope of my contribution were as follows:

Chapter	Nature of contribution	Extent of contribution (%)
Chapter 3	This chapter was published as a journal article (Harris & Van Niekerk 2019) and was co-authored by my supervisor who helped in the conceptualisation and writing of the manuscript.	D Harris 85% A van Niekerk 15%
Chapter 4	This chapter was published as a journal article (Harris & Van Niekerk 2018) and was co-authored by my supervisor who helped in the conceptualisation and writing of the manuscript.	D Harris 85% A van Niekerk 15%
Chapter 5	This chapter was published as a journal article (Harris, Vlok & Van Niekerk 2018) and was co-authored by Jan Vlok, who helped in the problem formulation and data collection, and my supervisor who helped in the conceptualisation and writing of the manuscript.	D Harris 70% J Vlok 15% A van Niekerk 15%

Signature of candidate:      Declaration with signature in possession of candidate and

Supervisor

Signature of supervisor:      Declaration with signature in possession of candidate and

Supervisor

Date:

Copyright © 2019 Stellenbosch University

All rights reserved

## SUMMARY

Widespread degradation of subtropical thicket (South Africa) by poorly managed pastoralism has led to substantial decreases in ecological functioning and biodiversity. Once degraded, thicket does not recover after the removal of livestock, but requires active intervention for restoration. It has been shown that planting spekboom (*Portulacaria afra*), a dominant and keystone thicket species, increases biomass, improves soil health and creates conditions that support the natural regeneration of biodiversity. Spatial data, especially spekboom canopy cover maps, are required to inform and support large-scale restoration. This research aimed to develop and demonstrate a semi-automated spekboom canopy cover mapping method. A large study area in the Little Karoo in South Africa was selected to encompass the ecological heterogeneity of the wider region. Following a literature review, very high resolution (VHR) multi-spectral aerial imagery was identified as a viable data source for the fine-scale discrimination of spekboom. A set of 2228 aerial images covering the study area was subsequently acquired from Chief Directorate: National Geo-spatial Information (NGI). Techniques for (1) radiometric correction and (2) feature selection were devised to address specific challenges of regional canopy cover mapping. These techniques then formed components of the spekboom canopy cover mapping method.

The need for the first technique, called radiometric homogenisation, arose from the presence of problematic radiometric variation in the aerial imagery. Radiometric homogenisation corrects for varying atmospheric and bidirectional reflectance distribution function (BRDF) effects by calibration with concurrent and collocated satellite surface reflectance data. In contrast to other radiometric correction methods, manual placement or acquisition of reflectance targets is not required. Moreover, it is not necessary to have detailed knowledge of atmospheric conditions at the time of capture. An experiment was conducted to establish the efficacy and accuracy of the technique. Homogenised images of the study area were validated by visual inspection and statistical comparison to surface reflectance reference data. Recognisable anomalies such as hot spots and seam lines were removed, and statistical results compared well to competing methods. While the technique was developed in the context of the spekboom canopy cover mapping problem, it could also be applied to general radiometric correction of VHR imagery. Radiometric homogenisation is especially applicable to large study areas where radiometric uncertainty can prevent accurate classification.

The second technique, called feature clustering and ranking (FCR), was devised to address problems of sub-optimality and instability that often arise when applying feature selection to redundant data. Unlike other feature selection approaches, FCR allows for the optional inclusion

of factors other than relevance (such as computation and measurement cost) in the selection criteria. An experiment was conducted to compare the effects of redundancy on popular feature selection approaches and FCR. Results confirmed that redundancy has a negative impact on commonly used ranking and greedy search (stepwise) feature selection methods. FCR provided the best accuracy and stability performance, confirming its value for selecting stable, informative features from high dimensional data containing redundancy.

Finally, the radiometric homogenisation and FCR techniques were incorporated into a method for VHR spekboom canopy cover mapping. Per-pixel spectral, textural and vegetation index features were generated from imagery that had been processed with the radiometric homogenisation technique. FCR was subsequently used to select a reduced set of informative and computationally efficient features. The core of the spekboom mapping method consisted of supervised classification of selected features, followed by morphological post-processing of classifier output maps to remove noise and smooth boundaries. An experiment was carried out to test the accuracy of popular classifiers by comparing canopy cover estimates to ground truth data. A decision tree provided the best performance of the tested classifiers. Canopy cover maps exhibited some variation between different habitats, but provided good accuracy overall, with a mean absolute (canopy cover) error (MAE) of 5.85%.

Regional vegetation maps are urgently required to inform responses to global issues, such as climate change. While there is a known operational need for large-area VHR vegetation maps, there are surprisingly few studies that address the cost, computation and classifier transferability challenges associated with large spatial extents. This research contributes to the important field of regional vegetation mapping through the development of the radiometric homogenisation and FCR techniques. In the context of thicket restoration, a viable method for regional mapping of spekboom canopy cover was demonstrated, providing a valuable foundation for future expansion of maps to the rest of the thicket biome. The techniques developed in this study will be useful for the mapping of other thicket vegetation traits, such as biomass.

## **KEYWORDS**

Spekboom, *Portulacaria afra*, subtropical thicket, canopy cover, aerial image, very high resolution, radiometric correction, atmospheric effects, BRDF, surface reflectance, feature selection, clustering, supervised classification, machine learning, vegetation mapping

## OPSOMMING

Wydverspreide agteruitgang van subtropiese struikgewas (Suid-Afrika) as gevolg van swak bestuurde weiding het gelei tot aansienlike afname in ekologiese funksionering en biodiversiteit. Wanneer dit eers gedegradeer het, herstel struikgewasse nie sonder aktiewe ingryping ná die verwydering van lewende hawe nie. Daar is getoon dat die plant van spekboom – 'n dominante en sleutelstruikgewasspesie – biomassa verhoog, die grondgesondheid verbeter en toestande skep wat die natuurlike regenerasie van biodiversiteit ondersteun. Ruimtelike data, veral spekboomblaardakdekkingskaarte, word vir inligting oor en ondersteuning van grootskaalse restourasie benodig. Hierdie navorsing se doel was om 'n semi-geoutomatiseerde karteringsmetode vir spekboomblaardakdekking te ontwikkel. 'n Groot studiegebied in die Klein Karoo in Suid-Afrika is gekies om die eienskappe van die ekologiese heterogeniteit van die breër streek in te sluit. Na 'n literatuuroorsig is baie hoë resolusie (BHR) multispektrale lugfoto's as 'n lewensvatbare databron vir die fynskaalse diskriminasie van spekboom geïdentifiseer. 'n Stel van 2228 lugfoto's wat die studiegebied dek, is daarna van die Hoofdirektoraat: Nasionale Georuimtelike Inligting (NGI) verkry. Tegnieke vir (1) radiometriese korreksie en (2) eienskapseleksie is bedink om spesifieke uitdagings van die kartering van streeksblaardakdekking aan te spreek. Hierdie tegnieke het dan komponente van die karteringsmetode van spekboomblaardakdekking gevorm.

Die behoefte aan die eerste tegniek, genaamd radiometriese homogenisering, het as gevolg van die teenwoordigheid van problematiese radiometriese variasie in die lugfoto's ontstaan. Radiometriese homogenisering korrigeer vir verskillende atmosferiese en effekte van tweekrigtingweerkaatsingverspreidingsfunksie (TWVF) deur kalibrasie met samelopende en gegroepeerde data oor satelliet-oppervlakweerkaatsing. In teenstelling met ander radiometriese regstellingsmetodes, is dit nie nodig om teikens vir weerkaatsing per hand te plaas of te verkry nie. Daarbenewens word gedetailleerde kennis oor atmosferiese toestande ten tye van die vaslegging nie benodig nie. 'n Eksperiment is uitgevoer om die doeltreffendheid en akkuraatheid van die tegniek te bepaal. Gehomogeneerde beelde van die studiegebied is deur visuele inspeksie en statistiese vergelyking met verwysingsdata van oppervlakweerkaatsing gevalideer. Herkenbare onreëlmatighede soos verligte punte en naatlyne is verwyder, en statistiese resultate het goed met mededingende metodes vergelyk. Al is die tegniek in die konteks van die probleem van die kartering van spekboomblaardakdekking ontwikkel, kan dit ook op algemene radiometriese regstelling van BHR-beelde toegepas word. Radiometriese homogenisasie is veral van toepassing op groot studiegebiede waar radiometriese onsekerheid akkurate klassifikasie kan voorkom.

Die tweede tegniek, genoem kenmerkzaamgroepering en -rangordening (KSR), is ontwerp om probleme van suboptimaliteit en onstabiliteit aan te spreek wat dikwels voorkom wanneer kenmerkseleksie op oortollige data toegepas word. In teenstelling met ander eienskapseleksiebenaderings, maak KSR voorsiening vir die opsionele insluiting van ander faktore buiten relevansie (soos berekenings- en metingskoste) in die siftingskriteria. 'n Eksperiment is uitgevoer om die effekte van oortolligheid op gewilde eienskapseleksiebenaderings en KSR te vergelyk. Resultate het bevestig dat oortolligheid 'n negatiewe impak op die algemeen gebruikte rangskikking en gulsige soektog (stapsgewys) eienskapseleksiemetodes het. KSR het die beste akkuraatheids- en stabiliteitsprestasie gelewer, wat sy waarde vir die kies van stabiele, insiggewende eienskappe van hoë-dimensionele data met oortolligheid bevestig.

Ten slotte is die radiometriese homogenisasie- en KSR-tegnieke in 'n metode vir die kartering van BHR-spekboomblaardekking opgeneem. Per-piksel spektrale, tekstuur en plantegroei-indekskenmerke is van beelde wat met die radiometriese homogenisasie tegniek verwerk is, geskep. KSR is gevolglik gebruik om 'n verkorte stel informatiewe en berekeningsdoeltreffende kenmerke te kies. Die kern van die spekboomkarteringsmetode het uit gekontroleerde-klassifikasie van geselekteerde kenmerke bestaan, gevolg deur morfologiese naverwerking van klassifiseerderuitsetkaarte om geraas en gladde grense te verwyder. 'n Eksperiment is uitgevoer om die akkuraatheid van gewilde klassifiseerders te toets deur blaardakdekkingskattings met veldwaarnemings te vergelyk. 'n Beslissingsboom het die beste prestasie van die getoetsde klassifiseerders gelewer. Blaardakdekkingskaarte het 'n mate van variasie tussen verskillende habitate getoon, maar het goeie algehele akkuraatheid behaal met 'n gemiddelde absolute (blaardakdekking) fout (GAE) van 5,85%.

Streksplantegroei-kaarte word dringend nodig om reaksies op globale kwessies soos klimaatsverandering in te lig. Alhoewel die operasionele behoefte aan grootskaalse BHR-plantegroei-kaarte bekend is, is daar verrassend min studies wat die uitdagings met betrekking tot koste, berekeningstyd en klassifiseerderoordraagbaarheid wat met groot ruimtelike areas verband hou, aanspreek. Deur die ontwikkeling van die radiometriese homogenisasie- en KSR-tegnieke dra hierdie navorsing by tot die belangrike veld van die kartering van streksplantegroei by. In die konteks van die restourasie van struikgewasse is 'n lewensvatbare metode vir spekboomblaardekking op streeksvlak gedemonstreer. Dit bied 'n waardevolle grondslag vir toekomstige uitbreiding van kaarte na die res van die struikgewasbioom. Die tegnieke wat in hierdie studie ontwikkel is, sal ook vir die kartering van ander struikgewasplantegroei-eienskappe, soos biomassa, van nut wees.

## **SLEUTELWOORDE**

Spekboom, *Portulacaria afra*, subtropiese struikgewas, blaardakdekking, lugfoto, baie hoë resoluksie, radiometriese regstelling, atmosferiese effekte, TWVF, oppervlakweerkaatsing, eienskapseleksie, saamgroepering, gekontroleerde klassifikasie, masjienleer, plantegroekartering

## ACKNOWLEDGEMENTS

My sincere thanks to my supervisor, Adriaan van Niekerk, for his valuable technical input and guidance and for all the patient encouragement. Jan Vlok was instrumental in this research, and he is gratefully acknowledged for his vision, support and remarkable field expertise. Special thanks to Julie Verhulp, Raoul Duesimi and Aslam Parker of Chief Directorate: National Geospatial Information, for their generous provision of aerial imagery and related support. The Centre for Geographical Analysis, especially Garth Stephenson, Theo Pauw and Jascha Muller, graciously assisted with data, computing and software resources. Ilze Boonzaaier beautified some of the maps and Helene van Niekerk assisted with editing. Camera calibration software was kindly provided by Adrian Roos of Intergraph SA. Bernhard Jacobs of GeoSpace International assisted with valuable calibration discussions and data. Andrew Knipe, of the Gamtoos Irrigation Board, helped with funding administration. Finally, to Mireille van Reenen, for your moral support and great company – thank you.

This research was partially funded by the Department of Environmental Affairs via the Working for Natural Resources Programme.

In addition, the financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the NRF.



## CONTENTS

<b>DECLARATION .....</b>	<b>i</b>
<b>SUMMARY .....</b>	<b>ii</b>
<b>OPSOMMING .....</b>	<b>iv</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>ACRONYMS AND ABBREVIATIONS.....</b>	<b>xiv</b>
<b>CHAPTER 1: INTRODUCTION .....</b>	<b>1</b>
<b>1.1 BACKGROUND.....</b>	<b>2</b>
<b>1.1.1 Degradation of subtropical thicket .....</b>	<b>2</b>
<b>1.1.2 Spekboom and its role in subtropical thicket .....</b>	<b>4</b>
<b>1.1.3 Spekboom restoration .....</b>	<b>5</b>
<b>1.1.4 Monitoring of spekboom canopy cover .....</b>	<b>7</b>
<b>1.1.5 Remote sensing of canopy cover .....</b>	<b>8</b>
<b>1.2 PROBLEM FORMULATION.....</b>	<b>10</b>
<b>1.3 RESEARCH AIM AND OBJECTIVES .....</b>	<b>12</b>
<b>1.4 METHODOLOGY.....</b>	<b>13</b>
<b>CHAPTER 2: REMOTE SENSING OF VEGETATION .....</b>	<b>17</b>
<b>2.1 FUNDAMENTALS OF REMOTE SENSING .....</b>	<b>17</b>
<b>2.1.1 Electromagnetic spectrum.....</b>	<b>17</b>
<b>2.1.2 Atmospheric and anisotropic effects .....</b>	<b>18</b>
<b>2.1.3 Sensor type and platform .....</b>	<b>20</b>
<b>2.1.4 Characteristics of optical images .....</b>	<b>22</b>
<b>2.2 IMAGE CLASSIFICATION .....</b>	<b>23</b>
<b>2.2.1 Image preparation.....</b>	<b>24</b>
<b>2.2.2 Segmentation.....</b>	<b>26</b>
<b>2.2.3 Feature generation .....</b>	<b>27</b>
<b>2.2.4 Feature reduction .....</b>	<b>28</b>
<b>2.2.5 Classification.....</b>	<b>29</b>
<b>2.2.6 Accuracy assessment.....</b>	<b>30</b>

<b>2.3</b>	<b>SPECIES-SPECIFIC VEGETATION MAPPING .....</b>	<b>31</b>
2.3.1	Spectral properties of vegetation .....	31
2.3.2	Remote sensing of species specific plant traits.....	32
2.3.3	Features and classification methods .....	34
2.3.4	Regional VHR vegetation mapping .....	35
<b>2.4</b>	<b>LITERATURE EVALUATION .....</b>	<b>37</b>
 <b>CHAPTER 3: RADIOMETRIC HOMOGENISATION OF AERIAL</b>		
<b>IMAGES BY CALIBRATING WITH SATELLITE DATA.....</b>		
<b>3.1</b>	<b>ABSTRACT .....</b>	<b>39</b>
<b>3.2</b>	<b>INTRODUCTION.....</b>	<b>40</b>
<b>3.3</b>	<b>METHODS .....</b>	<b>42</b>
3.3.1	Formulation of the local linear model .....	42
3.3.2	Parameter estimation.....	44
3.3.3	Incorporation of viewing geometry and relative spectral response effects....	46
3.3.4	Surface reflectance homogenisation .....	47
3.3.5	Study site, data collection and preparation .....	48
3.3.6	Linearity of band averaged values.....	51
3.3.7	Accuracy assessment .....	51
<b>3.4</b>	<b>RESULTS AND DISCUSSION.....</b>	<b>53</b>
3.4.1	Band averaged relationships .....	53
3.4.2	Mosaicking .....	54
3.4.3	MODIS statistical comparison .....	56
3.4.4	SPOT 5 statistical comparison .....	57
<b>3.5</b>	<b>CONCLUSIONS.....</b>	<b>62</b>
 <b>CHAPTER 4: FEATURE CLUSTERING AND RANKING (FCR) FOR</b>		
<b>SELECTING STABLE FEATURES FROM HIGH DIMENSIONAL</b>		
<b>REMOTELY SENSED DATA.....</b>		
<b>4.1</b>	<b>ABSTRACT .....</b>	<b>65</b>
<b>4.2</b>	<b>INTRODUCTION.....</b>	<b>65</b>
<b>4.3</b>	<b>METHODS .....</b>	<b>68</b>
4.3.1	Related work.....	68
4.3.2	Feature clustering and ranking.....	70
4.3.3	Data sets .....	72
4.3.4	Experimental design.....	73

4.4	RESULTS AND DISCUSSION.....	77
4.5	CONCLUSIONS.....	80
<b>CHAPTER 5: REGIONAL MAPPING OF SPEKBOOM CANOPY COVER USING VERY HIGH RESOLUTION AERIAL IMAGERY.....</b>		<b>81</b>
5.1	ABSTRACT .....	81
5.2	INTRODUCTION.....	81
5.3	DATA.....	85
5.3.1	Study area .....	85
5.3.2	Imagery .....	85
5.3.3	Data collection .....	86
5.4	METHODS AND EXPERIMENTS .....	89
5.4.1	Radiometric homogenisation.....	89
5.4.2	Mapping methodology .....	90
5.4.3	Features .....	90
5.4.4	Feature selection.....	93
5.4.5	Classification and spekboom canopy cover estimation.....	94
5.4.6	Evaluation .....	96
5.5	RESULTS.....	97
5.5.1	Feature selection.....	97
5.5.2	Classification and spekboom canopy cover estimation.....	98
5.6	DISCUSSION .....	102
5.6.1	Feature selection.....	102
5.6.2	Classification and spekboom canopy cover estimation.....	103
5.7	CONCLUSIONS.....	104
<b>CHAPTER 6: DISCUSSION AND CONCLUSIONS.....</b>		<b>106</b>
6.1	SUMMARY.....	106
6.2	CRITICAL EVALUATION.....	108
6.3	RECOMMENDATIONS .....	110
6.4	CONCLUSIONS.....	111
<b>REFERENCES.....</b>		<b>112</b>

## TABLES

Table 2.1	Main divisions of the electromagnetic spectrum .....	18
Table 2.2	Representative sources of satellite and aerial imagery and their specifications .....	22
Table 3.1	Statistical comparison between MODIS and DMC surface reflectance images .....	57
Table 3.2	Statistical comparison between SPOT 5 and MODIS surface reflectance images .....	58
Table 3.3	Statistical comparison between SPOT 5 and DMC surface reflectance images .....	58
Table 4.1	Data sets .....	72
Table 4.2	Methods as combinations of search schemes and criteria.....	75
Table 4.3	Feature selection parameters .....	76
Table 4.4	Method cumulative execution time over all data .....	79
Table 4.5	Non-dominated ranking of methods by accuracy and stability.....	79
Table 5.1	In situ canopy cover data.....	87
Table 5.2	Class descriptions.....	88
Table 5.3	Details of labelled pixel data.....	88
Table 5.4	Features .....	92
Table 5.5	Classifier parameters .....	95
Table 5.6	Ranked clusters .....	98
Table 5.7	Classifier performance comparison.....	98
Table 5.8	Classifier computation times.....	99
Table 5.9	Decision tree three-class confusion matrix .....	99
Table 5.10	Decision tree two-class confusion matrix .....	99
Table 5.11	Decision tree canopy cover estimates .....	100

## FIGURES

Figure 1.1 Intact and degraded subtropical thicket .....	1
Figure 1.2 Arid thicket in the Little Karoo, with spekboom in the foreground .....	2
Figure 1.3 Extent of solid and mosaic thicket forms.....	3
Figure 1.4 Extent of arid thicket, valley thicket and thicket structural types.....	3
Figure 1.5 Arid thicket near Oudtshoorn, dominated by spekboom .....	5
Figure 1.6 Research design and thesis structure .....	16
Figure 2.1 Direct ( $I_S$ ) and scattered ( $I_O$ and $I_D$ ) radiation components reaching the sensor.....	19
Figure 2.2 BRDF scattering modes .....	20
Figure 2.3 Sensor types and spectral ranges .....	20
Figure 2.4 Airborne survey flight plan.....	21
Figure 2.5 Remote sensing image characteristics .....	23
Figure 2.6 Typical vegetation reflectance spectra.....	32
Figure 3.1 Study area orientation map .....	49
Figure 3.2 DMC's and MODIS's RSR .....	49
Figure 3.3 SPOT 5 scene and mosaic extents .....	52
Figure 3.4 DMC versus MODIS simulated band averaged relationship for typical surface reflectances ((a) NIR, (b) red, (c) green and (d) blue bands) .....	53
Figure 3.5 Uncalibrated mosaic on MODIS reference image background .....	54
Figure 3.6 Homogenised mosaic on MODIS reference image background .....	55
Figure 3.7 Reduction of hot spot and seam lines, with (a) showing raw DN images including hot spot and seam lines and (b) the corrected surface reflectance image.....	55
Figure 3.8 DMC DN values and MODIS surface reflectance correlation for the (a) NIR, (b) red, (c) green and (d) blue bands .....	56
Figure 3.9 DMC homogenised mosaic and MODIS surface reflectance correlation for the (a) NIR, (b) red, (c) green and (d) blue bands.....	57
Figure 3.10 DMC and SPOT 5 RSRs.....	59
Figure 3.11 DMC DN mosaic and SPOT 5 surface reflectance correlation for the (a) NIR, (b) red and (c) green bands.....	59
Figure 3.12 DMC homogenised mosaic and SPOT 5 surface reflectance correlation for the (a) NIR, (b) red and (c) green bands .....	59
Figure 3.13 MODIS and SPOT 5 surface reflectance correlation for the (a) NIR, (b) red and (c) green bands .....	60
Figure 3.14 Effect of sliding window size on SPOT 5 comparison.....	60

Figure 3.15 Comparison of DMC and SPOT 5 spectra for (a) water, (b) bright sand, (c) bare ground, (d) vegetation type 1 and (e) vegetation type 2 surfaces .....	61
Figure 3.16 DMC and SPOT 5 surface reflectance comparison with (a) DMC homogenised mosaic masked to SPOT 5 extent; (b) SPOT 5 surface reflectance image; and (c) contrast stretched absolute difference image.....	61
Figure 4.1 Method stability per data set (methods along the x axis are ordered by their mean stability over the data sets) .....	77
Figure 4.2 Method accuracy per data set (methods along the x axis are ordered by their mean accuracy over the data sets) .....	78
Figure 5.1 Little Karoo study area .....	85
Figure 5.2 Study area spekboom habitats and field ground truth sites .....	86
Figure 5.3 Matjiesvlei2 canopy cover ground truth site.....	87
Figure 5.4 Example image class labels .....	88
Figure 5.5 MODIS and DMC RSRs.....	89
Figure 5.6 Methodological flowchart.....	96
Figure 5.7 Clustering of correlated features.....	97
Figure 5.8 Example spekboom canopy cover maps showing classified spekboom boundaries (a) Matjiesvlei (Habitat: arid thicket with spekboom), (b) Groenfontein (Habitat: valley thicket with spekboom); (c) Rooiberg (Habitat: arid thicket with spekboom and fynbos mosaic); and (d) Grootkop (Habitat: arid thicket with spekboom and succulent Karoo mosaic).....	101
Figure 5.9 Global spekboom canopy cover map.....	102

## ACRONYMS AND ABBREVIATIONS

AOD	Aerosol optical depth
AP	Affinity propagation
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
ATCOR	Atmospheric/topographic correction
BE	Backward elimination
BRDF	Bidirectional reflectance distribution function
CAM	Crassulacean acid metabolism
CCBA	Climate Community and Biodiversity Alliance
CIR	Colour-infrared
CRF	Conditional random field
DEM	Digital elevation model
DGPS	Differential global positioning system
DMC	Digital mapping camera
DN	Digital number
FCBF	Fast correlation based filter
FCR	Feature clustering and ranking
FOV	Field of view
FS	Forward selection
GCP	Ground control point
GLCM	Grey level co-occurrence matrix
ICA	Independent component analysis
IR	Infrared
JMI	Joint mutual information
KSC	Kennedy Space Centre
LBP	Local binary patterns
LDA	Linear discriminant analysis

MAD	Mean absolute difference
MAE	Mean absolute error
MI	Mutual information
MISR	Multi-angle Imaging SpectroRadiometer
ML	Maximum likelihood
MMAIQ	Maximal minimal associated index quotient
MODIS	Moderate resolution imaging spectroradiometer
NAIP	National Agriculture Imagery Program
NBAR	Nadir BRDF-adjusted reflectance
NDVI	Normalised difference vegetation index
NGI	National Geo-spatial Information
NIR	Near-infrared
NRF	National Research Foundation
PCA	Principal component analysis
PPS	Post-processing software
RBF	Radial basis function
RGB	Red, green and blue
RMS	Root mean square
RSR	Relative spectral response
RTM	Radiometric transfer model
RVI	Ratio vegetation index
SAE	Standard deviation of absolute errors
SAM	Spectral angle mapping
SMA	Spectral mixture analysis
SMFS	Supervised multiview feature selection
SPOT	Satellite Pour l'Observation de la Terre
SRM	Statistical region merging



STRP	Subtropical Thicket Restoration Programme
SVM	Support vector machine
UAV	Unmanned aerial vehicle
VCS	Verified Carbon Standard
VHR	Very high resolution

## CHAPTER 1: INTRODUCTION

The subtropical thicket biome occurs predominantly in the southern and south-eastern parts of South Africa (Vlok & Euston-Brown 2002) and is characterised by a dense combination of small trees, spinescent shrubs, forbs and succulents in its pristine state. Poorly managed goat browsing has degraded this habitat over much of its range (Lechmere-Oertel, Kerley & Cowling 2005; Lloyd, Van den Berg & Palmer 2002; Mills et al. 2005), resulting in severe reductions in biodiversity and ecological functioning. Removal of browsing pressure does not bring about recovery, and restoration of these degraded habitats appears unlikely to occur without active intervention (Lechmere-Oertel, Kerley & Cowling 2005; Sigwela et al. 2009). The planting of spekboom (*Portulacaria afra*), a dominant and ecologically important thicket species, has been identified as a practical means of restoring degraded areas (Mills & Cowling 2006; Powell, Mills & Marais 2005). Figure 1.1 shows an example of degraded and intact thicket across a fenceline in the south-eastern Cape. Spatial data, including spekboom canopy cover maps, are required to facilitate the planning, monitoring and funding of this important restoration work (Powell 2009).



Figure 1.1 Intact and degraded subtropical thicket

Source: Restoration Research Group (2012)

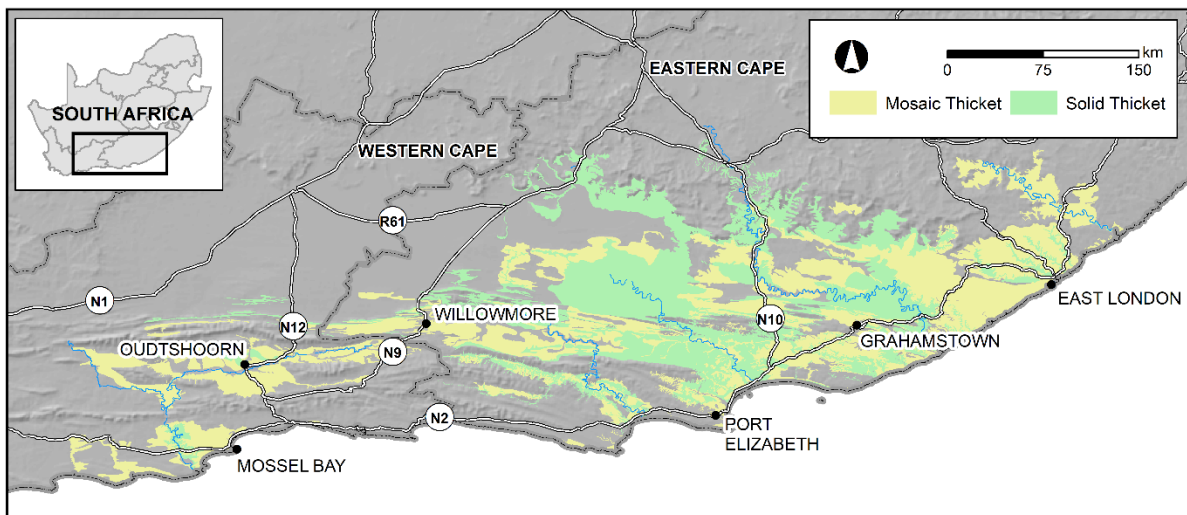
## 1.1 BACKGROUND

### 1.1.1 Degradation of subtropical thicket

In recent years, valuable work has been done to describe the character and extent of subtropical thicket (Cowling, ProcheŃ & Vlok 2005; Lechmere-Oertel, Kerley & Cowling 2005; Vlok, Euston-Brown & Cowling 2003). It is only fairly recently that subtropical thicket has been recognised as a biome (Hoare et al. 2006; Tinley 1975; Vlok & Euston-Brown 2002), and this complex habitat remains a topic of ongoing research. The general character of subtropical thicket can be seen in Figure 1.2, which shows an arid thicket habitat containing spekboom. Thicket vegetation comprises a high diversity of species and growth forms (Cowling, ProcheŃ & Vlok 2005). Thicket can occur as extensive solid stands or as scattered bush clumps in a background mosaic of different vegetation (Vlok & Euston-Brown 2002) (known as “solid” and “mosaic” forms respectively). Two main types of thicket are recognised: mainland and dune thicket (Vlok & Euston-Brown 2002). Mainland thicket is further sub-divided into arid thicket, valley thicket and thicket structural types, which occur along a gradient of increasing rainfall. The mainland thicket types can each occur in solid or mosaic form. Arid and valley thicket types occur in drier, arid to semi-arid conditions and often contain an abundance of spekboom (Vlok, Euston-Brown & Cowling 2003). Figure 1.3 shows the extents of the solid and mosaic thicket forms, and Figure 1.4 shows the extents of the arid thicket, valley thicket and other thicket types. There is an estimated 4.4 million ha of subtropical thicket habitat in southern and south-eastern South Africa (Lloyd, Van den Berg & Palmer 2002).

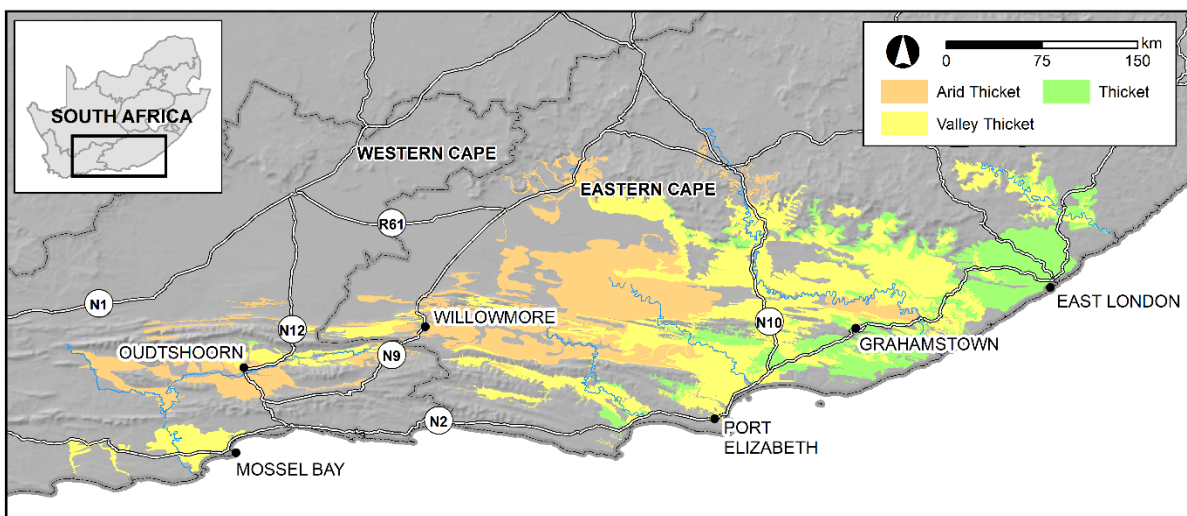


Figure 1.2 Arid thicket in the Little Karoo, with spekboom in the foreground



Adapted from Vlok &amp; Euston-Brown (2002)

Figure 1.3 Extent of solid and mosaic thicket forms



Adapted from Vlok &amp; Euston-Brown (2002)

Figure 1.4 Extent of arid thicket, valley thicket and thicket structural types

Over-stocking with domestic livestock has resulted in extensive degradation of subtropical thicket. While thicket is tolerant of browsing by indigenous herbivores, it is highly susceptible to over-browsing by goats (Moolman & Cowling 1994; Stuart-Hill 1992). Sustained goat browsing thins the canopy, altering the under-storey micro-climate and exposing the rich layer of litter to dispersal (Lechmere-Oertel, Kerley & Cowling 2005) and the topsoil to erosion (Mills & Cowling 2010). Most thicket species do not germinate in exposed conditions (Adie & Yeaton 2013) but require the favourable under-storey micro-environment in which to reproduce (Adie & Yeaton 2013; Sigwela et al. 2009). With little successful recruitment and sustained browsing, thicket is gradually transformed into sparsely scattered clumps, remnant trees and a covering of herbs (Lechmere-Oertel, Kerley & Cowling 2005). Degradation of this kind can occur in the space of a few decades, or possibly less (Hoffman & Cowling 1990; Kerley, Knight & De Kock 1995). It is thought that

the savannah-like degraded state is not a new stable state, but rather an intermediate step on a path towards a highly desertified landscape, consisting only of transitory (ephemeral) grasses and forbs (Lechmere-Oertel, Kerley & Cowling 2005). Thicket degradation is accompanied by associated losses in biodiversity (Lechmere-Oertel, Kerley & Cowling 2005), carbon storage (Mills et al. 2005), soil health (Lechmere-Oertel et al. 2008) and hydrological functioning (Van Luijk et al. 2013). An estimated 90% of subtropical thicket has been moderately or severely degraded (Lloyd, Van den Berg & Palmer 2002; Thompson et al. 2009).

### **1.1.2 Spekboom and its role in subtropical thicket**

Spekboom (*Portulacaria afra*), an evergreen succulent shrub with small fleshy leaves, is a dominant and ecologically important species in the valley and arid thicket types that can grow up to 2.5 m (Vlok, Cowling & Wolf 2005). It favours sunnier northern slopes and produces an unusually large quantity of leaf litter for a semi-arid plant (Lechmere-Oertel et al. 2008). The ability to shift between CAM (crassulacean acid metabolism), where plants absorb carbon dioxide at night to improve water efficiency, and C3 photosynthetic pathways, depending on water availability (Guralnick & Gladsky 2017; Guralnick & Ting 1986), likely contributes to its high productivity (Mills & Cowling 2006). Spekboom forms a dense, closed canopy which impedes rain throughfall (Cowling & Mills 2011), inhibiting decomposition of leaf litter and thereby improving incorporation of carbon into the soil. The water-holding capacity and fertility of the soil is substantially enhanced by the accumulation of soil organic matter (Adie & Yeaton 2013; Lechmere-Oertel et al. 2008; Mills & Cowling 2010; Mills & Fey 2004). The provision of shade, improved soil quality and moisture retention create a favourable environment for the germination of other plants (Adie & Yeaton 2013; Mills & Cowling 2010). Spekboom is considered to be a keystone species (Lechmere-Oertel et al. 2008) or “ecosystem engineer” (Van der Vyver et al. 2013) in subtropical thicket due to the role it plays in soil health and plant recruitment, but it is vulnerable to degradation. Due to its palatability, spekboom is the first canopy species to be over-browsed (Van Luijk et al. 2013; Vlok, Euston-Brown & Cowling 2003) and is often completely absent from degraded landscapes. Of the 2.1 million ha of solid thicket habitat, roughly 1.7 million ha supports spekboom (Lloyd, Van den Berg & Palmer 2002). An example of the growth form and dense canopy of spekboom can be seen in Figure 1.5, which shows a spekboom dominated slope in an arid thicket habitat.



Figure 1.5 Arid thicket near Oudtshoorn, dominated by spekboom

### 1.1.3 Spekboom restoration

Transformed subtropical thicket does not recover spontaneously, and restoration requires active intervention to reverse the desertification trajectory (Sigwela et al. 2009; Vlok, Cowling & Wolf 2005; Van der Vyver et al. 2013). Mass planting of spekboom has been identified as a viable means for restoring degraded arid and valley thicket habitats (Mills et al. 2007; Mills & Cowling 2006; Powell, Mills & Marais 2005). Spekboom propagates vegetatively and can be grown from truncheons (cuttings) (Stuart-Hill 1992), thereby avoiding nursery overheads (Mills & Cowling 2010). Restoration by planting spekboom is strongly motivated by its role as a keystone species (Adie & Yeaton 2013; Van der Vyver et al. 2013). It has been observed that over a period of 40 to 50 years, species composition in two spekboom restoration sites approached that of intact thicket (Van der Vyver et al. 2013). While these results are encouraging, Powell (2009) notes that the re-establishment of biodiversity is a critical and uncertain restoration outcome that warrants further research to confirm its realisation through spekboom planting.

In addition to biodiversity improvement, restoration is motivated by a number of other factors. Re-establishment of spekboom in degraded areas will have significant benefits for hydrological functioning, including prevention of soil erosion, reduction in flood severity and improved water quality (Van Luijk et al. 2013; Mills & Cowling 2006). Spekboom restoration also includes

benefits for local communities. Increased biodiversity could boost income from tourism and provide sustained access to wood, fruit and medicines (Mills et al. 2007). If restoration is implemented on a large scale, it can offer employment for thousands of workers (Mills et al. 2010). This is particularly relevant for the Eastern Cape where unemployment rates are the highest in South Africa at roughly 34 % (Statistics South Africa 2018).

For an arid region plant, spekboom is unusually effective at storing (sequestering) carbon. Measurements at established restoration sites in Mills & Cowling (2006) and Van der Vyver et al. (2013) indicate that average rates of carbon storage lie between 1.2 and 4.2 t C ha<sup>-1</sup> yr<sup>-1</sup>. Comparisons of degraded and intact sites show that restoration could sequester the order of 80 t C ha<sup>-1</sup> in combined above and below-ground carbon (Marais, Cowling & Powell 2009; Mills et al. 2005; Powell 2009). In one instance, estimated carbon lost through degradation was as high as 168 t C ha<sup>-1</sup> (Van der Vyver et al. 2013). These numbers compete with the carbon storage capacity of mesic forests (Mills et al. 2005). Aside from the benefits to ecosystem functioning, carbon sequestration through restoration has obvious relevance in the context of climate change adaptation and mitigation (Clarke, Shackleton & Powell 2012; Mills et al. 2010; Powell 2009).

Funding for spekboom restoration can be generated by verifying and trading stored carbon as credits on the international carbon market (Marais, Cowling & Powell 2009; Mills et al. 2007). This trade is facilitated by organisations such as Verified Carbon Standard (VCS) and Climate Community and Biodiversity Alliance (CCBA) who audit carbon accounting and environmental practice (Mills & Cowling 2014). Verified carbon credits can be purchased by countries or corporations wishing to offset their carbon emissions (they may want to do this voluntarily to meet carbon tax requirements or to satisfy international agreements such as the Kyoto Protocol and Paris Agreement). To date, 4.7 million ex ante carbon credits have been generated under VCS and CCBA by two thicket restoration projects (Mills et al. 2015).

The Subtropical Thicket Restoration Programme (STRP) was created by the South African government in 2004 to create a restoration economy through the collaboration of government, academia and implementation agencies (Mills et al. 2010). More than 10 000 ha of degraded thicket has been planted with spekboom cuttings as part of restoration experiments conducted under this programme (Mills et al. 2015). Unfortunately, these experiments suffered from a number of shortcomings, including poor plot placement, entrenched planting practice and lack of scientific oversight, resulting in low spekboom survivorship (Mills et al. 2015; Mills & Robson 2017). There is a pressing need to rectify these shortcomings and increase the pace and scale of thicket restoration. “The rate of restoration needs to increase by a factor of fifty if meaningful

progress towards restoring degraded thicket across the Eastern Cape is to be made over the next decade.” (Mills et al. 2015: 7)

#### **1.1.4 Monitoring of spekboom canopy cover**

Supporting spatial information is needed to aid in the planning and monitoring of subtropical thicket restoration. Given the focus on spekboom planting and its dominant and keystone role in thicket, species specific maps of spekboom canopy cover are of particular interest. In the restoration planning stage, spekboom canopy cover will help identify and prioritise candidate restoration areas. Measures of intact spekboom canopy cover would assist in carbon storage forecasting (Mills & Cowling 2006) and in formulating the financial case for restoration (Mills et al. 2007). Following the limited success of the STRP planting experiments, it has been recommended that future restoration protocols should be allowed to change dynamically based on emerging research and site specific results (Mills et al. 2015; Mills et al. 2018; Mills & Robson 2017). Regularly updated spekboom canopy cover can be used to monitor restoration progress and inform changes to protocol as necessary. Accurate canopy cover monitoring may also contribute to the formulation of a biomass estimation technique, that would assist with the carbon accounting required for trading credits on the international market (Mills et al. 2010).

Currently, spekboom canopy cover is measured manually in the field. A number of different approaches have been used. In Mills & Cowling (2006), carbon stocks were estimated using measured spekboom canopy cover at two sites. At the first site, canopy cover was derived from the average distance between plants and planted rows, as measured in 10 random samples. In the second site (a narrow strip of restored area), average spekboom canopy cover was measured along six line transects. Stuart-Hill (1992) established spekboom canopy cover by visual estimate in a number of small circular plots. Powell (2009) measured canopy dimensions of individual plants and approximated canopy area as the area of the circle whose diameter was the mean of the longest axis of the canopy and the axis perpendicular to it. At a much coarser resolution, regional vegetation maps have been produced by botanists through extensive field work (Vlok, Cowling & Wolf 2005; Vlok & Euston-Brown 2002). Source data for these maps is generated by in situ visual inspection of large areas (usually from elevated positions), in conjunction with hand drawing of habitat (or other) extents on printed satellite (Landsat) imagery. Vegetation habitat, degradation and spekboom canopy cover have been mapped in this way.

Field measurements are known to be time-consuming and costly (Eisfelder, Kuenzer & Dech 2012; Lu 2006). This is confounded in the subtropical biome by rugged terrain, dense vegetation and



complex growth forms (Powell 2009). In situ visual estimates can suffer from repeatability problems due to their subjective nature. Large-area field maps have been made at a necessarily coarse spatial resolution to reduce labour overheads (e.g. the Vlok & Euston-Brown (2002) map was made at a 1:100 000 and the Vlok, Cowling & Wolf (2005) map at a 1:50 000 scale). Monitoring small restoration sites and obtaining accurate spekboom canopy cover in small heterogeneous thicket clumps requires substantially finer spatial resolutions than those used for large-area field mapping. “Ideally one should employ a mapping scale that would resolve all the vegetation units to solid types. In the rather complex Little Karoo environment this probably implies a mapping scale of 1:5 000.” (Vlok, Cowling & Wolf 2005: 23) However, current field techniques for high spatial resolution canopy cover mapping are only viable for small areas (Mills & Cowling 2006; Powell 2009; Stuart-Hill 1992).

Generating a biome-wide, high spatial resolution canopy cover map using conventional field-based methods is highly impractical. A cost-effective and accurate means of high spatial resolution mapping of large areas of spekboom canopy cover is therefore needed.

### **1.1.5 Remote sensing of canopy cover**

Remotely sensed aerial or satellite images of the earth’s surface capture information that can be used for distinguishing vegetation species and determining their spatial extent. An effective remote sensing approach would offer a means of mapping spekboom canopy cover and substantially reduce the costs and repeatability concerns associated with current field-based methods.

A number of studies have applied optical remote sensing techniques to vegetation species mapping (i.e. species specific canopy cover mapping) problems. Plant spectral characteristics in the visible to short wave infrared (SWIR) range are affected by species’ distinguishing traits (Fassnacht et al. 2016; Niphadkar & Nagendra 2016). Photosynthetic absorption features occur in the visible (400–700 nm) region, while leaf and canopy traits, like internal cell scattering and water absorption, affect the near-infrared (NIR) (700–1300 nm) region (Fassnacht et al. 2016). Vegetation exhibits a sharp transition from absorption to reflection around the red-NIR boundary (690–720 nm) termed the “red-edge”. This feature has been found to be informative with regards to plant biochemical properties (Adjorlolo et al. 2012). The thermal infrared (TIR) region has also been found to be promising for species mapping (Fassnacht et al. 2016). TIR data has, however, seen little use in vegetation mapping due to availability and resolution constraints (Neinavaz 2017). While spectral properties are useful for distinguishing plant species, they are not necessarily sufficient. Plant

spectra are affected by numerous factors other than its species (Fassnacht et al. 2016). These complicating factors include health (Waser et al. 2014), age (Einzmann et al. 2014), surrounding vegetation (Clark & Roberts 2012) and view-illumination geometry.

The use of multi-spectral images for species mapping is common (Fassnacht et al. 2016; Vaz et al. 2018), due in part to the increasing availability and relative cost-effectiveness of this form of data. Hyperspectral data provide a finer spectral resolution and potentially more informative description of distinguishing spectral features compared to multi-spectral data and have been used quite successfully for species mapping in a number of cases (Ganivet & Bloomberg 2019; Nguyen et al. 2019). It is, however, seldom used for landscape or country scale mapping due to associated costs and computational requirements (Adjorlolo et al. 2012; Ganivet & Bloomberg 2019). Higher spatial resolutions provide more detailed plant information and can improve classification accuracy. Small pixel sizes allow for the identification of individual plants and help reduce the problem of spectral mixing between target species, other vegetation and soil (Fassnacht et al. 2016). VHR imagery also has the advantage of being able to describe plant spatial structure through the use of texture measures (Basu et al. 2015; Neigh et al. 2018; Niphadkar & Nagendra 2016).

Recently, a number of studies have used LiDAR (light detection and ranging) data for describing plant growth forms as a complement to spectral information (Fassnacht et al. 2016; Vaz et al. 2018). Measures of plant phenological characteristics derived from multi-temporal image sets have also been found to assist in species mapping (Niphadkar et al. 2017; Tarantino et al. 2019). The use of unmanned aerial vehicles (UAVs) for remote sensing of vegetation is an emerging area of research (Choi et al. 2016; Matese, Di Gennaro & Berton 2017). UAVs provide VHR imagery and flexible revisit times but have a limited flight range and therefore are not well suited to landscape or regional mapping (Vaz et al. 2018). While higher resolution images and complementary data sources can improve mapping accuracy, the choice of imagery is often limited by operational concerns such as cost, coverage, and available computational capacity. This is especially true for large-area studies which may require the use of thousands of images (Basu et al. 2015; Neigh et al. 2018).

Supervised classification is typically used for species mapping (Ganivet & Bloomberg 2019). In this approach, objects (image pixels or regions) with known class labels are used to build (“train”) a classifier that can predict labels for novel data. Objects are represented by a set of features or image measurements that describe spatial and spectral characteristics. To avoid overtraining, feature selection is used to reduce feature dimensionality to a salient minimum (Kononenko, Šimec

& Robnik-Šikonja 1997). By minimising the number of measured and computed features, feature selection can also assist in reducing computational overheads and measurement costs (Jain, Duin & Mao 2000). In the VHR species recognition and general vegetation mapping literature, feature selection is frequently omitted (Boyden et al. 2007; De Castro et al. 2012; Mehner et al. 2004), handled manually (Neigh et al. 2018; Tarantino et al. 2019) or executed with ranking methods (which are known to be sub-optimal and prone to instability in the presence of feature redundancy) (Basu et al. 2015; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Niphadkar & Nagendra 2016).

Radiometric and ecological variations increase over spatial and temporal scales. A major challenge in the remote sensing of vegetation is reducing or otherwise addressing these variations so that methods can generalise beyond the area(s) used to build them (Baraldi et al. 2010; Eisfelder, Kuenzer & Dech 2012; Fassnacht et al. 2016; Lu 2006). Examples of VHR vegetation mapping done at regional or country scales are very rare (Basu et al. 2015; Neigh et al. 2018). Radiometric variations arising from changing atmospheric and anisotropic effects (i.e. the dependence of reflectance on view-illumination geometry) are more pronounced when working with large-area mosaics of imagery, captured at varying times (Fassnacht et al. 2016; Gehrke & Beshah 2016; Lelong et al. 2008). Habitat, ecotype and phenological variations also increase over spatial and temporal scales and add uncertainty to the problem of classifying species (Fassnacht et al. 2016), especially in diverse and heterogeneous habitats (Bradley 2014; Ganivet & Bloomberg 2019). In the vast majority of cases, species mapping studies are applied to small areas (often consisting of a single image) (Boyden et al. 2007; Ghosh & Joshi 2014; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Mehner et al. 2004; Ouyang et al. 2011) and thus do not need to comprehensively address these sources of variation (Fassnacht et al. 2016). Despite numerous innovations, the effective separation of plant species from their surrounding environment remains a challenge (Bradley 2014).

## **1.2 PROBLEM FORMULATION**

Spekboom planting is a viable way of restoring vast areas of degraded subtropical thicket (Mills et al. 2007; Mills & Cowling 2006; Powell, Mills & Marais 2005). There is a pressing need to scale up restoration work for it to have a meaningful impact on the thicket biome (Mills et al. 2015). Spatial data, including spekboom canopy cover maps, are needed to support and inform these restoration efforts. Accurate high-resolution spekboom canopy cover will assist in prioritising areas for restoration, monitoring small planting stands, informing changes in planting protocol and providing source data for carbon accounting. Current methods of high-resolution

spekboom canopy cover mapping are costly, labour-intensive and only feasible for small areas. While larger canopy cover maps have been produced by in situ visual assessment, they are not suited to accurate restoration monitoring because of their coarse resolution and lack of reproducibility. A remote sensing technique for regional VHR mapping of spekboom canopy cover would be a valuable tool for restoration planning and monitoring.

Commercial VHR satellite imagery is costly at a regional scale and not a practical source for biome-wide spekboom canopy cover mapping (e.g. WorldView-2/3 imagery would cost in the region of R5 million to cover the approximately 1.7 million ha of spekboom rich thicket (LandInfo 2018)). The Chief Directorate: NGI, a component of the South African Department of Rural Development and Land Reform, acquires and supplies 0.5 m pixel<sup>-1</sup> multi-spectral aerial imagery with blue, green, red and NIR bands (National Geo-spatial Information 2012). Aerial surveys are repeated to achieve national coverage every three to five years. This free of charge, VHR imagery is a promising source of data for regional vegetation mapping. Unfortunately, as is typical with aerial imagery (Chandelier & Martinoty 2009; Collings et al. 2011; López et al. 2011), it contains pronounced radiometric variations due to atmospheric and anisotropic BRDF effects. These variations add uncertainty to spectral and spatial measures and are not conducive to quantitative remote sensing of large areas. Calibrating imagery to surface reflectance is beneficial, as reflectance is invariant to atmospheric and anisotropic effects.

Existing methods for radiometrically calibrating large mosaics of VHR imagery to surface reflectance require placement of known reflectance targets (Collings et al. 2011) or field reflectance measurements (López et al. 2011) at the time of survey. These field work components are time-consuming and costly (Chander, Meyer & Helder 2004; Gao et al. 2013; Liu et al. 2004), and not possible with historical imagery (such as the NGI archive). A method for calibrating aerial imagery to surface reflectance (without field reflectance measurements or placement of targets) is needed to prepare it for quantitative remote sensing. Such a method would be an important first step towards accurate spekboom canopy cover mapping of large spatial extents and possible repeat mapping for monitoring restoration progress. It would also be of benefit to large-area VHR remote sensing in general.

Computational efficiency affects operational effectiveness and flexibility and is an important consideration for large-area VHR applications (Basu et al. 2015; Neigh et al. 2018), requiring the processing of thousands of images. In addition to addressing the “curse of dimensionality” (Jain, Duin & Mao 2000), feature selection provides a means of reducing computational overhead by reducing dimensionality. Commonly used image features can carry different computational and

measurement costs, but conventional feature selection methods do not allow for the incorporation of these costs into the selection process. Relatively little is understood about how high dimensional redundant feature spaces affect stability and optimality of feature selection in a remote sensing context. The trend towards increasing dimensionality of images and extracted features requires a structured approach to feature selection that is robust to high dimensional, redundant feature spaces and considers feature computation and measurement costs.

Based on the importance of spekboom canopy cover maps for restoration and the need for related radiometric correction and feature selection investigations, the following research questions have been set:

1. To what extent can radiometric variation in mosaics of aerial imagery be reduced without the need for in situ target placement or reflectance measurements?
2. How do high dimensional, redundant feature spaces affect the stability and optimality of feature selection in remote sensing problems?
3. How can feature selection be achieved in a way that is robust to high dimensional redundant spaces and allows the consideration of feature computation and measurement costs?
4. Which features and classification techniques are suited to the regional mapping of spekboom canopy cover with VHR multi-spectral aerial imagery?
5. Can VHR multi-spectral aerial imagery be used to accurately map spekboom canopy cover over a large region in a semi-automated manner?

### **1.3 RESEARCH AIM AND OBJECTIVES**

The research aim is to develop a regional, cost-effective, semi-automated spekboom canopy cover mapping technique using multi-spectral VHR aerial imagery. The technique is intended for application in the Little Karoo, a large semi-arid region in South Africa.

To achieve the research aim, the objectives are to:

1. Review the literature on vegetation species (canopy cover) mapping, radiometric correction and feature selection.
2. Acquire suitable satellite and aerial imagery for radiometric correction. Acquire machine learning data sets for feature selection comparison. Collect in situ spekboom

canopy cover ground truth and create labelled image data for classifier training and validation.

3. Develop and evaluate a method to radiometrically correct mosaics of aerial imagery without the need for in situ reflectance measurements or placement of known reflectance targets.
4. Develop a feature selection method that addresses the problems of sub-optimality and instability associated with high dimensional redundant feature spaces and allows consideration of feature computation and measurement costs.
5. Compare the performance of the feature selection method to conventional approaches on typical high dimensional remote sensing datasets.
6. Using the established radiometric correction and feature selection methods, develop and evaluate a regional spekboom canopy cover mapping method.

## **1.4 METHODOLOGY**

The research was experimental and explorative in nature. An inductive approach was used to develop quantitative radiometric correction and feature selection methods. Machine learning techniques, in conjunction with the developed radiometric correction and feature selection methods, were used to generate a spekboom canopy cover map of the study area.

Image processing (interpolation), curve fitting (least squares regression) and statistical (regression analysis) techniques were used in the formulation of the radiometric correction technique, which contributed to research questions one and five. This technique utilises a reference satellite surface reflectance image, to which the aerial imagery is calibrated. An experiment was conducted to quantitatively and qualitatively evaluate the technique, which further contributed to research question one. The experiment calibrated NGI aerial imagery to a MODIS surface reflectance reference. Calibrated images were qualitatively evaluated by visual inspection and quantitatively evaluated by comparison to MODIS and SPOT 5 surface reflectance data. An additional experiment was conducted to test assumptions made in the formulation of the method. The linearity of the relationship between MODIS and Intergraph DMC (digital mapping camera) band averaged values was quantitatively analysed using regression analysis. The experiment used reflectance spectra from the ASTER spectral library (Baldrige et al. 2009), in conjunction with MODIS and Intergraph DMC relative spectral response (RSR) data.

The feature selection method, developed using statistical (mutual information and affinity propagation) and machine learning (naive Bayes classification) components, contributed to research questions three, four and five. The method aims to identify relevant features in high dimensional redundant feature spaces and to be stable in response to data perturbations. An evaluative experiment was performed to quantitatively compare stability and accuracy performance of the feature selection method to conventional approaches. A number of machine learning data sets (i.e. data with class labels) were used for this comparison. These data sets comprised two multi-spectral datasets from the UCI Machine Learning Repository (Lichman 2013), two publicly available hyperspectral data sets (GIC 2014) and two primary data sets (i.e. a synthetic data set and a spekboom canopy cover data set derived from aerial imagery).

The spekboom canopy cover mapping experiment aimed to answer research question four and five. The experiment was conducted on a large set of NGI aerial imagery, in conjunction with two empirical data sets, which were created to assist in training and evaluation. The first empirical data set (called “in situ canopy cover”) comprised in situ spekboom canopy cover measurements, which were acquired in collaboration with a botanist. The second empirical data set (called “labelled pixel data”) consisted of per-pixel class labels for selected regions in the aerial imagery. The aerial imagery was quantitatively prepared using the developed radiometric correction method. Subsequently, the developed feature selection method was applied to reduce features extracted from the labelled pixel data. Quantitative analysis of feature redundancy and relevance aimed to contribute to research question four. An experiment was conducted to quantitatively compare the performance of different machine learning techniques (support vector machine, random forest, decision tree, k-nearest neighbour and maximum likelihood classifiers) on the labelled pixel and in situ canopy cover data. This experiment contributed to research questions four and five. Lastly, the best performing classifier was applied to radiometrically calibrated aerial imagery of the study area, and the resulting spekboom canopy cover map was qualitatively evaluated by visual inspection. This final evaluation further contributed to research question five.

Figure 1.6 shows a block diagram of the research design and how it relates to the thesis structure. This chapter gives background on the real world problem, formulates the research problem and questions, and sets aims and objectives. Chapter 2 gives a general overview of optical remote sensing and pattern recognition, as well as a more specific review of radiometric correction, feature selection and classification, as background for the choice of methods used in the remaining chapters. Chapter 3 presents the development and evaluation of the radiometric correction method. The development of the feature selection method, and its comparison to conventional techniques,

is described in Chapter 4. The methods described in Chapters 3 and 4 are used in Chapter 5 to establish suitable features and classifiers for spekboom canopy cover mapping. Chapter 5 also describes the generation and evaluation of a regional spekboom canopy cover map. Chapters 3, 4 and 5 were published as independent research articles, which made some repetition of study area description and background motivation for choice of methods necessary. Study areas are described in Chapters 3 and 5 in the context of the methods presented in those chapters. Chapter 6 summarises and evaluates the research in the context of spekboom restoration. Results are related back to the aims and objectives and research questions. Recommendations are made for further research.



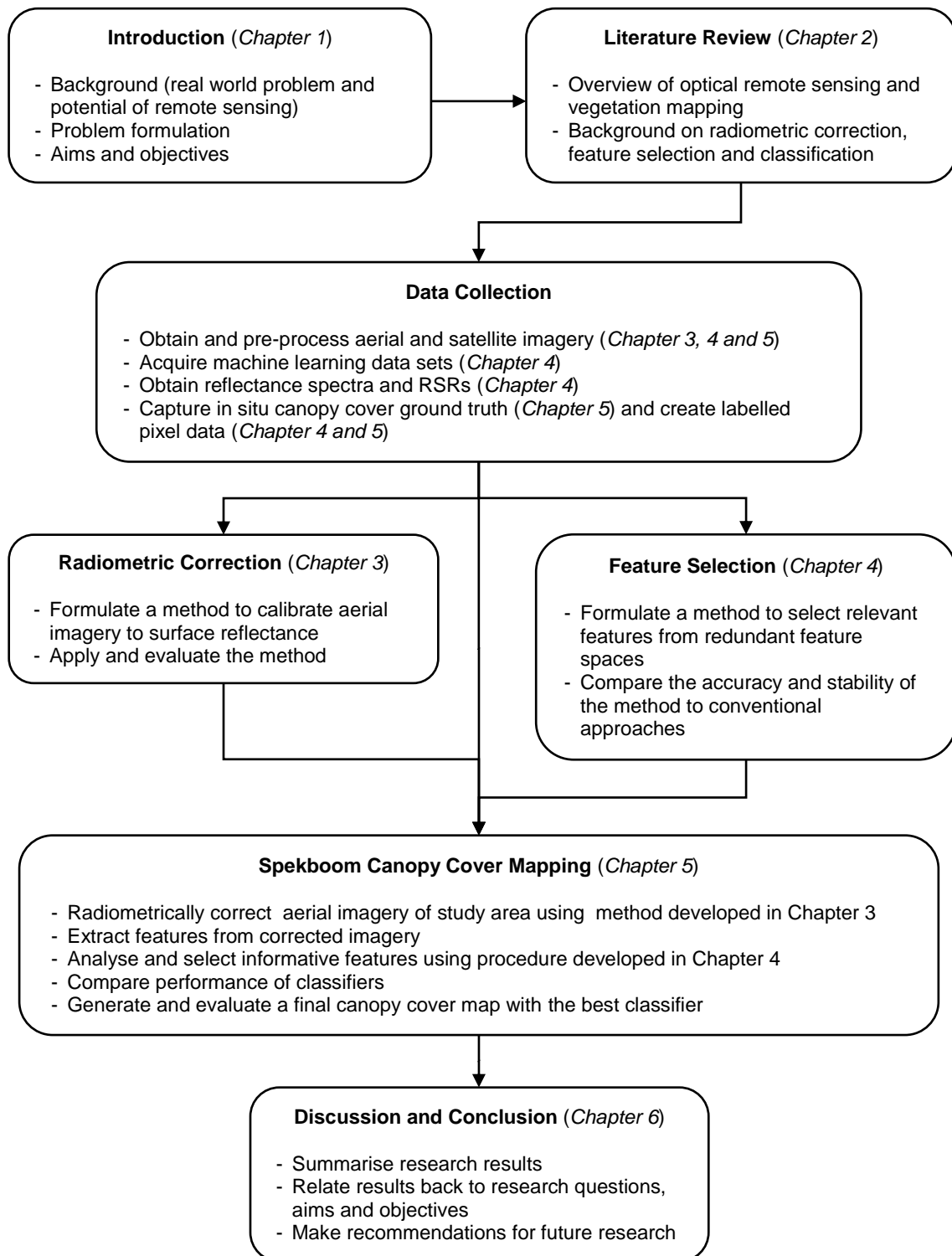


Figure 1.6 Research design and thesis structure

## CHAPTER 2: REMOTE SENSING OF VEGETATION

### 2.1 FUNDAMENTALS OF REMOTE SENSING

#### 2.1.1 Electromagnetic spectrum

Electromagnetic (EM) radiation is energy that is emitted and reflected by all objects. One way of modelling the behaviour of this energy is as a wave, which propagates as continuous oscillating electrical and magnetic fields (Tempfli et al. 2009). The wave model describes most of the EM radiation behaviour relevant to remote sensing. EM energy also behaves as if it were comprised of discrete particles (photons). The particle model is applicable to the process by which a sensor measures EM radiation. The EM spectrum refers to the wavelength ranges of different types of EM radiation (see Table 2.1). The human eye is sensitive to visible light (VIS), a very small part of the spectrum, occurring in the 0.38 to 0.72  $\mu\text{m}$  wavelength range. The optical spectrum consists of the UV, VIS and IR spectral regions (0.3  $\mu\text{m}$  to 1 mm wavelength range). EM radiation in this range obeys the laws of optics and can be focused and refracted with lenses. Wavelengths shorter than those of visible light are mostly scattered or absorbed by the atmosphere and generally not used for remote sensing of the earth surface (Mather & Koch 2011). The infrared portion of the spectrum is much wider than the visible portion and consists of wavelengths longer than those of the visible range. Together with visible light, the near-infrared (NIR) and short wave infrared (SWIR) portions form part of the “reflective spectrum” (0.38 to 3.0  $\mu\text{m}$ ) i.e. radiation that is subject to reflection from the earth’s surface. Thermal infrared (TIR) and far infrared spectral regions are emitted, rather than reflected, from the earth surface (Campbell & Wynne 2011). Microwaves occur in the 1 mm to 30 cm range and are the longest wavelengths used in remote sensing.

The sun is the main source of EM energy for Earth observation and emits radiation across most of the EM spectrum, with the bulk of the energy occurring in the visible light and NIR portions of the spectrum (Mather & Koch 2011). Objects in the atmosphere and on the earth’s surface interact with EM radiation so as to alter its spectral content in distinctive ways. The main basis of remote sensing is to use measurements of spectral and spatial patterns of EM radiation to distinguish and characterise objects based on these alterations (Campbell & Wynne 2011).

Table 2.1 Main divisions of the electromagnetic spectrum

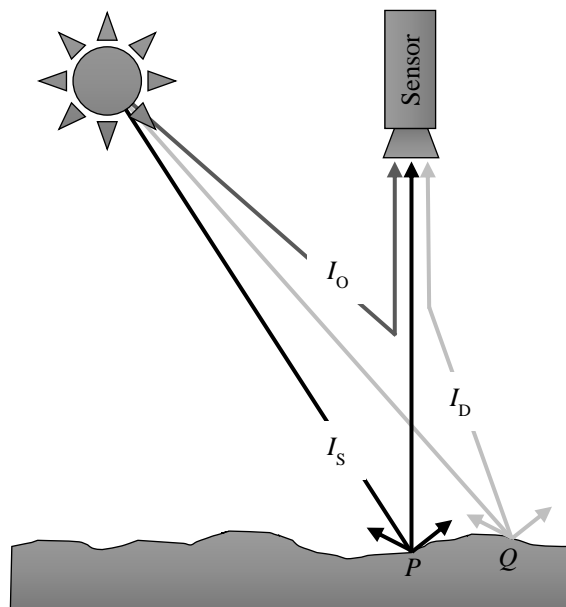
Division	Wavelength Range
Gamma rays	< 0.03 nm
X-rays	0.03 – 300 nm
Ultraviolet (UV)	0.30 – 0.38 $\mu\text{m}$
Visible light (VIS)	0.38 – 0.72 $\mu\text{m}$
Infrared	
Near-infrared (NIR)	0.72 – 1.30 $\mu\text{m}$
Short wave infrared (SWIR)	1.30 – 3.00 $\mu\text{m}$
Thermal infrared (TIR)	3.0 – 15 $\mu\text{m}$
Far infrared	15 – 1000 $\mu\text{m}$
Microwave	1 mm – 30 cm
Radio	$\geq$ 30 cm

Adapted from Liew (2001) and Campbell & Wynne (2011)

### 2.1.2 Atmospheric and anisotropic effects

Radiation interacts with the atmosphere and earth's surface through processes of scattering, absorption and reflection. Scattering deflects radiation from its path, while absorption reduces transmitted energy. These processes affect the signal measured by the sensor (Mather & Koch 2011). The amount of scattering and absorption varies with the wavelength of the radiation and the composition of atmosphere. Rayleigh scattering occurs off gaseous particles that are smaller than the wavelength of the radiation. Mie scattering occurs off larger particles, such as water droplets, dust and smoke, and manifests as haze in the measured image. As a general rule, shorter wavelengths of radiation are subject to more atmospheric scattering (Mather & Koch 2011). Areas of the spectrum that are relatively unaffected by atmospheric scattering and absorption are called "atmospheric windows" (Campbell & Wynne 2011). These windows define the usable areas of the spectrum for remote sensing.

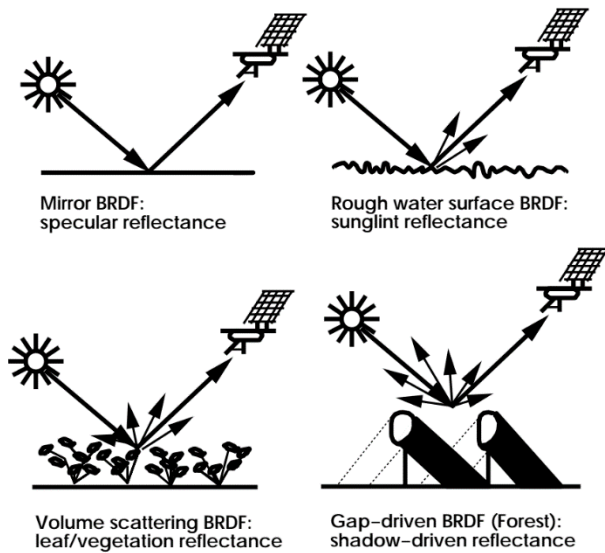
Figure 2.1 shows, in simplified form, how scattered components of radiation combine with direct path radiation to form the signal measured by the sensor. In Earth observation, the direct path radiation,  $I_S$ , is the important component as it contains information about the surface reflectance at the point of interest,  $P$ . The direct path radiation is obscured and interfered with by the scattered components,  $I_O$  and  $I_D$ , which also appear to be coming from  $P$ . The basic effect of scattering is to reduce image contrast.



Adapted from Mather & Koch (2011)

Figure 2.1 Direct ( $I_s$ ) and scattered ( $I_o$  and  $I_D$ ) radiation components reaching the sensor

Reflection is the re-direction of radiation by a surface. Smooth mirror-like surfaces exhibit specular reflectance. In this form of reflectance, all radiation is reflected in a single direction. Diffuse, or isotropic reflection, occurs when light is reflected more or less equally in all directions. This form of reflection occurs off surfaces that are rough relative to the wavelength of incident radiation (Campbell & Wynne 2011). The idealised form of this reflectance is called Lambertian reflectance, which is perfectly equal in all directions. In contrast to isotropy, anisotropy refers to a pattern of reflectance that is not uniformly distributed over all directions. Most natural surfaces exhibit some degree of anisotropy. Patterns of reflectance off the earth's surfaces change with the wavelength and orientation of the incident radiation (Campbell & Wynne 2011). Bidirectional reflectance distribution functions (BRDFs) are used to describe directional reflectance behaviour. A BRDF is a mathematical function that models the strength of reflectance in a particular direction, given the angle of illumination. Figure 2.2 shows the typical modes of scattering that cause BRDF effects. Specular reflectance occurs off smooth surfaces, like water, and manifests as “sun glint” when it is directed towards the observer. Scattering that occurs off homogenous leafy canopy is referred to as volume scattering. Gap-driven scattering, also known as geometric-optical scattering, occurs in more heterogenous vegetation where there are height differences and spaces between plants (Strahler & Muller 1999).

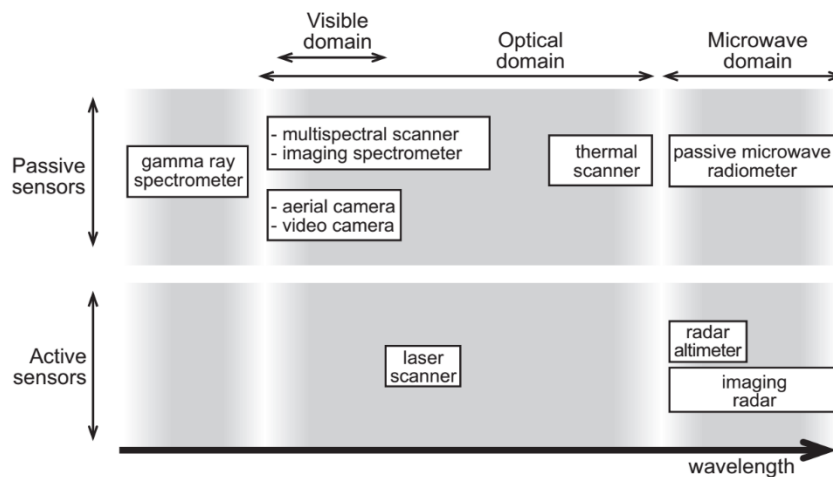


Source Strahler & Muller (1999: 45)

Figure 2.2 BRDF scattering modes

### 2.1.3 Sensor type and platform

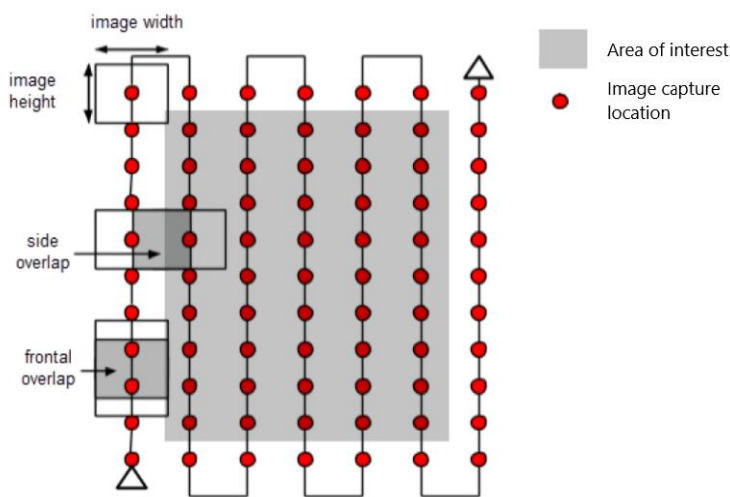
There are two basic types of remote sensing instruments: passive and active. In the passive type, the source of energy or illumination is external. Passive sensors measure solar radiation reflected from the earth's surface, or long thermal and microwave wavelengths that are emitted from the earth (Campbell & Wynne 2011). Active sensors provide their own source energy and measure the reflection of this controlled source from the earth's surface. Examples of active sensors include radar and LiDAR (light detection and ranging). Active sensors are less affected by atmospheric conditions than passive sensors and can operate in the day and night (Tempfli et al. 2009). Figure 2.3 shows examples of passive and active sensors and the spectral ranges in which they operate.



Source: Tempfli et al. (2009: 44)

Figure 2.3 Sensor types and spectral ranges

Sensors can be mounted on airborne or spaceborne (satellite) platforms. Airborne altitudes are substantially lower than those of spaceborne platforms. Due to their lower altitude, airborne sensors are capable of higher spatial resolutions than spaceborne sensors but require a much wider field of view (FOV) to cover the same ground area (Markelin et al. 2012). Lower altitudes also mean that airborne platforms are less affected by atmospheric interactions. Airborne surveys cover large areas by following a flight path of regularly spaced parallel transects (Campbell & Wynne 2011). Flight patterns and imaging intervals are chosen so that there is “frontal overlap” between successive images along a flight line and “side overlap” between neighbouring images in adjacent flight lines. An example of a flight plan is shown in Figure 2.4.



Adapted from Pix4D (2019)

Figure 2.4 Airborne survey flight plan

In contrast to airborne surveys that acquire images in a regularly spaced pattern, most Earth observation satellites follow near polar, sun-synchronous orbits (Liew 2001). A near polar orbit is one where the orbital track is slightly inclined relative to the earth’s axis. The combination of polar orbital movement with the earth’s rotation allows global or near-global coverage. Sun-synchronous orbits are designed so that successive passes over a particular location occur at the same local solar time (Campbell & Wynne 2011). This is beneficial as it reduces view-illumination variation by keeping the solar azimuth angle constant at a particular location (Mather & Koch 2011). The rate of coverage achieved by satellites is substantially higher than that of airborne platforms due to their higher ground speed (Tempfli et al. 2009). Table 2.2 shows key flight and image specifications for a popular selection of spaceborne and airborne image sources. It serves to highlight differences in spatial resolution, revisit time and image coverage between these different platforms and sensors.

Table 2.2 Representative sources of satellite and aerial imagery and their specifications

Image source	Spaceborne/ airborne	Height above Earth	Nadir swath width	Spectral bands	Spatial resolution	Revisit time
MODIS	Spaceborne	705 km	2330 km	36	Red, NIR (250 m), Blue, Green, IR (500 m), TIR (1 km)	0.5 day
Landsat-8 OLI	Spaceborne	705 km	185 km	9	Panchromatic (15 m), VIS, NIR, SWIR (30 m)	15 days
SPOT-5	Spaceborne	832 km	60 km	5	Panchromatic (2.5 m), VIS, NIR (10 m), SWIR (20 m)	26 days
Sentinel-2	Spaceborne	786 km	290 km	13	VIS, NIR (10 m), SWIR (20 m), other (60 m)	5 days
WorldView-3	Spaceborne	617 km	13 km	8	Panchromatic (0.3 m), VIS, NIR (1.2 m), SWIR (3.7 m)	1 – 5 days
NGI	Airborne	5 km	7 km	4	VIS, NIR (0.5 m)	3 – 5 years

### 2.1.4 Characteristics of optical images

A remote sensing image is formed by a regular grid of pixels. Each pixel represents a measurement of the intensity of radiation, reflected or emitted from a small portion of the earth's surface. Pixels usually consist of multiple measurements, called bands, where each band is the radiation intensity in a particular wavelength range. Figure 2.5 shows the construction of an image in terms of pixels and bands. The amount of detail or information contained in remote sensing images is determined by the spatial, spectral, radiometric and temporal resolutions. Higher resolutions require sophisticated sensor and platform hardware, as well as increased data storage, transmission and processing capacity. Practical trade-offs must therefore be made between the four resolutions when designing remote sensing systems (Campbell & Wynne 2011). Careful consideration of the characteristics of available image sources is necessary when selecting an image source for a specific application. Some form of compromise between the benefits of the four resolutions is usually required.

*Spatial resolution* refers to the size of the smallest object that can be resolved and determines the amount of spatial detail that is discernible in an image. It is specified by the size of a pixel on the earth's surface. Spatial resolutions of different image sources vary from centimetres (in the case of UAVs) to kilometres (in the case of weather satellites). For classification problems, the pixel size should ideally be smaller than the size of the object of interest to avoid spectral mixing between classes (Campbell & Wynne 2011).

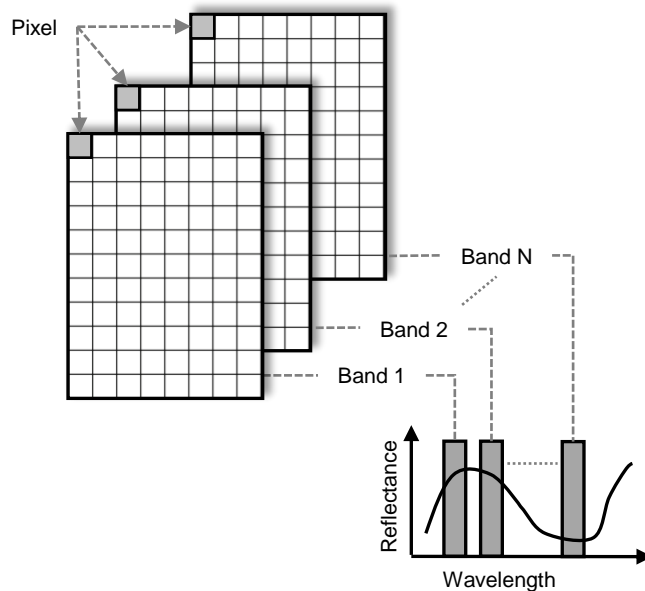


Figure 2.5 Remote sensing image characteristics

*Spectral resolution* refers to the number of image bands available and over what wavelength range those bands are distributed. Higher spectral resolution images provide a finer division of the EM spectrum. Multi-spectral images typically contain less than 20 bands and provide a relatively coarse division of the EM spectrum. Hyperspectral images may contain more than 100 bands, and provide a significantly more detailed division of the spectrum. Hyperspectral data allows for the discrimination of patterns based on subtle differences in spectra.

The smallest difference in reflectance or emittance that can be recorded by a sensor is termed the *radiometric resolution* (Tempfli et al. 2009). This resolution is determined by the sensor bit-depth (i.e. the number of digital bits used to store the value of a pixel).

*Temporal resolution*, also known as revisit time, is the time between successive images of the same area. For spaceborne sensors, it is determined by the orbit parameters and the ability of a satellite to point off-nadir. The temporal resolution of airborne surveys is variable. Surveys may be conducted once-off, as needed, or at regular intervals (e.g. as part of a government programme). A higher temporal resolution allows for the detection of rapid change on the earth's surface.

## 2.2 IMAGE CLASSIFICATION

Image classification refers to the process of assigning objects into classes based on the information contained in the image. The classification workflow follows the steps of image preparation (often called pre-processing), segmentation (optional), feature generation, feature reduction (if needed), classification and accuracy assessment. Preparing the image involves the application of pre-



processing steps to reduce unwanted geometric and radiometric variations that may detract from the accurate identification of object classes. Segmentation refers to the division of the image into meaningful objects or segments. If no segmentation is carried out, each pixel is considered an individual image object. In the feature generation stage, a set of descriptive measurements (features) are extracted from the image to represent each image object. Feature reduction is used to reduce the feature set to an informative minimum. In the classification step, the feature space is partitioned into classes by decision boundaries. Lastly, classification accuracy is evaluated by comparison to reference data. The following subsections describe these steps in more detail.

### **2.2.1 Image preparation**

Before being used for quantitative analyses, remote sensing images may require geometric and/or radiometric correction. Airborne and spaceborne images contain geometric distortions due to sensor perspective (position and orientation) and terrain elevation effects. The process of orthorectification adjusts for these effects and creates a planimetrically correct image, i.e. an image in which scale is consistent (Chuvieco & Huete 2009). Relative object positions in an orthorectified image correspond to their relative geodetic positions. Orthorectified imagery can consequently be used for spatial measurements. In some situations, geometric distortions due to orbital variations, Earth rotation and lens distortion also require correction (Mather & Koch 2011). Orthorectification requires both a description of terrain elevation, usually in the form of a digital elevation model (DEM), and a model of the sensor geometry as input. Sensor models describe how points in three-dimensional world space are transformed onto the two-dimensional image plane. These models may be defined in terms of physical principles or approximated by a rational polynomial function (Toutin 2004). Ground control points (GCPs), i.e. points whose image and geodetic co-ordinates are known, can be used to define or refine the sensor model (Tempfli et al. 2009). The use of GCPs is generally a requirement for the accurate orthorectification of high spatial resolution imagery.

Airborne and spaceborne images contain unwanted radiometric variations caused by atmospheric and BRDF effects. These variations distort or otherwise obscure spectral information that is useful for the identification or characterisation of objects. Aerial images are particularly prone to BRDF variations due to their relatively wider FOV, which results in greater viewing geometry variation across the image (Honkavaara et al. 2009; Mather & Koch 2011). In addition, long aerial survey times result in variations in illumination geometry and atmospheric effects over the captured images. Where images are used exclusively for qualitative or cartographic interpretation, these effects may be ignored or compensated for in a heuristic manner with the objective of creating an

aesthetically pleasing image (Tempfli et al. 2009). This cosmetic approach is often used for preparing mosaics of aerial imagery. Techniques such as dodging, colour balancing and histogram adaptation are used to reduce the appearance of “seam lines” (Gehrke 2010; Honkavaara et al. 2009; Laliberte et al. 2011), which are radiometric discontinuities between adjacent aerial images. Cosmetic approaches can modify (damage) spectral information and are not suited to quantitative analyses such as image classification.

A number of atmospheric correction approaches are based on the use of a radiometric transfer model (RTM), which describes scattering and absorption effects using physical principles. These approaches estimate absolute radiometric quantities, such as surface reflectance, from image pixel values (“digital numbers” (DNs)). The RTM is used in combination with knowledge of atmospheric conditions to compensate for absorption and separate the direct radiation component from its scattered counterparts (see Figure 2.1). Further details on radiative transfer can be found in Section 3.3.1. RTM approaches include atmospheric/topographic correction (ATCOR) (Richter 1997), moderate resolution atmospheric transmission (MODTRAN) (Berk et al. 1999) and second simulation of a satellite signal in the solar spectrum (6S) (Vermote et al. 1997). The benefits of correcting to absolute radiometric quantities is that these measurements have physical interpretations, can be fused with or compared to similar data and can be used in multi-temporal studies. A disadvantage of the RTM approach is that detailed knowledge of atmospheric conditions at the time of imaging is required, which is often not available. The RTM approach is also known for being complex and computationally demanding (Tempfli et al. 2009).

Another class of radiometric correction, sometimes called “image-based atmospheric correction” (Campbell & Wynne 2011), uses a linear model to approximate the relationship between image DN and ground reflectance (Tempfli et al. 2009). The parameters of the model are found using two or more reference areas of known reflectance and the DN from the corresponding image areas. The reference areas may be placed targets of known reflectance or in situ ground reflectance measurements made concurrent or near-concurrent to the time of capture. The linear model is a crude simplification of complex non-linear atmospheric and Earth surface interactions. It nevertheless suffices in practice for applications where atmospheric effects are a relatively insignificant source of error (Tempfli et al. 2009). A justification for the use of a linear model to approximate the relationship between surface reflectance and image DN can be found in Section 3.3.1.

The correction of BRDF effects is problematic, largely because BRDF varies with land cover and is seldom known a priori (Honkavaara et al. 2009; Mather & Koch 2011). BRDF correction is

most commonly achieved with the kernel-based approach (Roujean, Leroy & Deschamps 1992). The kernel approach approximates BRDF as a linear combination of weighted kernel functions that describe typical scattering modes (see Figure 2.2). Ground measurements from multiple view-illumination directions are required to estimate the kernel BRDF weights for a particular location. In practice, this is achieved using multi-temporal satellite data (e.g. MODIS (Strahler & Muller 1999)) or specialised instruments such as the multi-angle imaging spectroradiometer (MISR). MISR captures multi-directional measurements of the same ground area over a few minutes (Mather & Koch 2011). BRDF correction is important for image mosaicking, wide FOV image calibration and multi-temporal studies. It is, however, seldom addressed in practice because of the complexities involved (Vescovi & Minchella 2017).

### **2.2.2 Segmentation**

Image classification may adopt a per-pixel approach, where individual pixels are assigned to classes, or an object-based approach, where image regions or objects are the spatial units requiring classification. In remote sensing, the object-based approach is referred to as geographic object-based image analysis (GEOBIA) (Blaschke et al. 2014). Segmentation is the process of dividing the image into homogenous and meaningful image regions for later classification. In the per-pixel approach, the individual pixels are the objects to be classified, and segmentation is not required. Before the advent of VHR imagery, most image classification was conducted on a per-pixel basis as pixel sizes were generally larger than, or of a similar size to, the object of interest (Blaschke et al. 2014). The per-pixel approach is limited to extracting information from individual pixels or from fixed size sliding windows. By delineating related pixels, objects allow the extraction of richer and more specific information than is possible with individual pixels and sliding windows. The object-based approach represents a way of exploiting spatial and contextual information and has grown in popularity with the increasing availability of high spatial resolution imagery (Mather & Koch 2011). The development of the eCognition commercial software has also been instrumental in the growth of GEOBIA (Blaschke 2010).

A wide variety of segmentation approaches exist. The main challenge in identifying image objects is defining a homogeneity criterion that results in a segmentation meaningful to the problem being investigated (Hay et al. 2005). Machine learning approaches using supervised classifiers (i.e. classifiers trained with labelled example data) are increasingly being used due to their success on specialised object detection problems, such as building and vehicle detection (Cheng & Han 2016). Statistical region merging (SRM) (Nock & Nielsen 2004) is a computationally efficient region-growing approach that uses spectral differences to define objects. It was used successfully in a

ground-breaking study that mapped tree cover with VHR aerial imagery for the whole of Continental United States (Basu et al. 2015). A popular unsupervised segmentation method is the multiple resolution segmentation (MRS) algorithm (Baatz & Schäpe 2000), which is implemented in the eCognition software package (Trimble 2018). The algorithm segments at different scales using spectral and geometric properties to define homogenous regions. Watershed segmentation is a region-growing method that is sometimes used in GEOBIA (Carleer, Debeir & Wolff 2005; Hall et al. 2004). This method treats pixel values as elevations in a three-dimensional landscape, which is segmented into catchment basins (i.e. groups of pixels for which the path of steepest descent ends at the same minimum). Simpler approaches, such as Otsu segmentation (Otsu 1979), which derives thresholds from the intensity histogram, have also been applied successfully in demanding problems (Carleer, Debeir & Wolff 2005; Neigh et al. 2018).

### **2.2.3 Feature generation**

A good feature is one that is discriminative, i.e. it assists the assignment of objects into classes. Feature generation (in combination with feature reduction) can be thought of as a transformation of the raw image band space into a feature space, which improves class separability (Jain, Duin & Mao 2000). Feature design requires expert insight into the differentiating properties of the data and how these properties might be represented numerically. It is recognised as being a critical step in the classification workflow; information lost at this stage cannot be retrieved in later steps (Fassnacht et al. 2016; Pekalska & Duin 2012). Features may be generated per-object or per-pixel.

One approach for generating features is through the transformation of the raw spectral band values. Band arithmetic is used to create ratios and indices that highlight spectral features while normalising for spurious variations (Campbell & Wynne 2011). For example, the popular normalised difference vegetation index (NDVI) highlights the transition from the photosynthetically active region of the spectrum into NIR while normalising for intensity (i.e. overall illumination brightness) (Myneni et al. 1995). Transformation of raw red-green-blue (RGB) spectral bands into alternative colour spaces, such as intensity-hue-saturation, is used to suppress shadow and illumination variation by de-coupling intensity and colour information (Blauensteiner et al. 2006; Campbell & Wynne 2011). Feature extraction approaches, such as principal component analysis (PCA), are also used to generate new features as (non-)linear combinations of the raw band values (Mather & Koch 2011). More detail on feature extraction is given in Section 2.2.4. Spectral features are often initially found per-pixel. Statistics of the per-pixel features (such as mean and variance) are then extracted per-object where a GEOBIA paradigm is followed.

Another important class of features are texture descriptors. In contrast to band ratios and vegetation indices, which describe spectral relationships, texture descriptors describe spatial relationships and are calculated from a spatially connected group of pixels. Texture features are particularly useful for exploiting information in high spatial resolution imagery (Puissant, Hirsch & Weber 2005). The problem of numerically encoding complex and scale varying spatial patterns is challenging, and the development of novel texture descriptors is an active research field (Fassnacht et al. 2016; Mather & Koch 2011). At a simple level, the variance of the pixel values in a small region gives an indication of local image “roughness” (Mather & Koch 2011). More sophisticated methods for describing texture include statistics of the grey level co-occurrence matrix (GLCM) (Haralick, Shanmugam & Dinstein 1973), local binary patterns (Ojala, Pietikainen & Maenpaa 2002) and wavelet filters (Charalampidis & Kasparis 2002). Texture features may be found inside a sliding window (in the case of the per-pixel approach) or per-object (for GEOBIA).

GEOBIA provides the possibility of including features that describe object shape and size, which can be useful additions to spectral and texture descriptors. Examples of commonly used shape descriptors include area, circularity and eccentricity (Zhang & Lu 2004). Further details on typical spectral, textural and vegetation index features used in multi-spectral problems can be found in Section 5.4.3.

#### **2.2.4 Feature reduction**

In image classification, the amount of training data required to adequately represent class distributions in feature space increases exponentially as the number of features is increased – a phenomenon known as the “curse of dimensionality” (Bishop 2003). For finite training samples, increasing the features beyond a certain point results in overtraining and a decrease in the classifier accuracy. This so-called “peaking phenomenon” (Jain, Duin & Mao 2000) requires the size of the feature set to be reduced to a salient minimum in order to achieve an accurate classification. Two basic approaches to feature reduction exist: feature selection and feature extraction.

In feature extraction the feature set is mapped into a new feature space of reduced dimensionality (Webb 2002). Various criteria, such as separability and variance, are used to define the dimensions of the new ‘optimal’ space. PCA (Webb 2002) is an example of a popular feature extraction method. It uses a linear transform composed of the largest  $m$  eigenvectors of the covariance matrix to project the input features into a reduced space. Other methods such as projection pursuit and independent component analysis (ICA) also incorporate linear projections but do not rely on covariance and are thus better suited to non-Gaussian distributed data than PCA (Jain, Duin &

Mao 2000). Commonly used non-linear methods include kernel PCA (Schölkopf, Smola & Müller 1998) and multidimensional scaling (MDS) (Webb 2002). A major disadvantage of the feature extraction approach is that it requires measurements and computations to produce the full feature set, which can be prohibitively costly. Feature extraction also hinders interpretability, as it alters the original representation of the features.

Feature selection involves the selection of a subset of features from the original set according to some criterion of subset performance. A description of feature selection (its principles and typical approaches) can be found in Section 4.2. The choice between feature extraction and feature selection is data and application dependent. Feature selection retains the original meaning of the features and can provide computational and measurement savings. Feature extraction might provide a more informative feature subset but with possible increases to measurement cost and computation time. Note that feature extraction sometimes precedes feature selection, i.e. feature extraction can be used to generate features, some of which are subsequently discarded in a feature selection step (Jain, Duin & Mao 2000).

### **2.2.5 Classification**

A statistical pattern recognition approach to image classification has been used widely in remote sensing (Fassnacht et al. 2016; Ganivet & Bloomberg 2019; Vaz et al. 2018). In classification, statistical methods are used to partition the feature space into class regions in a training phase. Following this, objects are assigned class labels based on their position in feature space in a testing phase. There are two basic types of classification: 1) supervised, where objects are assigned to known classes, or 2) unsupervised, where the classes are unknown and must be learnt from the data. In the supervised approach, example labelled data is used to define a classifier that can then generalise to label unseen data. Popular supervised classifiers used in image classification are decision tree, random forest, support vector machine (SVM), maximum likelihood (ML) and *k*-nearest neighbours (kNN). These methods are described in more detail in Section 5.4.5. Unsupervised classification (also known as clustering) finds related groups of objects (i.e. classes) using object similarity measures. Affinity propagation (Frey & Dueck 2007), hierarchical clustering (Székely & Rizzo 2005) and *k*-means (Mather & Koch 2011) are examples of clustering algorithms. Image segmentation can also be understood as a form of unsupervised classification. Related pixels are grouped into image regions or objects, which, in effect, are the unknown classes that are learnt from the data. Clustering is a challenging problem, as both the form and number of the classes are often unknown (Jain, Duin & Mao 2000).

When the form of the class distributions in feature space is known, the problem is termed “parametric”. In this case, training becomes a parameter estimation problem and the class distribution parameters are fitted to the data. The ML classifier is an example of a common parametric classifier that assumes the classes are normally distributed (Mather & Koch 2011). A non-parametric classifier is one which makes no assumptions about the form of the class distributions or decision boundary (e.g. Parzen density estimation (Fukunaga & Hayes 1989)) but rather learns these from the data. Parametric and non-parametric methods can be further subdivided into generative and discriminative approaches. The generative approach models the class distributions in feature space. An object is then assigned to the class in which it has maximum probability, i.e. the decision boundary is determined by the modelled class distributions. ML and Parzen classifiers are examples of the generative approach. Discriminative classifiers fit the decision boundary directly to training data. The kNN (Bishop 2003) and SVM (Burges 1998) classifiers follow the discriminative approach.

### **2.2.6 Accuracy assessment**

Classification accuracy is evaluated by testing on labelled example data. The error rate, which is the percentage of available objects that were incorrectly classified, is the most common measure of classification performance (Jain, Duin & Mao 2000). It is, however, only a single figure representation and can hide important performance details (Campbell & Wynne 2011). The confusion matrix provides valuable and more detailed insight into classification performance. The  $(i, j)$  th element of the confusion matrix represents the number of actual class  $j$  objects that were classified as class  $i$  (Webb 2002). A number of useful accuracy metrics can be derived from the confusion matrix. The user’s accuracy for a class is the total number of objects correctly assigned to the class, as a percentage of the total number of objects assigned to the class. The producer’s accuracy for a class is the total number of objects correctly assigned to the class, as a percentage of the total number of objects belonging to the class. The Kappa statistic describes the accuracy of the classifier compared to pure chance (Campbell & Wynne 2011).

The performance of a supervised classifier on the training set is optimistically biased and not a reliable indication of operational performance (Jain, Duin & Mao 2000). This is especially true when the number of objects in the data is small relative to the feature dimensionality (Webb 2002). Classifiers should consequently be trained and tested on separate data to avoid biased error estimates. To obtain an accurate estimate of the classifier performance, the test set should be large enough to represent the underlying classes. Similarly, the training set is also required to represent the underlying classes in order to obtain a classifier with good generalisation ability. With limited

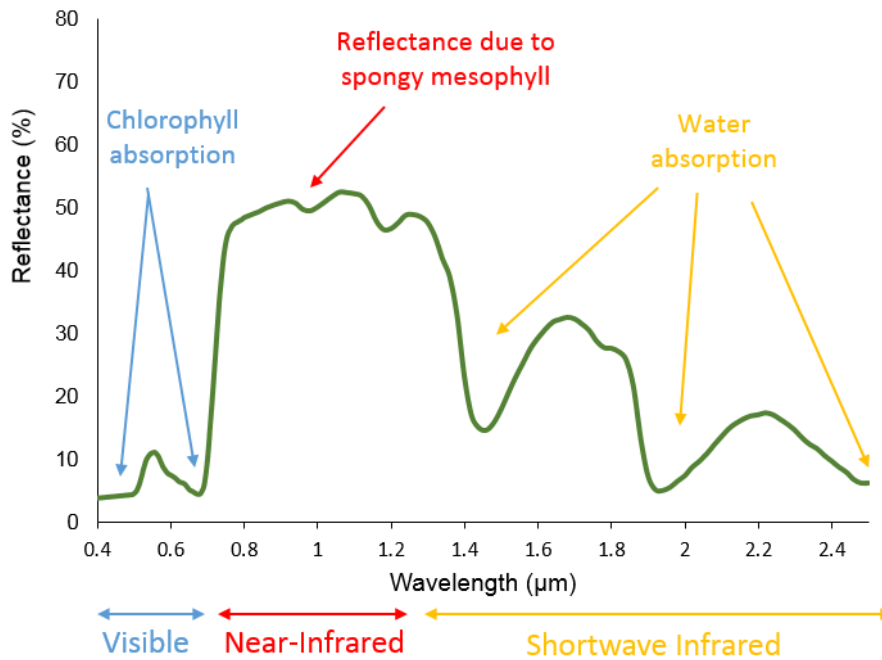
data this presents a trade-off in how the split into training and test sets is made (Jain, Duin & Mao 2000). In the holdout method, the data is simply split into two independent sets: one for training and one for testing. Leave-one-out and rotation versions of cross-validation improve on the holdout method by reducing error bias and simultaneously increasing the amount of data available for training (Bishop 2003). This is achieved by dividing the data into a number of subsets ( $P$ ), training on  $P-1$  subsets and testing on the remaining subset. This is repeated until testing has occurred on all  $P$  subsets. The bootstrap method uses resampling with replacement to generate many “new” data sets on which the classifier is trained and tested. This approach can improve on cross-validation error estimates at the price of increased computation (Webb 2002).

## **2.3 SPECIES-SPECIFIC VEGETATION MAPPING**

### **2.3.1 Spectral properties of vegetation**

The basic premise of spectral species recognition is that leaf chemistry and structure influence reflectance spectra in species specific ways (Clark & Roberts 2012). Vegetation reflectance spectra at the leaf scale are controlled by biochemistry and morphology (Asner 1998; Clark, Roberts & Clark 2005). In the visible portion of the spectrum, the reflectance of living vegetation is mainly affected by chlorophyll, the pigment that facilitates the absorption of sunlight for photosynthesis (Campbell & Wynne 2011). Chlorophyll absorbs the majority of incident blue and red light while reflecting some green light, giving living vegetation its characteristic green colour. Leaves are strongly reflective in the NIR spectrum due to scattering by internal plant structures, rather than reflection and absorption by pigments. Radiation passes through the leaf surface and is scattered by spongy mesophyll air-cell wall interfaces (Clark, Roberts & Clark 2005), resulting in both downward transmission and upwards reflectance. There is a sharp transition from absorption to reflectance around the red-NIR boundary, known as the “red-edge” (Campbell & Wynne 2011). In longer SWIR wavelengths, leaf reflectance diminishes and is dominated by absorption features resulting from water content (Asner 1998). TIR emissivity spectra are affected by chemicals in plant tissues that form external surfaces and cell walls (Ribeiro da Luz & Crowley 2007).





Source: HSU (2019)

Figure 2.6 Typical vegetation reflectance spectra

At the plant or canopy scale, vegetation reflectance is influenced by the physical arrangement of leaves and other structures such as flowers, stems, fruits and branches (Asner & Martin 2008; Clark, Roberts & Clark 2005). Small shadows cast amongst these canopy structures tend to decrease overall reflectance compared to leaf scale measurements. The relative strength of NIR reflectance is, however, increased due to repeated transmission and reflectance (scattering) of these wavelengths between and within canopy layers (Campbell & Wynne 2011). Canopy scale spectral effects can also contain species discriminating information, as they are influenced by plant morphology (Clark, Roberts & Clark 2005). While informative spectral features have been found throughout the VIS-SWIR region (Fassnacht et al. 2016), the NIR spectrum is recognised as being particularly important for species identification (Asner 1998; Campbell & Wynne 2011; Clark & Roberts 2012; Clark, Roberts & Clark 2005). Emissivity data in the TIR range has also been shown to be a promising source of species distinguishing information (Fassnacht et al. 2016). Typical vegetation reflectance is shown in Figure 2.6.

### 2.3.2 Remote sensing of species specific plant traits

Both hyperspectral and multi-spectral data has been used for remote sensing of vegetation species. Hyperspectral data is reported to provide superior classification performance due to its finer spectral resolution and ability to capture subtle spectral differences between species (Baldeck et al. 2015; Clark & Roberts 2012; Clark, Roberts & Clark 2005). Use of hyperspectral data is, however, limited by its cost and processing requirements (Adjorlolo et al. 2012; Ganivet &

Bloomberg 2019; Vaz et al. 2018). Multi-spectral imagery is increasingly being used due to its availability, coverage, high spatial resolution and relative cost-effectiveness. It has proved sufficient for species identification in many cases (Fassnacht et al. 2016; Ganivet & Bloomberg 2019; Vaz et al. 2018). Wavelengths in the red-NIR region are the most frequently used sources of species discriminating information (Clark, Roberts & Clark 2005; Fassnacht et al. 2016). TIR emissivity has also shown promise for species recognition (Neinavaz 2017; Ribeiro da Luz & Crowley 2007), but its use is very uncommon due to limited availability of data and challenges with calibration and atmospheric correction (Ribeiro da Luz & Crowley 2010).

Numerous studies have confirmed the usefulness of spectral measurements for identifying plant species (Fassnacht et al. 2016; Ganivet & Bloomberg 2019; Vaz et al. 2018). There are, however, a number of other factors, besides species, that contribute to plant spectral variation, making spectral species classification a challenging problem. Plant spectra change as a result of phenology (Clark & Roberts 2012), senescence (Einzmann et al. 2014) and stressors such as moisture shortage and disease (Campbell & Wynne 2011; Waser et al. 2014). Canopy scattering is highly anisotropic (Asner 1998), meaning that view-illumination geometry also impacts vegetation reflectance (Fassnacht et al. 2016; Strahler & Muller 1999). Finally, spectral mixing with surrounding soil and vegetation further contributes to reflectance variation (Clark, Roberts & Clark 2005; Fassnacht et al. 2016), especially in heterogeneous and sparsely vegetated habitats (Campbell & Wynne 2011; Eisfelder, Kuenzer & Dech 2012).

High spatial resolution imagery allows for the description of vegetation spatial structure with texture descriptors. Texture can capture aspects of plant morphology such as foliage size, foliage arrangement and branching patterns (Fassnacht et al. 2016). A number of species recognition studies have found texture descriptors to be useful additions to spectral information (Ghosh & Joshi 2014; Kollár, Vekerdy & Márkus 2013; Mustafa & Habeeb 2014; Neigh et al. 2018; Niphadkar et al. 2017). The encoding of complex scale-dependent patterns into meaningful texture descriptors is, however, a challenge (Mather & Koch 2011) that requires expert insight into the specific recognition task. Texture is sensitive to shadowing variations caused by changes in view-illumination geometry (Barbier et al. 2011) and is also prone to mixing between classes in heterogeneous habitats where a sliding window rather than object-based approach is followed (Blaschke 2010).

Approaches to the remote sensing of plant phenological traits have been used to assist with species identification. At a simple level, careful timing of snapshot (uni-temporal) imagery can help to capture distinguishing periods of seasonal plant growth (e.g. a target species may bloom while

others remain dormant) (Campbell & Wynne 2011; Niphadkar & Nagendra 2016). Compared to snapshot imagery, multi-temporal images provide more detailed descriptions of spectral variation relating to seasonal growth patterns. This form of data is seeing increasing use and can be useful for species identification where there are species discriminating phenological differences (Michez et al. 2016; Niphadkar et al. 2017; Tarantino et al. 2019). Some species exhibit phenology that depends on spatially varying environmental factors such as weather, geology and altitude (Fassnacht et al. 2016; Niphadkar & Nagendra 2016). These variations can be detrimental to classifier transferability.

LiDAR data has been used to characterise aspects of vegetation structure such as canopy architecture, branching pattern, foliage arrangement and foliage density (Coops et al. 2007; Ganivet & Bloomberg 2019; Korpela et al. 2010), which can be difficult to assess using spectral imagery. In a species identification context, LiDAR is often used complementary to passive optical multi-spectral or hyperspectral data (Alonzo, Bookhagen & Roberts 2014; Asner et al. 2008; Fassnacht et al. 2016). Promising results have been obtained using LiDAR waveform (intensity) data to describe features of tree structure such as crown and branch morphology for species recognition (Hovi et al. 2016; Korpela et al. 2010). Challenges with LiDAR data include multiple scattering within the canopy (Korpela et al. 2010), which creates uncertainty as to the true source of returns and intra-species variation of plant structure (Hovi et al. 2016), which can hamper classifier performance.

### **2.3.3 Features and classification methods**

Typical features generated for species identification include texture descriptors (Basu et al. 2015; Ghosh & Joshi 2014; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Mustafa & Habeeb 2014; Neigh et al. 2018; Niphadkar & Nagendra 2016), vegetation indices and band ratios (Basu et al. 2015; Boyden et al. 2007; De Castro et al. 2011; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Mehner et al. 2004; Mustafa & Habeeb 2014; Niphadkar & Nagendra 2016; Tarantino et al. 2019), spectral derivatives (Clark & Roberts 2012; Fassnacht et al. 2014), LiDAR metrics (Alonzo, Bookhagen & Roberts 2014; Coops et al. 2007; Hernández-Stefanoni et al. 2014; Hovi et al. 2016; Jones, Coops & Sharma 2011) and dimensionality reduction transformations such as PCA (Clark & Roberts 2012). Feature reduction is sometimes omitted or addressed using importance ranking (Basu et al. 2015; Clark & Roberts 2012; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Niphadkar et al. 2017) and manual band selection (Baldeck et al. 2015). The use of feature selection methods that consider the sub-optimal effects of feature redundancy, such as

stepwise selection (Clark, Roberts & Clark 2005; Ghosh & Joshi 2014), are uncommon in species identification.

Supervised classification is the most frequently used approach for identification of species based on remotely sensed features. Commonly used supervised classifiers are SVM (Baldeck et al. 2015; Ghosh & Joshi 2014; Tarantino et al. 2019), random forest (Clark & Roberts 2012; Nguyen et al. 2019) and normal distribution based ML and linear discriminant analysis (LDA) techniques (Boyden et al. 2007; De Castro et al. 2012; Clark, Roberts & Clark 2005; Mehner et al. 2004; Niphadkar et al. 2017; Tarantino et al. 2019). Algorithms implemented in the eCognition software package (Trimble 2018) such as the fuzzy and hierarchical approaches are also frequently used for vegetation identification (Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Niphadkar et al. 2017; Ouyang et al. 2011). Dissimilarity-based approaches such as spectral mixture analysis (SMA) and spectral angle mapping (SAM) are sometimes used with hyperspectral data (Asner et al. 2008; Clark, Roberts & Clark 2005). Unsupervised classification approaches have also been used (Neigh et al. 2018) but this is very unusual.

#### **2.3.4 Regional VHR vegetation mapping**

One of the main concerns in species recognition is that “there are few examples for tree species classifications over large geographic extents, and bridging the gap between current approaches and tree species inventories over large geographic extents is still one of the biggest challenges of this research field” (Fassnacht et al. 2016: 24). Regional scale vegetation mapping is most commonly done at coarse or moderate spatial resolutions (González-Roglich & Swenson 2016; Immitzer et al. 2018; Liu et al. 2017; Mathieu et al. 2018). Only two large area regional studies related to VHR vegetation species mapping were found in this literature survey. Neigh et al. (2018) used WorldView-1 VHR panchromatic imagery for mapping crop fields over a 4.1 million ha area in Ethiopia. A multi-level histogram thresholding (“Otsu multithresh”) method requiring no user input was employed for unsupervised classification. This simple and efficient algorithm was applied to per-pixel texture features and provided an overall accuracy of 94 %. Basu et al. (2015) mapped tree canopy cover over the Continental United States using multi-spectral VHR aerial imagery from the National Agriculture Imagery Program (NAIP) (National Agriculture Imagery Program 2019). An initial per-pixel classification on spectral and textural features was followed by automated segmentation of the classifier output using a conditional random field (CRF) method to capture contextual relationships between neighbouring pixels. An overall accuracy of 85 % was achieved.

Regional applications are limited by operational concerns such as cost, coverage and availability when selecting data sources. While higher spatial, spectral and temporal resolution data can potentially improve discriminative ability, cost and computational burden increases with an increase in the volume and complexity of the data (Ørka & Hauglin 2016). Regional studies generally need to compromise on data sophistication in order to remain operationally viable and cost-effective. Multi-spectral VHR satellite data represents a relatively cost-effective data source with global coverage. National programmes for the acquisition of aerial imagery such as NAIP in the United States (National Agriculture Imagery Program 2019) and NGI in South Africa (National Geo-spatial Information 2012) provide national coverage multi-spectral VHR aerial imagery free of charge. Multi-spectral VHR aerial or satellite imagery is currently the most practical option for budget constrained studies requiring regional coverage (Basu et al. 2015; Mirik & Ansley 2012; Neigh et al. 2018).

Due to the volume of data that needs to be processed, computational efficiency is also an important consideration for regional mapping applications. Complex image segmentation methods (such as MRS) are frequently used in small-area studies (Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Niphadkar et al. 2017; Ouyang et al. 2011). They are, however, computationally expensive and require user specification of parameters, and consequently do not scale well to projects requiring the automated processing of thousands of images (Basu et al. 2015; Neigh et al. 2018). Per-pixel classification provides adequate species identification accuracy in a number of cases (Boyden et al. 2007; De Castro et al. 2011; Mehner et al. 2004). Spatial and contextual information can still be incorporated into a per-pixel approach by automated post-processing and segmentation of the per-pixel classifier output (Baldeck et al. 2015; Basu et al. 2015). Computational efficiency of feature generation and classification is also relevant for regional methods (Basu et al. 2015; Neigh et al. 2018) but is often ignored in small-area studies (Boyden et al. 2007; Mehner et al. 2004; Mirik & Ansley 2012; Mustafa & Habeeb 2014; Ouyang et al. 2011).

Another crucial challenge of large-area methods is addressing sources of unwanted variation which hinder classifier generalisation ability (Fassnacht et al. 2016). Atmospheric and BRDF effects cause radiometric variations, which increase with the spatial and temporal extents of imagery due to changes in view-illumination geometry and atmospheric conditions (Fassnacht et al. 2016; Gehrke & Beshah 2016; Lelong et al. 2008). Mosaics of aerial images are particularly prone to radiometric variations due to wide image FOV and varying image capture times (Honkavaara et al. 2009; Korpela et al. 2014; Mather & Koch 2011). Radiometric corrections are sometimes not applied in small study areas (Boyden et al. 2007) or partially handled using conversion to top of

atmosphere radiance (Johansen et al. 2007; Mehner et al. 2004; Niphadkar et al. 2017; Tarantino et al. 2019), which is still subject to atmospheric and BRDF effects. Habitat, ecotype and phenological variations also increase over spatial and temporal scales and add uncertainty to the species recognition problem (Fassnacht et al. 2016), especially in diverse and heterogeneous habitats (Bradley 2014; Ganivet & Bloomberg 2019).

## 2.4 LITERATURE EVALUATION

Species-specific canopy cover maps are useful for restoration, forest management, invasive species management and conservation applications. There is a large body of research on the use of remote sensing for vegetation mapping in general and species recognition in particular. Remote sensing techniques have shown good potential for addressing these applications. The increasing availability of high spatial, spectral and temporal resolution data, in conjunction with the development of associated data preparation and classification algorithms, has seen a growth in species mapping applications. There nevertheless remain significant challenges to solving the species recognition problem. There is a known operational need for large-area vegetation maps, yet there are surprisingly few studies that address the cost, computation and classifier transferability challenges associated with large areas.

While LiDAR and hyperspectral data have been shown to improve species recognition accuracy, they are amongst the most expensive sources of data (Ørka & Hauglin 2016) and consequently are not viable data sources for this study. VHR satellite imagery is frequently used as a relatively cost-effective source of multi-spectral data, but it is still too expensive for most regional-scale mapping exercises. By using free of charge aerial imagery, the coverage and budget requirements of this study could be satisfied. This data has furthermore been used in other regional vegetation mapping work (Basu et al. 2015). Aerial imagery does, however, contain substantial radiometric variation that requires correction. Most vegetation mapping work has been applied to small study areas where radiometric variation due to atmospheric and BRDF effects is small enough to be ignored. The radiometric correction of large mosaics of VHR imagery remains an open and challenging problem.

Spectral and textural features are commonly used for vegetation mapping and are potential features for distinguishing spekboom. Rigorous feature reduction is a somewhat overlooked processing step in the species recognition literature. No studies comparing the effectiveness of established feature selection approaches for typical remote sensing data were found. The frequent use of importance ranking type feature selection ignores the sub-optimal effects of feature space

redundancy. With the majority of species recognition research being focused on small areas, issues of computational automation and efficiency are of limited importance and have consequently received little attention. While GEOBIA allows the exploitation of spatial and contextual information for providing discriminating features, it is not necessarily the best approach for regional studies, which require the automated processing of thousands of images. An alternative to traditional GEOBIA is an initial per-pixel classification, followed by post-processing of the classifier output to incorporate spatial context. Computationally efficient and automated implementations of this approach has been used successfully in demanding applications. The use of supervised classification for species identification is the norm. The best combination of features and classifier for distinguishing spekboom is not known and requires investigation.

This study aims to address the challenges of regional species recognition in the context of the spekboom canopy cover mapping application. It will investigate means of radiometric calibration of large mosaics of multi-spectral VHR imagery; assess the optimality and stability of common feature selection approaches in a remote sensing context; investigate the possibility of incorporating computational cost into a novel feature selection method; determine the most informative features and best supervised classifier for per-pixel spekboom identification; and investigate the benefit of incorporating spatial context in a post-processing step. Consideration will be given to computational efficiency in the selection and derivation of component methods and data.

## CHAPTER 3: RADIOMETRIC HOMOGENISATION OF AERIAL IMAGES BY CALIBRATING WITH SATELLITE DATA<sup>1</sup>

### 3.1 ABSTRACT

The use of very high resolution (VHR) aerial imagery for quantitative remote sensing has been limited by unwanted radiometric variation over temporal and spatial extents. In this paper we propose a simple yet effective technique for the radiometric homogenisation of the digital numbers of aerial images. The technique requires a collocated and concurrent, well-calibrated satellite image as surface reflectance reference to which the aerial images are calibrated. The bands of the reference satellite sensor should be spectrally similar to those of the aerial sensor. Using radiative transfer theory, we show that a spatially varying local linear model can be used to approximate the relationship between the surface reflectance of the reference image and the digital numbers of the aerial images. The model parameters for each satellite pixel location are estimated using least squares regression inside a small sliding window. The technique was applied to a set of aerial images captured over multiple days with an Intergraph digital mapping camera (DMC) system. A near-concurrent moderate resolution imaging spectroradiometer (MODIS) nadir bidirectional reflectance distribution function (BRDF) adjusted reflectance image was used as the reflectance reference dataset. The resulting DMC mosaic was compared to a near-concurrent satellite pour l'observation de la terre (SPOT) 5 reflectance image of a portion of the same area, omitting the blue channel from the DMC mosaic due to its absence in the SPOT 5 data. The mean absolute reflectance difference was found to be 3.43% and the mean coefficient of determination ( $R^2$ ) over the bands was 0.84. The technique allows the production of seamless mosaics corrected for coarse scale atmospheric and BRDF effects and does not require the manual acquisition (or provision) of ground reflectance references. The accuracy of corrections is limited by the resolution of the reference image, which is generally significantly coarser than VHR imagery. The method cannot correct for small scale BRDF or other variations not captured at the reference resolution. Nevertheless, results show a significant improvement in homogeneity and correlation with SPOT 5 reflectance.

---

<sup>1</sup> This chapter is published in the International Journal of Remote Sensing and consequently conforms to the prescribed structure of that journal



### 3.2 INTRODUCTION

Very high resolution (VHR) aerial and drone imagery is increasingly being used in remote sensing studies. The high spatial resolution of these images enables analyses on a finer spatial scale than most satellite-based platforms can provide and consequently allows the exploitation of information such as texture, object-based features and unmixed pixel spectra that is not available in lower resolution images (Chandelier & Martinoty 2009; Collings et al. 2011; Honkavaara et al. 2009; López et al. 2011; Markelin et al. 2012). Accurate geometric calibration techniques for producing orthorectified images are well established and form part of typical aerial imagery processing workflows (Chandelier & Martinoty 2009). Because aerial image mosaics are commonly produced for the purpose of visual interpretation, techniques such as dodging and lookup tables (LUTs) are often used to produce smooth and visually appealing results (López et al. 2011). This kind of adjustment can damage the spectral information content and is not suited to quantitative remote sensing. Also, spatial and temporal radiometric variations in aerial imagery limit the spatial and temporal extents to which quantitative remote sensing techniques can be successfully applied (Markelin et al. 2012). Ideally, quantitative analyses should be carried out on reflectance values. Atmospheric influences, bidirectional reflectance distribution function (BRDF) effects and sensor variations all contribute to radiometric variations in the imagery. To obtain surface reflectance, these radiometric variations must be removed, or reduced as far as possible. Transformation to surface reflectance is beneficial, as, unlike at-sensor quantities or surface radiance, surface reflectance is invariant to changes in atmospheric conditions and viewing geometry. This allows the surface reflectance data to be used in physical models, fused with other reflectance data and used in multi-temporal studies (Downey et al. 2010; Vicente-Serrano, Pérez-Cabello & Lasanta 2008).

There is some confusion and ambiguity around the use of reflectance terminology in the literature (Schaepman-Strub et al. 2006). In this paper, ‘surface reflectance’ is used to refer to the nadir BRDF-adjusted reflectance (NBAR) measurement provided by the moderate resolution imaging spectroradiometer (MODIS) MCD43 BRDF/albedo products. NBAR is the bidirectional reflectance factor normalised to local solar noon and viewed at nadir. It is worth noting that it is not possible or practical to correct for all the sources of radiometric variation in aerial imagery. The surface reflectance in most so-called ‘corrected’ or ‘calibrated’ images is only an approximation of the actual value.

A number of techniques for the correction of BRDF effects are available, including the popular kernel-based method (Roujean, Leroy & Deschamps 1992). Approaches based on radiometric

transfer modelling, such as atmospheric/topographic correction (ATCOR) (Richter 1997), moderate resolution atmospheric transmission (MODTRAN) (Berk et al. 1999) and second simulation of a satellite signal in the solar spectrum (6S) (Vermote et al. 1997) are used for atmospheric correction. While these atmospheric and BRDF correction methods are effective on single images (Markelin et al. 2012), blocks of multiple aerial images present unique challenges. The large field of view of aerial imaging cameras causes the viewing geometry to vary significantly within images (Lelong et al. 2008). Aerial campaigns are usually carried out over multiple days, resulting in significant variation in BRDF and atmospheric conditions. Each land cover also has its own unique BRDF and corrections should ideally model each of these covers separately (Collings et al. 2011; Honkavaara et al. 2009). Aerial campaigns can also consist of thousands of images, making it impractical to apply time-consuming atmospheric and BRDF correction models to every image (López et al. 2011). Even if it was practical, remnant radiometric variation due to the inexact nature of BRDF and atmospheric corrections will result in discontinuities, or seam lines, between adjacent images (Gehrke & Beshah 2016).

Approaches to calibrating mosaics of aerial imagery are receiving increasing attention (Chandelier & Martinoty 2009; Collings et al. 2011; Downey et al. 2010; Gehrke & Beshah 2016; López et al. 2011). Collings et al. (2011) introduced an empirical spatially varying model to perform combined atmospheric and BRDF correction. Land cover is assumed uniform in each image, resulting in a per-image BRDF parameterisation. The parameters of the model are solved by minimising a cost function that considers the internal accuracy of each image, similarity of overlapping image regions and smoothness (i.e. the lack of seam lines) of the mosaic. In a second stage the entire mosaic is calibrated to absolute reflectance using specially placed ground targets with known reflectance. In Chandelier and Martinoty (2009) a simple semi-empirical three parameter model of combined atmospheric and hot spot BRDF effects is fitted for each image by minimising the difference between ‘radiometric tie-points’, a selection of points in the overlapping image regions. It is a relative calibration method and no adjustment to absolute reflectance is made. López et al. (2011) apply theoretical atmospheric and semi-empirical BRDF kernel models using field-acquired spectral data. Gehrke (2010) uses standard atmospheric and BRDF methods, followed by a relative radiometric normalisation step using invariant points in overlapping regions to smooth the mosaic. A new radiometric normalisation method for heterogeneous image data, presented in Gehrke and Beshah (2016), improves on that of Gehrke (2010). The shortcomings of existing atmospheric and BRDF image adjustments are corrected using a spatially varying linear model. Model parameters are found at points in overlapping image regions and then interpolated into the remainder of the images.

A disadvantage of the aerial mosaic calibration techniques described above is their complexity and need for known ground references to achieve transformation to absolute surface reflectance. A number of the techniques also assume uniform BRDF characteristics within an image (Chandelier & Martinoty 2009; Collings et al. 2011; López et al. 2011). The options of placing targets of known reflectance to be captured as part of the mosaic or measuring the reflectance of suitably invariant sites on the ground are often not possible or practical. Many applications make use of archived imagery that had been captured prior to the commencement of the research and for which concurrent ground measurements are consequently not possible. Another approach is to make use of vicarious calibration involving knowledge of the spectral characteristics of specific ground sites, but this is recognised as being labour-intensive and costly (Chander, Meyer & Helder 2004; Gao et al. 2013; Liu et al. 2004).

In this paper, we propose a method of homogenising aerial imagery to coarse scale surface reflectance by calibrating to a concurrent and collocated satellite image that has been corrected for atmospheric and BRDF effects. Satellite programmes such as MODIS make such coarse-resolution surface reflectance products freely available. The proposed method avoids the need to perform atmospheric and BRDF corrections explicitly. It also does not require the placement of known reflectance targets or field spectral measurements, which can be impractical, error-prone and time-consuming in many instances. The technique was applied to a large set of aerial images captured with an Intergraph digital mapping camera (DMC). A near-concurrent MODIS MCD43A4 NBAR product was used as the reflectance reference dataset. The resulting DMC mosaic was compared to a near-concurrent Satellite Pour l’Observation de la Terre (SPOT) 5 reflectance image of the same area.

### **3.3 METHODS**

#### **3.3.1 Formulation of the local linear model**

The proposed method is based on approximating combined BRDF, atmospheric and sensor effects as a spatially varying linear relationship between surface reflectance and sensor measurement. Similar local linear relationships are used by Chandelier and Martinoty (2009), Collings et al. (2011) and Gehrke and Beshah (2016) for the radiometric correction of aerial image mosaics. In this section, we show that a spatially varying linear model is supported by radiative transfer theory.

Following the notation of López et al. (2011), the digital number (DN) measurement of an aerial sensor for each band can be expressed as:

$$\text{DN} = c_0 L_s + c_1 \quad \text{Equation 3.1}$$

where  $L_s$  is the radiance at the sensor and  $c_0$  and  $c_1$  are coefficients determined by the characteristics of the sensor. The proposed method requires the sensor radiance response to be linear; it should first be corrected for any non-linearity where this is not the case. The radiance at the sensor is expressed as:

$$L_s = \frac{\rho_s E_s \cos \theta}{\pi} \quad \text{Equation 3.2}$$

where  $\rho_s$  is the reflectance at the sensor,  $E_s$  is the irradiance at the sensor, and  $\theta$  is the solar zenith angle. The reflectance of a uniform Lambertian surface at the sensor is described by the radiative transfer equation (Vermote et al. 2006):

$$\rho_s = \rho_a + \frac{\rho_t}{1 - S\rho_t} \tau_u \tau_d \tau_g \quad \text{Equation 3.3}$$

where  $\rho_a$  is the intrinsic atmospheric reflectance,  $\rho_t$  is the surface reflectance and  $S$  is the atmospheric albedo.  $\tau_u$  and  $\tau_d$  are the atmospheric transmittances due to molecular and aerosol scattering between the surface and sensor and between the sun and the surface respectively, and  $\tau_g$  is the global atmospheric transmittance due to molecular absorption. It is common for aerial surveys to be conducted on clear days (Chandelier & Martinoty 2009). This was the case for the imagery used in our study. In clear sky conditions, the atmospheric albedo,  $S$ , is typically around 0.07 (Manabe et al. 1964). With a small value for  $S$  and a maximum value of one for  $\rho_t$ , the denominator in Equation 3.3 is approximately one and the reflectance at the sensor can be approximated as:

$$\rho_s \approx \rho_a + \rho_t \tau_u \tau_d \tau_g \quad \text{Equation 3.4}$$

Equations Equation 3.1, Equation 3.2 and Equation 3.4 express the relationship between the sensor measurement, atmospheric conditions and the surface reflectance. With the approximation of Equation 3.4, there is a linear relationship between surface reflectance and the sensor measurement. This linear relationship can be expressed as:

$$DN = M\rho_t + C \quad \text{Equation 3.5}$$

where

$$M = \frac{1}{\pi} c_0 \tau_u \tau_d \tau_g E_s \cos\theta \quad \text{Equation 3.6}$$

and

$$C = c_1 + \frac{1}{\pi} c_0 \rho_a E_s \cos\theta \quad \text{Equation 3.7}$$

The parameters  $M$  and  $C$  are spatially varying functions of the viewing geometry and atmospheric conditions. Implicit in any radiometric calibration technique is an approximation of these parameters so that the relationship can be inverted. A discussion of pertinent model assumptions is provided in the following section.

### 3.3.2 Parameter estimation

In our proposed method, we solve for  $M$  and  $C$  of the aerial sensor using a reference estimate for the surface reflectance parameter,  $\rho_t^{\text{ref}}$ , obtained from a well-calibrated satellite image. The reference surface reflectance image should have been captured at a similar time to the uncalibrated aerial image(s). The spatially varying property of the model allows the reduction of atmospheric effects that vary during aerial campaigns. The time difference between the reference and uncalibrated aerial images should, however, be small enough to avoid phenological or structural land cover changes. Such changes cannot be accounted for by the proposed method and will lead to errors (local to the area of change) in the calibrated result. In this study we use a 16-day composite reference image comprised of data from 25 January to 9 February 2010 and aerial images captured over multiple days between 22 January and 8 February 2010.

Least squares estimates of  $M$  and  $C$ , for the aerial sensor, are found for each pixel of the reference image inside a sliding window. Equation 3.5 can be rewritten in vector form, using the reference surface reflectance,  $\rho_t^{\text{ref}}$ , for the pixels inside a sliding window.

$$DN = M\rho_t^{\text{ref}} + CI \quad \text{Equation 3.8}$$

where  $\rho_t^{\text{ref}}$  and  $DN$  are column vectors of the  $N$  values inside the sliding window and  $I$  is a column vector of ones of length  $N$ .  $\rho_t^{\text{ref}}$  is obtained from the reference image and  $DN$  from the uncalibrated aerial image(s). The parameters can then be estimated as follows:

$$\begin{bmatrix} M \\ C \end{bmatrix} = [\rho_t^{\text{ref}} \quad I]^{-1}(DN) \quad \text{Equation 3.9}$$

In this form, the sliding window should consist of at least two pixels to solve for the two parameters. In order to accommodate the differing spatial resolutions,  $M$  and  $C$  must be found at the reference spatial resolution, resampled to the aerial spatial resolution, and then used to estimate surface reflectance at this resolution by inverting the relationship of Equation 3.5. The choice of the sliding window size involves a trade-off between the accuracy and effective spatial resolution of the estimated radiometric correction parameters,  $M$  and  $C$ . In essence, it is the typical parameter estimation trade-off between bias and variance or under- and overfitting (Webb 2002). Larger sliding windows will be less susceptible to overfitting on noisy data, while smaller sliding windows will provide higher spatial resolution correction parameters. The choice of sliding window size for the case study is discussed in Sections 3.3.5 and 3.4.4, and the effect of varying the sliding window size is investigated in Section 3.4.4.

The reference image will typically be at a substantially lower spatial resolution than the aerial imagery. BRDF corrected surface reflectance products, such as those produced from MODIS and multi-angle imaging spectroradiometer (MISR), have resolutions of the order of 500 m while aerial images usually have resolutions of 2 m or higher. This large resolution discrepancy affects the accuracy of the results. While the Equation 3.5 model applies to an aerial pixel,  $M$  and  $C$  are estimated at the reference resolution before resampling them to the aerial resolution. The reference resolution is only sufficient to capture gradual changes in BRDF and atmospheric conditions. Real BRDF can vary significantly over short distances where land cover is heterogeneous. This type of small scale BRDF variation is not captured at the resolution of the reference image and cannot be corrected for by the method. The theoretical formulation of Section 3.3.1 is intended to lend support to the use of the local linear model. It is acknowledged that the model involves a number of approximations and simplifications. These include ignoring adjacency effects and BRDF coupling with atmospheric effects (Vermote et al. 2006). We regard these approximations as necessary limitations of the method. Related methods make similar assumptions about BRDF homogeneity, often assuming a per-image BRDF model (Chandelier & Martinoty 2009; Collings et al. 2011; López et al. 2011) and also using simplified local linear models for approximating radiative transfer (Chandelier & Martinoty 2009; Collings et al. 2011; Gehrke & Beshah 2016).

### 3.3.3 Incorporation of viewing geometry and relative spectral response effects

The formulation of the local linear model in section 3.3.1 does not consider the effect of viewing geometry and relative spectral response (RSR) differences between the aerial and reference images. In this section, we show that these effects can also be modelled as locally linear relations, which can then be combined with Equation 3.5 to form a single linear relation between  $\rho_t^{\text{ref}}$  of the reference image and DN of the uncalibrated aerial image(s). These combined effects are thus still approximated by a spatially varying linear model with parameters as estimated with Equation 3.9.

The surface reflectance in Equation 3.4 represents the simplified case of a Lambertian reflector (i.e. it reflects equally in all directions), but in practice is subject to BRDF effects and so also varies with the viewing geometry (i.e. spatially) as in Equation 3.10.

$$\rho_t = f(\xi) \quad \text{Equation 3.10}$$

where  $f$  is the BRDF and  $\xi = (\theta, \vartheta, \phi)$  is the viewing geometry,  $\theta$  is the solar zenith angle,  $\vartheta$  is the view zenith angle and  $\phi$  the relative azimuth. To incorporate BRDF into the model, it is necessary to account for the viewing geometry differences between the reference and aerial images. It is not valid to simply replace the unknown aerial surface reflectance with the reference surface reflectance in Equation 3.5 when they occur at different viewing geometries. To describe the relationship between the two reflectances at a particular location, we express their ratio in Equation 3.11.

$$\frac{\rho_t^{\text{aerial}}}{\rho_t^{\text{ref}}} = \frac{f(\xi^{\text{aerial}})}{f(\xi^{\text{ref}})} \quad \text{Equation 3.11}$$

Here  $\rho_t^{\text{aerial}}$  and  $\rho_t^{\text{ref}}$  are the aerial and reference surface reflectances respectively, and  $\xi^{\text{aerial}}$  and  $\xi^{\text{ref}}$  are the aerial and reference viewing geometries respectively. The aerial surface reflectance can then be expressed as a multiple of the reference surface reflectance.

$$\rho_t^{\text{aerial}} = \frac{f(\xi^{\text{aerial}})}{f(\xi^{\text{ref}})} \rho_t^{\text{ref}} = F \rho_t^{\text{ref}} \quad \text{Equation 3.12}$$

$F$  is a spatially varying function of the aerial and reference viewing geometries. As this relation is locally linear, it can be incorporated into the model of Equation 3.5, maintaining the linearity and spatially varying properties. Coupling between atmospheric and BRDF effects (Vermote et

al. 2006) are not included in the model. The parameter  $F$  of Equation 3.12 is effectively estimated at the reference image resolution. The effects of viewing angle variations at the aerial image resolution are approximated by interpolating the coarse scale homogenisation parameters to the aerial resolution.

The model presentation in Section 3.3.1 ignored the effect of the spectral responses of the reference and uncalibrated sensors. The relation between surface reflectance and sensor measurement in Equation 3.5 becomes non-linear when including the spectral response effect. The surface reflectance in Equation 3.5 is a band averaged quantity, as represented by Equation 3.13.

$$\rho_t = \frac{\int \rho_t(\lambda)R(\lambda)d\lambda}{\int R(\lambda)d\lambda} \quad \text{Equation 3.13}$$

where  $\rho_t(\lambda)$  is the spectral surface reflectance and  $R(\lambda)$  is the sensor RSR for a particular band. Without knowledge of the surface reflectance spectra, it is not possible to completely calibrate for this effect. However, for real world surface reflectances it can often be shown that the relationship between the band averaged values for different sensors is approximately linear (Gao et al. 2013; Jiang & Li 2009). This means the relationship between surface reflectance and sensor measurement remains approximately linear even when the sensor spectral response is considered. We therefore approximate the effect of sensor spectral responses as a locally linear relationship that is incorporated into the linear model of Equation 3.5. This approximation is supported by simulations for the case study sensors in Sections 3.3.6 and 3.4.1.

### 3.3.4 Surface reflectance homogenisation

The homogenisation procedure follows these steps:

1. Resample uncalibrated aerial images to the reference image resolution and grid.
2. With the output from step (1), calculate sliding window estimates of  $M$  and  $C$  for each pixel of each band of the reference image using Equation 3.9. This forms two multi-band rasters  $\mathbf{M}$  and  $\mathbf{C}$  at the reference grid and resolution.
3. Resample  $\mathbf{M}$  and  $\mathbf{C}$  rasters to the aerial image resolution and grid.
4. Calculate estimated surface reflectance for each pixel of each band of the uncalibrated aerial image, using Equation 3.5.



The choice of resampling algorithms in steps 1 and 3 of the procedure are important, especially when there is a large difference in the spatial resolution of the aerial and reference images. Optical imaging systems are linear and thus subject to the superposition principle, which manifests as spectral mixing (Akhmanov & Nikitin 1997). Averaging the uncalibrated image over each reference pixel area is recommended when downsampling in step 1. This will approximate the spectral mixing that occurs in the larger reference image pixels.

It is necessary to produce smooth **M** and **C** rasters in step 3 to approximate slowly varying atmospheric and BRDF effects and to avoid discontinuities in the final image(s). Of the standard interpolation algorithms, cubic spline interpolation, with its constraints of continuity of the first and second derivatives, best satisfies this requirement (Hou & Andrews 1978). The Geospatial Data Abstraction Library (GDAL) (GDAL Development Team 2014) was used for implementing the resampling.

Since adjacent aerial images are calibrated to match the same reference image, overlapping image areas are similar. Blocks of aerial surface reflectance images generated with the procedure outlined above can generally be mosaicked without the need for additional colour balancing or normalisation procedures to reduce seam lines. Due to the disparity between reference and aerial image resolutions, there may be situations (such as uneven, shadowed terrain) where fine-scale differences between adjacent images produce slight seam lines. In these situations, one could use a feathering procedure to blend overlapping areas. In the case study however, we did not find this to be necessary.

### **3.3.5 Study site, data collection and preparation**

The surface reflectance homogenisation method proposed in this paper was tested in a 96 km × 107 km area (Figure 3.1) in the Little Karoo in South Africa. This particular study site was chosen as the calibration work forms part of a larger vegetation mapping study being done in the area.

The Chief Directorate: National Geo-spatial Information (NGI), a component of the South African Department of Rural Development and Land Reform, acquires and supplies national coverage aerial imagery. VHR multi-spectral 0.5 m pixel<sup>-1</sup> imagery of the study area was obtained from the NGI archive. The imagery was captured with a multi-spectral Intergraph DMC with red, green, blue and near-infrared (NIR) channels.

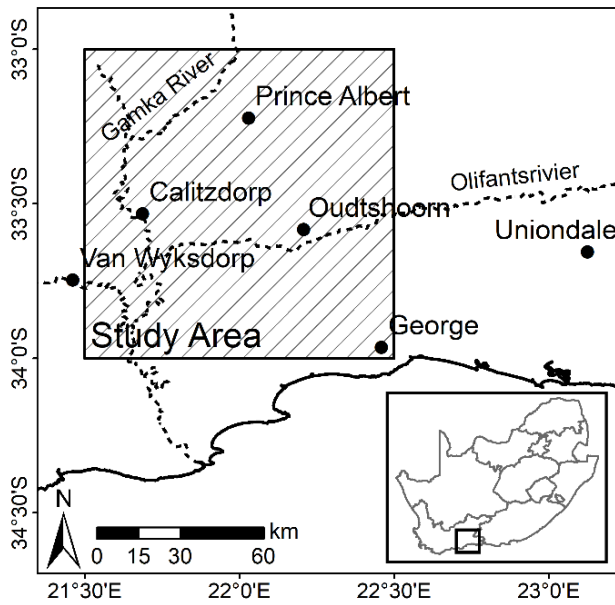


Figure 3.1 Study area orientation map

The RSRs of the DMC and MODIS sensors are shown in Figure 3.2. The peaks overlap well between the sensors in the red, green and blue bands, while the DMC NIR band is wider than that of MODIS.

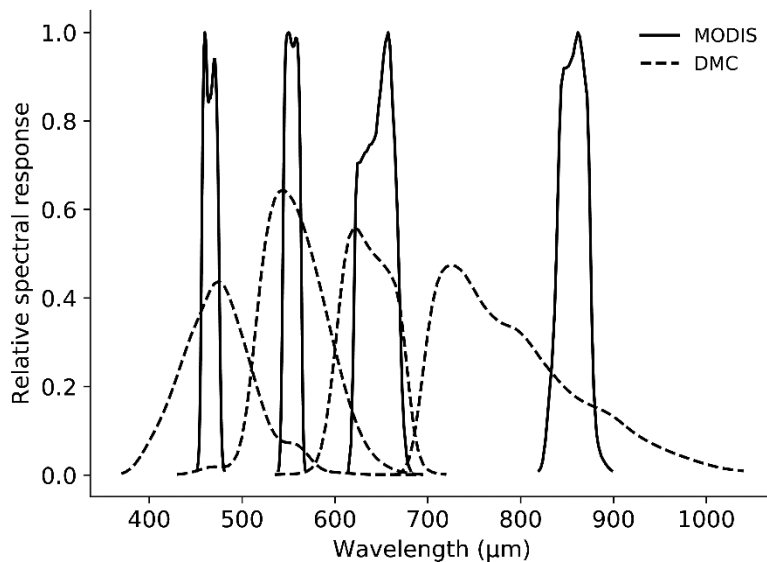


Figure 3.2 DMC's and MODIS's RSR

The study site is covered by 2228 images captured during four separate aerial campaigns on multiple days from 22 January 2010 to 8 February 2010. The inclusion of a large set of imagery acquired over multiple days allows an investigation into the robustness of the method to temporal variation (and the consequent BRDF and atmospheric variations). There is an average sidelap of 25% between images in adjacent flight lines and an average forward overlap of 60% between consecutive images in the same flight line (National Geo-spatial Information 2010). The images

were captured close to nadir, with a maximum tilt of  $5.0^\circ$ , at a height of 5000 m above ground. The DMC has a  $69.3^\circ$  cross track and  $42.0^\circ$  along track field of view (FOV). NGI campaigns are conducted on clear days such that images are free of clouds, cloud-shadows, smoke and excessive haze. Flight times are chosen to achieve a solar altitude of at least  $30.0^\circ$  in order to minimise shadowing. The mean aerosol optical depth (AOD) over the campaign days was 0.02 (the AOD provided by the MODIS MOD04 product was used for this calculation).

The raw aerial imagery was corrected for lens distortion, band spatial alignment, sensor non-linearity and dark current effects using the Intergraph Z/I Post-Processing Software (PPS). The PPS corrected imagery has a linear radiance response (as required by Equation 3.1), with zero offset (i.e.  $c_1 = 0$ ). This corrected imagery was orthorectified using existing aerotriangulation data supplied by NGI and a 5 m resolution digital elevation model (DEM) (Van Niekerk 2016).

A MODIS MCD43A4 NBAR product for the period from 25 January 2010 to 9 February 2010 was selected as a reference for the homogenisation. This product has a 500 m resolution and contains NBAR data composited from the best values over a 16-day period. The MODIS NBAR data has been processed with atmospheric and BRDF correction procedures (Strahler & Muller 1999) and is recognised as a reliable reference source for cross-calibration (Gao et al. 2013; Jiang & Li 2009; Li, Yang & Wang 2012; Liu et al. 2004). The NBAR data accuracy has been verified in a number of studies and is accurate to ‘well less than 5% albedo at the majority of the validation sites’ (MODIS Land Team 2014). MODIS was also selected as it has similar spectral bands to the Intergraph DMC. Bands 4, 1, 3 and 2 from the MODIS sensor were used to correspond to the red, green, blue and NIR bands from the DMC sensor respectively. While Landsat surface reflectance (Schmidt et al. 2012) could be a useful source of reference data due to its relatively high spatial resolution of 30 m, no cloud-free Landsat imagery concurrent (or near-concurrent) to the aerial imagery was available for testing in the case study.

The PPS processed imagery has zero offset, so the parameter  $c_1$  from Equation 3.7 was zero and the atmospheric reflectance,  $\rho_a$ , was small as the surveys were conducted on clear days, meaning that  $C$  was small. For the purposes of the case study,  $C$  was ignored and only the gain,  $M$ , was estimated. With only one parameter to estimate, a sliding window of one pixel was used to achieve the best possible spatial resolution in the  $\mathbf{M}$  raster. The gain-only simplification is given support by the results of Collings et al. (2011) who found the gain term in their linear model to carry the majority of the corrective effect. Nevertheless, it must be acknowledged that this is an initial approach and subsequent studies should be conducted to investigate the effects of including  $C$ .

The blue channel in particular could benefit from the inclusion of the offset ( $C$ ) term due to its haze sensitivity.

### 3.3.6 Linearity of band averaged values

In formulating the method in Section 3.3.1, it was assumed that the effect of sensor RSR on measured surface reflectance is locally linear and can be incorporated in the model of Equation 3.5. To investigate the validity of this assumption, MODIS and DMC band averaged values were simulated for typical surface reflectance spectra and statistically compared. Twenty surface reflectance spectra were selected from the ‘soil’, ‘vegetation’, ‘water’ and ‘man-made’ classes in the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) spectral library (Baldrige et al. 2009) to represent commonly encountered land covers. Band averaged values were then simulated for these representative spectra using Equation 3.13, with the MODIS and DMC RSRs as shown in Figure 3.2.

### 3.3.7 Accuracy assessment

Given that the DMC imagery was acquired in 2010, it was not possible to assess the accuracy of the reflectance retrieval method using ground-based spectral measures. Alternative methods for evaluating the results were consequently needed. First, the DMC DN and calibrated surface reflectance images were stitched into mosaics and the mosaics were visually compared to determine if discontinuities between adjacent images were reduced and to what extent the radiometric variations were corrected. Second, the DMC homogenised mosaic was resampled to the MODIS grid and resolution, and statistically compared to the MODIS reference image. Last, we quantitatively compared the DMC homogenised mosaic to a SPOT 5 scene, and the SPOT 5 scene to the MODIS reference image.

The 10 m resolution SPOT 5 level 1A image, acquired on 21 January 2010, covers portions of all four aerial campaigns as shown in Figure 3.3. The image was orthorectified using a 5 m resolution DEM (Van Niekerk 2016). The SPOT 5 DN image was converted to surface reflectance using the ATCOR 3 method (Richter 1997). While the ATCOR 3 correction did not include explicit BRDF correction, SPOT 5 is subject to substantially lower BRDF effects than the aerial imagery due to its narrow FOV (Mather & Koch 2011). Since the SPOT 5 sensor does not have a blue band, it was omitted from this comparison. The SPOT 5 resolution of 10 m allows the homogenised surface reflectance result to be checked at a resolution significantly closer to the aerial resolution than the reference MODIS resolution. This provides a useful check of the effect of approximating BRDF and atmospheric variations at the coarse scale of the reference image. While the MODIS

comparison checks the DMC surface reflectance against the reference it was fitted to, the SPOT 5 comparison uses an independent and ‘unseen’ source.

To establish the relative accuracy of the corrected SPOT 5 scene, it was downsampled (by averaging) to the MODIS resolution and grid and statistically compared to the MODIS reference image using Equation 3.14.

$$E(x, y) = |I^{\text{CMP}}(x, y) - I^{\text{SPOT}}(x, y)| \quad \text{Equation 3.14}$$

where  $I^{\text{CMP}}$  is the MODIS image,  $I^{\text{SPOT}}$  is the SPOT 5 image,  $(x, y)$  are the pixel coordinates and  $E$  is the difference image. Mean absolute difference (MAD) and root mean square (RMS) statistics were found to establish a benchmark against which similar statistics for DMC homogenised mosaic could be compared.

The DMC homogenised mosaic was then downsampled (by averaging) to the SPOT 5 resolution and grid, and statistically compared to the SPOT 5 surface reflectance image using Equation 3.14 with  $I^{\text{CMP}}$  as the DMC mosaic in this instance.

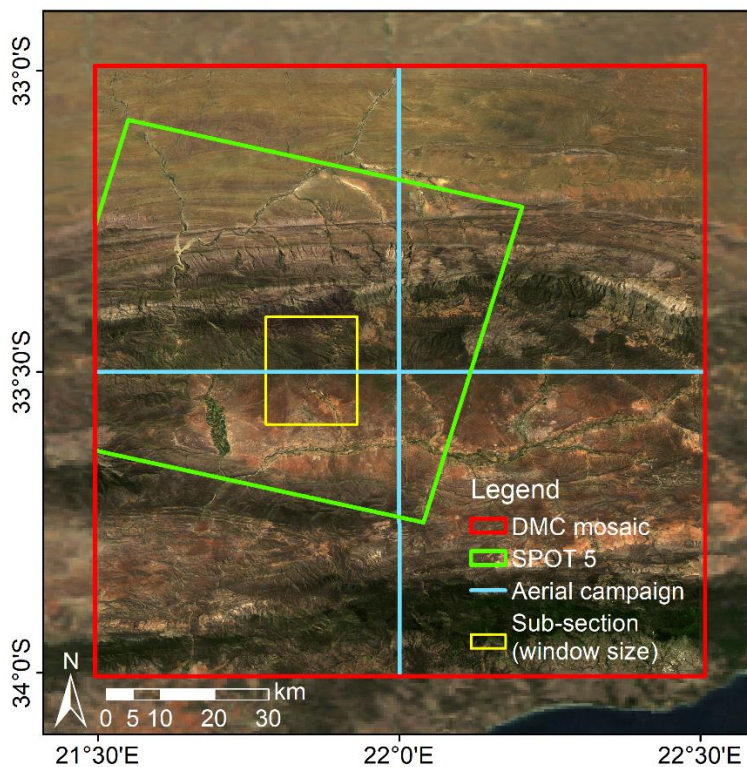


Figure 3.3 SPOT 5 scene and mosaic extents

The image resulting from the differencing process was used to identify spatial patterns in the discrepancies between the corrected SPOT 5 image and DMC mosaic. MAD and RMS statistics were found for the image to allow comparison with the results of similar studies. In addition to

the statistical evaluation, individual spectra from homogenous surfaces in the SPOT 5 surface reflectance image, and DMC homogenised and DN mosaics were compared.

The impact of the sliding window size was investigated by repeating the comparison between the SPOT 5 image and homogenised DMC mosaic with increasing sliding window sizes. A subsection of the full study area, as shown in Figure 3.3, was used for the sliding window size investigation, to expedite computation times.

### 3.4 RESULTS AND DISCUSSION

#### 3.4.1 Band averaged relationships

The simulated band averaged reflectance relationship for typical surface reflectances between the two sensors is shown in Figure 3.4 with coefficient of determination ( $R^2$ ) values.

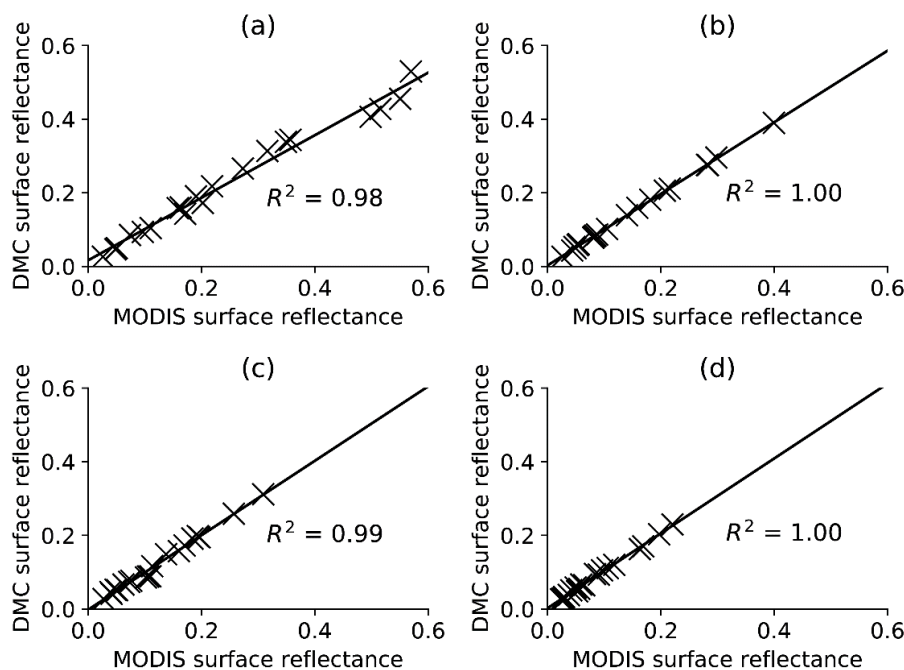


Figure 3.4 DMC versus MODIS simulated band averaged relationship for typical surface reflectances ((a) NIR, (b) red, (c) green and (d) blue bands)

The correlation between the DMC and MODIS simulated band averaged values (Figure 3.4) is surprisingly strong and supports the incorporation of the band averaging effect into the linear reflectance model of Equation 3.5. Similar linear relationships between different sensors for real world surface reflectances are reported in Gao et al. (2013) and Jiang and Li (2009). As the proposed method only requires the relationship to be locally linear, the variety of land covers simulated here is unlikely to be present inside the sliding window used to estimate the model parameters. For a small sliding window, the correlation of the band averaged values will

consequently be stronger than what is shown in Figure 3.4. The NIR channel has the lowest  $R^2$ , likely due to the relatively larger dissimilarity between MODIS and DMC RSRs, as evidenced in Figure 3.2. One can consequently expect higher surface reflectance errors in this channel compared to the others.

### 3.4.2 Mosaicking

Figure 3.5 shows a RGB (red, green and blue) mosaic of DMC DN images (bordered in red), against a background of the MODIS reference image. Seam lines between adjacent DMC images and radiometric variations over the set of images are clearly visible.

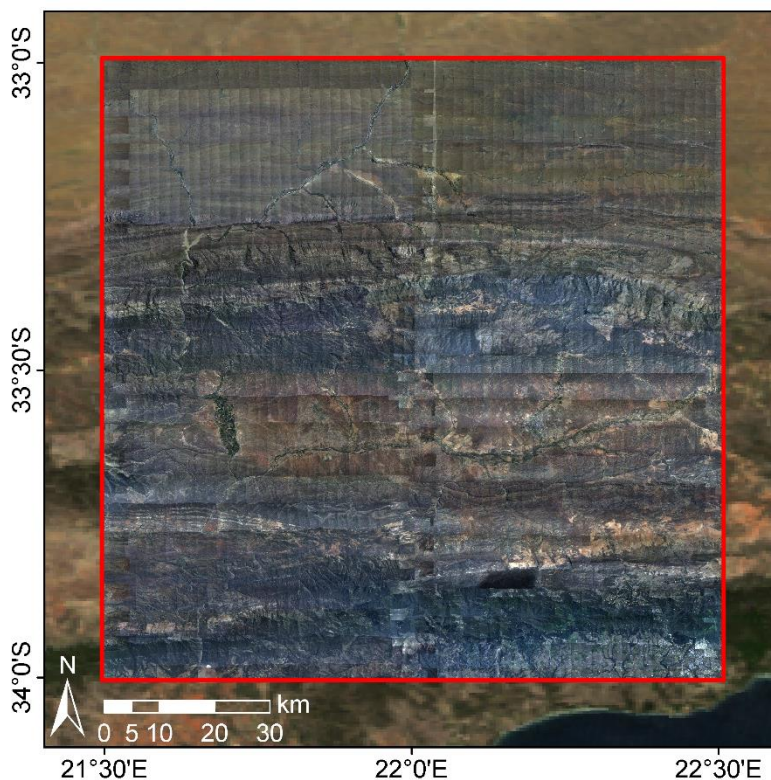


Figure 3.5 Uncalibrated mosaic on MODIS reference image background

Each DMC image was converted to surface reflectance using the proposed procedure. A RGB mosaic of the corrected images is shown in Figure 3.6, bordered in red, against a background of the MODIS reference image. No seam lines or radiometric anomalies (e.g. hot spots) are apparent at this scale, and the corrected images match the reflectance of the MODIS reference image.

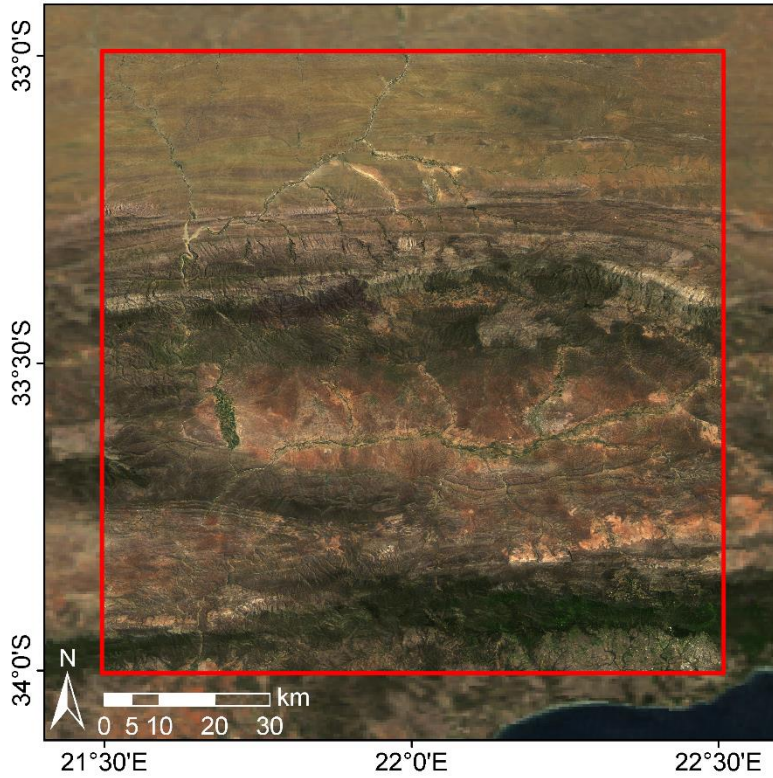


Figure 3.6 Homogenised mosaic on MODIS reference image background

Figure 3.7 (a) shows a close-up section of the DMC DN mosaic where a hot spot (i.e. a BRDF effect where sunlight is strongly reflected back into the camera) and seam lines between adjacent images are visible. Figure 3.7 (b) demonstrates the successful removal of the hot spot and seam lines after correction with the surface reflectance extraction method.

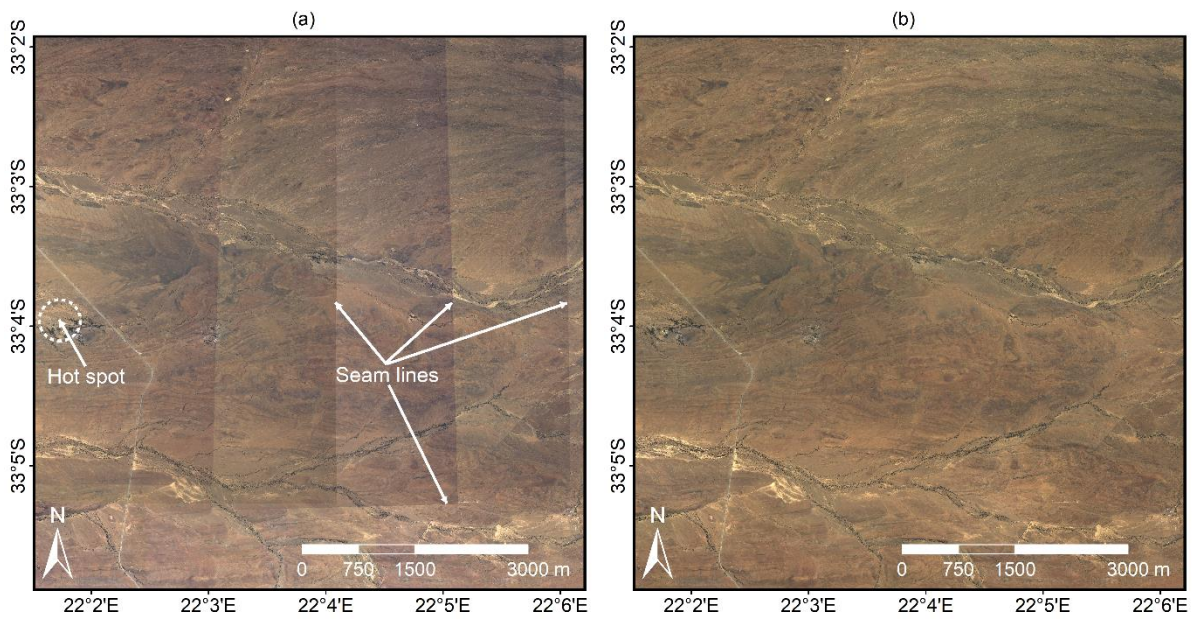


Figure 3.7 Reduction of hot spot and seam lines, with (a) showing raw DN images including hot spot and seam lines and (b) the corrected surface reflectance image



### 3.4.3 MODIS statistical comparison

Figure 3.8 shows scatter plots of the DMC DN and MODIS surface reflectance values with  $R^2$  coefficients indicating correlation strength. Figure 3.9 shows similar scatter plots for the DMC and MODIS surface reflectance values. Differences in the MODIS and DMC surface reflectance values at MODIS resolution are in part due to the use of the cubic spline interpolation to upsample the **M** and **C** rasters from the MODIS to DMC resolution. The spline interpolation is non-invertible (i.e. downsampling the upsampled rasters does not produce the original **M** and **C** rasters, but successively smooths the data at each application). As indicated by Figure 3.8 and Figure 3.9, the correlation of the DMC and MODIS values is significantly improved when using the homogenised DMC surface reflectance rather than DN values. This improvement in correlation is not unexpected, as Figure 3.9 is effectively comparing calibrated values to the values that were used for calibration. Nevertheless, this comparison serves as a general check on the validity of the method and as an indication of the effect of spline interpolation between the disparate MODIS and DMC resolutions.

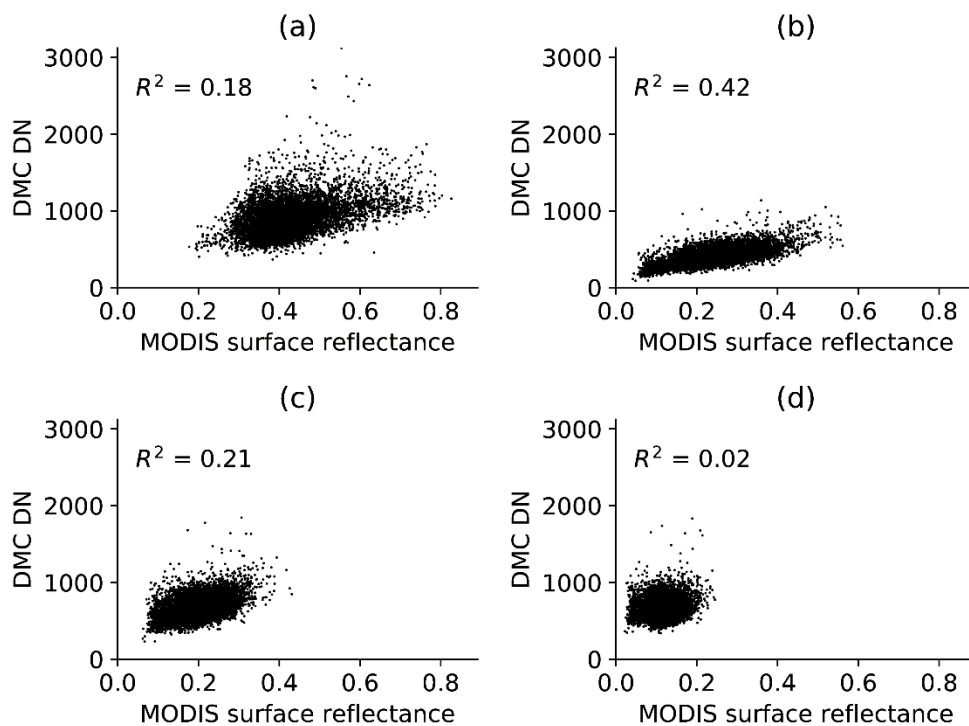


Figure 3.8 DMC DN values and MODIS surface reflectance correlation for the (a) NIR, (b) red, (c) green and (d) blue bands

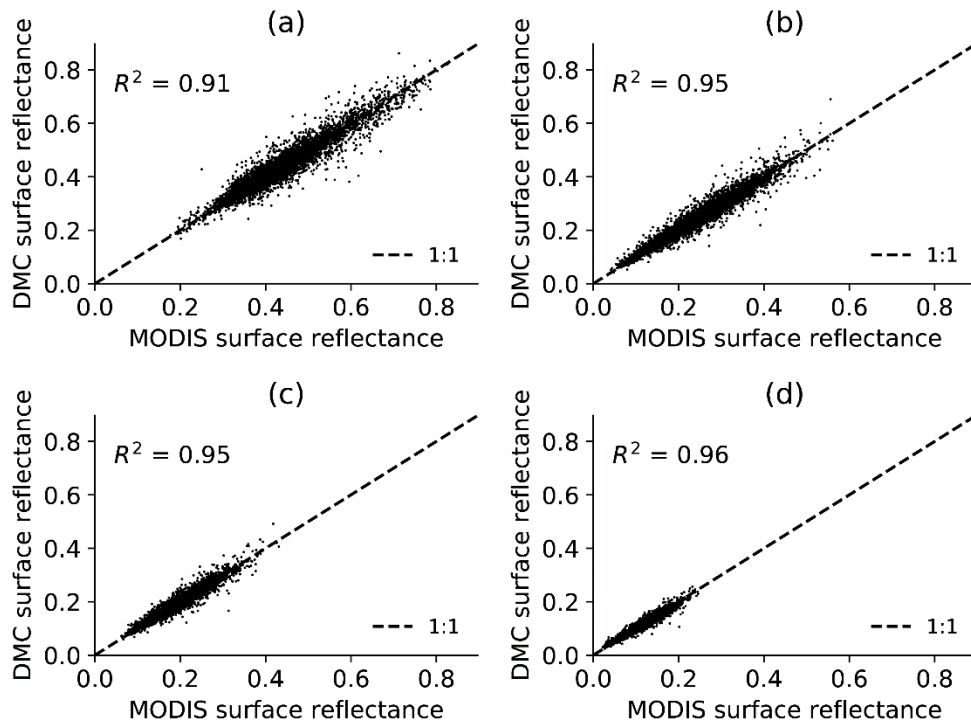


Figure 3.9 DMC homogenised mosaic and MODIS surface reflectance correlation for the (a) NIR, (b) red, (c) green and (d) blue bands

MAD, RMS and coefficient of determination statistics are given for the DMC and MODIS surface reflectance values in Table 3.1. Reflectance differences are the greatest in the NIR band, most likely due to the dissimilar MODIS and DMC RSRs in this band (Figure 3.2). This demonstrates the importance of using a reference image from a sensor with similar RSRs to those of the target imagery.

Table 3.1 Statistical comparison between MODIS and DMC surface reflectance images

Band	MAD (%)	RMS (%)	$R^2$
Near-infrared	1.70	2.50	0.91
Red	1.18	1.75	0.95
Green	0.79	1.16	0.96
Blue	0.48	0.69	0.96
All	1.04	1.67	0.94

### 3.4.4 SPOT 5 statistical comparison

An indication of magnitude of discrepancies in the SPOT 5 image is shown in Table 3.2, which shows the statistics for the difference between the SPOT 5 and MODIS images. The relatively low mean overall absolute reflectance difference of 3.35% between the SPOT 5 and MODIS values is consequently a good indication that the SPOT 5 surface reflectance extraction is effective. Note that the completeness of the SPOT 5 comparisons are limited by the missing blue band and partial coverage of the study area, as shown in Figure 3.3.

Table 3.2 Statistical comparison between SPOT 5 and MODIS surface reflectance images

Band	MAD (%)	RMS (%)	$R^2$
Near-infrared	4.81	5.79	0.86
Red	2.83	3.55	0.88
Green	2.40	2.91	0.87
All	3.35	4.27	0.87

Statistics for the reflectance difference between the corrected SPOT 5 image and the DMC homogenised mosaic are shown in Table 3.3. Not all of the reflectance differences can be attributed to errors in the homogenised DMC surface reflectances. Spatial misalignment of pixels due to orthorectification differences and BRDF errors in the SPOT 5 surface reflectances also contribute to the recorded differences. Despite this uncertainty due to the contribution of other error sources, these reflectance differences compare well to figures reported by other aerial image correction methods. Collings et al. (2011) achieved RMS reflectance errors of 1.37%–12.30% measured on placed targets of known reflectance for their aerial mosaic correction technique, and in the aerotriangulation approach of López et al. (2011), mean absolute reflectance differences of approximately 3.30%–5.00% were obtained on field measured test sites distributed throughout their study area. Similarly to the MODIS comparison, the largest reflectance differences occur in the NIR band. Again, this is likely due to dissimilarities in the RSRs of MODIS, DMC and SPOT 5 sensor NIR bands (see Figure 3.2 and Figure 3.10).

Table 3.3 Statistical comparison between SPOT 5 and DMC surface reflectance images

Band	MAD (%)	RMS (%)	$R^2$
Near-infrared	4.00	5.66	0.80
Red	3.11	4.27	0.86
Green	3.19	3.86	0.85
All	3.43	4.66	0.84

Scatter plots of DMC DN and SPOT 5 surface reflectance values are shown in Figure 3.11, and DMC surface reflectance and SPOT 5 surface reflectance values are shown in Figure 3.12, with  $R^2$  coefficients indicating correlation strength. The  $R^2$  values show a moderately strong correlation of 0.84 averaged over the bands. The lack of perfect correlation is due to, amongst other error sources, the effect of small scale land cover heterogeneity and differing SPOT 5 and DMC RSRs. The SPOT 5-MODIS scatter plots in Figure 3.13 show similar deviations from the one-to-one line as those for SPOT 5-DMC in Figure 3.12, especially for NIR. This suggests that the SPOT 5 image must be contributing, at least in part, to the SPOT 5-DMC differences. The effect of dissimilar SPOT 5, DMC and MODIS NIR RSRs is again evidenced in Figure 3.12 and Figure 3.13 as the marked deviations of the NIR scatters from the identity lines. Despite these disparities,

the homogenisation of DMC surface reflectance provides a substantial improvement in correlation between the DMC and SPOT 5 values.

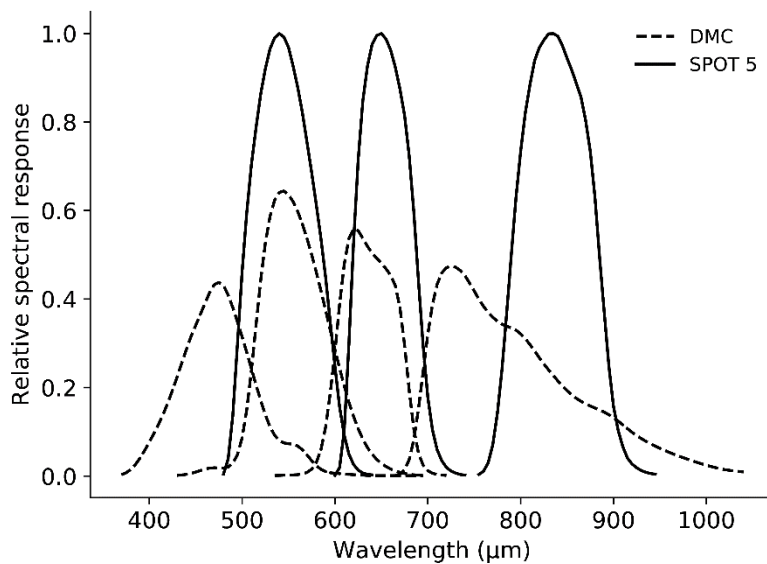


Figure 3.10 DMC and SPOT 5 RSRs

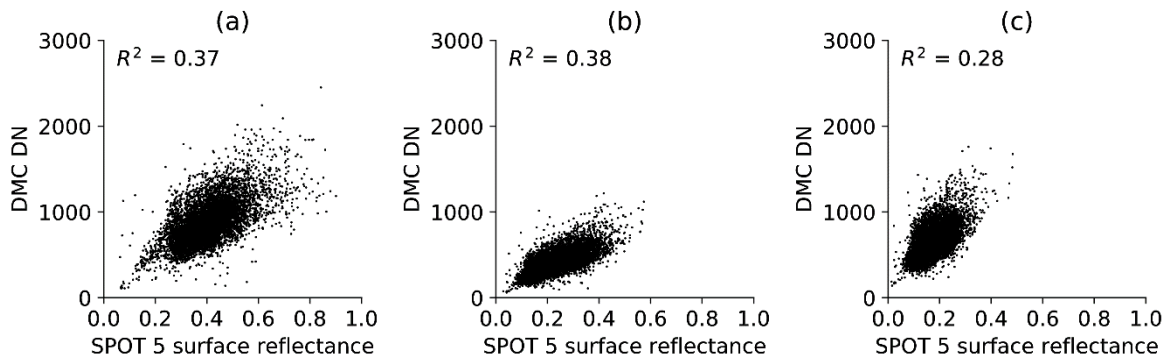


Figure 3.11 DMC DN mosaic and SPOT 5 surface reflectance correlation for the (a) NIR, (b) red and (c) green bands

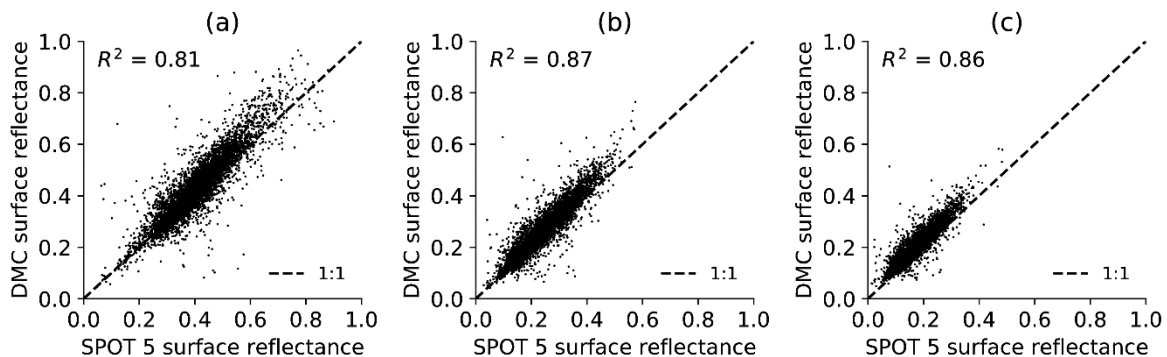


Figure 3.12 DMC homogenised mosaic and SPOT 5 surface reflectance correlation for the (a) NIR, (b) red and (c) green bands

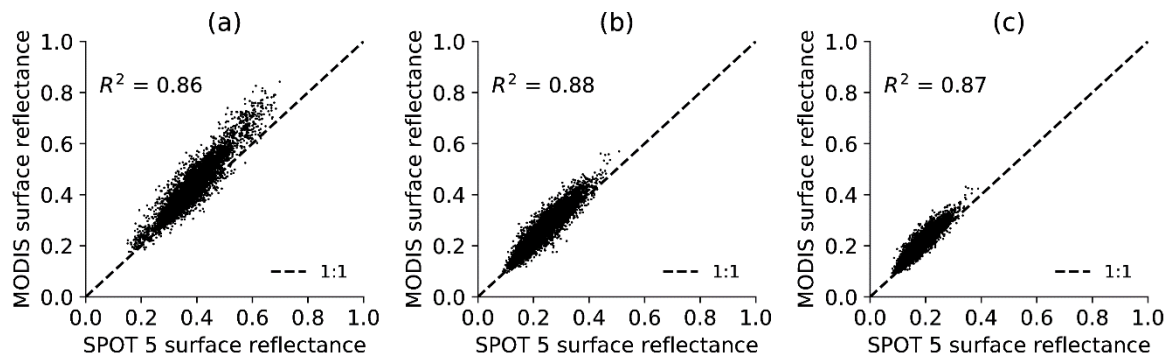


Figure 3.13 MODIS and SPOT 5 surface reflectance correlation for the (a) NIR, (b) red and (c) green bands

The effect of increasing the sliding window size on the reflectance difference between the SPOT 5 and DMC homogenised subsections is shown in Figure 3.14 (points are labelled with their corresponding window dimensions). The general characteristic is for the MAD to increase with sliding window size, suggesting that there is not an overfitting problem for small window sizes. The approximation of radiometric transfer and viewing geometry effects as locally linear relationships (Equation 3.5 and Equation 3.12) will be more precise for smaller sliding windows. As the sliding window size increases, the effective resolution of the radiometric homogenisation decreases, which likely results in the increasing MAD. This result supports the choice of a one pixel sliding window for the case study.

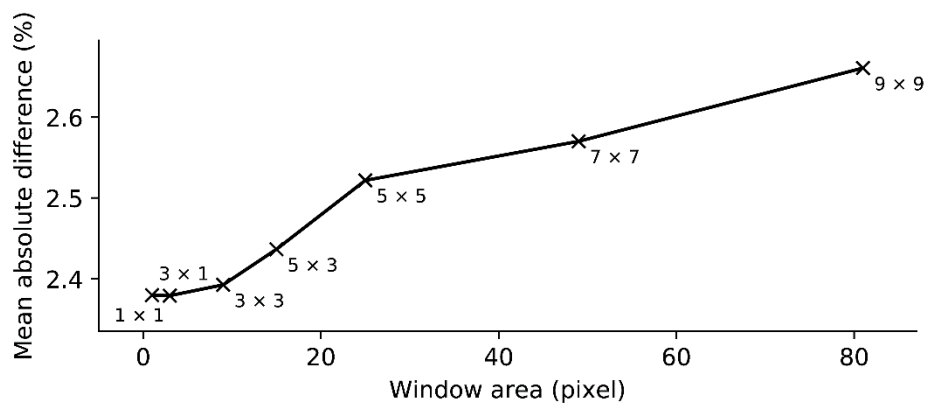


Figure 3.14 Effect of sliding window size on SPOT 5 comparison

Comparisons of diagnostic SPOT 5 and DMC spectra are shown in Figure 3.15. These spectra were manually selected from single pixels in homogenous areas. There is a marked improvement in the similarity of the DMC and SPOT 5 surface reflectance values after homogenisation for these critical cases. The MAD and RMS difference of the SPOT 5 and DMC diagnostic reflectance spectra are 4.52% and 5.70% respectively. While not representative of wider variation, these values are similar to the ones produced by the statistical image analysis (see Table 3.3).

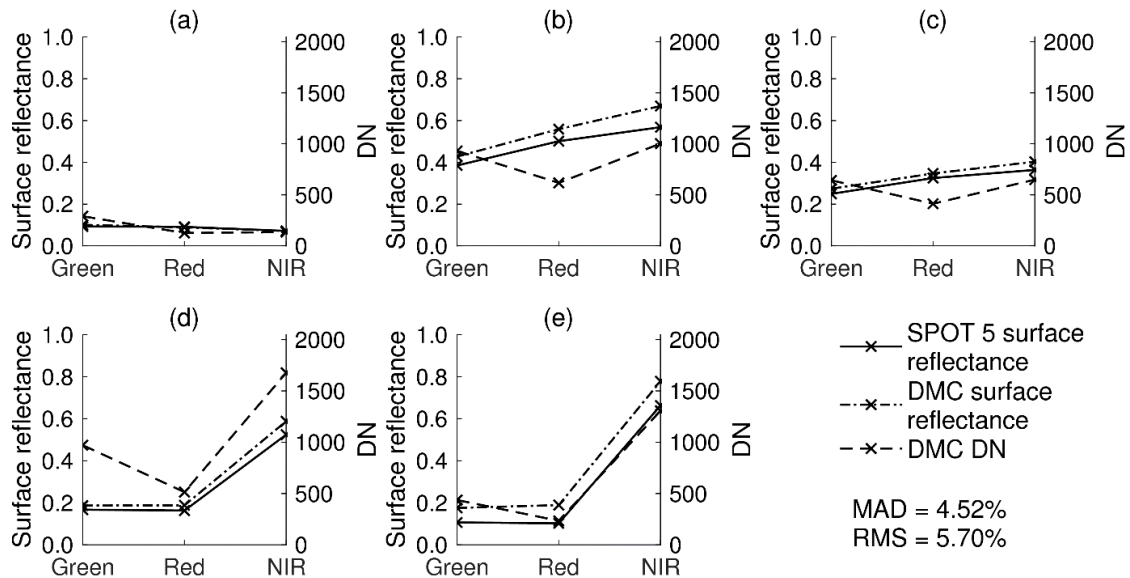


Figure 3.15 Comparison of DMC and SPOT 5 spectra for (a) water, (b) bright sand, (c) bare ground, (d) vegetation type 1 and (e) vegetation type 2 surfaces

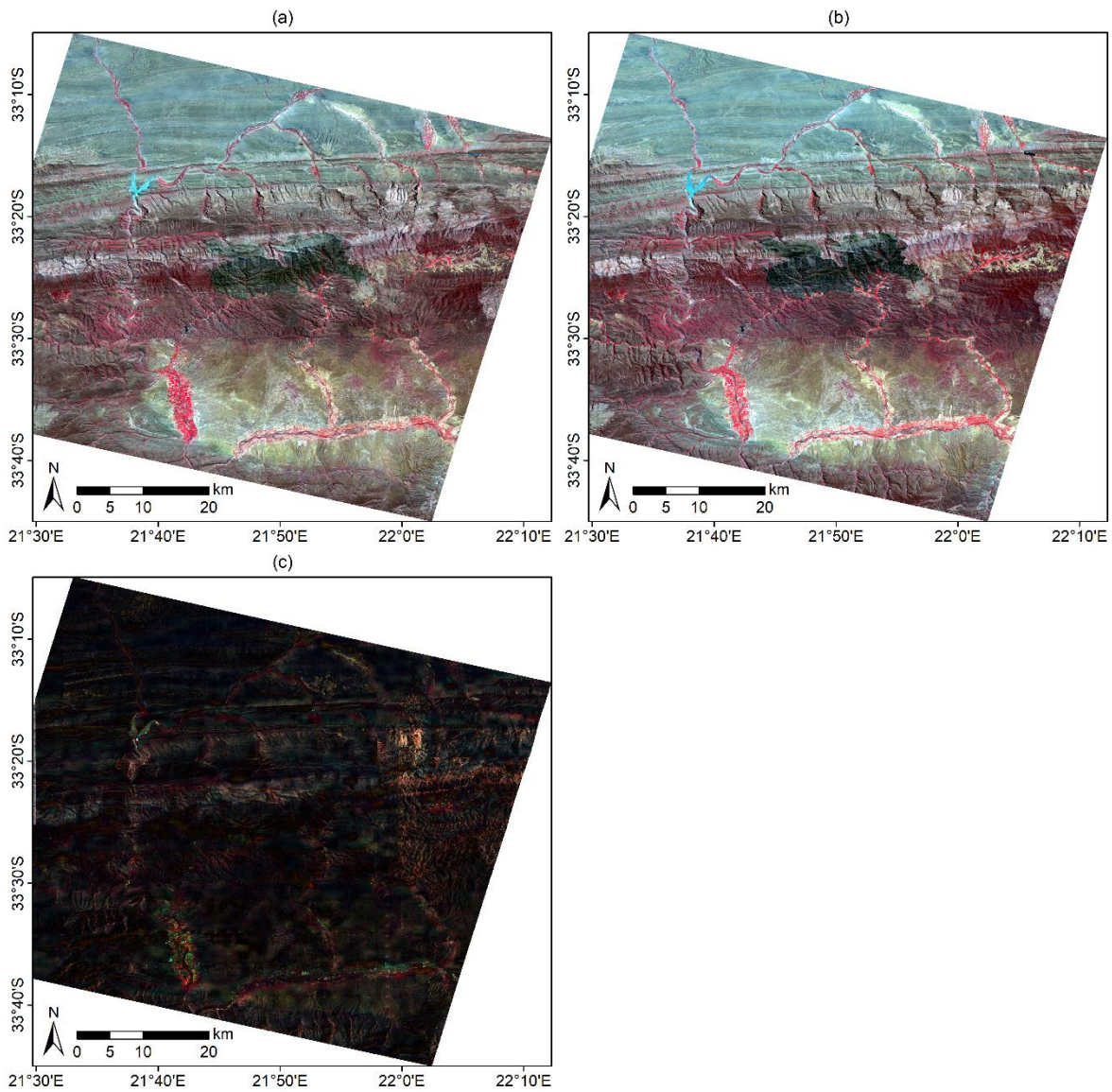


Figure 3.16 DMC and SPOT 5 surface reflectance comparison with (a) DMC homogenised mosaic masked to SPOT 5 extent; (b) SPOT 5 surface reflectance image; and (c) contrast stretched absolute difference image

False colour CIR (colour-infrared) renderings of the DMC, SPOT 5 and difference images are shown in Figure 3.16. The contrast stretched difference image shows that most discrepancies occur in the rugged mountainous areas that extend west to east in the northern section of the scene and in densely vegetated areas along river banks in the southern section of the scene. No shadow or terrain correction was performed on the SPOT 5 image. Shadow variations occurring in the DMC images below the scale of the reference resolution are not accounted for by the proposed method. Disparities in the mountainous areas are mainly due to differing, uncorrected shadow effects likely caused by variations in the time of day when the images were captured (the aerial images were captured throughout the day, while the SPOT 5 image was captured at 10:29). A particularly bright area is noticeable in the upper right corner of the difference image. This corresponds to an area of bare ground that is bright in both the DMC and MODIS images and likely corresponds to a BRDF correction failure. It is not possible to say if this failure occurs in the SPOT 5 and/or DMC corrections. The differences in the densely vegetated and cultivated areas are attributed to the differences in the MODIS, DMC and SPOT 5 sensor NIR RSRs being amplified by the known high NIR reflectivity of vegetation. Abrupt changes in BRDF may occur between adjacent fields in cultivated areas along the major rivers. As discussed in Section 3.3.2, these changes may not be captured at the MODIS resolution and could also be contributing to the NIR differences in these regions.

### 3.5 CONCLUSIONS

This study proposes a method of homogenising surface reflectance in aerial imagery by calibrating to a coarse-resolution, concurrent and collocated satellite image that has already been corrected for atmospheric and BRDF effects. It is shown that a spatially varying linear model can be used to approximate the relationship between the DN measured by the aerial sensor and the surface reflectance of the satellite image. The parameters of the model are estimated for each satellite pixel location using least squares regression inside a small sliding window. The method is limited by the low resolution of the satellite reference image. The effects of viewing geometry and land cover variations below the scale of a reference pixel are averaged out. Only gradual BRDF and atmospheric variations that can be captured by the coarser resolution of the reference image are compensated for.

The proposed surface reflectance homogenisation method was applied to 2228 Intergraph DMC images covering an area  $96 \text{ km} \times 107 \text{ km}$  in size, omitting the offset parameter,  $C$ , from the model. A MODIS MCD43A4 NBAR product was used as the surface reflectance reference. The DMC homogenised mosaic was free of visible seam lines and hot spots and matched the MODIS

reference well. The DMC homogenised mosaic was also compared to a concurrent SPOT 5 image in order to establish the method's efficacy at a spatial resolution closer to that of the DMC source resolution than the MODIS reference. The SPOT 5 image was corrected for atmospheric effects and converted to surface reflectance using the ATCOR 3 method. The mean  $R^2$  value and the mean absolute reflectance difference between the DMC mosaic and SPOT 5 image were 0.84 and 3.43% respectively. Despite the limitations and approximations inherent in the method, these statistics are considered supportive of the method's efficacy and are similar to figures reported by Collings et al. (2011) and López et al. (2011) for related correction techniques. While the method was evaluated on aerial imagery, it is not limited to this type of data and could be applied to any multi-spectral VHR imagery, including satellite and unmanned aerial vehicle (UAV) imagery.

The proposed technique does not require explicit BRDF and atmospheric correction; and mosaic normalisation techniques to reduce seam lines are not necessary. The spatially varying linear model allows for flexibility in the BRDF characteristics that can be corrected for. The method accuracy is limited by the accuracy of the reference surface's reflectance i.e. the accuracy of the homogenised images can at best be that of the MODIS reference. The method is also limited by the need for a reference image concurrent and spectrally similar to the aerial imagery. Such an image may not always be obtainable. The MODIS and DMC RSRs are quite different in the near-infrared region of the spectrum (see Figure 3.2). The surface reflectance homogenisation method approximates the effect of different sensor spectral responses with a linear relationship that is contained by the model of Equation 3.5. This approximation was supported by a simulation of MODIS and DMC measurements for typical land cover spectra. The relatively higher (4%) NIR reflectance difference between the DMC mosaic and the SPOT 5 values, and discrepancies in vegetated areas, are likely due to the more exaggerated differences in NIR RSRs between the MODIS, DMC and SPOT sensors.

While the results of the surface reflectance homogenisation technique were surprisingly good given the simplicity of the method, some aspects warrant further investigation. The effects of including the offset parameter,  $C$  should be investigated. The offset parameter may improve homogenised results where atmospheric effects like haze are relatively severe. Local terrain effects are poorly represented at the MODIS resolution. Landsat surface reflectance offers a higher resolution alternative but has the drawback of no BRDF correction (Schmidt et al. 2012). It could nevertheless be a useful homogenisation reference, as it will exhibit less BRDF variation than low altitude aerial imagery due to its narrower FOV. The MISR instrument is also a promising alternative to MODIS. MISR RSRs are a better match to those of the Intergraph DMC than the



MODIS bands, and it is possible to obtain 275 m reflectance products using MISR-HR (Verstraete et al. 2012). The MISR instrument captures data at nine different angles, which allows a more accurate modelling of the BRDF compared to the kernel-based approach followed in the calibration of the MODIS data (Strahler & Muller 1999). It would be informative to test the performance of the method with Landsat and MISR surface reflectance reference images.

## **CHAPTER 4: FEATURE CLUSTERING AND RANKING (FCR) FOR SELECTING STABLE FEATURES FROM HIGH DIMENSIONAL REMOTELY SENSED DATA<sup>1</sup>**

### **4.1 ABSTRACT**

High dimensional remote sensing data sets typically contain redundancy amongst the features. Traditional approaches to feature selection are prone to instability and selection of sub-optimal features in these circumstances. They can also be computationally expensive, especially when dealing with very large remote sensing datasets. This article presents an efficient, deterministic feature ranking method that is robust to redundancy. Affinity propagation is used to group correlated features into clusters. A relevance criterion is evaluated for each feature. Clusters are then ranked based on the median of the relevance values of their constituent features. The most relevant individual features can then be selected automatically from the best clusters. Other criteria, such as computation time or measurement cost, can optionally be considered interactively when making this selection. The proposed feature selection method is compared to competing filter approach methods on a number of remote sensing data sets containing feature redundancy. Mutual information and naive Bayes relevance criteria were evaluated in conjunction with the feature selection methods. Using the proposed method, it was shown that the stability of selected features improved under different data samplings, while similar or better classification accuracies were achieved compared to competing methods.

### **4.2 INTRODUCTION**

In image classification, the amount of training data required to adequately represent class distributions in feature space increases exponentially as the number of features (variables) is increased – a phenomena known as the “curse of dimensionality” (Bishop 2003). For finite training samples, increasing the features beyond a certain point results in overtraining and a decrease in the classifier accuracy. This so-called “peaking phenomenon” (Jain, Duin & Mao 2000) requires the size of the feature set to be reduced to a salient minimum in order to achieve an accurate classification. Support vector machine (SVM) (Burges 1998) and random forest (Breiman 2001) classifiers have become popular in remote sensing, partly because of their lack of sensitivity to the peaking phenomenon (Guyon et al. 2002), but several studies have demonstrated

---

<sup>1</sup> This chapter is published in the International Journal of Remote Sensing and consequently conforms to the prescribed structure of that journal

the benefits of feature reduction when these classifiers were applied to high dimensional data (Guyon et al. 2002; Strobl et al. 2008; Tolosi & Lengauer 2011). Reducing the number of features is also beneficial from the perspective of measurement costs and feature computation time. This is particularly relevant in large-scale remote sensing studies involving very high resolution (VHR) imagery, as vast quantities of data require processing.

Two basic approaches to feature set reduction exist: feature selection and feature extraction. In feature extraction the feature set is mapped into a new feature space of reduced dimensionality (Webb 2002). A major disadvantage of the feature extraction approach is that it requires measurements and computations to produce the full feature set, which can be prohibitively costly. Feature extraction also hinders interpretability as it alters the original representation of the features. A feature extraction approach was thus not followed in this study.

Feature selection involves the selection of a subset of features from the original set according to some criterion of subset performance. The number of possible subsets increases combinatorially with the size of the feature set and it is seldom practical to evaluate all possible subsets (Jain, Duin & Mao 2000). A variety of search schemes exist for reducing the portion of feature space searched. The fastest and most straightforward search scheme is simply to rank features based on their individual performance and select the best  $N$ . However, feature ranking approaches are problematic for data sets containing feature redundancy. In these situations, correlated features are ranked similarly, resulting in sub-optimal and redundant feature sets.

More advanced search schemes use greedy sequential approaches, such as forward selection and backward elimination. Compared to the feature ranking approach, greedy search procedures are more likely to find the globally optimal feature set as they explore more of the search space and are less inclined to select multiple redundant features (Webb 2002). The forward selection (FS) approach starts with an empty feature set and proceeds in a number of steps where one feature is added to the selected set at each step. The feature whose selection most improves an accuracy criterion is the one that is selected for that step. The selection process proceeds for a set number of steps or until a stopping criterion is reached (Bishop 2003). The backward elimination (BE) method starts with the full set of features and proceeds in a number of steps where one feature is eliminated from the selected set at each step. The feature whose removal produces the best accuracy according to some criterion is the one eliminated for that step. Again, the BE selection process proceeds for a set number of steps or until a stopping criterion is reached (Bishop 2003). BE is computationally more costly than FS as it begins evaluation on the full feature set. For the same reason, it also requires adequate data to represent the full feature set.

Feature selection methods can be grouped into filter, wrapper and embedded approaches. In the filter approach, generic measures of separability or importance are used to evaluate feature subsets, while in the wrapper approach, the accuracy of a specific classifier trained on the feature subset is used as the selection criterion (Duin & Tax 2005). An embedded approach is one where feature selection is incorporated into the classifier training procedure, such as with random forests (Breiman 2001). Filter approaches have the advantage over wrapper and embedded approaches of making feature selection independent of the classifier, thus allowing for greater flexibility in the choice of classifier (Brown et al. 2012). In general, filter approaches are also computationally more efficient than wrapper approaches. This is an important consideration for large and high dimensional data sets such as those often encountered in remote sensing. This study thus focuses on filter approaches.

High dimensional feature spaces typically contain feature redundancy (Cukur et al. 2015; Tolosi & Lengauer 2011; Yu & Liu 2004). Although feature correlation and redundancy are related, they are not strictly the same thing (Brown et al. 2012; Guyon & Elisseeff 2003). Features can help improve separability when the within class correlation is stronger than the between class correlation. We use the term “redundancy” to refer to correlation of features between classes. The raw bands of multi-spectral imagery often have significant spectral overlap and consequently are correlated with one another. This spectral overlap will exacerbate the redundancy amongst features derived from these raw bands (Cukur et al. 2015). Hyperspectral imagery is also well-known for containing redundancy amongst the bands (Yuan, Zhu & Wang 2015).

A number of authors have noted difficulties in selecting features from high dimensional data sets. Kononenko et al. (1997), Guyon et al. (2002), Yu & Liu (2004) and Yousef et al. (2007) noted that feature redundancy can have a negative impact on the optimality of feature selection. Feature redundancy not only leads to sub-optimal feature selection, but also makes selected features unstable and sensitive to small changes in the data used for selection (Guyon & Elisseeff 2003; Kalousis, Prados & Hilario 2007; Li, Harner & Adjero 2011; Tolosi & Lengauer 2011).

The increasing availability of high spatial and spectral resolution imagery necessitates computationally efficient feature selection techniques robust to high dimensional redundant spaces. In this article we propose a computationally efficient filter approach feature selection method for addressing the problems of sub-optimality and instability associated with high dimensional, redundant feature spaces of remotely sensed data. We adopt the filter approach due its separation of feature selection and classification tasks. The method employs affinity propagation (Frey & Dueck 2007) to identify clusters of redundant features and redundancy is

reduced by selecting a single representative feature from the most relevant clusters. The method requires no prior knowledge of the number of clusters. Correlation is used to measure feature similarity, which allows a broader encapsulation of feature redundancy than distance measures such as Euclidean distance (Chen et al. 2017). Assumptions of linear dependence between features and class labels made in structured sparsity regularisation approaches (Gui et al. 2017) are also avoided by selecting features with a relevance heuristic, based on the use of naive Bayes or mutual information criteria. The proposed method can be fully automated, or it can be used interactively to allow for consideration of computation time and measurement cost. The performance of the proposed method is compared to popular feature selection approaches, on a number of remote sensing data sets. The various feature selection methods are evaluated based on computation time, classification accuracy and stability of selected features under different data samplings.

## 4.3 METHODS

### 4.3.1 Related work

A number of feature selection approaches have been developed to address the issues of stability and sub-optimality encountered in high dimensional and redundant data. These methods consider the trade-off between feature relevance (i.e. how much information the feature contains about the class labels) and redundancy.

A means of selecting good features from redundant spaces was devised by Yousef et al. (2007), who used a  $k$ -means algorithm to produce a fixed number of clusters of correlated features. The accuracy of a SVM classifier is found for all the features of each cluster and the lowest performing clusters are eliminated, the remaining features combined, and the process is repeated until a desired number of clusters is reached. A related feature selection method that finds and removes redundancy by clustering features into similar groups was presented by Mitra, Murthy & Pal (2002). They used a novel clustering algorithm to group correlated features based on a similarity measure they call “maximal information compression index”, which is the smallest eigenvalue of the feature covariance. Sahu & Mishra (2011) also used  $k$ -means clustering to group redundant features. The best feature, according to an importance measure, is then selected from each cluster. Cukur et al. (2015) proposed a similar method, where redundant features are clustered and top ranked features selected from each cluster using an importance measure called “minimum redundancy maximum relevance” (mRMR).

A two-step procedure called the “Fast Correlation Based Filter” (FCBF), was developed by Yu & Liu (2004). It first creates a reduced set of relevant features, using a non-linear correlation

measure, called symmetrical uncertainty, that measures both feature relevance and redundancy. Relevance is measured by how well features are correlated with class labels and redundancy is measured by how well features are correlated with each other. Redundant features are then removed from this set using a search procedure based on Markov blanket filtering.

Wu et al. (2013) compared a number of filter approach feature selection methods for reducing redundancy in three hyperspectral data sets. They used a number of performance measures for comparison, including classifier accuracy, feature stability, and their own criterion called the maximal minimal associated index quotient (MMAIQ). MMAIQ uses Cramer's  $V$ -test values to trade feature relevance against redundancy and is applied in a FS type routine. While the authors concluded that MMAIQ provided the best overall performance, it did not provide good stability for high dimensional data.

A number of feature importance measures (including the FCBF) were incorporated by Brown et al. (2012) into a common theoretical framework. These measures all consider both relevance and redundancy in some way. A comprehensive empirical study was used to compare the performance (in terms of stability and classifier accuracy) of these measures. The study tested the criteria in a FS search scheme, under varying conditions, including redundancy in high dimensional feature spaces. They concluded that joint mutual information (JMI) (Yang & Moody 2000) provides the best feature selection performance overall.

In recent years, a number of feature selection approaches based on structured sparsity regularisation have been developed (Chen et al. 2017; Chen & Gu 2015; Gui et al. 2017; Nie et al. 2010). Structured sparsity regularisation modifies the traditional sparsity regularisation approach by incorporating prior knowledge of the group structure of features to improve performance (Gui et al. 2017). The supervised multiview feature selection (SMFS) method of Chen et al. (2017) uses a structured sparsity approach that groups features by similarity and uses this similarity structure to address the trade-off between feature relevance and redundancy. In SMFS, features are clustered into homogenous groups or "views" using affinity propagation (AP) with a squared Euclidean distance similarity measure. A sparse set of features is selected from these views by a joint  $\ell_{1,2}$ -norm minimisation of an objective function comprised of loss function and regularisation terms. The formation of the loss function assumes a linear dependence between features and class labels. Feature view structure is incorporated into the  $\ell_{1,2}$ -norms so as to encourage the sparsity of selected features within views, while retaining the information of multiple heterogeneous views. The objective function is minimised with quadratic programming, which is computationally expensive compared to greedy search type feature selection methods

such as FS and JMI. Feature weights produced by the optimisation procedure can be considered an importance measure that trades relevance against redundancy.

### 4.3.2 Feature clustering and ranking

Within the context of the related research overviewed in the previous section, our proposed feature selection method consists of the following three steps:

1. Perform AP clustering (Frey & Dueck 2007) of the feature set using the absolute value of the correlation coefficient as the similarity metric.
2. Rank each cluster's importance by finding the value of a relevance criterion for each individual feature and then finding the median of the feature relevance values in the cluster.
3. Select a single feature from each of the  $N$  clusters with best importance scores.

AP is a clustering technique that identifies cluster representatives (“exemplars”), and their corresponding clusters, by an iterative scheme of message passing between data points (Frey & Dueck 2007). A matrix of pairwise similarities and a “preference” parameter are required as inputs. The preference parameter affects the number of identified clusters and may be chosen automatically based on the values of the similarities. The proposed feature selection method sets the preference parameter to the median of the similarities, which results in a moderate number of clusters (Frey & Dueck 2007). Unlike clustering techniques such as  $k$ -means, AP does not require prior knowledge of the number of clusters.

Two kinds of messages, “availability” and “responsibility”, are passed between data points at each iteration. The values of these messages express the current affinity one point has for choosing another as its exemplar. The responsibility  $r(i, k)$  reflects the accumulated evidence that feature  $k$  is the exemplar for feature  $i$ , taking into consideration other possible exemplars for feature  $i$ . The availability  $a(i, k)$  reflects the accumulated evidence for how appropriate it would be for feature  $i$  to choose feature  $k$  as its exemplar, taking into consideration support from other features for choosing  $k$  as their exemplar.

To initialise, the availabilities are set to zero,  $a(i, k) = 0$ . In our method, the similarity  $s(i, k)$  between feature  $i$  and  $k$ , is set to the absolute value of the correlation coefficient, and the self-similarities,  $s(k, k)$ , are set to the preference value. At each iteration, the responsibilities are updated using the rule:

$$r(i, k) \leftarrow s(i, k) - \max_{k' \text{ s.t. } k' \neq k} \{a(i, k') + s(i, k')\} \quad \text{Equation 4.1}$$

The availabilities are correspondingly updated using the rule:

$$a(i, k) \leftarrow \begin{cases} \min \left\{ 0, r(k, k) + \sum_{i' \text{ s.t. } i' \notin \{i, k\}} \max\{0, r(i', k)\} \right\}, & i \neq k \\ \sum_{i' \text{ s.t. } i' \neq k} \max\{0, r(i', k)\}, & i = k \end{cases} \quad \text{Equation 4.2}$$

The exemplar for feature  $i$  is identified by the value of  $k$  that maximises  $a(i, k) + r(i, k)$ . The iterations continue until the clusters (and their corresponding exemplars) remain stable for ten consecutive updates.

We investigated the performance of two different feature relevance measures: the accuracy of a naive Bayes classifier and the mutual information (MI) between the feature and the class labels. The naive Bayes classifier, using a histogram to model class densities, was chosen primarily because it makes no assumptions about the form of the class distributions and can thus provide a generic measure of separability. It is simple, fast and recognised as being accurate for a variety of problems (Hand & Yu 2001). The “naive” assumption of feature independence is of no consequence when testing individual features.

MI is a measure of the dependence between two random variables (Brown et al. 2012). Given two random variables  $X$  and  $Y$ , with probability distributions  $p(x)$  and  $p(y)$  and joint probability distribution  $p(x, y)$ , the MI between  $X$  and  $Y$  is defined as:

$$I(X; Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad \text{Equation 4.3}$$

The MI between a feature and the class labels gives a useful indication of that feature’s relevance or importance (Brown et al. 2012). The probability distributions in Equation 4.3 are not known and are estimated using histograms.

The cluster importance measure for the  $k^{\text{th}}$  cluster is expressed as



$$C_k = \operatorname{median}_{X_j \in G_k} R(X_j, Y) \quad \text{Equation 4.4}$$

where  $G_k$  is the set of features in cluster  $k$  and  $R(X_j, Y)$  is the MI or naive Bayes feature relevance measure for feature  $X_j$  and class labels  $Y$ . Once the clusters of similarly relevant features have been ranked according to their importance measures, single features may be selected from the best clusters using an automatic procedure or by the user, taking measurement cost and computation time into account. The process of selecting a representative feature from each of the top ranked clusters reduces redundancy while retaining relevance. For automatic feature selection, the feature with the maximum relevance measure is selected from each of the  $N$  best clusters. When criteria of measurement cost or computation time require consideration, the user should hand select features minimising these values from the  $N$  best clusters.

The number of clusters to select,  $N$ , can be specified by the user based on the size of the training set or by using a grid search with the final classifier accuracy as performance measure. To avoid biased accuracy estimates, all classifier accuracy evaluations, for cluster ranking or selection of  $N$ , are done on unseen test data using a five-fold cross-validation (Bishop 2003).

### 4.3.3 Data sets

Five remote sensing and one synthetic data set (Table 4.1) were used for comparing the proposed method against popular existing feature selection methods. The “difficulty” in the last column of Table 4.1 is calculated as  $N/(mc)$ , as in Brown et al. (2012), where  $N$  is the number of objects,  $m$  the number of features and  $c$  the number of classes. Smaller values indicate that the data is less representative of the underlying class distributions, which results in more challenging feature selection and classification tasks.

Table 4.1 Data sets

Name	Abbreviation	Number of Features	Number of Objects	Number of Classes	Difficulty
Spekboom	Spekboom	46	57877	3	419.40
Synthetic	Synthetic	17	10000	2	294.12
Statlog Landsat	Landsat	36	3756	6	17.39
Urban land cover	Urban	147	261	9	0.20
Botswana	Botswana	145	1330	14	0.66
Kennedy space centre	KSC	176	1365	13	0.60

The spekboom set consists of 46 spectral and textural features derived from four band multi-spectral, 0.5 m spatial resolution aerial imagery. The classes represent three types of vegetation

found in the Little Karoo, a semi-arid region in South Africa. It was created as part of a vegetation mapping project being conducted by the authors.

The two-class synthetic data set was generated to have redundancy amongst the features. The first five features for class  $j$  were generated from a normal distribution,  $N(\mu_j, 1)$ , with mean  $\mu_j$  and standard deviation of one ( $j = 1..2$ ). The mean,  $\mu_j$ , of each distribution, was generated from the standard normal distribution,  $N(0,1)$ . The same number of objects were generated for each class. To introduce redundancy, an additional five features were generated by adding normally distributed noise,  $N(0,0.25)$ , to the original five features. A further five redundant features were similarly generated, but by adding normally distributed noise,  $N(0,0.5)$ , to the original features. Finally, two spurious features, sampled from a standard normal distribution,  $N(0,1)$ , were added to the data set.

The Statlog Landsat and Urban Land Cover data sets were obtained from the UCI Machine Learning Repository (Lichman 2013). The Statlog Landsat features are generated from six band multi-spectral pixel values in three by three neighbourhoods. The data set consists of six land cover classes. The features of the Urban Land Cover data set are comprised of multi-scale spectral, size, shape and textural measures, derived from high-resolution aerial imagery (Johnson & Xie 2013).

Kennedy Space Centre (KSC) and Botswana are public hyperspectral data sets with vegetation and land cover classes (GIC 2014). The Botswana data were acquired by the Hyperion sensor on board the NASA EO-1 satellite and consist of 145 bands in the 400-2500 nm portion of the spectrum, at a 30 m pixel resolution. The KSC data were acquired by the NASA AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor and consist of 176 bands in the 400-2500 nm range, acquired at a spatial resolution of 18 m.

#### **4.3.4 Experimental design**

The proposed feature clustering and ranking (FCR) method was compared to a number of other established and competing feature selection methods. We adopted a similar, although reduced, evaluation approach to that of Brown et al. (2012) and Wu et al. (2013). The compared methods included the standard search approaches of ranking, FS and BE. These standard approaches and FCR, were each evaluated with two different feature relevance criteria: MI and the naive Bayes classification accuracy. The MI relevance criterion for FCR and ranking approaches finds the MI between individual features and the class labels. To integrate the MI relevance criterion into FS

and BE, it is necessary to compute the MI of a set of multiple candidate features with the class labels. In this situation, the candidate features are first merged into a joint variable and then the MI of the class labels with this joint variable is computed (Brown et al. 2012). We used histograms with ten bins along each dimension to approximate probability densities for both the MI and naive Bayes criteria (Webb 2002), to avoid difficulties and inefficiencies associated with estimating probability densities for continuous variables (Brown et al. 2012).

In addition to the MI and naive Bayes criteria, two other criteria, namely JMI and maximum relevance minimum redundancy (mRMR), were included in our study to represent “state of the art” performance. Brown et al. (2012) compared the performance of several feature selection criteria in redundant high dimensional spaces, and found the JMI criterion gave the best overall performance in terms of classification accuracy and stability. The JMI measure for feature  $X_k$  is

$$J_{\text{JMI}}(X_k) = \sum_{X_j \in S} I(X_k X_j; Y) \quad \text{Equation 4.5}$$

where  $Y$  are the class labels and  $S$  is the set of previously selected features. JMI considers the MI between the class labels and the joint variables  $X_k X_j$ , which are the pairwise combinations of the candidate feature with each feature already selected. It measures how well the candidate feature complements selected features in describing the class labels.

The popular mRMR criterion, introduced by Peng, Long & Ding (2005), expresses the trade-off between feature relevance and redundancy using MI measures. The mRMR measure for candidate feature  $X_k$  is

$$J_{\text{mRMR}}(X_k) = I(X_k; Y) - \frac{1}{|S|} \sum_{X_j \in S} I(X_k; X_j) \quad \text{Equation 4.6}$$

The first term expresses relevance as the dependence between the candidate feature and class labels, while the second term approximates the redundancy as the mean pairwise dependencies between the candidate and previously selected features. The JMI and mRMR criteria are used in a FS search scheme. The evaluated methods are detailed in Table 4.2. In this evaluation, the term “method” is used to refer to a combination of search scheme, such as FS, and criterion, such as JMI.

Table 4.2 Methods as combinations of search schemes and criteria

Search scheme	Criterion	Method abbreviation
Forward selection	Joint mutual information	FS-JMI
Forward selection	Maximum relevance minimum redundancy	FS-MRMR
Feature clustering and ranking	Naive Bayes	FCR-NaiveBC
Ranking	Naive Bayes	Rank-NaiveBC
Forward selection	Naive Bayes	FS-NaiveBC
Backward elimination	Naive Bayes	BE-NaiveBC
Feature clustering and ranking	Mutual information	FCR-MI
Ranking	Mutual information	Rank-MI
Forward selection	Mutual information	FS-MI
Backward elimination	Mutual information	BE-MI

To quantify the stability of the selected features, we used the consistency index developed by Kuncheva (2007). If  $A$  and  $B$  are subsets of the full feature set  $X$ , with  $|A| = |B| = k$ ,  $r = |A \cap B|$  and  $0 < k < |X| = n$ , the consistency index is

$$C(A, B) = \frac{rn - k^2}{k(n - k)} \quad \text{Equation 4.7}$$

Its value lies in the range  $[-1, +1]$ , where positive values indicate similar sets, zero indicates a random relation and negative values indicate an anti-correlation between the feature sets (Kuncheva 2007). To evaluate stability for a particular method, we select features from bootstrap samples of the data. The consistency index is found for each pairwise combination of selected features over ten bootstraps of the data and averaged to give a measure of overall stability.

A  $k$ -nearest-neighbour ( $k$ -NN) classifier (with  $k = 3$ ) was used to evaluate the accuracy of the features selected by each method.  $k$ -NN is a generic classifier that makes no assumptions about the data and requires no tuning. While other classifiers may be more accurate in particular situations,  $k$ -NN allows a relative comparison of the feature selection methods, independent of the influence of classifier tuning for specific data. For each of the feature sets found from the bootstrap samples, the  $k$ -NN accuracy was found as the average per-class accuracy from a ten-fold cross-validation. For each method and data set combination, an overall accuracy was computed as the average of the bootstrap accuracies.

The number of features to select for each data set was fixed across methods. This parameter was selected by using the accuracy of a  $k$ -NN classifier ( $k = 3$ ), trained on the first  $N$  features selected by FS-NaiveBC, as the criterion in a grid search. A low value of  $N$  that achieved good accuracy

was selected for each data set. The number of features selected for each data set are detailed in Table 4.3.

Table 4.3 Feature selection parameters

Data Set	Number of Features to Select
Spekboom	6
Synthetic	5
Statlog Landsat	4
Urban land cover	4
Botswana	6
Kennedy Space Centre	7

The FCR methods (FCR-MI and FCR-NaiveBC) required some specific treatment to integrate them into the evaluation. After bootstrapping, clusters were assigned unique indices, ensuring identical clusters had the same index. The consistency index was then found using the selected cluster indices rather than feature indices. This was done to simulate hand-selection of preferred features from the best clusters for each bootstrap, while allowing the FCR algorithm to automatically choose the top ranked feature from each cluster (for the sake of simplicity and speed). In other words, the cluster index was used to represent the index of the preferred feature that could otherwise have been selected by hand from the cluster contents.

We followed a similar approach to that of Brown et al. (2012) for computing a single “non-dominated” ranking of the methods that considers stability and accuracy performance simultaneously. The concept of “Pareto optimality” was used to find a single optimal solution in terms of multiple criteria. In the context of our evaluation, the “Pareto front” is the set of methods on which no other method can improve without degrading either the accuracy or stability. The methods in this set are called “non-dominated” (Mishra & Harit 2010). Successive Pareto fronts can be formed iteratively by finding the current Pareto front of the set of methods that excludes members of the previous fronts. A method was thus given a non-dominated rank of  $N$  if it was a member of the  $N^{\text{th}}$  Pareto front. The average of the non-dominated ranks for each method over the six data sets was used to produce an overall ranking.

The bulk of the software implementation was done in Matlab<sup>TM</sup>, making use of the PRTools toolbox (TU Delft 2015). The MI, JMI and mRMR criteria were computed using the FEAST (FEAture Selection Toolbox) C++ implementation (Brown et al. 2012).

#### 4.4 RESULTS AND DISCUSSION

The results of the stability and accuracy evaluations for each method and data set combination are shown in Figure 4.1 and Figure 4.2 respectively. The methods appear along the x axis in order of their mean stability in Figure 4.1, and mean accuracy in Figure 4.2, over the six data sets. The wide range of stabilities confirms the sensitivity of some methods to variations in the data. The method accuracies span a smaller range than the method stabilities. Nonetheless, there were substantial differences in accuracy between the best and worst methods. Compared to the other data sets, the stability of the spekboom, Synthetic and Landsat data was noticeably superior. As reflected in the “difficulty” values in Table 4.1, these data sets are more representative of the underlying distributions and are thus less sensitive to disturbances.

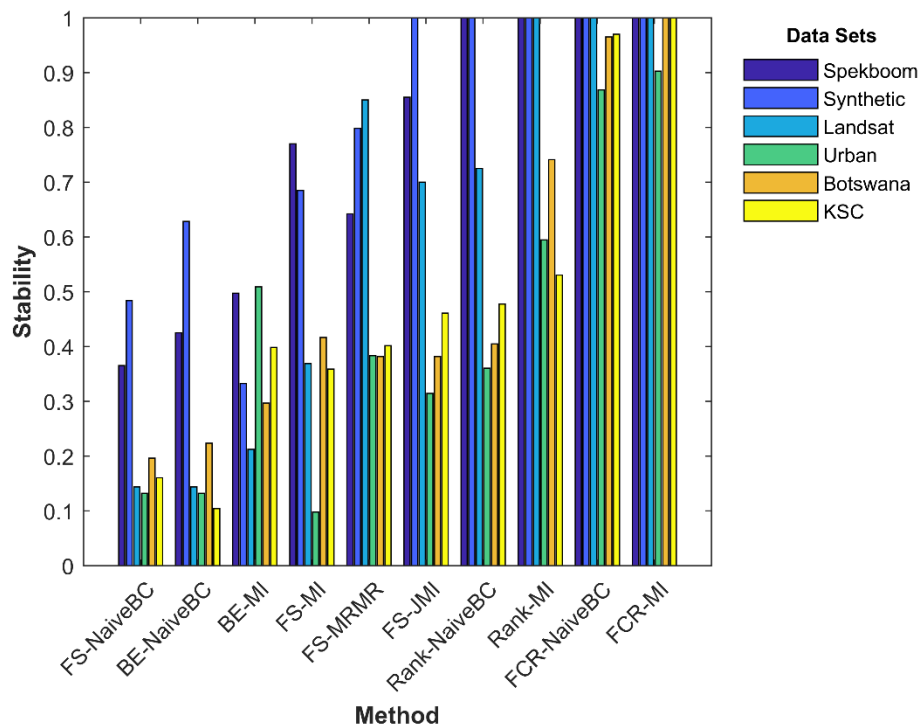


Figure 4.1 Method stability per data set (methods along the x axis are ordered by their mean stability over the data sets)

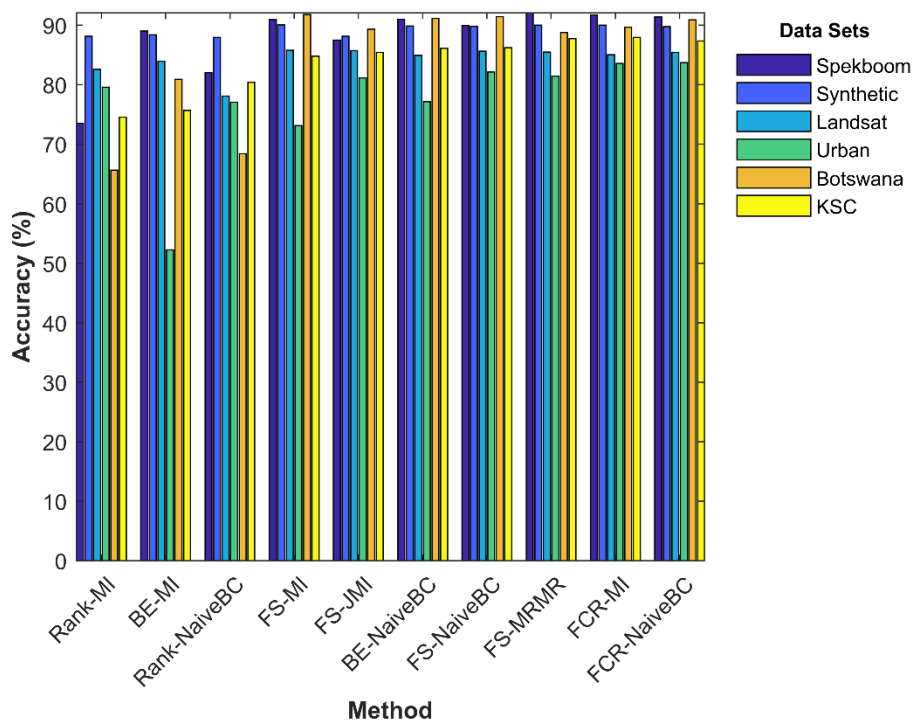


Figure 4.2 Method accuracy per data set (methods along the x axis are ordered by their mean accuracy over the data sets)

FCR-NaiveBC and FCR-MI occupy the top two positions for both performance measures. The ranking methods, Rank-MI and Rank-NaiveBC both had poor accuracy performance. This was expected as these methods do not consider feature complementarity and only measure relevance of features in isolation. The relatively poor accuracy and stability of FS-JMI was surprising in the context of the results of Brown et al. (2012), where it produced the best overall performance. The FS-JMI results nevertheless provide a benchmark that helps confirm the usefulness of the FCR method for the type of data investigated in our study.

As with classifier design, there is a “curse of dimensionality” problem with computing the MI of joint variables. As the number of features increases, the number of objects needed to adequately represent the feature distribution increases exponentially (Brown et al. 2012). For this reason, the MI criterion is not well suited for evaluating the BE search scheme, which requires computation of the relevance criterion for the full feature set. This likely explains the poor performance of BE-MI in terms of both accuracy and stability. Part of the motivation for the JMI and mRMR formulations is to circumvent this kind of representivity issue by using a low dimensional approximation to MI.

The method execution times, summed over the six data sets, are provided in Table 4.4. The execution time of FCR competed well with the other methods, although mRMR was the fastest overall. The naive Bayes criterion is slower to compute than the MI criterion as it uses a five-fold

cross-validation to evaluate the classification accuracy, while MI is computed once-off. Methods using the naive Bayes criterion are consequently slower than their MI counterparts. FS-JMI is faster than the related FS-MI method, as the criterion only requires MI computations between pairwise combinations of features and the class labels, while the MI criterion is evaluated on the combination all selected features. BE is known to be less efficient than FS (Guyon & Elisseeff 2003), and was the slowest of the tested search schemes.

Table 4.4 Method cumulative execution time over all data

Method	Time (s)
FS-MRMR	1.40
Rank-MI	2.11
FS-JMI	2.46
FCR-MI	3.01
FS-MI	12.08
FCR-NaiveBC	72.61
Rank-NaiveBC	73.05
BE-MI	242.39
FS-NaiveBC	429.76
BE-NaiveBC	3340.84

Table 4.5 presents the non-dominant ranking of the methods, in terms of both accuracy and stability. The best ranked method overall was FCR-MI, followed by FCR-NaiveBC. The Rank-NaiveBC, Rank-MI, BE-NaiveBC and BE-MI methods were ranked lowest due to the known limitations of these methods. The FS-MRMR method was competitive on all measures, and was the third ranked method overall.

Table 4.5 Non-dominated ranking of methods by accuracy and stability

Method	Rank
FCR-MI	1
FCR-NaiveBC	1.67
FS-MRMR	2
FS-JMI	2.50
FS-MI	2.83
FS-NaiveBC	2.83
Rank-MI	3
BE-NaiveBC	3.33
Rank-NaiveBC	3.33
BE-MI	4

If the clustering step were omitted, FCR-MI and FCR-NaiveBC would simplify to Rank-MI and Rank-NaiveBC respectively. FCR-MI and FCR-NaiveBC showed a substantial improvement in performance compared to Rank-MI and Rank-NaiveBC, which lends support to the effectiveness of the clustering step. Considering the combination of the MI and naive Bayes criteria with each search scheme in isolation, there was a general trend for MI to produce better stability and naive



Bayes to produce better accuracy. While FCR worked well with either criterion, the results favoured the use of MI as it is faster and produced a better non-dominant ranking than naive Bayes. On the whole, the evaluations demonstrate that the proposed FCR method is effective at selecting accurate and stable features from high dimensional remote sensing data containing redundancy.

## 4.5 CONCLUSIONS

Small changes in data sets containing redundancy can result in substantial changes in selected features. Feature redundancy is also known to cause selection of sub-optimal features. This study presented and evaluated a new method for selecting stable and informative features from redundant data by ranking correlated clusters of features. The method uses AP to identify a moderate number of clusters of correlated and similarly relevant features. It then ranks the clusters using an importance measure, calculated as the median of a relevance criterion evaluated on each individual feature in the cluster. By selecting an individual feature from each of the best clusters, a set of informative features is found while simultaneously removing redundancy from the data. These features may be selected automatically based on their relevance, or interactively, taking criteria such as computation time and measurement cost into account. The option to include factors other than relevance and redundancy in determining selected features distinguishes FCR from related feature selection methods; although this option is currently limited to a manual procedure. Future work will investigate ways of combining relevance with other criteria to allow feature selection without user input.

The effectiveness of the proposed FCR method was evaluated by comparing its accuracy, stability and execution time to a set of popular feature selection methods. Two criteria were tested for measuring feature relevance: the MI between the candidate feature(s) and the class labels, and the accuracy of a naive Bayes classifier trained on the candidate feature(s). Unlike structured sparsity regularisation approaches, these relevance criteria do not assume a linear dependence between features and class labels. FCR performed well overall, with both naive Bayes and MI criteria and was the highest ranked method when considering the accuracy and stability measures in combination. Another benefit of FCR is its relative speed compared to greedy search FS and BE type methods. Ever increasing quantities of high spatial and spectral resolution remote sensing data are being produced and require interpretation (Chi et al. 2016). In this context, instability and sub-optimality associated with feature selection from high dimensional redundant data will become increasingly important. Computationally efficient techniques, such as FCR, are required to address these challenges.

## CHAPTER 5: REGIONAL MAPPING OF SPEKBOOM CANOPY COVER USING VERY HIGH RESOLUTION AERIAL IMAGERY<sup>1</sup>

### 5.1 ABSTRACT

Very high resolution (VHR) canopy cover maps of spekboom are required to assist with the restoration of degraded habitat in the Little Karoo, a large semi-arid region in South Africa. Variations in habitat and level of degradation, in addition to radiometric variations in the imagery, make spekboom mapping at a regional scale a challenging problem. In this article, we present a per-pixel classification approach for canopy cover mapping of spekboom using multi-spectral 0.5 m resolution aerial imagery. The imagery was radiometrically homogenised with a technique that uses satellite data to convert digital numbers to estimated surface reflectance values. A feature selection procedure that is robust to redundancy was applied in order to select an informative feature subset from a typical set of spectral, textural and vegetation index features. Support vector machine (SVM), random forest, decision tree, k-nearest neighbour (kNN) and Bayes normal classifiers were evaluated against labelled pixel data and spekboom canopy cover ground truth acquired at 20 field sites. The results showed that all the classifiers, except the Bayes normal classifier, performed well. The decision tree produced the best results (mean absolute canopy cover error of 5.85% with a standard deviation of 4.65%).

### 5.2 INTRODUCTION

Spekboom (*Portulacaria afra*) is an evergreen succulent shrub with a dense canopy of small fleshy leaves that can grow up to 2.5 m in height. It occurs in the subtropical thicket biome in the semi-arid Eastern Cape and Little Karoo regions of South Africa (Vlok, Cowling & Wolf 2005). “The subtropical thicket habitat types are most easily recognised by the occurrence of woody trees, spinescent shrubs and a relative abundance of succulents. When occurring as solid stands the vegetation can form impenetrable dense thickets, but this is uncommon in the Little Karoo. In most of the area the subtropical thicket vegetation occurs as discrete bush clumps, usually in a matrix of Succulent Karoo vegetation” Vlok, Cowling & Wolf (2005: 37). While spekboom tolerates browsing by indigenous herbivores, it is highly susceptible to over-browsing by goats (Marais, Cowling & Powell 2009; Mills et al. 2007; Mills et al. 2005; Sigwela et al. 2009). Poorly managed

---

<sup>1</sup> This chapter is published in the Journal of Applied Remote Sensing and consequently conforms to the prescribed structure of that journal

goat grazing has transformed thicket over much of its range into sparsely scattered thicket clumps, isolated trees and a covering of herbs (Mills et al. 2005).

The benefits of restoring degraded thicket habitat are evident from a number of perspectives. Spekboom is unusually effective at storing carbon for an arid region plant (Mills et al. 2005). Subtropical thicket furthermore provides an important source of food for many herbivores, including domesticated livestock (Thompson et al. 2009; Vlok, Cowling & Wolf 2005). The re-establishment of spekboom in degraded areas will reduce erosion and flood severity and improve water quality (Van Luijk et al. 2013; Mills & Cowling 2006). The restoration of spekboom is also attractive from an employment perspective, since the restoration process could potentially create thousands of jobs in impoverished areas if implemented on a large scale. Currently, the most practical option for thicket ecosystem restoration is through the planting of spekboom cuttings (Mills et al. 2007; Van der Vyver et al. 2012). Spekboom is a keystone species and facilitates the creation of a favourable environment for the spontaneous recruitment of other plants (Adie & Yeaton 2013; Mills & Cowling 2010).

Spekboom canopy cover maps are required for assisting in the restoration process. There is a need for greater accuracy and repeatability than that provided by field-based mapping techniques. Field mapping is time consuming and costly and is not practical over large areas (Eisfelder, Kuenzer & Dech 2012; Lu 2006; Powell 2009). Manual field mapping is confounded by the density (inaccessibility), heterogeneous nature and complex growth forms of the subtropical thicket biome (Powell 2009).

Thompson et al. (2009) conducted a general degradation mapping study of the biomes occurring in the Little Karoo. A 1:50000 vegetation map, developed by Vlok, Cowling & Wolf (2005), was used to delineate different habitats so they could be treated separately. A coarse three-level degradation classification of subtropical thicket was derived by thresholding 250 m resolution MODIS normalised difference vegetation index (NDVI) data. The study was successful at estimating three degradation levels (intact, moderate and severe) of spekboom thicket at the 250 m MODIS resolution.

There is an initiative to involve private land owners in subtropical thicket restoration in order to broaden its impact (Curran et al. 2012; Mills et al. 2010). Spekboom is to be planted in stand sizes as small as three hectares on these lands. VHR imagery is required to produce maps of sufficient spatial detail for accurately monitoring spekboom canopy cover in these small stands. Spekboom often occurs in small clumps amongst a complex and varying mosaic of other vegetation. While

regional scale vegetation mapping is most commonly done at moderate to coarse spatial resolutions (González-Roglich & Swenson 2016; Immitzer et al. 2018; Liu et al. 2017; Mathieu et al. 2018), these low spatial resolutions suffer from spectral mixing of spekboom, other vegetation and soil in this heterogeneous biome (Thompson et al. 2009). VHR imagery is therefore necessary to facilitate fine-scale discrimination of spekboom from surrounding soil and vegetation. To achieve sufficient accuracy for restoration monitoring and carbon storage estimations, it is necessary to estimate spekboom canopy cover in finer detail than the three levels of degradation and 250 m resolution used in Thompson et al. (2009).

Multi-spectral VHR imagery has been used for vegetation mapping in a number of studies. A combination of spectral band, vegetation index, band ratio and textural features are commonly used to provide informative measures capable of distinguishing vegetation classes (Basu et al. 2015; Ghosh & Joshi 2014; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Mustafa & Habeeb 2014). Object-based approaches to image classification, where homogenous image objects are first generated through segmentation and then classified as a whole, have become popular in vegetation studies (Ghosh & Joshi 2014; Kollár, Vekerdy & Márkus 2013; Ouyang et al. 2011). These approaches are often favoured for VHR imagery (Basu et al. 2015; Ghosh & Joshi 2014; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Mustafa & Habeeb 2014; Ouyang et al. 2011) because they are potentially able to better exploit the additional spatial information and deal with unwanted variation when compared to the more traditional per-pixel approach (Ghosh & Joshi 2014; Ouyang et al. 2011). The segmentation problem is, however, recognised as being poorly posed, requiring manual adjustment of parameters and being difficult to solve (Baraldi et al. 2010). Per-pixel classification provided good and useful mapping accuracy in a number of studies (Boyden et al. 2007; De Castro et al. 2012; Mehner et al. 2004) and is a simpler and faster method, not requiring user specification of algorithms and associated parameters.

A variety of supervised approaches are used for classifying features derived from VHR imagery. Some authors found the Bayes normal (maximum likelihood (ML)) classifier to adequately model their class distributions (Boyden et al. 2007; De Castro et al. 2012; Mehner et al. 2004). Others adopted more sophisticated approaches such as SVMs (Ghosh & Joshi 2014) and neural networks (Basu et al. 2015; Mustafa & Habeeb 2014). Algorithms implemented in the eCognition software package (Trimble 2016), such as the fuzzy and hierarchical approaches, are also frequently used for VHR image classification (Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Ouyang et al. 2011).

As the number of features increases, the amount of data required to adequately represent class distributions in the increased feature space increases exponentially. This is known as the “curse of dimensionality” (Bishop 2003). For finite training samples, increasing the features beyond a certain point results in overtraining and a decrease in the classifier’s ability to generalise. This “peaking phenomenon” (Jain, Duin & Mao 2000) makes it necessary to apply feature selection to reduce the size of the feature set to a salient minimum in order to achieve an accurate classification. Feature selection by ranking, based on some separability or importance measure of individual features, is frequently used (Basu et al. 2015; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013). While fast, feature ranking is known to be sub-optimal for feature spaces containing redundancy (Tolosi & Lengauer 2011). Ghosh & Joshi (2014) used recursive feature elimination (also known as backward elimination) – a greedy search technique to select informative features. Of the reviewed studies, Ghosh & Joshi (2014) were the only ones to use a feature selection method that considers the effect of feature redundancy by evaluating features in combination.

The majority of the reviewed VHR vegetation mapping studies were applied to small areas, typically covered by a single satellite image (Boyden et al. 2007; Ghosh & Joshi 2014; Johansen et al. 2007; Kollár, Vekerdy & Márkus 2013; Mehner et al. 2004; Ouyang et al. 2011). Radiometric corrections are sometimes not applied in small study areas (Boyden et al. 2007) or partially handled using conversion to top of atmosphere radiance (Johansen et al. 2007; Mehner et al. 2004). These corrections do not compensate for varying atmospheric and bidirectional distribution function (BRDF) effects, characteristic of datasets containing hundreds or thousands of aerial images.

In this paper, we present a method for mapping spekboom canopy cover at a spatial resolution of 0.5 m. A total of 2228 multi-spectral aerial images, acquired over multiple days from 22 January to 8 February 2010, were used as input. Radiometric variations due to atmospheric and BRDF effects in the images were reduced using a simple yet effective technique for homogenising the digital numbers to approximate surface reflectance. This not only allows for the application of a single classification algorithm to the entire set of images, but also provides the possibility of extending the presented mapping technique spatially and temporally. An informative feature subset was selected from a typical set of spectral band, band ratio, vegetation index and textural features using a novel feature selection method that is robust to redundancy typically found in high dimensional feature sets. The selected features were used to evaluate a set of candidate classifiers.

## 5.3 DATA

### 5.3.1 Study area

The Little Karoo is a semi-arid, biodiverse area located in the Western Cape Province of South Africa (Figure 5.1). The subtropical thicket biome makes up 35.3% of the 23 439 km<sup>2</sup> area (Vlok, Cowling & Wolf 2005). A total of 54 habitat types are present, of which 10 support spekboom (Vlok, Cowling & Wolf 2005). It is estimated that about 90% of the spekboom thicket in the area is degraded (Thompson et al. 2009). In the Little Karoo, thicket usually occurs as scattered bush clumps in a background mosaic of different vegetation types including Fynbos, Succulent Karoo and Renosterveld (Vlok, Cowling & Wolf 2005). A 5893 km<sup>2</sup> (25%) representative portion of the Little Karoo was chosen as the study area, as shown in Figure 5.1. This area includes nine of the ten habitat types supporting spekboom. Details of study area spekboom habitats and mosaic vegetation types are shown in Figure 5.2.

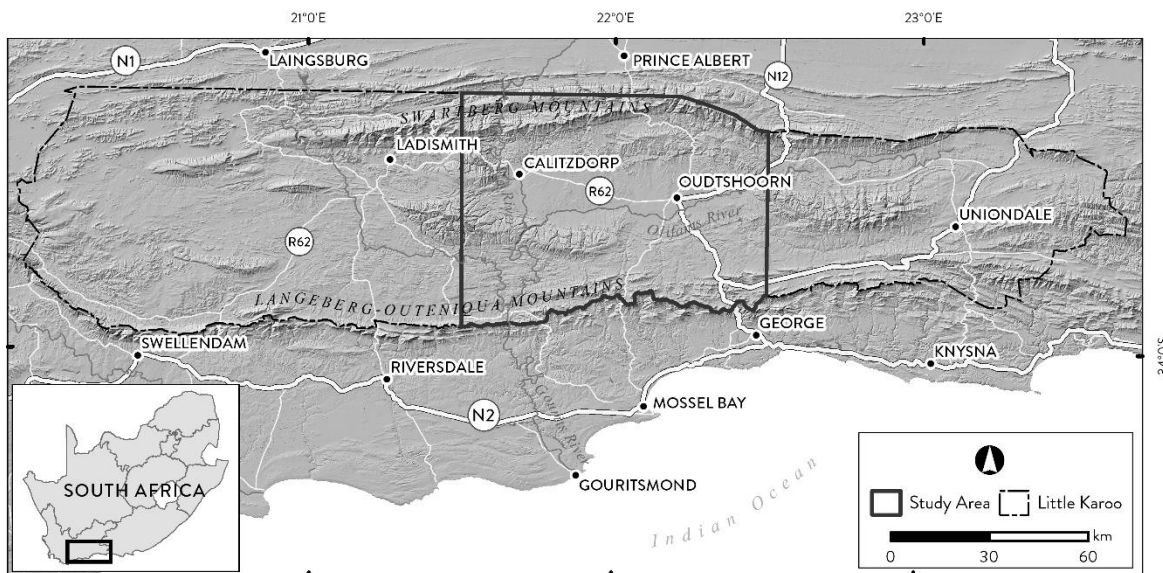


Figure 5.1 Little Karoo study area

### 5.3.2 Imagery

VHR aerial imagery of the study area was acquired from Chief Directorate: National Geo-spatial Information (NGI), a division of the South African Department of Rural Development and Land Reform. The images were captured at 0.5 m resolution with an Intergraph digital mapping camera (DMC) that provides multi-spectral red, green, blue and near-infrared (NIR) bands. The study area is covered by 2228 images, acquired over multiple days from 22 January to 8 February 2010, during the area's dry season. A distinguishing characteristic of spekboom is its year-round productivity (Ting & Hanscom 1977). The dry season imagery helped contrast the productive spekboom against the comparatively drier background vegetation.

### 5.3.3 Data collection

Two datasets were constructed: one for evaluating the fractional spekboom canopy cover estimates obtained from the classifier outputs, and a second for training and evaluating the spekboom classifier on a per-pixel basis. The first dataset consisted of in situ estimates of spekboom canopy cover acquired at 20 different sites, each of roughly one hectare. A botanist knowledgeable about the area provided expertise in the selection of sites to encompass variation in geology, habitat and level of degradation. Boundary polygons were recorded for each site by walking the perimeter with a differential global positioning system (DGPS) device. The DGPS coordinates were post-processed to provide about 30 cm accuracy. The characteristic lime-green colour and dense growth form of spekboom makes it relatively easy to visually distinguish from surrounding vegetation. During the field visits, visual estimates of spekboom canopy cover inside the site perimeters were made. The locations of the ground truth sites and their corresponding area names are shown in Figure 5.2. Details of the surrounding (“mosaic”) vegetation type, geology and estimated canopy cover of each site are given in Table 5.1. A three-level degradation measure, as used in Thompson et al. (2009), is reported for each site. This dataset is referred to as the “in situ canopy cover data” and was used for evaluating the accuracy of fractional spekboom canopy cover estimates, determined as the portion of pixels classified as spekboom over an area of interest.

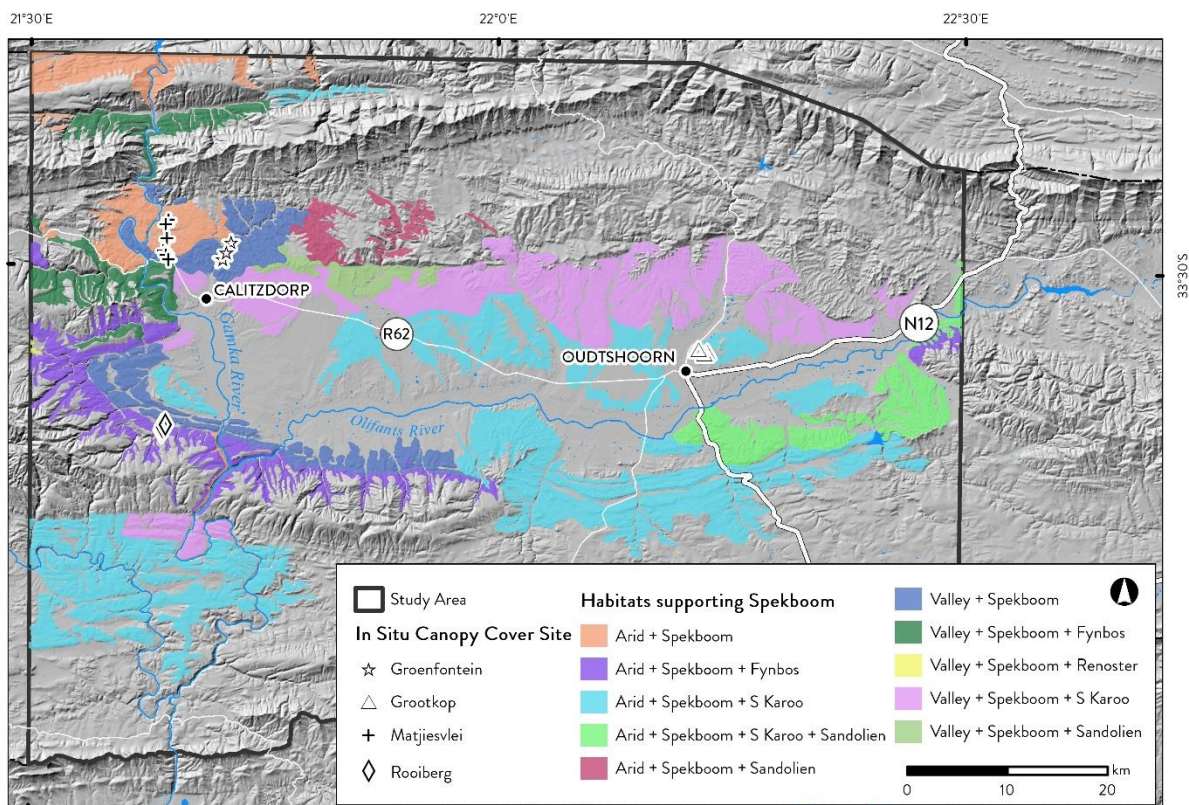


Figure 5.2 Study area spekboom habitats and field ground truth sites

Source: Vlok, Cowling & Wolf (2005)

Table 5.1 In situ canopy cover data

Area	Number	Geology	Mosaic	Degradation	Spekboom Cover (%)
Matjiesvlei	1a	Shale	Arid spekboom	Intact	6.0
	1b			Intact	22.5
	2			Intact	70.0
	3			Intact	85.0
	4			Intact	65.0
	5			Intact	37.5
	6			Intact	17.5
	7			Intact	15.0
Groenfontein	1	Shale	Arid spekboom	Severe	0.0
	2			Severe	4.0
	3			Severe	10.0
	4			Moderate	25.0
Grootkop	1	Enon conglomerate	Succulent Karoo	Intact	22.5
	2			Severe	0.5
	3			Moderate	42.5
	4			Intact	77.5
Rooiberg	1	Sandstone	Fynbos	Intact	20.0
	2			Moderate	11.0
	3			Severe	0.0

Figure 5.3 shows an example of a site perimeter on a background of the NGI imagery, rendered in RGB (red, green and blue). This in situ canopy cover data was gathered in November 2012, while the imagery was captured in January 2010. This time lag was unavoidable due to limited availability of the aerial imagery. Some change in spekboom canopy cover may have occurred during this time in recovering areas, but spekboom is relatively slow-growing (Vlok, Cowling & Wolf 2005), and these changes are not considered significant.



Figure 5.3 Matjiesvlei2 canopy cover ground truth site

For the second dataset, a labelling scheme of three classes was adopted as described in Table 5.2. While spekboom canopy cover mapping is in essence a two-class problem (spekboom and everything else), the addition of a third class, consisting of non-spekboom woody trees (labelled



as tree), was beneficial. By merging multiple classes into a single class, the ability to adjust the priors of individual merged classes is lost. The class priors, or “weights”, control the relative emphasis the classifier places on identifying each class correctly (Mingguo, Qianguo & Mingzhou 2009). Differentiation between spekboom and woody trees is a challenging part of the problem. Thus, the addition of the tree class is useful, as it allowed control over the classification accuracy of trees relative to the other classes. The size of the background class was reduced to be the same as the spekboom class by taking a random subsample. This was done to expedite classifier training times.

Table 5.2 Class descriptions

Class Name	Description
Spekboom	Spekboom
Tree	Any recognisable tree other than spekboom, but especially the darker Euclea and Papea trees commonly found intermingled in stands of spekboom
Background	Bare ground, small shrubs, herbs and anything else not included in the first two classes

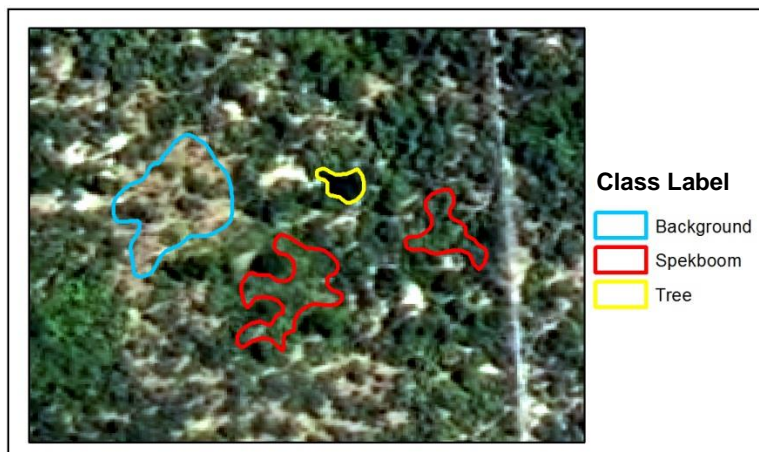


Figure 5.4 Example image class labels

Table 5.3 Details of labelled pixel data

Class Name	Polygons	Pixels
Spekboom	52	27260
Tree	64	3357
Background	44	182044
<b>Total</b>	<b>160</b>	<b>212661</b>

Due to the small (0.5 m) pixel size, limited DGPS and image spatial accuracy, dense vegetation growth and rugged terrain, it was not practical to obtain per-pixel class labels by in situ observation. Per-pixel class labels for the second dataset were therefore obtained by visual discrimination and the hand labelling of images. The distinctive colour and growth form of spekboom greatly facilitates this process (Thompson et al. 2009). Image areas belonging to the various classes were delineated as polygons in a GIS package. An example is shown in Figure 5.4. All pixels within a polygon object were assigned the same class label. Polygon areas were selected to encompass variability in habitat and level of degradation. The total number of polygons

and pixels labelled for each class is shown in Table 5.3. This dataset is referred to as the “labelled pixel data” and was used for training and evaluating the classifier on a per-pixel basis.

## 5.4 METHODS AND EXPERIMENTS

### 5.4.1 Radiometric homogenisation

While the imagery provided by NGI is orthorectified, no radiometric corrections were applied to it. The NGI imagery contains variations due to BRDF and atmospheric effects, which makes it poorly suited to quantitative remote sensing techniques. The imagery was consequently radiometrically homogenised through the application of a surface reflectance estimation technique (Harris & Van Niekerk 2019). This technique corrects for coarse scale atmospheric and BRDF effects using a well-calibrated, concurrent and collocated surface reflectance satellite image as a reference. We used a MODIS MCD43A4 composite image for the period of 25 January 2010 to 9 February 2010 for this purpose. This image has a 500 m resolution and contains nadir BRDF-adjusted reflectance data composited from the best values over a 16-day period. While Sentinel-2 (ESA 2015) or Landsat (Schmidt et al. 2012) surface reflectance could also serve as reference data, no cloud-free imagery concurrent (or near-concurrent) to the aerial imagery was available from those sources. The relative spectral responses (RSRs) of the DMC and corresponding MODIS bands are shown in Figure 5.5. Radiometric correction is important as it allows accurate snapshot mapping of large spatial extents and provides the possibility of repeating the spekboom canopy cover mapping to evaluate restoration progress.

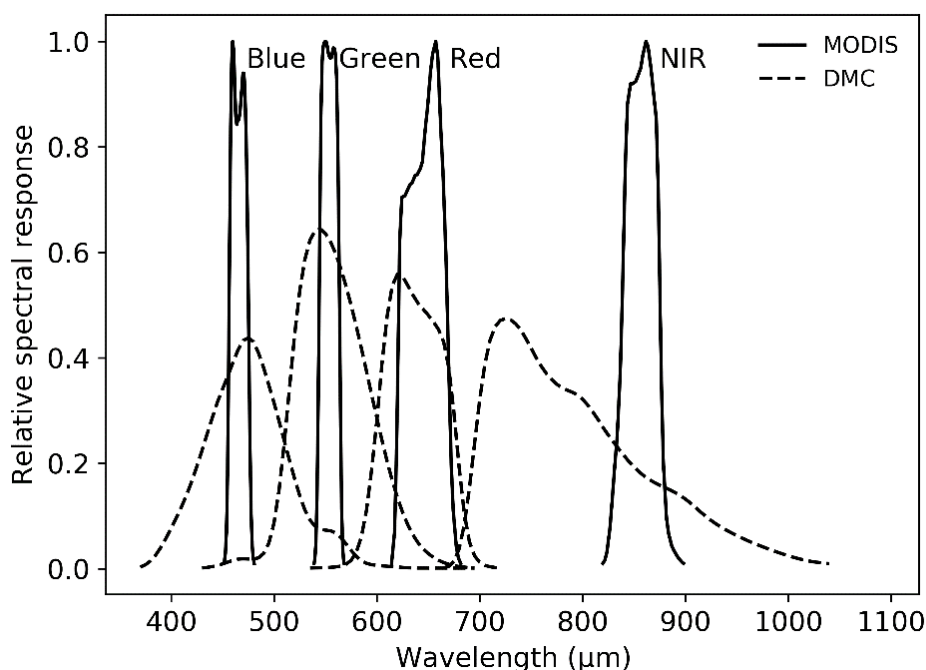


Figure 5.5 MODIS and DMC RSRs

## 5.4.2 Mapping methodology

The image resolution of 0.5 m, combined with the tendency of spekboom to grow in continuous stands, meant that there was little spectral mixing and that pixels covering spekboom were relatively pure. This supported a per-pixel classification approach to distinguish spekboom from the surrounding vegetation. The pixel-based approach also ensured that the complexities associated with segmentation could be avoided. The fractional canopy cover was determined as the portion of pixels classified as spekboom over an area of interest.

Given the large number of images, computation time was an important consideration in the formulation of our method. Radiometric homogenisation and classification software tools were developed using the GDAL (GDAL Development Team 2014) and OpenCV (Bradski 2000) software libraries. Careful consideration was given to computational efficiency in the selection of features and classification algorithm.

## 5.4.3 Features

A set of 46 features that would aid in describing the visual characteristics of spekboom were evaluated. The set included a typical combination of spectral features, vegetation indices and texture features. Similar features have been used in Li et al. (2010) and Trias-Sanz, Stamon & Louchet (2008). The features can be grouped into two broad categories: per-pixel and sliding window features. The per-pixel features are found with the spectral information from only one pixel, while the sliding window features are found from a statistic of the pixels inside a small local neighbourhood. While the spectral resolution of the VHR imagery is poor, the spatial resolution enables a description of the vegetation structure and spatial patterns, which is not possible with lower resolution satellite imagery. Texture features are a popular way of encapsulating spatial and structural information. Measures of vegetation texture are sensitive to shadow variations, an unavoidable phenomenon in aerial imagery caused by the long flight times and varying sun angle. Nevertheless, texture is recognised as an important feature in biomass estimation in complex habitats (Lu 2006; Singh, Malhi & Bhagwat 2014). The sliding window features were included to exploit this source of information when distinguishing the classes.

Although the imagery was calibrated to surface reflectance, it was done at a coarse spatial scale and fine resolution radiometric variations were not taken into account. A normalised colour space was consequently included in the features to reduce intensity variations not removed by the surface reflectance corrections. Colour is captured by the relative amounts of the raw colour bands rather than their absolute values. Normalised colour features are defined as (Blauensteiner et al. 2006):

$$c_i^n = \frac{C_i}{\sum_{i=1}^4 C_i} \quad \text{Equation 5.1}$$

where  $C_i$  are the raw R, G, B and NIR band values and  $i$  is the band number. The denominator normalises for intensity.

Green, living vegetation absorbs light in the photosynthetically active radiation region of the spectrum, which corresponds to the red band. There is a sharp transition from absorption to reflection around 700 nm (Gates 1980) which is captured by the DMC R and NIR wavelength ranges (see Figure 2.1). Vegetation is highly reflective in the near-infrared band as the energy in these wavelengths is insufficient for photosynthesis and potentially harmful due to its heating effects. Various vegetation indices exploit these spectral properties. The ratio vegetation index (RVI) is given by:

$$RVI = \frac{NIR}{R} \quad \text{Equation 5.2}$$

It has a range of zero to infinity and increases as the vegetation becomes denser and photosynthetically more active (Myneni et al. 1995). The well-known NDVI is defined as

$$NDVI = \frac{NIR - R}{NIR + R} \quad \text{Equation 5.3}$$

NDVI is limited to the range -1 to 1 and contains the same information as RVI, but is easier to visualise and interpret due to its limited range. Both indices are unaffected by intensity changes.

The tasselled cap transform is a linear transform of the raw band feature space to a new orthogonal co-ordinate system, similar to a principal component transform. It was designed for agricultural wheat classification and was intended to reduce variability in soil and wheat classes by removing variation due to topography, sun angle and wheat growth stage (Kauth & Thomas 1976). The tasselled cap transform was approximated in this study by using a principal component transform derived from the variance of the spekboom class. The first component was aligned with spekboom variation rather than wheat variation, as in the original tasselled cap transform. As it is simply a rotation of the raw band space, it is more useful as a dimensionality reduction technique (similar to principal components analysis (PCA)) than an extractor of novel features. The principal components of the normalised colours of Equation 5.1 were also included as features in the classification process.

Entropy is a statistic that describes the amount of randomness in a variable. It was included in our feature set as a texture feature to describe complexity in the local neighbourhood of a sliding window. The entropy of the values in the image window  $\mathbf{x}$  is defined as (Trias-Sanz, Stamon & Louchet 2008):

$$E(\mathbf{x}) = - \sum_i h_i(\mathbf{x}) \log_2 h_i(\mathbf{x}) \quad \text{Equation 5.4}$$

where  $h_i(\mathbf{x})$  is the probability in the  $i^{\text{th}}$  histogram bin of  $\mathbf{x}$ . A total of 256 bins were used in all cases.

Table 5.4 Features

No.	Name	Description
1	R	Red
2	G	Green
3	B	Blue
4	NIR	Near-infrared
5	rN	Normalised R
6	gN	Normalised G
7	bN	Normalised B
8	nirN	Normalised NIR
9	NDVI	Normalised difference vegetation index
10	RVI	Ratio vegetation index
11–14	tc1–4	Tasselled cap components
15–18	pc1–4	Principal components of raw bands
19–22	nc1–4	Principal components of normalised bands
23–26	Entropy##	Sliding window entropy of pc1, RVI, NDVI and gN
27–30	Std##	Sliding window standard deviation of pc1, RVI, NDVI and gN
31–34	Mean##	Sliding window mean of pc1, RVI, NDVI and gN
35–38	Median##	Sliding window median of pc1, RVI, NDVI and gN
39–42	Skewness##	Sliding window skewness of pc1, RVI, NDVI and gN
43–46	Kurtosis##	Sliding window kurtosis of pc1, RVI, NDVI and gN

In addition to the entropy, the median and the four central moment features (mean, standard deviation, skewness and kurtosis) of Li et al. (2010) were included as sliding window features. The first principal component, RVI, NDVI and normalised green channel were all used as inputs to the sliding window feature set. The complete feature set and their labels are listed in Table 5.4. A sliding window size of five was selected using a cross-validated grid search, with the accuracy of a naive Bayes classifier trained on the EntropyPc1 feature as the performance criterion. This size seemed sensible as it is comparable to that of a small spekboom clump.

#### 5.4.4 Feature selection

The bands of the imagery have significant spectral overlap (Intergraph 2008) and consequently are highly correlated. Given that the bands are the source data for all the derived features, the derived feature definitions also contain inter-dependencies. A number of authors have noted that feature redundancy can cause instability and sub-optimality in selected features when traditional approaches (such as ranking, FS and BE) are used (Sahu & Mishra 2011; Strobl et al. 2008; Tolosi & Lengauer 2011; Yousef et al. 2007). Redundancy can be reduced by using a feature extraction approach such as PCA, but requires computation of the full feature set and is not practical in computationally demanding applications such as ours.

A feature selection method, called feature clustering and ranking (FCR), was used to select relevant features in the presence of redundancy. The approach is based on Harris & Van Niekerk (2018) but uses hierarchical clustering instead of affinity propagation to group redundant features. It is described as follows:

1. Perform average-linkage hierarchical clustering (Székely & Rizzo 2005) of the feature set using the correlation coefficient as the dissimilarity metric.
2. Select a dissimilarity threshold at which to extract a natural number of clusters containing high correlation by visual inspection of the dendrogram.
3. Rank each cluster's importance by finding the value of a relevance criterion for each individual feature and then finding the median of the feature relevance values in the cluster.
4. Select a single feature from each of the  $N$  clusters with the best importance scores.

The number of clusters,  $N$ , was chosen using a grid search with the final classifier accuracy as performance measure. In this study, the accuracy of a naive Bayes classifier was used as the feature relevance criterion. The naive Bayes criterion makes no assumption about the form of the class distributions and can thus provide a generic measure of separability. It is simple, fast and recognised as being accurate for a variety of problems (Hand & Yu 2001). To avoid biased accuracy estimates, all classifier accuracy evaluations for feature relevance or selection of  $N$ , were done on unseen test data using a ten-fold cross-validation (Bishop 2003). The cluster-ranking method has the advantages of being quick and allowing hand-picking of the single features that represent each cluster. The flexibility to choose features enables the user to favour those features that are fastest to compute, or perhaps to choose those features that are more readily understood. The method was applied to the labelled pixel data.

### 5.4.5 Classification and spekboom canopy cover estimation

The decision tree, random forest, SVM, Bayes normal and k-nearest-neighbour (kNN) classifiers were evaluated in this study. A decision tree is a tree of binary decision nodes based on thresholds of different features. Data is recursively split at each branch node until a terminal representing a class label is reached (Breiman et al. 1984). Training is performed by a greedy procedure, which iteratively adds nodes and selects features producing the best split for each node. Criteria used for choosing the best feature at each node include the information content, node purity and Fisher's criterion (Jain, Duin & Mao 2000). Overtraining is a concern and trees can be pruned in a post-training step to reduce variance. Decision trees are known for their speed of execution and ease of interpretation. Node decisions can help provide insight into the problem. Decisions are usually binary and based on a single feature. As a result, the decision boundary is comprised of stepwise sections parallel to the feature axes and is at best an approximation of the optimal boundary (Jain, Duin & Mao 2000). Decision trees are flexible and broadly applied as they are non-parametric (i.e. they make no assumption about the form of class distributions) and can deal with categorical as well as continuous variables (Breiman et al. 1984).

Random forests are classifiers that use bootstrapped aggregation (bagging) (Breiman 1996) of a large collection of decision tree classifiers. Each tree is trained on a bootstrapped version of the dataset and the decision feature for each node is selected from a random subset of the full feature set (Breiman 2001). The bootstrapping and random feature subsets help introduce variation amongst the base tree classifiers. The uncorrelated decision trees, in combination, have greater predictive power than any single one. Importantly, a random forest is not prone to overtraining. Random forests are also robust to mislabelled training data. Both training and execution demand a moderate amount of computation time. The two main parameters for tuning a random forest are the number of trees and the number of features considered for each node.

Ground-breaking and widespread pattern recognition work has been done with the SVM (Amorós López et al. 2011; Li et al. 2010; Sahu & Mishra 2011; Yousef et al. 2007). The SVM was initially defined as the two-class linear decision boundary that maximised the distance to the nearest objects, called "support vectors" (Burges 1998). The decision boundary is determined only by the support vectors, not directly by features or generative descriptions of class distributions. The SVM minimises the Vapnik-Chervonenkis dimension, a measure of the complexity of the classifier. This is an important property of the SVM and explains how it effectively adapts its complexity to the data, is robust to overtraining and performs well in high dimensional feature spaces. The original formulation was extended to the case of overlapping multi-class problems using a penalty

term with user-defined multiplier  $C$ , which punishes class overlap. Using the kernel trick, the linear SVM was further extended to allow modelling of non-linear decision boundaries (Burges 1998). Different kernels such as polynomials or radial basis functions (RBF) may be chosen to suit the given problem. In kernel form, the SVM can be considered a non-parametric classifier. In our evaluation, an RBF kernel was used for the SVM classifier. The training is done by using a computationally demanding quadratic optimisation problem. However, execution is fast as it only requires an evaluation of the kernel function for the support vector – object vector pairs (Jain, Duin & Mao 2000).

The Bayes normal classifier, sometimes referred to as the maximum likelihood (ML) classifier, assumes that the classes are normally distributed. Mean and covariance parameters are estimated for each class from the data, usually with the ML criterion. Bayes' rule is then used to define the decision boundary (Duin & Tax 2005).

The kNN classifier labels test objects by finding the mode of classes of the closest  $k$  training objects (Bishop 2003). Any distance metric can be used for finding neighbours, but the Euclidean distance measure is prevalent and was used in our study. This classifier is a useful benchmark as it almost always performs reasonably well, requires only one parameter and is non-parametric (Jain, Duin & Mao 2000). It requires finding distances to the full training set, which can slow execution for large datasets.

User supplied tuning parameters for the classifiers were found with cross-validated grid searches. Table 5.5 details the parameter values selected for each classifier. Descriptions of the parameters can be found in the OpenCV documentation (OpenCV Development Team 2014).

Table 5.5 Classifier parameters

Classifier	Parameters
Decision tree	Maximum depth = 12, Use surrogates = false, Truncate pruned tree = true, Minimum sample count = 34, Priors = [0.33 0.33 0.33]
Random forest	Maximum number of trees = 5, Size of feature set = 4, Maximum tree depth = 10, Forest accuracy = 0.025, Priors = [0.2 0.4 0.2]
kNN	K = 5, Priors = [0.33 0.33 0.33]
SVM	SVM type = C Support vector classification, Kernel = RBF, Kernel width = 25, C = 1, Priors = [0.33 0.33 0.33]
Bayes normal	Priors = [0.33 0.33 0.33]

Morphological operators (Serra & Soille 1994) were applied as a post-processing step to the classifier produced maps to remove noise and smooth boundaries. Assuming that the majority of spekboom plants were big enough to cover more than one pixel, a morphological opening was applied to remove isolated spekboom pixels. Following this, spurious wrinkles and holes in the spekboom boundaries were removed with a morphological closing operation; the assumption



being that spekboom typically grows in solid clumps and any real gaps in these clumps would be more than a pixel wide. These operations can be seen as a way of further incorporating spatial context into the classification.

### 5.4.6 Evaluation

The per-pixel performance of the candidate classifiers on the selected features was evaluated with the labelled pixel data. To avoid biased estimates of performance, ten-fold cross-validation was used for classifier evaluation. The fractional spekboom canopy cover performance of the classifiers was tested on the in situ canopy cover data. After applying the classifiers and morphological operations to the relevant images, canopy cover estimates were extracted by evaluating the fractional portion of spekboom inside the areas of the field site polygons. These estimates were compared to the in situ canopy cover data. A flowchart illustrating the methodology for spekboom canopy cover estimation and evaluation is shown in Figure 5.6.

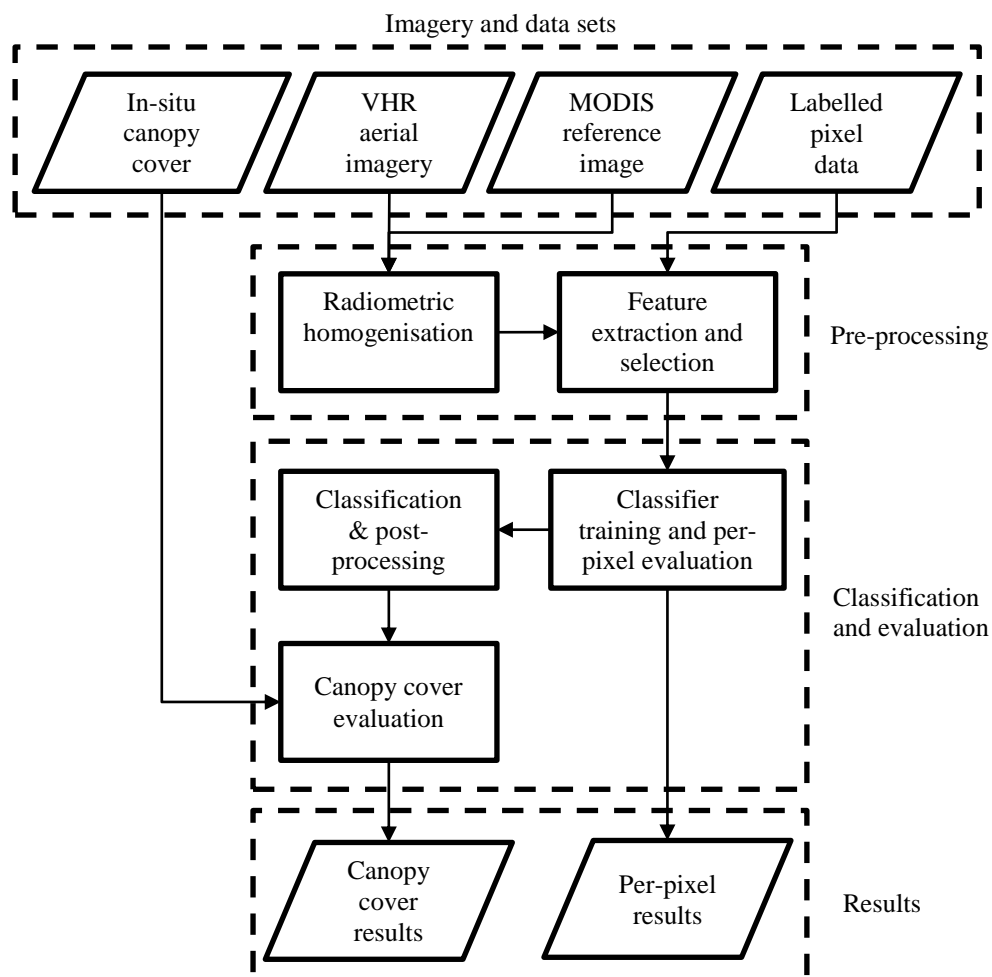


Figure 5.6 Methodological flowchart

## 5.5 RESULTS

### 5.5.1 Feature selection

The dendrogram showing the clustering of our feature set, is plotted in Figure 5.7. The dotted line shows the dissimilarity threshold value at which the feature clusters were extracted. This value was selected on the basis of being a relatively stable point in the hierarchy and being a point where the correlation amongst features is strong. Table 5.6 lists the clusters ordered by their importance, along with their component features. Using a grid search, it was found that a total of six features gave the best decision tree classification performance. The NDVI, pc1, EntropyPc1, gN, bN and nc2 features were selected from the top six clusters.

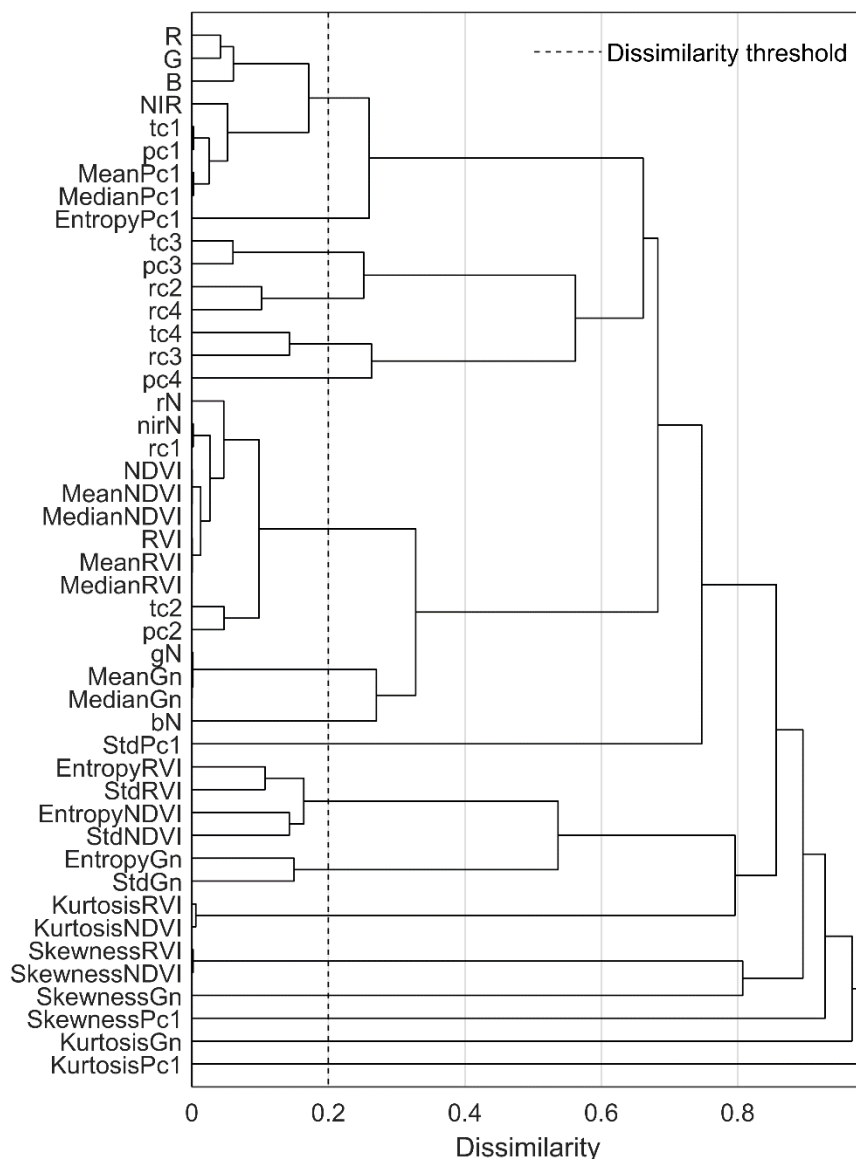


Figure 5.7 Clustering of correlated features

Table 5.6 Ranked clusters

Rank	Importance (%)	Features
1	68.27	rN, nirN, NDVI, RVI, tc2, pc2, nc1, MeanRVI, MedianRVI, MeanNDVI, MedianNDVI
2	61.38	R, G, B, NIR, tc1, pc1, MeanPc1, MedianPc1
3	60.41	EntropyPc1
4	55.23	gN, MeanGn, MedianGn
5	54.52	bN
6	53.57	nc2, nc4
7	50.57	tc4, nc3
8	49.34	pc4
9	47.93	EntropyRVI, StdRVI, EntropyNDVI, StdNDVI
10	43.96	StdPc1
11	43.62	EntropyGn, StdGn
12	42.65	tc3, pc3
13	41.29	SkewnessRVI, SkewnessNDVI
14	35.27	SkewnessGn
15	35.19	KurtosisRVI, KurtosisNDVI
16	35.03	SkewnessPc1
17	34.86	KurtosisGn
18	33.86	KurtosisPc1

### 5.5.2 Classification and spekboom canopy cover estimation

Table 5.7 compares the performance of the candidate classifiers. The table results are sorted according to the mean absolute canopy cover error (MAE) in the last column. Of the performance measures in the table, this is the only one evaluated against the in situ canopy cover data; the rest were evaluated against the labelled pixel data using cross-validation. Three- and two-class errors are reported as the class prior weighted errors i.e. the mean of the errors of omission. Cohen's Kappa and user's and producer's accuracies are given for the two-class case. Standard errors are given for all cross-validated performance measures that were evaluated on the labelled pixel data. The standard deviation of absolute errors (SAE) gives an indication of the variability in the spekboom canopy cover performance. Times taken for each classifier to process a single 12000 pixel  $\times$  8000 pixel image are listed in Table 5.8.

Table 5.7 Classifier performance comparison

Classifier	3 Class Error (%) <sup>a</sup>	2 Class Error (%) <sup>a</sup>	UA (Bg / Sb) <sup>a</sup>	PA (Bg / Sb) <sup>a</sup>	Kappa <sup>a</sup>	MAE (SAE) <sup>a</sup>
Decision tree	9.49 $\pm$ 0.47	3.59 $\pm$ 0.31	95.26 $\pm$ 0.44 / 98.05 $\pm$ 0.37	98.32 $\pm$ 0.32 / 94.50 $\pm$ 0.53	0.930 $\pm$ 0.006	5.85 (4.65)
Random forest	9.15 $\pm$ 0.95	2.69 $\pm$ 0.23	97.24 $\pm$ 0.25 / 97.45 $\pm$ 0.30	97.74 $\pm$ 0.27 / 96.89 $\pm$ 0.28	0.947 $\pm$ 0.005	7.09 (6.07)
kNN	10.28 $\pm$ 0.41	1.70 $\pm$ 0.13	98.95 $\pm$ 0.18 / 97.52 $\pm$ 0.25	97.05 $\pm$ 0.42 / 98.84 $\pm$ 0.20	0.965 $\pm$ 0.002	7.60 (6.20)
SVM	9.98 $\pm$ 1.51	2.47 $\pm$ 1.95	98.90 $\pm$ 0.21 / 96.08 $\pm$ 3.85	96.27 $\pm$ 3.90 / 98.80 $\pm$ 0.22	0.949 $\pm$ 0.041	7.99 (8.33)
Bayes normal	16.31 $\pm$ 0.86	9.03 $\pm$ 0.50	86.95 $\pm$ 0.62 / 98.09 $\pm$ 0.28	98.55 $\pm$ 0.21 / 83.38 $\pm$ 0.88	0.826 $\pm$ 0.010	8.08 (8.35)

<sup>a</sup>  $\pm$  = standard error of cross-validated performance measure, UA = User's accuracy (%), PA = Producer's accuracy (%), Bg = Background, Sb = Spekboom, MAE = Mean absolute canopy cover error on in situ canopy cover data (%), SAE = Standard deviation of absolute canopy cover errors on in situ canopy cover data (%)

Table 5.8 Classifier computation times

Classifier	Time (secs) <sup>a</sup>
Decision tree	47
Random forest	138
kNN	2067
SVM	788
Bayes normal	61

<sup>a</sup> computation time per 12000 pixel × 8000 pixel image

The decision tree three-class and two-class confusion matrices and performances, obtained from the labelled pixel data, are given in Table 5.9 and Table 5.10 respectively. The three-class confusion matrix shows that the tree class overlaps with both the spekboom and background classes, but that the overlap is larger with the background class.

Table 5.9 Decision tree three-class confusion matrix

	Background	Spekboom	Tree	Total	PA (%) <sup>a</sup>
<b>Background</b>	24776	330	2154	27260	90.89 ± 0.49
<b>Spekboom</b>	297	25762	1201	27260	94.50 ± 0.53
<b>Tree</b>	282	183	2892	3357	86.15 ± 1.33
<b>Total</b>	25355	26275	6247	57877	
<b>UA (%)<sup>a</sup></b>	97.72 ± 0.29	98.05 ± 0.37	46.34 ± 1.50		
<b>Kappa</b>	0.866 ± 0.007				
<b>Overall Error (%)</b>	9.49 ± 0.47				

<sup>a</sup> ± = standard error of cross-validated performance measure, UA = User's accuracy, PA = Producer's accuracy

Table 5.10 Decision tree two-class confusion matrix

	Background	Spekboom	Total	PA (%) <sup>a</sup>
<b>Background</b>	30104	513	30617	98.32 ± 0.32
<b>Spekboom</b>	1498	25762	27260	94.50 ± 0.53
<b>Total</b>	31602	26275	57877	
<b>UA (%)<sup>a</sup></b>	95.26 ± 0.44	98.05 ± 0.37		
<b>Kappa</b>	0.930 ± 0.006			
<b>Overall Error (%)</b>	3.59 ± 0.31			

<sup>a</sup> ± = standard error of cross-validated performance measure, UA = User's accuracy, PA = Producer's accuracy

Table 5.11 shows the spekboom canopy cover estimates obtained from the post-processed decision tree output for each of the in situ canopy cover sites. The mean of the absolute canopy cover error is 5.85%, with a standard deviation of 4.65%.

Table 5.11 Decision tree canopy cover estimates

Area	No.	Ground Truth (%)	Classifier (%)
Groenfontein	1	0.00	0.07
	2	4.00	0.47
	3	10.00	8.21
	4	25.00	17.44
Matjiesvlei	1a	6.00	7.21
	1b	22.50	31.37
	2	70.00	67.38
	3	85.00	73.12
	4	65.00	70.34
	5	37.50	35.95
	6	17.50	12.01
	7	15.00	25.74
8	2.00	5.42	
Rooiberg	1	20.00	6.03
	2	11.00	1.03
	3	0.00	0.00
Grootkop	1	22.50	8.05
	2	0.50	0.22
	3	42.50	34.38
	4	77.50	71.27
MAE (SAE) <sup>a</sup>		5.85% (4.65%)	

<sup>a</sup>MAE = Mean of absolute canopy cover errors (%), SAE = Standard deviation of absolute canopy cover errors (%)

The decision tree classifier was applied to the image mosaic of the study area to produce a spekboom canopy cover map that was morphologically post-processed. Figure 5.8 shows close-up examples of the resulting spekboom canopy cover map for each of the canopy cover ground truth areas (as described in Table 5.1). Spekboom boundaries, shown in blue, are overlaid on colour-infrared (CIR) aerial imagery. Figure 5.9 shows the final spekboom canopy cover map of the study area (the vegetation map of Vlok, Cowling and Wolf (Vlok, Cowling & Wolf 2005) was used to mask areas not supporting spekboom).

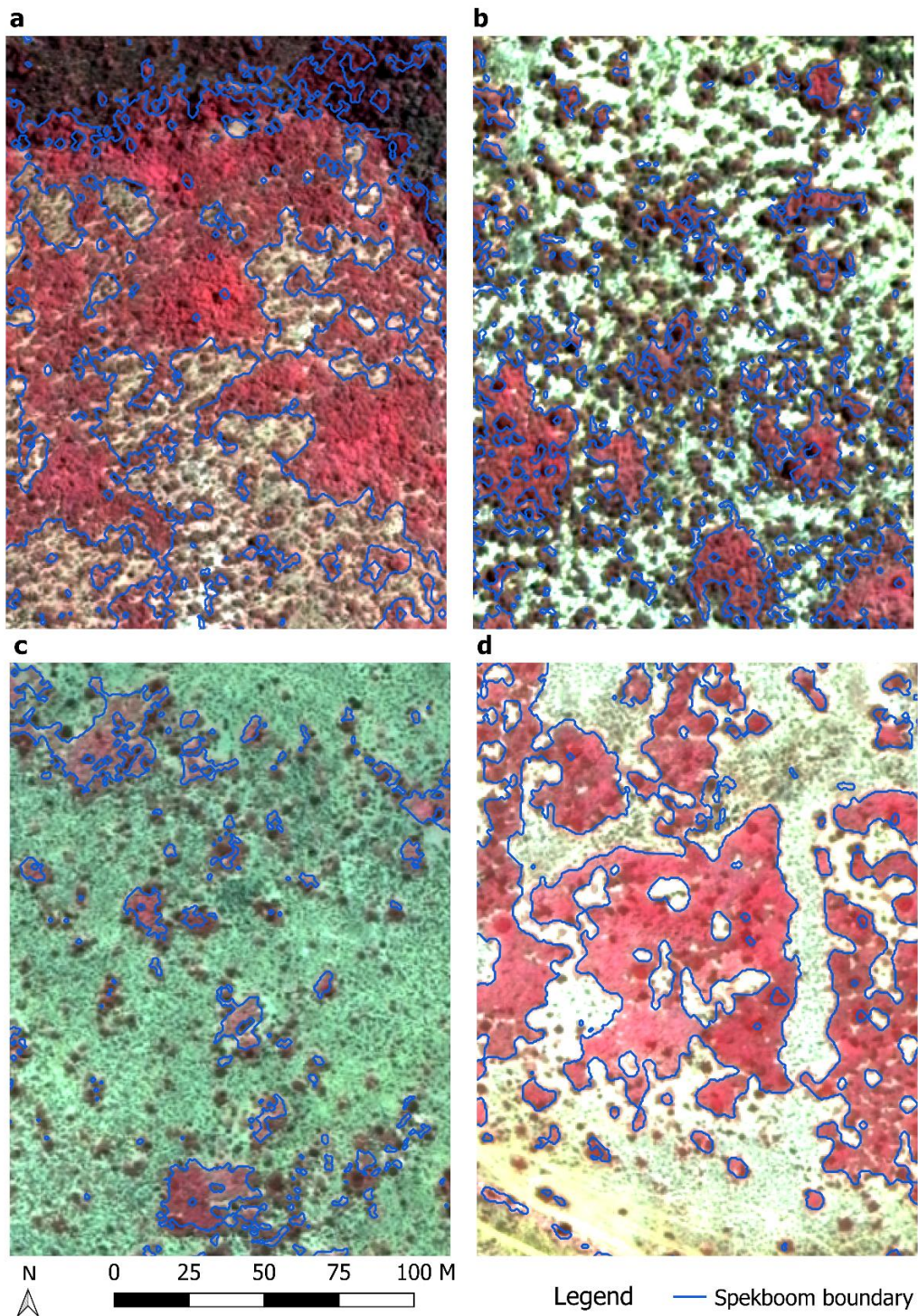


Figure 5.8 Example spekboom canopy cover maps showing classified spekboom boundaries (a) Matjiesvlei (Habitat: arid thicket with spekboom), (b) Groenfontein (Habitat: valley thicket with spekboom); (c) Rooiberg (Habitat: arid thicket with spekboom and fynbos mosaic); and (d) Grootkop (Habitat: arid thicket with spekboom and succulent Karoo mosaic)

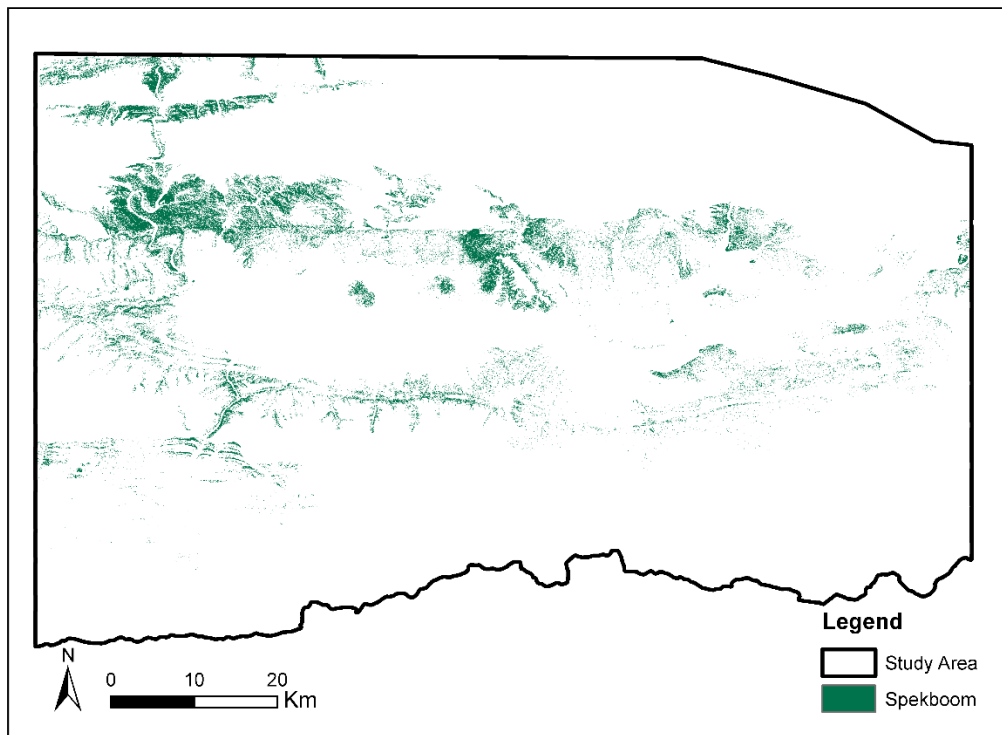


Figure 5.9 Global spekboom canopy cover map

## 5.6 DISCUSSION

### 5.6.1 Feature selection

Table 5.6 reveals a number of interesting properties of the features. First, it is clear that there is significant redundancy amongst the features. The correlation between the R, G, B and NIR bands is strong ( $>0.7$ ), likely due to strong coupling with intensity. The bands are consequently all grouped into a single cluster. While the definitions of the nirN, NDVI and RVI features are quite different, they are all describing the same spectral property of vegetation, namely high absorption in the red band and high reflectance in the NIR band. This is confirmed by their collection in the same cluster.

EntropyPc1 is ranked highly (third) in its own cluster, which supports the hypothesis that texture is an important property for mapping vegetation in VHR imagery. It is, however, the only texture feature in the best eight clusters. At the 0.5 m image resolution, texture will be descriptive of bush clumps more than individual spekboom plants. The bush clumps vary significantly in their composition and character with variation in habitat and level of degradation. We believe that the paucity of texture features in informative clusters is likely due to bush-clump and shadow variations.

The importance of bN was unexpected. The blue channel is particularly susceptible to haze effects and intuitively should not hold much discriminating power for vegetation. Inspecting bN images

shows an inversion of the topography shading seen in other channels. Sunlit northern slopes are dimmer and shaded southern slopes brighter. This occurs because the blue light in the shaded areas, which scatters more readily, is the dominant band of illumination. The contribution of bN is not fully understood but we believe its value lies in this property and that it helps to distinguish shaded vegetation from genuinely dark vegetation. In their tree mapping study, Key et al. (2001) also found the blue band to be valuable due to its insensitivity to shadowing issues.

The gN feature, its mean and its median form their own cluster. The mean sliding window feature, median sliding window feature and source feature operated on by those sliding windows are strongly correlated, as is expected.

The NDVI, pc1, EntropyPc1, gN, bN and nc2 features were selected from the top six clusters. Selection of sliding window features was avoided where possible as they are computationally more demanding than the per-pixel features. NDVI was selected from the first cluster simply because it is popular and easy to interpret. In the second cluster, pc1 was chosen as it is the first principal component of the raw bands and should therefore be more informative than any one of them in isolation. There is only one sliding window feature, EntropyPc1, in our final selection.

### **5.6.2 Classification and spekboom canopy cover estimation**

With the exception of the Bayes normal classifier, the classifiers' performance was good. The performances of the kNN and decision tree classifiers are as good as or better than the more complex SVM and random forest classifiers (see Table 5.7). The excellent performance of a diverse group of classifiers suggests that an informative feature set was selected. The notably poorer performance of the Bayes normal classifier implies that the classes are not normally distributed. The three-class errors are larger than the two-class errors due the tree class overlapping substantially with the background class. Errors due to tree samples being assigned to the background class, and vice versa, are negated when the tree class is lumped into the background class. The classifier performance variability on the labelled pixel data was low in general, as evidenced by the standard errors in Table 5.7, with the SVM producing the highest variability of the tested classifiers.

Of the performance measures in Table 5.7, the MAE is considered the most important for classifier comparison as it has the most direct relationship with actual spekboom canopy cover mapping accuracy over the study area. Taking the MAE and image ground truth performance into account, the decision tree was selected as the final classifier. It has the best canopy cover performance and



is the fastest option (see Table 5.8). While it is one of the poorer performers on the labelled pixel data, it is still very accurate when applied to this data.

The classifier performed well in the Groenfontein, Matjiesvlei and Grootkop areas, but underestimated spekboom canopy cover in all the Rooiberg sites. As a result of the sandstone/quartzite geology of the area, the spekboom plants at Rooiberg are smaller and have a canopy that is less dense than those in other sites. We believe this partially explains the canopy cover underestimation in this area.

A visual inspection of the spekboom canopy cover map revealed some spatial variation over the study area. Figure 5.8 shows close-up canopy cover map examples for each of the canopy cover ground truth areas (as described in Table 5.1). Arid areas, such as Rooiberg, seem more prone to underestimation, probably due to spectral mixing occurring with bare ground around the canopy borders and also due to the smaller and less dense stands occurring in these areas. Conversely, there tends to be a slight overestimation in more densely vegetated areas, likely the result of confusion due to spectral mixing with other green vegetation. In general, however, the spekboom canopy cover map of the study area appears accurate.

This study is one of few examples of vegetation mapping using VHR imagery over a large area (Eisfelder, Kuenzer & Dech 2012; Lu 2006). While the mapping accuracies achieved compare well with related studies (Baraldi et al. 2010; Ghosh & Joshi 2014; Heumann 2011; Johansen et al. 2007; Mehner et al. 2004), there are possible avenues for improvement. Ancillary information such as elevation, slope aspect or habitat could be incorporated into the classifier, similarly to Thompson et al. (2009) and Su (2009). This could be done by including it either as a feature or by designing separate classifiers for different ranges or categories of the ancillary variable.

## **5.7 CONCLUSIONS**

Accurate spekboom canopy cover estimates were obtained across the study area using a per-pixel classification approach. Homogenisation to surface reflectance by calibration with satellite data provided radiometric consistency and allowed application of a single classification algorithm over an extended area. Six distinguishing features of spekboom, consisting of a combination of spectral, textural and vegetation index type measures, were selected using a FCR method. Out of a set of candidate classifiers, a decision tree produced the best spekboom canopy cover accuracy, and was subsequently used to produce a map of the study area. A MAE of 5.85% over 20 ground truth sites was achieved.

Species distinguishing traits can be subtle, subject to variation and not easily captured with coarse spectral resolution imagery such as the NGI multi-spectral aerial imagery (Fassnacht et al. 2016). In this study, the distinctive lime-green colour and dense growth pattern of spekboom likely contributed to the successful separation of spekboom from background vegetation. While some variation in the canopy cover accuracy was observed over different habitats, the classifier's general performance was consistent over the study area. By incorporating ground truth from new areas, the techniques used to produce this map could be applied to the rest of the thicket biome. The availability of a spekboom canopy cover mapping technique will be a valuable starting point for developing measures of other environmental variables such as biomass and biodiversity (Eisfelder, Kuenzer & Dech 2012; Ludwig, Reynolds & Whitson 1975; Suganuma et al. 2006).

## CHAPTER 6: DISCUSSION AND CONCLUSIONS

### 6.1 SUMMARY

Spekboom is a dominant and ecologically important species in subtropical thicket. Its presence improves soil health, sequesters carbon and creates a suitable environment for the germination of other plants. It is, however, highly sensitive to browsing by domestic livestock and is one of the first species to disappear under sustained browsing pressure. Vast areas of the subtropical thicket biome have been degraded, mostly by goat over-browsing. In response, wide-scale planting of spekboom has been identified as a viable and cost-effective means of restoring degraded thicket ecosystems. Regional spekboom canopy cover maps are needed to assist with the planning and monitoring of the restoration process. This research set out to develop a regional, cost-effective spekboom canopy cover mapping technique using multi-spectral VHR aerial imagery.

A literature review of remote sensing, image classification and vegetation species mapping was provided in Chapter 2 to address Objective 1. The literature review identified a number of remotely sensed data sources suitable for vegetation species mapping. Of these, multi-spectral VHR aerial imagery was selected as a viable source for spekboom mapping that met constraints of the application (i.e. regional coverage, high spatial resolution and low cost). Challenges associated with regional vegetation species mapping that had not been adequately considered by previous work were highlighted. These challenges included the reduction of unwanted radiometric variation, generalised classification across heterogeneous ecologies and the computationally efficient processing of thousands of images. Another limitation of the reviewed species mapping methods was a lack of attention to the problem of sub-optimal feature selection resulting from feature space redundancy.

Chapter 3 presented the development of a radiometric homogenisation technique to alleviate the problem of unwanted radiometric variation in VHR aerial imagery (Objective 3). The technique adjusts images to approximate surface reflectance by fusion with a collocated and concurrent satellite surface reflectance image. A spatially varying linear model is used to model the relationship between the aerial image digital numbers (DNs) and satellite surface reflectance values. A set of 2228 0.5 m resolution multi-spectral aerial images of the study area, captured over multiple days, was obtained from Chief Directorate: National Geo-Spatial Information (NGI). These images contained substantial radiometric variations due to spatially and temporally varying atmospheric and BRDF effects. The homogenisation technique was applied to the aerial imagery using a MODIS nadir BRDF-adjusted reflectance (NBAR) product as surface reflectance

reference. A visual inspection of the homogenised mosaic confirmed that seam lines and BRDF effects that were apparent in the raw imagery had been removed. Performance was quantitatively measured by comparing the homogenised mosaic to a collocated and concurrent SPOT 5 surface reflectance image. Mean  $R^2$  and mean absolute difference (MAD) values of this comparison were 0.84 and 3.43% respectively. Overall, results compared well to related methods. The acquisition of the aerial, MODIS and SPOT 5 images partially addressed Objective 2.

Feature space redundancy is known to lead to sub-optimal feature selection. In addition, redundancy can make selected features unstable and sensitive to small changes in the data. The feature clustering and ranking (FCR) method overviewed in Chapter 4 was developed to address these problems (Objective 4). A secondary motivation for the development of FCR was to allow factors of feature computation and measurement cost to be included in the selection criteria. FCR uses affinity propagation to form a moderate number of clusters of correlated and similarly relevant features. The clusters are ranked by their importance and an individual feature is selected from each of the best  $N$  clusters. Features may be selected automatically based on their relevance or interactively, taking criteria such as computation time and measurement cost into account. Redundancy is effectively discarded by choosing a single feature from each cluster of correlated features.

Chapter 4 presented a comparison of the performance of the FCR method and popular feature selection methods on five remote sensing data sets and one synthetic data set (all containing redundancy). Results highlighted poor accuracy and stability performance of ranking type feature selection. Forward selection (FS) methods performed relatively well in terms of accuracy but also suffered from instability. The comparison demonstrated the effectiveness of the FCR approach for selecting stable and informative features from redundant spaces. Overall, FCR ranked highest in combined accuracy and stability measures. FCR provided substantially better stability and slightly better accuracy compared to competing approaches. The feature selection comparison addressed Objective 5 and the acquisition of the comparison data sets further contributed to Objective 2.

A per-pixel classification method for the regional mapping of spekboom canopy cover was developed and tested in Chapter 5 (Objective 6). The radiometrically homogenised multi-spectral aerial imagery produced in Chapter 3 was used to provide discriminating features. Classifier training data was produced by delineation and labelling of image areas. In situ canopy cover data, encompassing regional habitat and degradation variation, was gathered for purposes of map

validation. The acquisition of the training and validation data completed the requirements of Objective 2.

Common spectral, textural and vegetation index features were generated from the imagery. The FCR procedure of Chapter 4 was used to select six informative features from the generated set. FCR allowed computationally expensive features to be excluded where closely correlated alternatives existed. Five candidate classifiers (ML, decision tree, kNN, random forest and SVM) were trained on the labelled data. Per-pixel classifier outputs were post-processed with morphological operations to remove noise and smooth boundaries. The post-processing step served as a rudimentary means to further incorporate spatial context into the spekboom canopy cover map. Of the candidate classifiers, the decision tree produced the best spekboom canopy cover accuracy, with a mean absolute error (MAE) of 5.85%. Somewhat fortuitously, this classifier also had the fastest execution time of those investigated.

## **6.2 CRITICAL EVALUATION**

The radiometric homogenisation method developed in this study was effective at calibrating a large and radiometrically diverse set of imagery. Unlike other radiometric correction techniques, this method does not require manual field reflectance measurements (López et al. 2011), placement of targets of known reflectance (Collings et al. 2011) or detailed knowledge of atmospheric conditions at the time of capture (Berk et al. 1999; Richter 1997; Vermote et al. 1997). Adjustment of problematic BRDF variations is made implicitly. In contrast to existing mosaic correction methods, which adopt a uniform per-image BRDF model (Chandelier & Martinoty 2009; Collings et al. 2011; López et al. 2011), radiometric homogenisation can adjust for local BRDF variations within each image. The method can potentially be applied to any VHR imagery, including satellite and UAV imagery. It is particularly suited to regional mapping, where radiometric uncertainty is often detrimental to classification accuracy. Currently, the use of national coverage aerial imagery such as the NGI archive is often limited to visual interpretation (Campbell & Wynne 2011; López et al. 2011; Tempfli et al. 2009). Radiometric homogenisation offers a relatively simple means of extending the use of this valuable data to quantitative remote sensing applications. The method is, however, not without its limitations. The accuracy of the homogenised imagery is influenced by the accuracy of the reference surface reflectance data and radiometric variations occurring below the scale of the reference resolution cannot be accommodated. The method is also limited by the need for a reference image that is concurrent and spectrally similar to the aerial imagery. Results suggested that a lack of spectral similarity has a negative effect on accuracy.

FCR provides computationally efficient feature selection that is robust to redundancy typically found in high dimensional remote sensing data. The ability to include selection factors other than relevance (e.g. measurement and computation cost) distinguishes FCR from other methods (Brown et al. 2012; Jain, Duin & Mao 2000; Wu et al. 2013) and lends its use to computationally demanding or budget-limited applications. FCR provides a structured way of reducing these costs without affecting feature relevance. The consideration of other factors is, however, limited to a manual procedure, requiring user intervention. Another limitation arises from the clustering algorithm, which depends on the specification of a “preference” parameter to determine the number of clusters. While an automated specification of this parameter produced satisfactory results for the feature selection comparison, user adjustment of the “preference” parameter could give superior performance in some cases. The results of feature selection comparison experiments imply that best-practice feature selection should avoid the use of ranking methods, as they are particularly sensitive to data redundancy. FS methods produced comparatively good accuracy performance and could be used as a benchmark algorithm where feature stability is not a concern.

Unwanted radiometric variation, classifier transferability and computational efficiency are challenges for regional vegetation mapping (Basu et al. 2015; Neigh et al. 2018), which have received little attention in the literature (Fassnacht et al. 2016). The spekboom canopy cover mapping method developed in this study is distinguished by its regional applicability. In vegetation mapping, radiometric variation is sometimes ignored (Boyden et al. 2007) or dealt with by conversion to top of atmosphere radiance (Johansen et al. 2007; Mehner et al. 2004; Neigh et al. 2018; Niphadkar et al. 2017; Tarantino et al. 2019), which is still subject to atmospheric and BRDF effects. In this study, radiometric homogenisation sufficiently reduced spatially varying atmospheric and BRDF effects in the selected aerial imagery and allowed application of a single classifier over a large heterogeneous area. FCR provided stable informative features and improved the processing time of a large set of imagery by allowing computationally expensive (sliding window) features to be excluded.

The spekboom classifier exhibited some variation in performance between different habitats, with a 4.65% standard deviation of absolute errors (SAE). Generated maps revealed a tendency to overestimate spekboom canopy cover in densely vegetated areas and underestimate in sparsely vegetated areas. This was likely due to spectral mixing between spekboom and surrounding vegetation or bare ground. Despite this variation, useful canopy cover accuracies were obtained over diverse conditions, demonstrating the potential of the method for application to larger areas. Results of this study suggest that per-pixel approaches should not be disregarded for regional

mapping, confirming similar conclusions in Baldeck et al. (2015) and Basu et al. (2015). It is possible to incorporate spatial context into the per-pixel method with texture features and the post-processing of classifier output maps. The simplicity and computational efficiency of the per-pixel approach make it an attractive option for applications requiring automated processing of thousands of images.

There is a pressing need for large-area vegetation maps to inform responses to national and global issues such as climate change. One of the principal motivations for remote sensing is that it can provide cost-effective analyses over far greater spatial extents than field methods. This potential is, however, yet to be fully realised with VHR imagery where the majority of vegetation mapping studies are focused on small areas. VHR imagery provides discriminating spatial information not available in medium resolution imagery and is needed for plant scale vegetation mapping. This research adds to the important body of work addressing regional VHR vegetation mapping and its associated challenges.

### **6.3 RECOMMENDATIONS**

Local terrain effects are poorly represented at the 500 m resolution of the MODIS NBAR reference data used in this study and additional radiometric homogenisation experiments using alternate sources of surface reflectance data are recommended. Higher spatial resolution products such as Landsat, Sentinel-2 and MISR could improve fine-scale calibrations. While Landsat and Sentinel-2 surface reflectance products are not BRDF corrected, they are subject to significantly less BRDF variation than aerial imagery due to their narrower field of view (FOV). Consequently, they may produce improved fine resolution performance compared to the MODIS surface reflectance reference. MISR imagery provides a slight improvement (275 m) on the MODIS spatial resolution, with the benefit of rigorous BRDF correction. In addition to the use of alternate reference images, it is recommended that the effectiveness of including the offset parameter,  $C$  (which was excluded in the Chapter 3 experiments), is tested. The offset term may improve homogenised results where atmospheric effects are relatively severe.

Some means of automating the inclusion of computation and measurement costs into the selection criteria of FCR would be beneficial. This could be implemented as a weighted sum of criteria or by a combined ranking such as the non-dominated ranking that was used for comparing performance in Section 4.3.4. The correlation coefficient dissimilarity metric used for defining clusters is only capable of detecting linear associations between features. It would be worthwhile testing the use of metrics that can describe non-linear correlation, such as mutual information or

distance correlation (Székely, Rizzo & Bakirov 2007). While these metrics require increased computation, they could assist in further removal of redundancy due to their increased sensitivity.

The spekboom canopy cover map was produced using a simple classification approach and a number of opportunities exist for improving performance. Investigating a spatially stratified approach as a means of reducing performance variation due to ecological heterogeneity is suggested. Training separate classifiers for different strata (e.g. the habitat types delineated in the map of Vlok, Cowling & Wolf (2005)) could reduce ecological variation per stratum and allow classifiers to target stratum specific spekboom traits. Spekboom favours sunnier north-facing slopes (Vlok, Cowling & Wolf 2005) and inclusion of a slope aspect feature would likely improve general separation of spekboom from background vegetation. In their national tree mapping study, Basu et al. (2015) used a post-processing step that applied a conditional random field (CRF) (i.e. a supervised probabilistic technique for considering local spatial relationships) to per-pixel classifier outputs. This automated and computationally efficient method provided a valuable improvement to per-pixel accuracy. While the spekboom mapping method used texture features and morphological post-processing to exploit spatial information, a more sophisticated incorporation of spatial context through the post-processing technique of Basu et al. (2015) could also benefit mapping performance. Finally, it is recommended that spatial and temporal extensions of the method be tested through the acquisition of further imagery and ground truth.

## **6.4 CONCLUSIONS**

The planting of spekboom has been identified as an ecologically and operationally viable means of restoring vast areas of degraded subtropical thicket. Spatial information, including spekboom canopy cover maps, are needed for the effective co-ordination of biome-wide restoration.

This research developed a method for the semi-automated mapping of spekboom canopy cover using VHR multi-spectral aerial imagery and made contributions to the field of regional vegetation mapping. Novel techniques for radiometric homogenisation and feature selection were developed to assist with related components of the spekboom mapping problem.

The successful demonstration of the spekboom mapping method on a large heterogeneous study area provided a foundation for its extension to the thicket biome. The techniques developed in this research also present a starting point for the mapping of related thicket vegetation characteristics such as biomass. There is an urgent need for large-scale thicket restoration, which has myriad benefits, including the increasingly critical mitigation of climate change. The availability of spekboom canopy cover maps will greatly facilitate this important work.



## REFERENCES

- Adie H & Yeaton RI 2013. Regeneration dynamics in arid subtropical thicket, South Africa. *South African Journal of Botany* 88: 80-85.
- Adjorlolo C, Mutanga O, Cho M & Ismail R 2012. Challenges and opportunities in the use of remote sensing for C3 and C4 grass species discrimination and mapping. *African Journal of Range & Forage Science* 29, 2: 47-61.
- Akhmanov SA & Nikitin SY 1997. *Physical optics*. Oxford: Clarendon Press.
- Alonzo M, Bookhagen B & Roberts DA 2014. Urban tree species mapping using hyperspectral and LiDAR data fusion. *Remote Sensing of Environment* 148: 70-83.
- Amorós López J, Izquierdo Verdiguier E, Gómez Chova L, Muñoz Marí J, Rodríguez Barreiro JZ, Camps Valls G & Calpe Maravilla J 2011. Land cover classification of VHR airborne images for citrus grove identification. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 1: 115-123.
- Asner GP 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment* 64, 3: 234-253.
- Asner GP, Knapp DE, Kennedy-Bowdoin T, Jones MO, Martin RE, Boardman J & Hughes RF 2008. Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sensing of Environment* 112, 5: 1942-1955.
- Asner GP & Martin RE 2008. Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment* 112, 10: 3958-3970.
- Baatz M & Schäpe A 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In Strobl J, Blaschke T & Griesebner G (eds) *Angewandte Geographische Informations - Verarbeitung*, 12-23. Heidelberg: Wichmann Verlag.
- Baldeck CA, Asner GP, Martin RE, Anderson CB, Knapp DE, Kellner JR & Wright SJ 2015. Operational tree species mapping in a diverse tropical forest with airborne imaging spectroscopy. *PLoS ONE* 10, 7.
- Baldrige AM, Hook SJ, Grove CI & Rivera G 2009. The ASTER spectral library version 2.0. *Remote Sensing of Environment* 113, 4: 711-715.
- Baraldi A, Durieux L, Simonetti D, Conchedda G, Holecz F & Blonda P 2010. Automatic spectral-rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS,

- AVHRR/MSG, AATSR, IKONOS/QuickBird/OrbView/GeoEye, and DMC/SPOT-1/-2 imagery–Part I: System design and implementation. *IEEE Transactions on Geoscience and Remote Sensing* 48, 3: 1299-1325.
- Barbier N, Proisy C, Véga C, Sabatier D & Coueron P 2011. Bidirectional texture function of high resolution optical images of tropical forest: An approach using LiDAR hillshade simulations. *Remote Sensing of Environment* 115, 1: 167-179.
- Basu S, Ganguly S, Nemani RR, Mukhopadhyay S, Zhang G, Milesi C, Michaelis A, Votava P, Dubayah R, Duncanson L, Cook B, Yu Y, Saatchi S, DiBiano R, Karki M, Boyda E, Kumar U & Li S 2015. A semiautomated probabilistic framework for tree-cover delineation from 1-m NAIP imagery using a high-performance computing architecture. *IEEE Transactions on Geoscience and Remote Sensing* 53, 10: 5690-5708.
- Berk A, Anderson GP, Bernstein LS, Acharya PK, Dothe H, Matthew MW, Adler-Golden SM, Chetwynd, Jr. JH, Richtsmeier SC, Pukall B, Allred CL, Jeong LS & Hoke ML 1999. *MODTRAN4 radiative transfer modeling for atmospheric correction*. Proceedings of SPIE - The International Society for Optical Engineering. Denver: 348.
- Bishop CM 2003. *Neural networks for pattern recognition*. New York: Oxford University Press.
- Blaschke T 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 1: 2-16.
- Blaschke T, Hay GJ, Kelly M, Lang S, Hofmann P, Addink E, Queiroz Feitosa R, Van der Meer F, Van der Werff H, Van Coillie F & Tiede D 2014. Geographic object-based image analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing* 87: 180-191.
- Blauensteiner P, Wildenauer H, Hanbury A & Kampel M 2006. *On colour spaces for change detection and shadow suppression*. Proceedings of Computer Vision Winter Workshop 2006. Tele: 1-6.
- Boyden J, Walden D, Bartolo R, Bayliss P, Division SS & Territory N 2007. *Utility of VHR remote sensing data for landscape scale assessment of the environmental weed Para Grass (Urochloa mutica (Forssk.) T.Q. Nguyen) on a tropical floodplain*. Proceedings of 28th Asian Conference on Remote Sensing 2007. Kuala Lumpur: 551-556.
- Bradley BA 2014. Remote detection of invasive plants: a review of spectral, textural and phenological approaches. *Biological Invasions* 16, 7: 1411-1425.
- Bradski G 2000. The OpenCV library. *Dr. Dobb's Journal of Software Tools* 120: 122-125.

- Breiman L 1996. Bagging predictors. *Machine Learning* 24, 2: 123-140.
- Breiman L 2001. Random Forests. *Machine Learning* 45, 1: 5-32.
- Breiman L, Friedman J, Olshen R & Stone C 1984. *Classification and regression trees*. Belmont: Wadsworth.
- Brown G, Pocock A, Zhao M-J & Lujan M 2012. Conditional likelihood maximisation: A unifying framework for mutual information feature selection. *Journal of Machine Learning Research* 13: 27-66.
- Burges CJC 1998. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2, 2: 121-167.
- Campbell JB & Wynne RH 2011. *Introduction to remote sensing*. New York: The Guilford Press.
- Carleer AP, Debeir O & Wolff E 2005. Assessment of very high spatial resolution satellite image segmentations. *Photogrammetric Engineering & Remote Sensing* 71, 11: 1285-1294.
- De Castro AI, Jurado-Expósito M, Gómez-Casero MT, Gómez-Candón D, Caballero-Novella JJ & López-Granados F 2011. *Evaluation of aerial and quickbird images for mapping cruciferous weeds*. Precision Agriculture 2011 – Papers Presented at the 8th European Conference on Precision Agriculture 2011, ECPA 2011: 245-255.
- De Castro AI, Jurado-Expósito M, Peña-Barragán JM & López-Granados F 2012. Airborne multi-spectral imagery for mapping cruciferous weeds in cereal and legume crops. *Precision Agriculture* 13, 3: 302-321.
- Chandelier L & Martinoty G 2009. A radiometric aerial triangulation for the equalization of digital aerial images and orthoimages. *Photogrammetric Engineering & Remote Sensing* 75, 2: 193-200.
- Chander G, Meyer DJ & Helder DL 2004. Cross calibration of the Landsat-7 ETM+ and EO-1 ALI sensor. *IEEE Transactions on Geoscience and Remote Sensing* 42, 12: 2821-2831.
- Charalampidis D & Kasparis T 2002. Wavelet-based rotational invariant roughness features for texture classification and segmentation. *IEEE Transactions on Image Processing* 11, 8: 825-837.
- Chen X & Gu Y 2015. Class-specific feature selection with local geometric structure and discriminative information based on sparse similar samples. *IEEE Geoscience and Remote Sensing Letters* 12, 7: 1392-1396.
- Chen X, Zhou G, Chen Y, Shao G & Gu Y 2017. Supervised multiview feature selection exploring

- homogeneity and heterogeneity with L1,2-norm and automatic view generation. *IEEE Transactions on Geoscience and Remote Sensing* 55, 4: 2074-2088.
- Cheng G & Han J 2016. A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing* 117: 11-28.
- Chi M, Plaza A, Benediktsson JA, Sun Z, Shen J & Zhu Y 2016. Big data for remote sensing: challenges and opportunities. *Proceedings of the IEEE* 104, 11: 2207-2219.
- Choi SK, Lee SK, Jung SH, Choi JW, Choi DY & Chun SJ 2016. Estimation of fractional vegetation cover in sand dunes using multi-spectral images from fixed-wing UAV. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography* 34, 4: 431-441.
- Chuvieco E & Huete A 2009. *Fundamentals of satellite remote sensing*. Boca Raton: CRC Press.
- Clark ML & Roberts DA 2012. Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. *Remote Sensing* 4, 6: 1820-1855.
- Clark ML, Roberts DA & Clark DB 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment* 96, 3-4: 375-398.
- Clarke C, Shackleton S & Powell M 2012. Climate change perceptions, drought responses and views on carbon farming amongst commercial livestock and game farmers in the semiarid Great Fish River Valley, Eastern Cape province, South Africa. *African Journal of Range & Forage Science* 29: 13-23.
- Collings S, Caccetta P, Campbell N & Wu X 2011. Empirical models for radiometric calibration of digital aerial frame mosaics. *IEEE Transactions on Geoscience and Remote Sensing* 49, 7: 2573-2588.
- Coops NC, Hilker T, Wulder MA, St-Onge B, Newnham G, Siggins A & Trofymow JA 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees* 21, 3: 295-310.
- Cowling RM & Mills AJ 2011. A preliminary assessment of rain throughfall beneath *Portulacaria afra* canopy in subtropical thicket and its implications for soil carbon stocks. *South African Journal of Botany* 77, 1: 236-240.
- Cowling RM, Proche-Olivero J & Vlok JHJ 2005. On the origin of southern African subtropical thicket vegetation. *South African Journal of Botany* 71, 1: 1-23.
- Cukur H, Binol H, Uslu FS, Kalayci Y & Bal A 2015. *Cross correlation based clustering for*

- feature selection in hyperspectral imagery*. Proceedings of 9th International Conference on Electrical and Electronics Engineering (ELECO). Bursa: 232-236.
- Curran P, Smedley D, Thompson P & Knight AT 2012. Mapping restoration opportunity for collaborating with land managers in a carbon credit-funded restoration program in the Makana municipality, Eastern Cape, South Africa. *Restoration Ecology* 20, 1: 56-64.
- Downey M, Uebbing R, Gehrke S & Beisl U 2010. *Radiometric processing of ADS imagery: Using atmospheric and BRDF corrections in production*. Proceeding of ASPRS 2010 Annual Conference. San Diego: 1-8.
- Duin RPW & Tax DMJ 2005. Statistical pattern recognition. In Chen C & Wang P (eds) *Handbook of Pattern Recognition and Computer Vision*, 1-21. Singapore: World Scientific.
- Einzmann K, Ng W-T, Immitzer M, Pinnel N & Atzberger C 2014. Method analysis for collecting and processing in-situ hyperspectral needle reflectance data for monitoring Norway Spruce. *Photogrammetrie - Fernerkundung - Geoinformation* 2014, 5: 423-434.
- Eisfelder C, Kuenzer C & Dech S 2012. Derivation of biomass information for semi-arid areas using remote-sensing data. *International Journal of Remote Sensing* 33, 9: 2937-2984.
- ESA 2015. Sentinel-2 User Handbook [online]. Available from: [https://sentinel.esa.int/documents/247904/685211/Sentinel-2\\_User\\_Handbook](https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook) [Accessed 27 July 2018].
- Fassnacht FE, Latifi H, Stereńczak K, Modzelewska A, Lefsky M, Waser LT, Straub C & Ghosh A 2016. Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment* 186: 64-87.
- Fassnacht FE, Neumann C, Forster M, Buddenbaum H, Ghosh A, Clasen A, Joshi PK & Koch B 2014. Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7, 6: 2547-2561.
- Frey BJ & Dueck D 2007. Clustering by passing messages between data points. *Science* 315, 5814: 972-976.
- Fukunaga K & Hayes RR 1989. The reduced Parzen classifier. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11, 4: 423-425.
- Ganivet E & Bloomberg M 2019. Towards rapid assessments of tree species diversity and structure in fragmented tropical forests: A review of perspectives offered by remotely-sensed and field-

- based data. *Forest Ecology and Management* 432, August 2018: 40-53.
- Gao C, Jiang X, Li X & Li X 2013. The cross-calibration of CBERS-02B/CCD visible-near infrared channels with Terra/MODIS channels. *International Journal of Remote Sensing* 34, 9-10: 3688-3698.
- Gates DM 1980. *Biophysical ecology*. New York: Springer.
- GDAL Development Team 2014. Geospatial data abstraction library [online]. Available from: <http://www.gdal.org/> [Accessed 6 June 2014].
- Gehrke S 2010. *Radiometric processing of ADS imagery: Mosaicking of large image blocks*. Proceedings of Opportunities for Emerging Geospatial Technologies. San Diego: 184-195.
- Gehrke S & Beshah BT 2016. *Radiometric normalization of large airborne image data sets acquired by different sensor types*. Proceedings of 2016 XXIII ISPRS Congress. Prague: 317-326.
- Ghosh A & Joshi PK 2014. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. *International Journal of Applied Earth Observation and Geoinformation* 26: 298-311.
- GIC 2014. Hyperspectral remote sensing scenes [online]. Available from: [http://www.ehu.es/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes](http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes) [Accessed 29 September 2016].
- González-Roglich M & Swenson JJ 2016. Tree cover and carbon mapping of Argentine savannas: Scaling from field to region. *Remote Sensing of Environment* 172: 139-147.
- Gui J, Sun Z, Ji S, Tao D & Tan T 2017. Feature selection based on structured sparsity: A comprehensive study. *IEEE Transactions on Neural Networks and Learning Systems* 28, 7: 1490-1507.
- Guralnick LJ & Gladsky K 2017. Crassulacean acid metabolism as a continuous trait: variability in the contribution of Crassulacean acid metabolism (CAM) in populations of *Portulacaria afra*. *Heliyon* 3, 4: 1-15.
- Guralnick LJ & Ting IP 1986. Seasonal response to drought and rewatering in *Portulacaria afra* (L.) Jacq. *Oecologia* 70: 85-91.
- Guyon I & Elisseeff A 2003. An introduction to variable and feature selection. *Journal of Machine Learning Research* 3: 1157-1182.
- Guyon I, Weston J, Barnhill S & Vapnik V 2002. Gene selection for cancer classification using

- support vector machines. *Machine Learning* 46, 1-3: 389-422.
- Hall O, Hay GJ, Bouchard A & Marceau DJ 2004. Detecting dominant landscape objects through multiple scales: An integration of object-specific methods and watershed segmentation. *Landscape Ecology* 19, 1: 59-76.
- Hand DJ & Yu K 2001. Idiot's Bayes - not so stupid after all? *International Statistical Review* 69, 3: 385-398.
- Haralick RM, Shanmugam K & Dinstein I 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* SMC-3, 6: 610-621.
- Harris D & Van Niekerk A 2018. Feature clustering and ranking for selecting stable features from high dimensional remotely sensed data. *International Journal of Remote Sensing* 39, 23: 8934-8949.
- Harris D & Van Niekerk A 2019. Radiometric homogenisation of aerial images by calibrating with satellite data. *International Journal of Remote Sensing* 40, 7: 2623-2647.
- Harris D, Vlok J & Van Niekerk A 2018. Regional mapping of spekboom canopy cover using very high resolution aerial imagery. *Journal of Applied Remote Sensing* 12, 04: 1.
- Hay GJ, Castilla G, Wulder MA & Ruiz JR 2005. An automated object-based approach for the multiscale image segmentation of forest scenes. *International Journal of Applied Earth Observation and Geoinformation* 7, 4: 339-359.
- Hernández-Stefanoni JL, Dupuy JM, Johnson KD, Birdsey R, Tun-Dzul F, Peduzzi A, Caamal-Sosa JP, Sánchez-Santos G & López-Merlín D 2014. Improving species diversity and biomass estimates of tropical dry forests using airborne LiDAR. *Remote Sensing* 6, 6: 4741-4763.
- Heumann BW 2011. Satellite remote sensing of mangrove forests: Recent advances and future opportunities. *Progress in Physical Geography* 35, 1: 87-108.
- Hoare DB, Mucina L, Rutherford MC, Vlok JHJ, Euston-Brown DIW, Palmer AR, Powrie LW, Lechmere-Oertel RG, Procheş SM, Dold AP & Ward RA 2006. Albany Thicket Biome. In Mucina L & Rutherford MC (eds) *The vegetation of South Africa, Lesotho and Swaziland*, 540–567. Pretoria: SANBI.
- Hoffman MT & Cowling RM 1990. Desertification in the lower Sundays River Valley, South Africa. *Journal of Arid Environments* 19, 1: 105-118.
- Honkavaara E, Arbiol R, Markelin L, Martinez L, Cramer M, Bovet S, Chandelier L, Ilves R,

- Klonus S, Marshal P, Schläpfer D, Tabor M, Thom C & Veje N 2009. Digital airborne photogrammetry—A new tool for quantitative remote sensing?—A state-of-the-art review on radiometric aspects of digital photogrammetric images. *Remote Sensing* 1, 3: 577-605.
- Hou HS & Andrews H 1978. Cubic splines for image interpolation and digital filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26, 6: 508-517.
- Hovi A, Korhonen L, Vauhkonen J & Korpela I 2016. LiDAR waveform features for tree species classification and their sensitivity to tree- and acquisition related parameters. *Remote Sensing of Environment* 173: 224-237.
- HSU 2019. Vegetation spectral reflectance curves [online]. Available from: [http://gsp.humboldt.edu/OLM/Courses/GSP\\_216\\_Online/lesson2-1/vegetation.html](http://gsp.humboldt.edu/OLM/Courses/GSP_216_Online/lesson2-1/vegetation.html) [Accessed 25 April 2019].
- Immitzer M, Böck S, Einzmann K, Vuolo F, Pinnel N, Wallner A & Atzberger C 2018. Fractional cover mapping of spruce and pine at 1 ha resolution combining very high and medium spatial resolution satellite imagery. *Remote Sensing of Environment* 204: 690-703.
- Intergraph 2008. Digital mapping camera system [online]. Available from: [http://www.geospace.co.za/pdf/DMC Brochure.pdf](http://www.geospace.co.za/pdf/DMC%20Brochure.pdf)
- Jain AK, Duin RPW & Mao J 2000. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 1: 4-37.
- Jiang GM & Li ZL 2009. Cross-calibration of MSG1-SEVIRI infrared channels with Terra-MODIS channels. *International Journal of Remote Sensing* 30, 3: 753-769.
- Johansen K, Coops NC, Gergel SE & Stange Y 2007. Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment* 110, 1: 29-44.
- Johnson B & Xie Z 2013. Classifying a high resolution image of an urban area using super-object information. *ISPRS Journal of Photogrammetry and Remote Sensing* 83: 40-49.
- Jones TG, Coops NC & Sharma T 2011. Exploring the utility of hyperspectral imagery and LiDAR data for predicting *Quercus garryana* ecosystem distribution and aiding in habitat restoration. *Restoration Ecology* 19, 201: 245-256.
- Kalousis A, Prados J & Hilario M 2007. Stability of feature selection algorithms: A study on high-dimensional spaces. *Knowledge and Information Systems* 12, 1: 95-116.
- Kauth RJ & Thomas GS 1976. *The tasselled cap – a graphic description of the spectral-temporal*



- development of agricultural crops as seen by LANDSAT*. Symposium on Machine Processing of Remotely Sensed Data. Purdue University of West Lafayette, Indiana, USA. IEEE.: 4B41-4B51.
- Kerley GIH, Knight MH & De Kock M 1995. Desertification of subtropical thicket in the Eastern Cape, South Africa: Are there alternatives? *Environmental Monitoring and Assessment* 37, 1-3: 211-230.
- Key T, Warner TA, McGraw JB & Fajvan MA 2001. A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. *Remote Sensing of Environment* 75, 1: 100-112.
- Kollár S, Vekerdy Z & Márkus B 2013. Aerial image classification for the mapping of riparian vegetation habitats. *Acta Silvatica et Lignaria Hungarica* 9, 1: 119-133.
- Kononenko I, Šimec E & Robnik-Šikonja M 1997. Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence* 7, 1: 39-55.
- Korpela I, Mehtätalo L, Markelin L, Seppänen A & Kangas A 2014. Tree species identification in aerial image data using directional reflectance signatures. *Silva Fennica* 48, 3: 1-20.
- Korpela I, Ørka HO, Hyypä J, Heikkinen V & Tokola T 2010. Range and AGC normalization in airborne discrete-return LiDAR intensity data for forest canopies. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 4: 369-379.
- Kuncheva LI 2007. *A stability index for feature selection*. Proceedings of International Multi-conference: Artificial Intelligence and Applications. Innsbruck: 390-395.
- Laliberte AS, Goforth MA, Steele CM & Rango A 2011. Multispectral remote sensing from unmanned aircraft: Image processing workflows and applications for rangeland environments. *Remote Sensing* 3, 11: 2529-2551.
- LandInfo 2018. Pricing information for high resolution satellite imagery [online]. Available from: <http://www.landinfo.com/satellite-imagery-pricing.html> [Accessed 16 September 2019].
- Lechmere-Oertel RG, Kerley GIH & Cowling RM 2005. Patterns and implications of transformation in semi-arid succulent thicket, South Africa. *Journal of Arid Environments* 62, 3: 459-474.
- Lechmere-Oertel RG, Kerley GIH, Mills AJ & Cowling RM 2008. Litter dynamics across browsing-induced fenceline contrasts in succulent thicket, South Africa. *South African Journal of Botany* 74, 4: 651-659.

- Lelong CCD, Burger P, Jubelin G, Roux B, Labbé S & Baret F 2008. Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. *Sensors* 8, 5: 3557-3585.
- Li S, Harner EJ & Adjero D a 2011. Random KNN feature selection - a fast and stable alternative to Random Forests. *BMC Bioinformatics* 12, 1: 450.
- Li Z, Hayward RF, Zhang J, Jin H & Walker RA 2010. *Evaluation of spectral and texture features for object-based vegetation species classification using support vector machines*. Wagner W & Székely B (eds). ISPRS TC VII Symposium - 100 Years ISPRS. Vienna, Austria. IAPRS.: 122-127.
- Li L, Yang J & Wang Y 2012. *Cross-calibration of HJ-1B/CCD1 image based on Aqua/MODIS data*. Proceedings of 2012 Second International Workshop on Earth Observation and Remote Sensing Applications. Shanghai: 116-119.
- Lichman M 2013. UCI Machine learning repository [online]. Available from: <http://archive.ics.uci.edu/ml> [Accessed 15 August 2016].
- Liew S 2001. Principles of remote sensing [online]. Available from: <https://crisp.nus.edu.sg/~research/tutorial/rsmain.htm> [Accessed 11 February 2019].
- Liu J-J, Li Z, Qiao Y-L, Liu Y-J & Zhang Y-X 2004. A new method for cross-calibration of two satellite sensors. *International Journal of Remote Sensing* 25, 23: 5267-5281.
- Liu X, Liu H, Qiu S, Wu X, Tian Y & Hao Q 2017. An improved estimation of regional fractional woody/herbaceous cover using combined satellite data and high-quality training samples. *Remote Sensing* 9, 1: 1-13.
- Lloyd JW, Van den Berg EC & Palmer AR 2002. *Patterns of transformation and degradation in the Thicket Biome, South Africa*. Publication No GW/A/2002/30. Port Elizabeth: Terrestrial Ecology Research Unit, University of Port Elizabeth.
- López DH, García BF, Piqueras JG & Guillermo VA 2011. An approach to the radiometric aerotriangulation of photogrammetric images. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 6: 883-893.
- Lu D 2006. The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing* 27, 7: 1297-1328.
- Ludwig JA, Reynolds JF & Whitson PD 1975. Size-biomass relationships of several Chihuahuan desert shrubs. *American Midland Naturalist* 94, 2: 451-461.

- Manabe S, Strickler RF, Manabe S & Strickler RF 1964. Thermal equilibrium of the atmosphere with a convective adjustment. *Journal of the Atmospheric Sciences* 21, 4: 361-385.
- Marais C, Cowling RM & Powell M 2009. *Establishing the platform for a carbon sequestration market in South Africa: The Working for Woodlands Subtropical Thicket Restoration Programme*. Proceedings of XIII World Forestry Congress. Buenos Aires: 1-13.
- Markelin L, Honkavaara E, Schläpfer D, Bovet S & Korpela I 2012. Assessment of radiometric correction methods for ADS40 imagery. *Photogrammetrie - Fernerkundung - Geoinformation* 2012, 3: 251-266.
- Matese A, Di Gennaro SF & Berton A 2017. Assessment of a canopy height model (CHM) in a vineyard using UAV-based multispectral imaging. *International Journal of Remote Sensing* 38, 8-10: 2150-2160.
- Mather PM & Koch M 2011. *Computer processing of remotely-sensed images*. Chichester: John Wiley & Sons.
- Mathieu R, Wessels K, Main R, Naidoo L, van der Bergh F & Erasmus B 2018. A radar- and LiDAR-based earth observation system for monitoring savanna woody structure in southern Africa. *Biodiversity & Ecology* 6: 355-359.
- Mehner H, Cutler M, Fairbairn D & Thompson G 2004. Remote sensing of upland vegetation: the potential of high spatial resolution satellite sensors. *Global Ecology and Biogeography* 13, 4: 359-369.
- Michez A, Piégay H, Lisein J, Claessens H & Lejeune P 2016. Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. *Environmental Monitoring and Assessment* 188, 3: 146.
- Mills AJ, Blignaut J, Cowling RM, Knipe A, Marais C, Marais S, Pierce S, Powell M, Sigwela A & Skowno A 2010. *Investing in sustainability. Restoring degraded thicket, creating jobs, capturing carbon and earning green credit*. Port Elizabeth: Climate Action Partnership, Cape Town and Wilderness Foundation.
- Mills AJ & Cowling RM 2006. Rate of carbon sequestration at two thicket restoration sites in the Eastern Cape, South Africa. *Restoration Ecology* 14, 1: 38-49.
- Mills AJ & Cowling RM 2010. Below-ground carbon stocks in intact and transformed subtropical thicket landscapes in semi-arid South Africa. *Journal of Arid Environments* 74, 1: 93-100.
- Mills AJ & Cowling RM 2014. How fast can carbon be sequestered when restoring degraded

- subtropical thicket? *Restoration Ecology* 22, 5: 571-573.
- Mills AJ, Cowling RM, Fey M, Kerley G, Donaldson J, Sigwela A, Skowno A & Rundel P 2005. Effects of goat pastoralism on ecosystem carbon storage in semiarid thicket, Eastern Cape, South Africa. *Austral Ecology* 30: 797-804.
- Mills AJ & Fey M 2004. Transformation of thicket to savanna reduces soil quality in the Eastern Cape, South Africa. *Plant and Soil* 265, 1-2: 153-163.
- Mills AJ, Milton SJ, Taplin BD & Allen JL 2018. Viability of watering *Portulacaria afra* truncheons to facilitate restoration of subtropical thicket: Results from a nursery experiment and cost model. *South African Journal of Botany* 115: 58-64.
- Mills AJ & Robson A 2017. Survivorship of spekboom (*Portulacaria afra*) planted within the Subtropical Thicket Restoration Programme. *South African Journal of Science* 113, 1-2: 1-3.
- Mills AJ, Turpie JK, Cowling RM, Marais C, Kerley G, Richard G, Sigwela A & Powell M 2007. Assessing costs, benefits, and feasibility of restoring natural capital in Subtropical Thicket in South Africa. In Aronson J, Milton S & Blignaut J (eds) *Restoring Natural Capital: Science, Business and Practice (The Science and Practice of Ecological Restoration Series)*, 179-187. Washington DC: Island Press.
- Mills AJ, Van der Vyver M, Gordon IJ, Patwardhan A, Marais C, Blignaut J, Sigwela A & Kgope B 2015. Prescribing innovation within a large-scale restoration programme in degraded subtropical thicket in South Africa. *Forests* 6, 11: 4328-4348.
- Mingguo Z, Qianguo C & Mingzhou Q 2009. The Effect of Prior Probabilities in the Maximum Likelihood Classification on Individual Classes. *Photogrammetric Engineering & Remote Sensing* 75, 9: 1109-1117.
- Mirik M & Ansley RJ 2012. Utility of satellite and aerial images for quantification of canopy cover and infilling rates of the invasive woody species honey mesquite (*Prosopis glandulosa*) on rangeland. *Remote Sensing* 4, 7: 1947-1962.
- Mishra KK & Harit S 2010. A fast algorithm for finding the non-dominated set in multi objective optimization. *Multi-Objective Optimization using Evolutionary Algorithms* 1, 25: 35-39.
- Mitra P, Murthy CA & Pal SK 2002. Unsupervised feature selection using feature similarity. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI* 24, 3: 301-312.
- MODIS Land Team 2014. EOS validation status for MODIS BRDF/albedo: MCD43 [online]. Available from: <http://tinyurl.com/jyx8gjs> [Accessed 6 June 2014].

- Moolman HJ & Cowling RM 1994. The impact of elephant and goat grazing on the endemic flora of South African succulent thicket. *Biological Conservation* 68, 1: 53-61.
- Mustafa YT & Habeeb HN 2014. *Object based technique for delineating and mapping 15 tree species using VHR WorldView-2 imagery*. Proceedings of Remote Sensing for Agriculture, Ecosystems, and Hydrology XVI. Amsterdam: 92390G-1-13.
- Myneni RB, Hall FG, Sellers PJ & Marshak AL 1995. The interpretation of spectral vegetation indexes. *IEEE Transactions on Geoscience and Remote Sensing* 33, 2: 481-486.
- National Agriculture Imagery Program 2019. NAIP Imagery [online]. Available from: <https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/index> [Accessed 23 April 2019].
- National Geo-spatial Information 2010. Standard for the acquisition of digital aerial imagery [online]. Available from: [http://www.ngi.gov.za/images/stories/technical\\_info/standards/Standard\\_for\\_the\\_Acquisition\\_of\\_Digital\\_Aerial\\_Imagery.pdf](http://www.ngi.gov.za/images/stories/technical_info/standards/Standard_for_the_Acquisition_of_Digital_Aerial_Imagery.pdf) [Accessed 22 March 2012].
- National Geo-spatial Information 2012. National aerial photography and imagery programme [online]. Available from: <http://www.ngi.gov.za/index.php/what-we-do/aerial-photography-and-imagery> [Accessed 22 May 2012].
- Neigh CSR, Carroll ML, Wooten MR, McCarty JL, Powell BF, Husak GJ, Enenkel M & Hain CR 2018. Smallholder crop area mapped with wall-to-wall WorldView sub-meter panchromatic image texture: A test case for Tigray, Ethiopia. *Remote Sensing of Environment* 212: 8-20.
- Neinavaz E 2017. Sensing vegetation canopies in the thermal domain. Doctoral Dissertation. Enschede: University of Twente, Faculty of Geo-Information Science and Earth Observation.
- Nguyen U, Glenn EP, Dang TD & Pham LTH 2019. Mapping vegetation types in semi-arid riparian regions using random forest and object-based image approach: A case study of the Colorado River Ecosystem, Grand Canyon, Arizona. *Ecological Informatics* 50: 43-50.
- Nie F, Huang H, Cai X & Ding C 2010. *Efficient and robust feature selection via joint  $\ell_2, 1$ -norms minimization*. Proceedings of Advances in Neural Information Processing Systems 23. Vancouver: 1813-1821.
- Niphadkar M & Nagendra H 2016. Remote sensing of invasive plants: incorporating functional traits into the picture. *International Journal of Remote Sensing* 37, 13: 3074-3085.
- Niphadkar M, Nagendra H, Tarantino C, Adamo M & Blonda P 2017. Comparing pixel and object-based approaches to map an understory invasive shrub in tropical mixed forests. *Frontiers*

- in Plant Science* 8, May: 1-18.
- Nock R & Nielsen F 2004. Statistical region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 11: 1452-1458.
- Ojala T, Pietikainen M & Maenpaa T 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 7: 971-987.
- OpenCV Development Team 2014. OpenCV documentation [online]. Available from: <http://docs.opencv.org/>
- Ørka HO & Hauglin M 2016. *Use of remote sensing for mapping of non-native conifer species*. Ås: Norwegian University of Life Sciences.
- Otsu N 1979. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9, 1: 62-66.
- Ouyang Z-T, Zhang M-Q, Xie X, Shen Q, Guo H-Q & Zhao B 2011. A comparison of pixel-based and object-oriented approaches to VHR imagery for mapping saltmarsh plants. *Ecological Informatics* 6, 2: 136-146.
- Pekalska E & Duin RPW 2012. Representation and generalization [online]. Available from: <http://37steps.com/720/representation-and-generalisation/> [Accessed 25 March 2019].
- Peng H, Long F & Ding C 2005. Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27, 8: 1226-1238.
- Pix4D 2019. Designing the Image Acquisition Plan [online]. Available from: <https://support.pix4d.com/hc/en-us/articles/202557459-Step-1-Before-Starting-a-Project-1-Designing-the-Image-Acquisition-Plan-a-Selecting-the-Image-Acquisition-Plan-Type> [Accessed 19 April 2019].
- Powell MJ 2009. Restoration of degraded subtropical thickets in the Baviaanskloof Megareserve, South Africa. Master's Thesis. Grahamstown: Rhodes University, Department of Environmental Science.
- Powell MJ, Mills AJ & Marais C 2005. Carbon sequestration and restoration: Challenges and opportunities in subtropical thicket [online]. Available from: [https://www.daff.gov.za/daffweb3/Portals/0/GeneralReports/4271\\_\\_\\_Day2\\_session3\\_item1.pdf](https://www.daff.gov.za/daffweb3/Portals/0/GeneralReports/4271___Day2_session3_item1.pdf) [Accessed 29 December 2018].

- Puissant A, Hirsch J & Weber C 2005. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing* 26: 733-745.
- Restoration Research Group 2012. Intact and degraded spekboomveld [online]. Available from: <http://www.r3g.co.za/spekboom.htm> [Accessed 10 September 2012].
- Ribeiro da Luz B & Crowley JK 2007. Spectral reflectance and emissivity features of broad leaf plants: Prospects for remote sensing in the thermal infrared (8.0–14.0  $\mu\text{m}$ ). *Remote Sensing of Environment* 109, 4: 393-405.
- Ribeiro da Luz B & Crowley JK 2010. Identification of plant species by using high spatial and spectral resolution thermal infrared (8.0–13.5  $\mu\text{m}$ ) imagery. *Remote Sensing of Environment* 114, 2: 404-413.
- Richter R 1997. Correction of atmospheric and topographic effects for high spatial resolution satellite imagery. *International Journal of Remote Sensing* 18, 5: 1099-1111.
- Roujean J-L, Leroy M & Deschamps P-Y 1992. A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data. *Journal of Geophysical Research* 97, D18: 20455.
- Sahu B & Mishra D 2011. *A novel approach for selecting informative genes from gene expression data using signal-to-noise ratio and t-statistics*. Proceedings of 2nd International Conference on Computer and Communication Technology (ICCCCT-2011). Allahabad: 5-10.
- Schaepman-Strub G, Schaepman ME, Painter TH, Dangel S & Martonchik JV 2006. Reflectance quantities in optical remote sensing—definitions and case studies. *Remote Sensing of Environment* 103, 1: 27-42.
- Schmidt G, Jenkerson C, Masek J, Vermote E & Gao F 2012. *Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) Algorithm Description*. December.
- Schölkopf B, Smola A & Müller K-R 1998. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation* 10, 5: 1299-1319.
- Serra J & Soille P (eds) 1994. *Mathematical morphology and its applications to image processing*. 2nd International Symposium on Mathematical Morphology (ISMM'94). Computational imaging and vision. Kluwer Academic Publishers.: 383.
- Sigwela A, Kerley G, Mills AJ & Cowling RM 2009. The impact of browsing-induced degradation on the reproduction of subtropical thicket canopy shrubs and trees. *South African Journal of*

*Botany* 75, 2: 262-267.

- Singh M, Malhi Y & Bhagwat S 2014. Biomass estimation of mixed forest landscape using a Fourier transform texture-based approach on very-high-resolution optical satellite imagery. *International Journal of Remote Sensing* 35, 9: 3331-3349.
- Statistics South Africa 2018. *Quarterly labour force survey*. Quarter 2: 2018. Pretoria: Statistics South Africa.
- Strahler AH & Muller JP 1999. *MODIS BRDF/albedo product: Algorithm theoretical basis document*. Version 5.0. Greenbelt: NASA.
- Strobl C, Boulesteix A-L, Kneib T, Augustin T & Zeileis A 2008. Conditional variable importance for random forests. *BMC bioinformatics* 9: 307.
- Stuart-Hill GC 1992. Effects of elephants and goats on the Kaffrarian succulent thicket of the eastern Cape, South Africa. *The Journal of Applied Ecology* 29, 3: 699.
- Su L 2009. Optimizing support vector machine learning for semi-arid vegetation mapping by using clustering analysis. *ISPRS Journal of Photogrammetry and Remote Sensing* 64, 4: 407-413.
- Suganuma H, Abe Y, Taniguchi M, Tanouchi H, Utsugi H, Kojima T & Yamada K 2006. Stand biomass estimation method by canopy coverage for application to remote sensing in an arid area of Western Australia. *Forest Ecology and Management* 222, 1-3: 75-87.
- Székely GJ & Rizzo ML 2005. Hierarchical clustering via joint between-within distances: extending Ward's minimum variance method. *Journal of Classification* 22, 2: 151-183.
- Székely GJ, Rizzo ML & Bakirov NK 2007. Measuring and testing dependence by correlation of distances. *The Annals of Statistics* 35, 6: 2769-2794.
- Tarantino C, Casella F, Adamo M, Lucas R, Beierkuhnlein C & Blonda P 2019. *Ailanthus altissima* mapping from multi-temporal very high resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing* 147: 90-103.
- Tempfli K, Kerle N, Janssen LLF & Huurneman GC (eds) 2009. *Principles of remote sensing: An introductory textbook*. Enschede: ITC.
- Thompson M, Vlok JHJ, Rouget M, Hoffman MT & Cowling RM 2009. Mapping grazing-induced degradation in a semi-arid environment: A rapid and cost-effective approach for assessment and monitoring. *Environmental Management* 43: 585-596.
- Ting IP & Hanscom Z 1977. Induction of acid metabolism in *Portulacaria afra*. *Plant Physiology* 59, 3: 511-514.



- Tinley K 1975. Habitat physiognomy, structure and relationships. *University of Pretoria Publications* 97: 67-77.
- Tolosi L & Lengauer T 2011. Classification with correlated features: unreliability of feature ranking and solutions. *Bioinformatics* 27, 14: 1986-1994.
- Toutin T 2004. Review article: Geometric processing of remote sensing images: Models, algorithms and methods. *International Journal of Remote Sensing* 25, 10: 1893-1924.
- Trias-Sanz R, Stamon G & Louchet J 2008. Using colour, texture, and hierarchical segmentation for high-resolution remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 63, 2: 156-168.
- Trimble 2016. eCognition Developer user guide [online]. Available from: <http://www.ecognition.com/> [Accessed 8 July 2016].
- Trimble 2018. Trimble eCognition [online]. Available from: <http://www.ecognition.com/> [Accessed 1 December 2018].
- TU Delft 2015. PRTools [online]. Available from: <http://prtools.org/prtools/> [Accessed 20 October 2016].
- Van Luijk G, Cowling RM, Riksen MJPM & Glenday J 2013. Hydrological implications of desertification: Degradation of South African semi-arid subtropical thicket. *Journal of Arid Environments* 91: 14-21.
- Van Niekerk A 2016. Stellenbosch university digital elevation model (SUDEM). *ResearchGate* v16.
- Van der Vyver ML, Cowling RM, Campbell EE & Difford M 2012. Active restoration of woody canopy dominants in degraded South African semi-arid thicket is neither ecologically nor economically feasible. *Applied Vegetation Science* 15, 1: 26-34.
- Van der Vyver ML, Cowling RM, Mills AJ & Difford M 2013. Spontaneous return of biodiversity in restored subtropical thicket: *Portulacaria afra* as an ecosystem engineer. *Restoration Ecology* 21, 6: 736-744.
- Vaz AS, Alcaraz-Segura D, Campos JC, Vicente JR & Honrado JP 2018. Managing plant invasions through the lens of remote sensing: A review of progress and the way forward. *Science of The Total Environment* 642: 1328-1339.
- Vermote E, Tanré D, Deuzé JL & Herman M 2006. Second simulation of a satellite signal in the solar spectrum - vector (6SV) (user guide v3) [online]. Available from:

- [http://6s.ltdri.org/files/tutorial/6S\\_Manual\\_Part\\_1.pdf](http://6s.ltdri.org/files/tutorial/6S_Manual_Part_1.pdf) [Accessed 1 June 2019].
- Vermote E, Tanré D, Deuzé JL, Herman M & Morcette J-J 1997. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE Transactions on Geoscience and Remote Sensing* 35, 3: 675-686.
- Verstraete MM, Hunt LA, Scholes RJ, Clerici M, Pinty B & Nelson DL 2012. Generating 275-m resolution land surface products from the multi-angle imaging spectroradiometer data. *IEEE Transactions on Geoscience and Remote Sensing* 50, 10: 3980-3990.
- Vescovi F & Minchella A 2017. *Copernicus data quality control - technical note bidirectional reflectance distribution function*. 2.1. Didcot: European Space Agency.
- Vicente-Serrano S, Pérez-Cabello F & Lasanta T 2008. Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images. *Remote Sensing of Environment* 112, 10: 3916-3934.
- Vlok JHJ, Cowling RM & Wolf T 2005. *A vegetation map for the Little Karoo*. Unpublished maps and report for a SKEP project supported by CEPF grant no 1064410304.
- Vlok JHJ & Euston-Brown DIW 2002. *The patterns within, and the ecological processes that sustain, the subtropical thicket vegetation in the planning domain of the Subtropical Thicket Ecosystem Planning (STEP) project*. Publication No 40. Port Elizabeth: Terrestrial Ecology Research Unit, University of Port Elizabeth.
- Vlok JHJ, Euston-Brown DIW & Cowling RM 2003. Acocks' Valley Bushveld 50 years on: New perspectives on the delimitation, characterisation and origin of subtropical thicket vegetation. *South African Journal of Botany* 69, 1: 27-51.
- Waser L, Küchler M, Jütte K & Stampfer T 2014. Evaluating the potential of WorldView-2 data to classify tree species and different levels of ash mortality. *Remote Sensing* 6, 5: 4515-4545.
- Webb AR 2002. *Statistical pattern recognition*. Chichester, UK: John Wiley & Sons.
- Wu B, Chen C, Kechadi TM & Sun L 2013. A comparative evaluation of filter-based feature selection methods for hyper-spectral band selection. *International Journal of Remote Sensing* 34, 22: 7974-7990.
- Yang HH & Moody J 2000. Data visualization and feature selection: New algorithms for nongaussian data. *Advances in Neural Information Processing Systems* 12: 687-693.
- Yousef M, Jung S, Showe LC & Showe MK 2007. Recursive cluster elimination (RCE) for classification and feature selection from gene expression data. *BMC Bioinformatics* 8, 144.

Yu L & Liu H 2004. Efficient feature selection via analysis of relevance and redundancy. *Journal of Machine Learning Research* 5, 2004: 1205-1224.

Yuan Y, Zhu G & Wang Q 2015. Hyperspectral band selection by multitask sparsity pursuit. *IEEE Transactions on Geoscience and Remote Sensing* 53, 2: 631-644.

Zhang D & Lu G 2004. Review of shape representation and description techniques. *Pattern Recognition* 37, 1: 1-19.