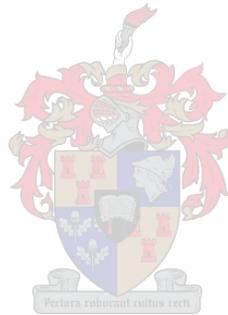


**AN INTEGRATED APPROACH TO GRASSLAND PRODUCTIVITY
MODELLING USING SPECTRAL MIXTURE ANALYSIS, PRIMARY
PRODUCTION AND GOOGLE EARTH ENGINE**

By LIEZL MARI VERMEULEN

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



Supervisor: Dr Zahn Munch

March 2020

DECLARATION

By submitting this report 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.

Date: March 2020-----

SUMMARY

Grassland degradation can have a severe impact on condition, productivity and consequently grazing potential. Current conventional methods for monitoring and managing grasslands are time-consuming, destructive and not applicable at large-scale. These constraints could be addressed using a remote sensing (RS)-based approach, however, current RS-based approaches also have technological and scientific limitations in the context of grassland management.

The inability of RS-based primary production models to discriminate between herbaceous and woody production at sub-pixel level poses constraints for use in grazing capacity (GC) calculation. The integration of fractional vegetation cover (FVC) is posed as a promising solution, specifically estimation using spectral mixture analysis (SMA). Current grassland monitoring approaches are limited by the technological constraints of traditional, desktop-based RS approaches, but the implementation of analysis in a Google Earth Engine (GEE) web application can address these limitations by providing dynamic, continuous productivity estimates.

Field data collection and analysis of biophysical parameters were performed to establish crucial relationships between vegetation productivity and RS signals. Biophysical parameters obtained include FVC, leaf area index (LAI), fraction of absorbed photosynthetically active radiation (fAPAR) and grass dry matter (DM) production. An important outcome was the improvement of the normalised difference vegetation index (NDVI) and fAPAR regression relationship, achieved by scaling fAPAR using the proportion of green, living biomass. The relationship proved useful in subsequent vegetation productivity modelling.

The potential of SMA for FVC estimation using medium resolution imagery (Landsat 8 and Sentinel-2) and relatively few field points, was explored. A linear spectral mixture model (LSMM) was calibrated, implemented and evaluated on accuracy and transferability. A number of bands and spectral indices were identified as core features, specifically the dry bare-soil index (DBSI). DBSI improved discrimination between bare ground and dry vegetation, a common challenge in semi-arid conditions. The calibrated LSMM performed well, with Sentinel-2 providing the most accurate results. The research proved the transferability of the LSMM approach, as accurate FVC estimates were obtained for both arid, dry season conditions and green, growing season conditions.

The LSMM-estimated FVC was combined with primary production to improve GC calculation for grassland and rangelands. Annual grassland production was calculated using the Regional Biosphere Model (RBM). Although a water stress factor is a well-known source of uncertainty, the research found its inclusion crucial to the transferability of the model between different climatic conditions. FVC was used to determine the grazable primary production from RBM

estimates, thus mitigating the effects of woody components on GC calculations. A comparison of model-estimated GC to the most recent national GC map showed good agreement. Slight discrepancies were likely due to the inability of the model to include species composition and palatability in GC calculations. The final FVC-integrated productivity model was implemented in a GEE web app to demonstrate the practical contribution of the research for continuous, dynamic, multi-scale and sustainable grassland management.

Overall, the findings of the research provide valuable insights into improving RS-based modelling of grassland condition and productivity. Operationalisation of this research can aid in identifying potential degradation, highlighting regions vulnerable to food shortages and establishing sustainable productivity levels. Recommendations include investigating alternative methods for estimating water stress and exploring the incorporation of species composition in GC calculation using RS.

KEY WORDS

grassland productivity, grassland condition, remote sensing, fractional vegetation cover, spectral mixture analysis, primary production modelling, grazing capacity, Google Earth Engine, sustainable grassland management

OPSOMMING

Agteruitgang van grasvelde kan 'n ernstige invloed op kondisie, produktiwiteit en gevolglik weidingspotensiaal hê. Huidige konvensionele metodes vir die monitering en bestuur van grasvelde is tydrowend, vernietigend en nie op groot skaal toepasbaar nie. Hierdie beperkinge kan met behulp van 'n afstandwaarnemings (AW)-gebaseerde benadering aangespreek word, maar huidige AW-metodes het egter ook tegnologiese en wetenskaplike beperkings, veral in die konteks van veldbestuur.

Die onvermoë van AW-gebaseerde primêre produksiemodelle om tussen kruidagtige en houtagtige produksie op sub-pixelvlak te onderskei, hou beperkings in vir die berekening van drakapasiteit (DK). Die integrasie van fraksionele plantegroeibedekking (FPB) word aangebied as 'n belowende oplossing. Beraming van FPB deur gebruik te maak van spektrale mengselanalise (SMA) het veral potensiaal. Huidige benaderings vir die monitering van grasvelde word beperk deur die tegnologiese beperkings van tradisionele, rekenaargebaseerde AW-metodes, maar die implementering van analise in 'n Google Earth Engine (GEE) webtoepassing kan hierdie beperkings aanspreek deur dinamiese, deurlopende produktiwiteitsramings te verskaf.

Velddata is ingesamel en analise van biofisiese parameters is uitgevoer om belangrike verwantskappe tussen plantproduktiwiteit en AW-seine te bepaal. Die biofisiese parameters sluit in FPB, blaaroppervlakte-indeks (BOI), fraksie van geabsorbeerde fotosintetiese aktiewe bestraling (fAFAB) en droë materiaal (DM) produksie. Die verbetering van die genormaliseerde verskil-plantegroei-indeks (NVPI) en fAFAB -regressie-verhouding, wat verkry is deur fAFAB te skaleer met behulp van die hoeveelheid groen, lewende biomassa was 'n belangrike uitkoms. Die verwantskap was nuttig in die daaropvolgende modellering van plantegroei.

Die potensiaal van SMA vir die bepaling van FPB deur middel van medium resolusiebeelde (Landsat 8 en Sentinel-2) met relatief min veldpunte is ondersoek. 'n Lineêre spektrale mengselmodel (LSMM) is gekalibreer, geïmplementeer en vir akkuraatheid en oordraagbaarheid geëvalueer. 'n Aantal bande en spektrale indekse is as kernkenmerke geïdentifiseer, spesifiek die droë kaal-grondindeks (DKGI). DKGI het die onderskeid tussen kaal grond en droë plantegroei, 'n algemene uitdaging in semi-droë landskappe, verbeter. Die gekalibreerde LSMM het goed gevaar, met Sentinel-2 wat die akkuraatste resultate gelewer het. Die navorsing het bewys dat die LSMM-benadering oorgedra kan word, aangesien akkurate FPB-ramings vir beide droë seisoen en groen, groeiseisoen toestande verkry is.

Die LSMM-beraamde FPB is met primêre produksie ramings gekombineer om die DK-berekening vir grasveld te verbeter. Die jaarlikse grasveldproduksie is met behulp van die Streeks Biosfeer

Model (SBM) bereken. Alhoewel 'n waterstresfaktor 'n bron van onsekerheid is, het die navorsing bevind dat dit die gebruik daarvan vir die oordraagbaarheid van die model tussen verskillende klimaatstoestande belangrik is. FPB is gebruik om die weibare primêre produksie volgens SBM-ramings te bepaal, en het die effekte van houtagtige komponente op DK-berekeninge verminder. 'n Vergelyking van die gemodelleerde DK met die nuutste nasionale DK-kaart het 'n goeie ooreenkoms getoon. Klein afwykings was waarskynlik te wyte aan die onvermoë van die model om spesiesamestelling en eetbaarheid by DK-berekeninge in te sluit. Die finale FPB-geïntegreerde produktiwiteitsmodel is in 'n GEE webtoep geïmplementeer om die praktiese bydrae van die navorsing vir deurlopende, dinamiese, meervoudige en volhoubare grasveldbestuur te demonstreer.

In die geheel bied die bevindinge van die navorsing waardevolle insigte in die verbetering van die AW-gebaseerde modellering van veldtoestand en produktiwiteit. Operasionalisering van hierdie navorsing kan tot die identifisering van potensiële agteruitgang, die uitlig van streke wat kwesbaar is vir voedseltekorte en die bepaling van volhoubare produktiwiteitsvlakke bydra. Aanbevelings sluit in die ondersoek van alternatiewe metodes vir die beraming van waterstres en die gebruik van spesiesamestelling in DK-berekening met behulp van AW.

TREFWOORDE

graslandproduktiwiteit, grasveldtoestand, afstandswaarneming, fraksionele plantbedekking, spektrale mengselanalise, primêre produksiemodellering, weidingskapasiteit, Google Earth Engine, volhoubare grasveldbestuur

ACKNOWLEDGEMENTS

I sincerely thank:

- My supervisor, Dr Zahn Munch, for her excellent guidance, invaluable advice and continuous encouragement throughout my academic and research career.
- My informal co-supervisor and mentor, Dr Anthony Palmer, for help with field work and extensive knowledge on all things grazing and grassland related.
- The South African Space Agency for awarding me a bursary to pursue this research.
- The staff at the Department of Geography and Environmental Sciences for helpful comments and constructive criticism during scheduled feedback sessions.
- My mother, Mrs Gene Vermeulen, for her continuous and unconditional support, patience and supply of coffee, as well as her incredible editing services.
- My father, Prof Johan Vermeulen, for continuously reminding that anything is possible with a little resilience.
- All my other friends who are not mentioned above for providing a great support system and ensuring my time at Stellenbosch University was a great experience.

CONTENTS

DECLARATION.....	ii
SUMMARY	iii
OPSOMMING.....	v
ACKNOWLEDGEMENTS	vii
CONTENTS	viii
TABLES	xii
FIGURES.....	xiii
APPENDICES	xvi
ACRONYMS AND ABBREVIATIONS.....	xvii
CHAPTER 1: INTRODUCTION	1
1.1 GRASSLAND CONVERSION AND DEGRADATION	1
1.2 REMOTE SENSING AND GRASSLAND PRODUCTIVITY.....	3
1.3 PROBLEM FORMULATION.....	4
1.4 RESEARCH AIM AND OBJECTIVES	5
1.5 STUDY AREA	6
1.6 RESEARCH METHODOLOGY & DESIGN	8
CHAPTER 2: DETERMINING GRASSLAND CONDITION AND PRODUCTIVITY	12
2.1 REMOTE SENSING.....	12
2.1.1 Surface reflectance properties and spectral signatures	13
2.1.2 Spectral indices.....	14
2.2 VEGETATION PRODUCTION IN TERRESTRIAL ECOSYSTEMS	16
2.2.1 Primary production models.....	17
2.2.1.1 Eddy covariance models.....	17
2.2.1.2 Vegetation index models.....	18
2.2.1.3 Light use efficiency models	18
2.2.1.4 Process-based models.....	20
2.2.1.5 Comparison of primary production models	20
2.2.2 Primary production modelling in southern Africa	21
2.2.2.1 Regional Biosphere Model.....	22
2.2.2.2 Biosphere Energy Transfer Hydrology model	25

2.3	FRACTIONAL VEGETATION COVER	27
2.3.1	Field-based methods.....	27
2.3.2	Remote sensing based methods	28
2.3.2.1	Vegetation indices	28
2.3.2.2	Spectral mixture analysis	29
2.3.2.3	Machine learning.....	29
2.3.3	Fractional vegetation cover products	30
2.4	GRASSLAND AND RANGELAND MANAGEMENT.....	31
2.4.1	Dry matter production.....	32
2.4.1.1	Clipping-and-weighing.....	32
2.4.1.2	Disc pasture meter.....	32
2.4.2	Grazing capacity.....	33
2.5	MODERN TECHNOLOGIES.....	35
2.5.1	Google Earth Engine.....	35
2.5.2	Existing applications	37
2.6	LITERATURE SUMMARY	38
CHAPTER 3: FIELD DATA COLLECTION AND ANALYSIS.....		40
3.1	DATA COLLECTION	40
3.1.1	Dry season field trip: Fort Beaufort.....	40
3.1.1.1	Fractional vegetation cover	41
3.1.1.2	Leaf area index	41
3.1.2	Growing season field trip: Cedarville.....	41
3.1.2.1	Fractional vegetation cover	43
3.1.2.2	Leaf area index and fAPAR	44
3.1.2.3	Green-dead biomass proportions.....	44
3.1.2.4	Dry matter production	45
3.2	DATA ANALYSIS	45
3.2.1	Dry season field trip: Fort Beaufort.....	45
3.2.2	Growing season field trip: Cedarville.....	47
3.2.2.1	NDVI-to-LAI	48
3.2.2.2	NDVI-to-fAPAR	54
3.2.2.3	NDVI-to-FVC (%)	56
3.2.2.4	DPM-to-DM production.....	57
3.3	DISCUSSION	57
CHAPTER 4: FRACTIONAL VEGETATION COVER ESTIMATION ...		59

4.1	METHODS & MATERIALS.....	59
4.1.1	Data selection & preparation	59
4.1.1.1	Satellite imagery selection and pre-processing	59
4.1.1.2	Endmember field points	61
4.1.2	Linear spectral mixture model.....	62
4.1.2.1	Feature selection.....	63
4.1.2.2	Endmember class refinement	63
4.1.2.3	Spectral signature extraction and refinement	64
4.1.2.4	Unmixing model implementation.....	64
4.1.2.5	Fractional vegetation cover evaluation	65
4.1.3	Model assessment	68
4.1.3.1	Comparison to existing fractional cover product	68
4.1.3.2	Validation against in-situ field data	69
4.1.4	Model transferability	69
4.1.5	Fractional vegetation cover change	69
4.2	RESULTS.....	69
4.2.1	Endmember spectral signatures	69
4.2.2	Fractional vegetation cover output maps.....	73
4.2.3	Model assessment	75
4.2.3.1	Comparison to existing fractional cover product	75
4.2.3.2	In-situ field data	77
4.2.4	Model transferability	77
4.2.5	Fractional vegetation cover change	78
4.3	DISCUSSION	78
CHAPTER 5: GRASSLAND PRODUCTIVITY MODELLING		82
5.1	METHODS & MATERIALS.....	82
5.1.1	Data selection & pre-processing.....	82
5.1.2	Primary production.....	83
5.1.2.1	APAR	85
5.1.2.2	Maximum light use efficiency.....	86
5.1.2.3	Effective light use efficiency.....	88
5.1.3	Grazing capacity.....	89
5.1.4	Accuracy assessment	90
5.1.4.1	Gross primary production.....	90
5.1.4.2	Grazing capacity.....	91

5.1.5	Application development.....	92
5.2	RESULTS.....	93
5.2.1	Gross primary production.....	93
5.2.1.1	Gross primary production output maps.....	93
5.2.1.2	Comparison to existing primary production products.....	94
5.2.1.3	Validation against in-situ field data	95
5.2.2	Grazing capacity.....	96
5.2.2.1	Comparison to long term grazing capacity map.....	96
5.2.2.2	Annual grazing capacity change	101
5.3	DISCUSSION	102
CHAPTER 6: DISCUSSION AND CONCLUSION.....		106
6.1	REVISITING AIM AND OBJECTIVES.....	106
6.2	MAIN FINDINGS AND VALUE OF THE RESEARCH	108
6.2.1	Field data collection and analysis for vegetation modelling.....	109
6.2.2	FVC estimation using SMA and medium resolution imagery	109
6.2.3	Grassland productivity modelling using FVC and primary production	111
6.2.4	GEE and application development for grassland management.....	113
6.3	LIMITATIONS AND FUTURE RESEARCH RECOMMENDATIONS	113
6.4	CONCLUSION.....	115
REFERENCES.....		116
APPENDICES		133

TABLES

Table 2.1 LUE subcomponent formulas	22
Table 2.2 Boolean and corresponding fuzzy operators	23
Table 2.3 RBM soil texture classes and corresponding fuzzy values	24
Table 3.1 Original and adjusted NDVI intervals for field sampling.....	43
Table 3.2 Summary of sample data for the dry season field trip to Fort Beaufort.....	46
Table 3.3 Results of regression analyses to determine the relationship between NDVI and LAI	51
Table 3.4 Results of regression analyses to determine the relationship between NDVI and fAPAR	54
Table 4.1 Comparison of Landsat 8 (L8) and Sentinel-2 (S2) corresponding bands.....	60
Table 4.2 Description of sensors, study area and analysis purpose of the different image stacks	61
Table 4.3 Final band and index combinations for Landsat 8 and Sentinel-2	63
Table 4.4 Final endmember class descriptions.....	64
Table 4.5 Field FVC values for field plots using discrete and fractional classification	67
Table 4.6 Comparison of model-estimated FVC and existing FVC product using RMSE and MAE	75
Table 4.7 Changes in field plot FVC from 2015 to 2018.....	78
Table 5.1 Image acquisition dates and composite time periods used for benchmark product comparison	91
Table 5.2 Comparison of model-estimated GPP to benchmark products using RMSE, MAE and mean	94
Table 5.3 Results of GPP validation against in-situ field measurements.....	95
Table 5.4 Comparison of model-estimated GC to LTGC map using zonal statistics	97

FIGURES

Figure 1.1 The two study areas, consisting of (a) Fort Beaufort and (b) Cedarville.....	8
Figure 1.2 Research design and thesis structure	10
Figure 2.1 The spectral signatures of soil, vegetation and water.	13
Figure 2.2 Ecosystem productivity variables.	16
Figure 2.3 BETHY model diagram	26
Figure 3.1 NDVI derived from Sentinel-2 imagery for the Cedarville region for (a) 2019/02/01 and (b) 2019/02/23	42
Figure 3.2 Example of changing FVC (%) with increasing transect length (m) with a 5 m stabilisation point	44
Figure 3.3 Sample points and corresponding bioregions for the dry season field trip to Fort Beaufort.....	46
Figure 3.4 Examples of different types and levels of woody encroachment, namely (a) no woody encroachment, (b) moderate woody encroachment of shrubs, (c) high woody encroachment by both shrubs and trees and (d) high woody encroachment by mainly trees	47
Figure 3.5 Sample sites for growing season field trip in Cedarville study area, namely (a) plains, (b) communal field, (c) farm 1, (d) farm 2, (e) farm 3 and (f) bare road.	48
Figure 3.6 Regression analysis for NDVI and unscaled LAI (blue), visually scaled LAI (orange) and DM mass scaled LAI (grey) for (a) 2019/02/01, (b) 2019/02/06, (c) 2019/02/08, (d) 2019/02/13, (e) 2019/02/18, (f) 2019/02/23, (g) 2019/02/26 and (h) 2019/02/28.	50
Figure 3.7 Mean NDVI and scaled LAI (DM mass) R^2 values for (a) daily rainfall (mmday ⁻¹) and (b) cumulative rainfall (mmday ⁻¹) for February 2019.....	52
Figure 3.8 Mean LAI per NDVI interval for (a) unscaled LAI, (b) scaled LAI (visual) and (c) scaled LAI (DM mass)	53
Figure 3.9 Mean fAPAR per NDVI interval for (a) unscaled fAPAR, (b) scaled fAPAR (visual) and (c) scaled fAPAR (DM mass).....	55
Figure 3.10 Regression analysis for NDVI and (a) bare cover (%), (b) grass cover (%) and shrub cover (%)	56
Figure 3.11 Calibration regressions for DPM height (cm) and (a) DM production (kg ha ⁻¹) and (b) the square root of DM production (kg ha ⁻¹).....	57
Figure 4.1 Comparison of Landsat 7 and 8 bands with Sentinel-2	60
Figure 4.2 Iterative model calibration process	62

Figure 4.3 Validation approach for CDLC fractional layers. Green corresponds to tree cover, orange to shrub cover and yellow to grass cover.	65
Figure 4.4 Comparison of field-classified FVC values using two classification methods, where (a) discrete classification and (b) fractional classification	66
Figure 4.5 Comparison of RMSE values for model-estimated FVC to (a) discrete field FVC and (b) fractional field FVC	67
Figure 4.6 Comparison of MAE values for model-estimated FVC to (a) discrete field FVC and (b) fractional field FVC	68
Figure 4.7 Mean spectral signatures for endmember classes Bare, Grassy and Woody for (a) Landsat 8 and (b) Sentinel-2.....	70
Figure 4.8 Mean index values for endmember classes Bare, Grassy and Woody for (a) Landsat 8 and (b) Sentinel-2	70
Figure 4.9 Mean spectral signatures for endmember subclasses Bare 1, Bare 2, Grassy, Shrubs and Trees (a) Landsat 8 and (b) Sentinel-2	71
Figure 4.10 Mean index values for endmember classes Bare 1, Bare 2, Grassy, Shrubs and Trees for (a) Landsat 8 and (b) Sentinel-2	71
Figure 4.11 Mean spectral signatures for Sentinel-2 for the (a) growing season and (b) dry season	72
Figure 4.12 Mean index values for Sentinel-2 for the (a) growing season and (b) dry season.....	72
Figure 4.13 Model-estimated FVC for the Fort Beaufort study area for 2018	73
Figure 4.14 Model-estimated FVC for the Cedarville study area for 2019	74
Figure 4.15 Visual comparison of (a) model-estimated FVC to (b) aerial imagery	74
Figure 4.16 Visual comparison of (a) aerial imagery, (b) model-estimated FVC at 30 m resolution, (c) resampled model-estimated FVC at 100 m resolution and (d) CDLC FVC at 100 m resolution	76
Figure 4.17 Comparison of model-estimated FVC to in-situ field data using Landsat 8 (grey) and Sentinel-2 (white) in terms of (a) RMSE and (b) MAE	77
Figure 4.18 Comparison of model-estimated FVC to in-situ field data using pre-developed mean spectral signatures (grey) and mean spectral signatures developed from field points (white) in terms of (a) RMSE and (b) MAE	77
Figure 5.1 RBM input and intermediary parameters.....	84
Figure 5.2 Comparison of relief factor (xrelief) for (a) Fort Beaufort and (b) Cedarville for 2018	86
Figure 5.3 Comparison of ϵ_x values (kg DM MJ ⁻¹) used by different GPP models in literature	87

Figure 5.4 GC calculation diagram	89
Figure 5.5 LTGC map grazing zones for (a) Fort Beaufort and (b) Cedarville	92
Figure 5.6 Comparison of (a) Fort Beaufort and (b) Cedarville study areas of (1) aerial imagery (2) model-estimated GPP (10 m), (3) CGDMP GPP (300 m) and (4) MOD17A2 GPP (500 m).....	93
Figure 5.7 Comparison of model-estimated GPP to MOD17A2 and CGDMP products using (a) RMSE and (b) MAE.....	94
Figure 5.8 Comparison of (a) LTGC map (2016) to (b) majority model-estimated GC (2016) for the Fort Beaufort study area	96
Figure 5.9 Comparison of (a) LTGC map (2016) to (b) majority model-estimated GC (2016) for the Cedarville study area	97
Figure 5.10 Frequency histograms for Fort Beaufort LTGC zones (a) FB1, (b) FB4, (c) FB11 and (d) FB14	99
Figure 5.11 Frequency histograms for Cedarville LTGC zones (a) C2, (b) C4, (c) C8 and (d) C11	100
Figure 5.12 Comparison of (a) poor grazing conditions to (b) ideal grazing conditions within LTGC zone FB3	101
Figure 5.13 Mean monthly rainfall (mm) and mean GC (ha/LSU) for (a) Fort Beaufort and (b) Cedarville).....	102

APPENDICES

- Appendix A: PAR calculation as defined by Swift (1976)
- Appendix B: Long Term Grazing Capacity Map for South Africa (2016)
- Appendix C Regression relationship between NDVI and fAPAR
- Appendix D Image acquisition dates of multi-temporal stacks for FVC estimation
- Appendix E Grazing capacity frequency histograms for Fort Beaufort and Cedarville
- Appendix F GrazeEngine example user interface screenshots

ACRONYMS AND ABBREVIATIONS

aDMP	Annual dry matter production
AET	Actual evapotranspiration
AI	Artificial intelligence
AMEE	Automated morphological endmember extraction
aNPP	Annual net primary production
APAR	Absorbed photosynthetically active radiation
API	Application programming interface
App	Application
ASB	Aboveground standing biomass
BETHY	Biosphere Energy Transfer Hydrology
BLM	Bureau of Land Management
BPLUT	Biome Parameter Lookup Table
BS	Bare soil
BW	Bandwidth
C/A	Coastal/aerosol
CDLC	Copernicus Dynamic Land Cover
CD: NGI	Chief Directorate: National Geospatial Information
CGDMP	Copernicus Gross Dry Matter Production
CHIRPS	Climate Hazards Group Infrared Precipitation with Stations
C/N	Carbon/nitrogen
CSIR	Council for Scientific and Industrial Research
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CW	Central wavelength
DBSI	Dry bare-soil index
DEA	Department of Environmental Affairs
DEM	Digital elevation model
DM	Dry matter
DOI	Department of Interior
DPM	Disc pasture meter
EC	Eddy covariance
EM	Electromagnetic
ET	Evapotranspiration

EVI	Enhanced vegetation index
FAO	Food and Agricultural Organisation
fAPAR	Fraction of absorbed photosynthetically active radiation
FVC	Fractional vegetation cover
GAE	Google App Engine
GC	Grazing capacity
GEE	Google Earth Engine
GEOGLAM	Group on Earth Observations Global Agricultural Monitoring
GPP	Gross primary production
HWSD	Harmonised World Soil Database
IDE	Integrated development environment
InSol	Potential insolation
LAI	Leaf area index
LC	Land cover
LCLU	Land cover/land use
LiDAR	Light detection and ranging
LSMM	Linear spectral mixture model
LST	Land surface temperature
LSU	Large stock unit
LTGC	Long Term Grazing Capacity
LUE	Light use efficiency
MAE	Mean absolute error
MODIS	Moderate Resolution Imaging Spectroradiometer
MSAVI	Modified soil adjusted vegetation index
NDVI	Normalised difference vegetation index
NEE	Net ecosystem exchange
NEP	Net ecosystem production
NIR	Near-infrared
NLC2017/18	National Land Cover 2017/2018
NPP	Net primary production
NPV	Non-photosynthetic vegetation
NRCS	Natural Resources Conservation Service
PAR	Photosynthetically active radiation
PET	Potential evapotranspiration
PPI	Pixel purity index

PV	Photosynthetic vegetation
RAP	Rangeland Analysis Platform
RAPP	Rangeland and Pasture Productivity
RBM	Regional Biosphere Model
RCMRD	Regional Centre for Mapping of Resources for Development
RDST	Rangeland Decision Support Tool
RE	Red edge
RGB	Red green blue
RMSE	Root mean square error
RS	Remote sensing
SAEON	South African Environmental Observation Network
SAR	Synthetic aperture radar
SAVI	Soil adjusted vegetation index
SMA	Spectral mixture analysis
SRTM	Shuttle Radar Topography Mission
SVAT	Soil-vegetation-atmosphere transfer
SWIR	Shortwave infrared
TEM	Terrestrial Ecosystem Model
TIR	Thermal infrared
USDA	United States Department of Agriculture
VCF	Vegetation Continuous Fields
VCI	Vegetation condition index
VI	Vegetation index
WPE	Woody plant encroachment
WV	Water vapour

CHAPTER 1: INTRODUCTION

Grasslands are one of the most important land uses in South Africa and are not only critical ecosystem service providers, but also play a vital role in water production and agriculture (Driver et al. 2012). Despite their importance, between 60 to 80% of South Africa's grasslands have been irreversibly transformed (Little, Hockey & Jansen 2015), and the biome has been identified as critically endangered (Wang et al. 2017). This grassland transformation can lead to land degradation, which consequently has a detrimental effect on grassland productivity. Such a decline in grassland productivity poses challenges for the livestock industry, a major food provider of the South African agricultural sector, which relies heavily on managed and indigenous grasslands for grazing of sheep, goat and cattle. To ensure sustainable economic and environmental management of grasslands, the livestock industry and producers require regular, high-quality condition and productivity data to support decision-making and to prevent degradation. This study thus focuses on the problem of grassland degradation, the detrimental effects it has on productivity and the consequent need for sustainable condition and productivity management of grasslands and rangelands.

1.1 GRASSLAND CONVERSION AND DEGRADATION

Grasslands can be defined as either cultivated or un-improved. Un-improved grasslands are composed of natural vegetation with few or no introduced plant species, where grazing, resting and fire are the dominant management actions that can be applied (Suttie, Reynolds & Batello 2005). Cultivated or improved grasslands (pastures), however, consist of planted grasses and legumes that are adapted for specific livestock and are managed consistently through seeding, mowing, fertilisation and irrigation (Suttie, Reynolds & Batello 2005). Grasslands can also be defined based on the grazing system implemented: traditional vs. commercial. Traditional grazing systems are typically aimed at subsistence, whereas commercial grazing systems are generally large-scale and aim to produce livestock products for the commercial market (Reynolds & Frame 2005). This study focuses predominantly on un-improved grasslands (implementing both traditional and commercial grazing systems), as these areas are more vulnerable to uncontrolled land conversion, subsequent land degradation and the consequent decline in grassland productivity. Further uses of the term 'grasslands' can thus be interpreted as referring to un-improved grasslands.

Being one of South Africa's most productive landscapes (Little, Hockey & Jansen 2015), grasslands are ideal land for grazing and therefore play a vital role in the South African livestock industry. In comparison to field crops and horticulture, livestock products have increased from 42

to 50.6% of South Africa's gross agricultural value from 2000 to 2018, establishing livestock production as a substantial contributor to food security (Directorate Statistics and Economic Analysis 2018). Livestock farming is also a major employer (Meissner, Scholtz & Palmer 2014) and plays a vital role in the local socio-economy by providing valuable commercial economic opportunities (Botha et al. 2014). Although the livestock industry relies heavily on the conservation of grasslands for commercial success, it is also one of the major contributors to the conversion and degradation of grasslands.

Large-scale land conversion of grasslands is a result of both anthropogenic factors (forestry, mining, dryland cultivation, urban expansion and over-grazing due to poor management), as well as natural causes (alien plant invasion and climate variation) (Le Maitre et al. 2016), and can lead to the irreversible degradation of the landscape. Land degradation is characterised by soil erosion, desertification and loss of biodiversity (Andrade et al. 2015) and can ultimately transform once arable grasslands into unproductive stretches of land. The implications of such degradation in productivity include the death of livestock, food shortage and famine resulting in various subsequent economic challenges for the agricultural industry (Kwon et al. 2016). One of the main drivers of large-scale grassland conversion and degradation is woody plant encroachment (WPE) (Shroder et al. 2016), defined as the proliferation of woody plants in grasslands, savannas and rangelands (Archer et al. 2017). The woody components can be either non-native species introduced purposefully or accidentally, or native species that have increased unexpectedly and expanded beyond their previous geographic range. A single main cause for increase in WPE is difficult to identify as proliferation occurs due to a range of interrelated and interacting factors that vary across tropical, arid, arctic and humid climates (Archer et al. 2017). The most prominent factors include overgrazing and browsing by livestock, fire frequency and intensity, climate change, soil properties and increases in atmospheric CO₂. Overgrazing is of particular interest to this study as the proliferation of woody components has historically coincided with global intensification of livestock grazing (Asner et al. 2004). The introduction of high quantities of grazers result in a decrease in fine fuels in the form of grasses. Consequently, less periodic fires are stimulated to allow the suppression and control of woody plants such as trees and shrubs. The outcome is increased development of woody plant communities and a decrease in grass cover, which threatens both the potential productivity of the grassland for livestock, as well as the ecosystems' vulnerability to grassland degradation in the form of wind and soil erosion, loss of biodiversity and desertification. The occurrence of WPE is therefore often associated with livestock grazing and the two interrelating factors should both be considered when investigating the productivity of un-improved grasslands such as savannas and rangelands.

Grassland conversion and degradation is a growing issue that particularly impacts the livestock industry. Sustainable grassland management systems are required to allow the identification, analysis and potential prevention of degradation occurrences and to provide quantification of the detrimental impacts on grassland productivity.

1.2 REMOTE SENSING AND GRASSLAND PRODUCTIVITY

Extensive research has been done assessing the degradation and productivity of grasslands using remote sensing (RS) and geospatial techniques (Adjorlolo & Botha 2015; Del Grosso et al. 2018; Le Houerou, Bingham & Skerbek 1988; Marsett et al. 2006; Palmer et al. 2016; Seaquist, Olsson & Ardö 2003; Vickery 1972; Zhang et al. 2017). The research indicates that grassland productivity is typically derived with the use of field-based methods (Bransby & Tainton 1977; Guevara et al. 1997; Zambatis et al. 2006) or RS-based primary production models (Palmer et al. 2016; Running et al. 2000; Seaquist, Olsson & Ardö 2003; Zhang et al. 2017; Zhu et al. 2018), where primary production is measured to serve as a precursor to grassland productivity estimates such as grazing capacity (GC). Biophysical characteristics of vegetation, such as normalised difference (NDVI), leaf area index (LAI), fraction of absorbed photosynthetic radiation (fAPAR), evapotranspiration (ET) and aboveground standing biomass (ASB) are extracted and used to describe energy and mass fluxes linked to landscape condition and productivity (Jiménez-Muñoz et al. 2009).

Fractional vegetation cover (FVC) is one of the main, but often overlooked, biophysical parameters relating to landscape surface processes and plays a vital role in deriving grassland condition and productivity (Guerschman et al. 2009). It involves determining the proportional area of green (e.g. leaves), dead (e.g. wood) and bare (e.g. soil) groundcover a pixel consists of based on various spectral characteristics of the pixel (Jiménez-Muñoz et al. 2009). This essentially allows discriminating between productive (grass) and unproductive (woody branches, bare soil) groundcover within a pixel, which in turn is potentially useful for deriving subsequent grassland productivity estimation.

Estimating an accurate FVC is a complex process, as it involves extracting sub-pixel information from temporally, spatially and spectrally variable phenomena such as vegetation (Li et al. 2014). Jimenez Munoz (2009) highlights the three most frequent techniques in literature for retrieving FVC: vegetation indices (VIs), spectral mixture analysis (SMA) and machine learning. In southern Africa, both vegetation indices (Scanlon et al. 2002) and spectral unmixing (Gessner et al. 2013), as well as a combination of these approaches (Sankaran et al. 2005), have been used to estimate FVC. A woody fractional canopy cover for South Africa was also developed by the Council for Scientific and Industrial Research (CSIR) Ecosystems Earth Observation Unit as part of the Carbon Sinks Atlas project administered by the Department of Environmental Affairs (DEA)

(Department of Environmental Affairs 2017). The 100 m and 1 km savanna woody fractional cover was created for the year 2011 with the use of Synthetic Aperture Radar (SAR) and Light Detection and Ranging (LiDAR) (Naidoo et al. 2016).

These studies all required either hyperspectral data, very high resolution imagery or very expensive active sensor technology (SAR, LiDAR), which introduces challenges with regards to accessibility, availability and affordability. Studies addressing this challenge have been performed in Australia, where rangelands and grasslands have very similar vegetational structure and climatic conditions to that of South Africa. Of particular interest is the fractional cover product developed by AusCover, which consists of an FVC layer for Australia produced from Landsat 8 imagery using SMA (Scarth, Roder & Schmidt 2010). In contrast to the previously mentioned studies, this study makes use of medium resolution satellite imagery to produce a fractional cover layer, thus eliminating the need for expensive hyperspectral or very high resolution data to produce accurate FVC results. The methodology has recently been adapted to determine FVC using Sentinel-2 imagery as well, which combines the advantages of shorter temporal resolution and finer spatial resolution. This approach proves promising, as access to both good spatial and temporal resolution is essential when analysing grassland productivity.

The investigated research confirms the success of using RS as an effective tool for observing the distribution and evolution of FVC and its potential use as an indicator of grassland degradation and productivity. FVC is a valuable biophysical parameter that can potentially improve productivity estimations if combined with current primary production models for GC calculations. Further research in applying these techniques to grassland condition monitoring could thus prove valuable for conservation planning and sustainable agricultural planning and aid in developing better grassland management systems.

1.3 PROBLEM FORMULATION

Although extensive research has been done assessing FVC and grassland productivity using RS and geospatial techniques, current literature and their derived products face both technological and scientific challenges when using them in the context of sustainable grassland and rangeland management. The majority of existing research follows a traditional Earth observation approach which entails using desktop hardware and software, downloaded imagery and extensive RS and geoprocessing experience and expertise. This limits the use by non-expert users involved in rangeland management, e.g. farmers. With respect to FVC estimation, the majority of studies, specifically those performed in southern Africa, typically use hyperspectral or very high resolution data to produce results. Current approaches are thus limited in providing easily accessible land condition data to the agricultural industry and local farmers. These approaches typically produce

results in the format of a country-wide static map using desktop-based hardware and software, which is useful when analysing and exploring grassland condition and productivity trends over time. However, for the purpose of effective management and continuous validation and calibration of production models, results must also be available at a local, in-field scale i.e. via a web or mobile application. A potential solution is the use of geoprocessing cloud platforms such as Google Earth Engine (GEE) which provides geoprocessing functionality at various scales, thus allowing continuous, dynamic results at both a national and local, in-field scale.

Scientifically, current RS-based grassland productivity approaches do not take into account the effect of vegetation structure (e.g. trees, shrubs, grasses) on productivity dynamics, e.g. the impact of increased woody components on GC. Models depict areas of tree cover as very high primary production and thus favourable GC, whereas in reality the plantations, forests and shrub-encroached fields are not suitable for cattle grazing. A gap can thus be identified, as there is no existing comprehensive model that explains both the coexistence and relative productivity of groundcover components across the grassland biome.

To address the identified gaps, an approach must be developed that efficiently and accurately estimates FVC and integrates the product in a productivity model to provide dynamic, continuous grassland condition and productivity estimations at various spatial scales. Such an approach must make use of publicly accessible, free satellite imagery, as demonstrated by Scarth, Roder & Schmidt (2010), to essentially provide the agricultural industry and local farmers with consistent and high quality grassland condition data, thus aiding in the sustainable monitoring and management of grasslands.

To address the research problem, the following research questions were formulated:

1. How can FVC be dynamically estimated for grasslands using medium resolution imagery?
2. How can FVC aid in determining accurate grassland productivity estimates?
3. How can grassland condition and productivity dynamics be disseminated at various spatial scales using evolving geoprocessing technologies?

1.4 RESEARCH AIM AND OBJECTIVES

This study aims to develop an FVC-integrated productivity model that estimates grassland productivity and condition at various spatial scales using Google Earth Engine.

To achieve the research aim, the following objectives have been set:

1. Carry out a literature review to investigate existing approaches for calculating grassland productivity and FVC, current grassland and rangeland management practices and tools available for geoprocessing cloud computing.
2. Develop a sampling scheme and collect field data on relevant biophysical parameters relating to FVC and productivity that can be used to assess and calibrate the model.
3. Estimate an FVC for a selected study area, with a possibility of extending to the rest of South Africa, using a suitable method identified in literature.
4. Develop a model that integrates FVC with productivity calculations to produce grassland condition and productivity estimates.
5. Develop and implement a web application that can geo-locate the user, estimate an FVC from satellite imagery, implement the FVC in a productivity model and give a resulting estimation of grassland condition and productivity.
6. Quantitatively and qualitatively evaluate the model and its resulting estimates.
7. Synthesise the results of the research analysis to make recommendations on modelling grassland condition and productivity using RS-based approaches.

1.5 STUDY AREA

The study area consists of two study sites, Fort Beaufort and Cedarville, both situated in the Eastern Cape Province of South Africa (Figure 1.1). The study sites are delineated based on a collection of quaternary catchments that expand the extent of the grasslands and rangelands in the region.

The Fort Beaufort study area is situated in the southern part of the Eastern Cape and spans the Q92D, Q92E, Q92G, Q94E and Q94F catchments. The altitude ranges from 198 to 1520 m above sea level and the terrain is characterised by a small plateau and low hills. The Fort Beaufort study area covers an area of 2381.9 km² and has a semi-arid climate (Conradie 2012). The mean annual rainfall of the area according to Schulze & Lynch (2007) is 400 mm, however recent drought conditions have likely resulted in a decrease. The mean monthly rainfall in the growing, wet season (November to April) is 33 mm, with mean daily temperatures of 22°C. The mean monthly rainfall in the dry season (May to October) is 15 mm, with mean daily temperatures of 14°C. The area is made up of the Albany Thicket, Drakensberg Grassland, Sub-Escarpment Grassland and Sub-Escarpment Savanna bioregions and evidently contains a wide range of different types of vegetation (Mucina & Rutherford 2006). The area has been subjected to high levels of grassland

conversion due to drought, erosion and woody plant encroachment and the landscape is very heterogeneous.

The Cedarville study area is located close to the KwaZulu-Natal border in the north-east part of the Eastern Cape, comprising the T31F and T31G catchments. The altitude ranges from 1500 to 2233 m above sea level, which is notably higher than Fort Beaufort. The valley consists of a large, flat plain, which is subject to flooding and marsh-like conditions from the Mzimvubu River during heavy rainfall periods. The Cedarville study area covers an area of 8.1 km² and has a temperate climate, with dry winters and long, cool summers (Conradie 2012). The mean annual rainfall of the area is 800 mm (Schulze & Lynch 2007), of which most occurs during the summer. The mean monthly rainfall in the growing wet, season (November to April) is 90 mm, with a mean daily temperature of 18°C. The mean monthly rainfall in the dry season (May to October) is 20 mm, with a mean daily temperature of 12°C. The area only consists of one bioregion, Sub-Escarpment Grassland (Mucina & Rutherford 2006). Grassland conditions can be described as pristine, lush and green and the landscape is very homogenous. Figure 1.1 shows the (a) Fort Beaufort and (b) Cedarville study area.

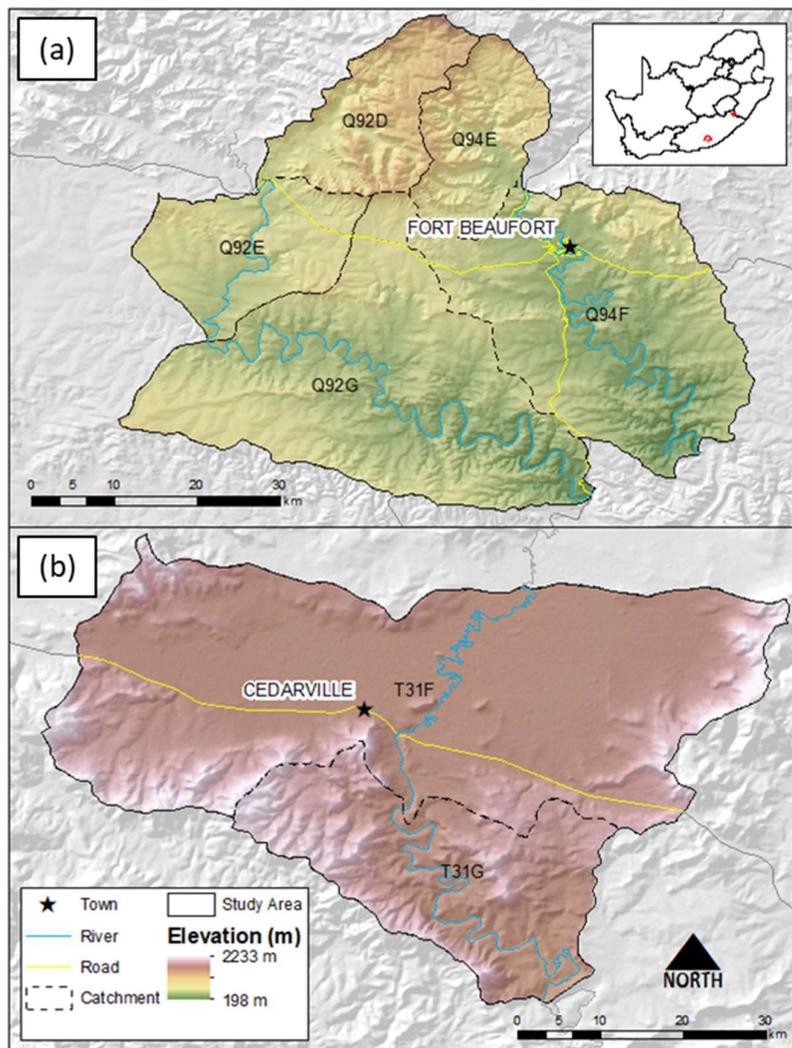


Figure 1.1 The two study areas, consisting of (a) Fort Beaufort and (b) Cedarville

Both areas have been subject to drought since 2015, however, the Cedarville region has since experienced significant increases in rainfall while the Fort Beaufort study area has not. The two study areas essentially represent opposite grazing conditions, as they contrast with respect to vegetation type and distribution, vegetation structure, degree of grassland conversion, rainfall etc. This allowed exploring the techniques and methods discussed in this research in different scenarios, providing a better perception of the transferability of this approach.

1.6 RESEARCH METHODOLOGY & DESIGN

The purpose of this study is to assist in establishing and improving sustainable grassland management systems by providing the livestock industry with continuous, dynamic estimates of grassland condition and productivity at various spatial scales. The real-world problem in this research relates to the increasing impact of grassland degradation in South Africa and its negative impact on land use productivity and consequently the livestock industry. Commercial and local

livestock farmers require country-wide, as well as in-field, grassland condition data for identifying, monitoring, managing and preventing potential grassland degradation. The research problem requires developing a productivity model that estimates and integrates FVC to provide accurate, dynamic, continuous land condition and productivity estimations. The research problem also includes providing these estimations at various spatial scales with the use of a web application. This provides the livestock industry with the means to develop sustainable productivity levels and to identify regions of potential degradation and food shortage, which is significant with regards to food security and the economy.

The answer to the first research question contributes to providing an FVC for South Africa using accessible, available and affordable data, unlike the approaches explored in current literature. A solution to the second question provides improved grassland condition and productivity estimations by developing a production model that integrates FVC into productivity calculations. Addressing the third question aids in providing quality, continuous and dynamic land condition data to the livestock industry at various spatial scales using geoprocessing cloud platforms and web application technologies, thus ensuring the establishment of more sustainable grassland management systems.

This study is evaluative and model-building in nature as it involves evaluating an existing method of FVC estimation for the purpose of developing a comprehensive model that explains both the coexistence and relative productivity of groundcover components across the grassland biome. The research approach is deductive, as the study makes use of existing algorithms and theories to estimate FVC and subsequent grassland productivity derivatives. The data used in this research is empirical and quantitative, comprising of digital satellite imagery and in situ field measurements of relevant biophysical variables. The resulting estimations of the application were assessed both quantitatively and qualitatively. The quantitative analyses involved comparing the FVC results and productivity estimates of the application to field validation data and existing products, whereas the qualitative evaluation consisted of visual inspection of results.

Figure 1.2 shows the research design for this study. This chapter (Chapter 1) introduced the research problem and provided background on grassland degradation and the need for sustainable grassland management in South Africa. Chapter 1 also described the potential of providing multi-scale grassland condition and productivity data to the livestock industry using web applications for improving the monitoring, management and prevention of potential grassland degradation. The aim and objectives were defined and the layout of the thesis (in the form of the research design diagram) discussed.

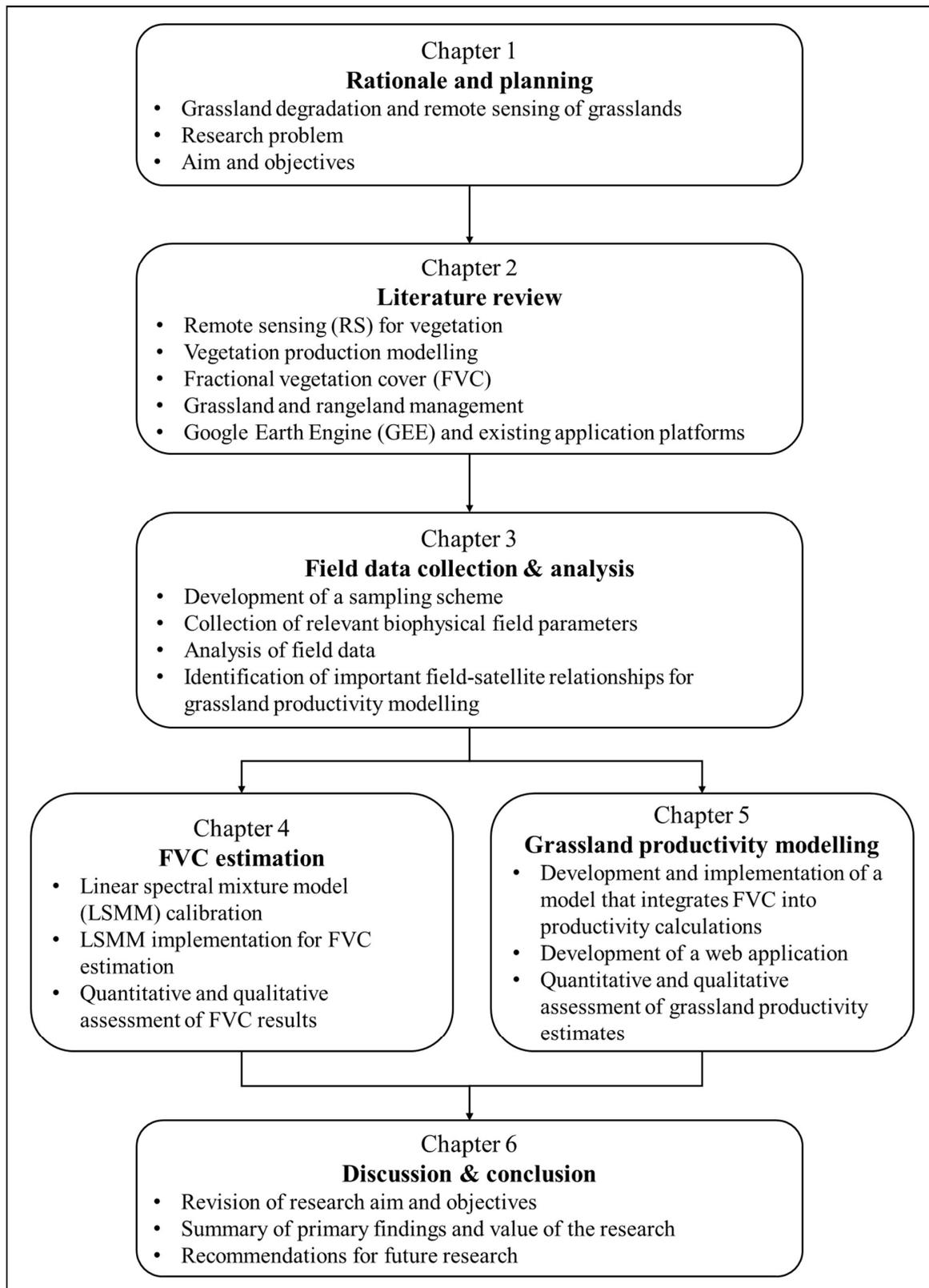


Figure 1.2 Research design and thesis structure

Chapter 2 provides an overview of the relevant literature with respect to grassland degradation, its impact on productivity and current research on grassland condition and productivity estimation

using RS. A discussion on the potential of geoprocessing cloud computing platforms i.e. Google Earth Engine is discussed, as well as the different appropriate platforms available for application development. Chapter 3 provides details on developing a sampling scheme to collect the appropriate field data for validation and calibration purposes, as well as describe the field data acquisition and subsequent analysis. Relevant biophysical field-satellite relationships are identified for use in Chapter 4 and 5. Chapter 4 involves the estimation of an FVC using an appropriate linear spectral mixture model (LSMM) and the evaluation of FVC results, while Chapter 5 investigates developing a productivity model that implements these FVC estimations. Chapter 5 further quantitatively and qualitatively assesses the grassland productivity estimations and describes the development of a web application that practically implements the productivity model to provide grassland condition and productivity estimations at various spatial scales. The findings of Chapters 3, 4 and 5 are summarised in Chapter 6, where conclusions are drawn, research aim and objectives are revisited, the value and limitations of the research discussed and recommendations for further research presented.

CHAPTER 2: DETERMINING GRASSLAND CONDITION AND PRODUCTIVITY

This chapter reviews the data and methods associated with grassland condition and productivity estimation, thereby addressing Objective 1. A brief background of the fundamentals of remote sensing (RS) is given, followed by an in-depth review of RS-based techniques for determining vegetation production in terrestrial ecosystems. Fractional vegetation cover (FVC) estimation and products are discussed, as well as current approaches for grassland and rangeland management. The chapter concludes with current cloud geocomputing technologies, as well as existing web applications for grassland management. These could aid in addressing current challenges in the discussed literature.

2.1 REMOTE SENSING

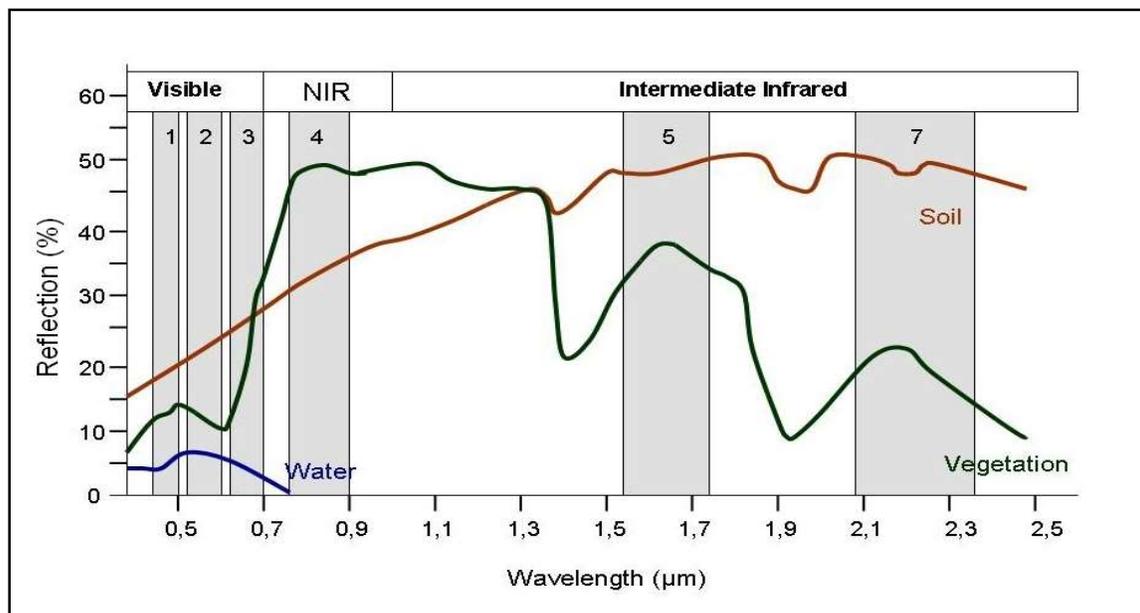
Remote sensing (RS) is generally defined as the acquisition of information about an object without making direct, physical contact with it (Campbell 2002). For this research, it specifically refers to obtaining information about the earth using passive, optical satellite sensors for measuring the electromagnetic (EM) radiation reflected from the earth's surface. The main source of EM energy is the sun, which produces a full spectrum of EM radiation, i.e. the EM spectrum. As the levels of energy reflection and absorption vary from object to object, these interactions can be recorded and used to acquire knowledge of the characteristics of the earth's surface (Campbell 2002).

RS has a wide range of applications, including flood extent mapping, sea ice monitoring, land use/land cover (LULC) change and fire scar mapping. One of the most common and useful application areas for RS is vegetation monitoring and modelling. RS is used widely for mapping extents of vegetation activity, investigating vegetation-climate interactions, modelling carbon sequestration, identifying forest degradation and deforestation, predicting crop production etc. Vegetation has a very unique response to EM energy, which is often characterised using spectral signatures (Section 2.1.1) and indices (Section 2.1.2). A large body of work exists on methodologies and algorithms for quantifying vegetation dynamics and structure using these derived signatures and indices. The techniques focussed on in this research is vegetation productivity modelling and spectral mixture analysis (SMA), discussed in Section 2.2 and Section 2.3 respectively. As this section provides only a brief overview of the core RS principles relating to vegetation dynamics, refer to Campbell (2002) for more in-depth theory relating to RS history, processes and techniques.

2.1.1 Surface reflectance properties and spectral signatures

The most relevant regions of the EM spectrum for RS is the visible, near-infrared (NIR), shortwave infrared (SWIR) and thermal infrared (TIR) regions. The visible region corresponds to energy that the human eye can observe (red, green and blue), with wavelengths ranging from 0.4 to 0.7 micrometres (μm). The NIR region ranges from wavelengths 0.7 to 1.2 μm and is not visible to the human eye. The 1.2 to 3 μm wavelengths correspond to the SWIR region, while wavelengths longer than 3 μm represent the TIR region.

The different regions of the EM spectrum can interact with the Earth's surface in three ways: reflection, absorption and transmission. Reflection occurs when light is redirected when interacting with an object, while transmission refers to EM energy passing through an object without attenuation. The amount of reflection, absorption and transmission depends on the characteristics of the surface, the wavelength of the EM energy (e.g. visible, NIR, SWIR or TIR) and the angle of incoming radiation. Different surface objects thus have different surface reflectance properties, resulting in a unique spectral signature for all material on Earth (Campbell 2002). An object's spectral signature can be visualised using spectral reflectance curves, which illustrates the percentage reflectance as a function of EM radiation wavelength. Figure 2.1 shows the typical spectral signatures of soil, vegetation, and water.



Adapted from Siegmund and Menz (2005)

Figure 2.1 The spectral signatures of soil, vegetation and water.

The surface reflectance properties and spectral signature of vegetation are influenced by leaf pigment, cell and canopy structure and plant physiology (Chuvieco & Huete 2009). Chlorophyll, the pigment found in leaves, strongly absorbs EM radiation in the blue (0.45 μm) and red (0.65

μm) wavelengths, thus these regions are known as the chlorophyll absorption bands (Chuvieco & Huete 2009). Absorption decreases at $0.55 \mu\text{m}$, i.e. the green wavelength, thus explaining the green appearance of vegetation. Cell and canopy structure strongly reflects radiation in the NIR region (0.75 to $1.35 \mu\text{m}$) (Campbell 2008), resulting in a sharp increase in reflection between 0.65 and $0.76 \mu\text{m}$ known as the red edge (RE).

These spectral signatures allow discrimination between different land cover (LC) objects during RS image analysis e.g. urban, water, soil and vegetation. Due to the unique spectral signature of vegetation, it is relatively easy to discriminate it from other LC classes using the RE. Challenges arise, however, when attempting to identify different vegetation types using solely spectral signatures, e.g. shrubs, grass and trees, as their spectral surface reflectance properties and signatures are quite similar (Price 1994). Arid and drought-stricken areas also prove difficult as dry, sparse vegetation and bare soil have very similar spectral responses (Leprieur et al. 2000). Both issues can be addressed using very high resolution or hyperspectral imagery (Guerschman et al. 2009; Zhang et al. 2012), which is often expensive and time-consuming to process. RS approaches thus often also include the use of spectral indices during analysis to provide additional features for discriminating between surface objects.

2.1.2 Spectral indices

Spectral indices are a method of image transformation and consist of combinations of spectral reflectance at two or more wavelengths (Jackson & Huete 1991). Spectral indices have been developed for water, soil, urban and fire scar applications, with the most common group being vegetation indices (VIs). A thoroughly documented and well-known VI is the normalised difference vegetation index (NDVI) introduced by Tucker (1979):

Equation 2.1

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)}$$

where $NDVI$ is the normalised difference vegetation index;
 NIR is the near-infrared band pixel value; and
 RED is the red band pixel value.

NDVI is a general indicator of plant “vigour” and is expressed using values ranging from -1 to 1. Generally, higher values indicate healthier vegetation and lower values correlate to less or no vegetation. Although NDVI is widely used and evaluated in literature and practice, it is subject to a number of limitations. NDVI tends to “saturate” (Liu & Huete 1995), which involves the loss of sensitivity to change in the amount of vegetation in very green, dense, high biomass conditions.

NDVI is also very sensitive to the effect of soil reflectance, limiting its use in arid and semi-arid areas consisting of sparse vegetation with exposed rock and soil. These limitations led to the development of several other VIs, including the enhanced vegetation index (EVI) and various soil adjusted vegetation indices (SAVI).

EVI, a modification of the NDVI, aims to improve vegetation mapping in areas of high biomass where NDVI “saturation” might occur (Liu & Huete 1995). Equation 2.2 describes the EVI calculation, where $C_1 = 6.0$, $C_2 = 7.5$ and $L = 1$ to produce values ranging from 0.0 to 1.0:

Equation 2.2

$$EVI = 2.5 \times \frac{(NIR - RED)}{(NIR + C_1 \times RED - C_2 \times BLUE + L)}$$

where

- EVI is the enhanced vegetation index;
- NIR is the near-infrared band pixel value;
- RED is the red band pixel value;
- C_1, C_2 are coefficients to correct aerosol scattering;
- $BLUE$ is the blue band pixel value; and
- L is the soil adjustment factor.

To address the limitation of NDVI in areas with high degrees of exposed soil (Qi et al. 1994), thus potentially improving discrimination between vegetation and bare ground, a wide range of soil indices have been proposed in literature. A popular choice is the modified soil-adjusted vegetation index (MSAVI). The newer version of the index, MSAVI2, allows the implementation of the index without specifying a soil brightness correction factor, eliminating the need for prior knowledge of the area’s vegetation cover. MSAVI2 values range from -1.0 to +1.0 and is defined by the following equation (Qi, Kerr & Chehbouni 1994):

Equation 2.3

$$MSAVI2 = \frac{(2 \times NIR + 1 - \sqrt{(2 \times NIR + 1)^2 - 8 \times (NIR - RED)})}{2}$$

where

- $MSAVI2$ is the modified soil adjusted vegetation index 2;
- NIR is the near-infrared band pixel value; and
- RED is the red band pixel value.

Another newly developed index that also improves discrimination between sparse vegetation and bare soil is the dry bare-soil index (DBSI) (Rasul et al. 2018). The index was specifically developed to identify bare areas in dry climates (Rasul et al. 2018) and can improve differentiation

between dry vegetation and bare soil in areas impacted by drought. DBSI values range from -2.0 to +2.0, with higher values corresponding to barer areas:

Equation 2.4

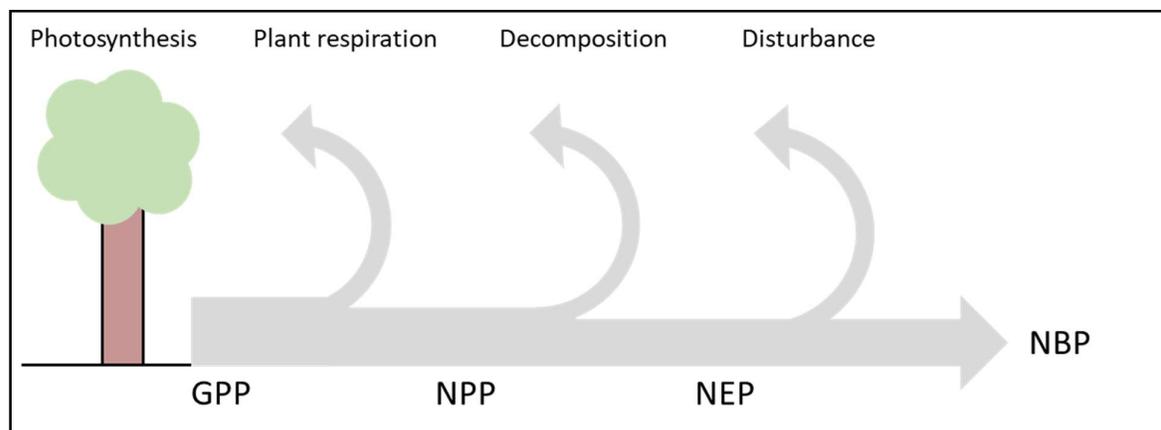
$$DBSI = \frac{(SWIR - GREEN)}{(SWIR + GREEN)} - NDVI$$

where $SWIR$ is the short wave infrared band pixel value;
 $GREEN$ is the green band pixel value; and
 $NDVI$ is the normalised difference vegetation index value.

Spectral indices play an important role in characterising ground cover features from spectral information. They are implemented in a wide range of applications, including LC classification, vegetation modelling, spectral unmixing, soil mapping etc. Although there are many more spectral indices defined in literature, only those most relevant to the research were discussed.

2.2 VEGETATION PRODUCTION IN TERRESTRIAL ECOSYSTEMS

Ecosystem productivity is an essential ecological variable and can serve as an indicator of the condition of a landscape (Swinnen & Van Hoolst 2018). Different methodologies for estimating ecosystem productivity have been defined and implemented in literature and in practice. Figure 2.2 shows four common ecosystem productivity variables, namely gross primary production (GPP), net primary production (NPP), net ecosystem production (NEP) and net biome production (NBP).



Adapted from Valentini (2003)

Figure 2.2 Ecosystem productivity variables.

GPP, gross primary production; NPP, net primary production; NEP, net ecosystem production; NBP, net biome production.

GPP is the amount of chemical energy as biomass that primary producers create in a given length of time, whereas NPP can be defined as the rate at which all the plants in an ecosystem produce net useful chemical energy (Lieth 1973). NPP is thus essentially GPP minus the energy used for respiration by plants and provides an indication of the net carbon captured by living vegetation per unit area for a given time period (Hanan et al. 1998). NEP is the net amount of primary production after the costs of both autotrophic plant respiration and heterotrophic decomposition of soil organic matter have been taken into account (Valentini 2003). NBP is the amount of carbon that remains in the vegetation after, autotrophic respiration, heterotrophic decomposition and losses due to anthropogenic disturbances (Valentini 2003). GPP and NPP are the most widely used in RS and earth observation and play a vital role in ecological, environmental and climate change monitoring (Eisfelder et al. 2013). They often serve as precursor to agronomic indicators of productivity, e.g. grazing capacity and stocking rate and have been estimated in literature at both global and regional scale using a wide variety of different primary production models.

2.2.1 Primary production models

Primary production refers to the production of chemical energy in organic compounds by living organisms (Lieth 1973) and can be modelled as GPP or NPP in various different ways. As NPP is the difference between GPP and autotrophic respiration (Figure 2.2), most primary production models provide an estimate of GPP which can subsequently be converted to NPP. NPP is often expressed on an annual basis and is thus described by the following equation (Running et al. 2000):

Equation 2.5

$$NPP = \varepsilon_{AR} \sum GPP$$

where NPP is the net primary production [$\text{g C m}^{-2}\text{year}^{-1}$];
 GPP is the gross primary production [$\text{g C m}^{-2}\text{day}^{-1}$]; and
 ε_{AR} is the autotrophic respiration conversion factor [scalar].

The prevalent approaches that utilise RS derived parameters for GPP modelling are eddy covariance (EC) models, vegetation index (VI) models, light use efficiency (LUE) models and process-based models.

2.2.1.1 Eddy covariance models

Eddy covariance (EC) models follow a micro-meteorological approach that uses direct EC flux tower measurements for primary production estimation (Jung et al. 2017; Jung et al. 2011; Liu et

al. 2016; Tramontana et al. 2016; Wei et al. 2017). Measurements are acquired from FLUXNET (Baldocchi et al. 2001), a global network of micrometeorological flux measurement sites that uses EC techniques to measure ecosystem variables like water, carbon, energy and nutrient fluxes. Net ecosystem CO₂ exchange (NEE) can be derived from measurements, which is equal to GPP minus ecosystem respiration and inorganic flows of CO₂. The towers capture data at a high temporal frequency (typically every 30 minutes) but are spatially restricted to each tower's flux footprint i.e. a few 100 m² upwind (Baldocchi et al. 2001). Machine learning techniques are thus implemented to extrapolate the GPP estimates to the appropriate scale for regional and global modelling. The direct EC measurements are often used as validation data for GPP model estimates. EC measurements are also frequently combined with other types of models, e.g. LUE or VI-based, to calibrate specific input parameters or GPP outputs (Tramontana et al. 2016; Veroustraete, Sabbe & Eerens 2002; Wu et al. 2010; Yuan et al. 2007).

2.2.1.2 Vegetation index models

Vegetation index (VI) models rely on empirical relationships between field measurements and spectral indices (Gitelson et al. 2006; Li et al. 2013; Liu, Wang & Wang 2014; Sims et al. 2008; Wu et al. 2010). Statistical correlations are built between the field-measured vegetation productivity and RS data and used to estimate vegetation productivity in other areas. The RS data is typically in the form of VIs (as discussed in Section 2.1.2). Although these models can provide accurate estimations, the established relationships are site-specific and can often not be transferred to other regions without recalibration of parameters (Liang, Li & Wang 2012). This modelling approach is also subject to certain limitations relating to VIs, e.g. saturation of indices with very high productivity and the effect of soil background on spectral reflectance values. In addition, VI-based statistical models do not include the effect of certain biophysical parameters as input, e.g. rainfall and temperature, and can thus not reflect the effect of variations in these parameters on productivity (Liang, Li & Wang 2012). These models are thus not suitable for predictive research and are only applicable when assessing existing vegetation production.

2.2.1.3 Light use efficiency models

Light use efficiency (LUE) models are data-driven, semi-empirical approaches that require limited input data compared to more complex biophysical models. They are thus widely used and several variations have been developed and implemented (Coops et al. 2010; Goetz et al. 2000; Jia et al. 2015; King, Turner & Ritts 2011; Liu 2008; Potter et al. 1993; Running et al. 2000; Veroustraete, Sabbe & Eerens 2002; Verstraeten, Veroustraete & Feyen 2006; Xiao et al. 2005; Yuan et al. 2007; Zhao et al. 2005). LUE models calculate primary production based on the light intercept concept

of Monteith (1972), where incoming radiation is linked to agricultural production through an empirical biophysical conversion factor. The Monteith (1972) approach states that vegetation growth is determined by the part of incoming solar radiation absorbed by the plant, i.e. the absorbed photosynthetically active radiation (APAR), and the ability of the plant to convert this radiation into plant biomass, i.e. LUE, as defined by Equation 2.6.

Equation 2.6

$$GPP = APAR \cdot LUE$$

where GPP is the gross primary production [$\text{g C m}^{-2}\text{day}^{-1}$];
 $APAR$ is the absorbed photosynthetically active radiation [$\text{MJ m}^{-2}\text{day}^{-1}$]; and
 LUE is the light use efficiency [g C MJ^{-1}].

According to Sellers (1985), APAR can be calculated as a product of the fraction of absorbed photosynthetically active radiation (fAPAR) and photosynthetically active radiation (PAR):

Equation 2.7

$$APAR = fAPAR \cdot PAR$$

where $APAR$ is the absorbed photosynthetically active radiation [$\text{MJ m}^{-2}\text{day}^{-1}$];
 $fAPAR$ is the fraction absorbed photosynthetically active radiation [scalar]; and
 PAR is the photosynthetically active radiation [$\text{MJ m}^{-2}\text{day}^{-1}$].

PAR represents the component of incoming solar radiation that can be absorbed by the plant chlorophyll (Monteith 1972) and can be calculated using a variety of complex models. According to literature (Potter et al. 1993; Running et al. 2000), the fAPAR can be derived from RS data using an empirical or theoretical relationship between fAPAR and NDVI. The fAPAR is strongly linked to leaf area index (LAI), defined as the one-sided green leaf area per unit ground surface area (Palmer et al. 2017). Both are important biophysical variables for vegetation modelling and they are often measured together.

Equation 2.6 and Equation 2.7 form the basis of all LUE-based primary production models. The principal difference between different LUE-based models is the way LUE is defined and calculated. Generally, a maximum gross LUE is identified and scaled based on a variety of stressors. Even though the LUE plays a critical role in GPP modelling, the various factors that define it and stressors that influence it are still not fully understood (Swinnen & Van Hoolst 2018) and contributes to the majority of uncertainty in primary production estimates. Models differ in both the way the maximum gross LUE is identified, as well as the stressors used to scale it. The maximum gross LUE can be defined based on biome (Goetz et al. 2000; King, Turner & Ritts

2011; Running et al. 2000; Veroustraete, Sabbe & Eerens 2002; Verstraeten, Veroustraete & Feyen 2006), as a function of soil nitrogen content (Coops et al. 2010) or as a global fixed value (Xiao et al. 2005; Yuan et al. 2007), where the biome-specific approach is the most widely used. Often these LUE models incorporate EC flux tower measurements to estimate or optimise LUE values (Garbulsy et al. 2010). This approach can account for some of the uncertainty in LUE models, but in turn, introduces limitations relating to the spatial restriction of flux footprints. Stressors of maximum LUE include surface and air temperature, plant age, soil moisture, cloudiness, leaf phenology, evapotranspiration, water vapour deficit, canopy water content, terrain etc. (Swinnen & Van Hoolst 2018). The choice of model thus depends heavily on access to input data for the variety of different stressors implemented by different models. In addition, model performance for a specific region will also depend on how suitable the included LUE stressors are for the climate, terrain and vegetation type of the area.

2.2.1.4 Process-based models

Process-based models are more complex compared to other approaches and attempt to model GPP using the biophysical processes of plant photosynthesis (Eisfelder et al. 2013; Jiang & Ryu 2016; Liu et al. 1997; Ryu et al. 2011; Zhang et al. 2018; Zhu et al. 2018). Melillo et al. (1993) published the first simulation results derived from a process-based, mechanistic model, the Terrestrial Ecosystem Model (TEM), which estimated the response of terrestrial NPP to elevated CO₂ and climate change. Subsequently, a multitude of process-based models have been developed and presented in literature (Jiang & Ryu 2016; Liu et al. 1997; Ryu et al. 2011; Zhang et al. 2018). These models adopt the same approximation for ecosystem structure but differ in parameterisation of carbon flows (e.g. photosynthesis, respiration, allocation, litterfall, and decomposition), which are complex functions of various biological and environmental factors. A common approach is the soil-vegetation-atmosphere transfer (SVAT) scheme, which explicitly considers the role of vegetation in affecting the ecosystem water and energy (Eisfelder et al. 2013). Similar to LUE-based models, the difference in parameterisation among the models result in a large difference in estimated carbon dynamics (Cramer et al. 1999) and inherent uncertainty that is not yet well-documented. In contrast to EC-based and VI-based models, the mechanistic structure of process-based models is suitable for making predictive estimations e.g. potential change in GPP due to climate change.

2.2.1.5 Comparison of primary production models

As good performance have been reported for all these models, the choice of primary production model is dependent on how sensitive the study area vegetation is to temporal and environmental

variations, the required scale of GPP estimates and access to necessary input data. Process-based models provide accurate results and can be used in predictive research, however, they are highly complex and require a large amount of field, meteorological and RS input data to effectively simulate plant physiological ecology principles (Eisfelder et al. 2013). VI-based models are simple and can achieve similar accuracies to complex process models (Sun et al. 2019), however, due to their empirical nature, they are often very site-specific and difficult to transfer spatially and temporally (Eisfelder et al. 2017). This is often also a challenge when modelling GPP using EC models, as accurate EC measurements are restricted to each tower's flux footprint (Baldocchi et al. 2001). EC measurements thus need to be extrapolated to regional and global scales using artificial intelligence (AI) and machine learning methods, which in itself has several challenges (Swinnen & Van Hoolst 2018). LUE-based models are theoretically more transferable than VI-based models and EC models, while also requiring fewer input parameters than process-based models. In addition, an element of mechanistic realism is incorporated in the modelling approach, which improves its potential for making predictive estimates. The determination of LUE itself, however, is a prominent source of uncertainty in GPP modelling. Despite this inherent uncertainty of LUE, it is still the most popular choice of GPP model type and is used to produce the majority of publicly available RS-generated primary production products e.g. the Moderate Resolution Imaging Spectroradiometer (MODIS) 17 GPP/NPP product (Running, Mu & Zhao 2015) and Copernicus Global Land Dry Matter Production product (Swinnen & Van Hoolst 2018).

Currently, there is no optimal modelling approach identified for GPP estimation across different ecosystems and enviro-climatic conditions (Sun et al. 2019). A recent study by Sun et al. (2019) comparing and evaluating the response of 14 GPP models to climate variability and CO₂ trends found no model performed consistently better across different ecosystems and under various external conditions. This is supported by the vast number of GPP models in literature, each developed based on specific user requirements relating to application, environment, climate region and available resources.

2.2.2 Primary production modelling in southern Africa

A large number of GPP models have been developed at global scale for various different applications. Although spatially and temporally transferable, global models provide coarse resolution results that do not capture the complex vegetation structure dynamics in savannas and woody encroached grasslands (Eisfelder et al. 2013; Richters 2005a). Only a few regional-scale models have been developed and implemented for the southern African landscape. This section discusses two such models: the Regional Biosphere Model (RBM) (Richters 2005b) and the Biosphere Energy Transfer Hydrology (BETHY) model (Niklaus, Tum & Günther n.d.).

2.2.2.1 Regional Biosphere Model

The RBM is an LUE-based model derived from Running et al. (2000)'s approach. It was developed for and implemented in north-western Namibia to describe the production and consumption of biomass by cattle, game and natural decomposition to monitor occurrences of overgrazing and degradation (Richters 2006). As the area consists of small-scale vegetation patterns with fractions of savannas, woody components, shrublands and grasslands (Richters 2005b), generally-defined biomes used in global GPP models (Running et al. 2000) do not provide satisfactory results.

As in Equation 2.6 and Equation 2.7, RBM calculates GPP as a product of LUE and APAR, which in turn is a function of fAPAR and PAR. To estimate PAR, the potential insolation (InSol) ($\text{cal cm}^{-2}\text{day}^{-1}$) is calculated according to Swift (1976) and scaled using a ratio coefficient. For more information on formulas and input components used for calculating InSol, refer to Appendix A.

In RBM, the biophysical conversion factor LUE (g C MJ^{-1}) consists of two subcomponents: maximum LUE (LUE_{max}) and effective LUE (LUE_{eff}). The LUE_{max} is derived yearly and depends on constant abiotic factors such as relief and soil, as well as the potential LUE (LUE_p) determined by vegetation type (i.e. herbaceous or woody). The LUE_{eff} is more variable and incorporates the effects of climatic conditions (water and temperature stress) on the LUE_{max} . Table 2.1 shows the formulas used to calculate the two LUE subcomponents.

Table 2.1 LUE subcomponent formulas

LUE Component	Formula	Equation
$\text{LUE}_{\text{max}}^{\text{a}}$	$\text{LUE}_{\text{max}} = \text{LUE}_p \cdot x_{\text{relief}} \cdot x_{\text{soil}}$	Equation 2.8
$\text{LUE}_{\text{eff}}^{\text{b}}$	$\text{LUE}_{\text{eff}} = \text{LUE}_{\text{max}} \cdot T_{\varepsilon} \cdot W_{\varepsilon}$	Equation 2.9

^a LUE_p , potential LUE; x_{relief} , relief factor; x_{soil} , soil factor. ^b T_{ε} , temperature stress factor; W_{ε} , water stress factor.

The LUE_{max} represents the best possible efficiency with which plants can convert absorbed sunlight (APAR) into biomass and is calculated for each grid cell using Equation 2.8. It incorporates the vegetational, topographical and pedological characteristics of the cell using a LUE_p , a relief factor (x_{relief}) and a soil factor (x_{soil}). As soil and relief are not as temporally variable, LUE_{max} is calculated annually and kept constant throughout the model year.

LUE_p is highly influenced by the metabolic pathway of carbon sequestration of the plant (i.e. C3 vs. C4 plants) (Liang, Li & Wang 2012) and thus differs between different vegetation growth forms (woody or herbaceous). In ecosystems with a co-existence of woody and herbaceous

vegetation, e.g. savannas and grasslands experiencing woody plant encroachment (WPE), it is thus particularly important to distinguish between growth forms such as grasses, shrubs and trees. LUE_p is calculated in the RBM with the aid of an FVC:

Equation 2.10

$$LUE_p(x) = \frac{\varepsilon_{bare}FVC_{bare}(x) + \varepsilon_{herb}FVC_{herb}(x) + \varepsilon_{tree}FVC_{tree}(x)}{100}$$

where $LUE_p(x)$ is the potential LUE [$g C MJ^{-1}$] for cell x ;
 ε is the biophysical conversion factor per cover type [$g C MJ^{-1}$]; and
 $FVC(x)$ is the proportion of cover type in cell x (%).

The biophysical conversion factors for bare soil (ε_{bare}), herbaceous vegetation (ε_{herb}) and woody vegetation (ε_{tree}) are based on empirical values found in literature (Seaquist, Olsson & Ardö 2003) and provided to the RBM in the form of a look-up table.

The x_{relief} and x_{soil} is determined using the concept of fuzzy logic (Zimmermann 2001), where the specific factor input is scaled based on a set of conditions to produce a unitless, real number between 0 and 1, where a higher value translates to less limitation on LUE. In contrast, Boolean logic operators produce a unitless, discrete value of either 0 or 1 depending on the input variable (Zimmermann 2001). Table 2.2 shows popular Boolean operators and their fuzzy equivalents.

Table 2.2 Boolean and corresponding fuzzy operators

Boolean Operator	Fuzzy Operator
AND(x,y)	MIN(x,y)
OR(x,y)	MAX(x,y)
NOT(x)	1 - x

x_{relief} incorporates the effects of two subcomponents, elevation and slope, which is combined using a “fuzzy OR” (Table 2.2) to ensure that an unfavourable relief (e.g. steep slope) can be compensated for by a favourable elevation (e.g. low elevation) and vice versa. Equation 2.11 and Equation 2.12 respectively show the empirical equations used to calculate the influence of elevation (h) and slope (s):

$$x(h) = -0.0005833 \cdot H + 1.5833 \quad \text{for } 1000m < H < 2200 \quad \text{Equation 2.11}$$

$$x(s) = -0.08 \cdot SL + 1.24 \quad \text{for } 3^\circ < SL < 8^\circ \quad \text{Equation 2.12}$$

where $x(h)$ is the elevation scaling factor; and
 H is the height above sea level (m).

$x(s)$ is the slope scaling factor; and

SL is the slope ($^{\circ}$).

An increase in elevation results in a smaller relief factor, which in turn translates to a reduction in LUE. This is due to the limiting effect of high elevation on plant growth (Machwitz et al. 2015). Similarly, an increase in slope steepness results in a smaller slope factor, which in turn translates to a reduction in LUE. In contrast, shallow, sloping basins and valleys result in no slope-related limitations on LUE.

x_{soil} combines the effects of two subcomponents: texture and carbon/nitrogen (C/N) ratio. As with the x_{relief} , the two scaling factors are combined using a “fuzzy OR” (Table 2.2) to ensure that an unfavourable texture (e.g. too rough texture, low water storage capacity) can be compensated for by a favourable C/N ratio (e.g. good nutrient supply). Soil texture provides a measure of the soil type’s water storage capacity (Potter et al. 1993), whereas the C/N ratio correlates with the soil nutrient supply (Richters 2005b). Only the topsoil texture and nutrient supply are evaluated, as the density and degree of soil development are usually so low that soil formation has hardly penetrated areas deeper than 1 m (Richters 2005b). Soil is divided into five texture classes according to the global texture distribution data published by Zabler (1986) and assigned a fuzzy value. Table 2.3 shows RBM texture classes and their corresponding fuzzy values. Generally, rough-textured soils have a lower water storage capacity resulting in soil-related constraints on the plant LUE.

Table 2.3 RBM soil texture classes and corresponding fuzzy values

Soil Texture	Clay (%)	Silt (%)	Sand (%)	Fuzzy Value
Rough	9	8	83	0.2
Rough/ Medium	20	20	60	0.3
Medium	30	33	37	0.4
Medium/ Fine	48	25	27	0.5
Fine	67	17	17	0.6

The dimensionless ratio of organic carbon (C) to nitrogen (N) reflects how well the essential nitrogen compounds are provided by humic substances for plants, thus providing an indirect measure of plant nutrition. The influence of soil C/N ratio is described using an empirical equation:

$$x(C/N) = -0.004 \cdot CN^2 + 0.0971 \cdot CN + 0.4$$

where $x(C/N)$ is the C/N ratio scaling factor; and

CN is the C/N ratio.

Optimal C/N ratio is 12, with higher and lower levels posing constraints for plant growth. Although it would be expected that a higher C/N ratio would improve nutrient supply, it must be noted that high C/N levels reflect a greater amount of undecomposed plant material within the soil, thus less nutrient-dense humus is available (Richters 2005b).

LUE_{eff} is determined by temperature stress (T_{ε}) and water stress (W_{ε}). As T_{ε} and W_{ε} reflect current climatic conditions, they are calculated for each time step. T_{ε} is derived from daily temperature data and water stress is calculated from actual evapotranspiration (AET) and potential evapotranspiration (PET):

$$T_{\varepsilon} = 0.8 \cdot 0.02T - 0.0005T^2 \quad \text{Equation 2.14}$$

$$W_{\varepsilon} = 0.5 + 0.5 \frac{AET}{PET} \quad \text{Equation 2.15}$$

where T_{ε} is the temperature stress;

W_{ε} is the water stress;

AET is the actual evapotranspiration; and

PET is the potential evapotranspiration.

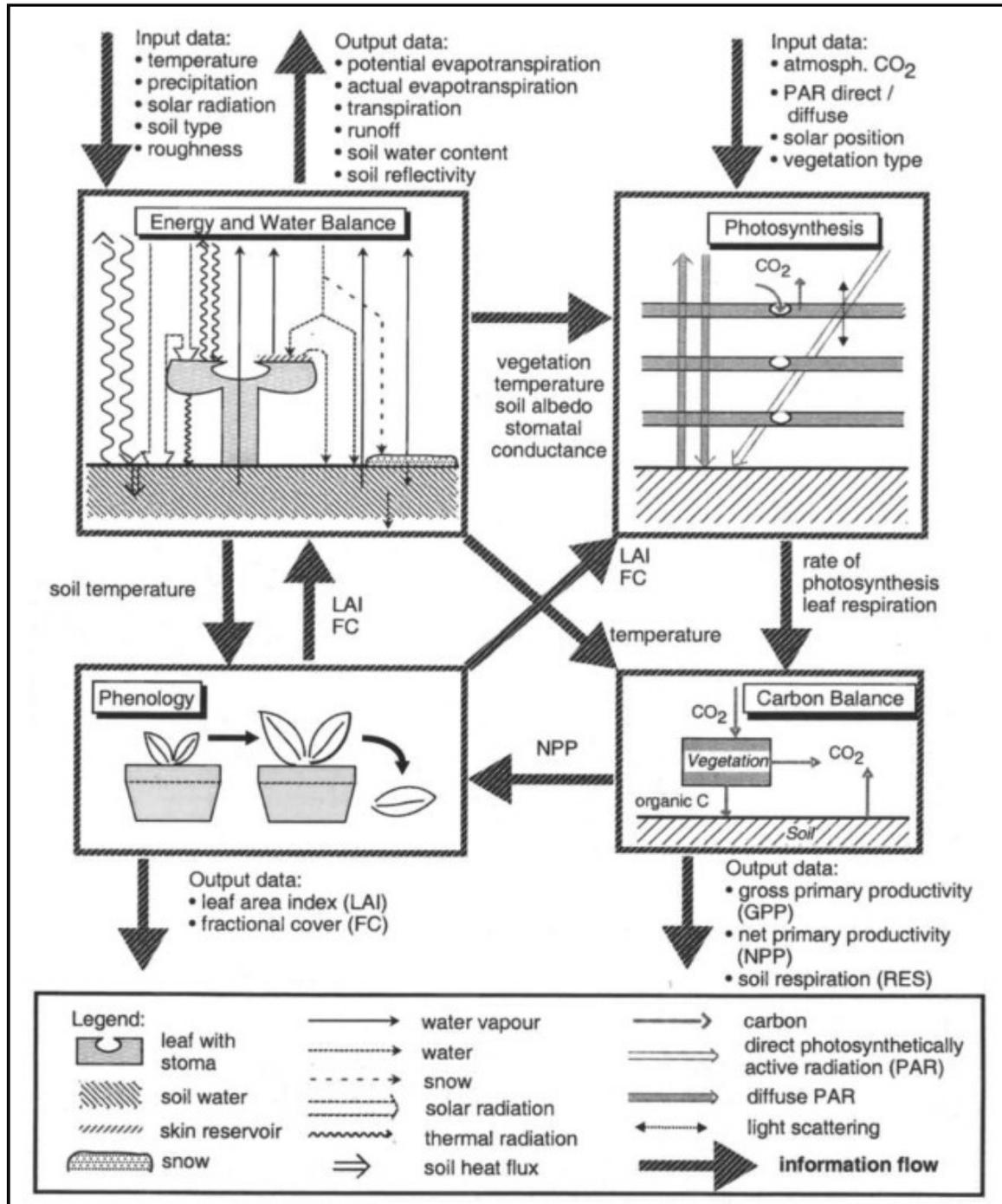
The W_{ε} is arguably the most crucial stressor to include in arid, semi-arid and drought-affected regions. The calculation of AET and PET is fairly complex (Carlson, Capehart & Gillies 1995; Courault, Seguin & Oliosio 2005; Glenn et al. 2008) and typically require meteorological data for accurate measurements. For further descriptions of AET and PET calculations in the RBM, refer to Richters (2005b).

The RBM has been used in various studies (Eisfelder et al. 2013; Machwitz et al. 2015; Richters 2006), particularly in southern-African study areas, where a global model would not be suitable for the fractioned vegetation patterns and low availability of additional input data.

2.2.2.2 Biosphere Energy Transfer Hydrology model

The BETHY model is a popular process-based model developed by Knorr & Heimann (2001) and has been used in southern Africa for carbon sink modelling and degradation analysis (Niklaus, Tum & Günther 2010). The model follows the SVAT scheme (Section 2.2.1.4) and consists of

four model components: energy and water balance, photosynthesis, phenology and carbon balance. Model inputs include biophysical parameters similar to that of RBM, as well as various other inputs relating to climatic, atmospheric and soil conditions, and the interaction between the four components are quite complex. Figure 2.3 shows the structure of the BETHY model with input and output data, as well as the information flow between the four components.



Source: Niklaus, Tum & Günther (2010)

Figure 2.3 BETHY model diagram

Similar to the model developed by Running et al. (2000), BETHY was originally designed for global GPP modelling. The model has subsequently been adapted for regional-scale GPP modelling by Wißkirchen et al. (2013) and named BETHY/DLR. Changes include improvement of the spatial resolution of model outputs, use of a continuous time-series of daily climatic input and, unlike the original model, phenology is calculated with the use of satellite-derived LAI data and not within the model itself (Wißkirchen et al. 2013). It must be noted, however, that satellite-derived LAI data typically do not provide very reliable results (Tian et al. 2002) and could potentially have a negative impact on subsequent productivity derivatives.

2.3 FRACTIONAL VEGETATION COVER

Fractional vegetation cover (FVC) is one of the main, but often overlooked, biophysical parameters relating to landscape surface processes and plays a vital role in deriving grassland condition and productivity (Guerschman et al. 2009). It involves determining the proportional area of specific types of groundcover a pixel consists of, based on various spectral characteristics of the pixel (Jiménez-Muñoz et al. 2009). The types of groundcover differ depending on the FVC application, but usually include green (e.g. leaves), dead (e.g. wood) and bare (e.g. soil). This essentially allows discriminating between productive (grass) and unproductive (woody branches, bare soil) groundcover within a pixel, which in turn is potentially useful for deriving subsequent productivity estimates.

2.3.1 Field-based methods

Field-based FVC is often measured using visual inspection due to its simplicity and operational ease. As this approach depends on the estimator's experience, there is an inherent element of subjectivity (Liang, Li & Wang 2012) and error is easily introduced depending on how meticulous the estimator is. The simplest form of visual estimation consists of selecting several sample plots and determining FVC based solely on observation. In some cases, vertical photographs are captured of the plots and visually estimated using reference images (Hu et al. 2007; Liang, Li & Wang 2012). When using the photographic method, an additional procedure involves generating a reference image series depicting the various levels of FVC that could be present in the sample plots (Liang, Li & Wang 2012). Estimators then receive training on image interpretation using the reference series, increasing processing time but decreasing potential error.

In an effort to improve precision and mitigate subjectivity based on estimator experience, visual inspection is often coupled with a form of sampling (Liang, Li & Wang 2012). The most common example is the line transect method, which entails unrolling a transect line along the gradient identified and recording species using either continuous sampling or systematic sampling

(Buckland et al. 2007). Continuous sampling is performed by recording all cover types touching the line along the entire length of the transect, whereas systematic sampling involves recording the presence or absence of a species at each marked point (Buckland et al. 2007). These measurements are then summed per groundcover type to determine the respective total proportions of each groundcover type within the sample plot.

2.3.2 Remote sensing based methods

Deriving an accurate FVC from RS data is a complex process, as it involves extracting sub-pixel information from temporally, spatially and spectrally variable phenomena such as vegetation (Li et al. 2014). The two most common techniques for retrieving FVC is VIs and spectral mixture analysis (SMA) (Jiménez-Muñoz et al. 2009). A less common method also includes the use of machine learning algorithms e.g. artificial neural networks, fuzzy classifiers, maximum likelihood classifiers and regression trees (Li et al. 2014).

2.3.2.1 Vegetation indices

When using VIs to estimate FVC, NDVI (Equation 2.1) is a popular choice. Gutman & Ignatov (1998) demonstrated a linear relationship between NDVI and FVC using a scaled NDVI:

Equation 2.16

$$FVC = \frac{NDVI - NDVI_s}{NDVI_v - NDVI_s}$$

where FVC is the fractional vegetation cover; and

$NDVI$ is the normalised difference vegetation index.

In this relationship, $NDVI_s$ correspond to representative values of NDVI for bare soil ($FVC=0$) and $NDVI_v$ correspond to representative values of NDVI for vegetation ($FVC=1$). The principal problem with this approach is identifying the correct values for $NDVI_s$ and $NDVI_v$, as these are region-, season- and vegetation-specific (Jiménez-Muñoz et al. 2009).

Due to the limitations of NDVI with regards to soil background and saturation (Section 2.1.2), the use of various other VIs for FVC estimation has been explored in literature (Gitelson et al. 2002). As with NDVI, particular care must be taken in identifying corresponding values for bare soil and vegetation for the specific VI. In addition to the linear relationship defined in Equation 2.16, non-linear regression models are also widely implemented (Liang, Li & Wang 2012). The type of regression method depends largely on the cover types explored, the study area and the VIs used for estimation.

2.3.2.2 Spectral mixture analysis

The SMA approach allows extraction of land-cover information from satellite imagery at a sub-pixel level. Endmembers, which represent the pure spectral characteristics of a specific LC feature (Plaza et al. 2004), are used to divide the pixel into its different components. Various models exist for performing SMA, of which the linear spectral mixture model (LSMM) is the simplest and most common. Relatively few endmembers are required to describe the surface composition per pixel, thus an LSMM is an appropriate choice when field data is limited. A general LSMM can be defined by the following equation (Sabot et al. 2002):

Equation 2.17

$$\rho_i = \sum_{em=1}^{N_e} F_{em} \rho_{em,i} + E_i ; \quad \sum_{em=1}^{N_e} F_{em} = 1$$

where

- ρ_i is the reflectivity for each channel i ;
- N_e is the number of endmembers;
- F_{em} is the fraction of endmember; and
- E_i is the unmodeled residual.

When performing SMA, the main problem is identifying and extracting high quality, pure endmember spectra that accurately reflect the spectral profile of each endmember. Various methods of endmember extraction have been implemented, including using a land use map (Jiménez-Muñoz et al. 2009), semi-supervised endmember extraction (Boardman, Kruse & Green 1995) and fully automatic endmember extraction (Plaza & Chang 2006). The first method involves manual selection of homogenous pixels relating to each endmember from a land use map, whereas semi-supervised extraction makes use of the pixel purity index (PPI) algorithm to find the most spectrally pure pixels in the image. Fully automatic endmember extraction is performed using the automated morphological endmember extraction (AMEE) approach, which makes simultaneous use of both spatial and spectral information via multi-channel morphological processing (Chang 2013). In a study comparing the three methods of endmember extraction in deriving FVC for an agricultural region in Spain (Jiménez-Muñoz et al. 2009), the manual method of using a land use map produced the best results.

2.3.2.3 Machine learning

As computer technology improves, the machine learning approach for FVC estimation has become an increasingly popular alternative (Jia et al. 2015; Shiferaw, Bewket & Eckert 2019; Yang et al. 2016). Common machine learning algorithms include support vector machines, decision trees and neural networks. When implemented to estimate FVC, machine learning methods typically involve

identifying training points corresponding to different groundcover types and using them to train a classification model. Classification is performed on satellite imagery and the probability of each pixel being a specific cover type is used as the FVC value. Machine learning methods require no parameterisation and can be easily automated, facilitating their extensive application in RS-based analysis in recent years (Liang, Li & Wang 2012). Successful application, however, requires a large number of training points (Campbell 2002), which is often time-consuming and/or expensive to obtain.

2.3.3 Fractional vegetation cover products

A number of public global-scale FVC products have been created, most well-known the MODIS Vegetation Continuous Fields (VCF) product (MOD44B collection 6) with a spatial resolution of 250 m (Townshend et al. 2017). The product is created annually using a machine learning approach. Another popular product is the Copernicus Dynamic Land Cover (CDLC) product (Tsenbazar et al. 2018) created on a once-off basis for 2015 at 100 m. CDLC was also estimated using a machine learning approach, where random forest regression was applied to determine the proportion of groundcover type within each pixel. The two products differ not only with respect to spatial resolution, but also in the groundcover types and definitions used. CDLC provide fractional cover layers for 10 groundcover types: shrubland, forest, moss and lichen, snow, bare/sparse vegetation, permanent water, seasonal water, cropland, built-up and herbaceous vegetation. MOD44B, however, provides only three i.e. tree, non-tree and non-vegetated (bare), and shrubs is included in the non-tree class.

FVC has been estimated in southern Africa using both VIs (Scanlon et al. 2002) and spectral unmixing (Gessner et al. 2013), as well as a combination of these approaches (Sankaran et al. 2005). The Council for Scientific and Industrial Research (CSIR) Ecosystems Earth Observation Unit developed a woody fractional canopy cover for South Africa as part of the Carbon Sinks Atlas project (Department of Environmental Affairs 2017). A once-off 100 m and 1 km savanna woody fractional cover was created for the year 2011 with the use of Synthetic Aperture Radar (SAR) and Light Detection and Ranging (LiDAR) (Naidoo et al. 2016). The South African Environmental Observation Network (SAEON) has made the 1 km spatial resolution layer available online to be viewed and downloaded, along with other layers related to carbon sinks and sequestration e.g. land degradation index, savanna aboveground woody biomass and soil organic carbon. This product could thus potentially be used in conjunction with current productivity models when conducting a national-scale, once-off grassland productivity assessment. However, the use of these layers for dynamic, continuous assessment of grassland productivity at a local scale is not ideal due to the coarse spatial and temporal resolution. A spatial resolution of 1 km will not be sufficient when

attempting to model fractional cover components at a larger scale, as more detail is required to distinguish between less discrete cover types e.g. grassland and shrubland. In addition, due to the static nature of the approach, new FVCs can only be created every 5 to 10 years. LiDAR data is expensive and not easily accessible, thus FVCs cannot be generated continuously to model the rapid change due to increasing WPE.

More extensive studies have been done in Australia, where the vegetational structure and climatic conditions of rangelands and grasslands are similar to that of South Africa. Of particular interest is the FVC product developed by AusCover based on the methodology of Scarth, Roder & Schmidt (2010). Scarth, Roder & Schmidt (2010) developed a non-negative constrained least squares linear spectral unmixing model to derive an accurate FVC from Landsat 8 imagery for the extent of Australia. In contrast to the previously mentioned studies, this study makes use of medium resolution satellite imagery (30 m spatial resolution) to produce a fractional cover layer, thus eliminating the need for expensive hyperspectral or very high resolution data to produce accurate FVC results. The Australian state governments of Queensland and New South Wales have recently adapted the Landsat-based methods to determine FVC using Sentinel-2 imagery (10 m spatial resolution), which combines the advantages of both a high temporal and high spatial resolution. Access to both high spatial and temporal resolution is essential when analysing grassland productivity, as grassland ecosystems tend to be unstable and temporally variable (Zhang et al. 2017). A drawback of this product for grazing management is the FVC class definitions, namely bare soil (BS), photosynthetic vegetation (PV) and non-photosynthetic vegetation (NPV) (Scarth, Roder & Schmidt 2010). The PV class contains both tree and shrub leaves as well as herbaceous vegetation, thus combining non-grazeable and grazeable material in the same band.

Many of these studies required either hyperspectral data, very high resolution imagery or very expensive active sensor technology (SAR, LiDAR) to produce accurate FVC estimates, which introduces challenges with regards to accessibility, availability and affordability. Those that did not typically provide FVC estimates at global scale, which is too coarse for identifying complex vegetational patterns and dynamics in savannas and woody encroached grasslands. In addition, the groundcover types and definitions used as fractional cover classes differ widely, complicating the process of comparing different FVC products.

2.4 GRASSLAND AND RANGELAND MANAGEMENT

Grassland and rangeland monitoring and management play an important role in preventing degradation caused by overgrazing and WPE. The majority of management approaches still rely primarily on field-based methods to determine various aspects of grassland condition and productivity, including dry matter (DM) production. These field measurements are typically used

for subsequent grazing capacity (GC) calculations, an important agro-statistical indicator for the sustainable grazing of rangelands.

2.4.1 Dry matter production

Dry matter (DM) production refers to the dry weight of material (e.g. grass) produced per unit area during a specified time period (Vickery 1972). DM production can be related to RS-based GPP estimates (Section 2.2.1) using a simple unit conversion (Swinnen & Van Hoolst 2018). Within the livestock industry, however, farmers primarily make use of field-based methods to determine DM production, of which the two most common are clipping-and-weighing and the disc pasture meter (DPM). In addition to the DM production of vegetation, the species composition of plots are usually also recorded and used for further determination of grassland condition and productivity.

2.4.1.1 Clipping-and-weighing

This method involves clipping, drying, and weighing samples to estimate the average DM production of an area (Jones, J. B. & Case 1990). Grass clippings are sampled in quadrants based on a sampling scheme of choice and packaged in paper bags. The fresh, green samples are weighed and dried in a lab. A forced-air oven at 60 °C is used to correct for water content using an appropriate drying method. The most common is the “drying to constant mass” method, where clippings are weighed every 24 hours until no changes have occurred between two weigh-ins (Jones, J. B. & Case 1990). The clippings are removed from the oven and weighed again to determine DM mass. Using the size of the clipping quadrants and measured DM mass, DM yield or production can be calculated for an area. If a large area needs to be surveyed, only a selected number of samples are taken to the lab and dried. The ratio of “fresh weight” to “dry weight” is then calculated and used to estimate the DM production of the remaining clipping samples that were not dried. The precision of this method depends largely on pasture variability and sampling efficiency. Hand clipping is the most precise and direct method (Harmony et al. 1997; Sharrow 1984), but time consuming and destructive, which makes routine use impractical.

2.4.1.2 Disc pasture meter

The DPM method was developed by Bransby & Tainton (1977). It is the most practical and efficient method (Joubert & Myburgh 2014) and, unlike clipping-and-weighing, it is non-destructive. The DPM consists of three components: a 180 cm long aluminium rod; a 120 cm long aluminium sleeve attached to a base-plate that slides freely along the central rod and an aluminium disc with a 45,8 cm diameter (disc diameters may vary) bolted to the base-plate. The central rod

is marked at 0.5 cm intervals in an upward direction. When taking DPM measurements, the central rod is held perpendicular to the ground and the sleeve (and attached disc) is released onto the grass sward. The settling height of the disc is read off the rod, thus essentially measuring the compressed grass height of the grass sward. A calibration equation is used to translate the compressed DPM measurement to DM production.

Calibration of the instrument was first described by Bransby & Tainton (1977) and their procedure has since been applied in various areas in South Africa (Danckwerts & Trollope 1980; Joubert & Myburgh 2014; Zambatis et al. 2006). The process involves dropping the disc, reading the DPM height and clipping, drying and weighing all grass materials under the disc (similar to Section 2.4.1.1). The mass is then converted to an appropriate unit (usually kg ha^{-1}) and used as the dependent variable in regression of mass against the DPM height readings (Bransby & Tainton 1977). This process is repeated for enough sample sites to formulate a reliable regression equation. The regression relationship can then be used for future surveys to derive DM production in a non-destructive manner.

Numerous studies in rangelands, grasslands and savannas of southern Africa have resulted in a number of universal calibration equations (Bransby & Tainton 1977; Danckwerts & Trollope 1980; Zambatis 2003; Zambatis et al. 2006). If a grassland area is similar to regions for which calibrations have already been done, no calibration is required and the method is entirely non-destructive. Joubert & Myburgh (2014), however, advises using universal calibration equations cautiously by ensuring they are suitable for the specific area's vegetation type, management practices and climate. In addition to unsuitable calibration equations, error can also be introduced during the measurement itself. Measurements must be taken with great care as disc height might be affected by rocks, rather than grass (Zambatis et al. 2006). In addition, the instrument also frequently settles above the grass bulk in very dense, tall grasslands, e.g. savanna in the Kruger National Park, resulting in a nonrepresentative DPM height reading (Zambatis et al. 2006). Accuracy of measurements thus depends highly on the quality of the DPM measurements, as well as the reliability of the calibration equation used.

2.4.2 Grazing capacity

Grazing capacity (GC) (ha LSU^{-1}) is defined as the number of hectares of productive land required to sustain one large stock unit (LSU). In southern Africa, GC is usually assessed with respect to cattle, thus one LSU relates to one cow. Determining GC is an important, yet complex, calculation in semi-arid grassland and rangeland management (Guevara et al. 1997). Once the GC is calculated, farmers can use this information for decision-making with regards to applying

appropriate grazing systems, identifying fields that require resting and adjusting the number of cattle per field for sustainable grazing (Guevara et al. 1997).

The calculation of GC is explained by Guevara et al. (1997) using the following scenario. A cow consumes approximately 2.5% of its weight per day, thus an adult 400 kg cow would consume approximately 10 kg of grass per day, i.e. 3650 kg grass annually. As the overall proper-use factor for un-improved native grasslands is approximately 30%, of an annual dry matter production (aDMP) of i.e. 634 kg DM⁻¹ha⁻¹year⁻¹, only 190 kg DM⁻¹ha⁻¹year⁻¹ can be used. Using this logic, the GC would then be estimated as 3650/190 = 19 ha LSU⁻¹. This GC calculation for grasslands can be formalised as:

Equation 2.18

$$GC = \frac{0.025\sigma}{0.3aDMP}$$

where σ is the average weight of a large stock unit [kg]; and

$aDMP$ is the annual dry matter production [kg DM ha⁻¹year⁻¹].

The aDMP used as input for GC calculation is typically estimated using field-based methods discussed in Section 2.4.1. Based on a number of systematic DM production field measurements a year, an annual estimate is calculated. This annual vegetation production, as well as observed species composition, palatability and accessibility, is used to characterise the grazing conditions of the area. Although field-based methods provide the most direct and accurate estimates, they are not applicable large-scale due to their destructive and time-consuming nature. Another source of GC estimates for rangeland management is provided in the form of the Long Term Grazing Capacity (LTGC) maps. These knowledge-based maps have been created for South Africa for the years 1993, 2009 and 2016. The most recent map (2016) consists of grazing zones representing homogenous grazing conditions, classified into 0.5 ha LSU⁻¹ GC intervals (Appendix B). The delineations of the zones were based on the Mucina & Rutherford (2006) vegetation type map and farm cadastral boundaries, and classified using expert opinion (Avenant 2017). The static map aims to serve as a guideline and does not indicate current veld condition and vegetation production (Avenant 2017).

A limited number of studies have explored the use of RS for GC estimation (Adjorlolo & Botha 2015; Espach, Lubbe & Ganzin 2009; Fajji, Palamuleni & Mlambo 2018; Yu et al. 2010). Theoretically, the primary production can be estimated using RS (Section 2.2.1), converted to agronomical units and used in Equation 2.18 to calculate GC. Current RS-based primary production approaches, however, have two main constraints. Firstly, it is not yet possible to

accurately derive species composition using medium resolution (10 to 30 m spatial resolution) satellite imagery. This hinders the determination of veld condition and palatability of grasslands using RS data and techniques. Another challenge is the effect of tree cover at sub-pixel level on the overall primary production of a pixel. A pixel with a mix of trees, shrubs and grass will typically have a higher primary production estimate than one consisting of only herbaceous vegetation (Liang, Li & Wang 2012). When subsequently using these primary production estimates in GC calculations, the woody encroached pixel will be characterised as having very low GC values i.e. favourable grazing conditions, even though it realistically has less grazable organic material.

Espach, Lubbe & Ganzin (2009) and Adjorlolo & Botha (2015) both attempted to address this constraint using a combination of extensive field sampling and RS-derived woody quantification. Espach, Lubbe & Ganzin (2009) corrected RS-based vegetation production using a combination of a discrete LC map and field surveys to generate a woody cover spatial mask. Drawbacks of this approach include the discrete, static nature of the LC map used, as well as the extensive field surveys required. Adjorlolo & Botha (2015) successfully identified woody fraction indications using RS-based measurement of tree density and total leaf mass. The method, however, involved integration of a complex biophysical model and random forest regression, which required a large number of field points. In addition, vegetation production was estimated using a conventional field-based survey (Section 2.4.1) instead of a RS-based approach. Although these approaches allowed incorporating the effect of tree and shrub cover on GC in grasslands, both required substantial field work to determine certain input parameters in the GC calculation.

2.5 MODERN TECHNOLOGIES

This section discusses the potential of modern technologies, i.e. geoprocessing cloud platforms and web applications (apps), to address the technological limitations of current RS approaches. The benefits and drawbacks of Google Earth Engine (GEE) are investigated, as well as available options for creating GEE web apps. The section concludes with a review of existing grassland and rangeland management apps.

2.5.1 Google Earth Engine

GEE is a geoprocessing cloud platform that combines a catalogue of online satellite imagery and geospatial datasets with global-scale analysis capabilities (Hansen et al. 2013). Users are allowed to run geospatial analysis on Google's Cloud infrastructure using either the web-based JavaScript integrated development environment (IDE) or the locally installed Python application programming interface (API) (Gorelick et al. 2017).

GEE is utilized across a wide variety of disciplines, including topics such as mapping global forest change (Hansen et al. 2013), global surface water change (Pekel et al. 2016), crop yield estimation (Lobell et al. 2015), rice paddy mapping (Dong et al. 2016), urban mapping (Zhang et al. 2015), flood mapping (Coltin et al. 2016), fire recovery (Soulard et al. 2016) and malaria risk mapping (Sturrock et al. 2014). This increasing popularity of GEE for large-scale processing of imagery and geospatial analysis is due to the many benefits it holds compared to traditional, desktop-based approaches.

In addition to some of the obvious benefits of GEE over traditional approaches i.e. free access, requiring no specialised desktop software, intensive processors or downloaded imagery, the platform itself also has a wide range of processing-related advantages. Mapping over large areas at high resolution is computationally expensive and it can take weeks to obtain final maps and to visually evaluate prediction models. In contrast, map generation in GEE is 40-100 times faster, as the platform allows users to take advantage of Google's computing power (Padarian, Minasny & McBratney 2015). The platform is designed to make planetary-scale earth observation easy, especially when executing extremely large computations (Navarro 2017). To achieve this, parallel processing is implemented by the underlying Google infrastructure when running GEE commands, thus relieving the users from having to configure parallelisation themselves. Although this allows users with limited computer science experience to make use of large-scale processing advantages in GEE (Padarian, Minasny & McBratney 2015), it also poses challenges for certain types of analyses. Gorelick et al. (2017) highlight that by design, many geoprocessing analyses are non-parallel. This means that some large-scale non-parallelisable operations cannot yet be performed effectively in the GEE environment e.g. recursive algorithms or long-running iterative processes, as large computations can only be expressed by using the parallel processing primitives provided in the GEE library. The platform also performs poorly with data-intensive models that require large volumes of data not already available in the GEE archive, as it requires substantial additional effort to continuously import large files (Navarro 2017). Despite these disadvantages, GEE is well-suited to per-pixel and neighbourhood operations e.g. band math, spectral unmixing and texture analysis, as well as chaining these operations together (Gorelick et al. 2017). It is also highly optimised for statistical operations, such as computing geostatistics on a large time-series stack of images (Padarian, Minasny & McBratney 2015).

In addition to the processing benefits of GEE, it is relatively easy to integrate the geoprocessing cloud functionality with app development platforms. Widely used GEE web apps include e.g. Map of Life for analysing habitat ranges of species (MOL 2016), Climate Engine for climate monitoring (Desert Research Institute 2016) and Collect Earth for land use change assessment

(Bey et al. 2016). GEE apps typically involve writing a backend server using the Earth Engine Python API and Google App Engine (GAE), which interacts with the user via a frontend interface (usually in the form of a web page built using HTML, JavaScript and CSS). The backend server authenticates and communicates with GEE to allow running the various geoprocessing analyses required by the web app in the cloud. New functionality within the GEE code editor allows the creation of web apps directly using existing scripts within the developer's repository, thus bypassing the time, maintenance and programming skills required to build your own GEE-GAE backend server. Simple JavaScript user interfaces are built within the script to allow input from users, limiting user access to the analysis script itself. Although the new functionality allows the creation of apps with relative ease and little programming experience, customisation regarding the user interface, backend functionality and additional features (e.g. database access) are limited.

The use of GEE for large-scale processing of imagery and geospatial analysis has become increasingly popular not only for traditional earth observation scientists, but also a wider audience that lacks the resources needed to perform large-scale geospatial analyses on a desktop-basis. New app functionality allows live demonstration of GEE analysis results to non-GEE users and individuals with little or no RS and spatial analysis experience. The platform, however, has its limitations and many users still prefer desktop-based analysis.

2.5.2 Existing applications

A number of existing applications for grassland and rangeland management exist. Most well-known is the Rangeland and Pasture Productivity (RAPP) app, a Group on Earth Observations Global Agricultural Monitoring (GEOGLAM) initiative. GEOGLAM RAPP aims to establish a global monitoring system of land use productivity for pastures and rangelands (GEO 2016) and was established and is funded by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the Australian Government's National Landcare Programme. Various layers are provided in the app, including FVC, climate data and cattle densities, where FVC is derived from Landsat 8 imagery at a 30 m spatial resolution (Guerschman et al. 2009). Another platform is the Rangeland Decision Support Tool (RDST), developed by the Regional Centre for Mapping of Resources for Development (RCMRD) and SERVIR: Eastern & Southern Africa (RCMRD 2018). The main objective of the application is to facilitate near real-time assessment and monitoring of rangeland resources and layers include NDVI, NDVI anomalies and vegetation condition index (VCI) using MODIS at 250 m spatial resolution. The final app is the Rangeland Analysis Platform (RAP), which was developed by the United States Department of Agriculture's Natural Resources Conservation Service (USDA NRCS), the Department of Interior's Bureau of Land Management (DOI BLM) and the University of Montana (USDA NRCS & DOI BLM 2018).

It is an interactive web application designed to assist in managing and monitoring America's rangelands and provides fractional cover layers for perennial, shrub, tree and bare cover. These FVCs are created using machine learning, Landsat 8 imagery, a vast database of more than 30 000 field points and GEE (Jones et al. 2018).

Although the discussed apps provide valuable information for rangeland management, results are either provided at a too coarse resolution (RDT), or are not available in the South African context (RAP). In addition, none of the applications provide estimations past basic vegetation condition, dynamics and structure e.g. NDVI and FVC.

2.6 LITERATURE SUMMARY

The literature reviewed in the chapter proves RS-based estimation of vegetation production (Section 2.2) and FVC (Section 2.3) shows great potential for sustainable and effective grassland and rangeland management. Various primary production modelling approaches have been developed and evaluated (Section 2.2.1) to allow accurate characterisation of vegetation productivity within terrestrial ecosystems. These models all have their benefits and limitations and no optimal modelling approach can be identified in literature, thus the choice of model relies heavily on available input data, analysis scale, vegetation type and structure, processing power and application. For an application that requires regional-scale and transferable results, few input parameters and an element of mechanistic realism, an LUE modelling approach shows the most potential.

Current grassland and rangeland management approaches primarily make use of field-based techniques to estimate GC (Section 2.4). It is evident that field data plays an important role in monitoring and managing of vegetation, both for general characterisation of vegetation dynamics and the modelling of productivity. Field-based methods, however, are not suitable for routine practice and measurement of large areas as they are time-consuming and destructive. The combination of field data with RS analysis, e.g. spectral signatures and indices (Section 2.1), can potentially address this challenge, however, these approaches have their own limitations in the context of sustainable grazing management. A prominent constraint for RS-based approaches is the inability of primary production models to distinguish tree production from herbaceous production at sub-pixel level. Pixels with high tree cover are thus characterised with lower GC values i.e. better grazing conditions, even though it realistically has less grazable organic material. The integration of FVC estimated from RS data (Section 2.3.2) could potentially address this limitation by allowing the productivity model to differentiate between herbaceous vegetation, woody components and bare soil. VIs, SMA and machine learning approaches for FVC estimation using RS are common in literature, although they typically require very high resolution or

hyperspectral imagery. Machine learning models have the potential to provide very accurate results at medium resolution, however, a large number of training points are required. SMA is thus an attractive option, as relatively few field points are required. When implementing this technique, however, the quality of field points play an important role. This could prove challenging for medium resolution imagery, as pure pixels of 10 and 30 m spatial resolution can be difficult to find.

Currently there are a number of grassland and rangeland management tools available (Section 2.5.2), however, none provide estimates of actual vegetation production and subsequent agronomical indicators e.g. GC, stocking rate etc. This is understandable in light of the scientific limitations of current RS-based approaches. The combination of FVC and primary production estimates, as well as cloud computing geoprocessing and app functionality of GEE (Section 2.5.1), can potentially allow dynamic, in-field and locally-tailored productivity estimates when implemented in a web app. This combination can thus not only address technological challenges in current literature, but also scientific limitations.

The next chapter (Chapter 3) focuses on field data collection and analysis, which plays a vital role in establishing relationships between satellite imagery and the complex biophysical parameters required for modelling vegetation dynamics. The results are then used in Chapter 4 and Chapter 5, which focus on FVC estimation and grassland productivity modelling respectively. Chapter 5 also investigates the implementation of an FVC-integrated grassland productivity model in a GEE app to investigate the scientific and technological gaps identified in current literature.

CHAPTER 3: FIELD DATA COLLECTION AND ANALYSIS

Data collection and analysis is crucial in understanding the relationship between vegetation productivity and remote sensing (RS) signals. The establishment of such relationships allow the development of robust, transferable production models that do not rely on extensive field sampling (Eisfelder, Kuenzer & Dech 2012). Due to the complex nature of vegetation productivity modelling and the general lack of sufficient field data, the collection and analysis of various biophysical parameters relating to grassland health, structure and productivity comprised a key component of the research. The collected field data was used for both calibration and validation purposes. The calibration results established a baseline that explains how changes in vegetation at ground-level is reflected and observed in the satellite imagery. KoboToolBox and KoBoCollect, a free, open-source field data collection package based on an OpenDataKit system (KoBoToolBox 2012), was used to record all pertinent field data. The field data was analysed qualitatively and quantitatively to identify relevant relationships used to estimate fractional vegetation cover (FVC) (Chapter 4) and model productivity (Chapter 5), thereby addressing Objective 2.

3.1 DATA COLLECTION

In-situ field data was collected during two field trips to cover both a dry season (Fort Beaufort) and a growing season (Cedarville) (Section 1.5). A growing season field trip was originally planned for the Fort Beaufort study area, but due to drought (Baudoin et al. 2017), little or no vegetation growth took place. It would thus not have been possible to collect any biomass for the purpose of modelling grassland condition during the wet season.

3.1.1 Dry season field trip: Fort Beaufort

The dry season field sampling took place from 25 to 29 June 2018 in Fort Beaufort and surrounds (Eastern Cape, South Africa) and was explorative in nature. Qualitative information was gathered to characterise the land cover (LC), vegetation and terrain of different levels of woody encroached grasslands, and to determine the range of biophysical field parameters required when modelling grassland productivity. The region has experienced drought conditions since 2015 (Baudoin et al. 2017), as well as high levels of woody plant encroachment (WPE), thus the grasslands were in relatively poor condition.

Different sampling sites were identified based on vegetation type and structure (e.g. forest, thicket, grasslands), resulting in eight sample points describing the different vegetation biomes and levels of woody encroachment within the study area. To establish relationships between vegetation

condition, level of encroachment and satellite imagery, two biophysical parameters were measured at each sample point: FVC and leaf area index (LAI).

3.1.1.1 Fractional vegetation cover

The FVC per sample site was measured using a simple visual inspection, as described in Section 2.3.1. The fraction of soil, grass, shrubs and trees at each site was assessed and recorded, thus reflecting the level of encroachment within each sample site.

3.1.1.2 Leaf area index

An AccuPAR LP-80 Ceptometer (METER Inc. 2018) was used to measure LAI. Five measurements were taken within each sample plot and averaged to determine an average LAI value for the different vegetation types and levels of encroachment. The sample points and collected field information are visualised and summarised in Section 3.2.1.

3.1.2 Growing season field trip: Cedarville

The growing season field sampling took place from 24 to 29 February 2019 in the Cedarville area, Eastern Cape, South Africa. Unlike the Fort Beaufort study site, the region had started recovering from drought conditions and the study site received an average rainfall of 280 mm in the month of February 2019 alone. Grasslands were lush and green with little to no woody encroachment.

To ensure a wide range of grassland condition field data, potential field sample sites were selected prior to the field trip within five normalised difference vegetation index (NDVI) classes, specified in Table 3.1. However, due to heavy rains during the month of February, NDVI values within the region increased notably until the period of collection. Figure 3.1 illustrates this increase, showing NDVI calculated using Sentinel-2 imagery for the beginning of the month (2019/02/01; Figure 3.1a) compared with two days before data collection (2019/02/23; Figure 3.1b).

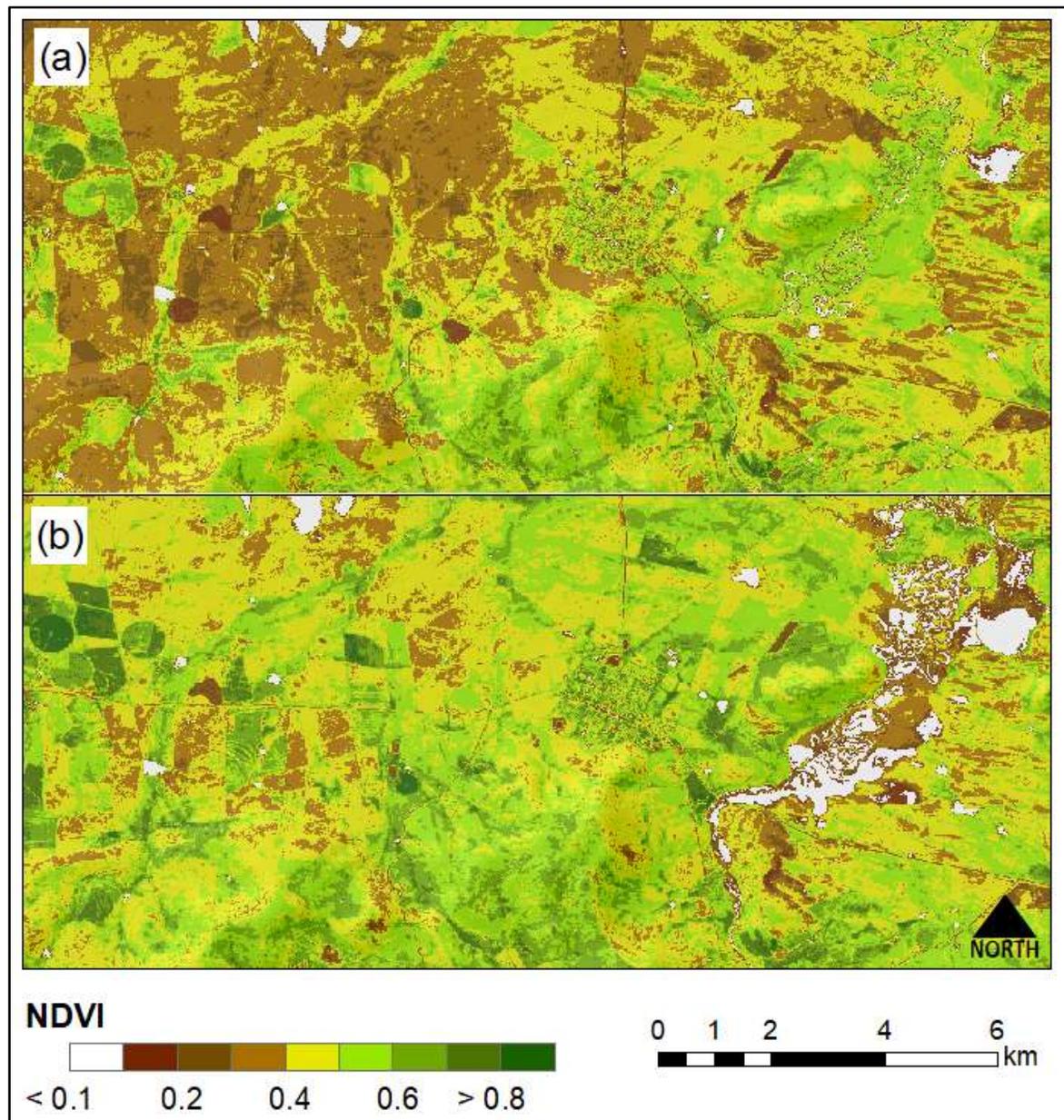


Figure 3.1 NDVI derived from Sentinel-2 imagery for the Cedarville region for (a) 2019/02/01 and (b) 2019/02/23

This heavy rainfall also resulted in flooding and wetland-like conditions in the low, flat grass plains situated around the Mzimvubu River, as shown by the increased number of NDVI pixels with values lower than 0.1 in the East of Figure 3.1b. To account for the overall increase in NDVI, sampling intervals had to be adjusted accordingly. Table 3.1 shows the original NDVI intervals (from Sentinel-2 image 2019/02/01), the adjusted intervals (from Sentinel-2 image 2019/02/23) and the description of grassland condition for the NDVI interval.

Table 3.1 Original and adjusted NDVI intervals for field sampling

Original NDVI Intervals	Adjusted NDVI Intervals	Grassland Condition
0.1 – 0.2	0.1 – 0.3	Bare; degraded
0.2 – 0.3	0.3 – 0.4	Recently grazed; overgrazed
0.3 – 0.4	0.4 – 0.6	Good condition; most common
0.4 – 0.5	0.6 – 0.7	Very productive; well-rested
0.5 – 0.6	0.7 – 0.8	Very dense grasslands; areas where water run-off accumulates

New sample point locations were identified based on the adjusted NDVI intervals (Table 3.1) and accessibility. To ensure consistency between satellite and ground-level, each sampling point's corresponding Sentinel-2 pixel bounds were recorded to locate and delineate the 10 m by 10 m pixel at ground-level during in situ field acquisition. All relevant biophysical parameters were then recorded within the field-delineated pixel. Four biophysical parameters were measured at each sample point: FVC, LAI, fraction of absorbed photosynthetically active radiation (fAPAR) and grass height. Two additional parameters were derived from grass height, namely green-dead biomass proportions and dry matter (DM) production.

3.1.2.1 Fractional vegetation cover

FVC was measured for each plot using the line transect method with continuous sampling (as described in Section 2.3.1). A 10 m tape measure was used as transect line to correspond to the spatial resolution of Sentinel-2 pixels. Patches of soil, grass, shrub and forb were recorded along the transect line and then summed to determine the respective total proportions of each cover type representing the pixel. Cover fractions for trees were not recorded as there were no trees present. For the first few sample sites, the full 10 m transect was performed. The change of respective cover proportions with increasing transect length was plotted to determine at what length proportions within the sample site stabilised, as shown in Figure 3.2. For each cover type, a stabilisation point was reached, where cover proportions no longer changed significantly.

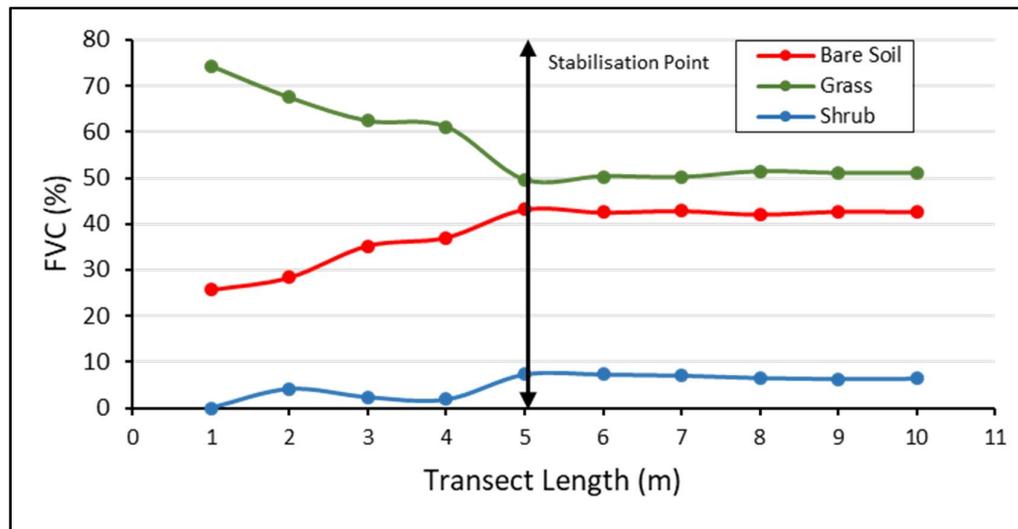


Figure 3.2 Example of changing FVC (%) with increasing transect length (m) with a 5 m stabilisation point

For this study, the stabilisation point was found to be 5 m. Subsequent FVC samples were thus taken using 6 m line transects to decrease sampling time while still ensuring accurate representation of the proportion of vegetation cover.

3.1.2.2 Leaf area index and fAPAR

An AccuPAR LP-80 Ceptometer (METER Inc. 2018) was used to measure LAI and fAPAR. Five measurements were taken within each sample plot and averaged. The LAI and fAPAR measurements were acquired to use in linear regression analysis to investigate relationships with NDVI, FVC and green-dead biomass proportions within the sample site.

3.1.2.3 Green-dead biomass proportions

Green-dead biomass proportions were measured using the plot harvest clipping method. Within selected sample sites, a grass clipping of 0.2 m² was taken (the size of the disc pasture meter's disc). The grass clippings were sorted into living biomass (green, productive grass) and dead biomass (brown, dead grass/twigs) an visual inspection was used to record the green-dead proportions per sample point. The proportion of green, living biomass was subsequently used to scale both LAI and fAPAR, henceforth known as "green scaling". The scaled LAI essentially determined the proportion of LAI contributed to by green, living leaves, whereas the fAPAR scaling allowed identifying the proportion of fAPAR absorbed only by photosynthesising components. Sorted grass clippings were packaged in paper bags, weighed and dried in a forced-air oven at 60 °C to correct for water content. The length of drying time was determined using the "drying to a constant mass" method (Section 2.4.1.1). Once dried, clippings were weighed to determine DM mass and used for calibration during calculation of DM production (described in the subsequent section). The DM mass of the green-dead sorted clippings were also used for green

scaling, thus scaling of LAI and fAPAR was performed based on both visual inspection and DM mass.

3.1.2.4 Dry matter production

The DM production per sample point was derived with the use of a disc pasture meter (DPM), grass clippings and empirical relationships in literature (Bransby & Tainton 1977; Danckwerts & Trollope 1980; Zambatis et al. 2006). The standard procedure was followed for DPM measurements (Section 2.4.1.2), which involves releasing the disc onto the grass sward and reading the settling height off the central rod. Five DPM height measurements were taken and averaged for each sample site. At selected sites, all grass materials under the disc were clipped and placed in paper bags, oven-dried to constant mass and weighed (as described in Section 3.1.2.3). These mass measurements were used for calibration of the instrument (Bransby & Tainton 1977), where DM mass is converted to an appropriate unit (usually kg ha^{-1}) and used as the dependent variable in regression of mass against DPM height readings. The resulting regression equation was subsequently used to determine DM production for sample points where no clippings were collected. The final DM production estimates provided field data for the validation of model-estimated primary production estimates (Chapter 5). The sample points and collected field data for the wet season are visualised and summarised in Section 3.2.2.

3.2 DATA ANALYSIS

3.2.1 Dry season field trip: Fort Beaufort

The spatial distribution of the eight sample points for the dry season field trip to Fort Beaufort is shown in Figure 3.3, while Table 3.2 summarises the qualitative data collected for each sample point, with ID corresponding to the label of each point in Figure 3.3. Qualitative data include LAI, FVC and level of WPE.

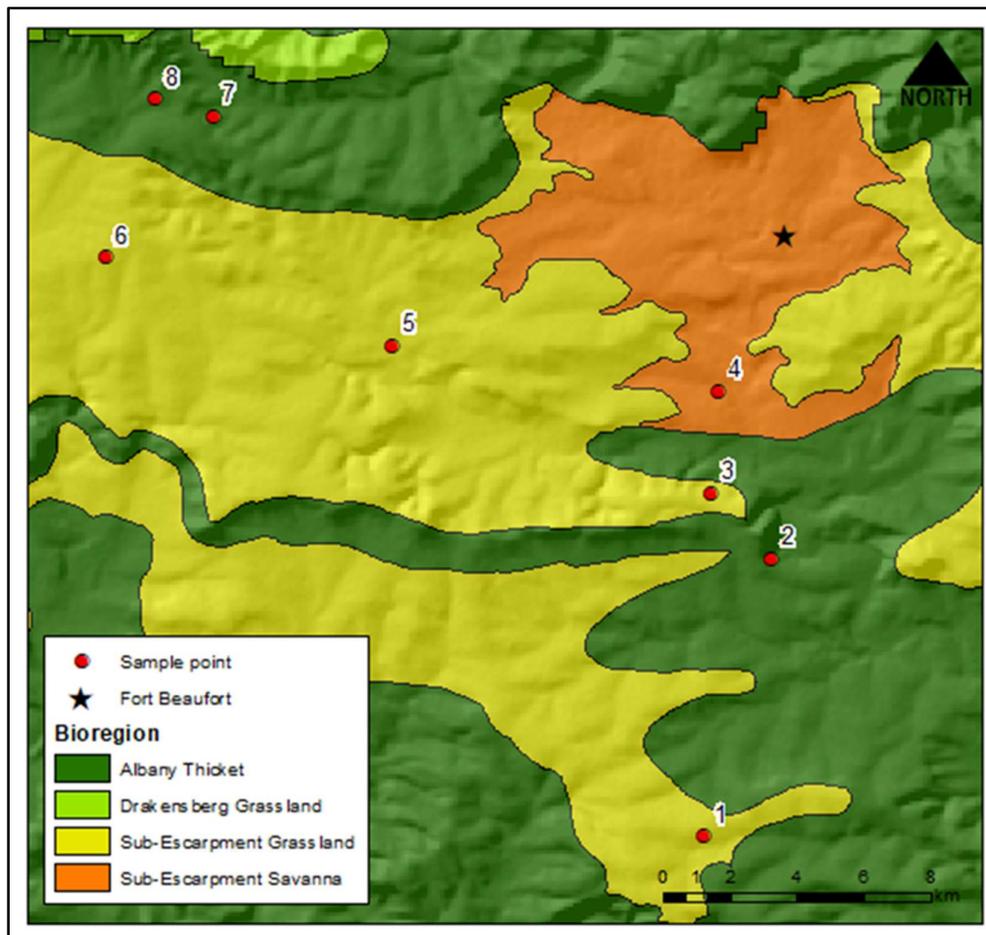


Figure 3.3 Sample points and corresponding bioregions for the dry season field trip to Fort Beaufort

Table 3.2 Summary of sample data for the dry season field trip to Fort Beaufort

ID	Bioregion	Mean LAI	Bare Cover (%)	Grass Cover (%)	Shrub Cover (%)	Tree Cover (%)	Level of WPE
1	Sub-Escarpment Grassland	0	100	0	0	0	Low
2	Albany Thicket	4.32	40	0	0	60	High
3	Sub-Escarpment Grassland	3.10	0	60	40	0	Moderate
4	Sub-Escarpment Savanna	3.22	0	45	15	30	High
5	Sub-Escarpment Grassland	1.17	30	40	25	5	Low
6	Sub-Escarpment Grassland	3.84	0	100	0	0	None
7	Albany Thicket	4.23	0	40	10	50	High
8	Albany Thicket	6.9	15	20	30	25	High

The grazing fields with the highest percentage of woody components (shrubs, trees) were those located in the Albany Thicket and Sub-Escarpment Savanna bioregions. Grasslands in the Sub-Escarpment Grassland all had low, moderate or no encroachment (Figure 3.3). The different types and increasing levels of woody encroachments are shown in Figure 3.4. Each photo corresponds to a sample point identified in Table 3.2 ranging from (a) no encroachment at point 6, (b) moderate

woody encroachment of shrubs at point 3, (c) high woody encroachment by both shrubs and trees at point 5 and (d) high woody encroachment by mainly trees at point 4.

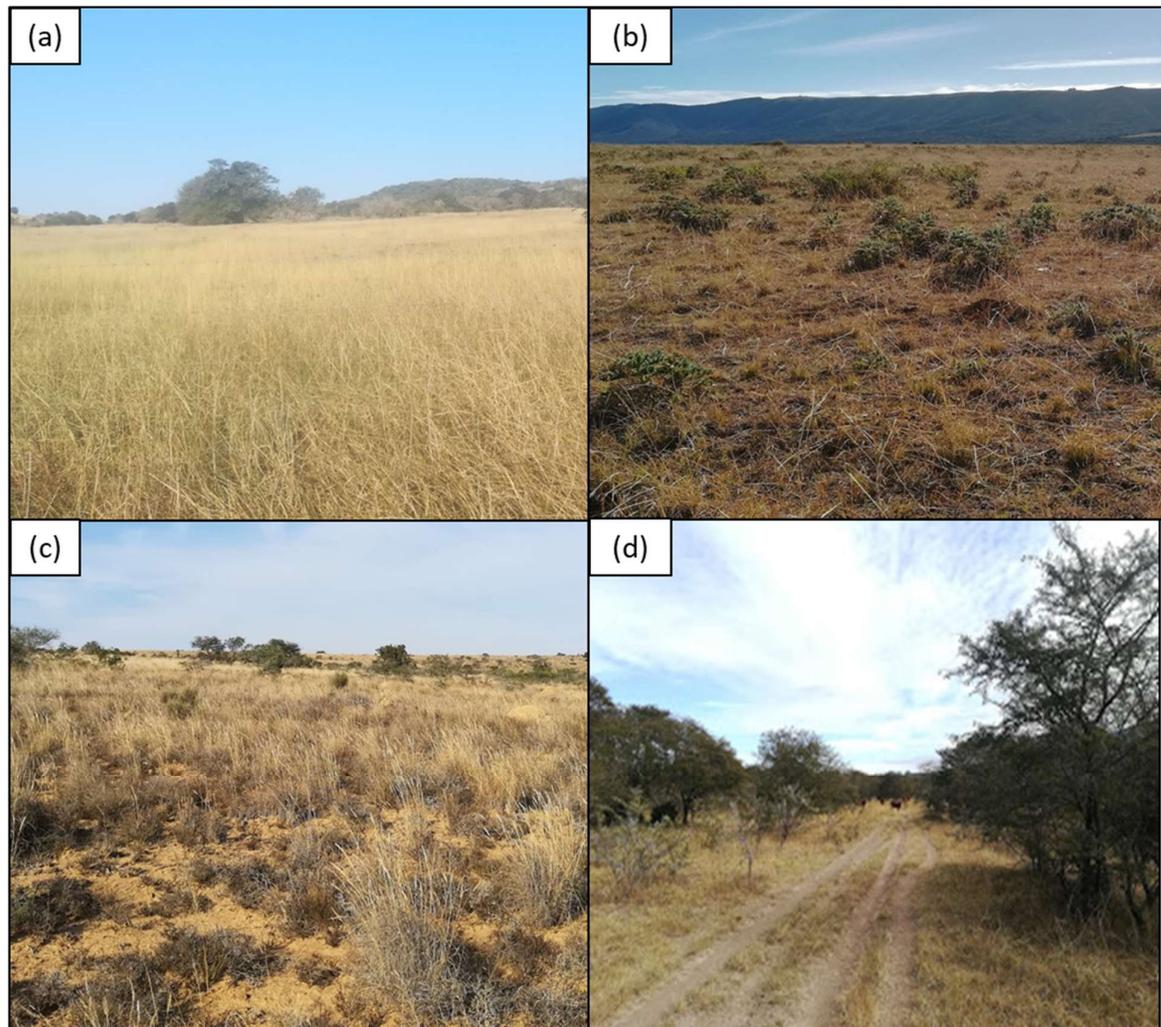


Figure 3.4 Examples of different types and levels of woody encroachment, namely (a) no woody encroachment, (b) moderate woody encroachment of shrubs, (c) high woody encroachment by both shrubs and trees and (d) high woody encroachment by mainly trees

This qualitative information was useful for determining different types and levels of encroachment, and identifying areas that are more prone to WPE (e.g. grazing fields located in the Albany Thicket and Sub-Escarpment Savanna bioregions). The FVC recorded for each point was used to validate FVC estimates (Chapter 4).

3.2.2 Growing season field trip: Cedarville

The distribution of sampling points for the growing season field trip to Cedarville is shown in Figure 3.5. Sample points were divided into homogeneous sample sites based on accessibility and NDVI intervals, including (a) the lower, flat plains, (b) communal, overgrazed field, (c/d/e) respective farms and (f) bare soil.

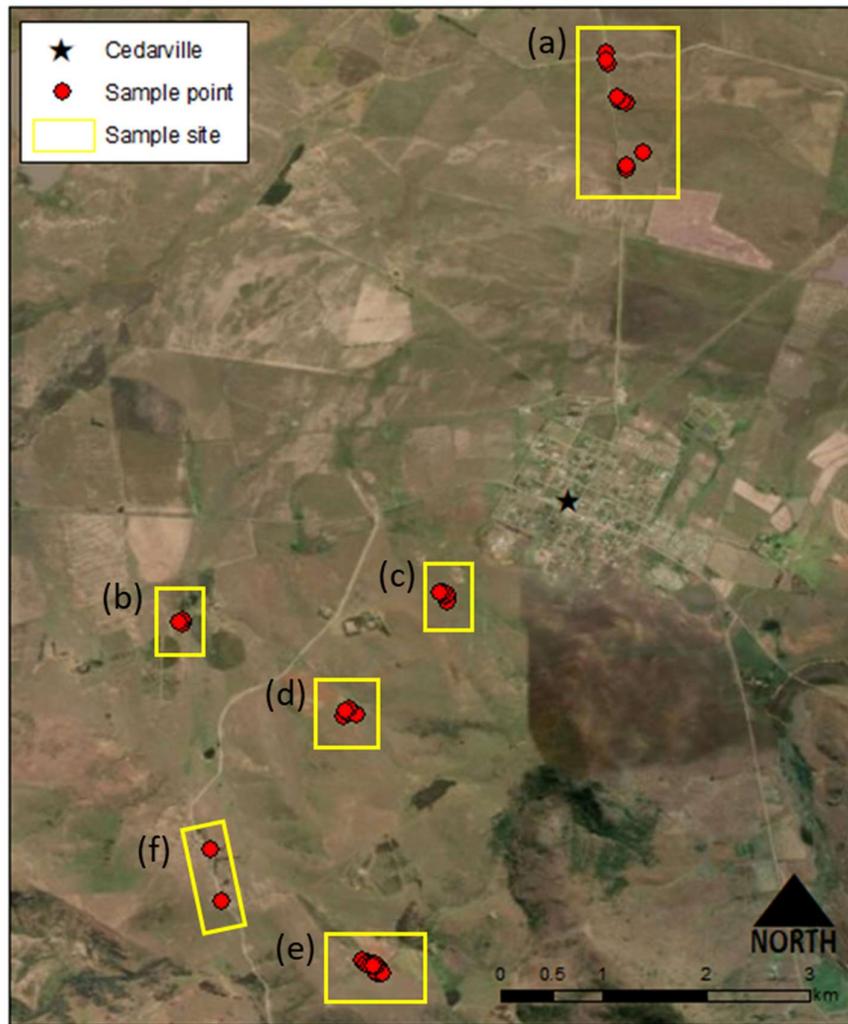


Figure 3.5 Sample sites for growing season field trip in Cedarville study area, namely (a) plains, (b) communal field, (c) farm 1, (d) farm 2, (e) farm 3 and (f) bare road.

Regression analysis was performed to determine links between NDVI and the various collected biophysical parameters, namely LAI, fAPAR, FVC and DM production (kg ha^{-1}). As heavy rainfall led to notable changes in vegetation between the beginning and the end of the month, NDVI was calculated for all cloud-free Sentinel-2 imagery for February 2019 to develop relationships in both dry and wet conditions. NDVI was thus calculated for the following dates in February 2019: 1, 6, 8, 13, 18, 23, 26 and 28.

3.2.2.1 NDVI-to-LAI

All points for which grass clippings were collected were plotted. Outlier points, located in the lower, flat plains of the study area or at the foot of a steep hill where water run-off accumulates, were identified and removed. Due to heavy rainfall before and during the field trip period, these locations experienced flooded, marsh-like conditions and/or unnaturally dense and high grass growth, which led to distorted NDVI and LAI values. Neither type of location would be suitable

for grazing of cattle, thus these outliers were removed. For the regression analysis, NDVI served as independent variable and LAI (unscaled and scaled) as dependent variable. The green scaled LAI essentially represent the proportion of LAI contributed to by green, living leaves. Figure 3.6 shows regression relationships between NDVI and unscaled LAI (blue), LAI scaled using visual inspection (orange) and LAI scaled using DM mass (grey) (Section 3.1.2.3). Regressions are shown for the specified dates, starting from the beginning of the month (Figure 3.6a) to the end of the month (Figure 3.6h).

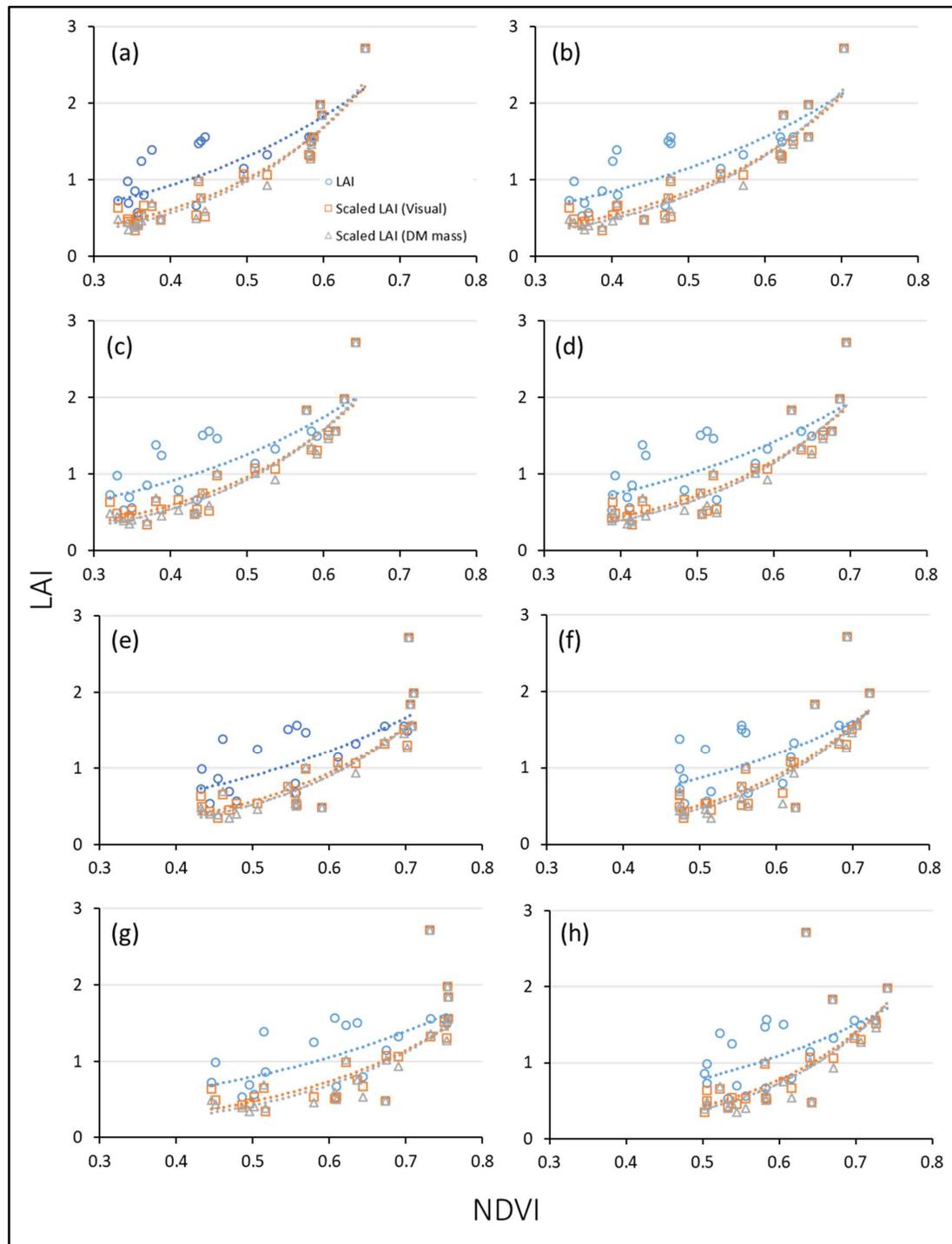


Figure 3.6 Regression analysis for NDVI and unscaled LAI (blue), visually scaled LAI (orange) and DM mass scaled LAI (grey) for (a) 2019/02/01, (b) 2019/02/06, (c) 2019/02/08, (d) 2019/02/13, (e) 2019/02/18, (f) 2019/02/23, (g) 2019/02/26 and (h) 2019/02/28.

The overall increase of NDVI visualised in Figure 3.1b is reflected in the regression analysis, with the lowest NDVI values changing from 0.35 (Figure 3.6a) to 0.5 (Figure 3.6h) during the sampled

month. Table 3.3 shows the resulting R^2 values and equations for the regression analysis of NDVI and LAI visualised in Figure 3.6.

Table 3.3 Results of regression analyses to determine the relationship between NDVI and LAI

NDVI by date	R^2			Equation		
	LAI	Scaled LAI (Visual)	Scaled LAI (DM mass)	LAI	Scaled LAI (Visual)	Scaled LAI (DM mass)
2019/02/01	0.61	0.87	0.91	$y = 0.239e^{3.3926x}$	$y = 0.0818e^{5.0414x}$	$y = 0.0632e^{5.4835x}$
2019/02/06	0.60	0.86	0.91	$y = 0.2522e^{3.0382x}$	$y = 0.0886e^{4.5159x}$	$y = 0.0682e^{4.9329x}$
2019/02/08	0.58	0.85	0.89	$y = 0.2466e^{3.2593x}$	$y = 0.0843e^{4.8787x}$	$y = 0.0646e^{5.3273x}$
2019/02/13	0.53	0.83	0.87	$y = 0.2171e^{3.1317x}$	$y = 0.0633e^{4.8684x}$	$y = 0.0478e^{5.2948x}$
2019/02/18	0.46	0.79	0.81	$y = 0.192e^{3.0892x}$	$y = 0.0459e^{5.0312x}$	$y = 0.0344e^{5.4346x}$
2019/02/23	0.34	0.73	0.73	$y = 0.183e^{3.1128x}$	$y = 0.0306e^{5.6314x}$	$y = 0.0234e^{5.9955x}$
2019/02/26	0.42	0.68	0.73	$y = 0.1985e^{2.7829x}$	$y = 0.052e^{4.4181x}$	$y = 0.0376e^{4.847x}$
2019/02/28	0.29	0.64	0.64	$y = 0.1604e^{3.2013x}$	$y = 0.022e^{5.9355x}$	$y = 0.0165e^{6.3188x}$

Overall, the green scaling of LAI notably improved the NDVI to LAI relationships, as seen by the increased R^2 values (Table 3.3). The LAI scaled using the weighed DM mass performed the best, with a max R^2 of 0.91 compared to 0.61 and 0.87 of unscaled and visually scaled LAI respectively. Regressions for the first half of the month performed better compared to later dates, with decreasing R^2 values as the month comes to an end. Total rainfall for the Cedarville study site was extracted using Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) daily rainfall data (Funk et al. 2015). Figure 3.7 plots the daily rainfall (mmday^{-1}), mean NDVI and R^2 values for the scaled LAI (DM mass) to show the impact of rainfall on the results.

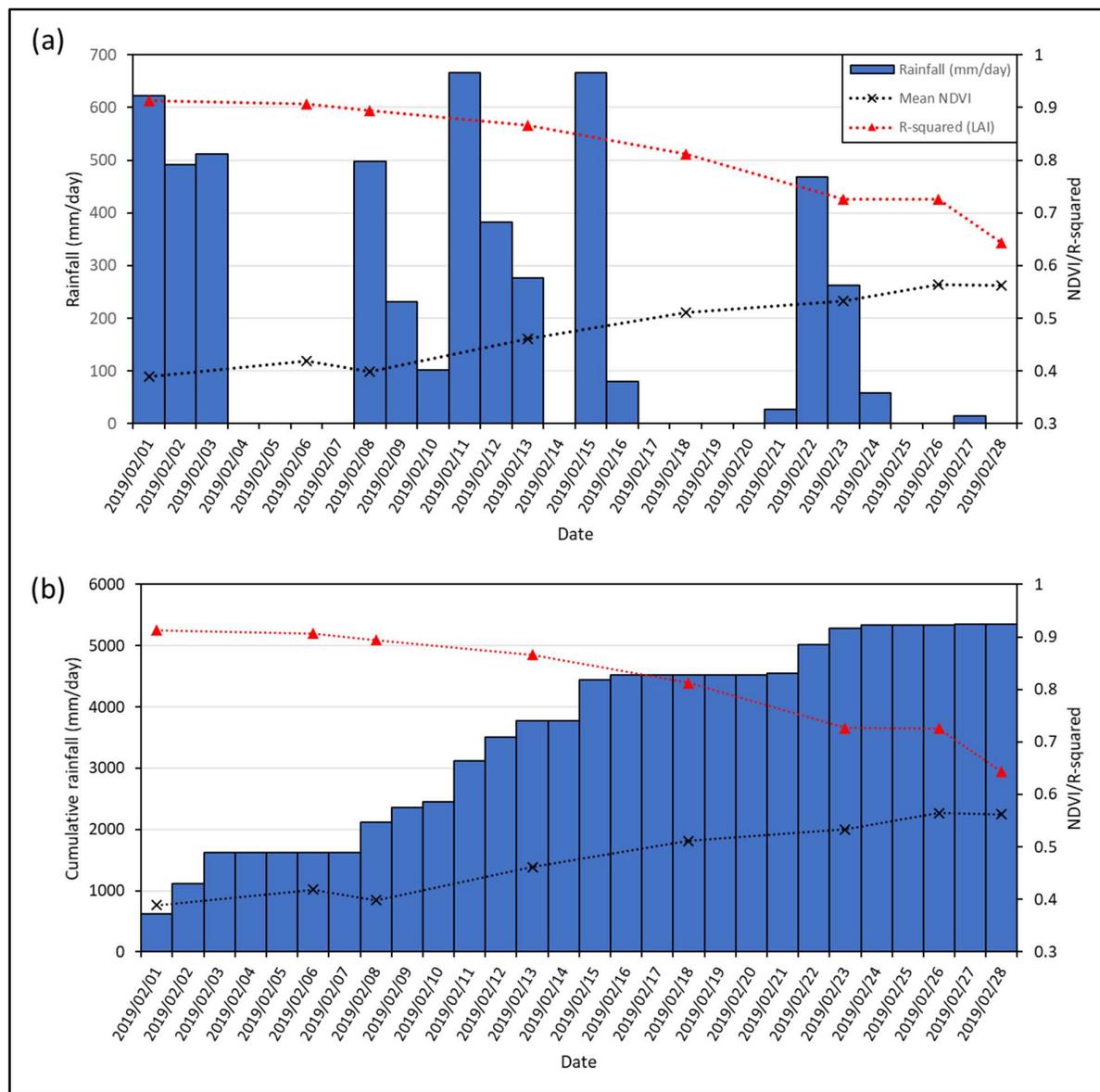


Figure 3.7 Mean NDVI and scaled LAI (DM mass) R^2 values for (a) daily rainfall (mmday^{-1}) and (b) cumulative rainfall (mmday^{-1}) for February 2019

When comparing daily rainfall (mmday^{-1}) with mean NDVI (black) and NDVI-to-LAI R^2 values (red) (Figure 3.7a), there seems to be no noticeable connection. However, the comparison between cumulative rainfall (mmday^{-1}) and the specified variables (Figure 3.7b) show an increase in mean NDVI (black) and a decrease in NDVI-to-LAI R^2 values (red) with increasing cumulative rainfall. As rainfall accumulates (Figure 3.7b), the region becomes wetter and greener, with not enough days between rainfall events for soil to drain and conditions to dry out (Figure 3.7a). This potentially leads to flooding in fields. It is also known that NDVI “saturates” at high values (Liu & Huete 1995), causing poorer NDVI to LAI relationships in the latter half of February.

Figure 3.8 shows the mean sampled LAI per NDVI interval for the temporally closest Sentinel-2 image (2019/02/26), with error bars representing one standard deviation. The error bars can be

used to compare variation within NDVI intervals for unscaled LAI (Figure 3.8a), visually scaled LAI (Figure 3.8b) and LAI scaled using DM mass (Figure 3.8c). Results were only plotted for NDVI values 0.4 and higher, as lower NDVI intervals correlated with overgrazed, degraded and bare areas.

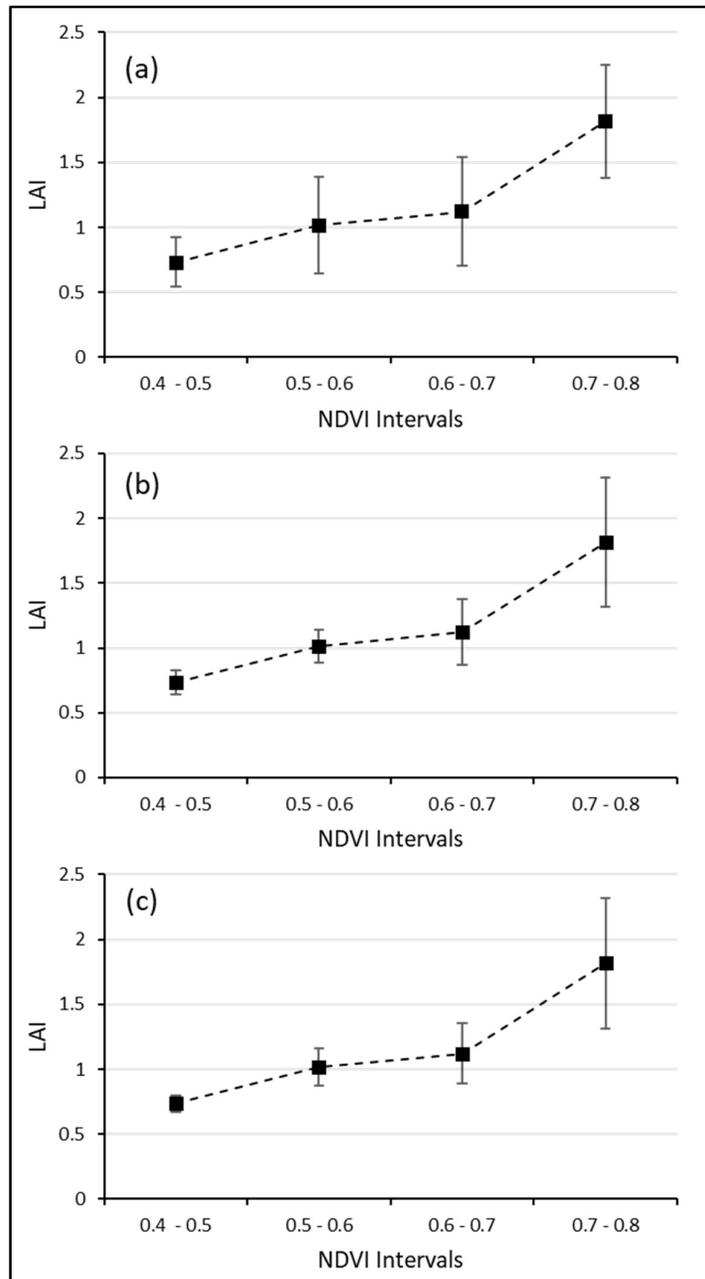


Figure 3.8 Mean LAI per NDVI interval for (a) unscaled LAI, (b) scaled LAI (visual) and (c) scaled LAI (DM mass)

Variation within almost all NDVI intervals decreases with green-dead biomass scaling of LAI, as shown by the smaller error bars (Figure 3.8c). However, variation in the interval 0.7 – 0.8 is the highest and does not show a notable decrease with scaling. The green scaling of LAI thus has its limitations in improving NDVI-to-LAI relationships when working with highly productive, potentially “saturated” NDVI values higher than 0.7.

3.2.2.2 NDVI-to-fAPAR

As with LAI, all points for which grass clippings were collected were plotted and the same outliers were identified and removed. NDVI served as independent variable and fAPAR (unscaled and scaled) as dependent variable. The green scaled fAPAR essentially represent fAPAR absorbed only by photosynthesising components. Regression relationships between NDVI and fAPAR (scaled and unscaled) were very similar to those observed for NDVI and LAI (Figure 3.6) and are included in Appendix C. Table 3.4 shows the resulting R^2 values and equations for the regression analysis of NDVI and fAPAR.

Table 3.4 Results of regression analyses to determine the relationship between NDVI and fAPAR

NDVI by date	R^2			Equation		
	fAPAR	Scaled fAPAR (Visual)	Scaled fAPAR (DM mass)	fAPAR	Scaled fAPAR (Visual)	Scaled fAPAR (DM mass)
2019/02/01	0.55	0.73	0.84	$y = 0.1518e^{2.5335x}$	$y = 0.0515e^{4.2082x}$	$y = 0.0401e^{4.626x}$
2019/02/06	0.56	0.74	0.84	$y = 0.1593e^{2.2449x}$	$y = 0.0563e^{3.7078x}$	$y = 0.0441e^{4.0837x}$
2019/02/08	0.56	0.74	0.84	$y = 0.1565e^{2.4054x}$	$y = 0.0545e^{3.9802x}$	$y = 0.0428e^{4.3744x}$
2019/02/13	0.50	0.73	0.82	$y = 0.1428e^{2.3097x}$	$y = 0.0428e^{3.9913x}$	$y = 0.0331e^{4.369x}$
2019/02/18	0.45	0.71	0.79	$y = 0.1282e^{2.3108x}$	$y = 0.0319e^{4.1737x}$	$y = 0.0245e^{4.5317x}$
2019/02/23	0.36	0.69	0.75	$y = 0.1165e^{2.4348x}$	$y = 0.0208e^{4.8321x}$	$y = 0.0162e^{5.1621x}$
2019/02/26	0.39	0.60	0.69	$y = 0.1311e^{2.0932x}$	$y = 0.0342e^{3.7386x}$	$y = 0.0251e^{4.1373x}$
2019/02/28	0.33	0.64	0.69	$y = 0.1007e^{2.5716x}$	$y = 0.0153e^{5.1288x}$	$y = 0.0118e^{5.4689x}$

Overall, the green scaling of fAPAR notably improved the NDVI to fAPAR relationships, as seen by the increased R^2 values (Table 3.4). The scaling of fAPAR using DM mass performed the best, with a max R^2 of 0.84 compared to 0.56 and 0.74 of unscaled and visually scaled LAI respectively. Similar to LAI, regressions for the first half of the month performed better compared to later dates, with decreasing R^2 values as the month comes to an end. It can be assumed that the poorer relationships are also due to wetter and greener conditions caused by heavy rainfall, as discussed in Section 3.2.2.1.

Figure 3.9 shows the mean fAPAR per NDVI interval for the temporally closest Sentinel-2 image (2019/02/26), with error bars representing one standard deviation. The error bars can be used to compare variation within NDVI intervals for unscaled fAPAR (Figure 3.9a), visually scaled fAPAR (Figure 3.9b) and fAPAR scaled using DM mass (Figure 3.9c). Results were only plotted for NDVI values 0.4 and higher, as lower NDVI intervals correlated with overgrazed, degraded and bare areas.

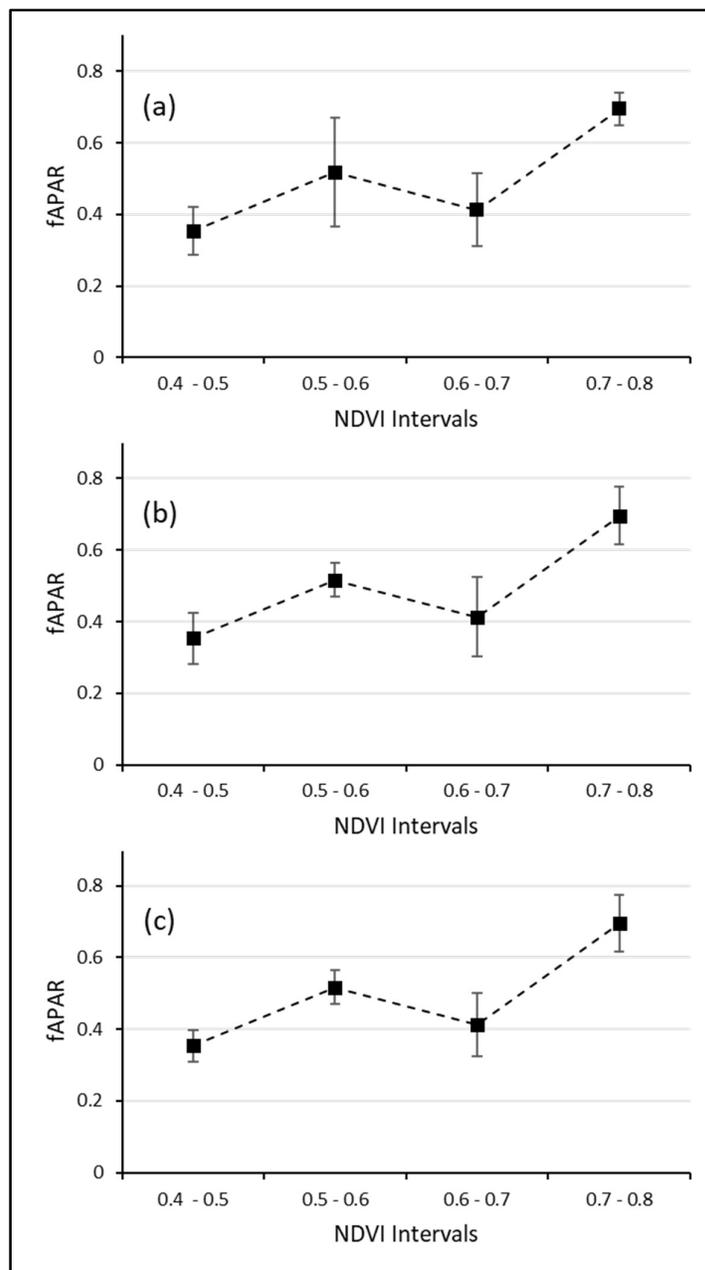


Figure 3.9 Mean fAPAR per NDVI interval for (a) unscaled fAPAR, (b) scaled fAPAR (visual) and (c) scaled fAPAR (DM mass)

Variation of fAPAR within NDVI intervals are initially already small (Figure 3.9), compared to the variation of LAI within the NDVI intervals (Figure 3.8). Variation within almost all NDVI intervals decreases even more with green scaling of fAPAR, as shown by the smaller error bars (Figure 3.9). However, variation in the interval 0.6 - 0.7 remains high even after scaling. As fAPAR calculation from the AccuPAR Ceptometer relies on the ratio of below canopy PAR measurements (Tau) (METER Inc. 2018), heavy cloud cover can influence Tau measurements and subsequent fAPAR calculations. This could explain the larger variation in fAPAR within the 0.6 - 0.7 NDVI interval, which was sampled on a cloudy day.

3.2.2.3 NDVI-to-FVC (%)

Regression analysis was used to determine the potential relationship between NDVI (independent variable) and the proportion of bare soil, herbaceous vegetation and woody vegetation within a pixel (dependent variables), termed FVC in this study. Figure 3.10 shows the regression analyses for NDVI (calculated from temporally closest Sentinel-2 image i.e. 2019/02/26) and bare cover (%) (Figure 3.10a), grass cover (%) (Figure 3.10b) and shrub cover (%) (Figure 3.10c).

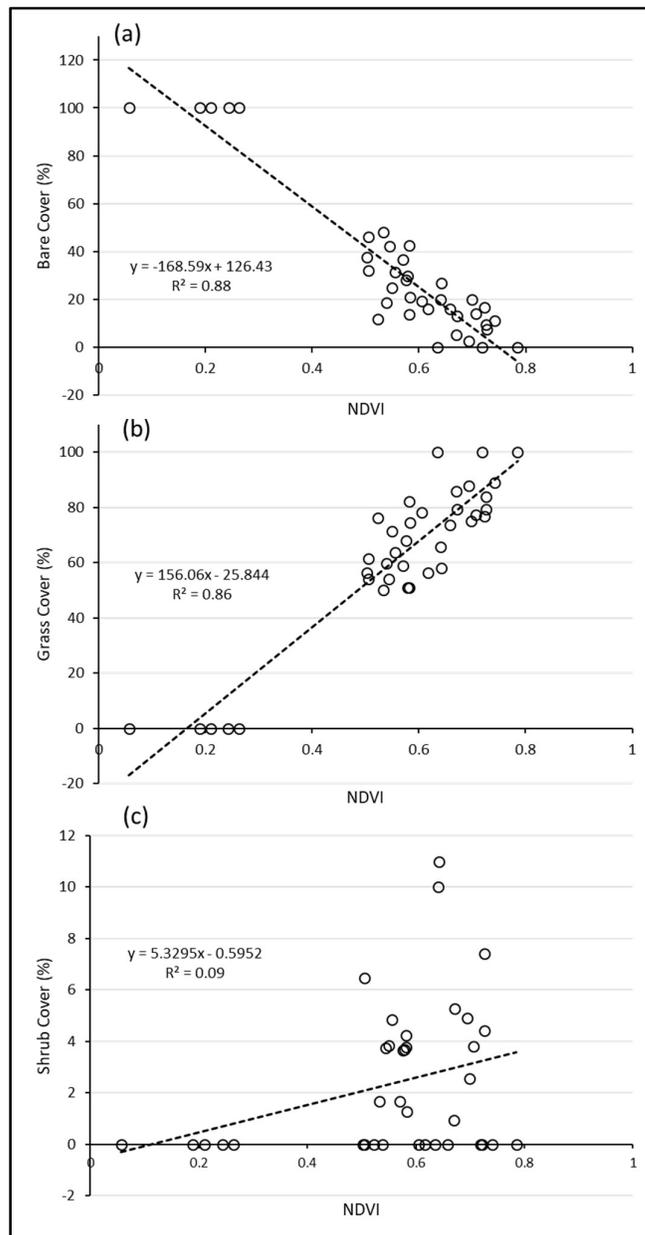


Figure 3.10 Regression analysis for NDVI and (a) bare cover (%), (b) grass cover (%) and shrub cover (%)

Strong linear relationships were identified between NDVI and bare and grass cover (Figure 3.10a, Figure 3.10b). This supports the use of NDVI for discrimination between fraction of bare soil and

herbaceous vegetation within a pixel during FVC derivation, discussed in Chapter 4. However, no strong relationship was found between NDVI and shrub cover (Figure 3.10c).

3.2.2.4 DPM-to-DM production

The calibration regression for the DPM height (cm) to DM production (kg ha^{-1}) is shown in Figure 3.11. Square root transformation was applied to DM production to improve the linear relationship with DPM height measurements (Danckwerts & Trollope 1980; Zambatis et al. 2006).

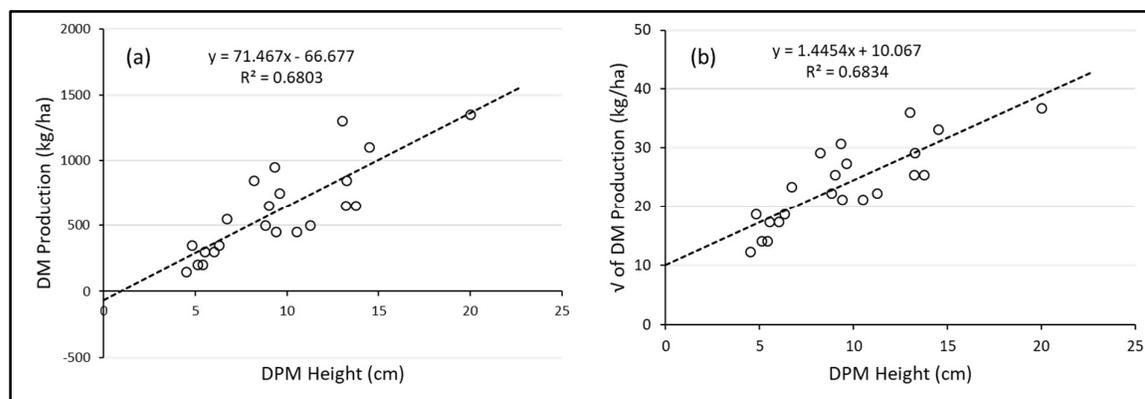


Figure 3.11 Calibration regressions for DPM height (cm) and (a) DM production (kg ha^{-1}) and (b) the square root of DM production (kg ha^{-1})

Although the R^2 value improved slightly with a square root transformation of DM production (Figure 3.11b), the linear relationship between the sampled DPM heights and DM production was not as strong as desired. This is most likely due to the low number of samples and can be improved in future DPM calibration procedures by collecting more grass clippings per field trip.

3.3 DISCUSSION

This chapter described the various biophysical parameters collected and analysed during two conducted field trips. Important relationships between field data and satellite observation were investigated and established, which forms a crucial component of building transferable, dynamic models for characterising vegetation dynamics using RS.

Analysis of the qualitative information collected during the dry season field trip in Fort Beaufort characterised different types and levels of encroachment and identified areas at higher risk of WPE. Grasslands in the Albany Thicket and Sub-Escarpment Savanna bioregions had higher percentages of woody components compared to those located in the Sub-Escarpment Grassland (Table 3.2), illustrating the increased risk of WPE and need for continuous monitoring. The qualitative information, combined with the statistical analysis of the Cedarville field data, established important biophysical links for FVC estimation and vegetation production modelling.

An important link for FVC estimation is the NDVI-to-FVC relationship, explored in Section 3.2.2.3. Although a strong relationship exists between NDVI and bare (Figure 3.10a) and grass cover (%) (Figure 3.10b), the R^2 values for NDVI-to-shrub cover (%) was very poor (0.09) (Figure 3.10c). As it is essential to discriminate between herbaceous productivity and woody productivity when modelling grazing systems, additional indices and band combinations must be explored to distinguish shrubby components in grasslands. In addition, a simple VI-based FVC estimation model (Section 2.3.2.1) is not suitable for describing complex, heterogeneous landscapes such as savannas and WPE grasslands.

Another important link is the relationship between NDVI and the biophysical parameters LAI and fAPAR. The scaling of LAI and fAPAR using proportions of green, living vegetation improved their regression relationships with NDVI (Figure 3.6; Table 3.3; Table 3.4) and decreased variation in measurements per NDVI interval (Figure 3.8; Figure 3.9). As fAPAR and LAI is notoriously difficult to model (Tian et al. 2002), the use of green scaling is a promising technique for improving vegetation modelling. Many vegetation productivity models rely on fAPAR as a key input parameter (Section 2.2.1). It is either included in the form of a modelled product (e.g. MODIS LAI/fAPAR product) or derived from NDVI using a standard, general empirical equation (Running et al. 2004). The improvement of NDVI-to-fAPAR relationship using green scaling is thus a key outcome of this study, as it will allow replacement of a general empirical equation with a calibrated regression relationship to calculate fAPAR from NDVI.

Although NDVI-to-LAI/fAPAR relationships were relatively good overall, R^2 showed a notable decrease with the accumulation of rainfall (Figure 3.7). This decline in R^2 is likely a combination of two factors: flooded, marsh-like fields distorting spectral reflectance and increasing greenness of vegetation leading to NDVI “saturation”. This illustrates the detrimental effect of wet conditions on field-to-satellite relationships, especially when using NDVI for modelling and mapping vegetation dynamics. It supports the need for incorporating some form of water stressor in vegetation production modelling (Section 2.2.1.3), to compensate for the effects of increased rainfall (and drought) on spectral reflectance and biophysical relationships.

The results of this chapter were used for calibration and validation purposes in subsequent chapters. The FVC field data (Section 3.1.1.1; Section 3.1.2.1) played an important role in evaluating model-estimated FVC (Chapter 4), with respect to both the accuracy and transferability of the model. The relationship between NDVI and fAPAR (scaled and unscaled) (Section 3.2.2.2) formed an important component of the grassland productivity model (Chapter 5), while the DM production measurements (Section 3.1.2.4) provided validation data for model-estimated primary production estimates (Chapter 5).

CHAPTER 4: FRACTIONAL VEGETATION COVER ESTIMATION

Fractional vegetation cover (FVC) estimation involves determining the proportional area of different land cover (LC) types a pixel consists of, based on various spectral characteristics of the pixel (Jiménez-Muñoz et al. 2009). FVC was estimated from remote sensing (RS) data and evaluated to demonstrate its potential use in grassland productivity modelling. This chapter discusses the linear spectral mixture model (LSMM) used to estimate FVC, thereby addressing Objective 3. The resultant FVC will be used in Chapter 5 to improve grassland productivity modelling. Firstly, the input data collection and preparation is described, followed by the model calibration, implementation and validation, concluding with results and discussion.

4.1 METHODS & MATERIALS

4.1.1 Data selection & preparation

4.1.1.1 Satellite imagery selection and pre-processing

Both Landsat 8 and Sentinel-2 imagery were used for FVC estimation. Sentinel-2 was selected for its higher spatial and temporal resolution and is acquired from 2016 onwards. Unfortunately, existing FVC products used for comparison (Buchhorn et al. 2019) are only available for the year 2015. Landsat 8 was therefore included to enable comparison of the LSMM results to existing FVC products (Section 4.1.3.1).

Image collections were filtered in Google Earth Engine (GEE) to include only images with less than 5% cloud cover. Urban areas, water, crops and mines were masked out using the National Land Cover 2017/2018 (NLC2017/8) product (Land Resources International 2018). No additional image pre-processing (e.g. orthorectification, radiometric correction, atmospheric correction etc.) was required, as the GEE data catalogue provides both raw and processed imagery for the various satellite sensors. The Landsat 8 Surface Reflectance Tier 1 and Sentinel-2 Level 2A (Surface Reflectance) datasets were used for further processing.

The centre wavelength (CW), minimum and maximum wavelength, bandwidth (BW) and spatial resolution (Res.) for corresponding bands differ notably between Landsat 8 (L8) and Sentinel-2 (S2) (Table 4.1). Figure 4.1 visually illustrates these differences in spectral placements of Landsat 8 and Sentinel-2 bands.

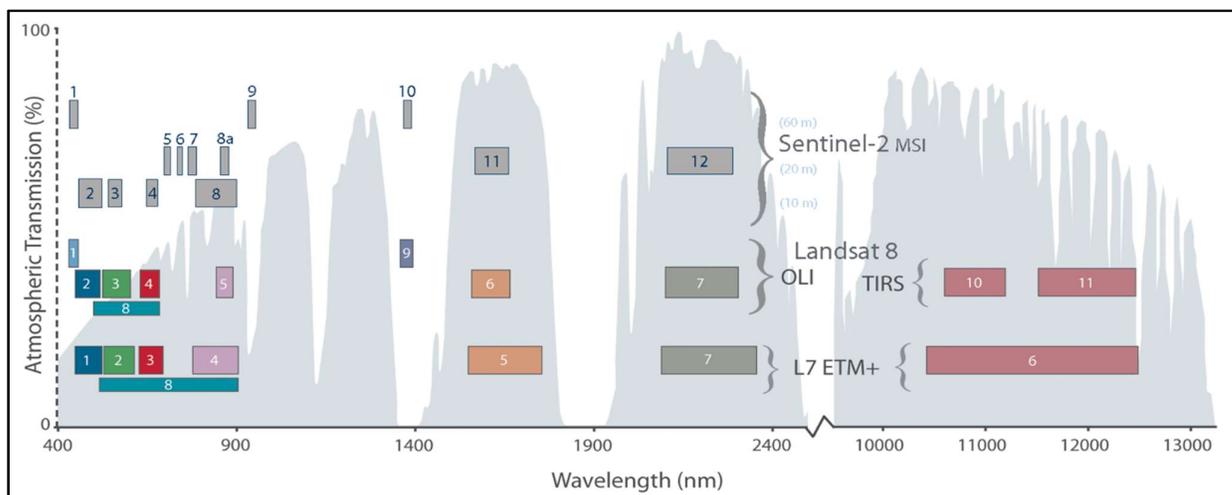
Table 4.1 Comparison of Landsat 8 (L8) and Sentinel-2 (S2) corresponding bands

L8 ^a	CW (µm)	Wavelength (min-max)	BW	Res. (m)	S2 ^b	CW (µm)	Wavelength (min-max)	BW	Res. (m)	
1	C/A	0.443	0.445-0.451	0.016	30	1 C/A	0.443	0.421-0.457	0.036	60
2	Blue	0.482	0.452-0.512	0.06	30	2 Blue	0.494	0.439-0.535	0.096	10
3	Green	0.561	0.533-0.590	0.057	30	3 Green	0.56	0.537-0.582	0.045	10
4	Red	0.655	0.636-0.673	0.037	30	4 Red	0.665	0.646-0.685	0.039	10
						5 VRE	0.704	0.694-0.714	0.02	20
						6 VRE	0.74	0.731-0.749	0.018	20
						7 VRE	0.781	0.768-0.796	0.028	20
						8 NIR	0.834	0.767-0.908	0.141	10
5	NIR	0.865	0.851-0.879	0.028	30	8 a NIR	0.864	0.848-0.881	0.033	20
						9 WV	0.944	0.931-0.958	0.027	60
9	Cirrus	1.373	1.363-1.384	0.02	30	10 Cirrus	1.375	1.338-1.414	0.076	60
6	SWIR	1.609	1.567-1.651	0.085	30	11 SWIR	1.612	1.539-1.681	0.142	20
7	SWIR	2.201	2.107-2.294	0.187	30	12 SWIR	2.194	2.072-2.312	0.24	20
8	Pan	0.59	0.503-0.676	0.172	15					
10	TIRS	10.895	10.60-11.19	0.59	100					
11	TIRS	12.005	11.50-12.51	1.01	100					

^a C/A, coastal/aerosol; NIR, near-infrared; SWIR, shortwave infrared; Pan, panchromatic; TIRS, thermal infrared

^b C/A, coastal/aerosol; VRE, vegetation red edge; NIR, near-infrared; WV, water vapour; SWIR, shortwave infrared

Adapted from ESA (2013); USGS (2016)



Source: USGS (2015)

Figure 4.1 Comparison of Landsat 7 and 8 bands with Sentinel-2

Landsat 8 generally has wider bandwidths than Sentinel-2 (Table 4.1; Figure 4.1), with the exception of the NIR band where the commonly used band 8 of Sentinel-2 is wider than band 5 of Landsat 8. A range of vegetation and soil indices discussed in Section 2.1.2 were calculated to include in LSMM calibration. The widely used normalised difference vegetation index (NDVI) (Tucker 1979) was chosen to distinguish between vegetated and non-vegetated surfaces (Equation 2.1). The enhanced vegetation index (EVI), a modification of the NDVI, was also calculated to potentially improve FVC estimation in areas of high biomass (Liu & Huete 1995) (Equation 2.2). To address the limitation of NDVI in areas with high degrees of exposed soil (Qi et al. 1994), thus potentially improving discrimination between vegetation and bare ground, a modified soil adjustment vegetation index (MSAVI) was calculated. The newer version of the index, MSAVI2 (Equation 2.3) allows the implementation of the index without specifying a soil brightness correction factor, which requires prior knowledge of vegetation cover. The final index calculated was the newly developed dry bare-soil index (DBSI) (Equation 2.4). The index was specifically developed to identify bare areas in dry climates (Rasul et al. 2018) and can improve differentiation between dry vegetation and bare soil in areas impacted by drought.

A multi-temporal stack of all selected image bands and indices was generated per analysis year for both Landsat 8 and Sentinel-2. Table 4.2 shows the different analysis years, the sensors used for each year, the study area and the analysis purpose of the different stacks. All image acquisition dates for the respective stacks for the different sensors and analysis years are summarised in Appendix D.

Table 4.2 Description of sensors, study area and analysis purpose of the different image stacks

Analysis Year	Sensors	Study Area	Purpose
2015	Landsat 8	Fort Beaufort	Model assessment (Section 4.1.3.1): comparison of LSMM FVC to existing FVC product
2018	Landsat 8, Sentinel-2	Fort Beaufort	Model calibration (Section 4.1.2.5): calibration of LSMM using field data Model assessment (Section 4.1.3.2): validation of LSMM FVC using in-situ field data
2019	Sentinel-2	Cedarville	Mode transferability (Section 4.1.4): evaluation of LSMM spatial and temporal transferability

4.1.1.2 Endmember field points

The supervised approach to endmember extraction (Jiménez-Muñoz et al. 2009) was implemented, which involves manually selecting endmember pixels based on a land use map (Section 2.3.2.2). Field points collected for a land cover/land use (LCLU) classification (10 m spatial resolution) of the Eastern Cape, South Africa in 2013/2014 (Lück et al. 2010), were inspected to identify potential endmember points. Firstly, LC cover classes corresponding to the endmember class

definitions of this study were identified. FVC endmember classes are most commonly defined as Bare Soil (BS), Photosynthetic Vegetation (PV) and Non-Photosynthetic Vegetation (NPV), where NPV includes both grassy vegetation and the leaves of trees (Scarath, Roder & Schmidt 2010) (Section 2.3). This definition is not ideal when estimating an FVC for grazing productivity, as cattle feed almost exclusively on grass biomass. To exclude vegetation not grazed by cattle, endmember classes for this research was defined as Bare (bare soil/rock), Grassy (herbaceous vegetation) and Woody (vegetation with woody components e.g. shrubs and trees). This definition allows for differentiation between the grazing potential of grass and woody components. All field points corresponding to the relevant LC classes were thus selected to represent endmembers. Furthermore, the points were manually inspected using the NLC2017/8 product and Google Earth aerial imagery to ensure all points corresponded to pure endmember pixels at both 30 m (Landsat 8) and 10 m (Sentinel-2) spatial resolution.

4.1.2 Linear spectral mixture model

The non-negative constrained least squares linear unmixing model (Scarath, Roder & Schmidt 2010) was calibrated from the selected satellite imagery to estimate FVC. Model calibration was performed using the Fort Beaufort study site (Section 1.5) using imagery acquired for the year 2018 (Table 4.2; Appendix D). Fort Beaufort was chosen due to the heterogeneous grasslands in the study area, caused by high levels of grassland degradation and woody plant encroachment (WPE). The relatively homogenous Cedarville area, consisting mainly of pristine grasslands, was unsuitable to accurately identify bare soil and woody components at sub-pixel level. The model was calibrated using an iterative process. In each iteration features were selected, endmember classes were defined, endmember spectral signatures were extracted and refined and the LSMM model was applied using these refined spectral signatures. This was followed by an evaluation of the estimated FVC results (Figure 4.2).

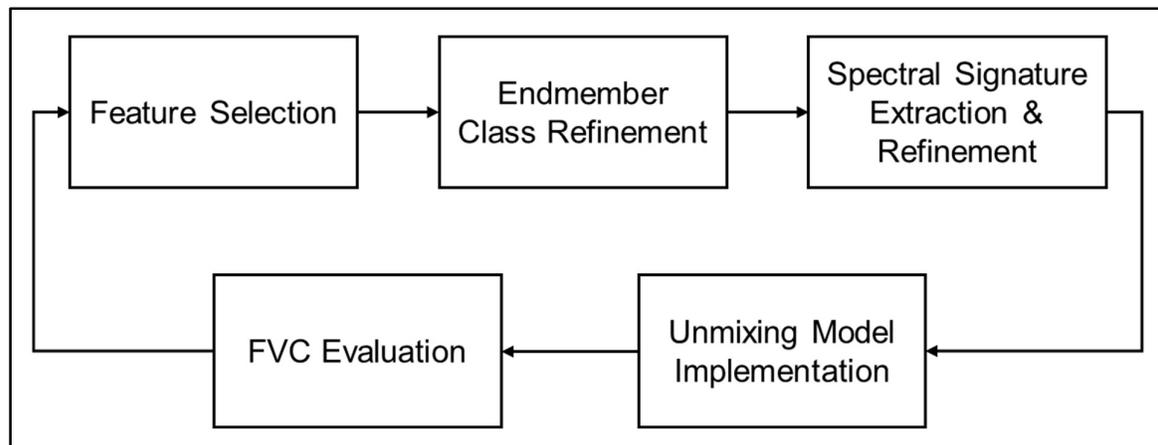


Figure 4.2 Iterative model calibration process

These steps are explained in detail in the subsequent sections. The results of the LSMM calibration and modelling process are summarised and visualised in Section 4.2.

4.1.2.1 Feature selection

Features were selected from calculated indices and band combinations. The Blue, Red, Green, near-infrared (NIR) and two shortwave infrared (SWIR) bands (Table 4.1) were selected as the initial band combination for unmixing. Different band and index combinations were used in each iteration to identify the optimal band combination for maximum separability, until a combination producing an FVC of suitable accuracy was found. Due to the notable differences in band characteristics for Landsat 8 and Sentinel-2 (Table 4.1; Figure 4.1), the same band and index combinations could not be used for both. Table 4.3 summarises the final band and index combinations for the respective satellite sensors.

Table 4.3 Final band and index combinations for Landsat 8 and Sentinel-2

	Landsat 8	Sentinel-2 Images
Bands	Blue, Green, Red, SWIR1, SWIR2	SWIR2
Indices	NDVI, EVI, MSAVI2, DBSI	NDVI, EVI, MSAVI2, DBSI

Sentinel-2 required fewer bands and indices to produce an FVC of equal or higher accuracy than Landsat 8. This is likely due to the coarser spatial resolution of Landsat 8, which necessitates more features to accurately discriminate between endmembers at sub-pixel level.

Band 8A, instead of band 8, was used as the NIR band for Sentinel-2. Although band 8 is more commonly used, it has a much wider bandwidth than the Landsat 8 NIR band (band 5) and resulted in flatter spectral signatures with lower separability between endmember classes (shown in Section 4.2.1). Based on bandwidth and centre wavelength, band 8A is more comparable to Landsat 8's band 5 (Table 4.1; Figure 4.1) and was consequently used as the NIR band during Sentinel-2 NDVI calculation. It must be noted that the Sentinel-2 band 8A has a 20 m spatial resolution (Table 4.1), resulting in a trade-off between increased separability and finer spatial resolution.

4.1.2.2 Endmember class refinement

Endmember classes were also split into subclasses to ensure greater separability between mean spectral signatures. Table 4.4 provides a description for each subclass. Different subclass combinations were identified and iteratively assessed. With each iteration, spectral characteristics of classes were plotted and analysed to identify spectral response patterns within the respective endmember classes.

Table 4.4 Final endmember class descriptions

Endmember Class	Endmember Subclass	Description
Bare	Bare 1	Loose, eroded bare soil and ground (typically in degraded grassland at low elevations)
	Bare 2	Bright, rocky bare ground (typically on mountain slopes)
Grassy	Grassy	Herbaceous, grassy vegetation
Woody	Shrubs	Shrubs and small bushes
	Trees	Trees and dense forests/woods

The Bare endmember class, in particular, was found to be too broad as different bare areas in the study site (e.g. bare soil, rocky slopes) had widely different spectral characteristics and responses. A mean spectral signature derived using all the different Bare endmember points did not provide a good representation of the different Bare cover types, thus the Bare class was split into two subgroups. The subdivision of Woody into Shrubs and Trees also ensured improved estimation of woody components in grasslands. Introduction of the Shrubs subclass did not only improve identifying woody components within pixels, but can also potentially be used to monitor the early stages of WPE in grasslands. The final and best performing subclass differentiation was Bare 1, Bare 2, Grassy, Shrubs and Trees. The improved separability between endmember spectral signatures with subdivision of endmember classes is shown in Section 4.2.1.

4.1.2.3 Spectral signature extraction and refinement

The LSMM model requires a mean spectral signature for each endmember class as input (Scarth, Roder & Schmidt 2010). These mean spectral signatures were derived by extracting spectral responses for all field points from the multi-temporal stack and calculating the mean reflectance per band, per endmember class. The mean spectral signatures were refined by inspecting the extracted spectral signatures and identifying outlier points that could potentially skew mean values. Field points with consistently extreme high or low values for multiple iterations were identified and consequently removed. As the most difficult part of SMA is finding quality, pure endmember spectra (Jiménez-Muñoz et al. 2009), this refinement process was thus essential in defining spectral signatures that best represent the spectral profile of each endmember. The mean spectral signatures for the respective satellite sensors and endmember classes are shown in Section 4.2.1.

4.1.2.4 Unmixing model implementation

A non-negative constrained least squares linear unmixing model (Scarth, Roder & Schmidt 2010) for the defined endmember classes was applied to the to estimate FVC. The input stack was

unmixed to produce a raster image, where the number of bands correlates to the number of endmember subclasses. The final unmixed product consisted of a 5-band layer, where each pixel showed the fraction of Bare 1, Bare 2, Grass, Shrubs and Trees. The two Bare bands (Bare 1, Bare 2) and two Woody bands (Shrubs, Trees) were summed to respectively create a Bare and a Woody band. The final FVC thus consisted of an RGB image, with red corresponding to fraction of Bare, green corresponding to fraction of Grassy and blue corresponding to fraction of Woody. The FVC output maps are shown in Section 4.2.2.

4.1.2.5 Fractional vegetation cover evaluation

The estimated FVC for each iteration was evaluated with field-classified FVC using a similar approach followed during the validation of the fraction layers of the Copernicus Dynamic Land Cover (CDLC) product (Tsenbazar et al. 2018). The CDLC cover fractions were evaluated using field data, where the sampling scheme consisted of a number of 100 m² plots divided into equal-sized 10 x 10 cell sample grid. Each sample cell was classified as a discrete LC class based on its majority LC. The number of cells per plot classified as a particular LC was then counted to determine that LC class's fractional cover value for the overall 100 m² plot. The CDLC validation scheme is illustrated in Figure 4.3, where a 100 m² field plot was classified as 53% tree cover (green), 18% shrub cover (orange) and 29% grass cover (yellow).



Source: Tsenbazar et al. (2018)

Figure 4.3 Validation approach for CDLC fractional layers. Green corresponds to tree cover, orange to shrub cover and yellow to grass cover.

A similar method of validation was simulated to compare the model results with field-classified FVC. Ten 100 m² plot sites were identified based on their vegetational structure so that a variety of grassy, woody, bare and mixed landscapes could be evaluated. The 100 m² plot size was chosen

to match the CDLC 100 m spatial resolution. These plot sites were then split into a 10 x 10 grid of 100 sample cells each. The free and open source land monitoring software Collect Earth (Bey et al. 2016) was used to classify each cell as either Bare, Grassy or Woody. This involved creating the grid sampling scheme, setting up a Collect Earth survey and importing the survey into the Collect Earth tool to compare each sampling cell to high resolution aerial imagery. As performed for the CDLC validation, a discrete LC class was recorded for each sample cell. In addition, the fraction of Bare, Grassy and Woody of each sample cell was also recorded to determine how the different classification methods might influence final FVC estimates. An FVC value was then estimated per plot using both the discrete and fractional LC classifications. For the discrete classifications, the same process as in Figure 4.3 was followed where the number of cells per plot is summed for each LC. For the fractional classifications, the fractional value per LC was averaged for all the cells within the plot. Figure 4.4 shows one of the field-classified FVC 100 m plots, where (a) represents the discrete classification and (b) represents the fractional classification.

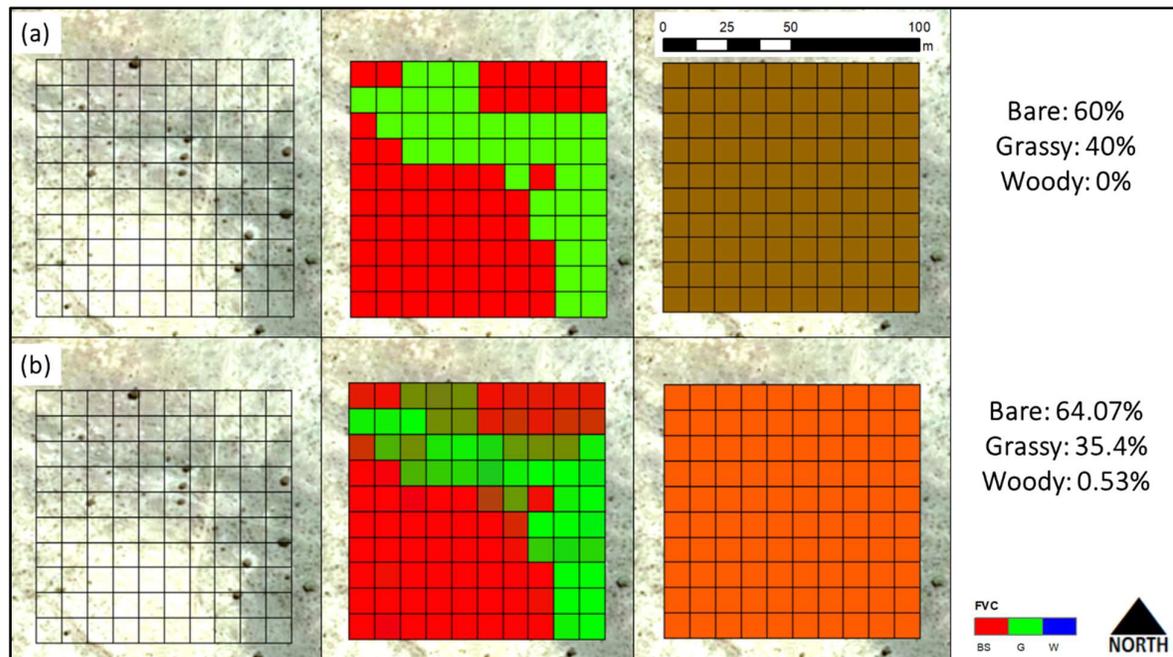


Figure 4.4 Comparison of field-classified FVC values using two classification methods, where (a) discrete classification and (b) fractional classification

The fractional field classification incorporates more detail into the overall FVC value, as shown by the inclusion of 0.53% woody (Figure 4.4b). As the discrete classification method involves assigning an LC class based on the majority cover within the 10 m² cell, the sparse trees and shrubs were not reflected in the overall FVC value for the plot (Figure 4.4a). The 10 field plots and their corresponding discrete and fractional cover classifications are shown in Table 4.5.

Table 4.5 Field FVC values for field plots using discrete and fractional classification

Plot #	Discrete			Fractional		
	Bare	Grassy	Woody	Bare	Grassy	Woody
1	0	9	91	0	13.15	86.85
2	0	4	96	0.	5.3	94.7
3	0	0	100	0	0	100
4	0	47	53	0	45.25	54.75
5	0	100	0	0	100	0
6	17	83	0	22.8	77.15	0.05
7	9	83	8	9.53	79.58	10.94
8	100	0	0	100	0	0
9	60	40	0	64.07	35.4	0.53
10	0	41	59	0.75	36.5	62.75

Once all field plots were classified, the discrete and fractional field FVC values could be used to evaluate the model-estimated FVC. For each iteration, the model-estimated FVC was resampled to the 100 m spatial resolution of the sampling plots and quantitatively compared to the discrete and fractionally classified field FVC using root mean square error (RMSE) and mean absolute error (MAE) (Congalton & Green 2009). This process was repeated (using various band and index combinations and endmember class definitions) until suitable RMSE and MAE values were achieved for both Landsat 8 and Sentinel-2. The final RMSE and MAE values for the model-estimated FVC vs. the discrete and field classified FVC is shown in Figure 4.5 and Figure 4.6 respectively.

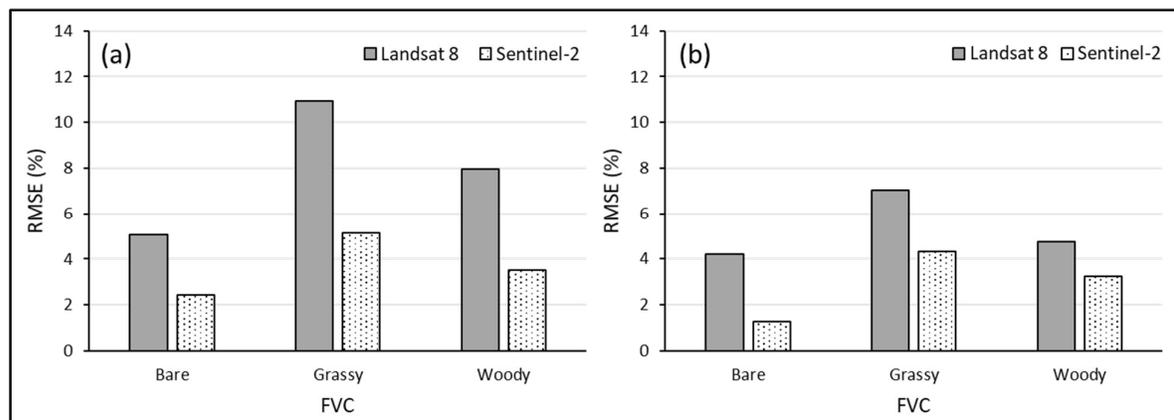


Figure 4.5 Comparison of RMSE values for model-estimated FVC to (a) discrete field FVC and (b) fractional field FVC

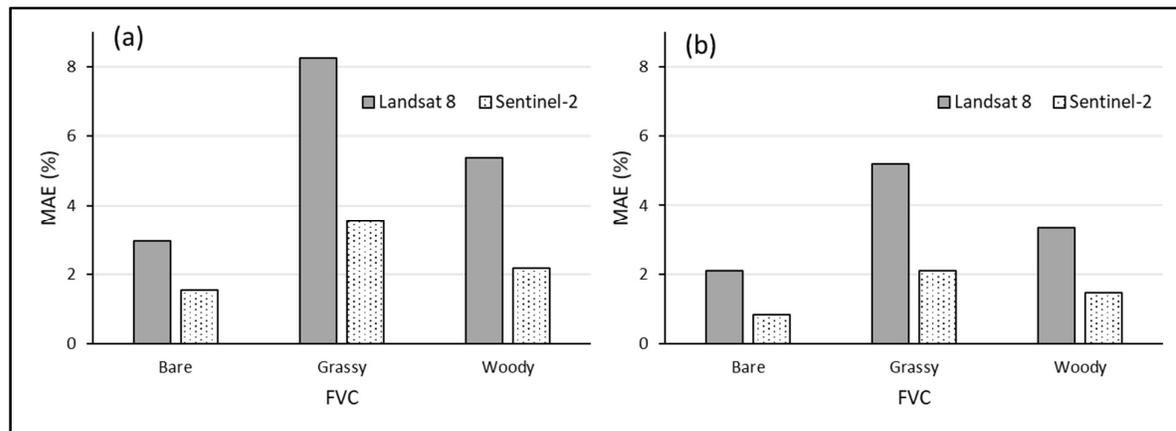


Figure 4.6 Comparison of MAE values for model-estimated FVC to (a) discrete field FVC and (b) fractional field FVC

FVC estimated using Sentinel-2 has lower RMSE and MAE values compared to Landsat 8 (Figure 4.5; Figure 4.6), as expected due to the finer spatial resolution and shorter temporal resolution. Both Landsat 8 and Sentinel-2 show more agreement with the fractional field FVC, with Landsat 8 especially showing higher RMSE and MAE values for the model-estimated vs. discrete field FVC values. Overall, the grassy class has the highest RMSE and MAE values for both Landsat 8 and Sentinel-2.

4.1.3 Model assessment

The LSMM model results were assessed using two different techniques: comparison to existing FVC product and validation against in-situ data.

4.1.3.1 Comparison to existing fractional cover product

The model-estimated FVC was compared to a benchmark FVC product (Section 2.3.3). The CDLC fraction layers product was selected as benchmark FVC (Buchhorn et al. 2019) as its endmember class definitions corresponded well with the model-estimated FVC. As the MODIS Vegetation Continuous Fields (VCF) product combines shrubs and grass within the same fractional cover layer (Dimiceli et al. 2015), the product is not suitable for comparison in this study. The product provides 10 cover fraction layers at 100 m spatial resolution, including a bare/sparse vegetation layer, herbaceous layer, shrubland layer and trees layer (Section 2.3.3). As the CDLC product was estimated for the year 2015, Landsat 8 imagery was used for unmixing (refer to Appendix D for the image acquisition dates). The relevant pre-processing steps were performed (Section 4.1.1.1) and the refined endmember points (Section 4.1.2.3) were used to extract mean spectral signatures for 2015. Using Landsat 8's optimal band combination (defined in Section 4.1.2.1) an FVC (30 m spatial resolution) for 2015 was estimated using the LSMM. The model-

estimated FVC was resampled to the CDLC spatial resolution (100 m) for qualitative and quantitative (RMSE, MAE) comparison.

4.1.3.2 Validation against in-situ field data

The model-estimated FVC was validated using independent in-situ field data collected in June 2018 for the Fort Beaufort study area during the dry season field trip (Section 3.1.1.1). FVCs were estimated using both Landsat 8 and Sentinel-2 imagery for June 2018 (refer to Appendix D for image acquisition dates) and agreement was assessed quantitatively using RMSE and MAE.

4.1.4 Model transferability

The temporal and spatial transferability of the LSMM was assessed by applying the LSMM in the Cedarville study area (Section 3.1.2) during the growing season (February 2019) using Sentinel-2 imagery (refer to Appendix D for image acquisition dates). The model was implemented using two different sets of endmembers: 1) pre-developed mean spectral signatures (Section 4.1.2.3); and 2) mean spectral signatures extracted from a number of endmember points selected by manual inspection of the NLC2017/18 product and aerial imagery. Five points were selected for each of the endmember subclasses: Bare 1, Bare 2, Grassy, Shrubs, Trees. The LSMM was then implemented using the optimal band combination identified for Sentinel-2 (Section 4.1.2.1) with the relevant mean spectral signatures as input. The results were validated using the in-situ field FVC data collected during the growing season field trip (Section 3.1.2.1).

4.1.5 Fractional vegetation cover change

Change in FVC was calculated to provide a quantitative estimate of grassland degradation and WPE in the study area. The 10 field FVC plots in the Fort Beaufort study area, used during calibration of the LSMM (Section 4.1.2.5), was classified using 2015 and 2018 aerial imagery. The difference between field-estimated FVC values for the two years was used to assess change in cover fractions.

4.2 RESULTS

This section summarises the results of the model calibration, implementation and validation as well as the evaluation of model transferability and the quantification of FVC change in the Fort Beaufort study area.

4.2.1 Endmember spectral signatures

In Figure 4.7, the spectral response from the selected bands is shown as reflectance (%) on the Y-axis. The error bars in the spectral signature plots represent one standard deviation and can be used

to assess the variation within each band, as well as overlap between bands. Figure 4.7 compares the mean spectral signatures for the endmember classes Bare, Grassy and Woody for (a) Landsat 8 and (b) Sentinel-2, while Figure 4.8 shows the mean values for each index.

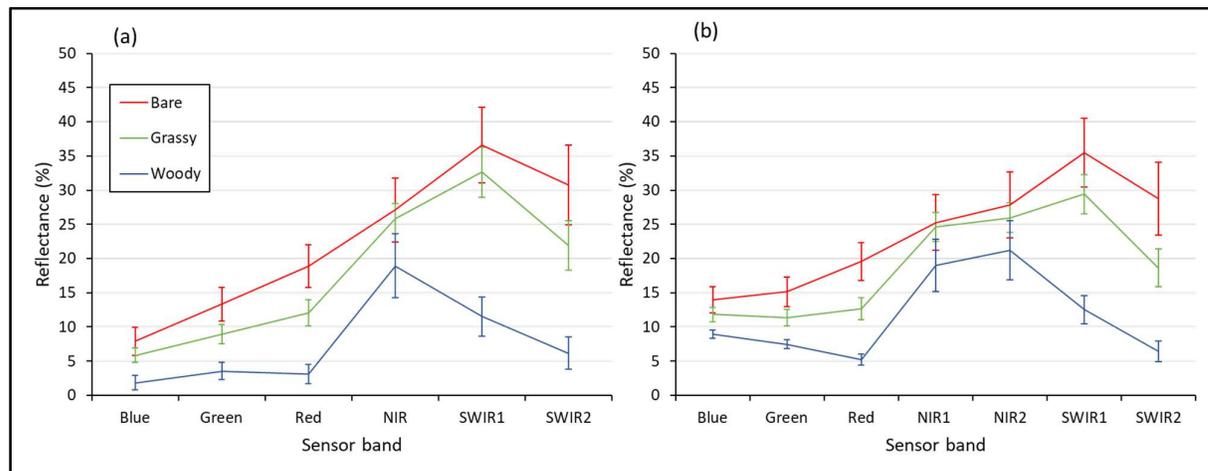


Figure 4.7 Mean spectral signatures for endmember classes Bare, Grassy and Woody for (a) Landsat 8 and (b) Sentinel-2

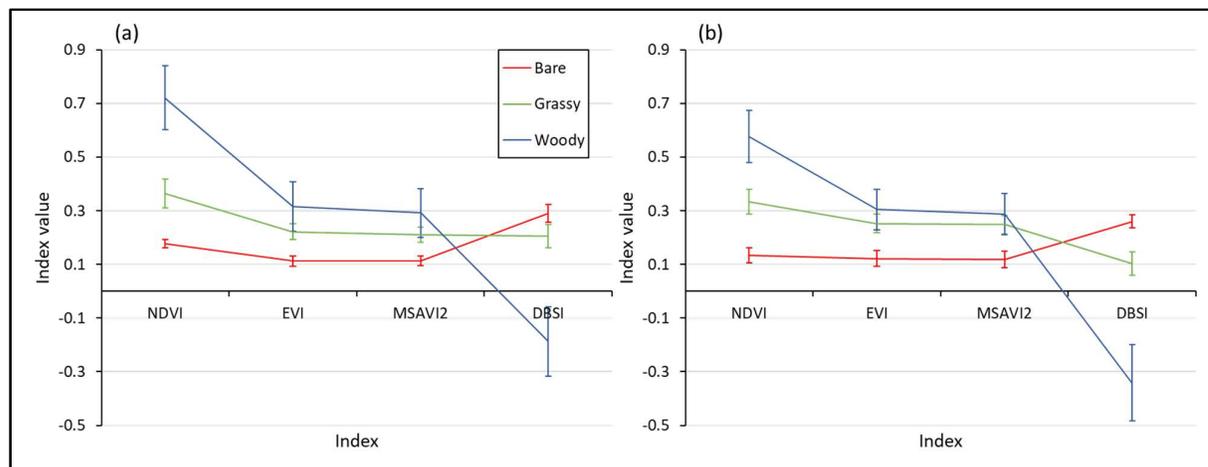


Figure 4.8 Mean index values for endmember classes Bare, Grassy and Woody for (a) Landsat 8 and (b) Sentinel-2

Variation of recorded reflectance values within the Bare (red) and Woody (blue) class was quite high, especially for the SWIR bands (Figure 4.7). Variation within the Bare class improved with the use of vegetation and soil indices as seen by the smaller error bars in Figure 4.8. However, the Woody class still showed notable variation in recorded index values, especially for the DBSI.

Figure 4.9 compares the mean spectral signatures for the endmember subclasses Bare 1, Bare 2, Grassy and Shrubs and Trees for Landsat 8 (Figure 4.9a) and Sentinel-2 (Figure 4.9b), followed by the mean values for each index in Figure 4.10.

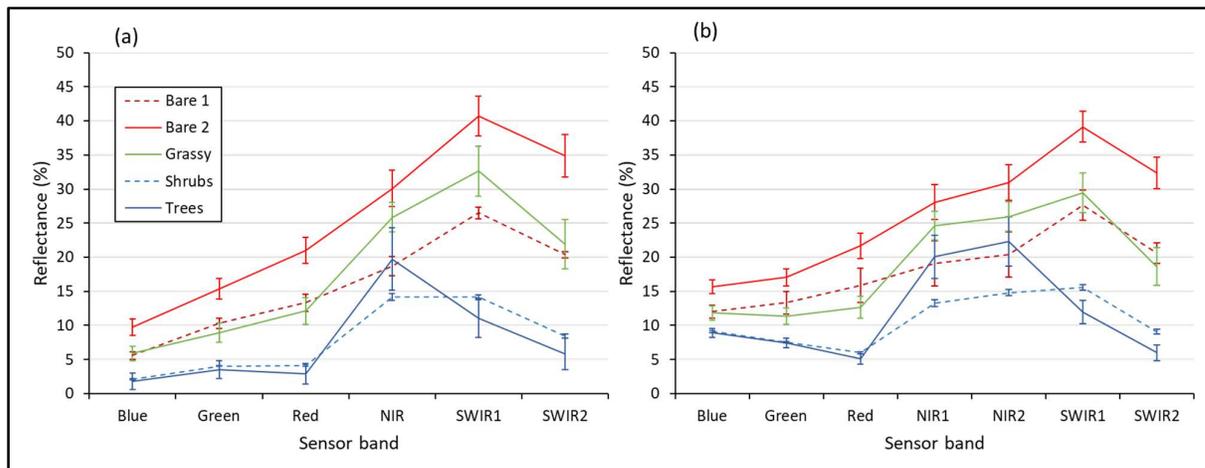


Figure 4.9 Mean spectral signatures for endmember subclasses Bare 1, Bare 2, Grassy, Shrubs and Trees (a) Landsat 8 and (b) Sentinel-2

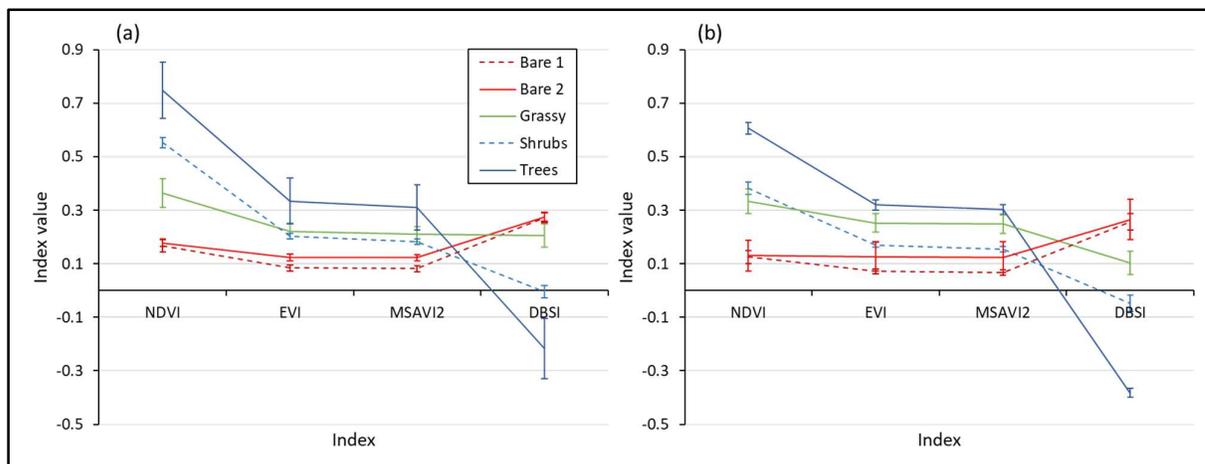


Figure 4.10 Mean index values for endmember classes Bare 1, Bare 2, Grassy, Shrubs and Trees for (a) Landsat 8 and (b) Sentinel-2

Splitting the endmember classes into subclasses decreased variation in recorded reflectance values, as shown by the significantly smaller error bars in Figure 4.9 and Figure 4.10. The only spectral profile that did not show notable improvement in variation after the subdivision of endmember classes was the Trees (blue) reflectance values extracted from Landsat 8 (Figure 4.10a).

There is substantial overlap between the Bare and Grassy class in the NIR spectral range for both Landsat 8 (Figure 4.7a; Figure 4.9a) and Sentinel-2 (Figure 4.7b; Figure 4.9b). This band was thus not included as a feature during spectral unmixing (Section 4.1.2.1). Furthermore, the NIR2 (band 8A) values for Sentinel-2 are slightly higher than that of the NIR1 (band 8) (Figure 4.7b). As there is overlap between the Bare and Grassy class in the range, the slightly higher reflectance recorded for NIR2 improves separability between these two classes and supports its use as NIR band in the Sentinel-2 NDVI calculation. Substantial overlap was also noted in the SWIR range for Bare 1 (eroded ground and soil) and Grassy for Sentinel-2, specifically SWIR1 (band 11). This band was thus not included as a feature during spectral unmixing (Section 4.1.2.1). The visual spectral bands

(Blue, Red, Green) were also not included for Sentinel-2 as their inclusion did not improve FVC estimation accuracy. Mean values recorded for EVI and MSAVI2 were very similar (Figure 4.8; Figure 4.10). Both indices were still included, however, as the region was relatively dry and EVI might still improve separability in areas with higher biomass production e.g. Cedarville.

To investigate the relationship between season and spectral response, mean spectral signatures were derived for Sentinel-2 for the dry and growing season respectively. Figure 4.11 and Figure 4.12 compare the Sentinel-2 recorded reflectance values and mean index values for the different endmember subclasses for the (a) growing season and (b) dry season.

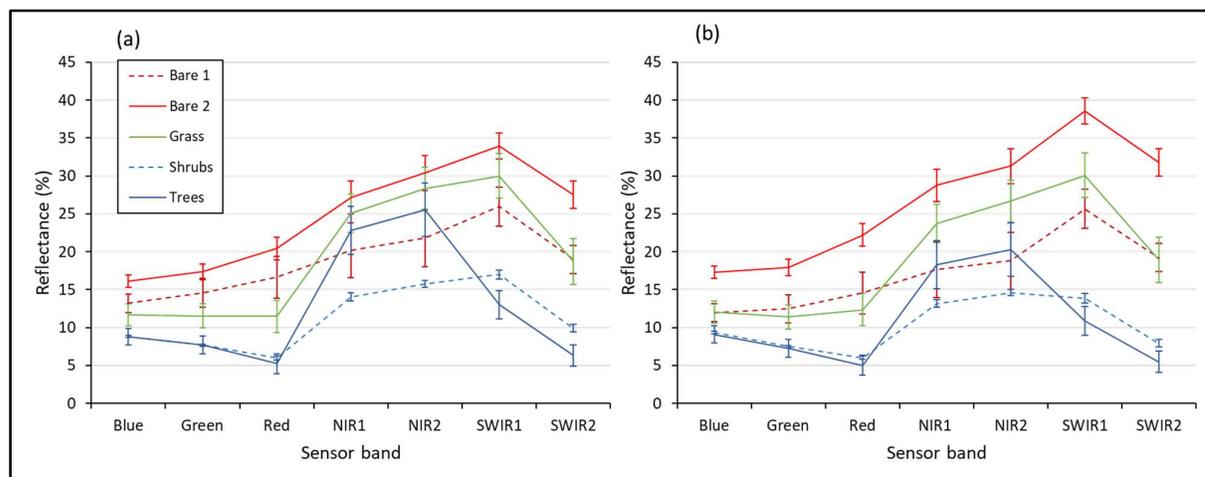


Figure 4.11 Mean spectral signatures for Sentinel-2 for the (a) growing season and (b) dry season

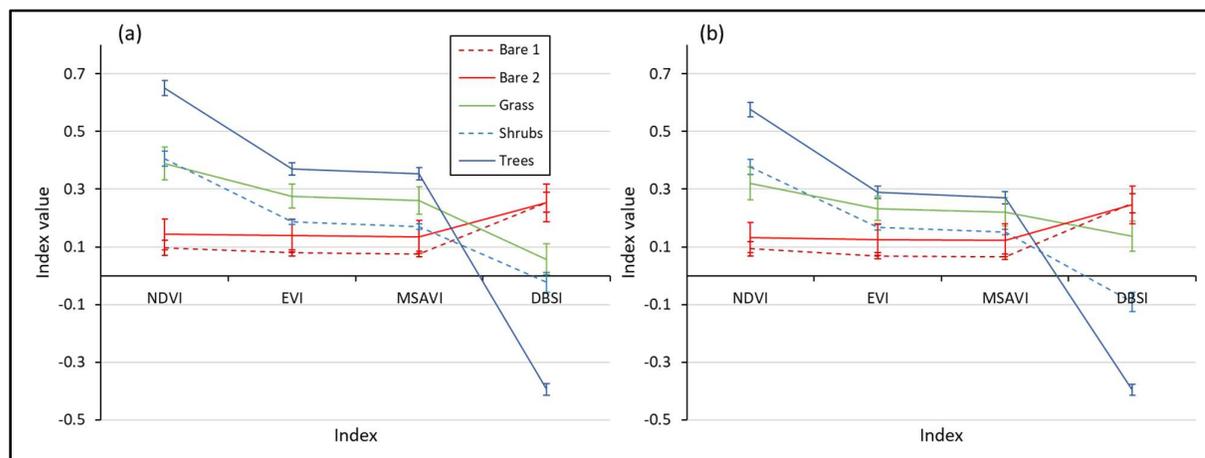


Figure 4.12 Mean index values for Sentinel-2 for the (a) growing season and (b) dry season

The NIR and SWIR bands show the biggest difference in spectral response between the seasons, with NIR reflectance for Trees decreasing by 5% and SWIR reflectance for Bare 2 increasing by 5% (Figure 4.11b). The mean index values for the two seasons are relatively similar, however, DBSI values for Grassy are higher during the dry season (Figure 4.12b) with more overlap between the Bare and Grassy endmember classes. This can potentially complicate discriminating between dry grass and bare ground in the dry season.

4.2.2 Fractional vegetation cover output maps

Figure 4.13 and Figure 4.14 show the final FVC map for Fort Beaufort (2018) and Cedarville (2019) respectively. The heterogeneous LC of the Fort Beaufort study area is reflected in the FVC (Figure 4.13), with only a small area of grassland and rangeland shown by the green pixels. In contrast, the Cedarville study area is majority grassland, with patches of scattered woods, bare soil and crops (masked out in grey) (Figure 4.14).

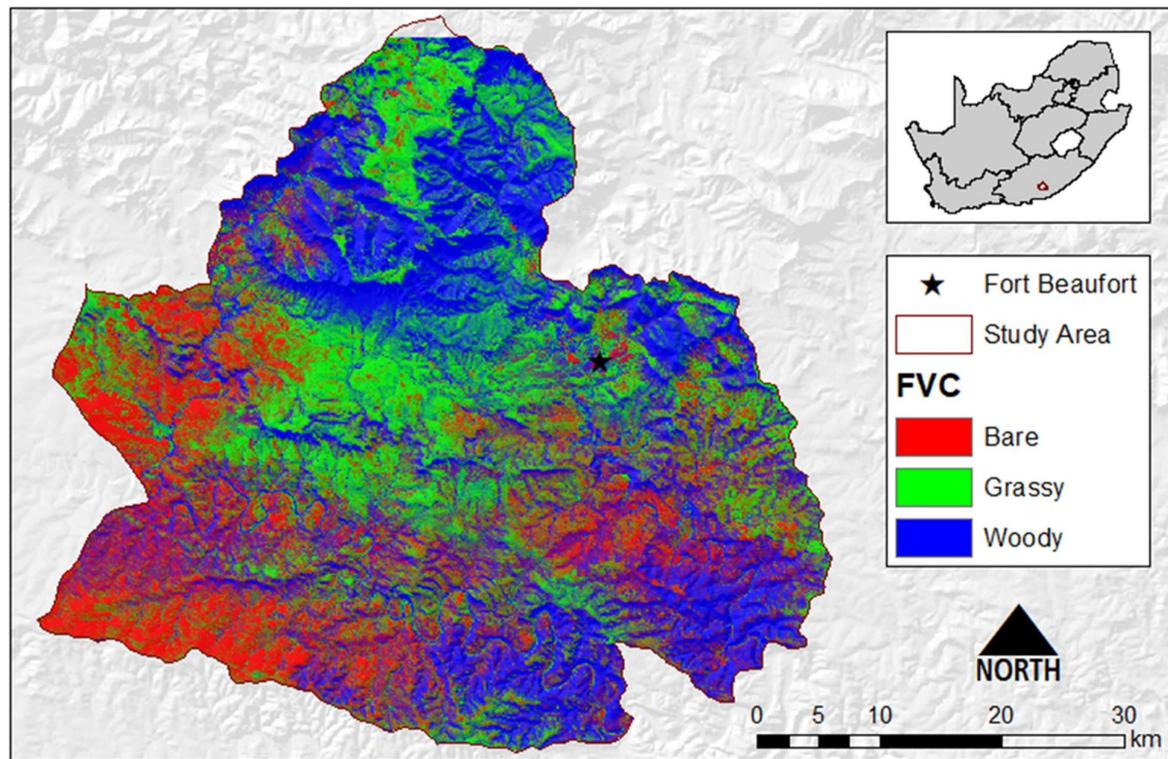


Figure 4.13 Model-estimated FVC for the Fort Beaufort study area for 2018

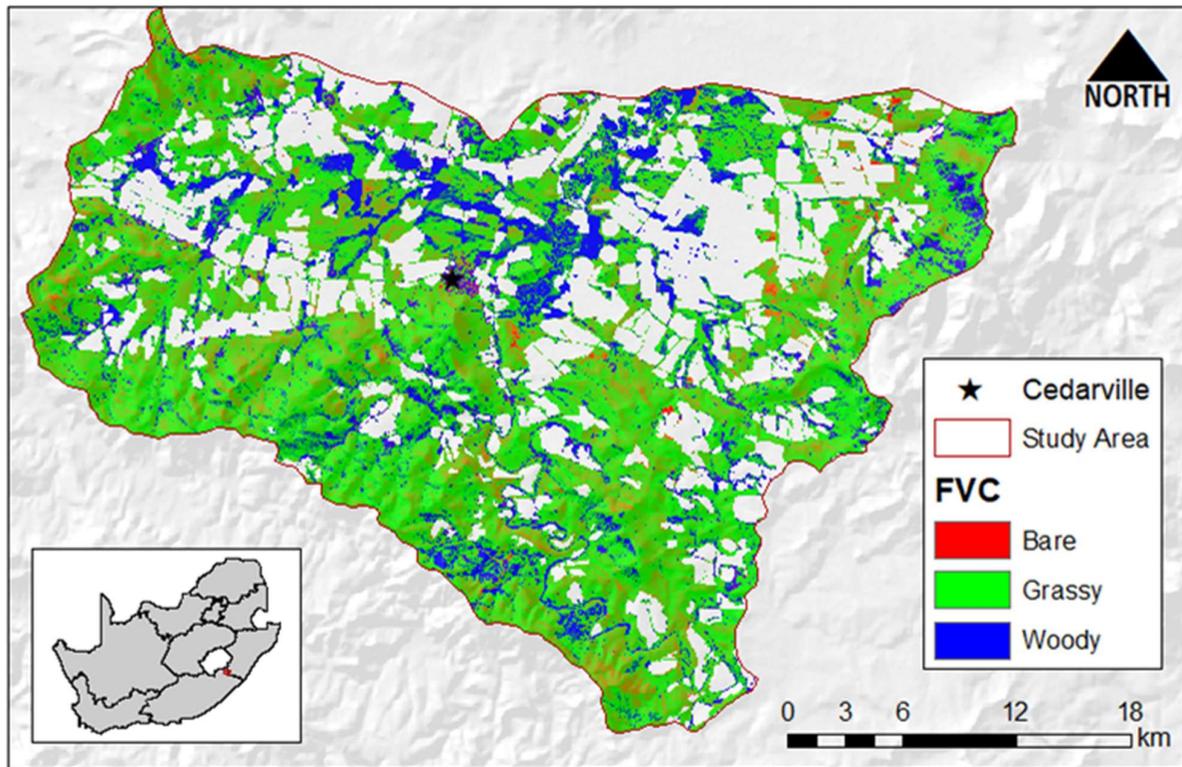


Figure 4.14 Model-estimated FVC for the Cedarville study area for 2019

Figure 4.15 compares the (a) model-estimated FVC to (b) aerial imagery using a degraded grassland field in the Fort Beaufort region. A good visual agreement is observed and the FVC reflects the grassland erosion (red) and WPE (blue) within the field well.

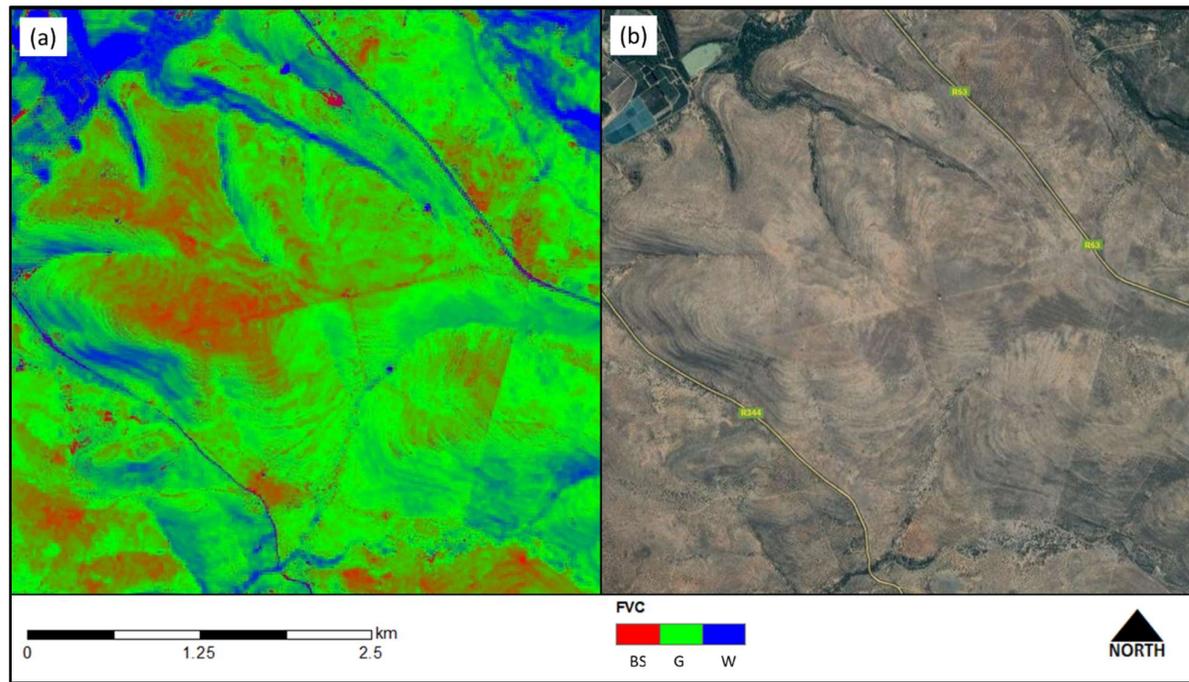


Figure 4.15 Visual comparison of (a) model-estimated FVC to (b) aerial imagery

4.2.3 Model assessment

4.2.3.1 Comparison to existing fractional cover product

Table 4.6 summarises the RMSE and MAE values for the comparison of the model-estimated FVC and CDLC fraction layers (Section 4.1.3.1). The Grassy class estimation performed the worst, with an RMSE of 10.92% and an MAE of 13.82%.

Table 4.6 Comparison of model-estimated FVC and existing FVC product using RMSE and MAE

FVC	RMSE (%)	MAE (%)
Bare	5.04	14.05
Grassy	10.92	13.82
Woody	7.96	11.99

A visual inspection of agreement was also performed using the 10 FVC plot sites identified in Section 4.1.2.5. Figure 4.16 shows an example of this comparison using the (a) aerial imagery, (b) model-estimated FVC (30 m), (c) resampled model-estimated FVC (100 m) and (d) CDLC FVC (100 m) using one of the 10 plots (plot nr. 9).

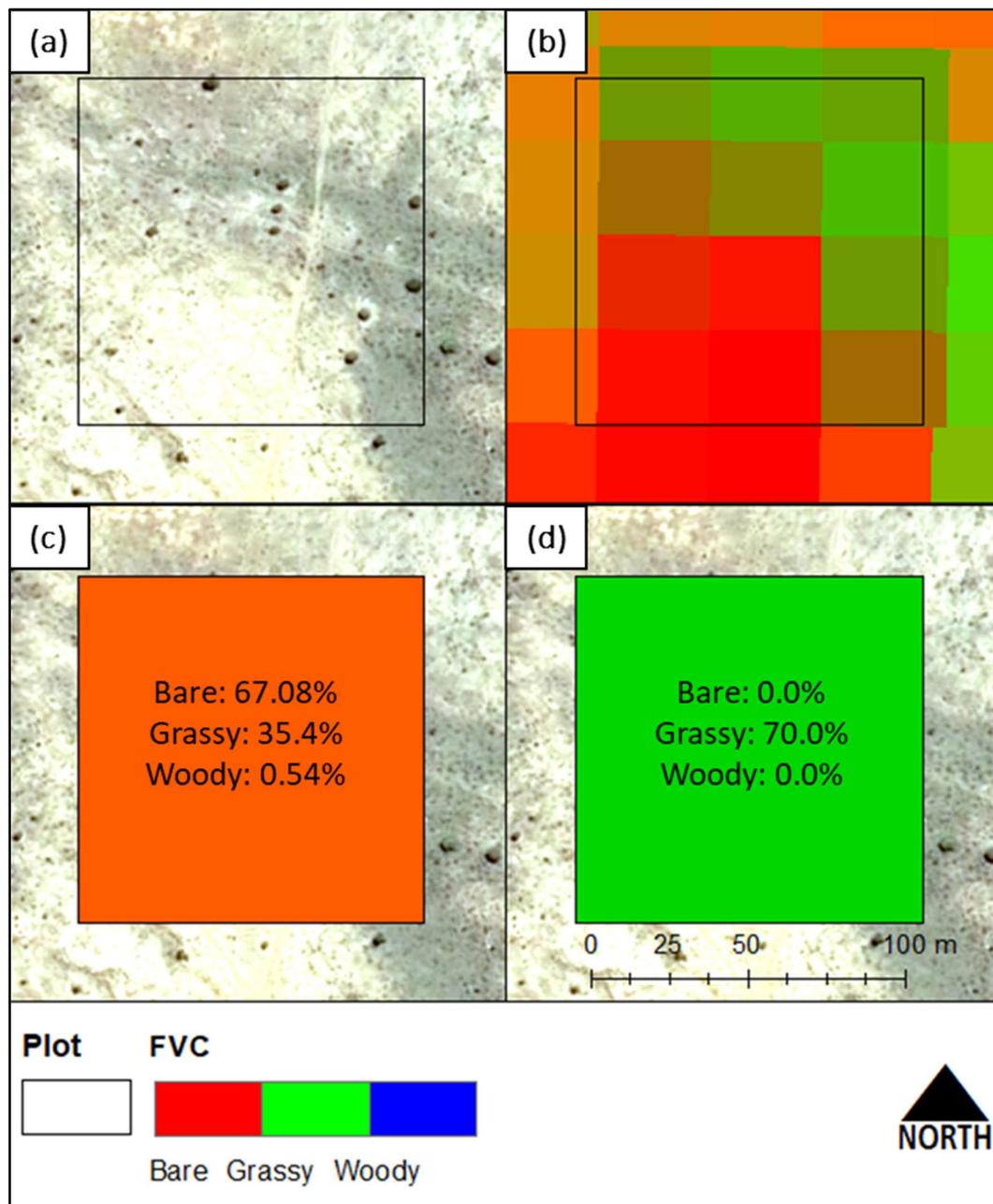


Figure 4.16 Visual comparison of (a) aerial imagery, (b) model-estimated FVC at 30 m resolution, (c) resampled model-estimated FVC at 100 m resolution and (d) CDLC FVC at 100 m resolution

A big difference is observed between the resampled model-estimated FVC (Figure 4.16c) and CDLC FVC (Figure 4.16d). The CDLC FVC estimates 0% for Bare and Woody, 70% for Grassy and the remaining 30% as Crops (not included in the map). Based on visual inspection the model-estimated FVC shows a better agreement with the aerial imagery, with a good representation of the bare ground (67.08%) and woody components (0.54%) in the plot. Other plots also showed better agreement between the model-estimated FVC and aerial imagery, while CDLC generally overestimated grass cover and underestimated woody components and bare soil.

4.2.3.2 In-situ field data

The model-estimated FVC was validated using in-situ FVC field data collected during the dry season field trip (Section 3.1.1.1). Figure 4.17 shows the (a) RMSE and (b) MAE values obtained during the model-estimated FVC and in-situ FVC field data comparison.

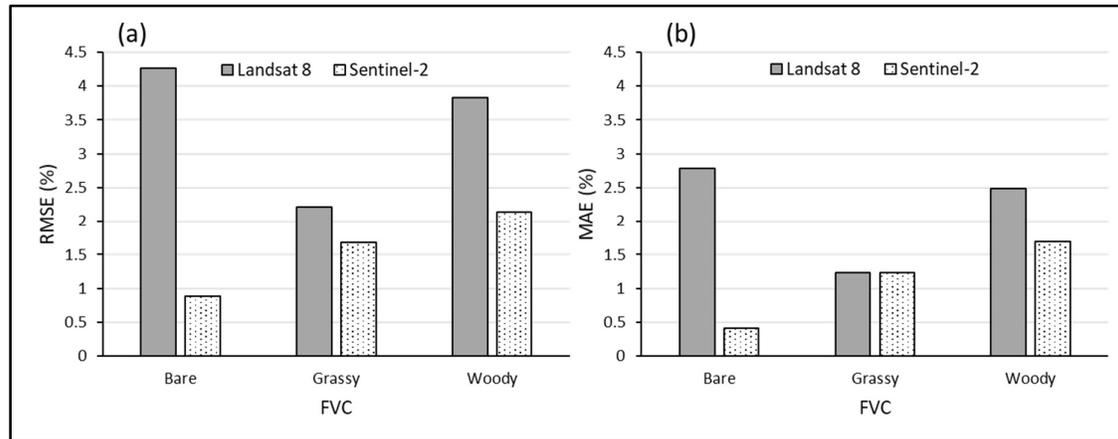


Figure 4.17 Comparison of model-estimated FVC to in-situ field data using Landsat 8 (grey) and Sentinel-2 (white) in terms of (a) RMSE and (b) MAE

Landsat 8 performed poorer overall, with higher RMSE (Figure 4.17a) and MAE (Figure 4.17b) values. This is to be expected due to the poorer spatial resolution compared to Sentinel-2. The Sentinel-2 FVC performed well, with a maximum RMSE of 2.2% and MAE of 1.7%.

4.2.4 Model transferability

The transferability of the LSM was evaluated by implementing the model in a spatially and temporally different scenario i.e. the Cedarville study area using Sentinel-2 imagery for the 2019 growing season. Figure 4.18 shows the (a) RMSE and (b) MAE values obtained during comparison of the model-estimated FVC and in-situ FVC field data comparison.

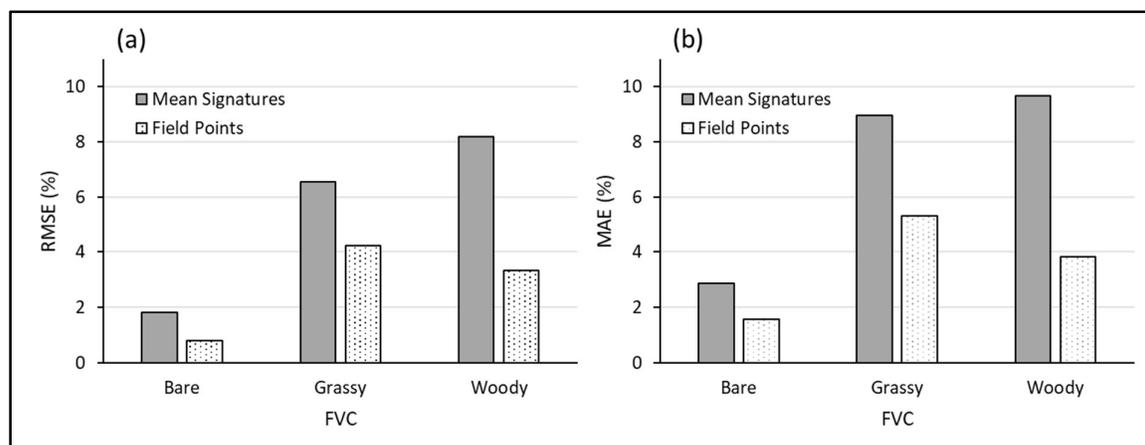


Figure 4.18 Comparison of model-estimated FVC to in-situ field data using pre-developed mean spectral signatures (grey) and mean spectral signatures developed from field points (white) in terms of (a) RMSE and (b) MAE

The LSMM performed better when endmember spectral signatures were derived from field points (white) compared to using fixed, pre-developed mean spectral signatures (grey) (Figure 4.18). This was expected as the pre-developed mean spectral signatures were derived for the much drier Fort Beaufort study area. Drier grass equates to lower NDVI, EVI and MSAVI values. When using dry grass mean spectral signature in a very green, productive region, much of the grass will have higher vegetation index values than the mean and would consequently be identified by the model as trees or shrubs. This explains the high RMSE and MAE values for Woody (grey) when using the pre-defined mean spectral signatures.

4.2.5 Fractional vegetation cover change

Table 4.7 summarises the change in FVC (%) for each field plot from 2015 to 2018. Green corresponds to an increase in cover, red to decrease in cover and grey to no change.

Table 4.7 Changes in field plot FVC from 2015 to 2018

Plot #	FVC (%)		
	Bare	Grassy	Woody
1	0	-8.65	8.65
2	0.3	-1.3	1
3	0	0	0
4	0	0	0
5	0	0	0
6	23.5	-23.45	-0.05
7	17	-18.05	1.05
8	0	0	0
9	9	-10	1
10	0	7.5	-7.5
Mean	4.98	-5.395	0.415

Overall there is an increase in Bare and Woody, with a decrease in Grassy. Increases in bare ground could be linked to the drought in the region, with increases in woody components showing the effects of WPE in the study area.

4.3 DISCUSSION

The LSMM-estimated FVC performed well based on visual inspection and validation against in-situ field data. Although high RMSE and MAE values were obtained when model-estimated FVC

was compared to CDLC fractions (Table 4.6), Figure 4.16 shows that the CDLC did not always correlate with cover fractions observed on aerial imagery. In addition to classification error (e.g. classifying grassland as crops in Figure 4.16), the main difference in final FVC estimates can be attributed to the different classification approaches followed. For comparison, the model-estimated FVC is directly resampled to a 100 m FVC, while the CDLC FVC estimation approach corresponds to a discrete field FVC classification (Section 4.1.2.5). This difference introduces a high level of uncertainty into the comparison and explains the disagreement between the model-estimated FVC and CDLC FVC.

The use of endmember subdivision notably improved FVC estimation using Sentinel-2 (Figure 4.9b; Figure 4.10b) and produced a final model-estimated FVC with RMSE and MAE values lower than 2.5% when compared to in-situ field data (Figure 4.17). A mean spectral signature derived for Bare using all Bare endmember points did not adequately represent the different Bare cover types, thus splitting the Bare class into Bare 1 (bare rocky ground) and Bare 2 (bare soil) improved separability. The subdivision of Woody into Shrubs and Trees also ensured improved estimation of woody components in grasslands. By introducing a Shrubs subclass, early stages of WPE can potentially be monitored. Subdividing the classes did not have a meaningful impact on the variation in reflectance and index values extracted from Landsat 8, particularly for the Woody subclasses (Figure 4.9a; Figure 4.10a). Although RMSE and MAE values for Landsat 8 were still relatively low, the finer spatial resolution of Sentinel-2 improved unmixing results in a highly heterogeneous area. Due to the coarser spatial resolution of Landsat 8, it was more difficult to find pure pixels for the respective endmember classes. There is a higher possibility of the presence of other LC components within a seemingly pure pixel, which consequently affects the range of recorded reflectance values for each endmember class. Spatial resolution thus plays a key role in spectral unmixing.

The Bare endmember class presented the lowest RMSE and MAE (Table 4.6; Figure 4.17). Validation of Grassy estimates produced RMSE values between 1.24% and 2.25%, with Woody performing the poorest (maximum RMSE 3.3%). Although the Bare class quantitatively performed the best, visual inspection showed slight overestimation in the bare soil extent (Figure 4.15). This is likely due to the dry conditions in the study area, resulting in difficulty in discriminating between dry grass and bare soil and ground. The effects of the drought in the region are also reflected in the mean spectral signatures, with substantial overlap in spectral response between Grassy and Bare 2 subclasses in the NIR and SWIR ranges (Figure 4.11) and similar DBSI values (Figure 4.12). The poor performance of the Woody class, specifically using Landsat 8, is likely due to the overestimation of shrubby vegetation in areas where the grassland is slightly

greener, as the Grassy and Shrubs subclasses have very similar NDVI values (Figure 4.11). Within the grasslands a large variation in spectral response was noted, as different grass species and levels of growth (grazed, dense, dry etc.) have different spectral characteristics. For a highly heterogeneous area such as Fort Beaufort, a potential improvement would be to define subclasses within the Grassy endmember class based on grass species and condition. The mean spectral signatures also reflect seasonal variability (Figure 4.11; Figure 4.12). The difference in spectral response between the dry season and growing season suggests that an annual FVC might not sufficiently capture the effect of the woody vegetation and bare areas on grazing capacity. Bi-annual FVC estimations would potentially provide better temporal information on grassland condition and how seasonal changes in FVC can impact grassland productivity.

The LSMM, calibrated for Fort Beaufort in 2018, could be transferred both spatially and temporally to the Cedarville region in 2019 with relative success (Figure 4.18). The FVC estimation performed better when endmember spectral signatures were derived from limited, manually selected field points compared to using fixed, pre-developed mean spectral signatures (Figure 4.18). This was expected as the pre-developed mean spectral signatures were derived for the much drier Fort Beaufort study area. Drier grass equates to lower NDVI, EVI and MSAVI values. When using such a dry grass mean spectral signature in a very green, productive region, most of the grass will have higher vegetation index values than the mean. The model would consequently identify the grass as trees or shrubs. This explains the particularly high RMSE and MAE values for Woody when using the pre-defined mean spectral signatures (Figure 4.18). Mean spectral signatures are thus not transferable between highly different regions, such as the heterogeneous, dry Fort Beaufort and the homogenous, wet, green Cedarville. The LSMM is, however, transferable when a limited number of points for each endmember class is available. Although the RMSE and MAE values are higher than achieved for the final FVC estimation for Fort Beaufort (Figure 4.18), RMSE values ranging from 0.81% to 4.22% were still acceptable for use in subsequent analysis.

The need for continuous FVC data is supported by the FVC change from 2015 to 2018 quantified in Table 4.7. The Fort Beaufort region has experienced a substantial increase in Bare and decrease in Grassy, with slight increases in Woody also recorded. The majority of available FVC products, specifically with endmember definitions suitable for grazing productivity estimation, have been created on a once-off basis e.g. the CDLC cover fractions for 2015. Other more recent FVC products, e.g. the MODIS VCF product, have very coarse spatial resolutions (250 m). As shown in Figure 4.16, products with such coarse spatial resolutions do not provide sufficient detail in the context of the real-world problem. The FVC estimation approach discussed in this chapter thus

has the potential to provide updated, continuous FVCs using freely available, medium resolution (spatial and temporal) imagery i.e. Sentinel-2. The transferability of the LSMM using a limited number of pure endmember points supports its use for this research, as other methods of FVC estimation, e.g. random forest regression, require a large number of field points to produce accurate results (Section 2.3.2.3). In an application where FVC must be estimated continuously and dynamically, this approach has the potential of providing accurate cover fractions at high spatial and temporal resolution. The results of this chapter will be used in Chapter 5 to determine the effects of woody and bare components at sub-pixel level on grassland productivity.

CHAPTER 5: GRASSLAND PRODUCTIVITY MODELLING

Grassland productivity was modelled using fractional vegetation cover (FVC) (Chapter 4), a remote sensing (RS)-based primary production model and formulas identified in literature for grazing capacity (GC) calculation. This chapter discusses the implementation of the Regional Biosphere Model (RBM) (Richters 2005b) to calculate gross primary production (GPP) and subsequently GC, thereby addressing Objective 4. Firstly, the input data and data preparation is described, followed by the GPP model implementation, GC estimation and validation, concluding with results and discussion. This chapter also investigates the implementation of the model in a Google Earth Engine (GEE) web application (app), with the aim of providing continuous, dynamic, multi-scale grassland condition and productivity estimates, thereby addressing Objective 5.

5.1 METHODS & MATERIALS

5.1.1 Data selection & pre-processing

Sentinel-2 Level 2A (Surface Reflectance) imagery was used as input to the primary production model. All image collections used during processing were filtered to only include images with less than 5% cloud cover. All areas that were not rangelands or grasslands, i.e. urban, water, crops and mines, were masked out using the National Land Cover 2017/2018 (NLC2017/8) product created for the Chief Directorate: National Geospatial Information (CD: NGI) by Land Resources International (Land Resources International 2018). No additional pre-processing of the imagery was required (e.g. orthorectification, radiometric correction, atmospheric correction etc.), as the Google Earth Engine (GEE) data catalogue provides both raw and processed imagery for various satellite sensors. Maximum-value normalised difference vegetation index (NDVI) composites (both 8-day and 10-day) were created as input for GPP modelling. Maximum-value compositing was chosen to minimise cloud contamination (Holben 1986), while composite time periods of 8-days and 10-days are short enough to capture changes in vegetation and climatic conditions (Richters 2005b). In addition to sensitivity to vegetation changes, these composite time periods are also comparable to the temporal resolutions of existing GPP products, namely the MODIS GPP product (MOD17A2) and the Copernicus Gross Dry Matter Production (CGDMP) product (Running, Mu & Zhao 2015; Swinnen & Van Hoolst 2018). GPP estimation was done for multiple analysis years (2016, 2017, 2018), to compare subsequent GC results to existing products (Long Term Grazing Capacity Map for 2016) and to quantify changes in grazing conditions from 2016 to present.

5.1.2 Primary production

Various primary production models exist in literature (Section 2.2.1) and the choice of model depends largely on the application area and available additional datasets. The RBM (Section 2.2.2.1), a light use efficiency (LUE) model developed specifically to describe the production and consumption of biomass by cattle and game in semi-arid areas (Richters 2005a), was chosen for this research. What distinguishes the RBM from other LUE-based production models (Coops et al. 2010; Goetz et al. 2000; Jia et al. 2015; King, Turner & Ritts 2011; Veroustraete, Sabbe & Eerens 2002; Verstraeten, Veroustraete & Feyen 2006; Xiao et al. 2005; Yuan et al. 2007; Zhao et al. 2005) is the combination of model transferability with localised results. Transferable models commonly produce generalised, global-scale results at coarse resolution (Running et al. 2004). In contrast, regional-scale, high resolution models mainly rely on non-transferable, locally-calibrated empirical equations derived from field data and/or flux tower measurements (Xiao et al. 2005; Yuan et al. 2007). The chosen RBM combines the benefits of both by including easily accessible input datasets that can be used to calculate regional-scale GPP results. The model requires relatively few input datasets and all are available in the public domain. This transferable model structure makes it ideal for implementation in the GEE environment.

Figure 5.1 reviews the input and intermediary products used by the RBM to calculate GPP, as discussed in Section 2.2.2.1. GPP ($\text{g C m}^{-2}\text{day}^{-1}$) is calculated as a function of APAR ($\text{MJ m}^{-2}\text{day}^{-1}$) and LUE (g C MJ^{-1}) (Equation 2.1), where LUE consists of two subcomponents: maximum LUE (LUE_{max}) and effective LUE (LUE_{eff}). As seen in Figure 5.1, LUE_{max} is calculated as an annual constant while LUE_{eff} is calculated for each time step. LUE_{max} incorporates the influence of soil texture and carbon/nitrogen (C/N) ratio; relief in the form of height and slope; and FVC to determine potential LUE (LUE_{p}) based on vegetation type and structure. LUE_{eff} incorporates the influence of water stress (W_{ϵ}) using evapotranspiration (ET); and temperature stress (T_{ϵ}) using land surface temperature (LST). APAR is calculated using the normalised difference vegetation index (NDVI), photosynthetically active radiation (PAR) and the fraction of absorbed photosynthetically active radiation (fAPAR).

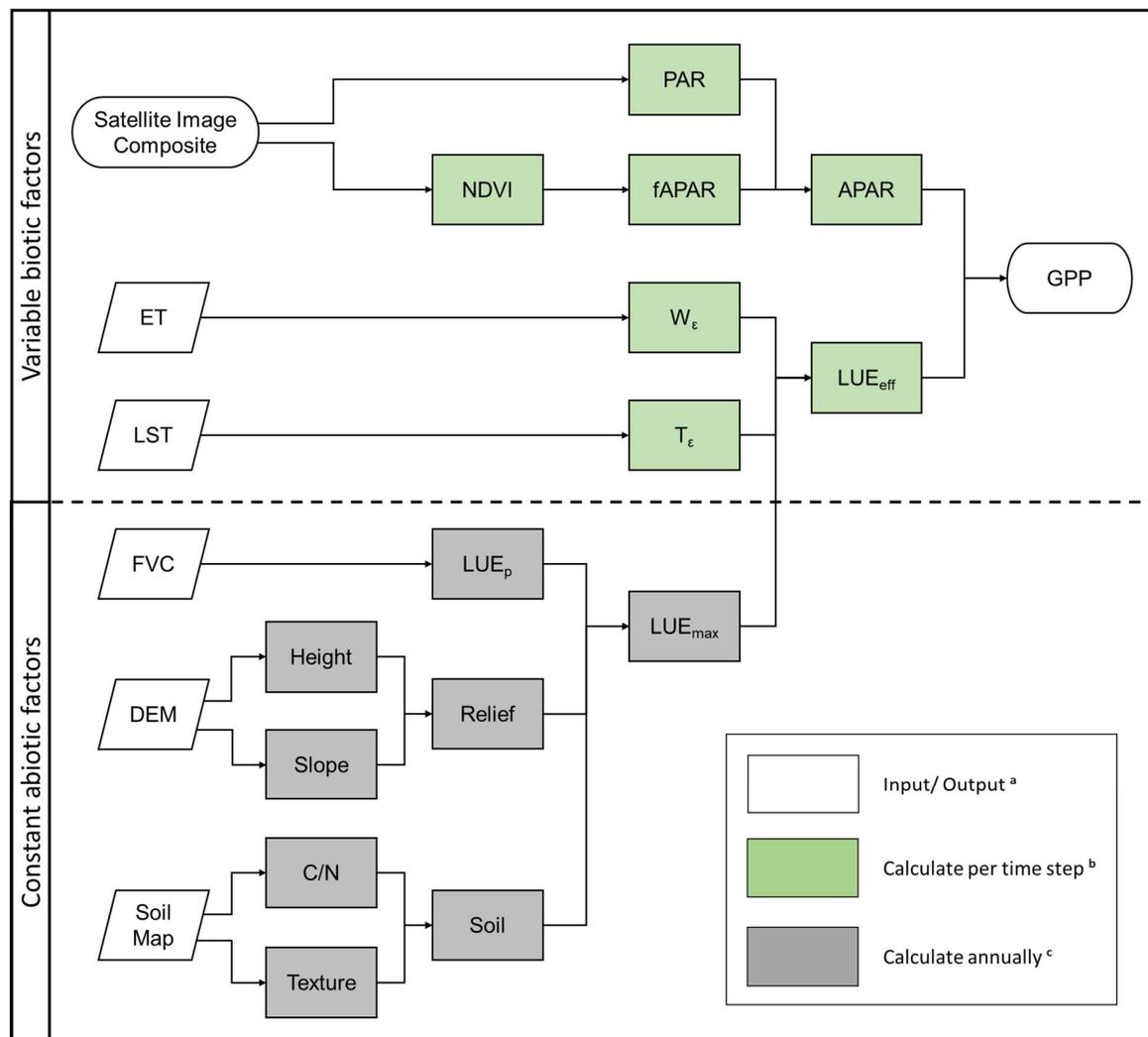


Figure 5.1 RBM input and intermediary parameters.

^a ET, evapotranspiration; LST, land surface temperature; FVC, fractional vegetation cover; DEM, digital elevation model; GPP, gross primary production ^b NDVI, normalised difference vegetation index; PAR, photosynthetically active radiation; fAPAR, fraction of absorbed photosynthetically active radiation; APAR, absorbed photosynthetically active radiation; W_{ϵ} , water stress factor; T_{ϵ} , temperature stress factor; LUE_{eff} , effective light use efficiency; ^c LUE_p , potential light use efficiency, LUE_{max} , maximum light use efficiency; C/N, carbon/nitrogen ratio.

The majority of the input parameters were implemented as in the original RBM (Richters 2005b), while FVC and fAPAR were modified to improve the spatial resolution of GPP estimates. These adaptations were implemented to downscale the RBM from regional-scale to local-scale to potentially improve the accuracy of GPP results and subsequent GC estimates. As a number of input parameters of various different spatial resolutions were used, the adapted RBM implemented followed a multi-resolution image analysis approach (Chen & Stow 2003). Although this allowed improving the detail of certain input variables and consequent model results, the overall efficacy of the final model is still limited by the coarsest input sources. The calculation of APAR, LUE_{max}

and LUE_{eff} and their contributing factors are discussed in Sections 5.1.2.1, 5.1.2.2 and 5.1.2.3 respectively.

5.1.2.1 APAR

APAR ($MJ\ m^{-2}day^{-1}$) was calculated as a function of fAPAR (unitless) and PAR ($MJ\ m^{-2}day^{-1}$) (Equation 2.7) at each time step. The original RBM model (Richters 2005b) used the MODIS Leaf Area Index (LAI)/FPAR product (MOD15) (1 km spatial resolution) as input dataset for fAPAR. This product has since been superseded by the MOD15A2H product (Myeni, Knyazikhin & Park 2015) to provide LAI/fAPAR information at 500 m spatial resolution every 8 days. Although the spatial resolution has improved, the heterogeneity of the Fort Beaufort grasslands required more detailed input data to ensure accurate GPP results. In addition, a multi-temporal evaluation using in-situ measurements has shown the product tends to overestimate fAPAR (Fensholt, Sandholt & Rasmussen 2004). An alternative approach for calculating fAPAR from NDVI was thus implemented. The field data analysis conducted for this research (Chapter 3) found a strong relationship between NDVI and fAPAR, specifically the fAPAR scaled using the green, living proportion of the collected grass field samples (Table 3.4). As this relationship essentially allows the fraction of PAR absorbed only by photosynthesising components to be derived, the use of green scaled fAPAR instead of the traditional fAPAR could potentially provide better primary production results. The Sentinel-2 maximum-value NDVI composites (Section 5.1.1) were used to estimate fAPAR using the NDVI-to-scaled fAPAR relationships developed during regression analysis of collected field data (Section 3.2.2.2). Regressions were performed for multiple dates, thus the NDVI-to-scaled fAPAR relationship with the highest R^2 was chosen for implementation in the model. The relationship is described by Equation 5.1:

Equation 5.1

$$fAPAR = 0.0428e^{4.3744NDVI}$$

where fAPAR is the fraction of absorbed photosynthetically active radiation [scalar]; and

NDVI is the normalised difference vegetation index [scalar];

PAR was integrated according to the original RBM (Richters 2005b). Potential insolation (InSol) ($cal\ cm^{-2}day^{-1}$) was calculated for each raster cell as a function of latitude, longitude, day of the year and solar angles (Swift 1976). InSol was scaled using an appropriate ratio coefficient (0.48) from literature (Monteith 1972) to transform potential radiation into actual values. The PAR results were then converted from $cal\ cm^{-2}day^{-1}$ to $MJ\ m^{-2}day^{-1}$ using Equation 5.2 to ensure comparability of units between PAR and LUE during GPP calculation:

$$PAR_i = \frac{(PAR_j \cdot 4.186 \cdot 1000)}{1000000}$$

where PAR_i is the photosynthetically active radiation [$\text{MJ m}^{-2}\text{day}^{-1}$]; and
 PAR_j is the photosynthetically active radiation [$\text{cal cm}^{-2}\text{day}^{-1}$].

5.1.2.2 Maximum light use efficiency

The LUE_{\max} (g C MJ^{-1}) is influenced by less variable, abiotic factors and was calculated as an annual constant. The relief factor (x_{relief}), soil factor (x_{soil}) and fractional vegetation cover (FVC) were used (Equation 2.8) to determine the effects of terrain, soil type and vegetation structure on the LUE.

The x_{relief} combines the effects of slope and elevation on LUE using a “fuzzy OR” (refer to Table 2.2). High slope or elevation can limit plant growth, thus increase in slope or elevation results in a reduction of LUE (Equation 2.11; Equation 2.12). By using fuzzy logic, favourable elevation could compensate for unfavourable slope, resulting in a unitless scaling value between 0 and 1. Slope and elevation were derived from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) (30 m spatial resolution) (Farr et al. 2007) freely available in the GEE archive, as implemented in the original RBM. Figure 5.2 shows that the x_{relief} for the Cedarville study area would have a much larger impact on the GPP estimates (Figure 5.2b) than in the Fort Beaufort study area (Figure 5.2a) due to steeper slopes and a higher overall elevation.

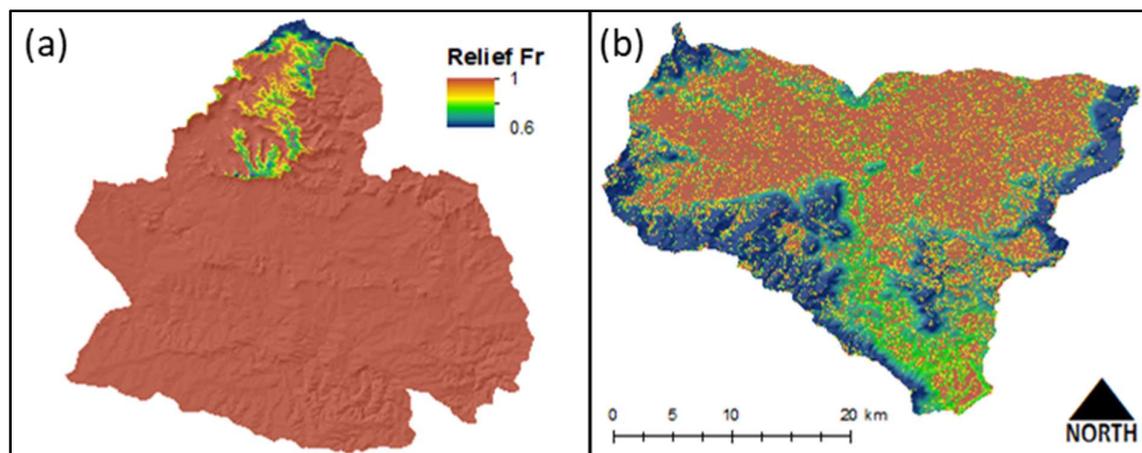


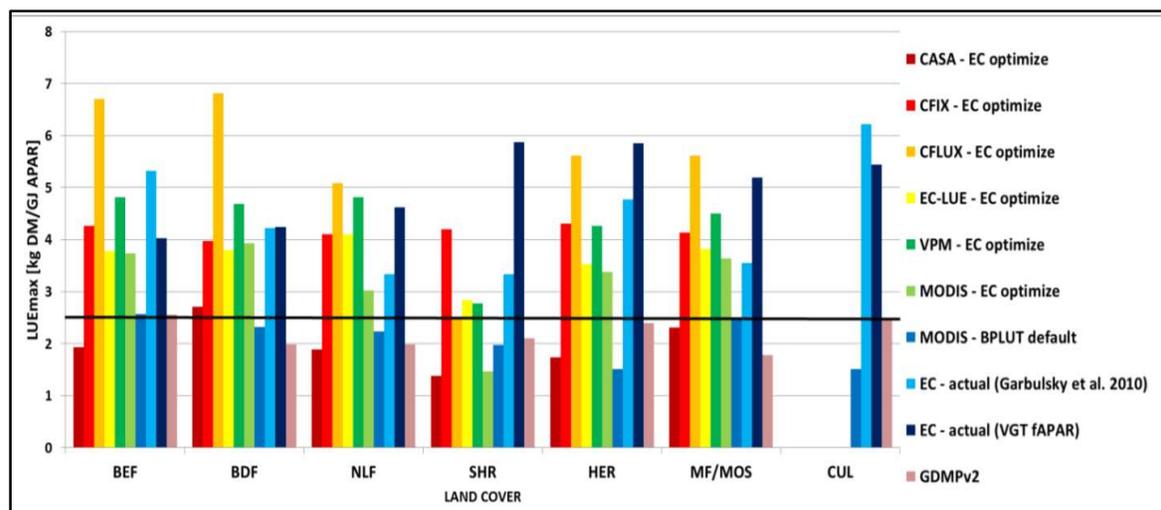
Figure 5.2 Comparison of relief factor (x_{relief}) for (a) Fort Beaufort and (b) Cedarville for 2018

The x_{soil} combines the effects of carbon/nitrogen (C/N) and texture on LUE using a “fuzzy OR” (refer to Table 2.2). C/N ratio and texture relate to nutrient supply and water storage capacity and was combined to create a unitless value between 0 and 1. The original RBM made use of the Food and Agricultural Organisation (FAO) Digital Soil Map, which was superseded by the Harmonised World Soil Database (HWSD) (1 km spatial resolution) (FAO/IIASA/ISRIC/ISS-CAS/JRC

2012), not yet available in the GEE archive. The HWSO was processed using ArcGIS 10.6.1 (ESRI 2018) to create soil texture and C/N ratio raster datasets of respective soil types. The two layers were ingested as assets into the GEE environment and used during modelling to determine the effects of soil type on LUE (Equation 2.13). As both study areas had relatively homogenous soil types with favourable C/N ratios and textures, the x_{soil} did not have a significant impact on GPP estimates for this research.

FVC, the final component of the LUE_{max} , plays a key role in determining the LUE_p and subsequently the LUE_{max} . The original RBM used the MODIS Vegetation Continuous Fields (VCF) product with a spatial resolution of 1 km, since replaced by a newer version (MOD44B collection 6) with a spatial resolution of 250 m (Townshend et al. 2017). Although the spatial resolution has improved, it is still not detailed enough to accurately model ecosystem dynamics of highly heterogeneous environments such as savannas and woody encroached grasslands. Another drawback is the inclusion of shrubs in the herbaceous cover. The linear spectral mixture model (LSMM) described and discussed in Chapter 4 was consequently implemented to estimate an FVC using Sentinel-2 imagery. As the RBM includes FVC as an annual constant, FVC was estimated on a yearly basis for the respective study areas and analysis years.

Using Equation 2.10, LUE_p was calculated using FVC by scaling the theoretical LUE (ϵ_x) of each land cover (LC) type with the fraction of that LC within the pixel (FVC_x). Various different ϵ_x values have been developed and implemented in research. Figure 5.3 shows the variation in ϵ_x used in different GPP models for a range of LC classes as summarised by Swinnen & Van Hoolst (2018). Note that Swinnen & Van Hoolst (2018) use ϵ_x and LUE_{max} interchangeably (as seen on the y-axis) and specify LUE units in kg DM MJ^{-1} not in g C MJ^{-1} .



Source: Swinnen & Van Hoolst (2018)

Figure 5.3 Comparison of ϵ_x values (kg DM MJ^{-1}) used by different GPP models in literature

As seen by the wide range of ϵ_x values, this parameter is evidently very difficult to standardise and a source of great uncertainty in vegetation modelling (Swinnen & Van Hoolst 2018). The use of different ϵ_x units in different models (kg DM MJ^{-1} vs. g C MJ^{-1}) is also a source of uncertainty and creates challenges when attempting to compare the performance of different GPP models. The original RBM implements the widely used ϵ_x values defined in the Biome Parameter Lookup Table (BPLUT) (Running, Mu & Zhao 2015), illustrated in Figure 5.3 as medium blue (MODIS – BPLUT default). The same ϵ_x values were used in this research, namely $\epsilon_{\text{bare}} = 0.010 \text{ g C MJ}^{-1}$, $\epsilon_{\text{grassy}} = 0.680 \text{ g C MJ}^{-1}$ and $\epsilon_{\text{woody}} = 1.116 \text{ g C MJ}^{-1}$.

5.1.2.3 Effective light use efficiency

The LUE_{eff} (g C MJ^{-1}) is influenced by variable, biotic factors and was calculated at each time step. The LUE_{max} , temperature stress factor (T_ϵ) (Equation 2.14) and water stress factor (W_ϵ) (Equation 2.15) were used to determine the effects of climatic conditions on the LUE.

The T_ϵ aims to include the influence of weather conditions on the opening and closing of stomata (Richters 2005a). As in the original RBM, the MODIS Land Surface Temperature (LST) and Emissivity product (MOD11) (Wan, Hook & Hulley 2015) was used as input. The MOD11 product is provided daily at a spatial resolution of 1 km. The day and night temperature bands were averaged to calculate mean daily temperature per pixel. As the MOD11 product has a daily temporal resolution, mean composites were created to correlate with the Sentinel-2 maximum-value NDVI composites (Section 5.1.1). The mean daily temperature composites were converted to the appropriate unit (Kelvin to Celsius) and implemented in Equation 2.14 (Section 2.2.2.1) to calculate a unitless T_ϵ scaling value between 0 and 1, where a higher value correlates to less temperature stress.

Water stress plays an important role in photosynthesis and plant growth and is quantified in the W_ϵ using Actual Evapotranspiration (AET) and Potential Evapotranspiration (PET) (Equation 2.15). In the original RBM, AET was calculated as a function of the MODIS Total Precipitable Water product (MOD05) (Gao 2015; Gao & Goetz 1990; Gao & Kaufman 1998) and PET was estimated as a function of the MODIS LST (MOD11) product (Wan, Hook & Hulley 2015) and MODIS Albedo (MOD43) product (Schaaf & Wang 2015; Strahler et al. 1999) using the Priestly-Taylor method (Priestley & Taylor 1972). Using this same approach, a MODIS Evapotranspiration (ET) product (MOD16) (Running, Mu & Zhao 2017) has since been developed at a 500 m spatial and 8-day temporal resolution. The MOD16 product includes both AET and PET as bands and was used in the research to calculate a unitless W_ϵ scaling value between 0 and 1, where a higher value correlates to less water stress. A drawback, however, of using the MODIS ET product are

notable gaps in the data which correspond to areas where ET modelling could not be performed due to a variety of reasons, including cloud cover (Running, Mu & Zhao 2017).

In a recent implementation of the RBM, Machwitz et al. (2015) found an overestimation of water stress during the dry season (i.e. lower W_{ε} values) and an underestimation of water stress during the wet, growing season (i.e. higher W_{ε} values). A correction coefficient was consequently developed, where the W_{ε} was increased by one third during the dry season and decreased by one third during the wet season (Machwitz et al. 2015). To rectify potential overestimation and underestimation of water stress, these correction coefficients were applied to the W_{ε} layer depending on seasonality.

Water stress is simultaneously one of the most influential and most uncertain factors in vegetation productivity modelling. Many models choose not to include the W_{ε} due to this uncertainty. However, as this research explored GPP estimation in both the dry season (and drought) and wet season, the W_{ε} was identified as an important factor to include.

5.1.3 Grazing capacity

Grazing capacity (GC) (ha LSU^{-1}) is dependent on the annual primary production of a region, as well as characteristics relating to cattle species that graze in the area, e.g. weight and daily consumption (Guevara et al. 1997). Figure 5.4 shows the process of GC calculation from RBM GPP estimates, net primary production (NPP), annual net primary production (aNPP), annual dry matter production (aDMP), FVC and the average cattle weight (σ).

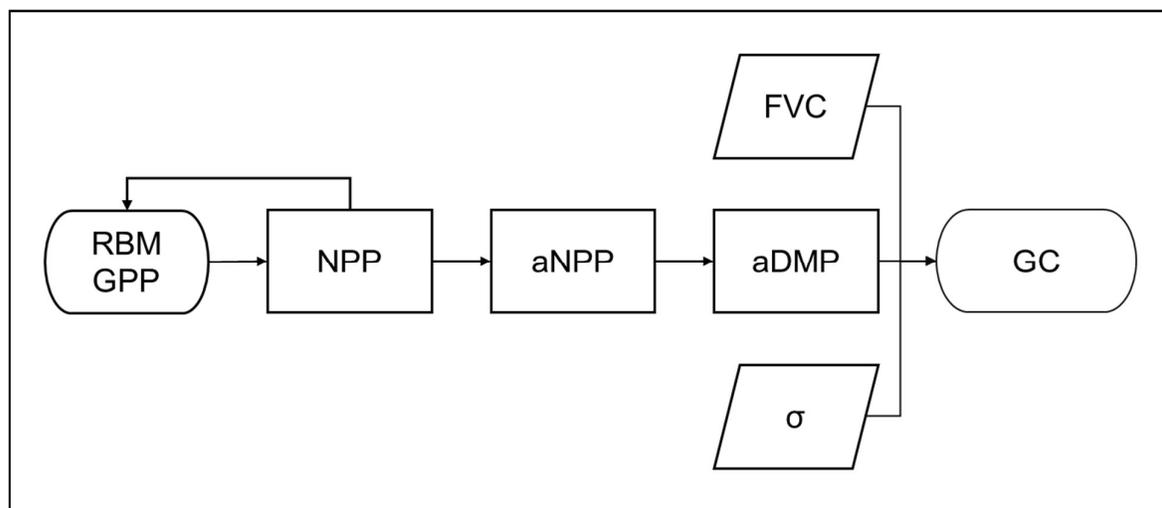


Figure 5.4 GC calculation diagram.

RBM GPP, Regional Biosphere Model gross primary production; NPP, net primary production; aNPP, annual net primary production; aDMP, annual dry matter productivity; FVC, fractional vegetation cover; σ , average cow weight; GC, grazing capacity.

GPP was accumulated annually and converted to NPP ($\text{g C m}^{-2}\text{year}^{-1}$) using a respiration conversion factor of 0.6 (Running, Mu & Zhao 2015), followed by conversion to dry matter production (DMP) ($\text{kg DM ha}^{-1}\text{year}^{-1}$), a primary production unit used for agro-statistical purposes (Swinnen & Van Hoolst 2018). To determine only the fraction of production contributed by grass and herbaceous vegetation, the Grassy component of the FVC (Chapter 4; Section 5.1.2.2) was used to scale the annual DMP (aDMP). The resulting grazable aDMP estimate was used in Equation 2.18 (Section 2.4.2), along with the average cattle weight (σ) of 450 kg, to calculate GC (ha LSU^{-1}). The use of FVC to identify grazable primary production for GC estimates has not been implemented in literature and is a novel approach to deriving grassland productivity in degraded and woody encroached rangelands.

5.1.4 Accuracy assessment

In the absence of extensive field data, which is almost impossible to acquire, the RBM GPP results, as well as the GC results, were assessed through comparison with other similar existing products. The RBM GPP results were also validated against in-situ field data.

5.1.4.1 Gross primary production

The RBM GPP results were assessed using two techniques: comparison to existing primary production products and validation against in-situ field data. The model-estimated GPP was compared to two benchmark primary production products. The MODIS GPP product (MOD17A2) (Running, Mu & Zhao 2015) has a spatial resolution of 500 m and temporal resolution of 8-days providing accumulated daily GPP which was divided by eight for correspondence to RBM estimated mean daily GPP. The CGDMP product (Swinnen & Van Hoolst 2018) provides mean primary production estimates at 10-day temporal resolution and has a spatial resolution of 300 m, thus no division was required. GPP product layers were converted to GDMP, the agro-statistical equivalent, with the appropriate units (kg DM ha^{-1} to g C m^{-2}) to allow comparison with RBM results.

A comparison was done for both the Fort Beaufort study area (June 2018) and Cedarville study area (February 2019) (Section 1.5) using the image acquisition dates summarised in Table 5.1. To ensure comparability between the model-estimated GPP and the respective benchmark products, the time period used to generate the input Sentinel-2 derived composites for the RBM (Section 5.1.1) were adapted to correspond to the temporal resolution of the benchmark products. Note that the CGDMP has an 10-day temporal resolution throughout the year with the exception of the end of February (21 to 28 February), where an 8-day modelling period is implemented to accommodate

for the shorter month. The corresponding composite date ranges were thus also adapted to ensure comparability between the model-estimated GPP and CGDMP product.

Table 5.1 Image acquisition dates and composite time periods used for benchmark product comparison

Benchmark Product	Study Area	Benchmark Product Images	RBM Composite Time Period	RBM Composite Date Range
MOD17A2	Fort Beaufort	10 June 2018	8 days	3 – 10 June 2018
		18 June 2018		11 – 18 June 2018
		26 June 2018		19 – 26 June 2018
	Cedarville	10 Feb 2019	8 days	3 – 10 Feb 2019
		18 Feb 2019		11 – 18 Feb 2019
		26 Feb 2019		19 – 26 Feb 2019
CGDMP	Fort Beaufort	10 June 2018	10 days	1 -10 June 2018
		20 June 2018		11 – 20 June 2018
		30 June 2018		21 – 30 June 2018
	Cedarville	10 Feb 2019	10 days	1 -10 Feb 2019
		20 Feb 2019		11 – 20 Feb 2019
		28 Feb 2019		21 – 28 Feb 2019

The model-estimated GPP results were resampled to suitable spatial resolutions (300 m; 500 m) and quantitatively compared to the benchmark products using mean absolute error (MAE) and root mean square error (RMSE). The model-estimated GPP results were also validated against in-situ field measurements of DMP collected during the Cedarville growing season field trip (Section 3.2.2.4). GPP was calculated using an 8-day input composite for the date range 21 to 28 Feb 2019. Field-measured DMP was converted to GPP (kg DM ha^{-1} to g C m^{-2}) for comparative purposes. RMSE and MAE were calculated to quantitatively assess the agreement between the model-estimated primary production and in-situ primary production estimates.

5.1.4.2 Grazing capacity

The GC results were compared to the Long Term Grazing Capacity (LTGC) map for South Africa (Section 2.4.2; Appendix B), created for the year 2016 (Avenant 2017). The LTGC map is based on expert opinion and aims to provide long term norms and guidelines for sustainable management. The map groups areas with similar GC into discrete grazing zones. Figure 5.5 shows the discrete LTGC grazing zones for (a) Fort Beaufort and (b) Cedarville labelled with unique zone IDs.

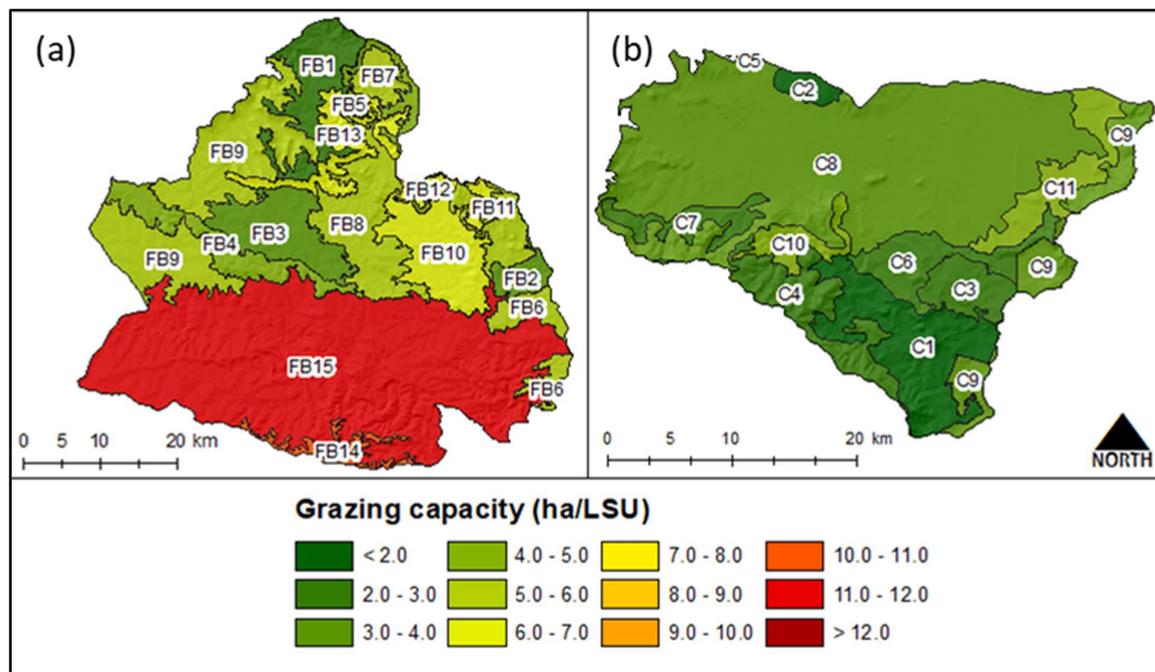


Figure 5.5 LTGC map grazing zones for (a) Fort Beaufort and (b) Cedarville

Zonal statistics were performed on the model-estimated GC maps to statistically compare the modelled GC to the knowledge-based LTGC map. Furthermore, the distribution of modelled GC within each LTGC zone was plotted using normalised frequency histograms. This statistical comparison was done for the year 2016 to ensure temporal consistency with the LTGC map. Annual GC change was also explored by comparing modelled GC for 2016, 2017 and 2018 for both Fort Beaufort and Cedarville.

5.1.5 Application development

A web app, GrazeEngine, was developed to demonstrate the potential practical implementation of the described models for grassland condition and productivity management. A web app was chosen as it is easier to update and maintain when compared to a mobile app. The web app was developed with GEE (Gorelick et al. 2017) serving as the core backend to allow dynamic and continuous cloud geoprocessing. The adapted RBM (Figure 5.1) was implemented in GEE to calculate GPP using an 8-day composite time period, using the LSMM for FVC estimation (Chapter 4). These GPP estimates were then accumulated annually, converted to agro-statistical units and scaled using the Grassy component of the FVC to determine grazable primary production (Figure 5.4), from which the GC per pixel was calculated. The web app provides both the LTGC map and the model-estimated GC as map layers. Other outputs include seasonal NDVI and the FVC, to give an estimate of grassland condition, as well as annual and seasonal vegetation primary production estimates. The app also allows the generation of per-pixel graphs and statistics generated when a location on the map is clicked.

The user interface was designed for use by farmers and management consultants within the livestock industry. As these individuals often have limited experience in RS, GIS and spatial analysis, input parameters were kept to a minimum and limited to the specification of location and time periods. More input parameters, i.e. breed of cattle, grass species, most recent burning etc. can be incorporated in future work. Currently the functionality of the implemented model is only demonstrated for two example locations, i.e. the Fort Beaufort and Cedarville study areas, as more extensive field validation is required to ensure model transferability. Refer to Appendix F for example screenshots of the user interface. Click on the following link to explore the functionality of the web app: <https://sites.google.com/view/graze-engine/home>

5.2 RESULTS

This section summarises the results of GPP and GC calculation and validation. The practical implementation of the models for grassland condition and productivity estimation is also presented by demonstrating the results in the web app.

5.2.1 Gross primary production

5.2.1.1 Gross primary production output maps

Figure 5.6 shows a side-by-side comparison of the (a) Fort Beaufort and (b) Cedarville study areas of (1) aerial imagery, (2) model-estimated GPP (10 m), as well as the two benchmark products, (3) CGDMP GPP (300 m) and (4) MOD17A2 GPP (500 m).

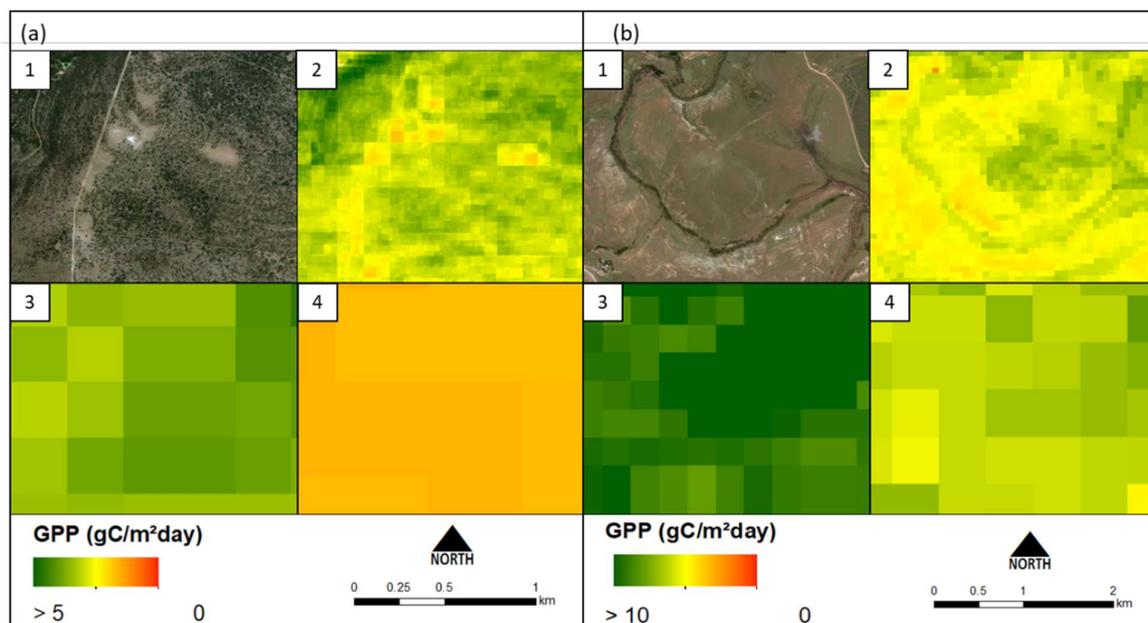


Figure 5.6 Comparison of (a) Fort Beaufort and (b) Cedarville study areas of (1) aerial imagery (2) model-estimated GPP (10 m), (3) CGDMP GPP (300 m) and (4) MOD17A2 GPP (500 m)

The maps highlight the variation of GPP within each of the benchmark product pixels. As seen in Figure 5.6b, degraded patches of grassland and bare soil are not reflected in the CGDMP and MOD17A2 products as the resolution is too coarse.

5.2.1.2 Comparison to existing primary production products

To assess the RBM model performance against the two benchmark products for similar seasonal and spatial conditions, RMSE and MAE values were compared for two correlating dates, 10 June 2018 and 10 Feb 2019 (Table 5.1). Figure 5.7 shows the comparison of agreement between the model-estimated GPP and the MOD17A2 and CGDMP products for 10 June 2018 (Fort Beaufort) and 10 February 2019 (Cedarville).

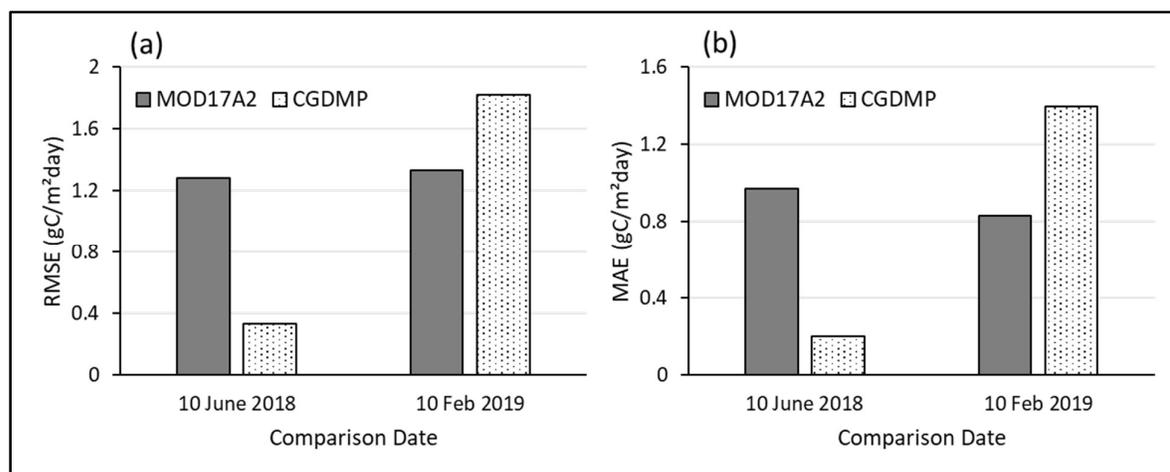


Figure 5.7 Comparison of model-estimated GPP to MOD17A2 and CGDMP products using (a) RMSE and (b) MAE. RMSE and MAE values for model-estimated GPP vs. MOD17A2 GPP were relatively similar for 10 June 2018 and 10 February 2019, with values ranging between 1.0 to 1.3 $\text{g C m}^{-2}\text{day}^{-1}$. RMSE and MAE values for model-estimated GPP vs. CGDMP GPP, however, increased notably between the two acquisition dates. Agreement for the 10 June 2018 (Fort Beaufort) was very good, with an RMSE and MAE less than 0.4 $\text{g C m}^{-2}\text{day}^{-1}$. An increase to more than 1.2 and 1.6 $\text{g C m}^{-2}\text{day}^{-1}$ for RMSE and MAE values respectively was observed for 10 February 2019 (Cedarville). Table 5.2 summarises the comparison results of the model-estimated GPP and benchmark products.

Table 5.2 Comparison of model-estimated GPP to benchmark products using RMSE, MAE and mean

Benchmark Product	Study Area	Comparison Date	RMSE	MAE	Mean
MOD17A2	Fort Beaufort	10 June 2018	0.43	0.24	1.36
		18 June 2018	1.19	0.89	1.40
		26 June 2018	1.48	1.27	1.06
	Cedarville	10 Feb 2019	1.85	1.31	6.12

		18 Feb 2019	1.66	1.12	6.25
		26 Feb 2019	1.46	1.05	6.41
		10 June 2018	0.33	0.20	2.14
	Fort Beaufort	20 June 2018	0.57	0.37	2.34
		30 June 2018	0.52	0.39	2.14
CGDMP		10 Feb 2019	1.82	1.39	7.95
	Cedarville	20 Feb 2019	3.16	3.63	9.77
		28 Feb 2019	4.08	3.90	10.64

The RMSE and MAE values for both MOD17A2 and CGDMP for Fort Beaufort (June 2018) is low, with error estimates under $1.5 \text{ g C m}^{-2}\text{day}^{-1}$. RMSE and MAE values remain low for the MOD17A2 product in the Cedarville area, however, agreement with the CGDMP product is poor for this study area. High RMSE and MAE values are achieved, specifically for the 20th and 28th of Feb 2019. This increase correlates directly with the increase in moisture in the area, as was discussed in Section 3.2.2. Comparing the mean GPP for MOD17A2 and CGDMP for the month of February 2019 (Cedarville), CGDMP estimates higher GPP values ($7.65 - 10.64 \text{ g C m}^{-2}\text{day}^{-1}$) than MOD17A2 ($6.12 - 6.41 \text{ g C m}^{-2}\text{day}^{-1}$).

5.2.1.3 Validation against in-situ field data

Table 5.3 summarises the results of the GPP validation against in-situ field measurements collected during the growing season field trip (Section 3.1.2.4). The model-estimated GPP was assessed, as well as the MOD17A2 GPP and CGDMP GPP, using RMSE and MAE.

Table 5.3 Results of GPP validation against in-situ field measurements

Product	RMSE	MAE	Mean
Model GPP	2.00	1.54	5.87
MOD17A2	3.04	2.65	7.06
CGDMP	5.48	5.90	11.98

The model-estimated GPP produced RMSE and MAE values of $2.00 \text{ g C m}^{-2}\text{day}^{-1}$ and lower, with an overall mean of $5.87 \text{ g C m}^{-2}\text{day}^{-1}$ for all field points (Table 5.3). Both MOD17A2 and CGDMP had RMSE, MAE and mean values higher than the model-estimated GPP. The CGDMP, in particular, produced a mean GPP of $11.98 \text{ g C m}^{-2}\text{day}^{-1}$ for all field points, which is notably higher than the average $6 - 8 \text{ g C m}^{-2}\text{day}^{-1}$ for grasslands during the growing season (Rossini et al. 2014).

5.2.2 Grazing capacity

5.2.2.1 Comparison to long term grazing capacity map

The model-estimated GC was compared to the LTGC map using zonal statistics and frequency histograms. Figure 5.8 and Figure 5.9 show a visual comparison of the (a) LTGC map and (b) model-estimated GC for Fort Beaufort and Cedarville, where the majority model-estimated GC per LTGC zone is shown.

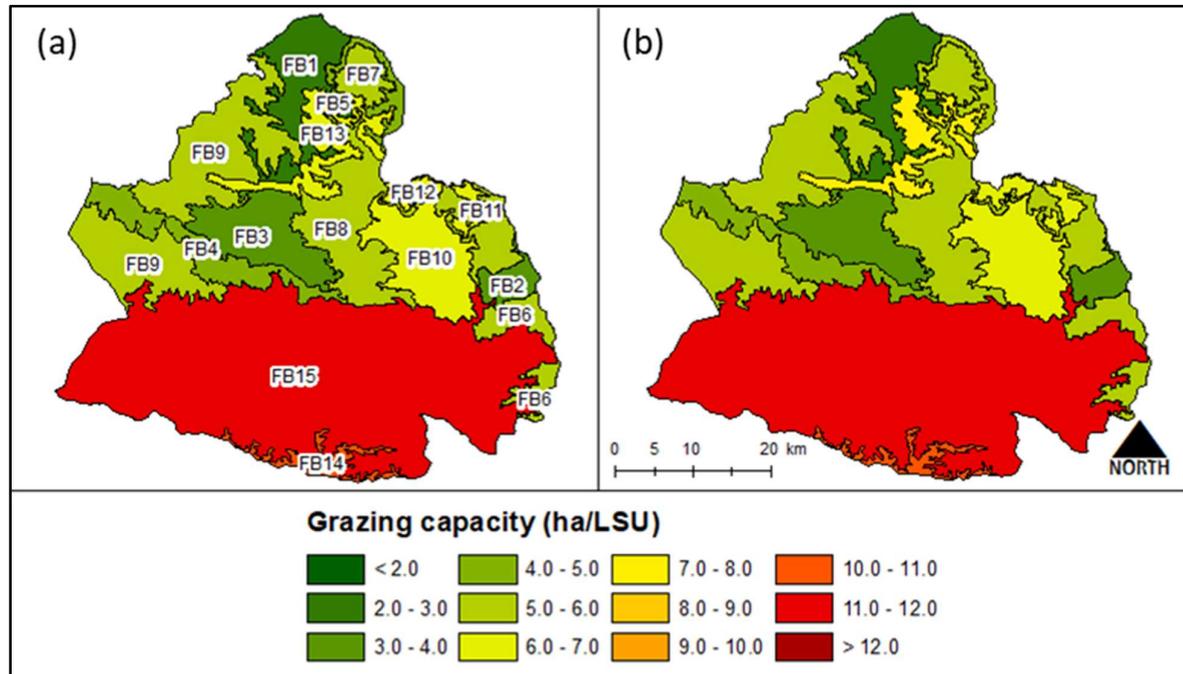


Figure 5.8 Comparison of (a) LTGC map (2016) to (b) majority model-estimated GC (2016) for the Fort Beaufort study area

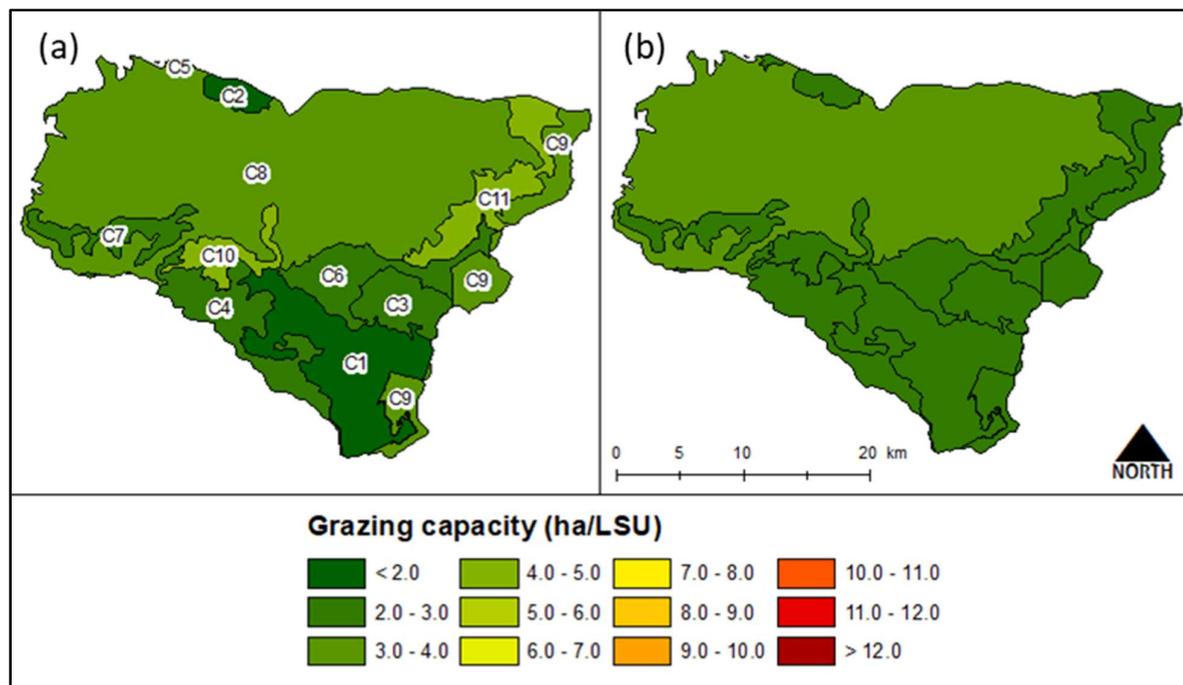


Figure 5.9 Comparison of (a) LTGC map (2016) to (b) majority model-estimated GC (2016) for the Cedarville study area

Table 5.4 summarises the statistical results of the model-estimated GC and LTGC map comparison. A variety of zonal statistics were calculated for each LTGC zone to explore the variation of model-estimated GC within each zone. The zone ID corresponds to the ID labels in Figure 5.8 and Figure 5.9.

Table 5.4 Comparison of model-estimated GC to LTGC map using zonal statistics

Study Area	Zone ID	GC (ha/LSU)	Mean	Majority	Median	Std Dev
Fort Beaufort	FB1	2.5	4.87	3	4	3.09
	FB2	4	6.22	3.5	4	2.45
	FB3	4	6.88	4	4	1.89
	FB4	5	4.50	4.5	4	2.53
	FB5	5	6.75	5.5	5	3.02
	FB6	6	5.96	6	5	2.58
	FB7	6	6.27	6	5	3.22
	FB8	6	7.34	6	5.5	2.76
	FB9	6	7.73	6	5.5	2.80
	FB10	7	6.98	6.5	5.5	2.53
	FB11	7	7.17	7	5.5	3.15
	FB12	7	6.93	7	6	2.62

	FB13	7	6.90	8	6	3.45
	FB14	11	9.23	11	6.5	3.17
	FB15	12	12.18	12	7.5	3.19
Cedarville	C1	2	3.19	2.5	3	1.48
	C2	2	2.87	3	3	0.49
	C3	2.5	2.88	2.5	3	0.67
	C4	2.5	3.23	2.5	3	1.38
	C5	2.5	3.30	3	3	0.98
	C6	3	3.11	3	3	0.78
	C7	3	3.07	3	3	0.85
	C8	3.5	3.11	3.5	3	1.05
	C9	3.5	3.19	3.5	3	1.36
	C10	5	3.40	2.5	3	1.27
	C11	5	3.20	2.5	3	1.05

Variation within the Fort Beaufort study area was much higher than in Cedarville, with a mean standard deviation of 2.49 ha LSU⁻¹ per zone compared to Cedarville's 1.03 ha LSU⁻¹. The majority statistic for model-estimated GC per zone correlated well with categorical zone GC from the LTGC map (Figure 5.8; Figure 5.9; Table 5.4), whereas the mean model-estimated GC was skewed by pixels of very high or low values within the zone. Overestimation of GC occurred in zone FB1, FB5, FB13, C1, C2 and C5 (Table 5.4). As these overestimations were small (less than 0.5 ha LSU⁻¹), they did not reflect visually in Figure 5.8 and Figure 5.9. Overestimated areas typically correlated to grasslands with large degraded patches of shrub encroachment, or contained small urban settlements that are not included in the NLC2017/18. Underestimation of GC occurred in zones FB2, FB4, FB10, C10 and C11 (Table 5.4). The model-estimated GC values for zones C10 and C11 were particularly low compared to the LTGC value of 5 ha LSU⁻¹, as seen in both Figure 5.9 and Table 5.4. This discrepancy is likely due to species composition and palatability, which was taken into account in the knowledge-based allocation of GC values to the LTGC map, but not for the model-estimated GC.

The distribution of model-estimated GC values within each LTGC zone was also plotted using normalised frequency histograms. Selected histograms are shown for Fort Beaufort and Cedarville in Figure 5.10 and Figure 5.11 respectively. Refer to Appendix E for the frequency histogram of

each grazing zone. The bins correlate to LTGC classification scheme, while the red dashed line corresponds to the categorical LTGC map value for that particular zone.

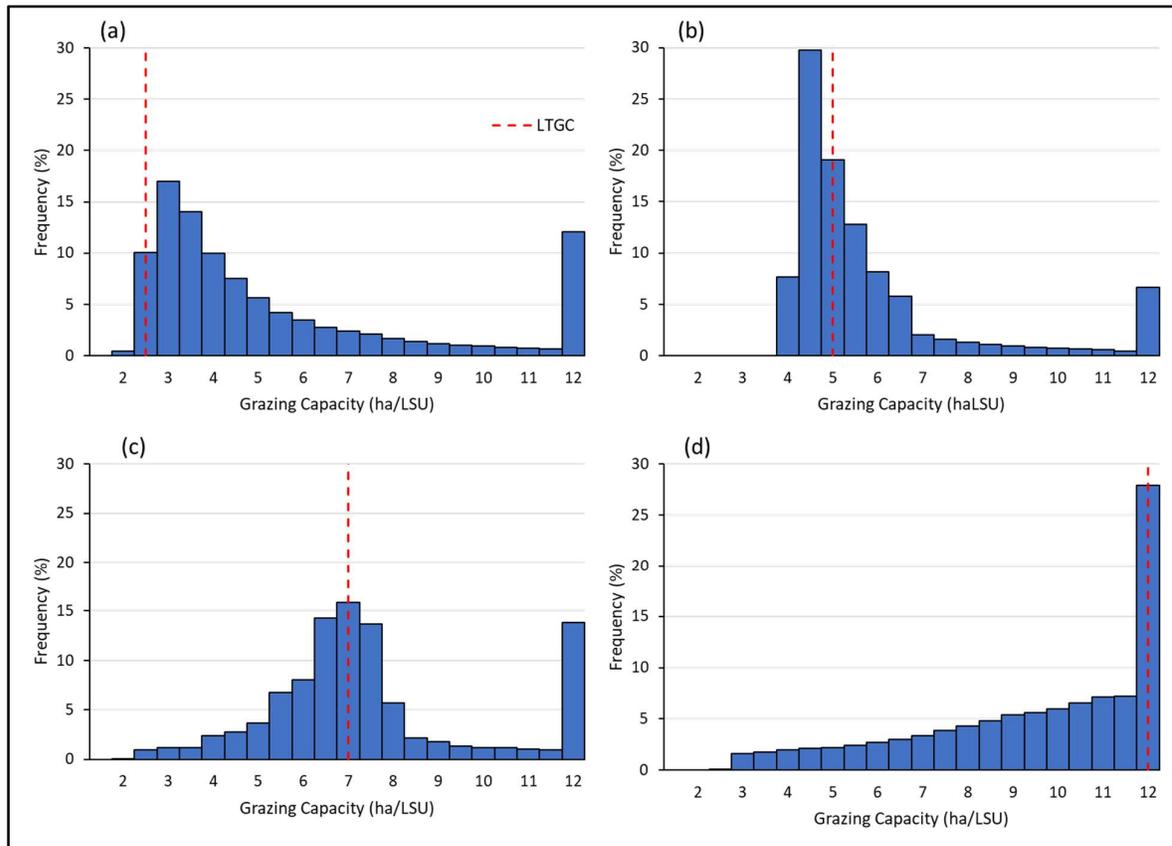


Figure 5.10 Frequency histograms for Fort Beaufort LTGC zones (a) FB1, (b) FB4, (c) FB11 and (d) FB14

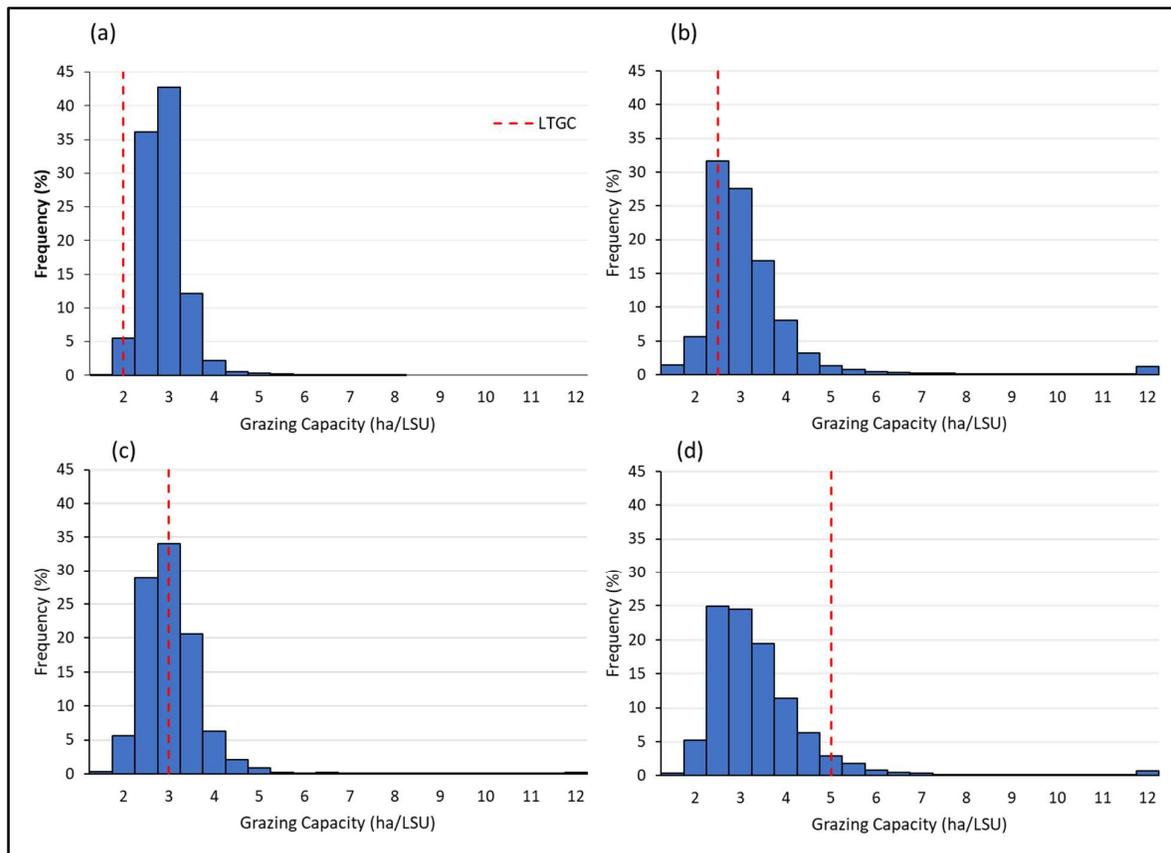


Figure 5.11 Frequency histograms for Cedarville LTGC zones (a) C2, (b) C4, (c) C8 and (d) C11

Although the model underestimated GC slightly in some zones (Figure 5.10d), model-estimated GC showed relatively good agreement with the LTGC zone map values for the Cedarville study area. The homogeneity of the region is reflected in the frequency distributions, with majority values ranging from 2 – 3.5 ha LSU⁻¹ (Figure 5.11). In contrast, the variation of model-estimated GC per zone in the Fort Beaufort area is reflected in the frequency distributions (Figure 5.10). Each zone consists of a wide range of GC values, including a large number of pixels estimated at 12 ha LSU⁻¹. Even though the visual comparison of model-estimated GC to the LTGC map shows very good agreement for the Fort Beaufort study area (Figure 5.8), the zonal statistical analysis (Table 5.4) and frequency distribution histograms (Figure 5.10) quantifies the variation in GC present within these zones. To further visualise this variation, Figure 5.12 shows two areas within the FB3 grazing zone.

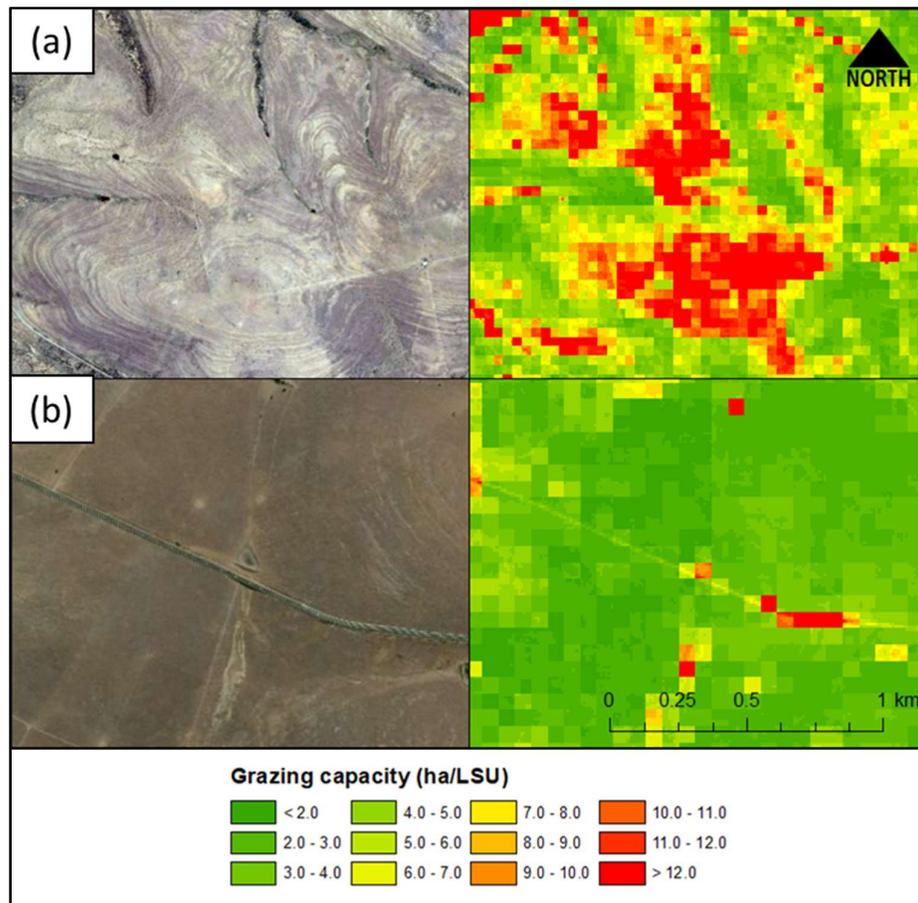


Figure 5.12 Comparison of (a) poor grazing conditions to (b) ideal grazing conditions within LTGC zone FB3

Although FB3 has an LTGC map value of 4 ha LSU⁻¹, it consists of a diverse combination of eroded and woody encroached fields (Figure 5.12a) and pristine grassland (Figure 5.12b). Grazing conditions thus range from very poor (12 ha LSU⁻¹) to ideal (2.0 – 0.3 ha LSU⁻¹) within a seemingly homogenous grazing zone.

5.2.2.2 Annual grazing capacity change

Figure 5.13 shows the mean monthly rainfall (mm) and mean GC (ha LSU⁻¹) for (a) Fort Beaufort and (b) Cedarville for the years 2016, 2017 and 2018.

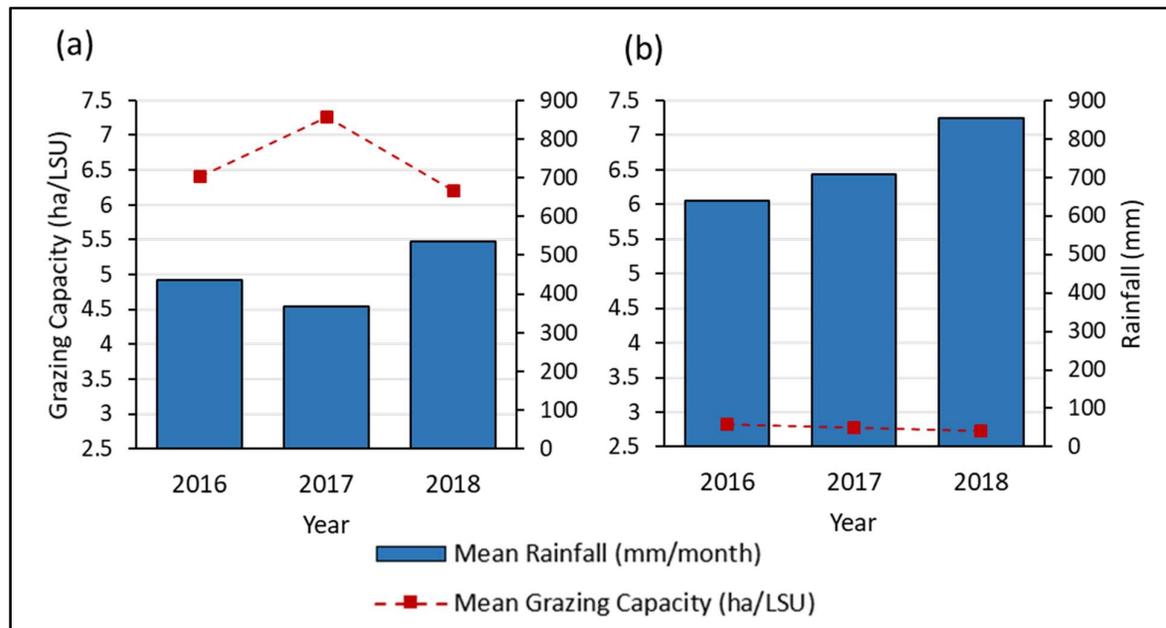


Figure 5.13 Mean monthly rainfall (mm) and mean GC (ha/LSU) for (a) Fort Beaufort and (b) Cedarville

The increase in model-estimated GC estimates corresponds to a decrease in mean monthly rainfall for the year 2017 in Fort Beaufort (Figure 5.13a). The subsequent year of 2018 shows a substantial increase in mean monthly rainfall, coupled with a notable decrease in GC estimates. The Cedarville study area experienced an increase in mean monthly rainfall and a decrease in GC estimates from 2016 to 2018 (Figure 5.13b). Figure 5.13 thus shows the sensitivity of GC to rainfall, as decreases in rainfall result in poorer grazing conditions and vice versa.

5.3 DISCUSSION

The adapted RBM, implemented to calculate GPP, performed well based on visual inspection and validation against in-situ field data (Section 5.2.1.1; Section 5.2.1.3). Comparison of the model-estimated GPP to existing primary production products (MOD17A2; CGDMP) showed good agreement (Section 5.2.1.2), with the exception of the model-estimated GPP vs. CGDMP comparison in the Cedarville study area during the February 2019 growing season (Table 5.2). Although RMSE and MAE values were relatively high for this comparison ($1.39 - 4.08 \text{ g C m}^{-2} \text{ day}^{-1}$), validation against the GPP in-situ field trip measurements (Table 5.3) suggests that the CGDMP overestimated GPP for Cedarville in the growing season. This statement is supported by the high mean CGDMP GPP values during both the benchmark product comparison ($7.95 - 10.64 \text{ g C m}^{-2} \text{ day}^{-1}$) and the in-situ field validation ($11.98 \text{ g C m}^{-2} \text{ day}^{-1}$) with respect to the model-estimated GPP, MOD17A2 GPP and in-situ field-measured GPP. The improved estimation of GPP by the RBM in wet conditions is likely due to the inclusion of water stress using the W_e , as well as a relief factor (x_{relief}), in scaling the maximum LUE (LUE_{max}). CGDMP does not include either of these factors in primary production calculation (Swinnen & Van Hoolst 2018),

thus no restrictions are placed on how much biomass can be produced during a particularly green, rainy growing season. Although the W_{ϵ} introduces uncertainty into GPP calculations through the use of the MODIS ET product (Section 5.1.2.3), the factor proved essential in areas affected by extreme climatic events e.g. drought or very heavy rainfall. Future research could potentially improve this uncertainty by implementing a different method of ET estimation when calculating W_{ϵ} . The difference in factors used for LUE_{\max} scaling, as well as the varying definitions and units of input parameters in literature, introduces inherent uncertainty when comparing the results from different primary production models. There is evidently a need for the standardisation of input stress factors, units, definitions etc. to allow better transparency and comparability in primary productivity modelling.

The model-estimated GC estimates show good correlation with the LTGC map. Underestimation of model-estimated GC compared to LTGC zones (i.e. lower GC values and better grazing conditions) was likely due to the inability of the model to include species composition and palatability in the GC calculations. These factors play a key role in the suitability of different grassland areas for cattle grazing (Raufirad et al. 2016) and were included in the LTGC creation through expert opinion (Avenant 2017). Overestimation of model-estimated GC compared to LTGC zones (i.e. higher GC values and poorer grazing conditions) is largely due to the finer spatial resolution of the model-estimated GC, which led to more variation within each grazing zone. As the model estimates GC on a per-pixel basis, very low GCs due to bare soil patches, shrub-encroached grasslands and dense bush/tree clumps is reflected in the statistics per LTGC grazing zone (Table 5.4). This is particularly prominent in the Fort Beaufort study area, which consists of a complex, heterogeneous combination of LC components. This high level of variation within the LTGC, shown by the zonal statistics (Table 5.4) and frequency histograms (Figure 5.10; Figure 5.11), suggests that the LTGC zonal delineations are not ideal for ecosystems with complex vegetation cover dynamics. Currently, these delineations are based primarily on vegetation type (Mucina & Rutherford 2006), with certain regions relying also on cadastral farm boundaries (Avenant 2017). As shown by the results of FVC estimation (Chapter 4), structural dynamics within study areas such as Fort Beaufort is increasingly complex due to grassland degradation and woody encroachment. This continual change in FVC suggests that static layers describing general vegetation type might not be suitable as the primary source for grazing zone delineation. As the model-estimated GC maps incorporate climate, relief and sub-pixel vegetation cover, this approach can potentially aid in both providing more detail within the current LTGC zones and adapting the current LTGC zone delineations to produce more homogenous grazing regions.

The annual change in GC estimates from 2016 to 2018 (Figure 5.13, Table 5.4) shows the need for more frequently updated GC maps. The impact of a low rainfall year on the overall GC of a region is reflected in Figure 5.13a, where grazing conditions worsen with a decrease in mean monthly rainfall. The model-estimated GC maps thus not only provide detailed estimates within the LTGC zones, but also allow comparing current, updated grazing conditions with the LTGC map to allow adjustment of livestock grazing practices as required. The practical implementation of this research in a web app (Section 5.1.5), therefore could be a useful tool for decision-making by providing broad, knowledge-based long term norms, as well as continuous, dynamically modelled GC estimates at an in-field level.

The research addresses one of the main constraints in literature relating to RS for GC estimation i.e. woody components in grasslands/rangelands reflecting as areas of very favourable GC. The inclusion of the model-estimated FVC (Chapter 4) provided LC condition information at sub-pixel level, which allowed for improved GPP calculation using the RBM. It also contributed to developing a novel approach to remotely-sensed GC estimation by scaling primary production according to the percentage of grassy (grazable) vegetation per pixel, thus addressing the aforementioned constraint. Although a limited number of studies have successfully addressed this constraint through the integration of RS data (Adjorlolo & Botha 2015; Espach, Lubbe & Ganzin 2009), the methods typically still require extensive field sampling for one or more input parameters of the GC calculation. Another constraint, however, is the inability to determine species composition in grasslands and rangelands without very high resolution satellite imagery. The effect of this constraint is reflected in the overestimation of GC in certain zones (Table 5.4), as the suitability of different grass species for grazing could not be mapped and included in calculations. A potential solution would be to include species composition as an input parameter within the web app. Users would be able to determine species composition in-field and provide suitable information as input, allowing the app to scale the model-estimated GC estimates appropriately. Although this solution will require the user to have expert, local knowledge of the area, it is a temporary solution until the constraint can be addressed in RS research.

To conclude, the FVC-integrated RBM shows potential in providing continuous, dynamic grassland and condition data for sustainable livestock management. Although field surveys of grassland biomass and GC are still the most accurate, these field methods cannot be applied at on a large scale as it requires destructive sampling (Eisfelder et al. 2017). RS-based approaches allow analysis of large areas, however, current RS-based primary production models produce coarse resolution results or lack transferability. The local-level resolution, spatial and temporal transferability of the discussed model, as well as its ability to distinguish between grassy (grazable)

productivity and woody productivity, thus provides a novel approach to monitoring and managing grasslands affected by degradation and woody plant encroachment.

CHAPTER 6: DISCUSSION AND CONCLUSION

This chapter presents a summary, discussion and critical evaluation of the collective results of the study. The specified aim and objectives are revisited, main findings of the research are discussed, suggestions for future research are made and conclusions are drawn.

6.1 REVISITING AIM AND OBJECTIVES

This research aimed to develop a model that integrates fractional vegetation cover (FVC) and primary production to estimate grassland productivity and condition at various spatial scales using Google Earth Engine (GEE). Grasslands are under continuous threat of degradation due to overgrazing and woody plant encroachment, which has a detrimental impact on productivity and condition. As grasslands play a crucial role in the grazing of livestock, there is a need for sustainable grassland management systems that provide identification, analysis, quantification and potential prevention of degradation occurrences and their detrimental impacts on grassland productivity. Although remote sensing (RS) has been used extensively and successfully for monitoring and mapping vegetation dynamics, current approaches have technological and scientific constraints. Traditional approaches require expensive hardware, software, imagery and training and typically provide results in the form of static maps. This limits the use of these approaches for continuous, dynamic and multi-scale estimation of grassland productivity, as required for the effective management of temporally variable ecosystems such as grasslands. Current RS-based grassland productivity approaches also do not take into account the effect of woody components (shrub/trees) on productivity dynamics i.e. grazing capacity (GC). These models can thus not be applied in the context of livestock management, as they incorrectly depict areas of woody cover as very high primary production and thus favourable GC, whereas in reality they are not suitable for cattle grazing. FVC was identified as a prospective solution to the scientific challenges of current approaches, while cloud computing and application development showed potential in addressing technological constraints. The research thus investigated the use of RS-based FVC estimation and vegetation production modelling, as well as current cloud computing technologies such as GEE, for accurate, transferable and robust modelling of grassland condition and productivity.

The first objective was to conduct a literature review on existing approaches for calculating grassland productivity and FVC, current grassland and rangeland management practices and tools available for geoprocessing cloud computing (Chapter 2). Literature supported RS as a valuable tool for the mapping and monitoring of vegetation dynamics, specifically in relation to vegetation production and structure. A vast range of primary production models have been developed,

implemented and evaluated to allow the quantification of vegetation productivity. The Regional Biosphere Model (Richters 2005b), a light use efficiency (LUE)-based gross primary production (GPP) model, was identified as most suitable for this research based on its semi-mechanistic approach, choice of input datasets and past successful implementations in southern African grasslands and savannas. A prominent constraint of current RS-based primary production models is the inability to distinguish tree production from herbaceous production at sub-pixel level. Current methods for grassland and rangeland management are thus still primarily field-based and involve destructive, time-consuming sampling. The integration of FVC was posed as a promising solution, where estimation using spectral mixture analysis (SMA) shows the greatest potential for transferable calculation using relatively few pure endmember points. Finally, the cloud computing geoprocessing and web application (app) functionality of GEE was highlighted as key component in providing continuous, dynamic RS analysis results. When combined with FVC and primary production estimates, a GEE web app can potentially allow dynamic, in-field and locally-tailored productivity estimates that address both the technological and scientific challenges in current literature.

The purpose of the second objective, the collection and analysis of relevant biophysical field data, was to establish crucial relationships between vegetation productivity and RS signals for the development of robust, transferable production models. This objective was addressed in Chapter 3, where data was collected on two field trips in two different study areas: during the dry season (June 2018) in Fort Beaufort and the growing season (February 2019) in Cedarville. The biophysical parameters measured and analysed include FVC, leaf area index (LAI), the fraction of absorbed photosynthetically active radiation (fAPAR) and grass clippings for dry matter (DM) production calculation. Results were used for both validation and calibration purposes in Chapter 4 and 5.

The estimation of FVC using SMA was performed in Chapter 4 (Objective 3). A linear spectral mixture model (LSMM) was calibrated and implemented to estimate FVC using freely available, medium resolution imagery i.e. Landsat 8 and Sentinel-2. The results were evaluated based on agreement with an existing FVC product (Copernicus Dynamic Land Cover (CDLC) product) and field data (Chapter 3), thus partially addressing Objective 6. The temporal and spatial transferability of the LSMM was also assessed by implementing the calibrated LSMM model in a different season and study area and the resulting FVC estimates were evaluated using field measurements (Chapter 3). The results of Chapter 4 thus also addressed the first research question: how can FVC be dynamically estimated for grasslands using medium resolution imagery?

The fourth objective aimed to develop a model that combines FVC (Chapter 4) with productivity calculations to improve grassland condition and productivity estimates. A FVC-integrated productivity model was consequently developed to provide grazing capacity estimates for grassland and rangelands (Chapter 5). Annual grassland production estimates was calculated using an adapted RBM. FVC was used to determine the grazable primary production from these RBM estimates, thus mitigating the effects of woody components on grazing capacity calculations. As a result, this chapter targeted the second research question: how can FVC aid in determining accurate grassland productivity estimates? Both the RBM primary production estimates and subsequent GC results were qualitatively and quantitatively assessed (Objective 6). Two methods were used to evaluate RBM primary production estimates: a comparison to existing products (MODIS GPP; Copernicus Gross Dry Matter Production (CGDMP)) and validation using field data (Chapter 3). The subsequent GC maps were visually and statistically compared to the most recent national product for South Africa i.e. the Long Term Grazing Capacity (LTGC) map for 2016.

Sustainable grassland and livestock management systems require access to continuous, dynamic condition and productivity data, therefore a web app was developed to provide transferable, multi-scale estimates of grassland productivity dynamics using RS data and analysis (Objective 5). The FVC-integrated productivity model (Objective 4) was implemented in the GEE environment and packaged as a web app. The tool can potentially address both the scientific and technological limitations identified in literature (Chapter 2; Objective 1) by providing dynamic estimates that incorporate the effects of woody components on productivity calculations. The model implementation in a GEE web app environment is documented in Chapter 5 and addresses the final research question: how can grassland condition and productivity dynamics be disseminated at various spatial scales using evolving geoprocessing technologies?

The sixth objective, quantitatively and qualitatively assessing the research results, is collectively addressed in both Chapter 4 and 5 using field data acquired in Chapter 3 and existing products in the public domain. Recommendations on the use of FVC, primary production estimates and RS for grassland condition and productivity modelling (Objective 7) is discussed in Section 6.3 of this chapter.

6.2 MAIN FINDINGS AND VALUE OF THE RESEARCH

Chapter 3, 4 and 5 investigated the role of field data in vegetation modelling, the use of SMA for transferable estimation of FVC and the potential of a FVC-integrated model to provide improved grassland condition and productivity estimates. The findings of these three components and their contributions to research are summarised in this section.

6.2.1 Field data collection and analysis for vegetation modelling

The collection and analysis of various biophysical field parameters formed a key component of this research. Vegetation modelling is a complex process that relies heavily on quality field data to accurately represent the relationship between vegetation dynamics and RS responses (Eisfelder et al. 2017). Various biophysical parameters were measured for calibration and validation purposes, in particular fAPAR. It is a key variable for assessing vegetation productivity and yield estimates (Fensholt, Sandholt & Rasmussen 2004), however, it is often difficult to establish a strong relationship between fAPAR and satellite observation (Tian et al. 2002). Although a number of fAPAR products are available in the public domain (i.e. MODIS LAI/fAPAR), these products usually model fAPAR at a global-scale, coarse spatial resolution. An essential result of the field data collection and analysis carried out in this research (Chapter 3) was establishing an improved relationship between satellite signals (NDVI) and fAPAR for subsequent vegetation productivity modelling.

Conventionally, a NDVI-to-fAPAR relationship is established using regression analysis, where NDVI is the independent variable and fAPAR is the dependent variable (Fensholt, Sandholt & Rasmussen 2004; Pickett-Heaps et al. 2014; Rahman & Lamb 2017; Wang et al. 2016). This research used the proportions of green, living vegetation (obtained from grass clippings) to scale fAPAR field measurements, thus essentially deriving the fAPAR absorbed only by photosynthesising components. This scaling of fAPAR improved regression relationships with NDVI and reduced overall variation of fAPAR per NDVI interval. Notable improvements in R^2 were achieved, with an increase of 0.56 to 0.84 for dry conditions and 0.33 to 0.69 for wet, rainy conditions. These results support the findings of Sakowska, Juszczak & Gianelle (2016), one of the few studies that have explored this approach to improve biophysical relationships. As the technique is not well-documented in literature, the results from this research provides a key contribution to further understanding and exploring the potential use of green scaling to improve field-to-satellite relationships for vegetation productivity modelling. For more information on the various biophysical parameters measured and analysed see Chapter 3.

6.2.2 FVC estimation using SMA and medium resolution imagery

FVC plays a crucial role in modelling landscape surface processes (Guerschman et al. 2009). Although FVC estimation using RS-based approaches is extensively documented in literature, the majority of approaches require very high resolution or hyperspectral imagery (Gessner et al. 2013; Guerschman et al. 2009; Zhang et al. 2012) or a large number of field points (Jia et al. 2015; Liu et al. 2018; Yang et al. 2016). In addition, current existing products typically have a coarse spatial resolution (e.g. MODIS Vegetation Continuous Fields (VCF) product) and/or have unsuitable

FVC class definitions that e.g. combine shrub/tree leaves and grass within the same class (Barnetson et al. 2017; Scarth, Roder & Schmidt 2010). Based on a review of literature (Chapter 2), a SMA approach (described in Section 2.3.2.2) was identified as a potential solution to the aforementioned constraints. SMA was thus used to estimate FVC using a LSMM, medium resolution imagery (Landsat 8 and Sentinel-2) and a limited number of field points. As expected, Sentinel-2 performed the best in the in-situ field validation, with low RMSE and MAE values of 0.4 - 2.2% compared to that of Landsat 8 (1.25 - 4.25%). The calibration of the LSMM identified key bands and spectral indices for FVC estimation using Sentinel-2, namely the second shortwave infrared band (SWIR2), normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), modified soil-adjusted vegetation index 2 (MSAVI2) and dry bare-soil index (DBSI). The inclusion of the newly developed DBSI (Rasul et al. 2018) notably improved discrimination between bare soil and dry grass, a common challenge in semi-arid, drought-stricken and degraded landscapes (Leprieur et al. 2000). The calibration process also highlighted the potential of subdividing endmember classes for improved FVC estimation. A general Bare mean spectral signature is, for example, not representative of the diverse spectral responses of different types of bare cover types. Both the Bare class and Woody class were consequently split into two subclasses (Bare 1 and Bare 2; Shrubs and Trees), which notably improved separability. The Grassy class was not subdivided for this research, however, the large variation in spectral response observed within the class suggests that subdivision based on species and condition could be useful for future estimation.

Although the iterative calibration process provided valuable insights, it was time-consuming and required a substantial amount of manual inspection (Section 4.1.2). Considerable effort went into refining mean endmember spectral signatures by identifying and removing outlier field points. Even though these outlier points were deemed “pure” visually, their spectral responses differed notably from others within the class and led to skewed mean signatures. This sensitivity of the LSMM to endmember outliers, and the consequent necessity of a lengthy calibration process, highlights one of the drawbacks of SMA for FVC estimation. Even though fewer endmember field points are required compared to e.g. a machine learning approach (Jiménez-Muñoz et al. 2009), the quality of the field points are crucial and substantial calibration is required to ensure accurate FVC estimates. This is especially true when using medium resolution imagery, as there is a higher risk of mixed cover types within a seemingly pure pixel.

The results of Chapter 4 supports the use of SMA, in particular a LSMM, for the transferable and accurate estimation of FVC using medium resolution imagery. Although many FVC products are created on a once-off basis every few years (e.g. CDLC product), FVC change quantification from

2015 to 2018 showed an increase in bare ground (4.98%) and woody components (0.415%) and decrease in herbaceous vegetation (5.395%). In addition, notable differences in spectral signatures of growing season and dry season grass were also observed, which suggests that bi-annual estimates of FVC can provide insight into seasonal changes of vegetation dynamics. This supports the need for continuous FVC data to monitor, manage and mitigate grassland degradation and conversion. The FVC estimation approach discussed in this chapter has the potential to provide updated, continuous FVCs using freely available, medium resolution imagery, thus addressing the first research question.

6.2.3 Grassland productivity modelling using FVC and primary production

Primary production can be modelled using several different approaches depending on the application, number of available input datasets and required scale of results. For the purpose of this research, the LUE-based RBM GPP model was adapted, implemented and evaluated. Adaptions included substituting the MODIS fAPAR 250 m input dataset with the improved NDVI-to-scaled fAPAR relationship (Chapter 3), as well as the MODIS VCF 250 m with a LSMM-estimated FVC (Chapter 4). Although the RBM has previously been adapted by Eisfelder et al. (2013) and Machwitz et al. (2015), the current study is the first to simultaneously improve multiple input parameters (fAPAR and FVC) and implement the model using medium resolution imagery i.e. Sentinel-2. The adapted RBM performed well based on visual inspection and validation against in-situ field data and provided accurate estimates for both semi-arid conditions (Fort Beaufort) and highly productive conditions (Cedarville). This transferability is likely attributed to the inclusion of a water stress factor in the RBM, which allowed incorporating the extreme meteorological conditions (drought vs. heavy rainfall) of both study areas into the model. Water stress is typically calculated using evapotranspiration (ET), which is a relatively complex variable to model using only RS data (Gwate et al. 2018). Many popular models, including the CGDMP product (Swinnen & Van Hoolst 2018) used for comparison in this study, do not include water stress to avoid the incorporation of additional uncertainty in model estimates. CGDMP, however, notably overestimated GPP for the Cedarville study area during the wet, growing season. In contrast, the adapted RBM performed well for the specific study area and time period, with much lower RMSE and MAE values ($2.00 \text{ g C m}^{-2} \text{ day}^{-1}$; $1.54 \text{ g C m}^{-2} \text{ day}^{-1}$) than the CGDMP ($5.48 \text{ g C m}^{-2} \text{ day}^{-1}$; $5.9 \text{ g C m}^{-2} \text{ day}^{-1}$) for the in-situ field validation. The inclusion of water stress is thus essential for robust, transferable GPP modelling of variable climatic and vegetational conditions, despite the uncertainty introduced through the use of the ET. These results support the findings of Gómez-Giráldez et al. (2018), who concluded that the incorporation of water stress and

meteorological attenuation (e.g. temperature stress) provided better GPP estimates compared to other formulations.

The subsequent combined use of FVC and RBM primary production estimates addresses one of the main constraints in literature relating to RS-based GC estimation i.e. woody components in grasslands/rangelands reflecting as areas of high production and consequently favourable GC conditions. The Grassy component of the LSMM-estimated FVC (Chapter 4) was used to scale the RBM primary production estimates (Chapter 5), thus identifying and extracting only the herbaceous, grazable primary production per pixel for the subsequent GC calculation. In the absence of GC field estimates, the model-estimated GC was compared to the most recent knowledge-based, discrete LTGC map (2016) using visual inspection, zonal statistics and frequency distribution histograms. The model-estimated GC estimates showed good visual and statistical agreement with the LTGC map, with isolated areas of underestimation and overestimation. Underestimation of model-estimated GC values compared to LTGC zones (i.e. lower GC values and better grazing conditions) was likely due to the inability of the model to include species composition and palatability in the GC calculations. This is another constraint of RS-based approaches for GC estimation (Avenant 2017) and falls outside the scope of this research. Overestimation of model-estimated GC values compared to LTGC zones (i.e. higher GC values and poorer grazing conditions) is largely due to the different natures of the estimation algorithms. The model provides pixel-based estimation of GC, thus areas of very low grazing capacities (bare patches, shrub-encroached grasslands) is reflected in the mean and majority statistics per LTGC grazing zone. In particular, high levels of variation in model-estimated GC per LTGC zone was observed for the heterogeneous Fort Beaufort landscape. This suggests that the LTGC zonal delineations are not ideal for ecosystems with complex vegetation cover dynamics and require further consideration. As both the knowledge-based LTGC map and model-estimated GC incorporate valuable information for sustainable rangeland management, the combination of the two approaches could potentially aid in adapting current LTGC zones based on climate, relief and sub-pixel vegetation cover to produce more homogenous grazing regions.

The approach discussed in this section aimed and succeeded in addressing the second research question: how can FVC aid in determining accurate grassland productivity estimates. Although a limited number of studies have explored the use of RS for grazing and browsing capacity, existing methods either do not compensate for the effects of woody components at sub-pixel level (Yu et al. 2010), or require extensive additional field sampling (Adjorlolo & Botha 2015; Espach, Lubbe & Ganzin 2009). The combination of robust and transferable FVC and primary production modelling consequently contributed to developing a novel approach to RS-based GC estimation

using little or no field data. Although the approach does not address all the constraints relating to RS-based GC calculation, this study has established a foundation for future research in this application area and has shown the unexplored potential of RS and geospatial analysis for sustainable livestock management.

6.2.4 GEE and application development for grassland management

Currently grassland and rangeland management rely primarily on field-based surveys and long-term static maps, neither of which provide easily-updateable, continuous estimates of grazing condition. There is thus a need for more frequent, continuous and dynamic information on grassland dynamics, as shown by the annual change in FVC (Section 4.2.5) and fluctuations in mean GC (Section 5.2.2.2) reported in this study. A number of grassland and rangeland management tools have been developed (GEO 2016; RCMRD 2018; USDA NRCS & DOI BLM 2018), but none provide estimates of actual vegetation production and subsequent agronomical indicators e.g. GC, stocking rate etc. Through the development, implementation and evaluation of a robust, transferable grassland productivity model, the research thus provides an approach for disseminating grassland condition and productivity dynamics at various spatial scales (Research Question 3). The practical implementation of this research in a web app (Section 5.1.5) can aid the livestock industry, local farmers and conservation ecologists by providing a range of grassland condition and productivity metrics. Annual FVC estimates provide crucial information on landscape changes at a sub-pixel level, enabling the continuous mapping of temporally variable phenomena e.g. invasive woody plant encroachment, soil erosion etc. With respect to grassland productivity, the dynamic, continuous, multi-scale modelling of GC allows comparing current, updated grazing conditions with the LTGC map to assist in adjustment of livestock grazing practices and establishment of sustainable productivity levels for livestock. The research thus has economic significance relating to food security, as the improvement of grassland productivity management will enable monitoring the degradation of grasslands and preventing potential decrease in productivity.

6.3 LIMITATIONS AND FUTURE RESEARCH RECOMMENDATIONS

An important limitation, for this research as well as vegetation modelling in general, is a lack of field data. Field data plays a crucial role in establishing relationships between ground phenomena and satellite observations (Eisfelder et al. 2017). To improve calibration of vegetation dynamics and productivity modelling, as well as to provide better validation of model results, more extensive biophysical field measurements relating to FVC, fAPAR, LAI and DM production need to be obtained. With respect to GC itself, additional field surveys are also needed for field-estimated

grazing conditions. The combination of field-based and RS-based approaches through the use of app technologies could address the general lack of sufficient field data available for vegetation modelling. The collection of field data and estimation of analysis results can be done simultaneously, aiding in the continuous calibration and validation of vegetation dynamics, condition and productivity models.

A substantial amount of time was spent on calibration of the LSMM for FVC estimation, specifically the refinement of endmember field points to ensure representative mean spectral signatures. The automation of this calibration process is thus recommended for future research. As refinement essentially involved identifying and removing outliers, statistical thresholds, proximity tests and spatial coherence (Alvera-Azcárate et al. 2012) can potentially be used to automate this step.

The adapted RBM implemented in this research uses the MODIS ET product (Running, Mu & Zhao 2017) for water stress calculation. Although the product is frequently used in various different ecosystem, carbon and hydrological models, literature suggests some drawbacks with respect to resolution and inherent uncertainty (Velpuri et al. 2013). As this research has shown the importance of water stress in GPP modelling, it is recommended that other approaches for ET modelling is investigated to further improve the accuracy and spatial resolution of RBM GPP calculations. Substitutes for other coarse resolution input variables should also be explored, namely the MODIS Land Surface Temperature (LST) product and Harmonised World Soil Database (HWSD) map, to ensure all input sources fall within the medium resolution range (i.e. 10 to 30 m). Another suggested RBM adaption is the incorporation of bi-annual FVC. Richters (2005) assumed FVC to be an abiotic, annual constant, which is disproved by the notable seasonal variability in spectral responses of the different FVC classes (Section 4.2.1). The seasonal estimation of FVC can therefore potentially improve grassland condition and productivity estimation.

An important limitation for GC estimation is the inability of RS-based approaches to incorporate species composition (Avenant 2017). There is a lack of research on classifying grassland and rangeland species using medium resolution satellite imagery, therefore the constraint is beyond the scope of this research. A potential practical solution, however, is including species composition as an input parameter within the web app to allow users to determine species composition in-field using expert opinion and personal experience. Although this solution will require the user to have expert, local knowledge of the area, it is a temporary solution until the constraint can be addressed in RS research. The incorporation of a measure of accessibility, an

indicator of how accessible forage is based on bush density (Espach, Lubbe & Ganzin 2009), is also recommended. This indicator can easily be derived from FVC estimates of woody cover.

Overall, the discussed approach shows great potential for increasing the use of RS within the livestock industry. Although this research focuses on cattle, the inclusion of a Shrub subclass in the FVC makes the calculation of production from both grass and shrubs, i.e. grazable and browsable production, feasible. Additional productivity estimates, such as the browsing capacity of a field for sheep (Adjorlolo & Botha 2015), can thus easily be implemented based on the specific needs of the farmer. The expansion and operationalisation of this research is recommended to provide decision-support for various stakeholders within the livestock industry.

6.4 CONCLUSION

Grassland degradation can have a severe negative impact on condition and productivity, and consequently on grazing potential for cattle. Current conventional methods of monitoring and managing grasslands and rangelands are time-consuming, destructive and not applicable at large-scale. Although RS-based approaches can address this constraint by monitoring large areas simultaneously, these approaches also have their technological and scientific limitations in the context of sustainable, continuous and dynamic grassland management.

The research presented in this thesis aimed to develop an FVC-integrated productivity model that estimates grassland productivity and condition at various spatial scales using Google Earth Engine, thereby aiding in the sustainable monitoring and management of grasslands and rangelands. The use of SMA is identified as a viable option for accurate, dynamic estimation of FVC using medium resolution imagery and a limited number of field points. The novel combination of SMA-derived FVC and primary production estimates addresses current constraints in RS-based GC calculation to provide improved continuous, dynamic, multi-scale grassland productivity estimates. The practical implementation of the transferable FVC-integrated grassland productivity model in a GEE web app presents a feasible addition to conventional, field-based grassland management systems.

The findings of the research provide valuable insights into improving RS-based modelling of grassland condition and productivity. The approach investigated in this research proved the potential of RS for sustainable grassland management and provides a point of departure for future work on RS-based grazing capacity modelling for the livestock industry. Operationalisation of this research can aid in identifying potential degradation, highlighting regions vulnerable to food shortages and establishing sustainable productivity levels. The research thus has both environmental and economic significance.

REFERENCES

- Adjorlolo C & Botha JO 2015. Integration of remote sensing and conventional models for modeling grazing/browsing capacity in southern African savannas. *Journal of Applied Remote Sensing* 9, 1: 096041.
- Alvera-Azcárate A, Sirjacobs D, Barth A & Beckers JM 2012. Outlier detection in satellite data using spatial coherence. *Remote Sensing of Environment* 119: 84–91.
- Andrade BO, Koch C, Boldrini II, Vélez-Martin E, Hasenack H, Hermann J-M, Kollmann J, Pillar VD & Overbeck GE 2015. Grassland degradation and restoration: A conceptual framework of stages and thresholds illustrated by southern Brazilian grasslands. *Natureza & Conservação* 13, 2: 95–104.
- Archer SR, Andersen EM, Predick KI, Schwinning S, Steidl RJ & Woods SR 2017. Woody plant encroachment: Causes and consequences. In Briske D (ed) *Rangeland systems*, 25–84. New York City: Springer.
- Asner GP, Elmore AJ, Olander LP, Martin RE & Harris AT 2004. Grazing systems, ecosystem responses, and global change. *Annual Review of Environment and Resources* 29, 1: 261–299.
- Avenant P 2017. *Long term grazing capacity norms for South Africa*. Pretoria: Department of Agriculture, Forestry and Fisheries.
- Baldocchi D, Falge E, Gu L, Olson R, Hollinger D, Running S, Anthoni P, Bernhofer C, Davis K, Evans R, Fuentes J, Goldstein A, Katul G, Law B, Lee X, Malhi Y, Meyers T, Munger W, Oechel W, Paw KT, Pilegaard K, Schmid HP, Valentini R, Verma S, Vesala T, Wilson K & Wofsy S 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society* 82, 11: 2415–2434.
- Barnetson J, Phinn S, Scarth P & Denham R 2017. Assessing Landsat fractional ground-cover time series across Australia’s arid rangelands: Separating grazing impacts from climate variability. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-3/W2: 15–26.
- Baudoin M-A, Vogel C, Nortje K & Naik M 2017. Living with drought in South Africa: Lessons learnt from the recent El Niño drought period. *International Journal of Disaster Risk Reduction* 23: 128–137.
- Bey A, Sánchez-Paus Díaz A, Maniatis D, Marchi G, Mollicone D, Ricci S, Bastin J-F, Moore R, Federici S, Rezende M, Patriarca C, Turia R, Gamoga G, Abe H, Kaidong E & Miceli G

2016. Collect Earth: Land use and land cover assessment through augmented visual interpretation. *Remote Sensing* 8, 10: 807.
- Boardman JW, Kruse FA & Green RO 1995. *Mapping target signatures via partial unmixing of AVIRIS data*. Proceedings of the fifth Annual JPL Airborne Earth Science Workshop held 23 - 26 January 1995. Pasadena: NASA Jet Propulsion Lab.
- Botha C, Burns A, Cadman M, Dini J, Fulton C, Germishuizen S, Ginsburg A, Lechmere-Oertel R, Prinsloo D, Stephens A, Walters D & Ward-Smith B 2014. *Grazing and burning guidelines managing grasslands for biodiversity and livestock production*. Pretoria: South African National Biodiversity Institute.
- Bransby DI & Tainton NM 1977. *The disc pasture meter: Possible applications in grazing management*, 12, 1: 115–118. Proceedings of the 12th Annual Congresses of the Grassland Society of Southern Africa. n.p. Taylor & Francis.
- Buchhorn M, Smets B, Bertels L, Lesiv M & Tsendbazar N-E 2019. *Moderate Dynamic Land Cover Product User Manual v2*. Publication No CGLOPS1_PUM_LC100_V2. Mol, VITO.
- Buckland ST, Borchers DL, Johnston A, Henrys PA & Marques TA 2007. Line transect methods for plant surveys. *Biometrics* 63, 4: 989–998.
- Campbell JB 2002. *Introduction to remote sensing*. 3rd ed. New York: The Guilford Press.
- Carlson TN, Capehart WJ & Gillies RR 1995. A new look at the simplified method for remote sensing of daily evapotranspiration. *Remote Sensing of Environment* 54, 2: 161–167.
- Chang C-I 2013. *Hyperspectral data processing: Algorithm design and analysis*. Hoboken: Wiley.
- Chen D & Stow D 2003. Strategies for integrating information from multiple spatial resolutions into land-use/ land-cover classification routines. *Photogrammetric Engineering & Remote Sensing* 69, 11: 1279–1287.
- Chuvieco E & Huete A 2009. *Fundamentals of satellite remote sensing*. Boca Raton: CRC Press.
- Coltin B, McMichael S, Smith T & Fong T 2016. Automatic boosted flood mapping from satellite data. *International Journal of Remote Sensing* 37, 5: 993–1015.
- Congalton RG & Green K 2009. *Assessing the accuracy of remotely sensed data: Principles and practices*. Boca Raton: CRC Press.

- Conradie DCU 2012. *South Africa's climatic zones: Today, tomorrow*. Proceedings of the International Green Building Conference and Exhibition held 25-26 July 2012, Sandton, South Africa.
- Coops NC, Hilker T, Hall FG, Nichol CJ & Drolet GG 2010. Estimation of light-use efficiency of terrestrial ecosystems from space: A status report. *BioScience* 60, 10: 788–797.
- Courault D, Seguin B & Oliosio A 2005. Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches. *Irrigation and Drainage Systems* 19, 3–4: 223–249.
- Cramer W, Kicklighter DW, Bondeau A, Iii BM, Churkina G, Nemry B, Ruimy A & Schloss AL 1999. Comparing global models of terrestrial net primary productivity (NPP): Overview and key results. *Global Change Biology* 5, S1: 1–15.
- Danckwerts JE & Trollope WSW 1980. *Assessment of the disc pasture meter on natural veld in the false thornveld of the eastern province*, 15, 1: 47–52. *Proceedings of the 15th Annual Congresses of the Grassland Society of Southern Africa*. n.p. Taylor & Francis.
- Department of Environmental Affairs 2017. *The South African Carbon Sinks Atlas 2017*. Pretoria: Department of Environmental Affairs.
- Desert Research Institute U of I 2016. Climate Engine [online]. Available from: <http://climateengine.org/> [Accessed 22 October 2018].
- Dimiceli C, Carroll M, Sohlberg R, Kim DH, Kelly M & Townshend JRG 2015. *MOD44B MODIS/Terra Vegetation Continuous Field yearly L3 global 250m SIN Grid V006*.
- Directorate Statistics and Economic Analysis 2018. *Economic Review of the South African Agriculture*. Pretoria: Department of Agriculture, Forestry and Fisheries.
- Dong J, Xiangming X, Menarguez MA, Zhang G, Yuanwei Q, Thau D, Biradar C & Moore B 2016. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sensing of Environment* 185: 142–154.
- Driver A, Sink J, Nel J, Holness S, Van Niekerk L, Daniels F, Jonas Z, Majiedt PA, Harris L & Maze K 2012. *National biodiversity assessment 2011: An assessment of South Africa's biodiversity and ecosystems*. Pretoria: South African National Biodiversity Institute and Department of Environmental Affairs.

- Eisfelder C, Klein I, Bekkuliyeva A, Kuenzer C, Buchroithner MF & Dech S 2017. Above-ground biomass estimation based on NPP time-series: A novel approach for biomass estimation in semi-arid Kazakhstan. *Ecological Indicators* 72: 13–22.
- 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.
- Eisfelder C, Kuenzer C, Dech S & Buchroithner 2013. Comparison of two remote sensing based models for regional net primary productivity estimation: A case study in semi-arid central Kazakhstan. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6, 4: 1843–1856.
- ESA 2013. *Sentinel-2 User Handbook*. Paris: European Space Agency.
- Espach C, Lubbe LG & Ganzin N 2009. Determining grazing capacity in Namibia with the aid of remote sensing. *African Journal of Range & Forage Science* 26, 3: 133–138.
- ESRI 2018. ArcGIS Desktop.
- Fajji NG, Palamuleni LG & Mlambo V 2018. A GIS scheme for forage assessment and determination of rangeland carrying capacity. *Journal of Remote Sensing & GIS* 7, 1: 233.
- FAO/IIASA/ISRIC/ISS-CAS/JRC 2012. *Harmonized World Soil Database (version 1.2)*.
- Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D & Alsdorf D 2007. The Shuttle Radar Topography Mission. *Reviews of Geophysics* 45, 2: RG2004.
- Fensholt R, Sandholt I & Rasmussen MS 2004. Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sensing of Environment* 91, 3–4: 490–507.
- Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A & Michaelsen J 2015. The climate hazards infrared precipitation with stations: A new environmental record for monitoring extremes. *Scientific Data* 2, 1: 150066.
- Gao B-C 2015. *MODIS Atmosphere L2 Water Vapor Product*.
- Gao B-C & Goetz AFH 1990. Column atmospheric water vapor and vegetation liquid water retrievals from Airborne Imaging Spectrometer data. *Journal of Geophysical Research* 95, D4: 3549.

- Gao B-C & Kaufman YJ 1998. *The MODIS Near-IR Water Vapor Algorithm Product ID: MOD05 - Total Precipitable Water*. Goddard Space Flight Center: NASA MODIS Adaptive Processing System.
- Garbulsky MF, Peñuelas J, Papale D, Ardö J, Goulden ML, Kiely G, Richardson AD, Rotenberg E, Veenendaal EM & Filella I 2010. Patterns and controls of the variability of radiation use efficiency and primary productivity across terrestrial ecosystems. *Global Ecology and Biogeography* 19, 2: 253–267.
- GEO 2016. GEOGLAM RAPP [online]. Available from: <https://www.geo-rapp.org/> [Accessed 29 October 2019].
- Gessner U, Machwitz M, Conrad C & Dech S 2013. Estimating the fractional cover of growth forms and bare surface in savannas: A multi-resolution approach based on regression tree ensembles. *Remote Sensing of Environment* 129: 90–102.
- Gitelson AA, Kaufman YJ, Stark R & Rundquist D 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment* 80, 1: 76–87.
- Gitelson AA, Viña A, Verma SB, Rundquist DC, Arkebauer TJ, Keydan G, Leavitt B, Ciganda V, Burba GG & Suyker AE 2006. Relationship between gross primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity. *Journal of Geophysical Research* 111, D8: D08S11.
- Glenn EP, Huete AR, Nagler PL & Nelson SG 2008. Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors* 8, 4: 2136–2160.
- Goetz SJ, Prince SD, Small J & Gleason ACR 2000. Interannual variability of global terrestrial primary production: Results of a model driven with satellite observation. *Journal of Geophysical Research* 105, D15: 20077–20091.
- Gómez-Giráldez PJ, Carpintero E, Ramos M, Aguilar C & González-Dugo MP 2018. *Effect of the water stress on gross primary production modeling of a Mediterranean oak savanna ecosystem*, 37–43. Proceedings of the International Association of Hydrological Sciences. Gottingen: Copernicus Publications.
- Gorelick N, Hancher M, Diixon M, Ilyushchenko S, Thau D & Moore R 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202: 18–27.

- Del Grosso SJ, Parton WJ, Derner JD, Chen M & Tucker CJ 2018. Simple models to predict grassland ecosystem C exchange and actual evapotranspiration using NDVI and environmental variables. *Agricultural and Forest Meteorology* 249: 1–10.
- Guerschman JP, Hill MJ, Renzullo LJ, Barrett DJ, Marks AS & Botha EJ 2009. Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 Hyperion and MODIS sensors. *Remote Sensing of Environment* 113, 5: 928–945.
- Guevara J., Cavagnaro J., Estevez O., Le Hou  rou H. & Stasi C. 1997. Productivity, management and development problems in the arid rangelands of the central Mendoza plains (Argentina). *Journal of Arid Environments* 35, 4: 575–600.
- Gutman G & Ignatov A 1998. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *International Journal of Remote Sensing* 19, 8: 1533–1543.
- Gwate O, Mantel SK, Palmer AR, Gibson LA & Munch Z 2018. Measuring and modelling evapotranspiration in a South African grassland: Comparison of two improved Penman-Monteith formulations. *Water SA* 44, 3: 482.
- Hanan NP, Kabat P, Dolman AJ & Elbers JA 1998. Photosynthesis and carbon balance of a Sahelian fallow savanna. *Global Change Biology* 4, 5: 523–538.
- Hansen MC, Potapov P V., Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman S V., Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO & Townshend JRG 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342, 6160: 850–853.
- Harmoney KR, Moore KJ, George JR, Brummer EC & Russell JR 1997. Determination of pasture biomass using four indirect methods. *Agronomy Journal* 89, 4: 665.
- Holben BN 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing* 7, 11: 1417–1434.
- Le Houerou HN, Bingham RL & Skerbek W 1988. Relationship between the variability of primary production and the variability of annual precipitation in world arid lands. *Journal of Arid Environments* 15, 1: 1–18.
- Hu Z, He F, Yin J, Lu X, Tang S, Wang L & Li X 2007. Estimation of fractional vegetation cover based on digital camera survey data and a remote sensing model. *Journal of China University of Mining and Technology* 17, 1: 116–120.

- Jackson RD & Huete AR 1991. Interpreting vegetation indices. *Preventive Veterinary Medicine* 11, 3–4: 185–200.
- Jia K, Liang S, Liu S, Li Y, Xiao Z, Yao Y, Jiang B, Zhao X, Wang X, Xu S & Cui J 2015. Global land surface fractional vegetation cover estimation using general regression neural networks from MODIS surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing* 53, 9: 4787–4796.
- Jiang C & Ryu Y 2016. Multi-scale evaluation of global gross primary productivity and evapotranspiration products derived from Breathing Earth System Simulator (BESS). *Remote Sensing of Environment* 186: 528–547.
- Jiménez-Muñoz J, Sobrino J, Plaza A, Guanter L, Moreno J & Martínez P 2009. Comparison between fractional vegetation cover retrievals from vegetation indices and spectral mixture analysis: case study of PROBA/CHRIS data over an agricultural area. *Sensors* 9, 12: 768–793.
- Jones, J. B. J & Case VW 1990. Sampling, handling and analyzing plant tissue samples. In *Soil testing and plant analysis*, 389–427. Madison: Soil Science Society of America Inc.
- Jones MO, Allred BW, Naugle DE, Maestas JD, Donnelly P, Metz LJ, Karl J, Smith R, Bestelmeyer B, Boyd C, Kerby JD & McIver JD 2018. Innovation in rangeland monitoring: Annual, 30 m, plant functional type percent cover maps for U.S. rangelands, 1984–2017. *Ecosphere* 9, 9.
- Joubert AJ & Myburgh WJ 2014. A comparison of three dry matter forage production methods used in South Africa. *International Journal of Ecology* 2014: 1–6.
- Jung M, Reichstein M, Margolis HA, Cescatti A, Richardson AD, Arain MA, Arneth A, Bernhofer C, Bonal D, Chen J, Gianelle D, Gobron N, Kiely G, Kutsch W, Lasslop G, Law BE, Lindroth A, Merbold L, Montagnani L, Moors EJ, Papale D, Sottocornola M, Vaccari F & Williams C 2011. Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research* 116, G3: G00J07.
- Jung M, Reichstein M, Schwalm CR, Huntingford C, Sitch S, Ahlström A, Arneth A, Camps-Valls G, Ciais P, Friedlingstein P, Gans F, Ichii K, Jain AK, Kato E, Papale D, Poulter B, Raduly B, Rödenbeck C, Tramontana G, Viovy N, Wang Y-P, Weber U, Zaehle S & Zeng N 2017. Compensatory water effects link yearly global land CO₂ sink changes to temperature. *Nature* 541, 7638: 516–520.

- King DA, Turner DP & Ritts WD 2011. Parameterization of a diagnostic carbon cycle model for continental scale application. *Remote Sensing of Environment* 115, 7: 1653–1664.
- Knorr W & Heimann M 2001. Uncertainties in global terrestrial biosphere modeling: 1. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. *Global Biogeochemical Cycles* 15, 1: 207–225.
- KoBoToolBox 2012. KoBoToolbox: Data collection tools for challenging environments [online]. Cambridge: Harvard Humanitarian Initiative. Available from: <https://www.kobotoolbox.org/> [Accessed 11 August 2019].
- Kwon H-Y, Nkonya E, Johnson T, Graw V, Kato E & Kihui E 2016. Global estimates of the impacts of grassland degradation on livestock productivity from 2001 to 2011. In Nkonya E, Mirzabaev A & Von Braun J (eds) *Economics of land degradation and improvement : A global assessment for sustainable development*, 197–214. Cham: Springer.
- Land Resources International 2018. *SSC WC 03 (2017/2018) National Land Cover Report*. Pretoria: Department of Rural Development.
- Leprieur C, Kerr YH, Mastorchio S & Meunier JC 2000. Monitoring vegetation cover across semi-arid regions: Comparison of remote observations from various scales. *International Journal of Remote Sensing* 21, 2: 281–300.
- Li F, Chen W, Zeng Y, Zhao Q & Wu B 2014. Improving estimates of grassland fractional vegetation cover based on a pixel dichotomy model: A case study in Inner Mongolia, China. *Remote Sensing* 6, 12: 4705–4722.
- Li F, Wang X, Zhao J, Zhang X & Zhao Q 2013. A method for estimating the gross primary production of alpine meadows using MODIS and climate data in China. *International Journal of Remote Sensing* 34, 23: 8280–8300.
- Liang S, Li X & Wang J (eds) 2012. *Advanced remote sensing*. Waltham: Academic Press.
- Lieth H 1973. Primary production: Terrestrial ecosystems. *Human Ecology* 1, 4: 303–332.
- Little IT, Hockey PAR & Jansen R 2015. Impacts of fire and grazing management on South Africa's moist highland grasslands: A case study of the Steenkampsberg Plateau, Mpumalanga, South Africa. *Bothalia - African Biodiversity & Conservation* 45, 1: 1–15.
- Liu D, Yang L, Jia K, Liang S, Xiao Z, Wei X, Yao Y, Xia M & Li Y 2018. Global fractional vegetation cover estimation algorithm for VIIRS reflectance data based on machine learning methods. 10: 1648.

- Liu HQ & Huete A 1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing* 33, 2: 457–465.
- Liu J, Chen JM, Cihlar J & Park WM 1997. A process-based boreal ecosystem productivity simulator using remote sensing inputs. *Remote Sensing of Environment* 62, 2: 158–175.
- Liu S, Zhuang Q, He Y, Noormets A, Chen J & Gu L 2016. Evaluating atmospheric CO₂ effects on gross primary productivity and net ecosystem exchanges of terrestrial ecosystems in the conterminous United States using the AmeriFlux data and an artificial neural network approach. *Agricultural and Forest Meteorology* 220: 38–49.
- Liu X 2008. Airborne LiDAR for DEM generation: Some critical issues. *Progress in Physical Geography* 32, 1: 31–49.
- Liu Z, Wang L & Wang S 2014. Comparison of Different GPP Models in China Using MODIS Image and ChinaFLUX Data. *Remote Sensing* 6, 10: 10215–10231.
- Lobell DB, Thau D, Seifert C, Engle E & Little B 2015. A scalable satellite-based crop yield mapper. *Remote Sensing of Environment* 164: 324–333.
- Lück W, Mhangara P, Kleyn L & Remas H 2010. *Land Cover Field Guide v2*. Stellenbosch: Centre for Geographical Analysis.
- Machwitz M, Gessner U, Conrad C, Falk U, Richters J & Dech S 2015. Modelling the Gross Primary Productivity of West Africa with the Regional Biomass Model RBM+, using optimized 250 m MODIS FPAR and fractional vegetation cover information. *International Journal of Applied Earth Observation and Geoinformation* 43: 177–194.
- Le Maitre DC, Forsyth GG, Dziki S & Gush MB 2016. Estimates of the impacts of invasive alien plants on water flows in South Africa. *Water SA*.
- Marsett RC, Qi J, Heilman P, Biedenbender SH, Carolyn Watson M, Amer S, Weltz M, Goodrich D & Marsett R 2006. Remote sensing for grassland management in the arid southwest. *Rangeland Ecology & Management* 59, 5: 530–540.
- Meissner H, Scholtz M & Palmer A 2014. Sustainability of the South African livestock sector towards 2050 part 1: Worth and impact of the sector. *South African Journal of Animal Science* 43, 3: 282–297.
- Melillo JM, McGuire AD, Kicklighter DW, Moore B, Vorosmarty CJ & Schloss AL 1993. Global climate change and terrestrial net primary production. *Nature* 363, 6426: 234–240.

- METER Inc. 2018. *AccuPAR PAR/LAI Ceptometer Model LP-80 Operator's Manual*. Pullman: METER Inc.
- MOL 2016. Map of Life [online]. Available from: <https://mol.org/> [Accessed 22 October 2018].
- Monteith JL 1972. Solar radiation and productivity in tropical ecosystems. *Journal of Applied Ecology* 9, 3: 747–766.
- Mucina L & Rutherford MC 2006. *The vegetation of South Africa, Lesotho and Swaziland*. 1: 1 000 000 scale sheet maps. Pretoria: South African National Biodiversity Institute.
- Myeni R, Knyazikhin Y & Park T 2015. *MOD15A2H MODIS Leaf Area Index/FPAR 8-day L4 global 500m SIN grid V006*.
- Naidoo L, Mathieu R, Main R, Wessels K & Asner GP 2016. L-band Synthetic Aperture Radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs. *International Journal of Applied Earth Observation and Geoinformation* 52: 54–64.
- Navarro JA 2017. *First experiences with Google Earth Engine*, 250–255. Proceedings of the 3rd International Conference on Geographical Information Systems Theory, Applications and Management held 27-28 April 2017, Porto, Portugal. Setubal: SCITEPRESS.
- Niklaus M, Tum M & Günther P 2010. *Modeling carbon sinks and sources in semi-arid environments for a land degradation assessment approach*, 648–656. Proceedings of the 24th International Conference of Informatics for Environmental Protection held 6-8 October 2010, Cologne/Bonn, Germany. Bonn: EnviroInfo.
- Niklaus M, Tum M & Günther KP *Estimating the carbon cycle of South Africa with BETHY/DLR*. Proceedings of the EGU General Assembly held 2-7 May 2010, Vienna, Austria. Gottingen: Copernicus.
- Padarian J, Minasny B & McBratney AB 2015. Using Google's cloud-based platform for digital soil mapping. *Computers & Geosciences* 83: 80–88.
- Palmer AR, Finca A, Mantel SK, Gwate O, Münch Z & Gibson LA 2017. Determining fPAR and leaf area index of several land cover classes in the Pot River and Tsitsa River catchments of the Eastern Cape, South Africa. *African Journal of Range & Forage Science* 34, 1: 33–37.
- Palmer AR, Samuels I, Cupido C, Finca A, Kangombe WF, Yunusa IA, Vetter S & Mapaure I 2016. Aboveground biomass production of a semi-arid southern African savanna: Towards a new model. *African Journal of Range & Forage Science* 33, 1: 43–51.

- Pekel J-F, Cottam A, Gorelick N & Belward AS 2016. High-resolution mapping of global surface water and its long-term changes. *Nature* 540, 7633: 418–422.
- Pickett-Heaps CA, Canadell JG, Briggs PR, Gobron N, Haverd V, Paget MJ, Pinty B & Raupach MR 2014. Evaluation of six satellite-derived Fraction of Absorbed Photosynthetic Active Radiation (FAPAR) products across the Australian continent. *Remote Sensing of Environment* 140: 241–256.
- Plaza A, Martinez P, Perez R & Plaza J 2004. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* 42, 3: 650–663.
- Plaza A & Chang C-I 2006. Impact of initialization on design of endmember extraction algorithms. *IEEE Transactions on Geoscience and Remote Sensing* 44, 11: 3397.
- Potter CS, Randerson JT, Field CB, Matson PA, Vitousek PM, Mooney HA & Klooster SA 1993. Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles* 7, 4: 811–841.
- Price JC 1994. How unique are spectral signatures? *Remote Sensing of Environment* 49, 3: 181–186.
- Priestley CHB & Taylor RJ 1972. On the assessment of surface heat flux and evaporation Using large-scale parameters. *Monthly Weather Review* 100, 2: 81–92.
- Qi J, Chehbouni A, Huete AR, Kerr YH & Sorooshian S 1994. A modified soil adjusted vegetation index. *Remote Sensing of Environment* 48, 2: 119–126.
- Qi J, Kerr YH & Chehbouni A 1994. *External factor consideration in vegetation index development*, 723–730. Proceedings of 6th International Symposium on Physical Measurements and Signatures in Remote Sensing held 17-22 January 2014, Val d'Iserre, France. Pasadena: NASA Jet Propulsion Lab.
- Rahman MM & Lamb DW 2017. The role of directional LAI in determining the fAPAR–NDVI relationship when using active optical sensors in tall fescue (*Festuca arundinacea*) pasture. *International Journal of Remote Sensing* 38, 11: 3219–3235.
- Rasul A, Balzter H, Ibrahim G, Hameed H, Wheeler J, Adamu B, Ibrahim S & Najmaddin P 2018. Applying built-up and bare-soil indices from Landsat 8 to cities in dry climates. *Land* 7, 3: 81.
- Raufirad V, Azadi H, Ebrahimi A & Bagheri S 2016. Determining rangeland species palatability: Application of principal component analysis. *Rangelands* 38, 3: 105–112.

- RCMRD 2018. Rangelands Decision Support Tool [online]. Available from: <http://tools.rcmrd.org/> [Accessed 29 October 2019].
- Reynolds SG & Frame J 2005. *Grasslands: Developments, opportunities, perspectives*. Rome: Food and Agricultural Organization of the United Nations.
- Richters J 2005a. Biomass changes on north-western Namibia: First results for a remote sensing modelling approach. *EARSeL eProceedings* 4: 157–170.
- Richters J 2005b. Entwicklung eines fernerkundungsgestützten Modells zur Erfassung von pflanzlicher Biomasse in NW-Namibia. Bonn: Rheinischen Friedrich-Wilhelms-Universität Bonn, Geographischen Institut.
- Richters J 2006. *Biomass production, pasture balance, and their ecologic consequences in NW Namibia*. Proceedings of the SPIE Optics and Photonics Conference held 13-17 August 2006, San Diego, USA. Bellingham: International Society for Optics and Photonics.
- Rossini M, Migliavacca M, Galvagno M, Meroni M, Cogliati S, Cremonese E, Fava F, Gitelson A, Julitta T, Morra Di Cella U, Siniscalco C & Colombo R 2014. Remote estimation of grassland gross primary production during extreme meteorological seasons. *International Journal of Applied Earth Observations and Geoinformation* 29: 1–10.
- Running S, Mu Q & Zhao M 2015. *MOD17A2H MODIS/Terra Gross Primary Productivity 8-Day L4 Global 500m SIN Grid V006*.
- Running S, Mu Q & Zhao M 2017. *MOD16A2 MODIS/Terra Net Evapotranspiration 8-Day L4 Global 500m SIN Grid V006*.
- Running S, Nemani RR, Heinsch FA, Zhao M, Reeves M & Hashimoto H 2004. A continuous satellite-derived measure of global terrestrial primary production. *BioScience* 54, 6: 547–560.
- Running S, Thornton PE, Nemani R & Glassy JM 2000. Global terrestrial gross and net primary productivity from the earth observing system. In *Methods in Ecosystem Science*, 44–57. New York: Springer.
- Ryu Y, Baldocchi DD, Kobayashi H, van Ingen C, Li J, Black TA, Beringer J, van Gorsel E, Knohl A, Law BE & Rouspard O 2011. Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales. *Global Biogeochemical Cycles* 25, 4: GB4017.
- Sabol DE, Gillespie AR, Adams JB, Smith MO & Tucker CJ 2002. Structural stage in Pacific Northwest forests estimated using simple mixing models of multispectral images. *Remote Sensing of Environment* 80, 1: 1–16.

- Sakowska K, Juszczak R & Gianelle D 2016. Remote sensing of grassland biophysical parameters in the context of the Sentinel-2 satellite mission. *Journal of Sensors* 2016: 1–16.
- Sankaran M, Hanan NP, Scholes RJ, Ratnam J, Augustine DJ, Cade BS, Gignoux J, Higgins SI, Le Roux X, Ludwig F, Ardo J, Banyikwa F, Bronn A, Bucini G, Caylor KK, Coughenour MB, Diouf A, Ekaya W, Feral CJ, February EC, Frost PGH, Hiernaux P, Hrabar H, Metzger KL, Prins HHT, Ringrose S, Sea W, Tews J, Worden J & Zambatis N 2005. Determinants of woody cover in African savannas. *Nature* 438, 7069: 846–849.
- Scanlon TM, Albertson JD, Caylor KK & Williams CA 2002. Determining land surface fractional cover from NDVI and rainfall time series for a savanna ecosystem. *Remote Sensing of Environment* 82, 2–3: 376–388.
- Scarth P, Roder A & Schmidt M 2010. *Tracking grazing pressure and climate interaction: The role of Landsat fractional cover in time series analysis*. Proceedings of the 15th Australasian Remote Sensing and Photogrammetry Conference (ARSPC) held 13-17 September 2010, Alice Springs, Australia.
- Schaaf C & Wang Z 2015. *MCD43A1 MODIS/Terra+Aqua BRDF/Albedo Model Parameters Daily L3 Global - 500m V006*.
- Schulze RE & Lynch SD 2007. *South African atlas of climatology and agrohydrology*.
- Seaquist J., Olsson L & Ardö J 2003. A remote sensing-based primary production model for grassland biomes. *Ecological Modelling* 169, 1: 131–155.
- Sellers PJ 1985. Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing* 6, 8: 1335–1372.
- Sharrow SH 1984. A simple disc meter for measurement of pasture height and forage bulk. *Journal of Range Management* 37, 1: 94.
- Shiferaw H, Bewket W & Eckert S 2019. Performances of machine learning algorithms for mapping fractional cover of an invasive plant species in a dryland ecosystem. *Ecology and Evolution* 9, 5: 2562–2574.
- Shroder JF, Sivanpillai R, Angerer JP, Fox WE & Wolfe JE 2016. Land degradation in rangeland ecosystems. *Biological and Environmental Hazards, Risks, and Disasters*: 277–311.
- Siegmund A & Menz G 2005. Fernes nah gebracht - satelliten- und luftbildeinsatz zur analyse von umweltveränderungen im geographieunterricht. *Geographie und Schule* 27, 154: 2–10.
- Sims DA, Rahman AF, Cordova VD, El-Masri BZ, Baldocchi DD, Bolstad P V., Flanagan LB, Goldstein AH, Hollinger DY, Misson L, Monson RK, Oechel WC, Schmid HP, Wofsy SC &

- Xu L 2008. A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS. *Remote Sensing of Environment* 112, 4: 1633–1646.
- Soulard C, Albano C, Villarreal M & Walker J 2016. Continuous 1985–2012 Landsat monitoring to assess fire effects on meadows in Yosemite National Park, California. *Remote Sensing* 8, 5: 371.
- Strahler AH, Lucht W, Schaaf CB, Tsang T, Gao F, Li X, Muller J-P, Lewis P, Barnsley MJ, Strugnell N, Hu B, Hyman A, D'entremont RP, Chen L, Liu Y, Mciver D, Liang S, Disney M, Hobson P, Dunderdale M & Roberts G 1999. *MODIS BRDF/Albedo Product ID: MOD43*.
- Sturrock HJ, Cohen JM, Keil P, Tatem AJ, Le Menach A, Ntshalintshali NE, Hsiang MS & Gosling RD 2014. Fine-scale malaria risk mapping from routine aggregated case data. *Malaria Journal* 13, 1: 421.
- Sun Z, Wang X, Zhang X, Tani H, Guo E, Yin S & Zhang T 2019. Evaluating and comparing remote sensing terrestrial GPP models for their response to climate variability and CO₂ trends. *Science of The Total Environment* 668: 696–713.
- Suttie J, Reynolds S & Batello C 2005. *Grasslands of the world*. Rome: Food and Agriculture Organization of the United Nations.
- Swift LW 1976. Algorithm for solar radiation on mountain slopes. *Water Resources Research* 12, 1: 108–112.
- Swinnen E & Van Hoolst R 2018. *Algorithm Theoretical Basis Document: Dry Matter Production (DMP) Collection 300 m Version 1*. Document No CGLOPS1_ATBD_DMP300m-V1. Mol: VITO.
- Tian Y, Woodcock C, Wang Y, Privette J, Shabanov N, Zhou L, Zhang Y, Buermann W, Dong J & Veikkanen B 2002. Multiscale analysis and validation of the MODIS LAI product I: Uncertainty assessment. *Remote Sensing of Environment* 83, 3: 414–430.
- Townshend J, Hansen M, Carroll M, DiMiceli C, Sohlberg R & Huang C 2017. *User guide for MODIS Vegetation Continuous Fields product collection 6, version 1*. Sioux Falls: NASA EOSDIS Land Processes DAAC
- Tramontana G, Jung M, Schwalm CR, Ichii K, Camps-Valls G, Ráduly B, Reichstein M, Arain MA, Cescatti A, Kiely G, Merbold L, Serrano-Ortiz P, Sickert S, Wolf S & Papale D 2016. Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences* 13, 14: 4291–4313.

- Tsenbazar N, Herold M, Lesiv M & Fritz S 2018. *Moderate Dynamic Land Cover Validation Report v1*. Document No CGLOPS1_QAR_LC100m_V1. Mol: VITO.
- Tucker CJ 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8, 2: 127–150.
- USDA NRCS & DOI BLM 2018. Rangeland Analysis Platform [online]. Available from: <https://rangelands.app/> [Accessed 29 October 2019].
- USGS 2016. *Landsat 8 data users handbook*. 2nd ed. Sioux Falls: EROS.
- USGS 2015. Comparison of Sentinel-2 and Landsat [online]. Sioux Falls: USGS. Available from: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-sentinel-2-comparison-sentinel-2-and-landsat?qt-science_center_objects=0#qt-science_center_objects [Accessed 20 August 2019].
- Valentini R (ed) 2003. *Fluxes of carbon, water and energy of European forests*. Berlin: Springer.
- Velpuri NM, Senay GB, Singh RK, Bohms S & Verdin JP 2013. A comprehensive evaluation of two MODIS evapotranspiration products over the conterminous United States: Using point and gridded FLUXNET and water balance ET. *Remote Sensing of Environment* 139: 35–49.
- Veroustraete F, Sabbe H & Eerens H 2002. Estimation of carbon mass fluxes over Europe using the C-Fix model and Euroflux data. *Remote Sensing of Environment* 83, 3: 376–399.
- Verstraeten WW, Veroustraete F & Feyen J 2006. On temperature and water limitation of net ecosystem productivity: Implementation in the C-Fix model. *Ecological Modelling* 199, 1: 4–22.
- Vickery PJ 1972. Grazing and net primary production of a temperate grassland. *The Journal of Applied Ecology* 9, 1: 307.
- Wan Z, Hook G & Hulley G 2015. *MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006*.
- Wang Y, Xie D, Liu S, Hu R, Li Y, Yan G, Li X, Liou Y-A, Liu Q, Atzberger C & Thenkabail PS 2016. Scaling of FAPAR from the field to the satellite. *Remote Sensing* 8, 310.
- Wang Z, Deng X, Song W, Li Z & Chen J 2017. What is the main cause of grassland degradation? A case study of grassland ecosystem service in the middle-south Inner Mongolia. *CATENA* 150: 100–107.
- Wei S, Yi C, Fang W & Hendrey G 2017. A global study of GPP focusing on light-use efficiency in a random forest regression model. *Ecosphere* 8, 5: e01724.

- Wißkirchen K, Tum M, Günther KP, Niklaus M, Eisfelder C & Knorr W 2013. Quantifying the carbon uptake by vegetation for Europe on a 1km-squared resolution using a remote sensing driven vegetation model. *Geoscientific Model Development* 6, 5: 1623–1640.
- Wu C, Munger JW, Niu Z & Kuang D 2010. Comparison of multiple models for estimating gross primary production using MODIS and eddy covariance data in Harvard Forest. *Remote Sensing of Environment* 114, 12: 2925–2939.
- Xiao X, Zhang Q, Saleska S, Hutyyra L, De Camargo P, Wofsy S, Frohking S, Boles S, Keller M & Moore B 2005. Satellite-based modeling of gross primary production in a seasonally moist tropical evergreen forest. *Remote Sensing of Environment* 94, 1: 105–122.
- Yang L, Jia K, Liang S, Liu J & Wang X 2016. Comparison of four machine learning methods for generating the GLASS fractional vegetation cover product from MODIS Data. *Remote Sensing* 8, 8: 682.
- Yu L, Zhou L, Liu W & Zhou HK 2010. Using remote sensing and GIS technologies to estimate grass yield and livestock carrying capacity of Alpine grasslands in Golog Prefecture, China. *Pedosphere* 20, 3: 342–351.
- Yuan W, Liu S, Zhou G, Zhou G, Tieszen LL, Baldocchi D, Bernhofer C, Gholz H, Goldstein AH, Goulden ML, Hollinger DY, Hu Y, Law BE, Stoy PC, Vesala T & Wofsy SC 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agricultural and Forest Meteorology* 143, 3–4: 189–207.
- Zambatis N, Zacharias P, Morris C & Derry J 2006. Re-evaluation of the disc pasture meter calibration for the Kruger National Park, South Africa. *African Journal of Range & Forage Science* 23, 2: 85–97.
- Zambatis N 2003. Determinants of grass production and composition in the Kruger National Park. University of KwaZulu-Natal.
- Zhang M, Lal R, Zhao Y, Jiang W & Chen Q 2017. Spatial and temporal variability in the net primary production of grassland in China and its relation to climate factors. *Plant Ecology* 218, 9: 1117–1133.
- Zhang Q, Li B, Thau D & Moore R 2015. Building a better urban picture: Combining day and night remote sensing imagery. *Remote Sensing* 7, 9: 11887–11913.

- Zhang S, Zhang J, Bai Y, Koju UA, Igbawua T, Chang Q, Zhang D & Yao F 2018. Evaluation and improvement of the daily boreal ecosystem productivity simulator in simulating gross primary productivity at 41 flux sites across Europe. *Ecological Modelling* 368: 205–232.
- Zhang X, Liao C, Li J & Sun Q 2012. Fractional vegetation cover estimation in arid and semi-arid environments using: HJ-1 satellite hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation* 21, 1: 506–512.
- Zhao M, Heinsch FA, Nemani RR & Running SW 2005. Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment* 95, 2: 164–176.
- Zhu X, Pei Y, Zheng Z, Dong J, Zhang Y, Wang J, Chen L, Doughty RB, Zhang G & Xiao X 2018. Underestimates of grassland Gross Primary Production in MODIS standard products. *Remote Sensing* 10: 1771.
- Zimmermann H-J 2001. *Fuzzy set theory and its applications*. Dordrecht: Springer Netherlands.
- Zobler L 1986. *A world soil file for global climate modeling*. Goddard Space Flight Center: NASA.

APPENDICES

Appendix A	PAR calculation as defined by Swift (1976)	134
Appendix B	Long Term Grazing Capacity Map for South Africa (2016)	136
Appendix C	Regression relationship between NDVI and fAPAR	137
Appendix D	Image acquisition dates of multi-temporal stacks for FVC estimation	138
Appendix E	Grazing capacity frequency histograms for Fort Beaufort and Cedarville	139
Appendix F	GrazeEngine example user interface screenshots	142

APPENDIX A

PAR calculation as defined by Swift (1976)

Figure A.1 shows the model diagram for the PAR calculation following Swift methodology, whereas Equation A.1 to Equation A.8 show the formulas used for the calculation of the different model components.

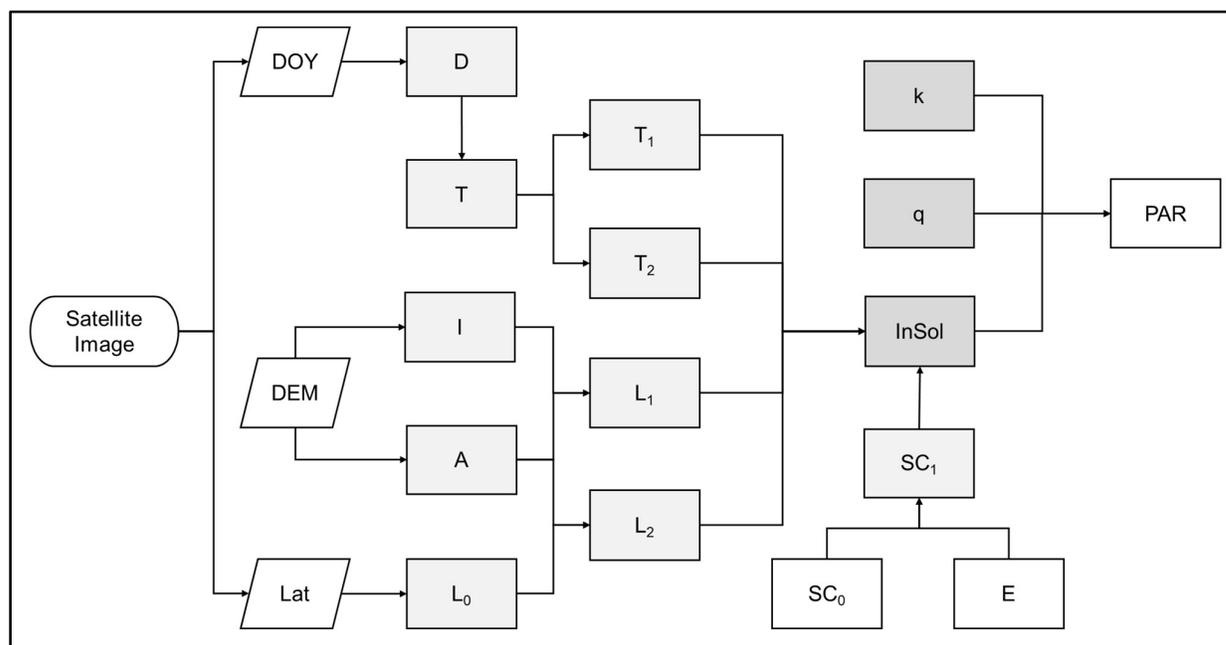


Figure A.1 PAR model diagram. A, aspect; D, declination of the Sun; DEM, digital elevation model; DOY, day of year; E, radius vector of the Sun; I, slope; InSol, potential insolation; k, proportion of photosynthetically usable radiation; Lat, latitude; L_0 , latitude of current slope; L_1 , latitude of the equivalent slope; L_2 , hour angle between the current slope and the equivalent slope; q, cloud transmission factor; SC_0 , solar constant; SC_1 , solar constant for 60 min.; T, hour angle; T_1 , hour angle of sunrise on the slope; T_2 , hour angle of sunset on the slope.

$$PAR = InSol \cdot q \cdot k \quad \text{Equation A.1}$$

$$InSol = SC_1 \cdot \left(\frac{\sin(D) \det \cdot \sin(L_1) \cdot (T_2 - T_1)}{15 + \cos(D) \cdot \cos(L_1) \cdot \left(\sin(T_2 + L_2) - \sin(T_1 + L_2) \right) \cdot \frac{12}{\pi}} \right) \quad \text{Equation A.2}$$

$$L_1 = \sin^{-1}(\cos(I) \cdot \sin(L_0) + \sin(I) \cdot \cos(L_0) \cdot \cos(A)) \quad \text{Equation A.3}$$

$$L_2 = \tan^{-1} \left(\frac{\sin(I) \cdot \sin(A)}{\cos(I) \cdot \cos(L_0) - \sin(I) \cdot \sin(L_0) \cdot \cos(A)} \right) \quad \text{Equation A.4}$$

$$D = \sin^{-1}(0.39785 \cdot \sin((278.9709 + 0.9856 \cdot DOY + 1.9163) \cdot \sin(356.6153 + 0.9856 \cdot DOY))) \quad \text{Equation A.5}$$

$$T = \cos^{-1}(-\tan(L_0) \cdot \tan(D)) \quad \text{Equation A.6}$$

$$E = 1.0 - 0.0167 \cdot \cos((DOY - 3) \cdot 0.0172) \quad \text{Equation A.7}$$

$$SC_1 = \frac{60 \cdot SC_0}{(E \cdot E)}$$

Equation A.8

where	<i>A</i>	is the aspect [°];
	<i>D</i>	is the declination of the Sun;
	<i>DOY</i>	is the Julian day of the year;
	<i>E</i>	is the radius vector of the Sun;
	<i>I</i>	is the slope [°];
	<i>InSol</i>	is the potential insolation [$\text{cal cm}^{-2}\text{day}^{-1}$]
	<i>k</i>	is the proportion of photosynthetically usable radiation;
	<i>L₀</i>	Is the latitude of current slope;
	<i>L₁</i>	is the latitude of equivalent slope;
	<i>L₂</i>	is the hour angle between the current and the equivalent slope;
	<i>PAR</i>	is the photosynthetically active radiation [$\text{cal cm}^{-2}\text{day}^{-1}$]
	<i>q</i>	is the cloud transmission factor;
	<i>SC₀</i>	is the solar constant [$\text{cal cm}^{-2}\text{min}^{-1}$];
	<i>SC₁</i>	is the solar constant for 60 min.;
	<i>T₁</i>	is the hour angle of sunrise on the slope; and
	<i>T₂</i>	is the hour angle of sunset on the slope.

APPENDIX B

Long Term Grazing Capacity Map for South Africa (2016)

Figure B.1 shows the Long Term Grazing Capacity Map for South Africa, used for comparative purposes in Chapter 5.

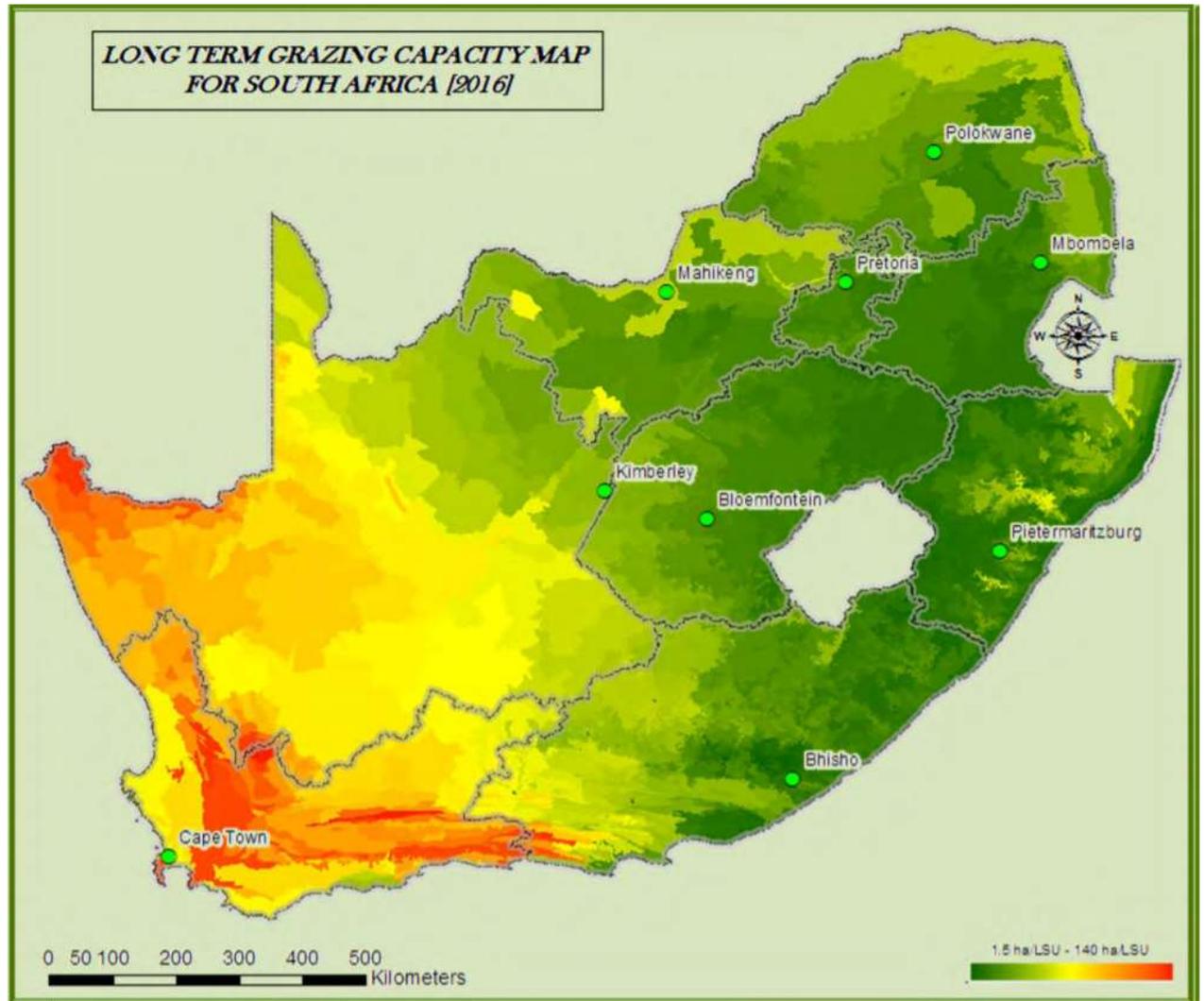


Figure B.1 Long Term Grazing Capacity Map for South Africa (2016)

APPENDIX C

Regression relationship between NDVI and fAPAR

Figure C.1 shows regression relationships between NDVI and unscaled fAPAR (blue), fAPAR scaled using visual inspection (orange) and fAPAR scaled using DM mass (grey) (Section 3.1.2.3). Regressions are shown for the specified dates, starting from the beginning of the month (Figure C.1a) to the end of the month (Figure C.1h). These results are discussed in Section 3.2.2.2.

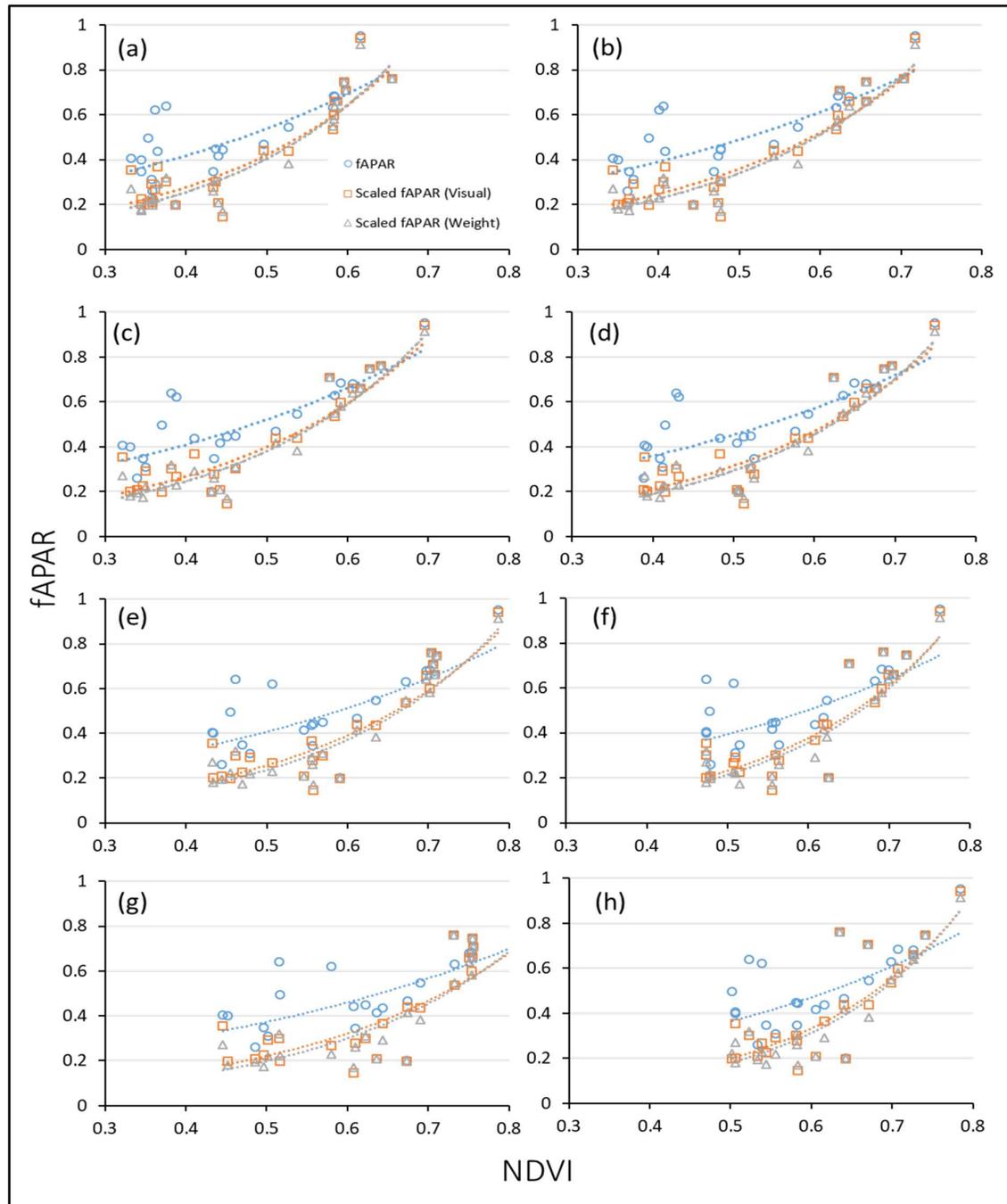


Figure C.1 Regression analysis for NDVI and unscaled fAPAR (blue), visually scaled fAPAR (orange) and DM mass scaled fAPAR (grey) for (a) 2019/02/01, (b) 2019/02/06, (c) 2019/02/08, (d) 2019/02/13, (e) 2019/02/18, (f) 2019/02/23, (g) 2019/02/26 and (h) 2019/02/28.

APPENDIX D**Image acquisition dates of multi-temporal stacks for FVC estimation**

Table D.1 Image acquisition dates for FVC multi-temporal stacks used for model comparison, calibration, validation and transferability assessment

Month	Comparison	Calibration		Validation		Transferability
	(Section 4.1.3.1)	(Section 4.1.2.5)		(Section 4.1.3.2)		(Section 4.1.4)
	2015	2018		2018		2019
	Landsat 8	Landsat 8	Sentinel-2	Landsat 8	Sentinel-2	Sentinel-2
January	3, 10, 19	2, 11, 18, 27	2, 5, 7, 10, 12, 15, 17	-	-	-
February	4, 11, 20, 27	3, 12, 19, 28	4, 6, 9, 26	-	-	6, 8, 18, 23, 28
March	8, 24, 31	7, 16, 23,	13, 26, 28	-	-	-
April	9, 16, 18, 25	1, 8, 17, 24,	7, 12, 15, 22, 30	-	-	-
May	2, 11, 18	3, 10, 19, 26	5, 10, 17, 27, 30	-	-	-
June	3, 12, 19, 28	4, 20, 27	1, 11, 14, 16, 21	4, 11, 20, 27	1, 11, 14, 16, 21	-
July	5, 14, 21, 30	6, 13, 22, 29	4, 6, 9, 14, 16, 19, 21, 31	-	-	-
August	6, 15, 31	7, 14, 23, 30	3, 18, 23, 25, 28, 30	-	-	-
September	7, 16, 23	8, 15, 24,	12, 14	-	-	-
October	2, 9, 18	1, 10, 17, 26	4, 7, 9, 19, 22, 24, 27, 29	-	-	-
November	3, 10, 19, 26	2, 11, 18, 27	1, 3, 8, 11, 16, 23	-	-	-
December	5, 12, 21	4, 13, 20, 29	1, 11, 21, 23, 26	-	-	-

APPENDIX E

Grazing capacity frequency histograms for Fort Beaufort and Cedarville

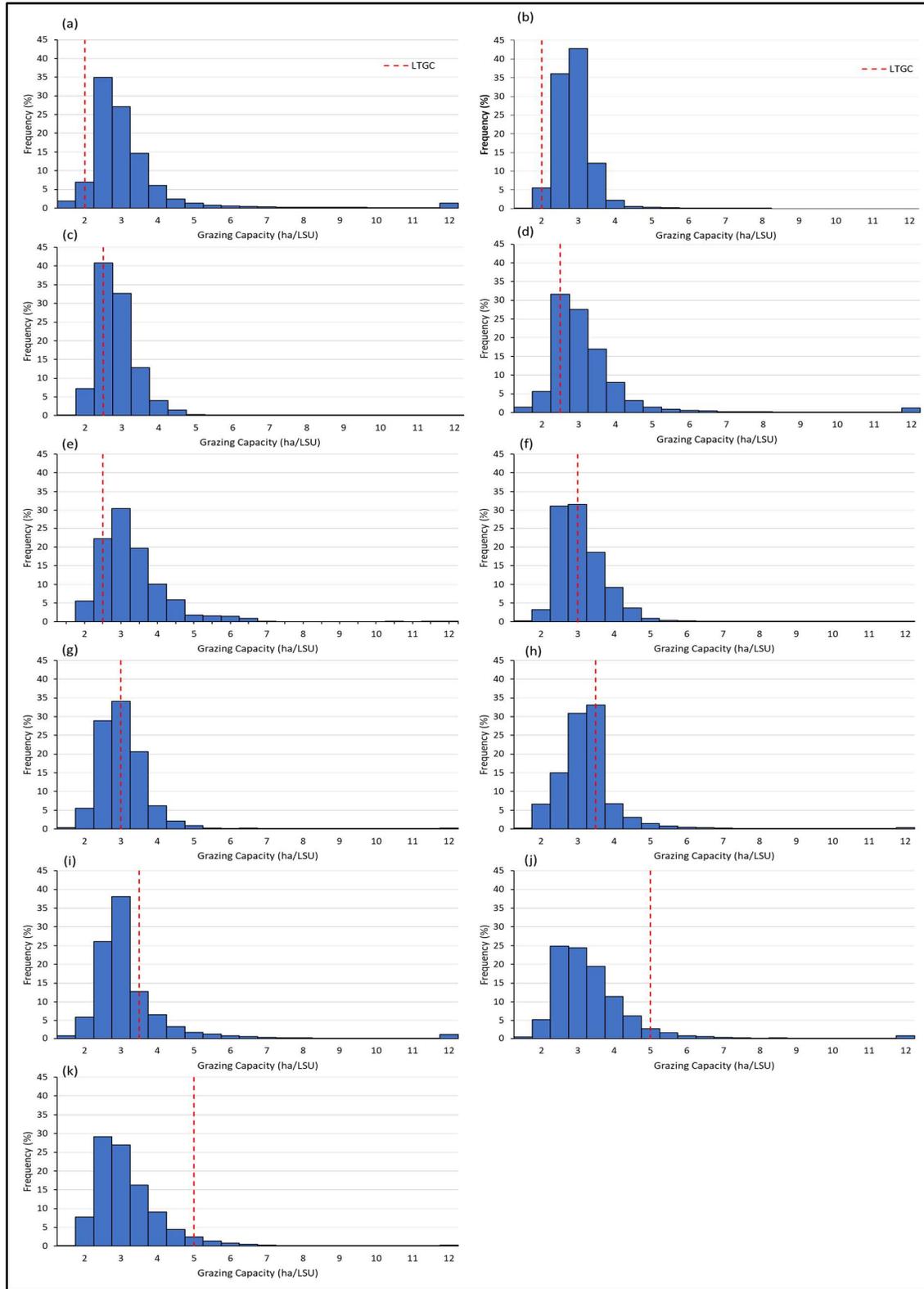
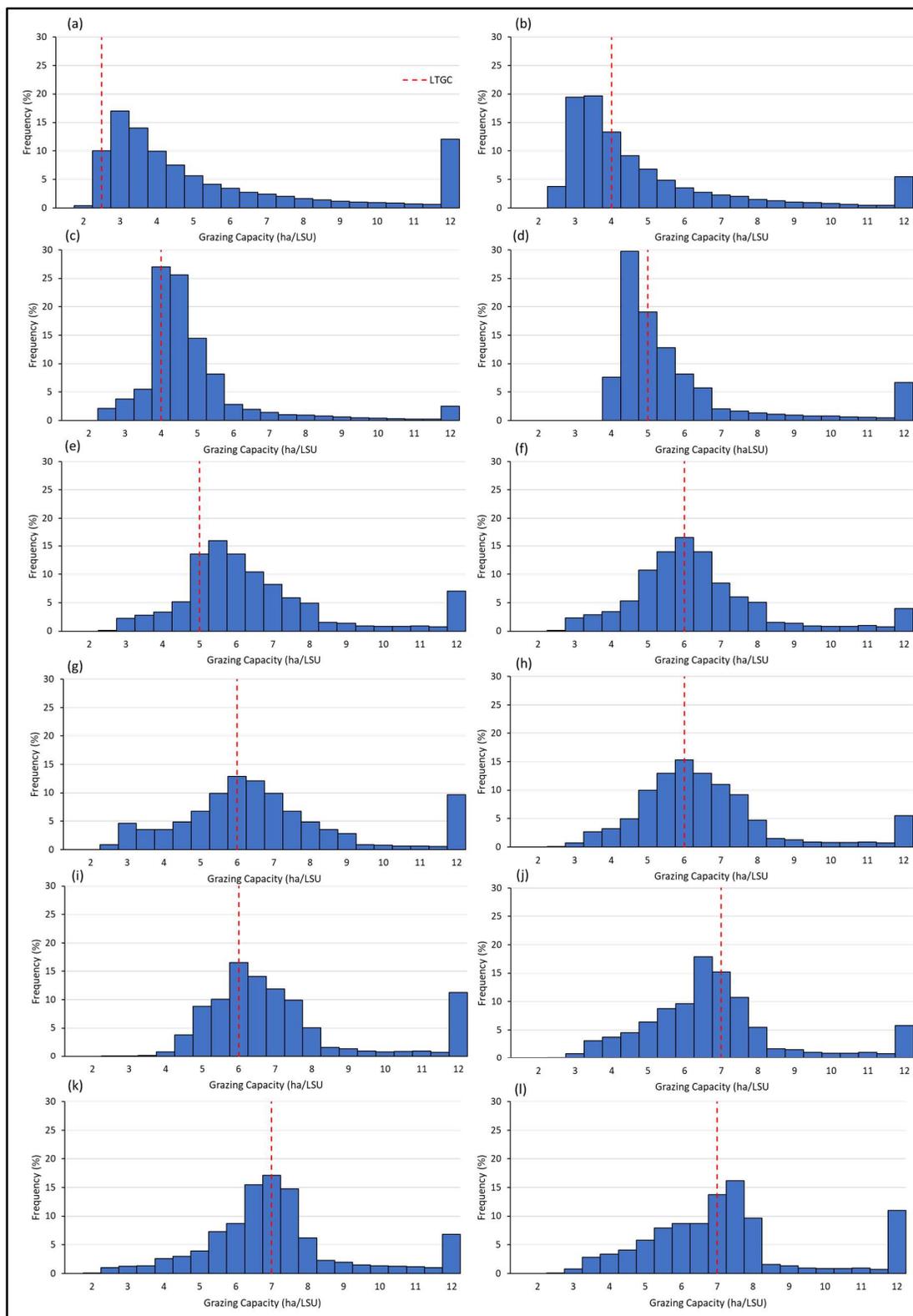


Figure E.1 Frequency histograms for Cedarville LTGC zones (a) C1, (b) C2, (c) C3, (d) C4, (e) C5, (f) C6, (g) C7, (h) C8, (i) C9, (j) C10 and (k) C11



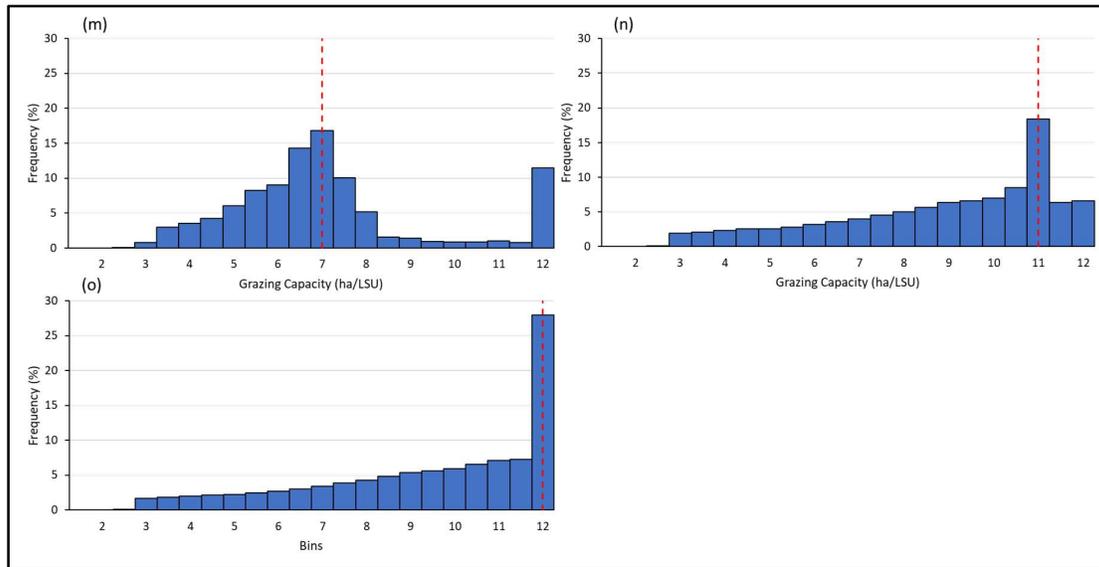


Figure E.2 Frequency histograms for Fort Beaufort LTGC zones (a) FB1, (b) FB 2, (c) FB 3, (d) FB 4, (e) FB 5, (f) FB 6, (g) FB 7, (h) FB 8, (i) FB 9, (j) FB 10, (k) FB 11(l), FB 12, (m) FB 13, (n) FB 14 and (o) FB 15

APPENDIX F

GrazeEngine example user interface screenshots

Figure F1. shows an example screenshot of the Graze Engine user interface when exploring fractional vegetation cover (FVC).

