

# Abstract

---

Savannah biomes are home to a great diversity of flora and fauna species whose survival depend on the inherent heterogeneity within the savannah landscape. In southern African savannahs, grazing lawn patches play a central role in ecosystem dynamics via their influence on herbivory and fire regimes. The extent and distribution of grazing lawns therefore have important cascading impacts on habitat heterogeneity, biodiversity and important ecosystem processes such as nutrient cycling. Despite the keystone nature of grazing lawns in savannah ecosystems, there is limited knowledge on their extent and distribution. There is substantial empirical evidence of the factors that drive the formation and persistence of grazing lawns. However, no broad-scale approach exist to monitor grazing lawns and enable investigations into changes in their distribution and the impacts on broader ecosystem dynamics. Ground-based monitoring of grazing lawns is laborious and inefficient over large spatial and temporal scales.

This research uses high-resolution satellite remote sensing to characterize grazing lawns and investigate dynamics in their cover and structure in southern African savannahs. This is achieved through a series of analysis that addresses three main objectives including (i) developing methods for accurate detection and mapping of savannah grazing lawn distribution using high-resolution satellite imagery and machine learning techniques; (ii) identifying changes in cover and structural distribution of grazing lawns over space and time; and (iii) Identifying the dominant drivers of change in grazing lawn cover and structure. For the analysis, a robust machine learning workflow is developed to identify grazing lawns in mesic and semi-arid savannah landscapes using WorldView-3 imagery; a cost effective approach for high resolution grazing lawn monitoring is developed via fusion of open access Planet and Sentinel-2 imagery; multi-temporal high-resolution satellite images from 2002, 2014 and 2019 are used to identify changes in grazing lawn cover and structure under different savannah landscape conditions, including the effect of drought stress; and lastly, spatio-temporal analysis of grazing lawn occurrence and change trajectory is used to identify the dominant drivers of grazing lawn dynamics.

High grazing lawn detection accuracies were achieved with all machine learning algorithms. Random Forest, Multilayer Perceptron and Support Vector Machine algorithms had similar accuracies and marginally outperformed Classification and Regression Trees algorithm. The results demonstrate the utility of high-resolution satellite images for overcoming savannah heterogeneity challenges to classification, leading to accurate grazing lawn detection. WorldView-3 with its high spatial resolution and broad array of vegetation sensitive spectral bands is particularly ideal for grazing lawn monitoring, and could be used for targeted investigations due to acquisition cost. Alternatively, fusion of open-access planet and Sentinel-2 images provide a cost effective option for operational management applications. In the absence of drought stress, grazing lawn extent increase uniformly, signaling to possible grazer population increase between 2002 and 2014. Most of the increase tend to occur as an expansion to existing patches or in close proximity to existing patches. The impact of drought stress on grazing lawns depends on local landscape characteristics, particularly related to water availability. Gains and losses in grazing lawn cover largely occur as transitions to and from tall grass swards, further highlighting tall grass vegetation as the main competitor for space. Woody encroachment was found not to be an immediate threat to grazing lawn cover, but could gain more significance with projected increase in drought frequency and intensity. In terms of the dominant drivers of grazing lawn dynamics, this study found the presence of water points as resource hot-spots to be an important determining factor of grazing lawn spatial distribution within the savannah landscape. Overall, grazing lawn dynamics was inferred to be primarily driven by grazers, the pattern and nature of which depends on factors that act to alter grazer population and behavior.

This research makes vital methodological contributions to grazing lawn and overall savannah vegetation monitoring. A pioneering remote sensing based methodology for monitoring grazing lawn dynamics is developed. Additionally this research contributes to literature on grazing lawn ecology. A greater knowledge of grazing dynamics has been achieved, contributing to a better understanding of habitat heterogeneity in southern African savannahs. Overall, this research provides important tools and novel ecological insights to guide conservation management in savannah ecosystems.

Keywords: savannahs, southern Africa, grazing lawns, remote sensing, image fusion, machine learning, image classification, change detection, herbivory, fire, drought.

---

## Lay Summary

---

Grazing lawn patches play a central role in savannah ecosystem dynamics via their influence on herbivory and fire. Dynamics in grazing lawn cover therefore have important cascading impacts on habitat heterogeneity, biodiversity, and ecosystem processes such as nutrient cycling. Yet, there is limited knowledge of their extent and distribution. Importantly, no broad-scale approach exists to monitor grazing lawns and enable investigations into changes in their distribution and impacts on broader ecosystem dynamics. Ground based monitoring of grazing lawns is laborious and inefficient over large spatial and temporal scales. This research uses satellite remote sensing to map and investigate dynamics in grazing lawn cover and structure in southern African savannahs. To achieve this, a robust machine learning workflow is developed to identify grazing lawns using WorldView-3 imagery. A cost-effective approach for high resolution grazing lawn monitoring is developed by fusing open access Planet and Sentinel-2 imagery. Multi-temporal very high-resolution satellite images are used to identify changes and drivers of change in grazing lawn cover and structure, including the effect of drought stress under different landscape conditions. This research provides pioneering evidence of the use of remote sensing for grazing lawn monitoring. High-resolution satellite imagery with an array of vegetation sensitive spectral bands helped to overcome savannah heterogeneity challenges to classification, leading to accurate grazing lawn detection. Gains and losses in grazing lawn cover largely occurred as transitions to and from tall grass swards. In the absence of drought stress, grazing lawn extent increased uniformly, as an expansion to existing patches or near existing patches. Drought impact on grazing lawns varied depending on local landscape characteristics, mainly, water availability. This research found the presence of water points as resource hot spots to be an important determining factor of grazing lawn spatial distribution. Overall, grazing lawn dynamics was inferred to be primarily driven by grazers and shaped by factors that act to alter grazer population and behavior. This research provides important tools and novel ecological understanding for conservation management in savannah ecosystems.

---

# Acknowledgements

---

It was great pleasure having the opportunity to work with different advisors and colleagues at various points along this remarkable journey. I'd like to express my heartfelt gratitude and appreciation to my primary supervisor Prof. Paul Aplin for giving me the opportunity to work on this topic towards my PhD and for his guidance and assistance throughout the preparation of this thesis. Thank you Paul for always having your doors open whenever I knocked. To my supervisor, Dr. Ian Powell, I say thank you for your insights and the invaluable discussions which helped shape the ecological focus of this research. I enjoyed our time at Kruger National Park collecting data and bird watching. I want to express my gratitude to Dr Christopher Marston who's initial ideas formed the foundation of this exciting project that I inherited and developed further. Chris was always available to discuss my queries and was instrumental in making sure I had the right data for my research. Along the line, a few administrative adjustments had to be made, which saw Dr Joaquín Cortés and Dr Marcio Pie join the supervisory team. Thank you Joaquín and Marcio for stepping in the to see to it that my PhD journey came to a successful end, *Meda mo ase pii*. Special thanks to former colleague Daniel Knight for his company and resourcefulness during field data collection. I am sincerely grateful to Edge Hill University, Department of Geography and Geology for funding my studies through their GTA studentship scheme and to the Royal Geographical Society for partly funding my fieldwork stay in Kruger National Park, South Africa. I am also grateful to South Africa National Parks (SANParks) for helping to facilitate my field data collection campaign, and particularly to Dr. Izak Smith for his rich ideas and contribution at different stages of my research. Finally, I am most grateful to my family, especially my wife Josephine for the remarkable support, encouragement and motivation throughout my studies.

---

## Declaration

---

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

---

**Kwame Tweneboa Awuah**

---

# Contents

---

<b>Abstract</b>	<b>iii</b>
<b>Lay Summary</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vi</b>
<b>Declaration</b>	<b>vii</b>
<b>Figures and Tables</b>	<b>xii</b>
<b>1 General Introduction</b>	<b>1</b>
1.1 Savannahs in Africa . . . . .	1
1.2 Drivers of Vegetation Structure in African Savannahs . . . . .	2
1.3 Global Change and Impacts on Savannah Ecosystem Structure and Functioning . . . . .	3
1.4 Grazing Lawns as Functional Components in African Savannahs . . . . .	6
1.5 Dynamics in Grazing Lawn Development and Persistence in African Savannahs . . . . .	7
1.6 Grassland Vegetation Monitoring . . . . .	8
1.6.1 Ground-based Methods . . . . .	9
1.6.2 Remote Sensing Methods . . . . .	10
1.7 Remote Sensing Time-series Change Detection . . . . .	13
1.8 Research Focus . . . . .	14
1.8.1 Aim and Objectives . . . . .	15
1.9 Study Area and Fieldwork Protocol . . . . .	16
1.9.1 Study Area . . . . .	16
1.9.2 In-situ Data Collection . . . . .	18
1.10 Thesis Outline . . . . .	19
1.10.1 Chapter 1: General Introduction . . . . .	19
1.10.2 Chapter 2: Probabilistic Mapping and Spatial Pattern Analysis of Grazing Lawns in Southern African Savannahs Using WorldView-3 Imagery and Machine Learning Techniques.....	20

1.10.3	Chapter 3: Spatio-temporal Analysis of Grazing Lawn Dynamics in Southern African Savannahs Using Multi-temporal High-resolution Satellite Images .....	20
1.10.4	Chapter 4: Multi-sensor Optical Image Fusion for Land Cover Classification in a Heterogeneous African Savannah: Toward Accurate and Cost Effective grazing Lawn Monitoring .....	20
1.10.5	Chapter 5: Synthesis and Conclusion .....	21
<b>2</b>	<b>Probabilistic Mapping and Spatial Pattern Analysis of Grazing Lawns in Southern African Savannahs Using WorldView-3 Imagery and Machine Learning Techniques</b>	<b>22</b>
2.1	Introduction .....	22
2.2	Materials and Methods .....	26
2.2.1	Study Sites.....	26
2.2.2	Land Cover and Classification Scheme .....	26
2.2.3	Data .....	28
2.2.4	Preparation of Image Features.....	32
2.2.5	Feature Selection.....	33
2.2.6	Machine Learning Algorithms .....	35
2.2.7	Algorithm Calibration and Evaluation .....	40
2.2.8	Spatial Analysis of Grazing Lawn Distribution.....	43
2.3	Results .....	43
2.3.1	Model Quality for Land Cover Classification.....	43
2.3.2	Grazing Lawn Occurrence Probability Prediction and Classification .....	46
2.3.3	Spatial Patterns in Grazing Lawn Cover .....	51
2.4	Discussion .....	53
2.4.1	Model Quality for Savannah Land Cover Classification.....	53
2.4.2	Grazing Lawn Detection and Model Comparison .....	56
2.4.3	Spatial Patterns in Grazing Lawn Distribution.....	57
2.5	Conclusions.....	59
<b>3</b>	<b>Spatio-temporal Analysis of Grazing Lawn Dynamics in Southern African Savannahs Using Multi-temporal High-resolution Satellite Images</b>	<b>60</b>
3.1	Introduction .....	60
3.2	Materials and Methods .....	63
3.2.1	Study Sites.....	63

3.2.2	Data . . . . .	65
3.2.3	Image Classification and Accuracy Assessment . . . . .	67
3.2.4	Post-classification Characterisation of Grazing Lawn Structure	69
3.2.5	Post-processing and Change Analysis . . . . .	69
3.3	Results . . . . .	71
3.3.1	Land Cover Classification . . . . .	71
3.3.2	Changes in Grazing Lawn Cover and Structure . . . . .	80
3.4	Discussion . . . . .	93
3.5	Conclusions . . . . .	96
<b>4</b>	<b>Multi-sensor Optical Image Fusion for Land Cover Classification in a Heterogeneous African Savannah: Toward Accurate and Cost Effective Grazing Lawn Monitoring</b>	<b>98</b>
4.1	Introduction . . . . .	98
4.2	Materials and Methods . . . . .	102
4.2.1	Study Site . . . . .	102
4.2.2	Land Cover Classification Nomenclature . . . . .	102
4.2.3	Reference Data . . . . .	104
4.2.4	Remote Sensing Data and Preprocessing . . . . .	104
4.2.5	Multi-sensor Image Fusion . . . . .	105
4.2.6	Image Quality Assessment . . . . .	106
4.2.7	Image Segmentation . . . . .	107
4.2.8	Feature Extraction . . . . .	108
4.2.9	Image Classification and Accuracy Assessment . . . . .	109
4.2.10	Feature Importance Estimation . . . . .	111
4.3	Results . . . . .	111
4.3.1	Image Fusion Accuracy . . . . .	111
4.3.2	Land Cover Classification . . . . .	113
4.3.3	Feature Importance . . . . .	118
4.3.4	Grazing Lawn Detection Accuracy . . . . .	119
4.4	Discussion . . . . .	121
4.4.1	Image Fusion for high-resolution Savannah Land Cover Classification . . . . .	121
4.4.2	Grazing Lawn Detection and Monitoring with Open-access Optical Satellite Imagery . . . . .	124
4.5	Conclusions . . . . .	125

<b>5 Synthesis and Conclusion</b>	<b>127</b>
5.1 Summary .....	127
5.1.1 Satellite remote sensing of savannah grazing lawns .....	128
5.1.2 Changes in grazing lawn cover and structure .....	131
5.1.3 Drivers of change in grazing lawn cover and structure .....	133
5.2 Contribution of this research.....	135
5.3 Limitations and future research.....	136
<b>References</b>	<b>138</b>
<b>Appendices</b>	
<b>A Probabilistic Mapping and Spatial Pattern Analysis of Grazing Lawns in Southern African Savannahs Using WorldView-3 Imagery and Machine Learning Techniques</b>	<b>168</b>
A.1 Multicollinearity and Feature Selection .....	168
A.2 Comparison of Grazing Lawn Area Estimates across Models in Each Landscape.....	171
A.3 Confusion Matrices for the Lower Sabie Landscape .....	171
A.4 Confusion Matrices for the Satara Landscape.....	172
<b>B Spatio-temporal Analysis of Grazing Lawn Dynamics in Southern African Savannahs Using Multi-temporal High-resolution Satellite Images.</b>	<b>174</b>
B.1 Confusion Matrices used for accuracy assessment .....	175
B.2 Transition matrices for land cover change analysis .....	177
<b>C Multi-sensor optical image fusion for land cover classification in a heterogeneous African savannah: toward accurate and cost effective grazing lawn monitoring.</b>	<b>181</b>
C.1 Confusion Matrices .....	181

---

# Figures and Tables

---

## Figures

1.1	Location of study sites in Kruger National Park, South Africa .....	17
1.2	Illustration of plot design for data collection. ....	18
2.1	Map of study area showing locations of the Satara and Lower Sabie field sites in Kruger National Park (KNP), with enlarged views of the WorldView-3 satellite scenes (False colour: NIR1, R, G) overlaid with hydrology. Inset map shows the location of KNP within South Africa. Geological information was obtained from South African National Parks data repository .....	27
2.2	Examples of Plant Functional Types (PFTs): (a) woody evergreen, (b) woody deciduous, (c) bunch grass and (d) grazing lawns .....	28
2.3	Sample display of land cover types from WorldView-3 satellite image scene (False colour: NIR1, R, G). (a) Woody evergreen; (b) Woody deciduous; (c) Bunch grass; (d) Grazing lawn; (e) Water body; (f) Bare; (g) Built-up; (h) Shadow .....	29
2.4	Conceptual workflow showing steps in machine learning model development and evaluation towards grazing lawn detection. ....	34
2.5	Land cover classification of the training region from (a) RF, (b) SVM, (d) CART and (e) MLP models. The WorldView-3 image scene (False colour: NIR1, R, G) of the training region is showed in panel (c). RF = Random Forest, SVM = Support Vector Machines, CART = Classification and Regression Trees and MLP = Multilayer Perceptron.....	44
2.6	Image feature weights derived from permutation feature importance estimates for Random Forest (RF), Support Vector Machines (SVM), Classification and Regression Trees (CART) and Multilayer Perceptron (MLP) models. Feature weights are sorted in an descending order across models to identify features on high predictive importance. For the detailed names and description of image feature acronyms, refer to Table 2.3.....	45
2.7	Grazing lawn occurrence probability surfaces (A); optimal probability threshold plot (B); and binary map of grazing lawn and other cover (C) derived from RF, SVM, CART and MLP models for the Lower Sabie landscape .....	48

2.8	Grazing lawn occurrence probability surfaces (A); optimal probability threshold plot (B); and binary map of grazing lawn and other cover (C) derived from RF, SVM, CART and MLP models for the Satara landscape.....	49
2.9	Landscape-scale summary of grazing lawn spatial characteristics. (A) Number of grazing lawn patches; (B) Proportion of total landscape covered by grazing lawns; (C) Physical connectedness of grazing lawn patches; (D) Distribution of grazing lawn patch size. Dashed horizontal line represents mean patch size.....	52
2.10	Plots of spatial metrics showing patterns in grazing lawn spatial structure and distribution with distance from water-points. CI = Cohesion Index, MPA = Maximum Patch Area, PL = Proportion of Landscape .....	54
3.1	Map showing study site locations (Sites 1, 2 and 3) overlaid on SRTM Digital Elevation Model (DEM). Inset map shows Kruger National Park (KNP) boundary and the study region in southern KNP .....	64
3.2	Total annual rainfall for Skukuza, Kruger National Park, South Africa, from 1991 to 2019. Major drought years are highlighted in red. Rainfall data was obtained from (South African National Parks, 2021).....	64
3.3	Land cover map for Site 1 from IKONOS image in 2002.....	72
3.4	Land cover map for Site 1 from IKONOS image in 2014.....	73
3.5	Land cover map for Site 1 from Planet image in 2019.....	74
3.6	Land cover map for Site 2 from QuickBird image in 2002.....	75
3.7	Land cover map for Site 2 from WorldView-2 image in 2014.....	76
3.8	Land cover map for Site 2 from WorldView-3 image in 2019.....	77
3.9	Land cover maps for Site 3 from IKONOS image in 2002.....	78
3.10	Land cover map for Site 3 from IKONOS image in 2014.....	79
3.11	Land cover map for Site 3 from WorldView-3 image in 2019.....	80
3.12	Grazing lawn change map for study site 1 for the period between 2002 and 2014.....	84
3.13	Grazing lawn change map for study site 1 for the period between 2014 and 2019.....	85
3.14	Grazing lawn change map for study site 1 for the period between 2002 and 2019.....	86
3.15	Grazing lawn change map for study site 2 for the period between 2002 and 2014.....	87

3.16 Grazing lawn change map for study site 2 for the period between 2014 and 2019 .....	88
3.17 Grazing lawn change map for study site 2 for the period between 2002 and 2019 .....	89
3.18 Grazing lawn change map for study site 3 for the period between 2002 and 2014 .....	90
3.19 Grazing lawn change map for study site 3 for the period between 2014 and 2019 .....	91
3.20 Grazing lawn change map for study site 3 for the period between 2002 and 2019 .....	92
3.21 Plots of spatial metrics showing dynamics in grazing lawn structure over time: (A) Patch Density (PD); (B) Edge Density (ED); and (C) Cohesion Index (CI). .....	93
 4.1 Map of study area showing location of the Lower Sabie field site in Kruger National Park (KNP), with enlarged view of the Planet image scene (False colour: R, NIR, G) from NICFI overlaid with the Nkhulu exclosure boundary (sourced from South African National Parks Scientific Services (Siebert & Eckhardt, 2008)). Inset map shows the location of KNP within South Africa .....	103
4.2 False colour (R: NIR, G: Red, B: Green) display of original and fused images scenes of a sub-area from Gram-Schmidt sharpening. A visual comparison shows higher spatial resolution of Planet and Fused images relative to the Sentinel-2A image .....	112
4.3 Spectral quality metrics from pixel-level fusion. The figure shows Correlation Coefficient from Sentinel-2 and fused image band pairs (left) and boxplot showing summary of Structural Similarity Index (SSIM) for different savannah land cover types (right). B = Blue, G = Green, R = Red, RE-1 = Red Edge-1, RE-2 = Red Edge-2, RE-3 = Red Edge-3, NIR = Near Infrared, SWIR-1 = Short Wave Infrared-1 and SWIR-2 = Short Wave Infrared-2 .....	113
4.4 Land cover map derived from Planet image.....	114
4.5 Land cover map derived from Fused image .....	115
4.6 Land cover map derived from Sentinel-2A image .....	115

4.7	General classification accuracies from stratified 10-fold cross-validation. Accuracy metrics were calculated as weighted averages from individual land cover class accuracies. Accuracy values represent fractions between 0 and 1.	116
4.8	Post-hoc comparison of Sentinel-2A, Planet and Fused datasets with Ne-menyi test, based on F-scores for Woody and Bunch grass classes. Connected groups are not significantly different at $\alpha = 0.05$ . Critical Difference (CD) = 1.048	118
4.9	Image feature importance estimates. Feature weights represent contribution to classification accuracy and are sorted in descending order. For feature acronyms, B = Blue, G = Green, R = Red, RE1 = Red Edge-1, RE2 = Red Edge-2, RE3 = Red Edge-3, RE4 = Red Edge-4, N = Near Infrared, SWIR1 = Shortwave Infrared-1, SWIR2 = Shortwave Infrared-2, GEMI = Global Environmental Monitoring Index, MSAVI2 = Modified Soil Adjusted Vegetation Index-2, Mean = Mean, Var = Variance, Savrg = Sum Average	119
4.10	Maps of grazing lawn cover derived from (A) Sentinel-2A, (B) Planet and (C) Fused images; and (D) Omission and (E) Commission error rates for grazing lawn classification against other land cover types. Error rates represent fractions between 0 and 1.	120
A.1	Feature selection results from VIF and RF-RFE analysis. (A) VIF of selected spectral indices. (B) VIF of selected texture features. (C) Importance scores of final selected image features following RF-RFE. Original bands and selected spectral indices and texture features from VIF served as input to RF-RFE. Final selection was based on number of features that retained optimal accuracy	169
A.2	Plot of accuracy versus number of features	170
C.1	Confusion matrices from 10-fold classification and accuracy assessment of the Sentinel-2A image. S2A = Sentinel-2A, K = Iteration. For the land cover classes considered, Woody = 1, Bunch grass = 2, Grazing lawns = 3, Water body = 4, Bare = 5 and Built-up = 6.	182
C.2	Confusion matrices from 10-fold classification and accuracy assessment of the Planet image. PS = Planet, K = Iteration. For the land cover classes considered, Woody = 1, Bunch grass = 2, Grazing lawns = 3, Water body = 4, Bare = 5 and Built-up = 6.	183

C.3	Confusion matrices from 10-fold classification and accuracy assessment of the Fused image. S2A+PS = Fused image, K = Iteration. For the land cover classes considered, Woody = 1, Bunch grass = 2, Grazing lawns = 3, Water body =4, Bare = 5 and Built-up = 6.....	184
-----	---	-----

---

## Tables

1.1	Major optical civilian Earth observation satellites with global coverage. . .	10
2.1	Description of land cover classification nomenclature and reference data. Numbers represent number of reference points, while figures in parenthesis represent area of training polygons in hectares. Lower Sabie and Satara validation points are separated by “/” (i.e., Lower Sabie/Satara). . .	30
2.2	Description of datasets used. ....	31
2.3	Initial image features serving as potential predictors.....	33
2.4	Results of optimal hyperparameter values used in each model, from 2x5 nested cross-validation using Scikit Learn python package. Refer to the work in (Pedregosa <i>et al.</i> , 2011) for more details on model hyperparameters.	36
2.5	Accuracy scores (F1 and Overall Accuracy) from 2 x 5 nested cross-validation showing a comparison of model performance. RF = Random Forest, SVM = Support Vector Machines, CART = Classification and Regression Trees, MLP = Multilayer Perceptron. ....	44
2.6	Summary of the first three most important image features from spectral bands, spectral indices and texture variables across all models. Image features that appear in at least three models are in bold. For the detailed names and description of image feature acronyms, refer to Table 2.3.....	47
2.7	Model Precision, Recall and F-score metrics of grazing lawn detection in both Lower Sabie and Satara landscapes. RF = Random Forest, SVM = Support Vector Machines, CART = Classification and Regression Trees, MLP = Multilayer Perceptron. ....	50
2.8	Accuracy adjusted area estimates of grazing lawn cover in Lower Sabie and Satara landscapes. Area estimates with different letters differ significantly and vice versa in each landscape. RF = Random Forest, SVM = Support Vector Machines, CART = Classification and Regression Trees, MLP = Multilayer Perceptron. ....	50

2.9	McNemar's chi-squared test ( $\chi^2$ ) of marginal homogeneity between model pairs. Values in parenthesis represent p-value. Model pairs that show statistically significant difference ( $p \leq 0.05$ ) in error rate are in bold. CART = Classification and Regression Trees, MLP = Multilayer Perceptron, RF = Random Forest, SVM = Support Vector Machines .....	51
2.10	Pearson Correlations (r) and Coefficients of Determination ( $r^2$ ) from the relationship between grazing lawn spatial metrics and distance from water source. PL = Proportion of Landscape, MPA = Maximum Patch Area and CI = Cohesion Index. Relationships are significant at $p < 0.0001$ ***, $p < 0.001$ ** and $p < 0.01$ *.....	53
3.1	Description of target land cover classes used for image classification. ....	66
3.2	Overview of satellite images used for land cover classification. Spectral bands: B = blue, C = cirrus, CA = coastal aerosol, CB = coastal blue, G = green, NIR = near infra-red, R = red, RE = red edge, SWIR = short-wave infrared, Y = yellow.....	67
3.3	. Summary of selected spatial metrics. $A$ is the total landscape area, $a_{ij}$ is area of patch $i j$ , $p_{i j}$ is the perimeter of patch $i j$ , $n_i$ is the number of patches for patch $i$ , and $Z$ is the total number of cells in the landscape, $e_{ik}$ is the total edge in landscape between patch types $i$ and $k$ , $m'$ is number of patch types present in the landscape including landscape border (Mcgarigal & Marks, 1995).....	70
3.4	Summary of land cover classification accuracy scores for individual land cover categories and overall classification. Overall accuracy scores are weighted averages of accuracy scores from individual land cover classifications. Acronyms P, R and F represent precision, recall and F-score respectively.....	81
3.5	Summary of land cover area and changes for all sites. Only land cover classes relevant for grazing lawn dynamics are considered. Changes (%) correspond to proportion of the total area for each study site .....	82
3.6	Relative change trajectories of grazing lawn cover expressed as proportion of the overall landscape area. ....	83
4.1	Description of land cover classification nomenclature and reference data. Reference samples are expressed as number of polygons and total area covered by polygons (in sq km) for each land cover class .....	103

4.2	Summary of classification accuracy scores for savannah land cover types showing a comparison across dataset used (Sentinel-2A, Planet, Fused). Accuracy scores represent mean $\pm$ standard deviation from 10-fold stratified cross-validation. Accuracy values represent fractions between 0 and 1.....	117
4.3	Summary of Friedman chi-square ( $\chi^2_F$ ) test for significant differences (at $\alpha = 0.05$ ) in accuracy scores for overall map and individual land cover types. Significance test was conducted on F-scores from stratified 10-fold cross-validation.....	118
A.1	Two-proportions Z-test comparing the proportions of estimated grazing lawn cover. Values in parenthesis represent p-value. Model pairs that show statistically significant difference ( $p \leq 0.05$ ) in proportion of grazing lawn cover are in bold. CART = Classification and Regression Trees, MLP = Multilayer Perceptron, RF = Random Forest, SVM = Support Vector Machines.....	171
A.2	Confusion matrix summarising results from Random Forest (RF) model classification of grazing lawn and other cover.....	171
A.3	Confusion matrix summarising results from Support Vector Machines (SVM) model classification of grazing lawn and other cover.....	171
A.4	Confusion matrix summarising results from Classification and Regression Trees (CART) model classification of grazing lawn and other cover .....	172
A.5	Confusion matrix summarising results from Multilayer Perceptron (MLP) model classification of grazing lawn and other cover.....	172
A.6	Confusion matrix summarising results from Random Forest (RF) model classification of grazing lawn and other cover.....	172
A.7	Confusion matrix summarising results from Support Vector Machines (SVM) model classification of grazing lawn and other cover.....	172
A.8	Confusion matrix summarising results from Classification and Regression Trees (CART) model classification of grazing lawn and other cover .....	173
A.9	Confusion matrix summarising results from Multilayer Peceptron (MLP) model classification of grazing lawn and other cover.....	173
B.1	Site 1 confusion matrix used for classification accuracy assessment for each time period (2002, 2014 and 2019). Columns represent classification labels and rows represent reference labels.....	175

B.2	Site 2 confusion matrix used for classification accuracy assessment for each time period (2002, 2014 and 2019). Columns represent classification labels and rows represent reference labels.....	176
B.3	Site 3 confusion matrix used for classification accuracy assessment for each time period (2002, 2014 and 2019). Columns represent classification labels and rows represent reference labels.....	177
B.4	Site 1 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2014. Columns represent 2014 entries and rows represent 2002 entries .....	177
B.5	Site 1 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2014 and 2019. Columns represent 2019 entries and rows represent 2014 entries .....	178
B.6	Site 1 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2019. Columns represent 2019 entries and rows represent 2002 entries .....	178
B.7	Site 2 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2014. Columns represent 2014 entries and rows represent 2002 entries .....	178
B.8	Site 2 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2014 and 2019. Columns represent 2019 entries and rows represent 2014 entries .....	179
B.9	Site 2 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2019. Columns represent 2019 entries and rows represent 2002 entries .....	179
B.10	Site 3 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2014. Columns represent 2014 entries and rows represent 2002 entries .....	179
B.11	Site 3 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2014 and 2019. Columns represent 2019 entries and rows represent 2014 entries .....	180
B.12	Site 3 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2019. Columns represent 2019 entries and rows represent 2002 entries .....	180

---

# Chapter 1

---

## General Introduction

---

### 1.1 Savannahs in Africa

Savannahs constitute the world's largest terrestrial biome and extends over nearly 33 million square-kilometers, accounting for approximately 20 % of the earth's land surface area (Devine *et al.*, 2015; Sankaran & Ratnam, 2013). They occur widely across Africa, Asia, Australia and South America and support a great diversity and density of flora and fauna communities (Sankaran & Ratnam, 2013). Savannahs also play a substantial role in human economies with the presence of rangelands that supports majority of the world's livestock biomass (Sankaran & Anderson, 2009).

Savannah biomes are widespread in Africa, covering about 50 % of the continent's land area (Osborne *et al.*, 2018). They extend from west to east Africa sandwiched between the Sahara desert to the north and the rainforest to the south, and then round the Congo basin to Namibia (Shorrocks & Bates, 2015). This extensive cover of savannah vegetation however exhibits considerable variation in structure creating diverse gradients of grass-shrub and / or grass-tree matrix (Sankaran & Ratnam, 2013). The gradient of wooded vegetation stretches to more open grasslands with increasing distance from the Congo basin both northwards towards the Sahara desert and southwards through the miombo woodland to the drier regions of Namibia and South Africa (Osborne *et al.*, 2018; Shorrocks & Bates, 2015). The grassy vegetation persists mainly due to natural boundaries to the spread of woody cover imposed by water availability (i.e. rainfall), fire and herbivory (Bond, 2008; Marston *et al.*, 2019; Sankaran *et al.*, 2005). Woody plant recruitment is also negatively impacted by competition from grasses for nutrients and light (Osborne *et al.*, 2018).

The woody component in the drier west (i.e. with 100-600 mm mean annual precipitation (MAP)) is dominated by shrub forms including *Acacia laeta* and *Balanites aegyptiaca*, while annual species like *Aristida stipoides* and *Cenchrus biflorus* dominates the grass layer (Sankaran & Ratnam, 2013). Towards the east around southern Kenya and the Maasai dry lands of Tanzania (with 600-800 MAP), perennial grasses such as *Panicum coloratum*, *Themeda triandra* and *Andropogon sp.* dominate (Huntley, 1982; Sankaran & Ratnam, 2013). In southern Africa, savannah vegetation forms a grass-shrub transition between the northern mopane and the southern deserts, with more open vegetation occurring towards the south (Sankaran & Ratnam, 2013). Karoo shrubs such as *Acacia newbrownii* and *Parkinsonia africana* dominate the shrub layer whereas tufted grasses dominated by the genus *Stipagrostis* forms the scattered grass layer (Huntley, 1982).

## 1.2 Drivers of Vegetation Structure in African Savannas

The factors that promote co-existence of grass and tree life forms and shape the relative distributions of vegetation structure has been a subject of much research. The availability of resources such as water and soil nutrients and the persistence of disturbance factors including fire and herbivory has been identified as critical in shaping the balance of this grass-tree matrix (Archibald, 2008; Sankaran & Ratnam, 2013). However, there is no consensus on the roles or relative importance of these factors to structural dynamics in savannah vegetation. Using data from 854 sites across Africa, Sankaran *et al.* (2005) showed that the proportion of woody cover in arid and semi-arid savannahs (i.e. with less than 650 mm MAP) is mainly controlled by water availability and increases with increasing MAP. The results further revealed that below the threshold MAP of 650 mm, fire, herbivory and soil properties interact to further reduce woody cover or suppress tree seedling recruitment, thus shifting the balance to grass dominance. In contrast, areas with more than 650 mm MAP were found to be dominated by woody vegetation, and grassy vegetation could only co-exist after disturbance events (herbivory and fire). In a related study, Sankaran *et al.* (2008) identified MAP as the most important predictor of woody cover, followed by fire return periods, soil characteristics and herbivory regimes. Levick *et al.* (2009) observed that herbivory predominantly exerted more influence on three-dimensional structure of vegetation than fire in Kruger National Park, where areas protected from

herbivory had more tree cover and tall canopy layers than control areas. Fire on the other hand mainly restricted the growth of tree seedlings within 0 - 3 m in height in the presence or absence of herbivory. The study however offered no control for the influence of soil nutrients which is known to influence forage quality and hence herbivory (Grant & Scholes, 2006). Other studies attribute the displacement of grasses by woody vegetation to global change factors, with Buitenwerf *et al.* (2012) showing the consistency of increasing  $CO_2$  concentration with increasing tree density when keeping all disturbance factors constant.

### **1.3 Global Change and Impacts on Savannah Ecosystem Structure and Functioning**

Earth's changing climate poses significant stress to natural ecosystems globally. For savannahs, current evidence points to major changes in ecosystem structure and functioning as a result of global change phenomena such as erratic precipitation regimes, increased atmospheric carbon dioxide ( $CO_2$ ) and atmospheric nitrogen (N) and phosphorus (P) deposition, which threatens ecological integrity (Midgley & Bond, 2015; Parr *et al.*, 2014).

Savannahs are inherently characterized by pronounced dry periods and highly seasonal rainfall regimes. As such, savannah vegetation has evolved diverse disturbance resistance and resilience traits as strategies to cope with such variability (Pennington *et al.*, 2018; Sankaran, 2019). Nevertheless, current and projected increase in drought severity and frequency could shift tree-grass competition dynamics and alter disturbance drivers such as fire and herbivory, with consequences for vegetation composition, structure and overall ecosystem functioning (Fensham *et al.*, 2009; Staver *et al.*, 2019). Drought stress coupled with warmer temperatures increase evaporative demand on savannah vegetation and hastens susceptibility to mortality (Sankaran, 2019).

Wide-spread tree die-back in response to drought stress has been extensively documented in savannahs (Case *et al.*, 2019; Fensham *et al.*, 2009; O'connor, 1998). Case *et al.* (2020) showed that trees that have relatively low investment in non-structural carbohydrates (e.g. *Dichrostachys cinerea*) suffered most from drought-induced mortality, resulting in significantly limited distribution and abundance in drought-frequent savannah landscapes. In contrast, drought tolerant tree species could benefit from

attenuated fire (frequency and intensity) and competition (for limited soil moisture) due to drought-induced reduction of surrounding grass biomass (February *et al.*, 2013). Even so, increasing drought frequency disrupts resistance and resilience mechanisms of drought-tolerant species including cavitation fatigue and reduced capacity to replenish stored reserves (Sankaran, 2019). Such physiological damage negatively impacts recovery through resprouting, reduces seed production which impacts recolonization (Hartmann *et al.*, 2018) and can lead to directional shifts in tree species composition through individual mortality and overall exclusion.

As in the case of trees, savannah grasses show a strong physiological response to water stress. Although grasses may have better recovery response to low-to-moderate drought episodes than trees (Sankaran, 2019), evidence suggests that they are equally or even more sensitive to intense droughts (Ghannoum, 2009), with substantial impact on mortality relative to woody plants (Walker *et al.*, 1987). Grasses also exhibit a range of physiological drought tolerance and avoidance strategies which varies among species. Studying 426 grass species distributed over a range of climatic and phylogenetic affinities, Craine *et al.* (2013) showed that species with higher photosynthetic rates ( $17.0 \pm 0.5 \text{ } \mu\text{mol m}^{-2}\text{s}^{-1}$ ) and higher stomatal conductance ( $0.185 \pm 0.008 \text{ } \mu\text{mol m}^{-2}\text{s}^{-1}$ ) were typically more tolerant to physiological drought. Further, leaf width also showed some correlation with drought tolerance, where wide-leaved grasses were largely drought-intolerant and narrow-leaved grasses had a range of tolerance and intolerance abilities (Craine *et al.*, 2013). Other significant traits that facilitate drought tolerance and avoidance in savannah grasses include extensive below ground biomass allocation (deep roots) with arbuscular mycorrhizal fungi associations (Nippert & Holdo, 2015), rapid leaf shedding and leaf rolling strategies (Bolger *et al.*, 2005).

Savannah vegetation response to drought is strongly mediated by the occurrence of disturbance events which vary both spatially and temporally (Macgregor & O'Connor, 2002; Swemmer *et al.*, 2018). Herbivory (including browsing and grazing) and fire are important disturbance mechanisms that regulate growth and survival of trees and grasses in savannah ecosystems (Archibald *et al.*, 2005; Archibald, 2008; Donaldson *et al.*, 2018). As such, the extent and intensity of such disturbances, which in themselves are influenced by droughts, significantly alter the effects of droughts on trees and grasses. Browsing suppresses tree growth by reducing starch reserves and increasing susceptibility to pathogen attacks with consequences for individual tree survival during droughts (Sankaran, 2019). Similarly, grazing influences spatial

variation in grass biomass and productivity during droughts. For example, Donaldson *et al.* (2020) observed that protracted drought decoupled grazers from already established grazing lawn patches - which are a key food resource for some mega and meso-herbivores - leading to colonization by tussock grasses during post-drought recovery. Additionally, Staver *et al.* (2019) reported severe grass declines both in droughted areas and non-droughted refugia due to grazing and drought-induced grazer movement. Interaction between drought and patterns in grazing and/or grazer behaviour can thus result in the eventual extension of the ecological impacts of droughts even to non-droughted areas. On the other hand, drought impacts on fire and their interactive effects on savannah ecosystem are largely contingent on the availability of grass fuel biomass (Sankaran, 2019). Fuel loads build up from moisture availability and may be cumulative from preceding years. As such more intense and widespread fires are expected early in drought seasons when droughts are preceded by above average rainfall conditions. Fire intensity is particularly pronounced in areas with low grazing pressure where there is high grass biomass build up, which significantly hastens tree mortality (Donaldson *et al.*, 2018). Conversely, protracted droughts reduce grass biomass due to significant reduction in productivity coupled with early drought fire events. This in turn minimizes frequency, intensity and extent of fires. Increasing drought frequency is thus expected to lead to longer fire return intervals and a decrease in total area burnt including size of individual burns (Balfour & Howison, 2002). This, coupled with reduced browsing pressure from drought-induced herbivore mortality (Walker *et al.*, 1987), will likely favour recruitment of woody vegetation leading to a shift in savannah vegetation structure and ecosystem functioning.

Global change phenomena such as increasing atmospheric CO<sub>2</sub> have been identified to potentially enhance widespread encroachment of woody plants in grassy savannahs. Short and regular fire intervals in savannahs (every 1 - 3 years (Bond & Midgley, 2000)) control woody plant recruitment, particularly in areas with low browsing pressure. As such, woody plants invest in carbon storage to survive future injury and carbon allocation for rapid stem growth in order to escape fires and reach full size (Bond & Midgley, 2000). Elevated CO<sub>2</sub> conditions enhances rapid resprouting and height growth (Ceulemans *et al.*, 1995) leading to enhanced woody plant recruitment rates and higher stem densities in previously open savannahs (Stevens *et al.*, 2017), with cascading impacts on ecosystem structure and functioning including biodiversity (Ratajczak *et al.*, 2012), soil carbon (Berthrong *et al.*, 2012) and grazing potential (Angassa & Baars, 2000).

## 1.4 Grazing Lawns as Functional Components in African Savannahs

Grazing lawns are distinct grass communities structurally typified by short-statured grasses and exhibit stoloniferous growth pattern (Hempson *et al.*, 2015). Their formation, growth and persistence is stimulated by constant grazing which creates dense and palatable swards that in turn attract more grazers. This cycle of regular grazing is necessary to prevent invasion by tall grass species that would otherwise out-compete grazing lawns for light. In southern African savannahs, grazing lawns represent key food resources for large mammal grazers such as wildebeest (*Connochaetes taurinus*), zebra (*Equus burchellii* ), and white rhino (*Ceratotherium simum*), and reveal a co-evolutionary history of grasses and a diversity of large herbivores (Cromsigt & Olff, 2008; Hempson *et al.*, 2015).

Due to their stoloniferous growth adaptation, much of the structural growth parts of grazing lawns including stem and buds escape grazing, leaving only a rich build-up of leaf material that is accessible to grazers. In comparison to their tall grass counterparts, grazing lawn swards have higher foliar nitrogen (N) levels and low carbon to nitrogen ratio (C:N), making them more easily digestible (Chaves *et al.*, 2006) and attractive to grazers. For grazers that utilize grazing lawns, seasonal availability could play a significant role in grazer behaviour (e.g. movement) and population dynamics (Hempson *et al.*, 2015). For example, access to high quality forage resources on grazing lawns during the wet season may provide females with the nutritional requirements to meet high energy demands of late pregnancy and lactation (Parker *et al.*, 2009). Additionally, improved nutrition increases the likelihood of survival in juveniles during their first dry season (Gaillard *et al.*, 2000). As such the presence of grazing lawns could enhance grazer recruitment rates. In addition to the nutritional benefits, grazing lawn patches reduces the risk of predation for grazers due to the low amount of herbaceous cover for predators (Hempson *et al.*, 2015).

The structural characteristics of grazing lawns also have important consequences for fire behavior, with cascading impacts on overall ecosystem dynamics, including herbivore movement and vegetation structure (Donaldson *et al.*, 2018; Hempson *et al.*, 2019). Unlike tall grass swards, grazing lawns do not hold on to moribund growth parts and thus, have very limited fuel biomass that could hasten fire propagation

through the savannah landscape (Archibald *et al.*, 2005). Thus, grazing lawn patches effectively serve as fire breaks and influence the spread, intensity and frequency of fires. Given their influence on herbivores and fire regime, grazing lawn effects can extend well beyond their borders (Hempson *et al.*, 2015).

Ultimately, grazing lawns have significant conservation value as habitat for a broad diversity of fauna and flora species. Specialist short-grass fauna such as grasshoppers (Samways & Kreuzinger, 2001), spiders (Warui *et al.*, 2005), birds (Krook *et al.*, 2007) and mammalian grazers (e.g. rhino, wildebeest) thrive on grazing lawn patches. Additionally, grazing lawns shape tree community composition by supporting tree species that are adapted to withstand heavy browsing (Staver *et al.*, 2012) given the diversity of ungulates that utilize lawn patches, and are also associated with diverse grass and forb species, which altogether enhances plant biodiversity (Hempson *et al.*, 2015).

## 1.5 Dynamics in Grazing Lawn Development and Persistence in African Savannahs

Grazing lawn presence may be static - occurring on nutrient hot-spots (Hempson *et al.*, 2015) - or dynamic - shaped by the interaction between fire and herbivory (Donaldson *et al.*, 2018). Different parts of the savannah landscape may be predisposed to grazing lawn formation due to localized availability of resources that concentrate grazers. These include areas around water points and areas of mineral accumulation such as around termite mounds (Hempson *et al.*, 2015).

Grazing lawn patches are maintained by constant grazing, which results in a feedback of dense nutrient rich plant growth that in turn attract more grazers, creating a feedback loop (Hempson *et al.*, 2015). The initiation of a cycle of regular grazing thus appears to be a critical factor for grazing lawn development and persistence. Nonetheless, the rates and specific pathways of grazing lawn development likely depend on factors like rainfall, fire and soil types (Archibald, 2008). Rainfall has a strong influence on the rate of grass biomass accumulation and the height growth of tall grass stands. The relative proportion of grazing lawns to high biomass tall grasses is also influenced by dynamics in soil nutrients through its strong influence on grass productivity. Under high rainfall and soil nutrient conditions, increased grazing frequency is required to prevent the invasion of tall-grass light competitors. Too

infrequent grazing increases the vulnerability of a switch to upright growth forms by facultative lawn species (Archibald *et al.*, 2005; Hempson *et al.*, 2015). Conversely, exceptionally heavy grazing may lead to degradation due to trampling and soil compaction, or result in the predominance of wire-grass species like *Aristida* sp.

Fire also consumes grass biomass and has the potential to shift grass community composition and structure within different environmental constraints (Leonard *et al.*, 2010). Tall bunch grasses with low forage quality dominate fire-driven grassy systems (Hempson *et al.*, 2015). Additionally, post-fire regrowth can also attract grazers away from previously established grazing lawns causing them to be invaded by tall bunch grasses (Archibald *et al.*, 2005).

## 1.6 Grassland Vegetation Monitoring

Monitoring grassland vegetation is vital for the purposes of research, management planning and biodiversity conservation (Carlsson *et al.*, 2005). Depending on the available infrastructure, a number of methods exist for grassland vegetation monitoring (Ali *et al.*, 2016). Such approaches however, can be broadly categorized into ground-based and remote sensing-based methods (Ali *et al.*, 2016; Carlsson *et al.*, 2005; del Pozo *et al.*, 2006). Both methods have been widely applied complementarily to produce precise estimates of grassland vegetation variables over extensive spatial and temporal scales. For example, Jin *et al.* (2014) combined data from field sampling and remote sensing-based vegetation indices to estimate the spatio-temporal variations in biomass of China's temperate grassland. Additionally, Gaitán *et al.* (2013) calibrated a linear regression model using field data and multiple remote sensing indices to predict spatial variability in structure and functioning of Patagonian steppes. However, the feasibility of such data integration (including accuracy of estimates) likely depends on factors such as phenology, volume of field data, spatial resolution of the remote sensing data and the congruence in acquisition dates of field and remote sensing data (Ali *et al.*, 2016; Casady *et al.*, 2013; Lu, 2005).

### **1.6.1 Ground-based Methods**

Ground-based measurement involves in-situ data recording often at selected field locations or in delineated areas (i.e. plots) within the landscape. In most cases, the selection of field location is based on a predefined statistical sampling design (Fehrmann *et al.*, 2019). However, conditions such as cost, safety and spatial occurrence of target variables may necessitate arbitrary or purposive selection of field locations for data recording (Stehman & Overton, 1994; VandenBygaart, 2006). For example, VandenBygaart (2006) identified purposive non-probability sampling as more appropriate for monitoring soil organic carbon (SOC) due to the large inherent spatial variability and the cost of SOC analysis per sample. Unlike an arbitrary selection of field measurement locations, a design-based sampling approach allows unbiased point and interval estimates of target variables (Kleinn & Vilčko, 2006).

At sampling locations, methods for measuring grassland biophysical variables include visual assessment, clipping, use of rising plate meter (RPM) and field spectrometry (Ali *et al.*, 2016). Data retrieval by visual assessment is generally subjective, less accurate and lacks spatial and temporal continuity (Newnham *et al.*, 2010). A trained and experienced observer may achieve high accuracy in their estimates, however studies show as much as 50 % variation between observers (Anderson & Pearce, 2003; Millie & Adams, 1999). Clipping or cut-and-dry is a popular approach in grass vegetation assessment. This involves the harvest of grass samples which are then dried and weighed to get dry matter fraction. Clipping offers a very accurate measurement of grass biomass and is widely used in laboratory assessment of grass nutrient status and quality (Ali *et al.*, 2016). The method is however destructive, laborious and costly. The use of rising plate meters such as disc pasture meter (DPM) offers an efficient non-destructive alternative to clipping with comparable accuracy in estimating biomass as well as grass height (Hakl *et al.*, 2012). Another ground-based method is the use of spectrometers held at waist level above the surface. Field spectrometry can be used to record grassland biophysical variables such as above ground biomass, leaf area index and species composition based on reflectance spectra (Ali *et al.*, 2016; Flynn *et al.*, 2008). The recorded reflectance spectra however, need to be calibrated against in-situ samples. Flynn *et al.* (2008) found a good correlation between spectrometer-derived normalized difference vegetation index (NDVI) and biomass measurements from clipping ( $r^2 = 0.68$ ) and rising plate meter ( $r^2 = 0.54$ ).

While ground-based methods provides detailed and reliable information (also as input for remote sensing methods) for grassland vegetation monitoring, they are subjective, laborious, costly and are feasible only for small-scale assessment (Ali *et al.*, 2016; Xu *et al.*, 2008).

### 1.6.2 Remote Sensing Methods

Recent technological developments in aerial and space-borne sensors enables efficient, cost-effective and statistically robust wide area data collection for remote grassland vegetation monitoring. Using remote sensing methods, target variables of grassland vegetation can be monitored at spatial scales ranging from 25 cm to 1 km (Ali *et al.*, 2016), for satellite-based sensors, and can be finer in sensors aboard aerial platforms or drones. Additionally, the high temporal depth of satellite remote sensing archives such as that from the Moderate Resolution Imaging Spectrometer (MODIS) and Landsat allows historical data collection and change analysis (see Table 1.1 for timeline of major optical Earth observation satellites). Currently, space-borne sensors record data either actively or passively using optical (i.e. visible and near infrared) and radar technology (i.e. microwaves). A notable exception is the recently launched Global Ecosystem Dynamics Investigation (GEDI) lidar sensor by National Aeronautics and Space Administration (NASA) (Patterson & Healey, 2015).

**Table 1.1:** Major optical civilian Earth observation satellites with global coverage.

Access policy	Satellite mission	Operational timeline	Pixel size	Revisit period	Source
Open access	Landsat 1-3	1972 - 1983	80 m	18 days	NASA (USA)
	Landsat 4-5	1982 - 2013	30 m	16 days	NASA (USA)
	Landsat 7	1999 - present	30 m	16 days	NASA (USA)
	Landsat 8	2013 - present	30 m	16 days	NASA (USA)
	Landsat 9	2021 - present	30 m	16 days	NASA (USA)
	ASTER	2000 - present	15 m	tasked	NASA (USA)
	MODIS	2000 - present	250 - 1000 m	1 - 2 days	NASA (USA)
	Sentinel 2	2015 - present	10 m	5 days	ESA (Europe)
Commercial	SPOT 1-7	1986 - present	6 - 20 m	tasked	CNES/AIRBUS (France)
	IKONOS	2000 - 2015	1 - 4 m	tasked	Digital Globe
	QuickBird	2001 - 2015	2.4 m	tasked	Digital Globe
	WorldView 1-4	2007 - present	2 m	tasked	Digital Globe
	RapidEye	2009 - 2020	5 m	tasked	Planet Labs
	GeoEye 1	2008 - present	2 m	tasked	Digital Globe
	Pleiades	2011 - present	2 m	tasked	CNES/AIRBUS (France)
	KOMPSAT	1999 - present	2.8 m	tasked	KARI (Korea)
	SkySat	2013 - present	1 m	tasked	Planet Labs
	ALOS	2006 - 2011	10 m	tasked	JAXA (Japan)
	Gaofen	2013 - present	3.2 m	tasked	CNSA (China)

## **Optical Satellite Imagery**

Monitoring of grasslands with optical sensors typically combine original bands, texture features and spectral indices to quantify biophysical vegetation features (Ali *et al.*, 2016). To get meaningful estimates of vegetation variables, spectral features are fitted to field observations of the target variable using predictive models (Xie *et al.*, 2008). Depending on the strength of the correlation between field measurements and remote sensing features, the models can be applied to precisely estimate target variables over different spatial and temporal scales. He *et al.* (2009) developed methods to retrieve biophysical properties of vegetation using ground-based LAI measurements and satellite-based vegetation indices. In doing so, the authors built linear regression models between two satellite-derived vegetation indices (i.e. renormalized difference vegetation index (RDVI) and the adjusted transformed soil-adjusted vegetation index (ATSAVI)) and the in-situ LAI dataset. The results showed a significant relationship ( $p < 0.01$ ) between predicted LAI and independent in-situ LAI values. In a related study, Huang *et al.* (2013) found exponential function as optimal for describing relationships between Landsat and MODIS vegetation indices and soil respiration in the Tibetan alpine grasslands. Similarly, Xu *et al.* (2007) used exponential models to estimate grass production using MODIS NDVI and field data. Other studies explored optimal regression models (Long *et al.*, 2010), power regression (Xu *et al.*, 2008) and logarithmic regression models (Vescovo & Gianelle, 2008) for spatio-temporal estimation of grassland biophysical variables using satellite derived vegetation indices.

Multispectral sensors that are designed to record in a higher number of wavelength bands provide details for more robust characterization of grassland vegetation (Schuster *et al.*, 2015). For example, in addition to the visible and near infrared bands, yellow and red-edge bands in WorldView-3 imagery, provides the capability of reliably detecting ripening or dying plants and foliar chlorophyll content respectively. Phenological stages of grasslands can therefore be monitored (Schuster *et al.*, 2015). Aside enhanced spectral resolution, increased temporal availability of imagery is vital for monitoring grassland phenology which in turn could enhance accuracy of grass cover mapping (Huang & Geiger, 2008). Butterfield & Malmström (2009) demonstrated enhanced understanding of biomass-NDVI relationship when looking at different phenological stages.

Another aspect of grassland monitoring is categorical mapping which has a long history with the use of optical remote sensing data (Ali *et al.*, 2016). Different grass-dominated habitats such as pastures and rangelands have been successfully delineated through land cover classification of optical imagery. For example, Munyati & Makgale (2009) used multi-temporal Landsat data to map and quantify degraded areas in the Bahurutshe communal grazing lands in southern Africa. Different approaches have been explored to discriminate grassland habitat types, including the use of statistical, object-based and machine learning methods (Ali *et al.*, 2016). For a period, maximum likelihood classifier was the most widely used method. The maximum likelihood classifier assigns pixels to predefined classes based on their probability distribution (Otukei & Blaschke, 2010). Generally, statistical classifiers perform well on normally distributed datasets and in less complex spectral feature space. However the spectral similarity of different grassland vegetation types limits the performance of these classification methods. Unlike statistical classifiers, machine learning methods make no statistical assumptions regarding data distribution and are very efficient in non-linear classification problems (Abburu & Golla, 2015). Otukei & Blaschke (2010) compared the performance of maximum likelihood, support vector machine and decision trees classifiers for land cover mapping, and found that the machine learning algorithms gave more accurate results. Generally, object oriented approaches are preferred in mapping grassland habitats due to their ability to deal with spectral confusion and return land cover information at the scale of real-world objects (Ali *et al.*, 2016; Aplin & Smith, 2011). Object-based classification has also been proven superior particularly at very high image spatial resolution. Maxwell & Warner (2015) used object-based machine learning classification to retrieve mine-replaced grasslands with above 80 % user's and producer's accuracy. Xu *et al.* (2018), showed that a combination of multitemporal images and auxiliary data (i.e. digital elevation model (DEM) and NDVI) improved grassland type separability using object-based classification.

## 1.7 Remote Sensing Time-series Change Detection

Accurate detection and identification of habitat changes provides valuable information for conservation planning and management and allows general insights into key drivers of changes. Different methods exist for the detection of abrupt and subtle changes in land cover using time-series remote sensing data, which could be adopted to quantify changes in grassland habitats. A popular approach is the comparison of bi-temporal imagery based on simple arithmetic operations such as subtraction or division. The results of such operations depicts radiometric differences which could be linked to the magnitude of changes on the ground (Tewkesbury *et al.*, 2015). A common practice is the transformation of image spectra into vegetation indices prior to the application of arithmetic operations. For example, Liu *et al.* (2018) identified vegetation losses in aerial imagery using NDVI image ratios, while Gandhi *et al.* (2015) quantified changes in vegetation cover by differencing time-series NDVI. Further to this, arithmetic comparisons can be enhanced beyond radiometric differences by considering other contextual information such as texture and morphology (Tewkesbury *et al.*, 2015). For example, Klaric *et al.* (2013) developed an automated change detection method for very high-resolution satellite imagery based on a combination of neighborhood spectral, textural and morphological features. In a similar approach using Quickbird imagery, Falco *et al.* (2010) used morphological attribute profiles as a basis for change detection, showing that temporal changes can be detected outside of spectral differences.

Another established and widely used change detection approach is post-classification change detection which involves the comparison of thematic maps from different time periods. Peterson *et al.* (2004) compared historical and more recent land cover maps alongside an analysis of transition matrix to detect the magnitude and direction of land cover change. Post-classification techniques provide a thematic context and can be used to locate changes in target land cover types. For example, when comparing multi-temporal Landsat thematic maps, Basnet & Vodacek (2015) identified forest cover losses of 216.4 and 130.5 thousand hectares in central Africa, between 1988-2001 and 2001-2011 respectively. The advantage of this method is that no radiometric normalization is required since both images are processed separately (Coppin *et al.*, 2004; Tewkesbury *et al.*, 2015). However issues related to map production such as misclassification or classification errors present a major constraint to post-classification change detection. As such, the quality of input maps must be considered if any meaningful change results are to be realized (Coppin *et al.*, 2004).

Direct classification approach, where multi-temporal images are stacked and treated as one set of features in a classification process has also been used in detecting changes in land cover (Tewkesbury *et al.*, 2015). The method involves only one step of image classification which is advantageous in terms of processing time and can be particularly efficient for targeted land cover classes as demonstrated by Chehata *et al.* (2011) in an object-based forest change detection study. Other change detection methods include transformation by principal component analysis (PCA) (Deng *et al.*, 2008) or multivariate alteration detection (MAD) (Doxani *et al.*, 2012), change vector analysis (CVA) (Bovolo & Bruzzone, 2007) and hybrid methods (combination of methods) (Lu *et al.*, 2004).

A recent trend in time series-based change detection is the transition from the conventional endpoint approach (i.e., comparing temporal image pairs several years apart) to continuous monitoring with dense time-series imagery (annual, sub-annual and seasonal) (Woodcock *et al.*, 2020). This has been shown to allow more precise characterization of the timing and drivers of change (Arévalo *et al.*, 2020). Such approaches typically use break detection algorithms which monitor changes in residuals of forecasts from statistical models fitted to time series surface reflectance data (Holden *et al.*, 2016; Zhu & Woodcock, 2014). Popular continuous change detection algorithms include the Landsat-based detection of Trends in Disturbance and Recovery (LandTrendr) (Kennedy *et al.*, 2010) and the Continuous Change Detection and Classification (CCDC) algorithms (Zhu & Woodcock, 2014). An important challenge of such approaches is the computational resources required for preprocessing and dense time series analysis. However, the availability of cloud-based processing platforms such as Google Earth Engine has become instrumental to their implementation for land change monitoring.

## 1.8 Research Focus

Grazing lawn demographics influence fire frequency, intensity and spread; and impact herbivore densities, movement, reproductive success (Hempson *et al.*, 2015; Owen-Smith, 2004) and recruitment rate (Gaillard *et al.*, 2000). Consequently, variation in the area covered by these grass communities have a profound impact on broader savannah ecosystem functioning, including, nutrient cycling, plant community composition and predation; as well as conservation efforts supported through game viewing. Yet, there is limited understanding of grazing lawn distribution and of spatial and temporal

changes in structure. More importantly, we lack a robust, broad-scale mechanism for identifying and monitoring grazing lawn vegetation to enable investigation of distribution and structural changes, their response to broader environmental change over space and time, and the implications for overall ecosystem dynamics.

A more comprehensive broad-scale analysis of dynamics in grazing lawns and key drivers of their distribution pattern is vital to build an enhanced understanding of the constraints on grassland habitat distributions and heterogeneity in African savannahs. Specifically, a remote sensing-based analysis spanning the pathways through which proposed factors like fire, grazing, resource hot-spots and global change interact to regulate grazing lawn dynamics is necessary to provide spatially informed indicators for managing habitat heterogeneity in the context of biodiversity conservation. Moreover, such information will aid adoption of effective management strategies for maintaining an optimal resource ratio (i.e. grazing lawns to bunch grass ratio) necessary to support high grazer densities especially in a range land context.

### **1.8.1 Aim and Objectives**

The overall aim of this research is to map grazing lawns and investigate dynamics in grazing lawn cover and structure in southern African savannahs using satellite remote sensing. The three main objectives are to:

1. Develop methods for accurate detection and mapping of savannah grazing lawn distribution using high-resolution satellite imagery and machine learning techniques. The research questions addressed under this objective are:
  - How do current state-of-the art machine learning algorithms compare in discrimination of grazing lawns from other savannah vegetation components using high-resolution WorldView-3 imagery?
  - How can the spatial and spectral advantages of current open access satellite imagery be harnessed for accurate high resolution mapping of grazing lawns in order to ensure cost effective operational monitoring?
2. Identify changes in cover and structural distribution of savannah grazing lawns over space and time. The research questions addressed under this objective are:
  - What are the key transition pathways between grazing lawn cover and other savannah habitat states?
  - How does grazing lawn structure (spatial configuration) evolve under different landscape conditions?

- How does drought stress impact changes in grazing lawn cover and structure?
- 3. Identify the dominant drivers of change in grazing lawn cover and structure. The research questions addressed under this objective are:
  - What are the key factors driving grazing lawn dynamics?
  - How do water points as resource hot-spots in the savannah landscape shape spatial distribution of grazing lawns.

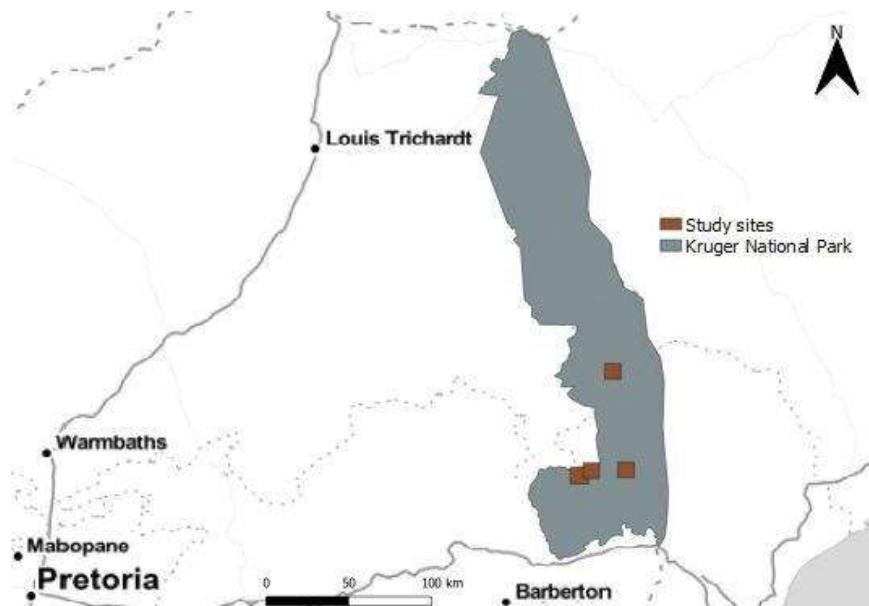
## 1.9 Study Area and Fieldwork Protocol

### 1.9.1 Study Area

Figure 1.1 shows locations of the selected study sites in KNP used in this research. Kruger National Park (KNP) lies between  $30^{\circ}53'18''E$ ,  $22^{\circ}19'40''S$  and  $32^{\circ}01'59''E$ ,  $25^{\circ}31'44''S$  in South Africa. The park spans approximately  $20,000\text{ km}^2$ , extending about 360 km from north to south. The significant diversity in the landscape is expressed in its climate, soil, flora and fauna, making it a globally important site for ecological studies. The climate is subtropical with maximum mean annual rainfall between 500 mm and 700 mm in the northern and southern parts of the park, respectively (Venter *et al.*, 2003). Geologically, KNP is divided into granitic soils to the west and basaltic soils to the east, which are separated by a narrow band of shale from the south to the mid portion of the park and rhyolite on the eastern extreme (van Wilgen *et al.*, 2014). This, coupled with spatial and temporal rainfall gradients, as well as disturbance events, exert enormous influence on vegetation type distribution across the landscape (Cromsigt & Beest, 2014). More open, productive grasslands occur on the basalt, while denser bushland savannah occupy the granite. Mopane (*Colophospermum mopane*), red bushwillow (*Combretum apiculatum*) and silver clusterleaf (*Terminalia sericea*) constitute some of the dominant vegetation types in the northern half of the park (Venter *et al.*, 2003). The open grasslands of the eastern plain are dominated by species like blue buffalo grass (*Cenchrus ciliaris*), red grass (*Themeda triandra*), stinking grass (*Bothriochloa radicans*) and finger grass (*Digitaria eriantha*), dotted with knob-thorn acacia (*Acacia nigrescens*) and marula trees (*Sclerocarya birrea*) (Venter, 1990). Mixed broadleaf woodlands of bushwillow (*Combretum sp*) with corridors of grassland cover the central-western part of KNP, while thorn thickets (e.g., *Acacia robusta*), silver clusterleaf (*Terminalia*

*sericea*) and sour grasses (*Hyparrhenia filipendula*) form a dominant part of the higher rainfall southern landscape (Venter, 1990). Alongside variations in abiotic factors which influences vegetation type distribution within the KNP landscape, the presence of a great diversity of herbivores exert significant impact on vegetation structure. For example, high population density of the African elephant (*Loxodonta africana*) has been suggested to be the major driver of woody vegetation change in KNP Marston *et al.* (2017); Munyati & Sinthumule (2016). The dominant grass consumers (with >50% of grass in diet) includes impala (*Aepyceros melampus*), blue wildebeest (*Connochaetes taurinus*), zebra (*Equus quagga*), buffalo (*Syncerus caffer*) and white rhino (*Ceratotherium simum*) (Kleynhans *et al.*, 2011).

Management of KNP is generally focused on maintaining habitat heterogeneity through adaptive fire management regimes (van Wilgen *et al.*, 2014) alongside natural fire events. Natural burns vary in frequency and intensity depending on rainfall patterns and the prevalence of high grass biomass (Govender *et al.*, 2006; Venter *et al.*, 2003). The consequence of varying fire regimes is the different spatial configurations of grass productivity and biomass accumulation (Govender *et al.*, 2006). For example, the distribution of short grass grazing lawns whose persistence depends on positive feedback loops associated with frequent grazing has been observed to be highly influenced by variation in burn size and frequency (Archibald *et al.*, 2005; Donaldson *et al.*, 2018).

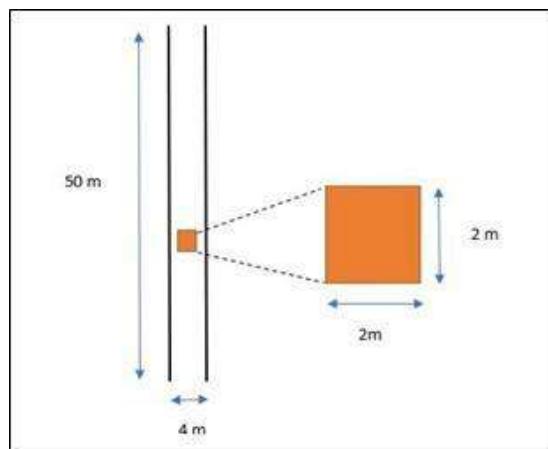


**Figure 1.1:** Location of study sites in Kruger National Park, South Africa.

## 1.9.2 In-situ Data Collection

### Sampling and plot design

In-situ field observations and measurements were recorded at sampled locations distributed within the study sites based on a systematic sampling approach. To avoid the influence of edge effects on recorded data, sample points were selected in areas 100 meters away from access roads. At each sample point, a 50x4 meters south-facing transect with 2x2 meters quadrats placed at every 10-meter distance along the length of the transect was installed for data collection. An illustration of the plot design used for data collection is provided in Figure 1.2.



**Figure 1.2:** Illustration of plot design for data collection.

### Data collection procedure

Data recording was completed for each sample plot according to the following steps:

1. Navigate to the sample point using the GPS device, record the plot number, point coordinate and take photographs of the area in the north, east, south and west cardinal directions.
2. Roll out the two graduated 50 m ropes and set their starting pegs at a separating distance of 4m in the eastern direction. Stretch out the ropes while walking for 50m in the southern direction and install their ending pegs 4m apart. Record the coordinate and take photographs of the area as in step 1.

3. Walk the 50 x 4 m transect back towards the start point while recording tree attributes (DBH, crown diameter, height of tallest tree), shrub attributes (base diameter of main shoot, crown diameter, number of shoots/branches, height of tallest shrub) and number of burnt points (at every 2m marking on ropes) along each rope.
4. Walk the 50 x 4 m transect again towards the end-point and record herbivore dung density by counting dung droppings per herbivore species.
5. Take the last walk along the 50 x 4 m transect towards the start point, place the 2x2 m quadrat at every 10 m along the transect and record fractions of bunch grass, short grass and forbs as well as the maximum and minimum grass heights in each quadrat.
6. Record the dominant land cover type.

## **1.10 Thesis Outline**

This thesis is organized into five chapters. This includes an introductory chapter; three analysis chapters, each of which addresses one or more of the research objectives; and a synthesis and conclusion chapter.

### **1.10.1 Chapter 1: General Introduction**

Chapter one introduces the overarching theme investigated in this research. It provides general background information and a review of literature relevant to the research objectives. General contextual information is provided on African savannahs; drivers of vegetation structure; and impacts of global change on savannah ecosystem structure and functioning. Background information on grazing lawns as functional components of African savannahs is provided, including dynamics in the development and persistence and how that influences savannah ecosystem processes. Grassland vegetation monitoring approaches and remote sensing in particular are introduced as valuable tools for grazing lawn monitoring. The primary focus of this research including the problem addressed, justification, aim, objectives and research questions are outlined. Finally, a description of the study area and the field data collection protocol used are provided.

### **1.10.2 Chapter 2: Probabilistic Mapping and Spatial Pattern Analysis of Grazing Lawns in Southern African Savannahs Using WorldView-3 Imagery and Machine Learning Techniques**

Chapter two is published in *MDPI Remote Sensing Journal*. The chapter explores machine learning methods for accurate detection of grazing lawns in heterogeneous semi-arid and mesic savannah landscapes and provides an analysis of spatial variations in grazing lawn occurrence and structure. Using spectral and texture features derived from WorldView-3 imagery, four machine learning algorithms are parameterised and assessed for general discrimination of plant functional types; the models are subsequently compared for grazing lawn detection. Spatial metrics are used to analyse spatial patterns in grazing lawn distribution along a gradient of distance from waterbodies. Chapter two addresses the first research question of objective one and the second research question of objective three.

### **1.10.3 Chapter 3: Spatio-temporal Analysis of Grazing Lawn Dynamics in Southern African Savannahs Using Multi-temporal High-resolution Satellite Images**

Chapter three uses a post-classification change detection approach to identify long-term changes in grazing lawn cover and structure and the dominant drivers of change. Additionally, it includes a specific analysis of the influence of protracted drought events on grazing lawn distribution. Chapter three addresses all the research questions under objective two and the first research question under objective three.

### **1.10.4 Chapter 4: Multi-sensor Optical Image Fusion for Land Cover Classification in a Heterogeneous African Savannah: Toward Accurate and Cost Effective grazing Lawn Monitoring**

The fourth chapter explores the utility of open-access Sentinel-2 imagery and Planet mosaics for accurate discrimination of grazing lawns from other vegetation components at high spatial scales through an effective image fusion approach. Gram-Schmidt transformation is used to fuse Sentinel-2 and Planet images. The datasets (i.e. Sentinel-

2, Planet and Fused images) augmented with two spectral indices and three Haralick texture features are compared in a multi-layer perceptron neural network classification for grazing lawn detection. Chapter four addresses the second research question under objective one.

### **1.10.5 Chapter 5: Synthesis and Conclusion**

Finally, chapter five summarises and consolidates the research findings for each objective in the context of existing literature. Additionally, chapter five provides commentary on the contributions and impact of this research, research limitations, an outlook for future research and overall concluding remarks.

---

# Chapter 2

## Probabilistic Mapping and Spatial Pattern Analysis of Grazing Lawns in Southern African Savannahs Using WorldView-3 Imagery and Machine Learning Techniques

---

### 2.1 Introduction

Savannah ecosystems inherently exhibit a considerable degree of variability in structural and physical attributes across their range of occurrence (Sankaran *et al.*, 2008). In Southern Africa, they feature the coexistence of grasses and an overstorey layer of trees with varying gradients of dominance and spatial formations (Shorrocks & Bates, 2015). Within the grassy layer, plant forms are typified structurally by tall bunch grasses and short grass grazing lawns (Cromsigt & Kuijper, 2011), which form a significant component of the heterogeneity in Southern African savannah grasslands (Cromsigt & Olff, 2008).

The relative proportions and distribution of grazing lawns and tall bunch grass resources have been directly linked to important ecosystem changes such as fluctuations in herbivore density (Cromsigt & Beest, 2014; Owen-Smith, 1987) and changing fire regimes (Gill *et al.*, 2009; Waldram *et al.*, 2008). For example, the amount of high-quality lawn grasses has been suggested to be the primary natural limiting factor to population size of mega-herbivores such as the white rhinoceros (*Ceratotherium simum*) and the hippopotamus (*Hippopotamus amphibius*) (Cromsigt & Beest, 2014; Owen-Smith, 1992; Waldram *et al.*, 2008). Additionally, the persist-

ence of grazing lawns creates natural barriers to the spread of fire due to limited above ground fuel biomass that may serve as fuel for the spread of fire (Donaldson *et al.*, 2018; Hempson *et al.*, 2015; Waldram *et al.*, 2008). By contrast, tall bunch grasses keep their moribund growth forms and increase savannah grassland fuel load (Archibald *et al.*, 2005; Waldram *et al.*, 2008). As such, changes in grazing lawn coverage and distribution could potentially alter the size, frequency and intensity of fire within the landscape (Waldram *et al.*, 2008), with cascading effects for nutrient cycling, plant community composition, habitat structure and biodiversity. Monitoring the occurrence and spatial patterns of grazing lawns is therefore fundamental to understanding the ecology of these vital grassland systems.

Grazing lawns are dynamic and maintained by constant grazing, resulting in a feedback of dense nutrient-rich plant growth that in turn attracts more grazing (Hempson *et al.*, 2015). Different parts of a landscape may also be predisposed to grazing lawn formation due to localised availability of resources and nutrient hotspots that concentrate grazers. These include areas around water bodies and areas of mineral accumulation (e.g., sodium) (Hempson *et al.*, 2015). The initiation of a cycle of regular grazing thus appears to be a critical factor for their development and persistence (Archibald *et al.*, 2005; Donaldson *et al.*, 2018; Hempson *et al.*, 2015). Nonetheless, the rates and specific pathways of their development likely depend on factors such as rainfall, fire and soil types (Archibald, 2008). Rainfall has a strong influence on the rate of grass biomass accumulation and the height of tall grass stands (Veldhuis *et al.*, 2016). The relative proportion of grazing lawns to high biomass tall grasses is also influenced by dynamics in soil nutrients through their strong influence on grass productivity. Under high rainfall and soil nutrient conditions, increased grazing frequency is required to prevent the invasion of tall-grass competitors (Hempson *et al.*, 2015). Too infrequent grazing increases the vulnerability of a switch to tall bunch grasses (Donaldson *et al.*, 2018). Fire also consumes grass biomass and has the potential to shift grass community composition and structure within different environmental constraints (Leonard *et al.*, 2010). Tall bunch grasses with low forage quality dominate fire-driven grassy systems (Archibald *et al.*, 2005; Donaldson *et al.*, 2018; Hempson *et al.*, 2015). Additionally, post-fire regrowth can also attract grazers away from previously established grazing lawns causing them to be invaded by tall bunch grasses (Archibald *et al.*, 2005; Donaldson *et al.*, 2018). The varying spatial and temporal nature of the key interacting factors that drive grazing lawn dynamics urges for a robust landscape-scale approach to better understand their variation over space and time.

Ground-based monitoring of grazing lawn responses to the complex top-down and bottom-up ecological processes is challenging due to the large areas involved. Although ground-based methods can provide more detailed and valuable local insights (Helman *et al.*, 2015), they are not efficient in capturing regional-scale dynamics—due to the high cost involved—nor do they provide any retrospective information beyond the start of monitoring activities (Marston *et al.*, 2017). Remote sensing technology is able to overcome the spatial and temporal limitations, which in combination with ground-based observations, offers valuable tools for accurate, efficient and cost-effective ways for vegetation monitoring (Jensen, 2015; Khorram *et al.*, 2016).

Medium resolution satellite imagery such as the Landsat Operational Land Imager (OLI) (Wulder *et al.*, 2012) and Sentinel-2 missions (Drusch *et al.*, 2012) are freely available, with extensive temporal coverage which offers enormous benefits for monitoring vegetation dynamics. Further, recent advances in very high spatial resolution (VHR) satellite imagery such as WorldView-3 presents opportunities to partially overcome limitations in spatial resolution associated with medium resolution imagery, particularly in heterogeneous savannah landscapes (Marston *et al.*, 2017). At nadir spatial resolution of 1.24 m (Vajsova *et al.*, 2017), the WorldView-3 sensor is able to identify and discriminate between different sized vegetation components such as trees, shrubs and grass patches (Kaszta *et al.*, 2016; Marston *et al.*, 2017). Additionally, the yellow, red-edge and two near-infrared bands in WorldView-3 imagery provide the capability of reliably detecting photosynthetically active or dying plants and foliar chlorophyll content (Schuster *et al.*, 2015). As such, various phenological stages of vegetation can be monitored, which is instrumental in dealing with spectral similarity of different savannah vegetation composition (Kaszta *et al.*, 2016; Whiteside *et al.*, 2011).

In parallel with advances in remote sensing imaging technology, free open source software packages and increased computational power have been developed to facilitate image analysis. The combination of these factors has advanced the use of machine learning algorithms in land cover classification (Abdi, 2020). Among the most popular machine learning classification algorithms are Random Forest (RF) (Breiman, 2001), Support Vector Machines (SVM) (Cortes & Vapnik, 1995), Decision Trees (DT) (Loh, 2011) and Artificial Neural Network (ANN) (Del Frate *et al.*, 2007). RF, SVM, DT and ANN are nonparametric classifiers and are very efficient in dealing with nonlinear classification problems (Abdi, 2020; Camargo *et al.*, 2019), having been proven to be effective in different savannah ecosystems. Camargo *et al.* (2019)

demonstrated the utility of RF, SVM, DT and ANN in classifying land cover in the Brazilian Tropical Savannah biome. RF is well known for its flexibility of application on both continuous and categorical datasets, either as a regression or classification algorithm, respectively (Abdi, 2020; Breiman, 2001). Symeonakis *et al.* (2018) used RF to classify different land cover types in southern African savannahs and reported a maximum accuracy of 91.1%. The SVM classifier is popular for its strong ability to generalise in complex nonlinear feature space (Cortes & Vapnik, 1995). SVM was used in seasonal separation of vegetation components in southern African savannahs and gave the highest accuracy score under dry leaf-off conditions compared to k-Nearest Neighbour (k-NN), Maximum Likelihood Classifier (MLC), RF and DT classifiers (Kaszta *et al.*, 2016). The application of DT in an eastern African savannah resulted in an increased mapping accuracy over MLC and SVM, with over 93% overall accuracy (Otukei & Blaschke, 2010). ANN classifiers have been widely used in satellite image classification, due to the capability to adapt and generalise different input data structures (Abdi, 2020). The successful application of ANN has been well demonstrated in different remote sensing contexts, including classification of endangered tree species (Omer *et al.*, 2015) and dynamic modelling of land cover changes in semi-arid landscapes (Silva *et al.*, 2020).

Much of the literature on monitoring grazing lawn dynamics in southern African savannahs focuses on localised and controlled experimental studies of responses to the mechanisms that induce their establishment and persistence (Archibald *et al.*, 2005; Cromsigt & Olff, 2008; Donaldson *et al.*, 2018; Hempson *et al.*, 2015). There is limited evidence on whether the proposed pathways translate into broad-scale spatial patterns in grazing lawn occurrence. Among the few empirical studies is the work of Archibald *et al.* (2005) who mapped grass structural distribution and found that the extent of grazing lawns was directly related to fire return interval. Though there is substantial information on how different biotic and abiotic factors shape grazing lawns, knowledge of their present cover and distribution, which is critical to understanding habitat heterogeneity, is lacking. More importantly, there is no robust, broad-scale approach for detecting and monitoring grazing lawns to enable comprehensive investigation into their dynamics over space and time, and the implications for broader ecosystem dynamics. Against this backdrop, this study seeks to develop a robust machine learning framework for mapping grazing lawns in southern African savannahs by (i) parameterising and assessing the quality of Random Forest (RF), Support VectorMachines (SVM), Multilayer Perceptron (MLP) and Classification and Regression Trees (CART) models for savannah land cover classification in a localised context,

and (ii) comparing model performance for probabilistic mapping of grazing lawns on a wider scale. Additionally, spatial patterns in grazing lawn distribution along a gradient of proximity to water bodies, which has been hypothesised to influence grassland spatial structure (Hempson *et al.*, 2015; Smit & Archibald, 2019) is analysed.

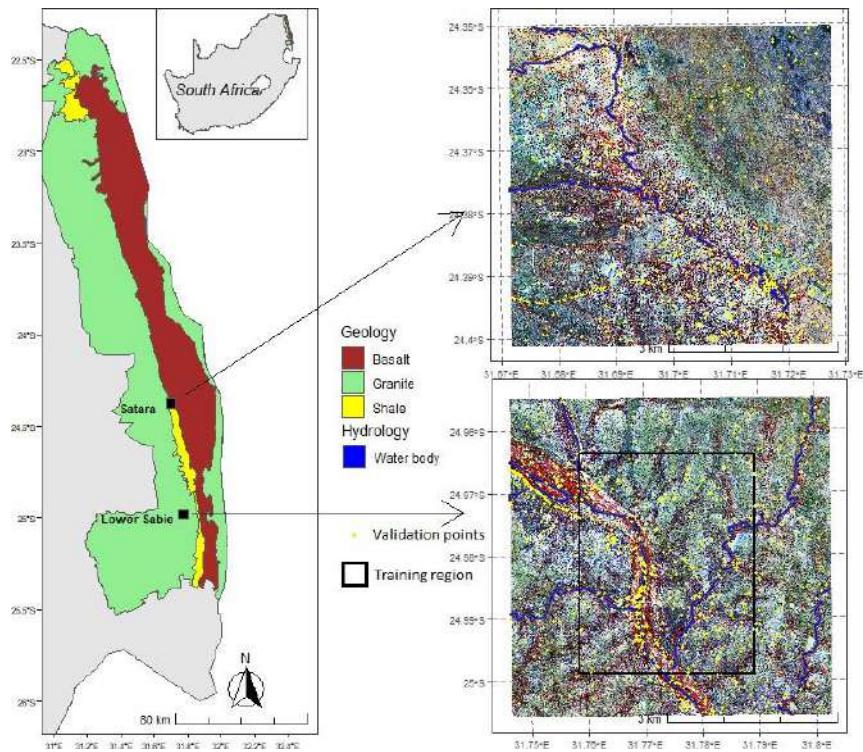
## 2.2 Materials and Methods

### 2.2.1 Study Sites

Two study sites located within the Satara and the Lower Sabie regions of the park (Figure 2.1), each covering 5.7 x 5.7 km, and extending over a range of habitat conditions including rainfall, geology and vegetation type were used. The Satara site is a well-studied grazing system close to the latitudinal center of KNP, and covers both granitic and basaltic soil types interspersed with a strip of ecca shales. The landscape is semiarid with mean annual rainfall of 400-500 mm. The granite areas to the west are generally more wooded and undulating than the flat and more open and grassy basaltic plains. In contrast, the Lower Sabie site falls under mesic landscape conditions with mean annual rainfall of 600-700 mm. The area has an underlying granite geology and encompasses portions of the Sabie River catchment. A number of sodic sites are also present within the Lower Sabie study site. Sodic sites typically occur at footslopes of catenas and are known to have high soil and vegetation sodium content which concentrates grazers and aids the formation and maintenance of continuous grazing lawn patches (Hempson *et al.*, 2015).

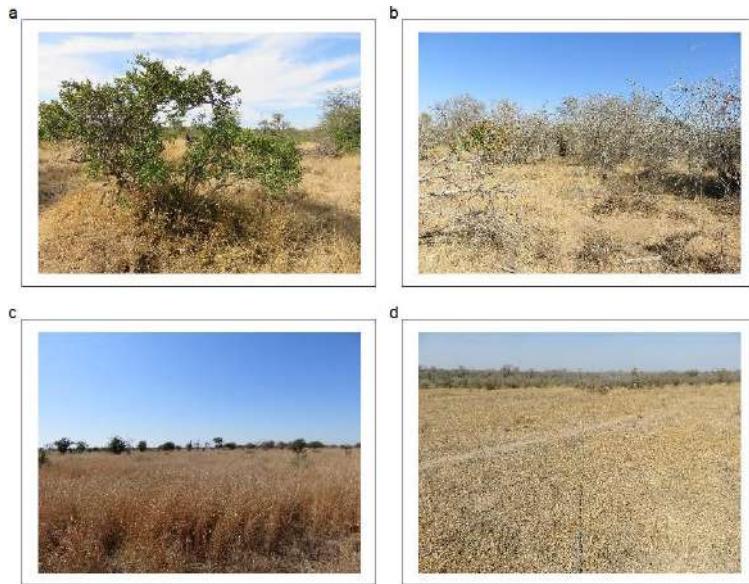
### 2.2.2 Land Cover and Classification Scheme

The study sites were selected to exclude as much anthropogenic influence as possible due to the natural ecological focus of the study. Thus, natural and semi-natural land cover features such as vegetation patches, bare soil surfaces and waterbodies dominate the selected study sites, with the only artificial surfaces being roads and isolated structures which serve as rest stops and picnic sites for tourists. Four plant functional types (PFTs) were identified in order to distinguish grazing lawns from other vegetation types (Figure 2.2). These included evergreen woody components, deciduous woody components, bunch grasses and short grass grazing lawns.



**Figure 2.1:** Map of study area showing locations of the Satara and Lower Sabie field sites in Kruger National Park (KNP), with enlarged views of the WorldView-3 satellite scenes (False colour: NIR1, R, G) overlaid with hydrology. Inset map shows the location of KNP within South Africa. Geological information was obtained from South African National Parks data repository

The PFT categories were finalised following the vegetation nomenclature provided in Marston *et al.* (2017) and were modified based on knowledge from dry season field survey and the spectral reflectance properties of the different vegetation components contained within satellite imagery. Within the landscape, woody components are mainly trees and shrubs, which in many cases were challenging to objectively differentiate. This is a well-known dilemma in savannah landscapes due to structural complexities such as multiple stems, varying disturbance adaptations and height limitations (Zizka *et al.*, 2014). The common practice has been to use arbitrary morphological traits like diameter and height thresholds depending on research objectives and ecological relevance. For example, Marston *et al.* (2017) used a main trunk diameter threshold of 7 cm to distinguish between trees and shrubs, where woody components with >7 cm diameter were classified as trees and those with <7 cm diameter as shrubs. Moreover, optical satellite imagery only records planar-view spectral reflectance information from surface cover with little structural detail. This presents a further challenge for successful differentiation of vegetation



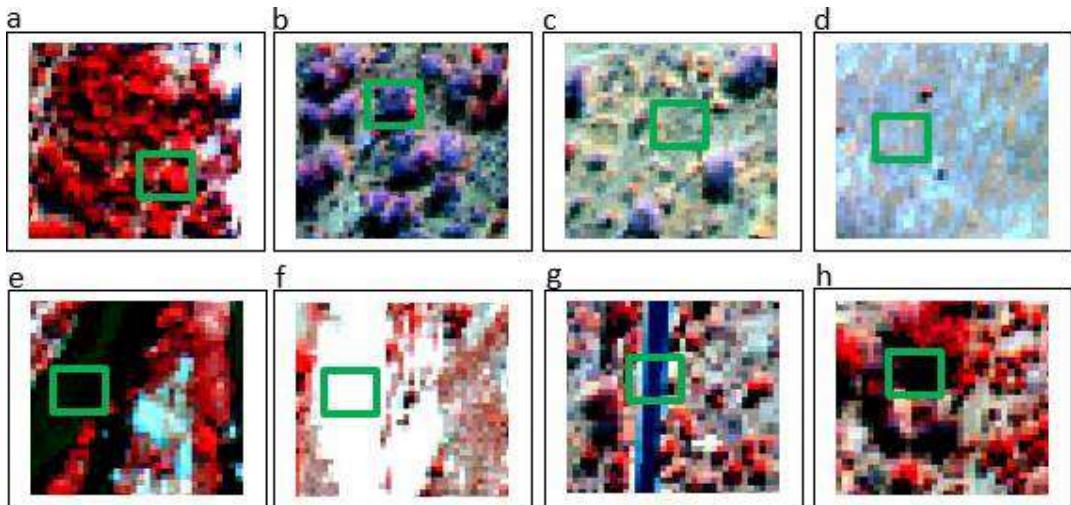
**Figure 2.2:** Examples of Plant Functional Types (PFTs): (a) woody evergreen, (b) woody deciduous, (c) bunch grass and (d) grazing lawns.

structure. The woody components were thus differentiated based on dry season phenological differences (leaf on/off), which allowed for objective delineation both in situ and on satellite imagery. Bunch grasses were identified as tall grass patches with height  $>20$  cm. In contrast, grazing lawns were identified as short grass areas with stoloniferous growth forms and height  $<20$  cm. Unlike bunch grasses which generally occurred as dense patches, grazing lawns within the study sites often had a sparse distribution and exhibit a relatively smooth texture in appearance, which aided visual interpretation of VHR satellite imagery. In addition to the PFTs, waterbodies, bare soil surfaces, built-up features and shadows on satellite imagery were identified, which constituted the classes used in this study (Figure 2.3). Table 2.1 provides a summary of the land cover classification nomenclature used.

### 2.2.3 Data

#### Satellite Imagery

Multispectral VHR imagery from the WorldView-3 satellite sensor was used in this study. Ortho-ready standard 8-band multispectral scenes (in UTM/WGS 84 projection) were acquired which had been processed to level 2A by the vendor (Table 2.2). The images were acquired in the dry season, on July 1, 2019 (at  $28.97^\circ$  Sun Azimuth and  $19.49^\circ$  off Nadir) and July 7, 2019 (at  $30.98^\circ$  Sun Azimuth and  $1.85^\circ$  off Nadir)



**Figure 2.3:** Sample display of land cover types from WorldView-3 satellite image scene (False colour: NIR1, R, G). (a) Woody evergreen; (b) Woody deciduous; (c) Bunch grass; (d) Grazing lawn; (e) Water body; (f) Bare; (g) Built-up; (h) Shadow.

for the Lower Sabie and Satara sites, respectively, under cloud-free conditions. Dry season imagery has previously been used to successfully discriminate vegetation types in similar contexts (Brandt *et al.*, 2017; Marston *et al.*, 2017). Apart from the reduced persistence of cloudy conditions in the dry season, which is an advantage to optical satellite remote sensing particularly in the tropics (Symeonakis *et al.*, 2018), spectral differentiation is maximized due to phenological differences among different vegetation types (Bucini *et al.*, 2009; Marston *et al.*, 2017). Image acquisition was timed to coincide with our field survey season (June 26–July 21, 2019), which allowed for the collection of consistent reference information for land cover classification and validation.

### Reference Data

Reference data on the different land cover types were generated from georeferenced field survey locations, and were extended via further interpretation of VHR images augmented by field photos and Google Earth satellite scenes. Overall, data from (i) 111 predefined field locations, systematically distributed within 200 m buffer beyond 100 m distance from access roads, and (ii) 5122 randomly distributed points from augmented visual interpretation, formed the reference data points (i.e., total of 5233 points) for training (3807–i.e., 73%–reference points) and validation (1426–i.e., 27%–reference points). Polygons of spectrally homogeneous areas were manually

**Table 2.1:** Description of land cover classification nomenclature and reference data. Numbers represent number of reference points, while figures in parenthesis represent area of training polygons in hectares. Lower Sabie and Satara validation points are separated by “/” (i.e., Lower Sabie/Satara).

ID	Name	Description	Land Cover		Reference Samples	
			Model Training	Map Validation		
1	Woody ever-green	Woody vegetation components that are adapted to retain their leaves all year round. Classified based on dry season field observations.	863 (3.94)		100 / 80	
2	Woody deciduous	Woody vegetation components that are adapted to retain their leaves in the wet season and shed them in the dry season. Classified based on dry season field observations.	1047 (3.26)		100 / 65	
3	Bunch grass	Tall grass patches with height >20 cm, and often occur as dense patches with upright growth form.	680 (10.12)		100 / 114	
4	Grazing lawn	Short grass patches with height <20 cm, and often occur in sparse distribution with stoloniferous growth form.	465 (7.99)		100 / 103	
5	Water body	Water bodies occurring within the landscapes including rivers, streams and reservoirs.	58 (3.18)		100 / 38	
6	Bare	Bare surfaces occurring as patches of exposed soil and includes dusty trails and rocky outcrops.	464 (4.29)		100 / 74	
7	Built-up	Built artificial structures within the landscape as well as asphalt and concrete coated surfaces such as roads and bridges.	37 (0.75)		100 / 64	
8	Shadow	Shadows of trees and other tall structures falling on adjacent surfaces which results in very dark or low brightness values.	193 (0.63)		100 / 88	

digitised and labelled according to land cover class IDs (see Table 2.1) using the locations of training points. The polygon extents (Table 2.1) were then used to extract image pixels for model training. Of the many potential approaches that could be used to extract training pixels, polygon objects have been shown to provide the most accurate classification outcomes (Corcoran *et al.*, 2015; Ma *et al.*, 2017). Spectral plots

**Table 2.2:** Description of datasets used.

Dataset	Description	Temporal Coverage	Source
WorldView-3 imagery	Multi-spectral 8-band satellite imagery with 1.24 m spatial resolution. Bands include: Coastal (C: 400-450 nm), Blue (B: 450-510 nm), Green (G: 510-580 nm), Yellow (Y: 585-625 nm), Red (R: 630-690 nm), Red Edge (RE: 705-745 nm), Near Infrared 1 (NIR1: 770-895 nm), Near Infrared 2 (NIR2: 860-1040 nm).	July 2019	European Space Imaging
Reference data	Input image pixels labeled according to land cover classification nomenclature. Pixels were extracted from reference polygon and point features.	June 2019 - July 2019	Georeferenced field survey locations; Field photos; and Google Earth and VHR scenes
Auxiliary data	OpenStreetMaps watercourses data sourced as line vector layer for streams and rivers, and polygon vector layer for reservoirs.	November 2019	Openstreetmap

of the different land cover classes were examined and reviewed alongside Jefferies-Matusita distance measures to ascertain adequate spectral separability prior to model training (Van Niel *et al.*, 2005). Models were parametrised and trained for prediction based on data from a sub-area within the Lower Sabie field site (see training region in Figure 2.1). This was necessary to ensure high model quality as a greater proportion of the georeferenced field sample locations was concentrated within the training region, while reducing computational cost. In contrast, map validation was conducted using site-specific reference data.

### Auxiliary Data

Multiple buffer distances (100 m divisions) from water source were used to analyse spatial pattern in grazing lawn distribution. Water points represent significant resources and important predictors of grazer movement (Smit *et al.*, 2007) and spatial heterogeneity in general within semi-arid savannah landscapes (Marston *et al.*, 2019). The data (Table 2.2) was downloaded from OpenStreetMaps surface water archive (streams, rivers and reservoirs) (Haklay & Weber, 2008) and was validated against

a drainage network and stream order data obtained from the Scientific services of South African National Parks. The OSM surface water layer contributed in November 2019 had the closest temporal coverage to the acquisition dates of satellite imagery and field data, and was selected for spatial analysis.

#### 2.2.4 Preparation of Image Features

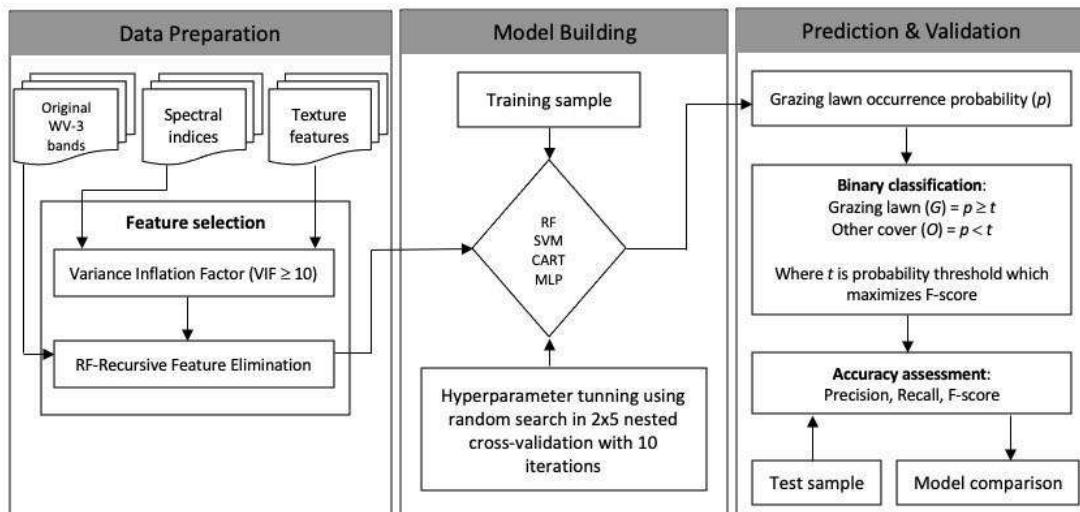
Following acquisition, a cubic convolution resampling approach was used to upsample images to 2 m spatial resolution, which is reflective of the average minimum patch size of short grass grazing lawns in southern African savannahs (Cromsigt & Beest, 2014; Waldram *et al.*, 2008). A series of spectral indices were calculated highlighting greenness, moisture and soil properties, in order to increase utility of the spectral information contained in the original image bands (Table 2.3). Greenness, moisture and soil indices are derived from arithmetic combination of spectral information recorded in visible and near-infrared image bands and exhibit high correlation with vegetation characteristics such as phenology (Balzarolo *et al.*, 2016; Liu *et al.*, 2017; Munyati *et al.*, 2020), biomass (Fajji *et al.*, 2017; Guerini Filho *et al.*, 2020; Yin *et al.*, 2018) and moisture content (Hunt Jr *et al.*, 2016; Roberto *et al.*, 2016). To complement the spectral information, spatial heterogeneity measures were calculated as a selection of simple and advanced Haralick texture features based on Gray Level Co-occurrence Matrix (GLCM) (Haralick *et al.*, 1973). The GLCM variables (Table 2.3) were calculated on the near-infrared band (NIR1) (see Table 2.2 for details on image bands), which contains valuable spectral information for differentiating vegetation characteristics. A probabilistic quantizer, with 32 quantization levels in a 3x3 moving window, at an offset distance of 1 pixel in all directions ( $0^\circ$ ,  $35^\circ$ ,  $90^\circ$  and  $135^\circ$ ) was used (Pratt, 2013; Symeonakis *et al.*, 2018). In total, 27 spectral indices and 18 texture features were processed using the Orfeo-Toolbox remote sensing image processing software (Inglada & Christophe, 2009) (Table 2.3). The spectral indices as well as texture features in combination with the original image bands served as input data in the machine learning models and analysis workflow (summarised in Figure 2.4). Incorporating spectral and textural image features is well known to enhance discrimination space for more accurate land cover mapping particularly in heterogeneous savannah landscapes (Johansen & Phinn, 2004; Paneque-Gálvez *et al.*, 2013; Symeonakis *et al.*, 2018).

**Table 2.3:** Initial image features serving as potential predictors.

Data (abbreviation)	Description
<b>Spectral features from individual bands (B):</b>	
B_C, B_B, B_G, B_Y, B_R, B_RE, B_NIR1, B_NIR2	Coastal, Blue, Green, Yellow, Red, Red Edge, Near Infrared-1, Near Infrared-2
<b>Spectral features from vegetation (V), moisture (M) and soil (S) indices:</b>	
V_NDVI, V_TNDVI, V_RVI, V_SAVI, V_TSABI, V_MSAVI, V_MSAVI2, V_GEMI, V_IPVI, V_LAI, M_NDWI, M_NDWI2, M_MNDWI, S_BI2, S_BI, S_CI, S_RI, S_NDSI, S_SI1, S_SI2, S_SI3, S_SI4, S_SI5, S_SI6, S_SI7, S_SI8, S_SI9	Normalized Difference Vegetation Index, Transformed Normalized Vegetation Index, Ratio Vegetation Index, Soil Adjusted Vegetation Index, Transformed Soil Adjusted Vegetation Index, Modified Soil Adjusted Vegetation Index, Modified Soil Adjusted Vegetation Index-2, Global Environment Monitoring Index, Infrared Percentage Vegetation Index, Leaf Area Index, Normalized Difference Water Index, Normalized Difference Water Index-2, Modified Normalized Difference Water Index, Brightness Index-2, Brightness Index, Color Index, Redness Index (Inglada & Christophe, 2009), Normalized Difference Salinity Index, Salinity Index-1, Salinity Index-2, Salinity Index-3, Salinity Index-4, Salinity Index-5, Salinity Index-6, Salinity Index-7, Salinity Index-8, Salinity Index-9 (Elhag, 2016)
<b>Haralick texture features (T):</b>	
T_Ener, T_Ent, T_Corr, T_IDM, T_Iner, T_CS, T_CP, T_HCorr, T_Mean, T_Var, T_Diss, T_SAvrg, T_SVar, T_SEnt, T_Dent, T_DVar, T_IC1, T_IC2	Energy, Entropy, Correlation, Inverse Distance Moment, Inertia, Cluster shade, Cluster prominence, Haralick correlation, Mean, Variance, Dissimilarity, Sum average, Sum variance, Sum entropy, Difference of Entropies, Difference of variances, Information correlation-1, Information correlation-2 (Inglada & Christophe, 2009)

## 2.2.5 Feature Selection

Remote sensing image features such as spectral indices (vegetation, moisture and soil indices) as well as texture variables tend to exhibit high levels of collinearity. Highly correlated features increases data redundancy and risk of overfitting, which could have adverse consequences for algorithm performance especially for high-dimensional datasets (Alonso *et al.*, 2011), a problem that results from the Hughes



**Figure 2.4:** Conceptual workflow showing steps in machine learning model development and evaluation towards grazing lawn detection.

phenomenon (Hughes, 1968). Although nonparametric machine learning algorithms are thought to be less susceptible to Hughes phenomenon, recent findings shows they benefit from dimensionality reduction nonetheless (Pal & Foody, 2010). In this study, the most robust predictors were targeted while reducing prohibitive computational efforts.

Image variables were selected by combining two procedures. First, collinearity check was conducted with the Variance Inflation Factor (VIF) using the “usdm” package (Naimi *et al.*, 2014) within R-programming environment (R Core Team, 2017). This was done separately for the spectral indices (i.e., vegetation, moisture and soil) and the Haralick texture features derived from the WorldView-3 imagery (see Table 2.3). VIF measures the degree to which predictor variables are correlated. For example, given  $k$  independent predictor variables, each variable is regressed with the remaining  $k - 1$  variables and coefficient of determination ( $R^2$ ) is estimated. The VIF of the dependent variable is thus computed as

$$VIF = \frac{1}{1 - R^2} \quad (2.1)$$

Large values of VIF implies a corresponding high degree of collinearity and vice versa. Following VIF analysis, correlated variables were subsequently removed by considering a stepwise elimination threshold of  $VIF \geq 10$  (Dormann *et al.*, 2013). The VIF assessment resulted in six spectral indices and 15 Haralick texture features being retained.

Second, the less correlated spectral indices and texture variables were combined with the original image bands to select final image feature subset using Random Forest-Recursive Feature Elimination (RF-RFE) (Granitto *et al.*, 2006; Guyon *et al.*, 2002). Recursive Feature Elimination (RFE) is an iterative process that uses some measure of feature importance to rank and select features by backward elimination (Guyon *et al.*, 2002). The technique basically builds a model with the entire feature set, computes an importance score for each feature, removes the least important features and repeats the process until a user-defined number of features subset is reached. Feature importance scores derived from random forest out-of-bag (OOB) error estimates were used for ranking features in the RF-RFE process. The final subset of features were then determined by analysing the relationship between number of features and accuracy scores derived from a stratified 10-fold cross-validation assessment. Overall, 26 WV-3 image features achieved optimal accuracy (See Figure A.1 and Figure A.2 in Appendix A). All steps in the RF-RFE process were implemented using Scikit-learn python library (Pedregosa *et al.*, 2011).

### 2.2.6 Machine Learning Algorithms

There is a proliferation of machine learning algorithms, which, coupled with the conflicting reports of their performance in remote sensing classification literature (Maxwell *et al.*, 2018), makes it challenging to select the optimal method for any specific application. The optimal classification algorithm is generally context-specific and in most cases depends on the landscape and classes mapped (Lawrence & Moran, 2015), parameter settings (Huang *et al.*, 2002; Maxwell *et al.*, 2018; Shi & Yang, 2016), nature of training data (Foody *et al.*, 2016; Ghimire *et al.*, 2012; Li *et al.*, 2014; Rodriguez-Galiano *et al.*, 2012) and data dimensionality (Maxwell *et al.*, 2018,1). Lawrence & Moran (2015) recommend prior experimentation with multiple classifiers to determine optimal performance.

For this study, four state-of-the-art nonparametric machine learning algorithms were tested including: Random Forest (RF) (Breiman, 2001), Support Vector Machines (SVM) (Cortes & Vapnik, 1995), Classification and Regression Trees (CART) (Loh, 2011) and Multilayer Perceptron (MLP) (Del Frate *et al.*, 2007). All have been shown to achieve high performance in many remote sensing applications, and in particular, land cover mapping (Maxwell *et al.*, 2018). Their superiority in handling complexity and high-dimensional data makes them ideal for application in highly heterogeneous savannah landscape conditions (Kaszta *et al.*, 2016). The selected algorithms were configured and implemented in the python programming environment using Scikit-learn python library (Pedregosa *et al.*, 2011). Optimal parameter values (Table 2.4) from hyperparameter tuning were used in each model. Summary descriptions of how the algorithms work are presented below.

**Table 2.4:** Results of optimal hyperparameter values used in each model, from 2x5 nested cross-validation using Scikit Learn python package. Refer to the work in (Pedregosa *et al.*, 2011) for more details on model hyperparameters.

Model	Optimal Hyper-parameter Value	Description
RF	n_estimators = 2000, max_features = 'auto', max_depth = 20, min_samples_split = 2, min_samples_leaf = 1	n_estimators = number of trees in the forest. max_features = number of features to consider for the split, 'auto' takes $\sqrt{No. \ features}$ . max_depth = maximum depth of the tree. min_samples_split = minimum number of samples required to split an internal node. min_samples_leaf = minimum number of samples required to be at a leaf node.
MLP	hidden_layer_sizes = (150,100,50), activation = 'logistic', solver = 'adam', max_iter = 100, alpha = 0.0000001	hidden_layer_sizes = number of neurons in each hidden layer (three layers in this case). activation = activation function of the hidden layer. solver = solver for weight optimization, 'adam' is based on the stochastic gradient optimizer. max_iter = maximum number of iterations. alpha = regularization parameter.
CART	criterion = 'gini', max_depth = 80, min_samples_split = 20, min_samples_leaf = 5	criterion = function to measure quality of split. max_depth = maximum depth of tree. min_samples_split = minimum number of samples required to split an internal node. min_samples_leaf = minimum number of samples required to be at a leaf node.
SVM	C = 1000, gamma = 0.001, kernel = 'rbf'	C = regularization parameter. gamma = kernel coefficient. kernel = kernel type used, 'rbf' represents radial basis function.

## RF

The RF classifier is an ensemble of decision tree algorithms with demonstrated robustness in remote sensing image classification compared to single classifiers (Breiman, 2001; Rodriguez-Galiano *et al.*, 2012). The algorithm relies on unit vote contributions from each classifier within the ensemble to assign input vectors to different classes, where the most frequently voted class is retained (Breiman, 2001). The individual decision trees are parameterised using several independent random subsets of training data sampled through bootstrap aggregation or bagging. This reduces multicollinearity and generalization error (Abdi, 2020; Breiman, 2001). The input vectors that do not form part of the bootstrap sample (i.e., “out-of-bag” (OOB) sample) are used for evaluation and variable importance estimation (Breiman, 2001; Eisavi *et al.*, 2015). By design, decision tree classifiers require some measure for selecting suitable features per class, which maximizes dissimilarities between classes (Rodriguez-Galiano *et al.*, 2012). The RF algorithm uses Gini Index for feature selection at each node (Breiman, 2017). When assigning an input pixel to a class ( $C_i$ ), for a given training set ( $T$ ), the Gini Index measures feature impurity with respect to the different classes and is expressed as

$$\sum_{i \neq j} (f(C_i, T)/|T|)(f(C_j, T)/|T|) \quad (2.2)$$

where  $(f(C_i, T)/|T|)$  is the probability that the selected pixel belongs to class  $C_i$  (Pal, 2005; Rodriguez-Galiano *et al.*, 2012).

Each decision tree therefore grows to a maximum depth using a combination of features. The number of features used to grow a tree at each node and the number of decision trees are the required user-defined parameters to instantiate a RF prediction model (Pal, 2005).

## SVM

SVM was developed based on statistical learning theory (Wang & Zhong, 2003). The algorithm creates an optimal separating hyperplane based on the location of a small subset of training samples at class boundaries, the so-called “support vectors” (Cortes & Vapnik, 1995). Given a simple binary linear classification problem, the SVM

uses quadratic optimization techniques to select the optimum margin of separation between the two classes such that the distance to the hyperplane from the closest support vectors of both classes is maximal (Cortes & Vapnik, 1995; Foody *et al.*, 2016).

For a nonlinear classification problem, the algorithm selects the optimal margin by (i) allowing some misclassification errors and (ii) transforming the original input space into a higher dimensional feature space using nonlinear functions  $\varphi$  (Wang & Zhong, 2003), making linear separation possible in the new feature space. To reduce computational cost, kernel functions,  $K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$ , such as polynomials, radial basis and sigmoid functions, are used for the transformation (Camps-Valls & Bruzzone, 2009). The decision function is given by

$$f(x) = \text{sign} \sum_{i=1}^l \alpha_i y_i (\varphi(x_i) \cdot \varphi(x)) + b \quad (2.3)$$

where  $\alpha_i$  is a slack variable (Lagrange multiplier).

To classify new datasets, the algorithm uses learned parameters from the decision function based on training data. The trade-off between margin of class separation and misclassification errors is controlled by defining a regularisation parameter  $C$ , where  $C \in \mathbb{Z}$  and  $0 < C < \infty$  (Cortes & Vapnik, 1995).

## CART

CART is a decision tree algorithm that builds classification or regression trees based on categorical or numerical attributes, respectively (Breiman, 2017). The structure of the tree is typified by a root node and a series of internal nodes (splits) and terminal nodes (leaves). Within this framework, the algorithm builds a model by recursively partitioning the training dataset into increasingly homogeneous subsets using tests applied at each node to training features (Xie *et al.*, 2019). Given a continuous data set, the test performed at each node is of the form

$$x_i > c \quad (2.4)$$

for decision functions based on a single feature (i.e., univariate decision trees), where  $x_i$  is a measurement in  $n$  feature space ( $n = 1$  in this case) and  $c$  is a decision threshold estimated from the range of  $x_i$  measurements (Pal & Mather, 2003). The threshold ( $c$ ) value is determined using an impurity measure such as entropy (Quinlan,

2014) and the Gini index (Breiman, 2017). If the decision boundaries are defined by a combination of features (i.e., multivariate decision trees), the test takes the form

$$\sum_i a_i x_i \leq c \quad (2.5)$$

where  $a_i$  is a vector of coefficients of a linear discriminant function estimated from the training data (Pal & Mather, 2003). The series of testing outputs form the branches of the tree which proceeds sequentially through internal nodes until a terminal node is reached. At each terminal node, class labels are assigned based on maximum probabilities (Bittencourt & Clarke, 2003).

### MLP

The MLP is a feedforward artificial neural network (ANN) classifier which is trained using back-propagation (Goodfellow *et al.*, 2016). Learning in ANN is inspired by the functioning of neurons within the brain, which is based on parallel and distributed processing of information (Bischof *et al.*, 1992). Similarly, the MLP architecture is composed of multiple layers of fully connected processing units called neurons, which are arranged sequentially as a network of input, hidden and output layers. During training, each unit in a hidden layer receives data from the input/previous layer, processes it and feeds it forward to units in the next layer (Kanellopoulos *et al.*, 1992). This allows more abstract representations of the data to be learned until the output layer is reached (Abdi, 2020). The connections between units carry weights, which are modified iteratively to minimise a cost function. Apart from the input layer, the net input to each unit is therefore the weighted sum of outputs from the previous layer (Bischof *et al.*, 1992; Kanellopoulos *et al.*, 1992). The net input is wrapped in an activation function to produce the output for that unit. The output for each processing unit is expressed as

$$o_i = f \left( \sum_j w_{ij} * o_j + b_i \right) \quad (2.6)$$

where  $o_i$  is the output of a neuron in layer  $i$ ,  $w_{i,j}$  is the connecting weight between layers  $i$  and  $j$ ,  $o_j$  is output from layer  $j$  and  $b_i$  is bias and  $f$  is the activation function (Bischof *et al.*, 1992; Kanellopoulos *et al.*, 1992).

## 2.2.7 Algorithm Calibration and Evaluation

The machine learning algorithms, namely, RF, SVM, CART and MLP, were first calibrated and evaluated for general land cover classification using data from a sub-area within the Lower Sabie landscape (see Figure 2.1) via a nested cross-validation approach. For each algorithm, the combination of parameters that returned the best expected classification accuracies were then used in the prediction of grazing lawn occurrence probabilities in the broader Lower Sabie and Satara Landscapes. The steps employed are broadly summarised into (i) data preparation, (ii) parameterisation training and classification and (iii) accuracy assessment. All processing was done using Intel(R) Core(TM) i5-6200U CPU with 8GB RAM on 64 bit Windows 10 operating system, and was supplemented by leveraging the power of Google's free GPU hardware, using the Google Collaborator platform.

### Data Preparation

The post-RF-RFE spectral and texture variables were used as input predictors for modelling. The input dataset was then transformed by subtracting the mean and scaling to a unit variance to generate normalised scores per feature using Equation 2.7:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (2.7)$$

where  $x_{ij}$ ,  $\mu_j$  and  $\sigma_j$  are pixel value, mean and standard deviation of pixels in the  $j$ th feature, respectively, and  $z_{ij}$  is the transformed value of  $x_{ij}$  (Singh & Singh, 2019).

Normalising input features is a crucial preprocessing technique which approximately equalises dynamic data ranges in features for unbiased and improved learning (Singh *et al.*, 2015). Further, it is a common requirement prior to the training of machine learning estimators such as Support Vector Machines and Artificial Neural Networks (Singh & Singh, 2019).

### Parameterisation, Training and Classification

Each of the selected algorithms comes with a set of hyperparameters which has to be tuned to maximise performance during training. Algorithm training thus involved hyperparameter optimisation whereby optimal hyperparameter sets were selected for RF, SVM, CART and MLP algorithms from a predefined grid. The optimisation process and selection of best model parameters were performed using randomised

grid search in a 2x5 nested cross-validation approach with 10 iterations. Nested cross-validation incorporates optimal hyperparameter selection and unbiased estimation of model performance in inner and outer cross-validation loops respectively (Wainer & Cawley, 2018). The approach is mostly recommended against traditional “flat cross-validation” which results in biased accuracy estimates due to information leakage and the split sample method plagued by insufficient availability of training and test datasets (Cawley & Talbot, 2010). The chosen thresholds for tune-length and train-test splits were deemed appropriate to provide a reasonable trade-off between ensuring a robust model and computational time. Hyperparameters that returned the best expected classification accuracies were selected and used as input parameters in the machine learning algorithms. The algorithms were retrained with the full training data for landscape-wide prediction of land cover occurrence probabilities in both the Lower Sabie and Satara landscapes.

Individual image variable weights were computed using permutation feature importance estimates (Pedregosa *et al.*, 2011) in order to assess their relative contributions in each machine learning model. Permutation feature importance (PFI) generates variable weights based on an observed decrease in model score when a single variable is randomly shuffled (Breiman, 2001). The drop in model score thus represents the degree to which the model depends on the variable of interest. The PFI technique is model agnostic, which makes it suitable for comparison of feature importance estimates from RF, SVM, CART and MLP models used in this study.

The predicted occurrence probability surface for grazing lawns was selected and used as input in an optimised probability thresholding procedure. Optimised probability thresholding involved the selection of a single occurrence probability value (threshold) which maximises some measure of classification accuracy (Hird *et al.*, 2017) for the target class. We tested a series of probability values at 0.05 intervals to determine the threshold that maximises F-score of grazing lawn detection. Grazing lawn (G) and non-grazing lawn (O) classes were assigned using simple relational expressions represented by Equation 2.8 and Equation 2.9, respectively,

$$G = p \geq t \quad (2.8)$$

$$O = p < t \quad (2.9)$$

where  $p$  is occurrence probability and  $t$  is the optimal probability threshold,  $t \in p$ .

### Accuracy Assessment and Comparison

Model performance in discriminating different savannah land cover types during hyperparameter tuning was assessed using point and interval estimates of Overall Accuracy (OA) and F-score, based on a 2x5 nested cross-validation approach (see Section 2.2.7). Further, accuracy of grazing lawn/non-grazing lawn binary maps was assessed by confusion matrix (Congalton & Green, 2019), from which precision, recall, F-score and OA metrics were calculated using Equations 2.10 to 2.13. Accuracy-adjusted estimates of grazing lawn area coverage were obtained following Olofsson *et al.* (2014).

$$Precision = \frac{tp}{tp + fp} \quad (2.10)$$

$$Recall = \frac{tp}{tp + fn} \quad (2.11)$$

$$F\_score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (2.12)$$

$$OA = \frac{tp + tn}{tp + fp + tn + fn} \quad (2.13)$$

where tp, fp, tn and fn represent the number of true positive, false positive, true negative and false negative cases, respectively.

Marginal homogeneity between predictions from model pairs was tested at 5% level of significance using the McNemar chi-squared ( $\chi^2$ ) test (McNemar, 1947). The McNemar test compares the error matrices of two classification methods to test the null hypothesis that the two methods have the same error rate. The method is based on  $\chi^2$ -test and provides a robust statistical comparison of class-wise predictions between two algorithms (Roggo *et al.*, 2003). Additionally, the estimated proportion of grazing lawn cover (PGLC) was compared for model pairs using the two-proportion Z-test at 5% level of significance. The two-proportion Z-test follows a  $\chi^2$  distribution with one degree of freedom (Abdi, 2020), and was used to test the null hypothesis of no difference between PGLC for model pairs.

## **2.2.8 Spatial Analysis of Grazing Lawn Distribution**

Using spatial metrics, the characteristics of grazing lawn distribution were determined at landscape-scale and spatial patterns were analysed along a gradient of distance from water source in both Lower Sabie and Satara landscapes. Spatial metrics provide vital information on landscape configuration and composition (Herold, 2001). Spatial-contextual information such as density, shape, size and aggregation of land cover patches can be extracted from spatial metrics to better understand ecological processes at the landscape-scale (Herold, 2001; Mcgarigal & Marks, 1995). The classification output with the least error properties was selected as input in the calculation of (i) Number of Patches (NP), (ii) proportion of landscape covered by grazing lawns (PL), (iii) maximum patch area (MPA) and (iv) cohesion index (CI) (Mcgarigal & Marks, 1995), from which patterns in grazing lawn structure were determined. Further, Pearson's correlation and coefficients of determination were estimated in order to identify the nature and significance of the relationship between grazing lawn structural attributes and proximity to water source. Calculation of the selected spatial metrics was carried out using the “SpatialEco” package (Evans & Ram, 2019) in the R programming environment (R Core Team, 2017).

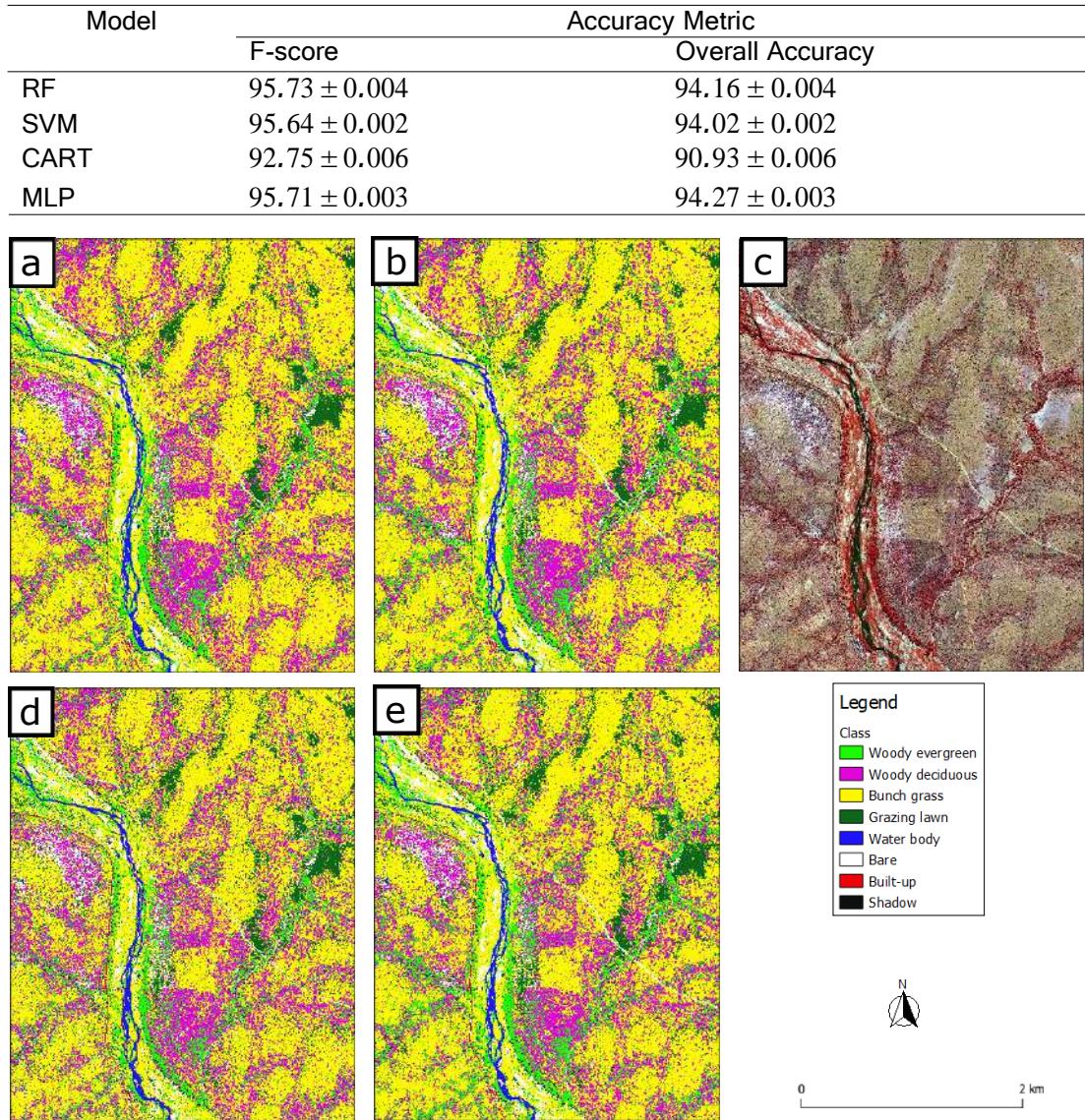
## **2.3 Results**

### **2.3.1 Model Quality for Land Cover Classification**

Cross-validation accuracy results for individual models using F-score and Overall Accuracy (OA) measures are presented in Table 2.5. Generally, all models achieved high accuracies in differentiating the different land cover classes, with median F-score and OA measures ranging between  $92.75 \pm 0.005\%$  and  $95.73 \pm 0.003\%$  and  $90.92 \pm 0.002\%$  and  $94.27 \pm 0.003\%$ , respectively. RF, SVM and MLP models had similar accuracy scores and marginally outperformed CART for both F1 and OA measures (Table 2.5).

Figure 2.5 shows land cover maps for the training region at 2 m spatial resolution. The maps show similar representation of savannah land cover types across all models, all of which were closely consistent with the reference satellite image scene (Figure 2.5c).

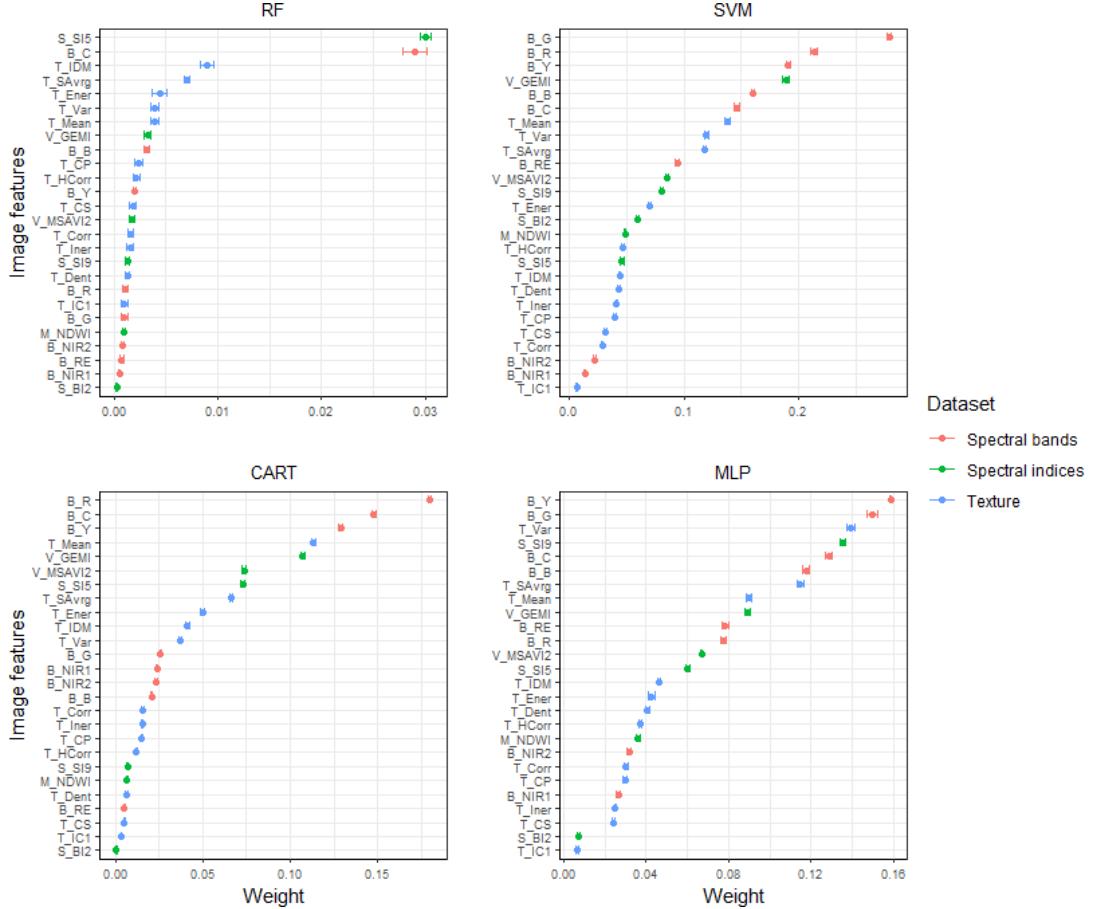
**Table 2.5:** Accuracy scores (F1 and Overall Accuracy) from  $2 \times 5$  nested cross-validation showing a comparison of model performance. RF = Random Forest, SVM = Support Vector Machines, CART = Classification and Regression Trees, MLP = Multilayer Perceptron.



**Figure 2.5:** Land cover classification of the training region from (a) RF, (b) SVM, (d) CART and (e) MLP models. The WorldView-3 image scene (False colour: NIR1, R, G) of the training region is showed in panel (c). RF = Random Forest, SVM = Support Vector Machines, CART = Classification and Regression Trees and MLP = Multilayer Perceptron.

Figure 2.6 shows permutation feature importance estimates across all models. A mix of image features from original spectral bands, spectral indices and texture variables showed high importance in each model. There was generally more agreement among SVM, CART and MLP models in assigning relatively more importance to original

spectral bands in terms of both magnitude of feature weight and number of features. However, image features that exhibited high importance in the RF model were largely dominated by texture variables (Figure 2.6). Image features that were of highest importance in RF, SVM, CART and MLP models were S\_SI5, B\_G, B\_R and B\_Y, respectively.



**Figure 2.6:** Image feature weights derived from permutation feature importance estimates for Random Forest (RF), Support Vector Machines (SVM), Classification and Regression Trees (CART) and Multilayer Perceptron (MLP) models. Feature weights are sorted in an descending order across models to identify features on high predictive importance. For the detailed names and description of image feature acronyms, refer to Table 2.3.

A summary of the most important predictors for each feature group (i.e., spectral bands, spectral indices and texture variables) is presented in Table 2.6. Following Kukunda *et al.* (2018), identification of the most important image features considered both the magnitude of feature weights and consistency of being assigned high importance across all models. In terms of magnitude, image features that were considered

highly important in each model were limited to the first three features, in each feature group (Figure 2.6). Conversely, features were deemed consistent if they were assigned high importance in at least three models (Table 2.6). Among the most important image features were **B\_C** and **B\_Y** for spectral bands, **V\_GEMI** and **V\_MSAVI2** for spectral indices and **T\_Mean** and **T\_SAvrg** for texture variables (see Table 2.6 for features in bold).

### 2.3.2 Grazing Lawn Occurrence Probability Prediction and Classification

The outputs of grazing lawn occurrence probability surfaces for RF, SVM, CART and MLP models are shown in Figure 2.7A and Figure 2.8A for Lower Sabie and Satara landscapes respectively. The general pattern of grazing lawn occurrence probability surfaces at both study sites is comparable among the four models. Within the Lower Sabie site, high grazing lawn occurrence probabilities were mostly confined to the eastern and north-eastern part of the landscape, and were similar across all models (Figure 2.7A). The obvious qualitative difference among models is the relative lack of many very low values in the CART probability surface compared to RF, SVM and MLP models.

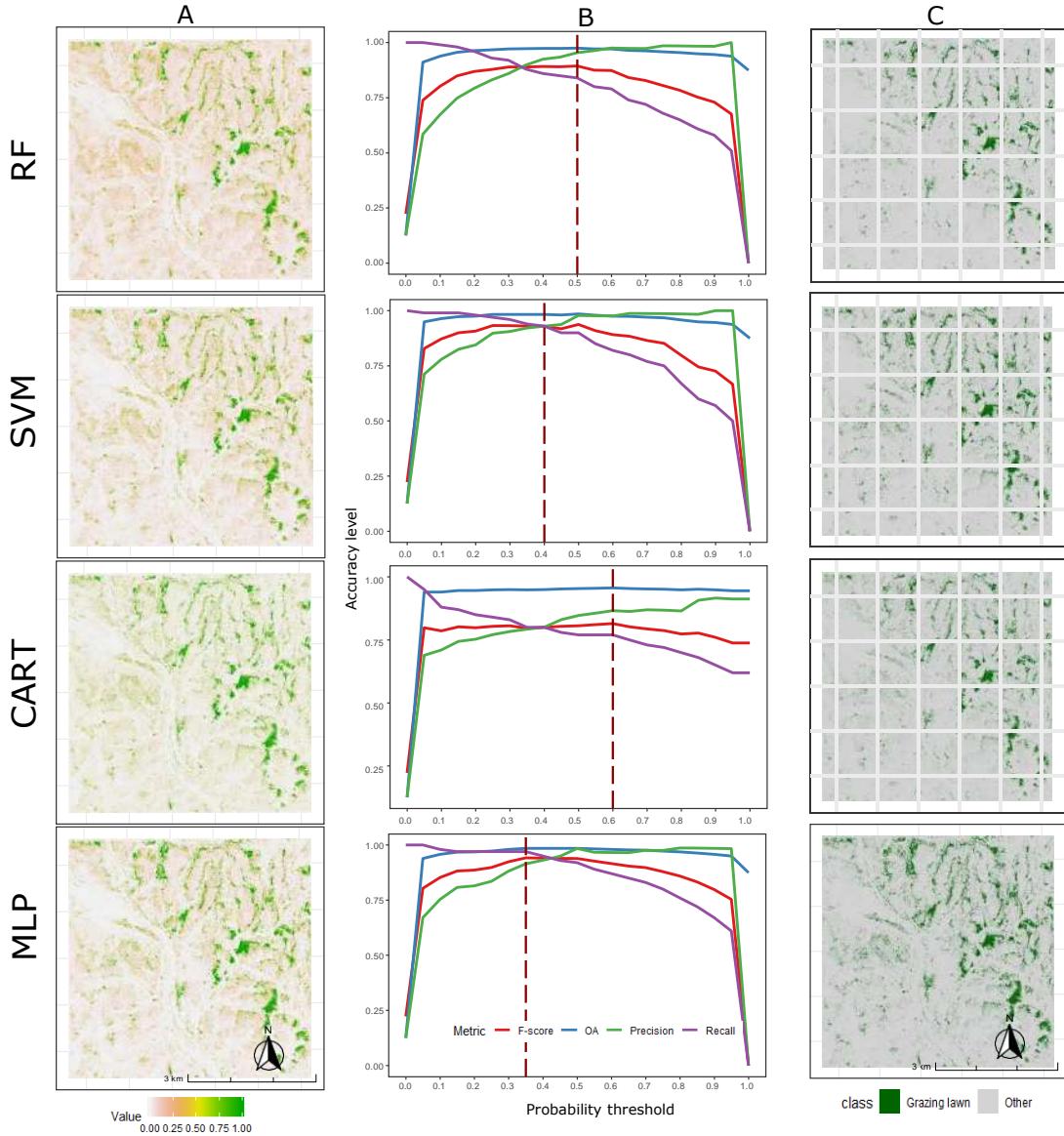
Within the Satara landscape, high grazing lawn occurrence probabilities mostly aligned along a diagonal stretch from northwest to southeast (Figure 2.8A), which is the interface of the granite and basalt geologies. Despite similarities in spatial distribution of high occurrence probability values, there were noticeable qualitative differences in range among the four models. The RF probability surface exhibited a relatively high prevalence of a continuous range of very low to medium probability values across the landscape, and very few distinctively high occurrence probabilities. In contrast, the CART model predicted relatively more medium to high probability values across the landscape, while MLP and SVM predictions were similar in the distribution of very low and very high occurrence probability values (Figure 2.8A).

Plots of model F-score, Precision, Recall and OA values generated over a series of predicted probabilities for the Lower Sabie and Satara landscapes are presented in Figure 2.7B and Figure 2.8B, respectively. Analysis of the relationship between F-score and predicted probabilities revealed the optimal threshold for classifying grazing lawns. The optimal threshold is the probability value which maximises model F-score of grazing lawn detection, and was found to coincide with or lie close to the equilibrium

**Table 2.6:** Summary of the first three most important image features from spectral bands, spectral indices and texture variables across all models. Image features that appear in at least three models are in bold. For the detailed names and description of image feature acronyms, refer to Table 2.3.

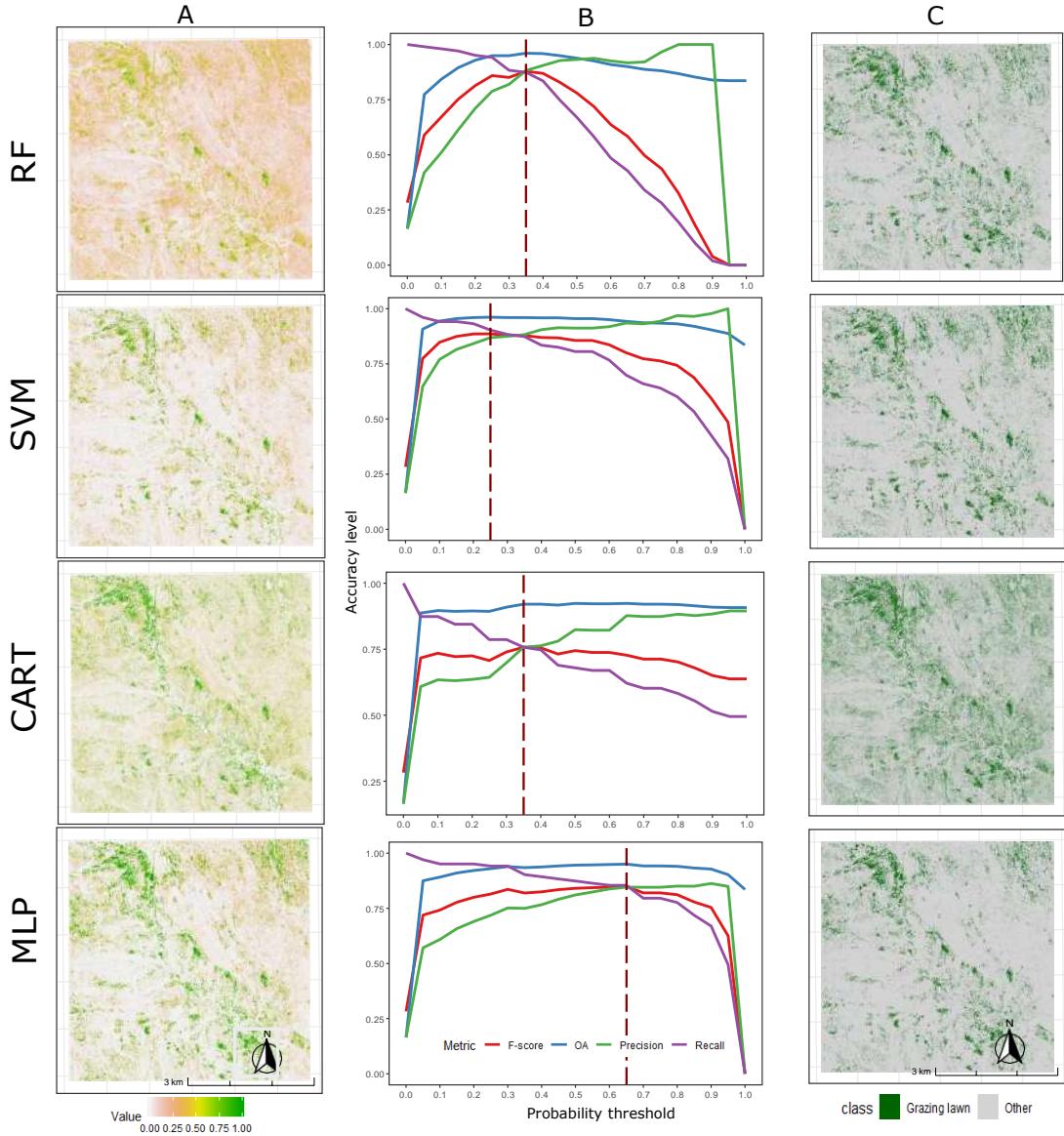
Dataset	Image feature	Model			
		RF	SVM	CART	MLP
Spectral band	<b>B_C</b>	☒		☒	☒
	B_B	☒			
	B_G		☒		☒
	<b>B_Y</b>	☒	☒	☒	☒
	B_R		☒	☒	
	B_RE				
	B_NIR1				
	B_NIR2				
Spectral index	<b>V_GEMI</b>	☒	☒	☒	☒
	<b>V_MSAVI2</b>	☒	☒	☒	☒
	M_NDWI				
	S_SI5	☒		☒	
	S_SI9		☒		☒
	S_BI2				
Texture	T_Ener	☒		☒	
	T_Corr				
	T_IDM	☒			
	T_Iner				
	T_CS				
	T_CP				
	T_HCorr				
	<b>T_Mean</b>		☒	☒	☒
	T_Var		☒		☒
	<b>T_SAvrg</b>	☒	☒	☒	☒
	T_Dent				
	T_IC1				

point between model Precision and Recall. The resulting values varied across models, with the F-score of RF, SVM, CART and MLP models peaking at 0.5, 0.4, 0.6 and 0.35,



**Figure 2.7:** Grazing lawn occurrence probability surfaces (A); optimal probability threshold plot (B); and binary map of grazing lawn and other cover (C) derived from RF, SVM, CART and MLP models for the Lower Sabie landscape.

respectively, for Lower Sabie as seen in Figure 2.7B. Similar analysis on predicted probability surfaces for the Satara landscape resulted in relatively lower thresholds for RF, SVM and CART (0.35, 0.25 and 0.35, respectively) and a higher threshold for MLP (0.6) where model F-scores were maximum Figure 2.8B.



**Figure 2.8:** Grazing lawn occurrence probability surfaces (A); optimal probability threshold plot (B); and binary map of grazing lawn and other cover (C) derived from RF, SVM, CART and MLP models for the Satara landscape.

The grazing lawn/non-grazing lawn binary maps resulting from applying corresponding thresholds to each of the four predicted probability surfaces are shown in Figure 2.7C and Figure 2.8C for both Lower Sabie and Satara landscapes, respectively. Analogous to the probability surfaces, patterns of grazing lawn distribution were similar for all classifications within both landscapes. However, local variations persisted and were consistent with the distribution of predicted probability values for each model in both landscapes. Overall, the Satara maps showed a considerable level of speckling (Figure 2.8C).

A summary of accuracy measures for grazing lawn detection is presented in Table 2.7. F-scores for Lower Sabie ranged between 0.81 for CART to 0.94 for MLP, while SVM and RF classifications achieved F-score of 0.93 and 0.89 respectively (Table 2.7). Grazing lawn detection accuracy results were high, but relatively lower for the Satara area compared to Lower Sabie. F-score ranged between 0.75 for CART to 0.88 for SVM, while RF and MLP achieved F-scores of 0.87 and 0.85, respectively (Table 2.7).

**Table 2.7:** Model Precision, Recall and F-score metrics of grazing lawn detection in both Lower Sabie and Satara landscapes. RF = Random Forest, SVM = Support Vector Machines, CART = Classification and Regression Trees, MLP = Multilayer Perceptron.

Landscape	Accuracy metric	Model score			
		RF	SVM	CART	MLP
Lower Sabie	Precision	0.95	0.93	0.87	0.92
	Recall	0.84	0.93	0.77	0.97
	F-score	0.89	0.93	0.81	0.94
Satara	Precision	0.88	0.87	0.76	0.85
	Recall	0.87	0.90	0.76	0.85
	F-score	0.87	0.89	0.76	0.85

Accuracy-adjusted estimates of area covered by grazing lawns within both landscapes are presented in Table 2.8. As expected, all model classifications gave comparable estimates of grazing lawn cover within the Lower Sabie site, ranging between 2.46 km<sup>2</sup> (for RF) and 2.98 km<sup>2</sup> (for CART) (Table 2.8). In contrast, estimates of grazing lawn cover were significantly different ( $p \leq 0.05$ ) for all models within the Satara landscape (see test results in Table A.1 of supplementary data).

**Table 2.8:** Accuracy adjusted area estimates of grazing lawn cover in Lower Sabie and Satara landscapes. Area estimates with different letters differ significantly and vice versa in each landscape. RF = Random Forest, SVM = Support Vector Machines, CART = Classification and Regression Trees, MLP = Multilayer Perceptron.

Landscape	Area Estimate (km <sup>2</sup> )			
	RF	SVM	CART	MLP
Lower Sabie	$2.46 \pm 0.18^a$	$2.64 \pm 0.13^a$	$2.99 \pm 0.22^a$	$2.96 \pm 0.11^a$
Satara	$3.82 \pm 0.22^a$	$3.61 \pm 0.21^b$	$5.54 \pm 0.35^c$	$3.13 \pm 0.24^d$

McNemar test results presented in Table 2.9 showed statistically significant differences ( $p \leq 0.05$ ) in grazing lawn detection error rate when comparing CART to RF, SVM and MLP models in both Lower Sabie and Satara landscapes. In contrast, no significant differences were observed for all the other model pairs (Table 2.9).

**Table 2.9:** McNemar's chi-squared test ( $\chi^2$ ) of marginal homogeneity between model pairs. Values in parenthesis represent p-value. Model pairs that show statistically significant difference ( $p \leq 0.05$ ) in error rate are in bold. CART = Classification and Regression Trees, MLP = Multilayer Perceptron, RF = Random Forest, SVM = Support Vector Machines.

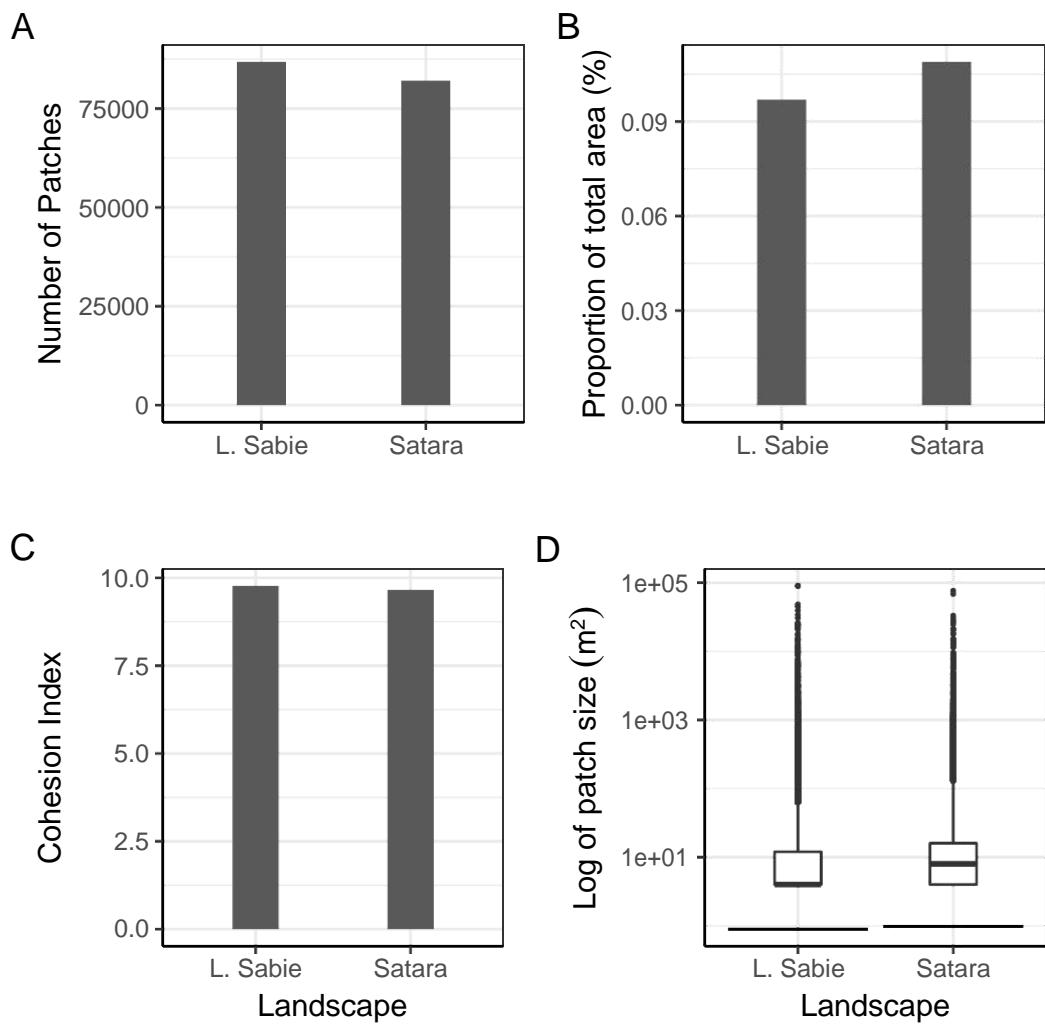
Lower Sabie		Satara	
Model pair	$\chi^2$ -test	Model pair	$\chi^2$ -test
CART v MLP	<b>14.667(0.000)</b>	CART v MLP	<b>5.891(0.015)</b>
CART v RF	<b>10.316(0.001)</b>	CART v RF	<b>13.395(0.000)</b>
CART v SVM	<b>16.000(0.000)</b>	CART v SVM	<b>11.574(0.000)</b>
MLP v RF	2.450(0.117)	MLP v RF	1.250(0.264)
MLP v SVM	0.100(0.752)	MLP v SVM	1.565(0.211)
RF v SVM	2.083(0.149)	RF v SVM	0.000(1.000)

### 2.3.3 Spatial Patterns in Grazing Lawn Cover

Landscape-scale summary of number of grazing lawn patches, total coverage, connectedness and patch size distribution are presented in Figure 2.9.

Number of grazing lawn patches was relatively higher in Lower Sabie compared to the Satara landscape (Figure 2.9A). However, analysis of patch size distribution revealed the Satara landscape as having relatively larger grazing lawn patches (Figure 2.9D), and higher area coverage compared to the Lower Sabie landscape (Figure 2.9B). Spatial connectedness of grazing lawn patches was however comparable in both landscapes (Figure 2.9C).

Further analysis of patterns in the proportion of landscape covered by grazing lawns (PL), maximum patch area (MPA) and cohesion (CI) revealed significant relationships with distance from water sources in both landscapes. Grazing lawn PL, MPA and CI showed an inverse relationship with distance from water source in both landscapes (see correlation coefficients in Table 2.10). However, the trends were relatively less distinct in the Lower Sabie landscape (Figure 2.10), as also suggested by the differ-



**Figure 2.9:** Landscape-scale summary of grazing lawn spatial characteristics. (A) Number of grazing lawn patches; (B) Proportion of total landscape covered by grazing lawns; (C) Physical connectedness of grazing lawn patches; (D) Distribution of grazing lawn patch size. Dashed horizontal line represents mean patch size.

ences in magnitude of correlation coefficients between both landscapes (Table 2.10). Overall, grazing lawn fractional cover, patch size and spatial connectedness were highest within 0.7 km from water sources in both Lower Sabie and Satara landscapes (Figure 2.10).

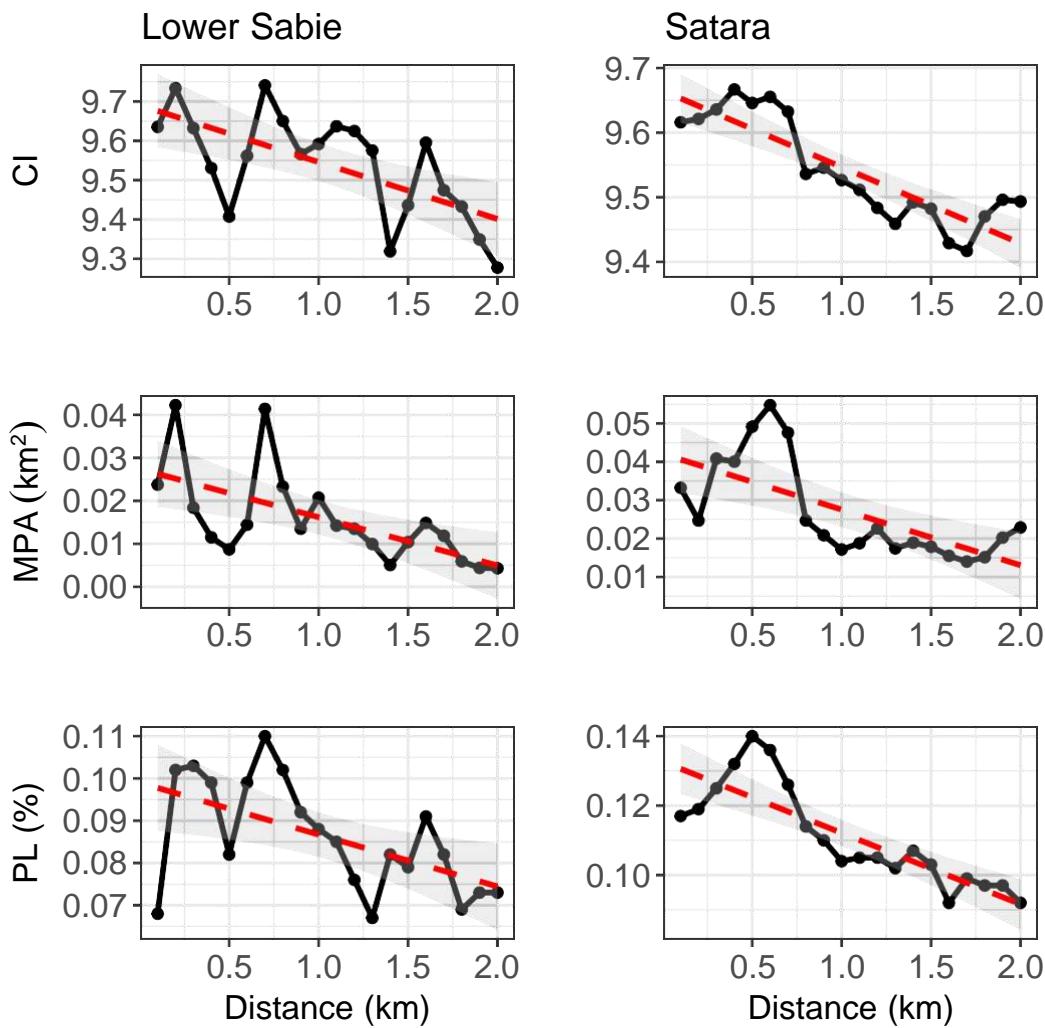
**Table 2.10:** Pearson Correlations ( $r$ ) and Coefficients of Determination ( $r^2$ ) from the relationship between grazing lawn spatial metrics and distance from water source. PL = Proportion of Landscape, MPA = Maximum Patch Area and CI = Cohesion Index. Relationships are significant at  $p < 0.0001$  ‘\*\*\*’,  $p < 0.001$  ‘\*\*’ and  $p < 0.01$  ‘\*’.

<b>Landscape metric</b>	<b>Lower Sabie</b>		<b>Satara</b>	
	<b>r</b>	<b><math>r^2</math></b>	<b>r</b>	<b><math>r^2</math></b>
PL	-0.55	0.30*	-0.84	0.70***
MPA	-0.62	0.39**	-0.68	0.46**
CI	-0.65	0.42**	-0.87	0.75***

## 2.4 Discussion

### 2.4.1 Model Quality for Savannah Land Cover Classification

In this study, efforts were focused on developing robust machine learning framework for grazing lawn detection by first assessing model performance for general classification of savannah land cover types. The convergence of remote sensing and data science techniques through machine learning offers unparalleled capacity for more accurate processing of satellite imagery, especially for the purposes of land cover monitoring. While this presents many advantages for remote sensing-based ecosystem monitoring, the choice of fit-for-purpose machine learning algorithms often requires some experimentation. This is partly due to the vast availability of options to select from, but most importantly also due to contextual differences in application such as varying landscape conditions, data and research objectives (Foody *et al.*, 2016; Lawrence & Moran, 2015). Robust evaluation of algorithm performance is therefore vital for the selection of optimal models for application. The nested cross-validation approach used here allowed the simultaneous tuning of hyperparameters and unbiased estimation of individual model performance. In so doing, the optimum combination of algorithm hyperparameters which enhanced model quality could be selected. Model quality evaluation via nested cross-validation has been proven effective in avoiding biased accuracy estimates common in “flat cross-validation” due to information leakage, while preventing poor model generalisation capabilities due to data paucity—a regular challenge of the split sample approach to model evaluation (Cawley & Talbot, 2010).



**Figure 2.10:** Plots of spatial metrics showing patterns in grazing lawn spatial structure and distribution with distance from water-points. CI = Cohesion Index, MPA = Maximum Patch Area, PL = Proportion of Landscape.

All machine learning models (RF, SVM, CART and MLP) demonstrated high performance in classifying the different savannah land cover categories. Even so, RF, SVM and MLP marginally outperformed the CART model. The relatively lower performance of CART observed in this study is consistent with widely reported inferiority of single decision tree (DT)-based classifiers relative to other machine learning algorithms for land cover classification (Camargo *et al.*, 2019; Kaszta *et al.*, 2016; Rodriguez-Galiano & Chica-Rivas, 2014). For example, similar findings were reported by Kaszta *et al.* (2016) in a comparative assessment of classification algorithms for seasonal separation of southern African savannah components. The authors recor-

ded CART as having the lowest accuracy score in both pixel-based and object-based approaches relative to SVM and RF algorithms. Camargo *et al.* (2019) compared the performance of RF, SVM, MLP and DT for classifying the Brazilian tropical savannah biome, and found that DT produced a relatively lower performance than RF, SVM and MLP classifiers. Similar to our findings, the authors recorded closely comparable performance for the latter three algorithms. In a related Mediterranean land cover monitoring study, Rodriguez-Galiano & Chica-Rivas (2014) reported significantly lower mapping accuracy for DT than SVM, ANN and RF algorithms.

Although CART is relatively more flexible and intuitive to implement, it is very sensitive to small variability in data, which, in this case, may have contributed to the relatively lower performance in a heterogeneous savannah landscape. Compared to single decision trees such as CART, the RF algorithm draws its higher generalisability from the contribution of multiple decision trees parameterised using random subsets and bootstrap aggregation (Breiman, 2001). RF is therefore highly adaptable to different data ranges and robust against multicollinearity. Similarly, the MLP architecture allows more abstract representations of data to be learned (Kanellopoulos *et al.*, 1992). However, the performance of MLP is strongly influenced by input data structure, and performs better when data ranges of all input features are equal (Singh & Singh, 2019). Thus, the inclusion of data normalisation during preprocessing likely aided gains in classification accuracy. In the case of SVM, the use of nonlinear vector mapping functions facilitates creation of decision boundaries for effectively dealing with nonlinearly separable classes (Camps-Valls & Bruzzone, 2009). The superiority of RF, MLP and SVM algorithms in dealing with the typical spectral homogeneity of the heterogeneous savannah landscape could thus be attributed to their relatively higher adaptive capacity in complex nonlinear classification problems. It should be noted that RF was used in an RF-RFE procedure for selecting final input image features for classification (see Section 2.2.5). This may have aided the performance of RF, although the RF-RFE algorithm was configured with different (default) hyperparameters during implementation.

As expected, a combination of original image bands, spectral indices and texture variables enhanced discrimination capacities of the machine learning models in savannah land cover classification. Across all models, the most important predictors—B\_C, B\_Y (original bands), V\_GEMI, V\_MSAVI2 (spectral indices), T\_Mean and T\_SAvrg (texture features)—highlight variations in photosynthetic status and structure of savannah vegetation. The high importance of the coastal blue (B\_C) and yellow (B\_Y)

WorldView-3 image bands could be attributed to their strong sensitivity to differences in foliar chrolophyll content (Schuster *et al.*, 2015), given that images were acquired in the dry season, which is when phenological differences are most pronounced (Bucini *et al.*, 2009; Symeonakis *et al.*, 2018). Several studies have reported the contribution of these bands and the red edge band in mapping vegetation components (Ghosh & Joshi, 2014; Immitzer *et al.*, 2012; Kaszta *et al.*, 2016). Unlike the reported studies, the red edge band was not very important in our classification of savannah land cover. The coastal blue wavelength is absorbed by chlorophyll in healthy plants while the yellow band detects dryness/“yellowness” of vegetation, both of which are instrumental in vegetative analysis. The high importance of the spectral indices could be explained by their strong correlation with vegetation biomass (Guerini Filho *et al.*, 2020; Yin *et al.*, 2018) and moisture content (Hunt Jr *et al.*, 2016; Roberto *et al.*, 2016), which helps to capture the varying characteristics of heterogeneous savannah vegetation that would otherwise be attenuated when using the original image bands alone. Additionally, the high importance of the texture features highlighted the chromatic variations in dry season savannah vegetation components. Both T\_Mean and T\_SAvrg measures inter-pixel average in brightness values which were sufficiently captured at 2 m resolution of the WorldView-3 in a heterogeneous savannah landscape.

#### **2.4.2 Grazing Lawn Detection and Model Comparison**

After ascertaining the optimal parameters for training, models were refitted with the entire training set for wide-scale prediction of land cover occurrence probabilities in both Lower Sabie and Satara landscapes. For each landscape, the general pattern of grazing lawn occurrence probabilities was comparable for all models, particularly the distribution of high occurrence probability values. Individual model outputs however exhibited local predictive variations in the distribution of low to medium occurrence probabilities, which could be reflective of differences in model complexity and uncertainties (Ferchichi *et al.*, 2018; Schulp *et al.*, 2014). Binary maps were derived from predicted grazing lawn occurrence probability surfaces using thresholds which maximised F-score. Hird *et al.* (2017) adopted a similar approach for large area classification of wetlands and drylands using True Skill Statistics (TSS), and achieved 85% OA score. As expected, derived maps showed more coherent representation of grazing lawn areas within the Lower Sabie landscape, while maps for the Satara landscape were characterised by relatively higher degree of noise, particularly for the CART-derived map. This was also reflected in F-score measures, where relatively higher

grazing lawn detection accuracies were recorded across all models for the Lower Sabie savannah landscape. It should be emphasised that models were trained using data from the Lower Sabie landscape—for reasons related to training data quality and computational cost—which allowed testing their spatial transferability. Possible differences in dry season reflectances due to the different rainfall regimes and underlying geologies as well as difference in image acquisition geometries, may have increased prediction bias of model spatial transfer and contributed to the observed differences both across the different landscapes and models. Further analysis is required to comprehensively ascertain the impacts of differences in environmental conditions and image acquisition characteristics on spatial transferability of the classification models. McMenar and two-proportion Z-test results showed significant differences ( $p \leq 0.05$ ) in error rates for pairs of of CART versus other models, further highlighting the relative preponderance of RF, SVM and MLP. Overall, differences in accuracy of grazing lawn detection were attenuated when models were applied in a different spatial context. This suggests the need to calibrate predictive models with local contextual information prior to application. Additionally, our findings re-emphasise recommendations from (Foody, 2004; Janssen & Vanderwel, 1994; Momeni *et al.*, 2016) to conduct statistically rigorous comparison of accuracy statements before drawing definitive conclusions on map quality assessment.

### 2.4.3 Spatial Patterns in Grazing Lawn Distribution

The formation and persistence of grazing lawns in southern African savannah landscapes has been identified to be dependent on a number of interacting top-down and bottom-up ecological processes. Among the most widely reported are continued grazing (Cromsigt & Olff, 2008; Grant & Scholes, 2006), which can be linked to fire, rainfall and nutrient hotspots concentrating grazers on specific areas (Archibald *et al.*, 2005; Archibald, 2008; Donaldson *et al.*, 2018). Spatial variations in such factors are thus expected to shape spatial patterns and distribution of grazing lawns Hempson *et al.* (2015).

Both fire and grazers consume grass biomass, and have the potential to shift grass communities into tall grass or short grass grazing lawn states (Donaldson *et al.*, 2018). However, the rate at which these alternate grassland states are established is strongly influenced by landscape productivity (Archibald *et al.*, 2005). Tall grasses are strong light competitors and are well adapted to fire-prone conditions due to their extensive rooting system, which makes them the dominant grass community under

high productivity conditions (Hempson *et al.*, 2015). On the other hand, short grass grazing lawns can withstand high grazing pressure due to their stoloniferous growth form which protect reproductive parts from being destroyed. Subject to similar grazing conditions, the proportion of grazing lawn cover within the savannah landscape is expected to be lower under high rainfall regimes. This is consistent with the relatively high grazing lawn coverage in the semiarid Satara landscape compared to the mesic Lower Sabie landscape.

Different parts of the savannah landscape may be predisposed to grazing lawn formation due to the presence of resource hot-spots that attract grazers (Hempson *et al.*, 2015). This study explored the relationship between water sources as resource hot-spots and patterns in grazing lawn distribution. Grazing lawn structural attributes expressed as fractional cover (PL), maximum patch area (MPA) and connectivity (CI) were examined relative to distance from water sources. Generally, patterns in grazing lawn structure significantly correlated with distance from water sources, and was similar in both mesic and semiarid landscapes. The largest contiguous grazing lawn patches were found within 0.7 km from water sources, which is suggestive of the prevalence of grazing lawns in close proximity to waterbodies. This could be attributed to increased grazer activity around water sources (Smit, 2011; Smit *et al.*, 2007) and is consistent with observations that landscape-scale distribution of grazers is generally biased towards areas around reliable rivers and permanent waterholes (Redfern *et al.*, 2003; Smit *et al.*, 2007). For example, Smit (2011) found that different grazers of varying body mass and digestive requirements had significantly strong association with both rivers and artificial waterholes in Kruger National Park. This phenomenon is especially evident during dry seasons when moisture content of graze is low (Berry & Louw, 1982) and surface water is spatially restricted (Redfern *et al.*, 2003). Additionally, sodic sites which are highly utilised by grazers and hence have extensive grazing lawn cover, occur close to waterbodies and drainage lines, and may have contributed to the observed patterns.

It is important to note that other landscape phenomena may be influencing spatial patterns in grazing lawn distribution. For example, the prevalence of open grasslands, which is typical of the Satara landscape, may lead to the formation of more grazing lawn patches. Burkepile *et al.* (2013) found that more open savannah grasslands with sparse woody cover make attractive habitats or grazing grounds for selection by herbivores such as zebra (*Equus quagga*) and blue wildebeest (*Connochaetes taurinus*), in part to mitigate the risk of predation.

## 2.5 Conclusions

Dynamics in grazing lawn communities in southern African savannahs have been directly linked with fluctuations in mega-herbivore densities and changing fire regimes, with cascading effects on ecological processes such as nutrient cycling, plant community composition and habitat structure. Knowledge of their coverage and distribution is therefore critical to understanding habitat heterogeneity and the overall ecology of these vital grassland systems. This study presents the first attempt to develop a broad-scale approach for grazing lawn detection using very high-resolution satellite images. Application of machine learning techniques for mapping grazing lawn occurrence from WorldView-3 satellite imagery, and further analysis of their spatial structure and distribution in southern African savannahs was successfully demonstrated. The RF, SVM and MLP models produced comparable accuracies in the classification of different plant functional types (PFTs) and other land cover, all of which outperformed the CART model. Differences in grazing lawn detection accuracy followed a similar trend particularly within the same landscape (Lower Sabie). Performance for all models however reduced when they were transferred to a different landscape (Satara) even though high accuracies were achieved. Analysis of grazing lawn spatial structure and distribution showed that the Satara savannah landscape supports a relatively higher proportion of grazing lawn cover than Lower Sabie. Additionally, larger and contiguous patches persist in close proximity to water sources, which concentrate grazers within the savannah landscape, irrespective of differences in underlying environmental conditions. The proposed approach provides a novel and robust workflow for accurate and consistent landscape-scale monitoring of grazing lawns. Additionally, the findings ascertain experimental and local-scale reports on grazing lawn dynamics at a wider landscape scale, and provide timely information critical for understanding habitat heterogeneity in southern African savannahs.

---

## Chapter 3

# Spatio-temporal Analysis of Grazing Lawn Dynamics in Southern African Savannahs Using Multi-temporal High-resolution Satellite Images

---

### 3.1 Introduction

Southern African savannahs feature a heterogeneous mix of C4 grass and C3 woody vegetation communities which occur in varying spatial formations. This inherent compositional and configurational heterogeneity translates into diverse habitat conditions shaped by different bottom-up and top-down factors including nutrient availability, precipitation, fire and herbivory (Archibald, 2008; Levick *et al.*, 2009; Marston *et al.*, 2019; Sankaran & Ratnam, 2013; Sankaran *et al.*, 2008). Grazing lawn patches constitute an important part of the habitat heterogeneity in southern African savannahs and play a critical role in ecosystem processes such as nutrient cycling, plant community composition, habitat structure (Hempson *et al.*, 2019), grazer movement and predation (Hempson *et al.*, 2015). The central role of grazing lawns in savannah ecosystem processes results from their direct impact on patterns in fire behaviour and herbivory, which are the main consumers of plant biomass in savannah ecosystems.

Grazing lawn patches directly influence frequency, intensity and spread of fire within the savannah landscape by curtailing ignition and fire propagation due to their limited fuel biomass (Archibald *et al.*, 2005). In parallel, seasonal availability of grazing lawns influences grazer movement and population dynamics (Hempson *et al.*, 2015). Compared to tall grass swards, grazing lawns have higher foliar nitrogen content and are easily digestible due to their low carbon to nitrogen ratio (C:N) (Chaves *et al.*, 2006),

making them very attractive to grazers. Additionally, by having low herbaceous cover, grazing lawn patches minimize the risk of predation for herbivores (Hempson *et al.*, 2015; Voysey *et al.*, 2021). Voysey *et al.* (2021) found high browser mesoherbivore abundance on short grasses, highlighting the significance of predator avoidance as a driver for herbivore use of grazing lawns. Information on the changes in grazing lawn extent and structure is therefore critical in better understanding the overall ecology of savannahs with important implications for conservation management.

There is substantial empirical evidence of factors that shape the formation and persistence of grazing lawns in savannah landscapes (Archibald *et al.*, 2005; Archibald, 2008; Cromsigt & Beest, 2014; Donaldson *et al.*, 2018,2; Hempson *et al.*, 2015). Hempson *et al.* (2019) defined a framework for alternate grass vegetation states in savannah ecosystems, as either tall or bunch grasses which dominate fire-driven landscapes or short grass grazing lawns which dominate grazer-driven landscapes. Central to the formation of grazing lawns is grazer activity, either solely (Cromsigt & Beest, 2014) or in interaction with fire (Archibald *et al.*, 2005; Archibald, 2008; Donaldson *et al.*, 2018). Grazing lawns may form in parts of the savannah landscape where resource hot-spots concentrate grazers, such as around water sources, sodic sites and termite mounds (Archibald, 2008; Awuah *et al.*, 2020; Cromsigt & Olff, 2008; Hempson *et al.*, 2015). Additionally, new regrowth from recovering burnt areas may attract grazers and lead to the formation of new grazing lawn patches over time, or return previously established grazing lawn patches to other vegetation types if grazing frequency is reduced (Donaldson *et al.*, 2018). Under high productivity conditions, frequent grazing is required to maintain grazing lawns and prevent colonization by facultative tall grass species (Archibald *et al.*, 2005; Archibald, 2008; Hempson *et al.*, 2015). Thus, any changes in conditions within the savannah landscape that directly or indirectly influence grazer numbers or activity could result in a switch from grazing lawns to different habitat types dominated by other vegetation components and vice versa. Investigations by Donaldson *et al.* (2020) showed that disturbance from a protracted drought event between 2014 and 2017 decoupled grazers and grazing lawns, leading to post-drought losses of grazing lawn cover through self-shading and eventual colonization by tussock grasses. It appears therefore that the most immediate transition pathway in grazing lawn dynamics is a switch to and from competitive tall grass species. Yet the significance of grazing lawn transitions to and from other habitats such as woody vegetation remains unknown, particularly in the phase of widespread woody encroachment in savannah ecosystems (Marston *et al.*,

2017; Stevens *et al.*, 2017). Additionally, it is unclear how extreme disturbance like intense drought events, which are projected to increase in frequency in southern Africa with increasing atmospheric CO<sub>2</sub>, will impact dynamics in grazing lawn cover and structure.

A comprehensive analysis of grazing lawn dynamics within a spatially and temporally dynamic landscape, therefore calls for a robust broad-scale approach with provisions for quality time series data. While in-situ field-based monitoring provides detailed local insights, it is costly, inefficient and limited in spatial and temporal coverage (Marston *et al.*, 2017). Satellite remote sensing-based land cover analysis complements ground-based methods and offers a valuable tool for vegetation monitoring at large geographical scales (Ali *et al.*, 2016). Additionally, the availability of historical satellite imagery allows accurate characterisation of vegetation dynamics and provides general insights into drivers of change using change detection methods. For example, Marston *et al.* (2017) used multi-temporal medium resolution and very high resolution satellite images to identify patterns in woody encroachment in Kruger National Park. A common remote sensing-based change detection approach is the post-classification comparison of thematic maps (Peterson *et al.*, 2004). Post-classification change detection offers valuable information on the magnitude and direction of change in target land cover types (Basnet & Vodacek, 2015; Peterson *et al.*, 2004). The process requires no radiometric normalization due to the separate processing of images (Tewkesbury *et al.*, 2015). However, the accuracy of change estimates is significantly influenced by the quality of input thematic maps (Coppin *et al.*, 2004).

Savannah ecosystems exhibit significant spatial heterogeneity defined by different mosaics of tree, shrub, grass and bare ground patches. This coupled with the spectral similarity of different vegetation components creates significant challenges for accurate vegetation classification and subsequent monitoring (Awuah *et al.*, 2020; Kaszta *et al.*, 2016; Symeonakis *et al.*, 2018). Compared to coarse and medium resolution images, very high-resolution (VHR) satellite imagery can help to overcome pixel mixing and discriminate between different vegetation components (Awuah *et al.*, 2020; Kaszta *et al.*, 2016; Marston *et al.*, 2017). In the same savannah landscape context, Marston *et al.* (2017) successfully demonstrated the utility of VHR IKONOS, QuickBird and WorldView-2 images for discriminating different vegetation components at very high accuracies (> 85 % overall accuracy).

Analysis of changes in savannah vegetation structure over space and time largely focuses on the woody vegetation components (Bond, 2008; Bond & Midgley, 2000; Case *et al.*, 2019,2; Fensham *et al.*, 2009; Marston *et al.*, 2017; Stevens *et al.*, 2017), with very little attention on the understorey grassy layer. In particular, no studies exist on the changes in grazing lawn extent and spatial configuration, the dominant pathways of transitions from grazing lawns to other habitats and vice versa, as well as how disturbance events influence these dynamics. Against this backdrop, this study aims to analyse spatial and temporal dynamics in grazing lawn cover using VHR satellite imagery by: (i) identifying changes in grazing lawn cover and structure; (ii) highlighting the dominant drivers of grazing lawn dynamics in the broader savannah landscape context; and (iii) investigating how drought impacts dynamics in grazing lawn cover and structure using the 2015/2016 drought as a case study.

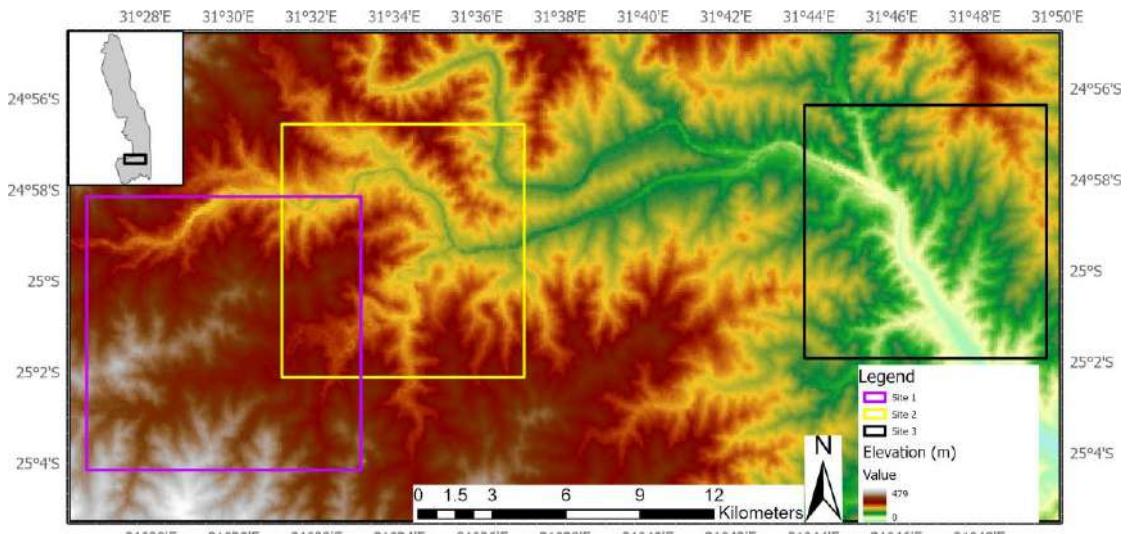
## 3.2 Materials and Methods

### 3.2.1 Study Sites

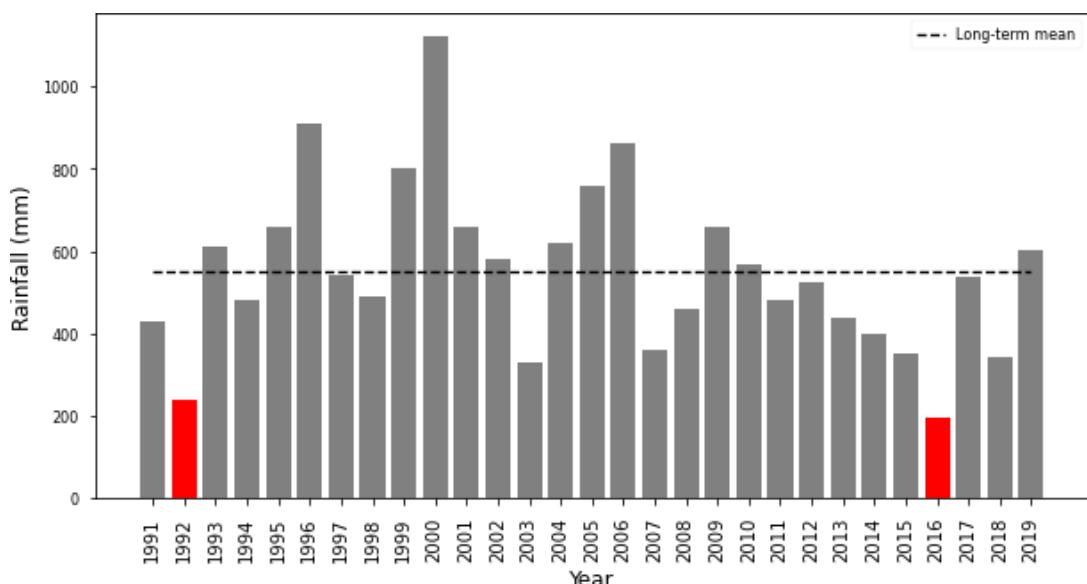
Three sites were used for this study (Figure 3.1), all located in the southern portion of Kruger National Park (KNP), South Africa, specifically within the Skukuza and Lower Sabie regions of KNP. The selected sites fall on the granitic substrate of KNP and cover a range of moderately undulating topographies, with elevation ranging from about 200m to 400m above sea level (Figure 3.1). KNP is a protected area spanning approximately 20,000 km<sup>2</sup> with minimal human footprint and a great diversity of flora and fauna shaped by their interactions with natural environmental phenomena, making it a globally important site for ecological studies. The selected sites are thus reflective of these natural savannah ecosystem dynamics which play out in the context of grazing lawn formation and persistence. Additionally, Site 3 encompasses the Nkhulu exclosure designed to examine the impact of mammalian herbivores and fire on savannah vegetation structure and distribution (Siebert & Eckhardt, 2008).

The study sites are characterized by semi-arid and mesic savannah conditions over the Skukuza and Lower Sabie regions respectively (Venter *et al.*, 2003), with corresponding mean annual precipitation (MAP) values of ~ 550 mm (Skukuza) and ~ 602 mm (Lower Sabie) (South African National Parks, 2021). The growing season in KNP occurs between the summer months of November and April, during which 84%

of total annual rainfall is received on average (MacFadyen *et al.*, 2018). The most recent major drought event, since that of 1991/1992, occurred in 2015/2016 (Figure 3.2) which coupled with satellite image availability presents an opportunity to assess the effect of drought on grazing lawn cover.



**Figure 3.1:** Map showing study site locations (Sites 1, 2 and 3) overlaid on SRTM Digital Elevation Model (DEM). Inset map shows Kruger National Park (KNP) boundary and the study region in southern KNP.



**Figure 3.2:** Total annual rainfall for Skukuza, Kruger National Park, South Africa, from 1991 to 2019. Major drought years are highlighted in red. Rainfall data was obtained from (South African National Parks, 2021)

### **3.2.2 Data**

#### **Reference Data and Classification Scheme**

Field surveys were conducted in November 2002 and July 2014 for all sites (i.e. Sites 1, 2 and 3) and between June and July 2019 for only Site 3. A random sampling approach was used to generate a total of 285 and 188 points which were visited during the 2002 and 2014 (including 66 points that were common to both years) fieldwork campaigns respectively (Marston *et al.*, 2017). On the other hand, 111 points systematically distributed within 200 m buffer region along access roads were used during the 2019 fieldwork campaign. To minimize edge effects on the surveyed vegetation cover, the systematic sample points were placed beyond 100 m distance from the access roads. During each of the field surveys, the predefined sample points were located using a handheld GPS device. At each sample location, cardinal photographs were taken in the directions of north, east, south and west successively. Descriptions of site condition and dominant land cover were taken at each sample location, including compass bearings to other significant surrounding land cover patches, which enabled the generation of further sample points based on satellite image interpretation. In addition to the point samples used during field surveys, further land cover information was recorded from additional randomly generated points based on visual interpretation of satellite images, augmented by contextual knowledge from field visits, field photos, very high-resolution (VHR) Google satellite scenes as well as from obvious image features (e.g. water body, bare land etc) visible within the VHR images used for classification. This allowed the retrieval of enough spectral variability within the same target land cover type, which is necessary for improved classification outcomes. In total, 1188, 1120 and 3398 reference points were generated for 2002, 2014 and 2019 respectively.

Together with water bodies, built-up and bare cover, three vegetation classes were identified based on structure in order to distinguish grazing lawns. All woody vegetation, primarily composed of trees and shrubs were collapsed into one vegetation class named woody. Grasses, which are the dominant herbaceous vegetation, were distinguished into bunch grasses and grazing lawns. Bunch grasses have upright growth form and occur as tall and often dense grass patches within the savannah landscape. In this study, grass patches greater than 20 cm in height were identified as bunch grasses. In contrast, grazing lawns are distinct short grass patches with

prostrate growth forms and spread laterally by means of stolons. Grass patches in this category, with height less than or equal to 20 cm, and often with smooth textured appearance in satellite imagery, were thus identified as grazing lawns. An overview of all target land cover types is provided in Table 3.1.

**Table 3.1:** Description of target land cover classes used for image classification.

ID	Land cover	Description
1	Woody	All woody vegetation components at sample locations, mainly composed of trees and shrubs.
2	Bunch grass	Tall grass patches greater than 20 cm in height, often showing dense growth with an upright growth pattern.
3	Grazing lawn	Short grass patches which exhibit stoloniferous growth pattern and less than or equal to 20 cm in height.
4	Water body	Rivers, lakes and ponds, or any collection of surface water visible within the landscape.
5	Bare	Barren areas with no vegetation cover and showing bare soil surface; also including bare rocky outcrops.
6	Built-up	Artificial structures including buildings, rooftops and asphalt or concrete coated surfaces.

### Image Data and Preprocessing

Multi-temporal very high-resolution satellite images acquired in the dry season were used in this study (Table 3.2). Phenological differences in vegetation cover peak during the dry season leading to maximum spectral differentiation of different plant functional types (Bucini *et al.*, 2009), particularly in heterogeneous landscapes (Marston *et al.*, 2017). Additionally, there is generally less prevalence of atmospheric effects on dry season image acquisition. As such, dry season satellite image acquisition is well-known to provide better classification outcomes in savannah environments (Brandt *et al.*, 2017; Marston *et al.*, 2017). In total, nine satellite images were acquired from five VHR sensors, with three images per study site, representing three different temporal acquisition windows. Table 3.2 shows a summary of image properties.

All images had been processed to surface reflectance by the vendor. Of all image acquisitions, only the IKONOS and QuickBird images from 2002 covering sites 1 and 2 respectively (Table 3.2) had some cloud cover. Pre-processing thus involved cloud and cloud shadow masking using a semi-automatic approach based on band thresholding (Marston *et al.*, 2017). This involved, initially, the creation of a cloud/non-cloud binary image through band thresholding. The process was repeated multiple

times to fine-tune the detection of cloudy pixels. Areas with cloudy pixels were then isolated via a vectorization process. A cloud shadow mask was created by applying a geographical offset to the cloud mask. Affected areas that were not captured in the process of cloud and cloud shadow masking were manually digitized to exclude them from further analysis. In addition to the cloud and cloud shadow masks, a mask for burnt areas was created using a burn map derived from initial unsupervised clustering of affected images using K-Nearest Neighbour (KNN). The burn map was refined through manual digitisation of missed burn pixels, and was later used in post-processing reclassification of burnt areas (see section 3.2.5). To ensure consistency in temporal change analysis, all images were co-registered to the 2019 acquisition for each site. Further, for each site, images were resampled to spatially match the image with the coarsest spatial resolution. For example, at Site 3, the 2019 WorldView-3 acquisition with 2m spatial resolution was resampled to 4m spatial resolution to be consistent with the IKONOS acquisitions from September 2014 and April 2002.

**Table 3.2:** Overview of satellite images used for land cover classification. Spectral bands: B = blue, C = cirrus, CA = coastal aerosol, CB = coastal blue, G = green, NIR = near infra-red, R = red, RE = red edge, SWIR = short-wave infrared, Y = yellow.

Site	Sensor	Acquisition Date	Spatial Resolution (m)	Spectral Bands
1	IKONOS	March 7, 2002	4 (resampled to 4.7)	B, G, R, NIR
1	IKONOS	September 27, 2014	4 (resampled to 4.7)	B, G, R, NIR
1	Planet	June 2019	4.7	B, G, R, NIR
2	QuickBird	December 9, 2002	2.8	B, G, R, NIR
2	WorldView-2	October 14, 2014	2 (resampled to 2.8)	CB, B, G, Y, R, RE, NIR-1, NIR-2
2	WorldView-3	July 1, 2019	2 (resampled to 2.8)	CB, B, G, Y, R, RE, NIR-1, NIR-2
3	IKONOS	April 28, 2002	4	B, G, R, NIR
3	IKONOS	September 27, 2014	4	B, G, R, NIR
3	WorldView-3	July 1, 2019	2 (resampled to 4)	CB, B, G, Y, R, RE, NIR-1, NIR-2

### 3.2.3 Image Classification and Accuracy Assessment

The reference point locations (see section 3.2.2) were used to generate polygons from spectrally homogeneous areas in close proximity to the sample points, to which the target land cover labels were assigned (see Table 3.1). The polygon extents were then used to extract reference pixels which were divided according to a 70:30 ratio for supervised classification and accuracy assessment respectively. Compared with other spatial objects such as points and lines, the use of polygon objects in training data extraction allows for the representation of enough intra-class spectral

variability leading to better classification outcomes (Corcoran *et al.*, 2015; Ma *et al.*, 2017). Prior to training data extraction, the input image features were normalized by subtracting the mean and scaling to a unit variance to minimize bias during training and classification (Singh *et al.*, 2015).

The Random Forest (RF) algorithm was used in the supervised classification process. The RF algorithm is an ensemble of decision trees which are parameterised with several independent random subsets of input vectors selected through bootstrap aggregation (or bagging) (Breiman, 2001). The remaining samples, or so called 'out-of-bag' samples (OOB) can be used for evaluation and variable importance estimation (Breiman, 2001; Eisavi *et al.*, 2015). Through bagging and random subspacing, multicollinearity and generalisation error is minimised (Abdi, 2020; Breiman, 2001), which gives the RF classifier its robustness compared to other algorithms. During classification, each decision tree within the ensemble contributes a unit vote to assign input vectors to different target classes, and the final classification is determined from the most frequently voted class (i.e. majority voting). The RF algorithm uses Gini Index as a measure for selecting suitable features per class at each node in order to maximize dissimilarities between classes (Rodriguez-Galiano *et al.*, 2012). A combination of features is therefore used to grow each decision tree within the ensemble to a maximum depth. The most important user-defined parameter required to instantiate the RF algorithm is the number of decision trees, which was set to 2000 following hyperparameter optimization based on randomised grid search.

Classification accuracy was assessed for overall classification and individual land cover classes using precision, recall and F-score accuracy metrics, all of which were calculated from a confusion matrix (Congalton & Green, 2019). The accuracy metrics were calculated using Equations 3.1, 3.2 and 3.3.

$$Precision = \frac{tp}{tp + fp} \quad (3.1)$$

$$Recall = \frac{tp}{tp + fn} \quad (3.2)$$

$$F\_score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3.3)$$

where tp, fp and fn represent the number of true positive, false positive, and false negative cases, respectively.

### **3.2.4 Post-classification Characterisation of Grazing Lawn Structure**

The spatial configuration of grazing lawn patches in the land cover maps derived for each site over time was quantified using spatial metrics (Herold, 2001; Mcgarigal & Marks, 1995), from which grazing lawn structure was determined. Spatial metrics provide landscape-scale information on density, aggregation and size distribution of land cover patches (Herold, 2001). In the case of grazing lawns, such information is vital in highlighting important ecological indicators such as habitat heterogeneity (Awuah *et al.*, 2020) and processes such as fire behavior (Van Langevelde *et al.*, 2003) with consequences for biodiversity. A combination of spatial metrics including Patch Density (PD), Edge Density (ED) and Cohesion Index (CI) (Table 3.3) were calculated in the R programming environment (Development Team, 2016) using the "SpatialEco" library (Evans & Ram, 2019).

### **3.2.5 Post-processing and Change Analysis**

Following classification, a majority filter with 8-pixels neighbourhood kernel (i.e. 3x3 window) was used to minimize noisy pixels (i.e. 'salt-and-pepper' effect), which is a common phenomenon in pixel-based classification of very high-resolution satellite images (Hirayama *et al.*, 2019). Additionally, burnt patches specifically in the 2014 imagery of Site 1 and Site 3 were reclassified into a 'burn scar' category using the burn-mask generated during preprocessing (section 3.2.2).

Bi-temporal changes in grazing lawn cover and structure were analysed for 2002 to 2014 and 2014 to 2019. Grazing lawn cover expressed in terms of persistence, gains, losses and transitions to and from other land cover types were analysed using a transition matrix (Pontius Jr *et al.*, 2004). For a pair of classifications of the same landscape derived from two periods (e.g. time-1 and time-2), the transition matrix compares the proportions of each land cover type within the landscape from which persistence and transitions in the different land cover categories can be identified. An important challenge for post-classification change detection is the reliance on the classification accuracy of the maps being compared. This was minimised by limiting the analysis to changes between grazing lawns and the other vegetation classes (i.e. woody and bunch grasses) and the bare category. Additionally, bare areas visibly identified as roads / trails and those in stream and river beds (due to low water levels or drying-up) were disregarded with up to a five meter buffer distance (for roads /

**Table 3.3:** . Summary of selected spatial metrics.  $A$  is the total landscape area,  $a_{ij}$  is area of patch  $ij$ ,  $p_{ij}$  is the perimeter of patch  $ij$ ,  $n_i$  is the number of patches for patch  $i$ , and  $Z$  is the total number of cells in the landscape,  $e_{ik}$  is the total edge in landscape between patch types  $i$  and  $k$ ,  $m'$  is number of patch types present in the landscape including landscape border (Mcgarigal & Marks, 1995).

Spatial Metric	Formula	Description	Value Range
Patch Density (PD)	$\frac{n_i}{A} * (10000) * (100)$	Number of patches per 100 ha for a patch type within a landscape.	$PD > 0$ . Increases without limit, with larger values showing more fragmentation.
Edge Density (ED)	$\frac{\sum_{k=1}^{m'} e_{ik}}{A} * (10000)$	Sum of lengths of boundary segments of the corresponding patch type divided by the total landscape area.	$ED \geq 0$ . Increases without limit. $ED = 0$ when the entire landscape and landscape border consists of only one patch of the corresponding patch type. Larger ED values depict a corresponding high fragmentation.
Cohesion Index (CI)	$h_i = 1 - \frac{\sum_{j=1}^n p_{ij}}{\sum_{j=1}^n p_{ij} / a_{ij}}$ $h_h = i_{-1} - \sqrt{\frac{1}{Z}}$	* Provides information on the physical connectedness of patches within a landscape.	$0 < CI < 100$ . Increasing CI correspond to a higher degree of patch connectedness or low fragmentation, whereas CI approaching zero means increasing patch subdivisions.

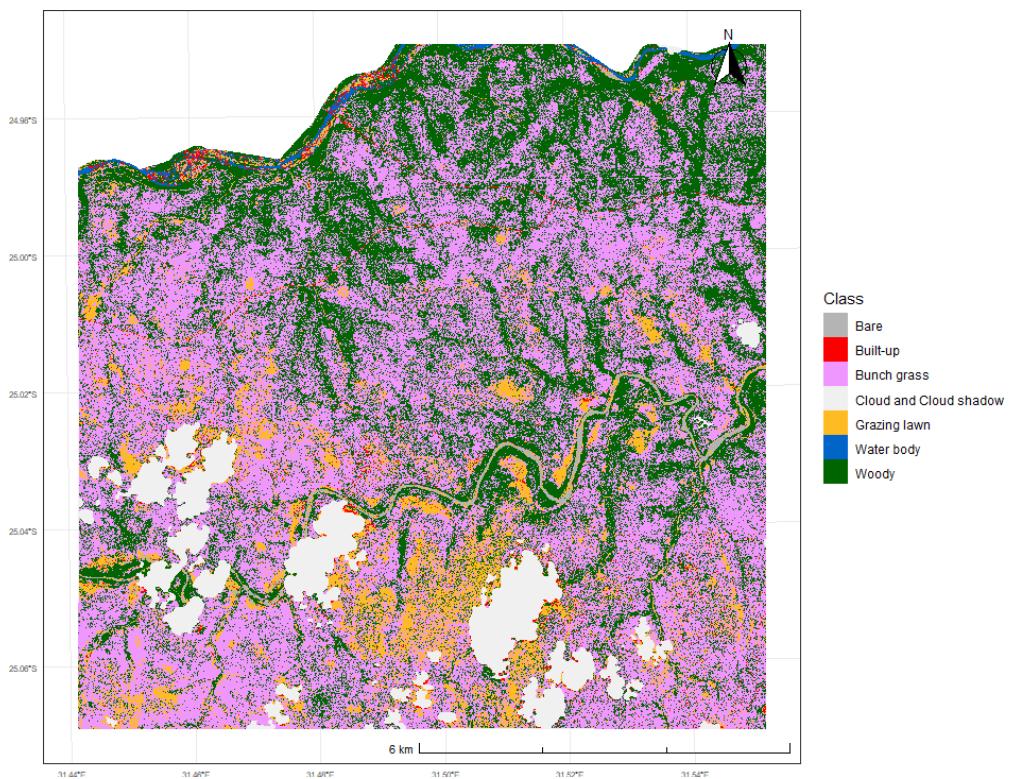
trails). 'Burn scars' generated from post-processing were also disregarded in the analysis since such static fire events do not give a true reflection of the influence of fire on vegetation dynamics, particularly when compared maps are years apart. Further, areas within and around dense human settlements, particularly the Skukuza camp, with recreational turfs (e.g. golf course) that are similar to grazing lawns in structure and spectra were manually digitized and excluded from the analysis. The above modifications helped to minimise the effects of classification errors due to the reduction in area considered. Further, it is consistent with the ecology of grazing lawn dynamics, including persistence; switch to other vegetation types such as tall bunch grass due to low grazing frequency; or switch to bare patches due to degradation from overgrazing and / or trampling (Hempson et al., 2015).

## **3.3 Results**

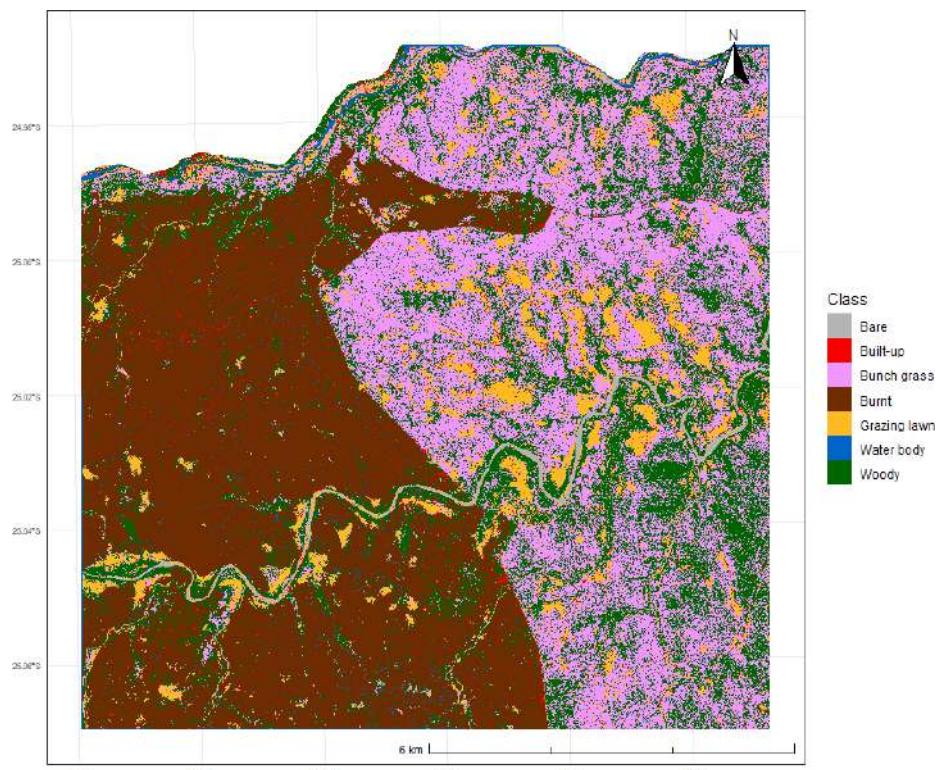
### **3.3.1 Land Cover Classification**

Multi-temporal land cover classifications are presented for Site 1 (Figures 3.3, 3.4 and 3.5), Site 2 (Figures 3.6, 3.7 and 3.8) and Site 3 (Figures 3.9, 3.10 and 3.11). A clear evolution in the extent of grazing lawn distribution can be observed for Site 1 and Site 3. In both landscapes, grazing lawns extended to cover areas that previously had no grazing lawn presence from 2002 to 2014 and 2014 to 2019. In contrast, no clear pattern in the spread of grazing lawns was observed for Site 2 ((Figures 3.6, 3.7 and 3.8). The widest distribution of grazing lawns at Site 2 was observed in 2014 (Figure 3.7), while grazing lawn patches were mostly confined to the lower left fringes of the landscape in 2019 (Figure 3.8). It should be noted that the 2002 image for Site 2 had a significant amount of cloud cover, and thus, did not give the true distribution of grazing lawns.

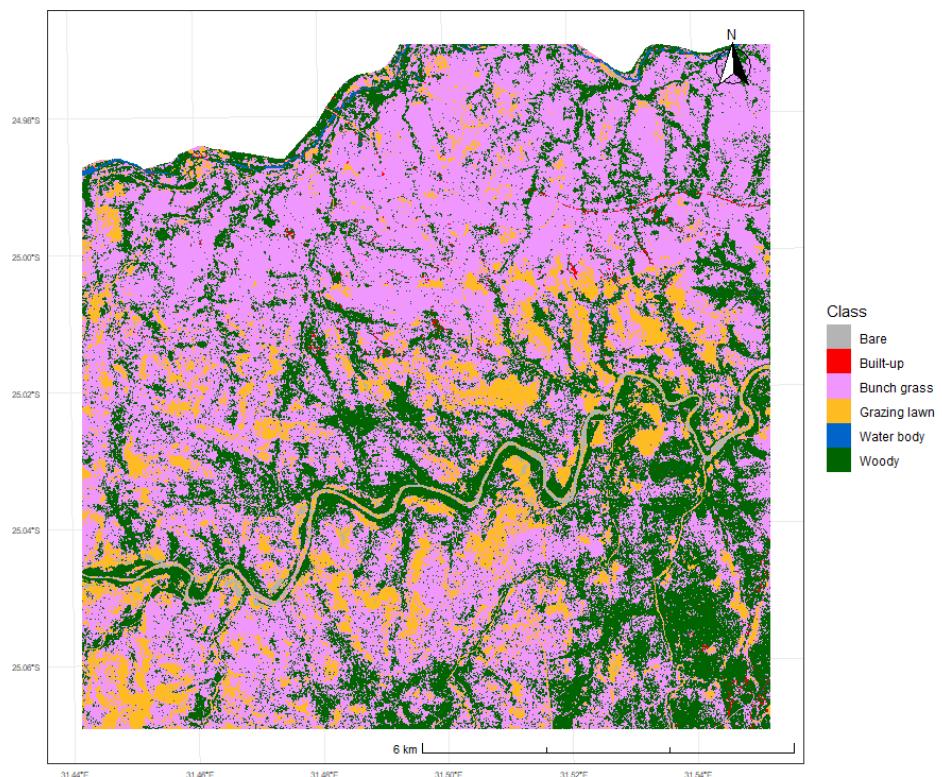
Table 3.4 shows accuracy scores for individual land cover categories and the overall classification derived from confusion matrices (Appendix B). Generally, land cover classification accuracies were very high with mean F-scores over the considered time periods of 0.91, 0.98 and 0.98 for Site 1, Site 2 and Site 3 respectively, for all considered time periods. High classification accuracy is critical to minimize errors in post-classification change detection. For the land cover categories considered in analysis of changes from grazing lawns to other classes and vice versa, average F-scores for Site 1, Site 2 and Site 3 respectively were: 0.92 (2002), 0.98 (2014) and 0.98 (2019) for woody; 0.89 (2002), 0.93 (2014) and 0.96 (2019) for bunch grass; 0.96 (2002), 0.96 (2014) and 0.97 (2019) for grazing lawns; and 0.91 (2002), 0.99 (2014) and 0.99 (2019) for bare.



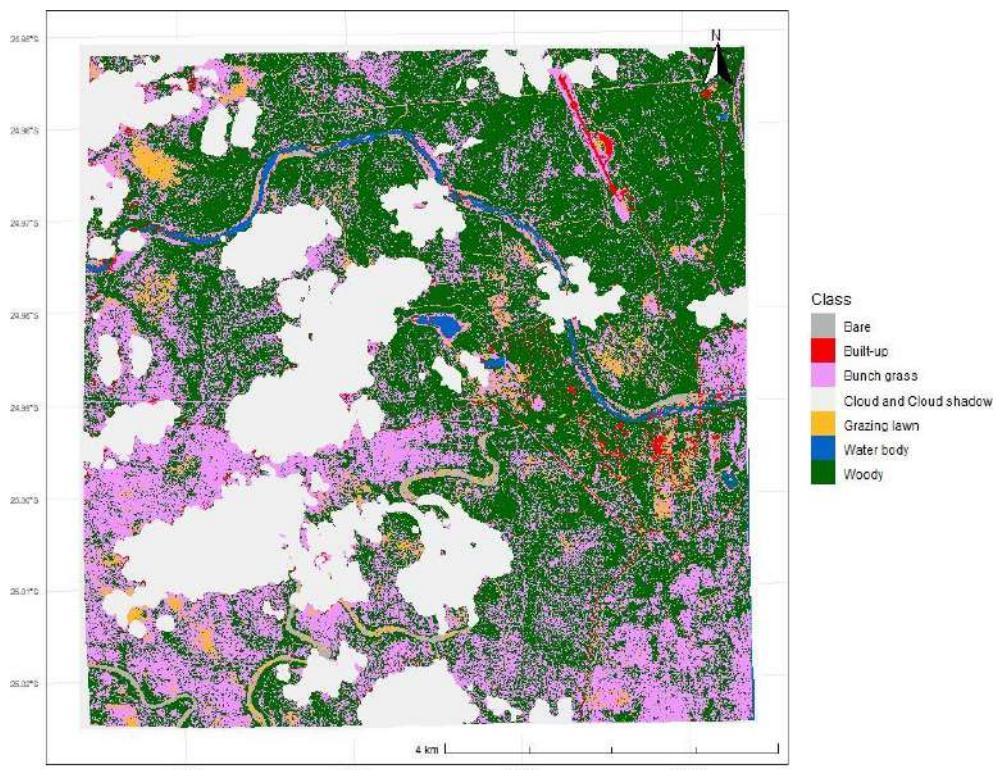
**Figure 3.3:** Land cover map for Site 1 from IKONOS image in 2002.



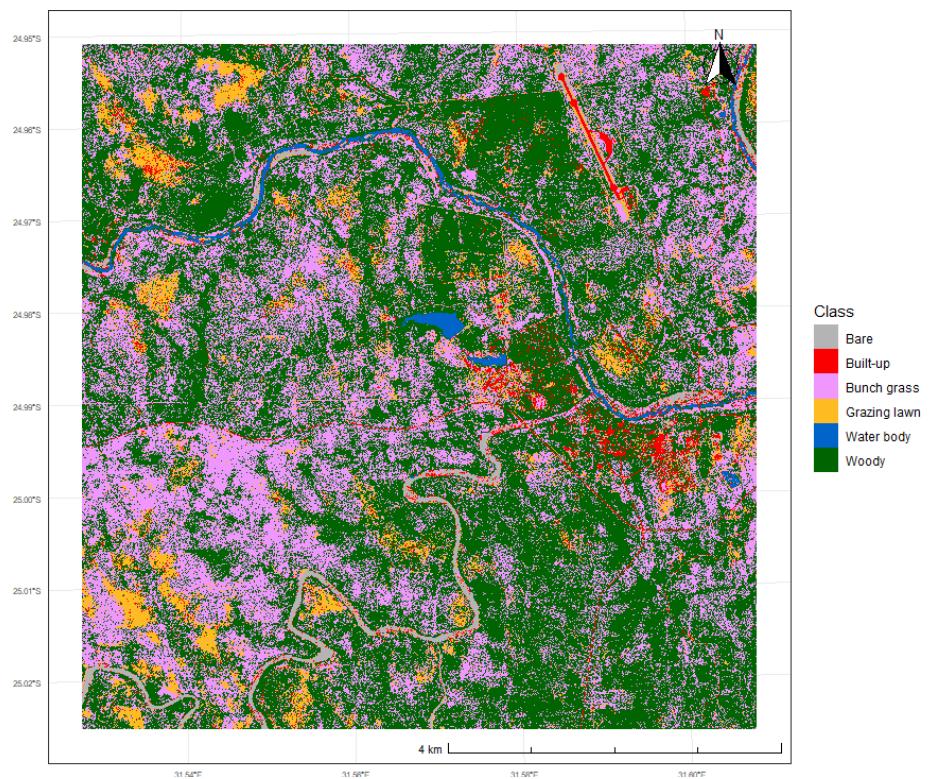
**Figure 3.4:** Land cover map for Site 1 from IKONOS image in 2014.



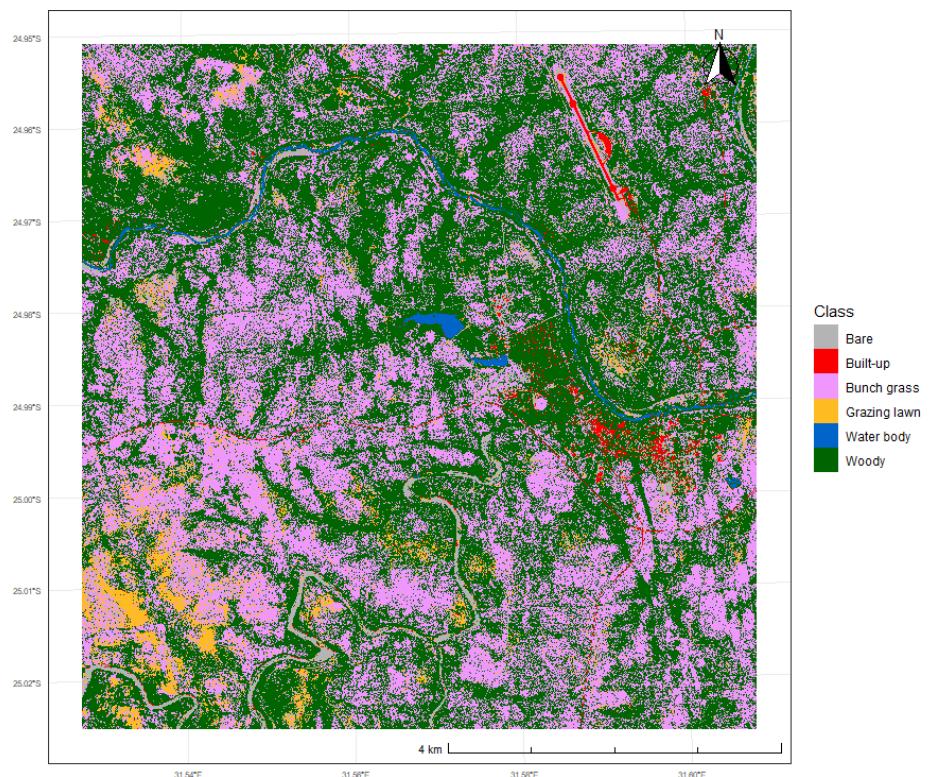
**Figure 3.5:** Land cover map for Site 1 from Planet image in 2019.



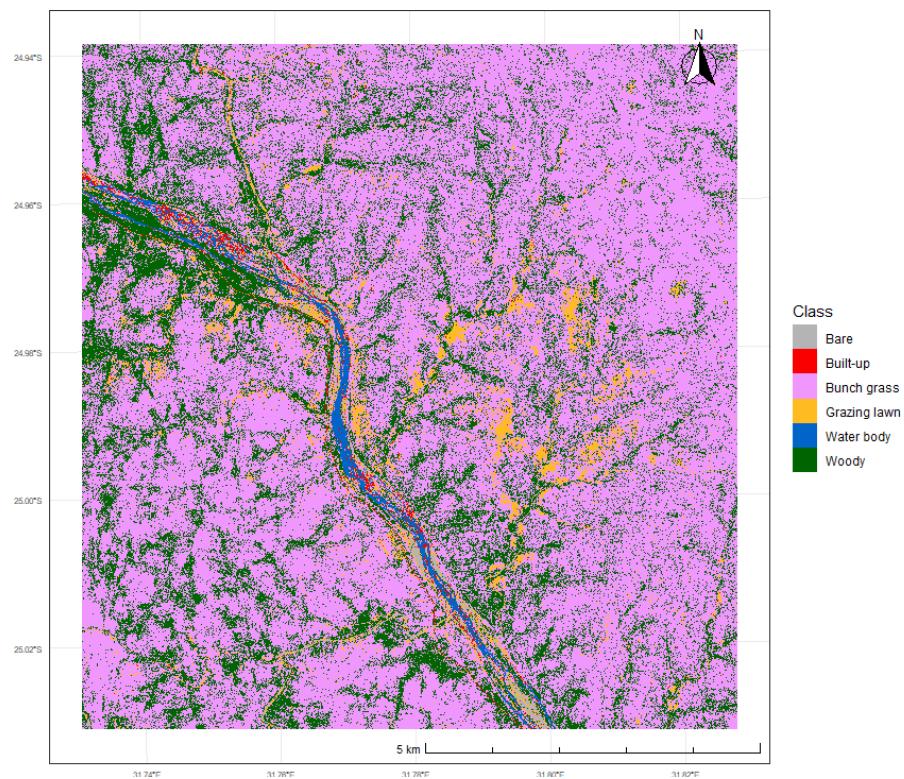
**Figure 3.6:** Land cover map for Site 2 from QuickBird image in 2002.



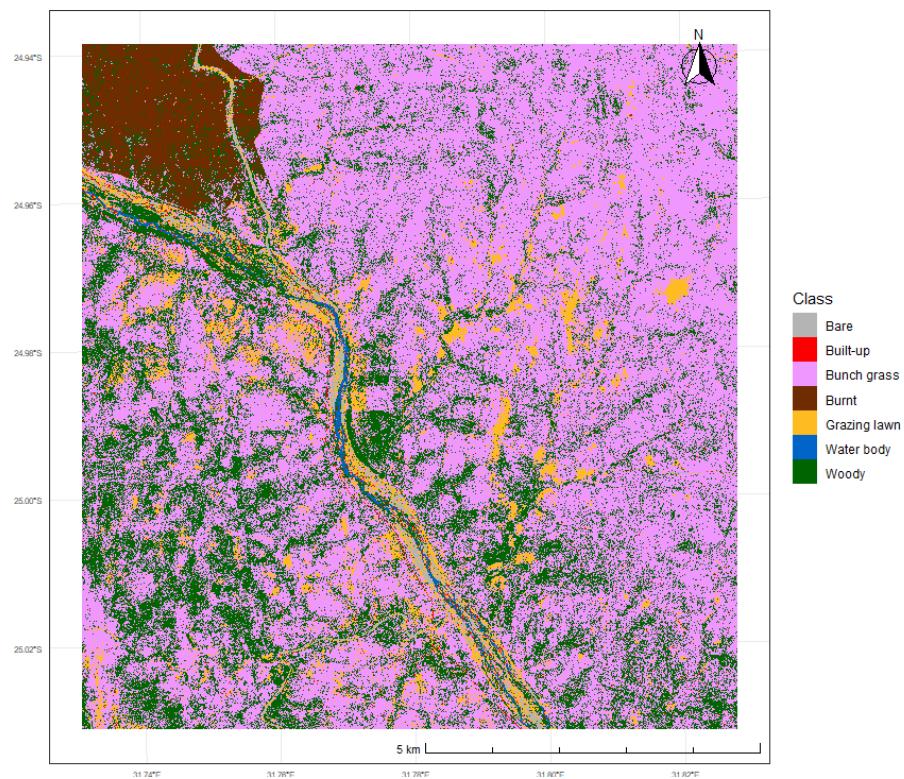
**Figure 3.7:** Land cover map for Site 2 from WorldView-2 image in 2014.



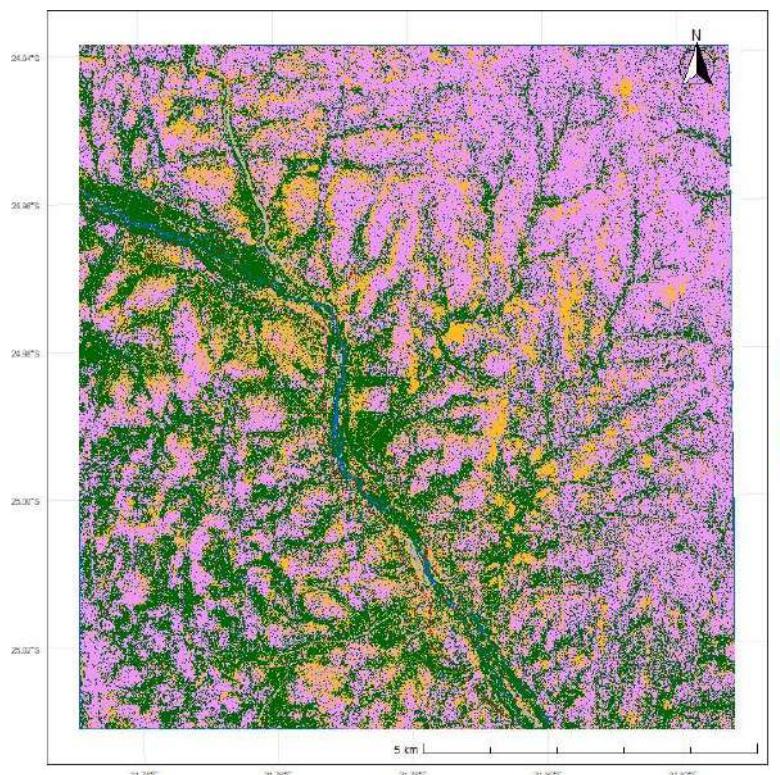
**Figure 3.8:** Land cover map for Site 2 from WorldView-3 image in 2019.



**Figure 3.9:** Land cover maps for Site 3 from IKONOS image in 2002.



**Figure 3.10:** Land cover map for Site 3 from IKONOS image in 2014.



**Figure 3.11:** Land cover map for Site 3 from WorldView-3 image in 2019.

### 3.3.2 Changes in Grazing Lawn Cover and Structure

Table 3.5 shows a summary of area coverage and changes for the land cover categories considered in grazing lawn change analysis. To ensure consistency in area comparison and change assessment, areas affected by cloud pollution and burnt areas were disregarded across all years. For example, for Site 2, the equivalent areas in 2014 and 2019 of cloud covered areas in 2002 were not included in the calculation of land cover area. For each site, changes in land cover area were expressed as a proportion of the entire landscape area.

Generally, dynamics in grazing lawn area showed a similar trend to the extent of grazing lawn distribution presented for Site 1 (Figures 3.3, 3.4 and 3.5), Site 2 (Figures 3.6, 3.7 and 3.8) and Site 3 (Figures 3.9, 3.10 and 3.11). For Site 1, grazing lawn area coverage increased sharply between 2002 ( $3.89 \text{ km}^2$ ) and 2014 ( $6.56 \text{ km}^2$ ) equivalent to 2.15% of the total landscape area and a marginal increase of 0.71% between 2014 and 2019 ( $7.44 \text{ km}^2$ ). Grazing lawn area coverage peaked in 2014 ( $2.48 \text{ km}^2$ ) for Site 2, with a 2.05% increase from 2002 ( $1.11 \text{ km}^2$ ). However, between 2014 and 2019

**Table 3.4:** Summary of land cover classification accuracy scores for individual land cover categories and overall classification. Overall accuracy scores are weighted averages of accuracy scores from individual land cover classifications. Acronyms P, R and F represent precision, recall and F-score respectively.

Land Cover	Site 1								
	2002			2014			2019		
	P	R	F	P	R	F	P	R	F
Woody	0.97	0.96	0.97	1.00	0.95	0.97	0.90	0.87	0.89
Bunch grass	0.95	0.95	0.95	0.88	0.96	0.92	0.88	0.83	0.85
Grazing lawns	0.97	0.99	0.98	0.97	0.99	0.98	0.91	0.96	0.93
Water body	0.96	0.98	0.97	1.00	1.00	1.00	0.86	0.86	0.86
Bare	0.99	1.00	0.99	0.99	0.99	0.99	0.83	0.77	0.80
Built-up	0.94	0.89	0.91	0.88	0.80	0.84	0.81	0.81	0.81
Overall	0.97	0.97	0.97	0.97	0.97	0.97	0.88	0.87	0.88
Land Cover	Site 2								
	2002			2014			2019		
	P	R	F	P	R	F	P	R	F
Woody	0.99	0.99	0.99	0.99	0.98	0.99	0.98	0.97	0.98
Bunch grass	0.96	0.97	0.96	0.91	0.98	0.94	0.89	0.89	0.89
Grazing lawns	0.94	0.98	0.96	0.99	0.98	0.99	0.91	0.92	0.92
Water body	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99
Bare	0.99	0.98	0.99	0.99	1.00	0.99	0.98	0.98	0.98
Built-up	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Overall	0.98	0.98	0.98	0.99	0.99	0.99	0.97	0.97	0.97
Land Cover	Site 3								
	2002			2014			2019		
	P	R	F	P	R	F	P	R	F
Woody	1.00	0.97	0.98	0.99	0.98	0.98	0.96	0.97	0.97
Bunch grass	0.96	0.99	0.98	0.97	0.99	0.98	0.92	0.92	0.92
Grazing lawns	0.97	0.99	0.98	1.00	0.99	0.99	0.97	0.98	0.98
Water body	1.00	1.00	1.00	0.93	1.00	0.97	0.98	0.98	0.98
Bare	1.00	1.00	1.00	0.98	1.00	0.99	0.98	1.00	0.99
Built-up	0.97	0.95	0.98	0.97	0.85	0.91	0.99	0.96	0.98
Overall	0.99	0.99	0.99	0.98	0.98	0.98	0.97	0.97	0.97

(1.63 km<sup>2</sup>), grazing lawn area coverage decreased by 1.27% of the total landscape area. Similar to Site 1, a progressive increase in grazing lawn coverage was recorded for Site 3. Between 2002 (3.08 km<sup>2</sup>) and 2014(5.39 km<sup>2</sup>), the grazing lawn extent of Site 3 increased by 2.29%, and by a significant proportion of 7.53% between 2014 and 2019 (12.99 km<sup>2</sup>) (Table 3.5). When considering the overall analysis period from 2002 to 2019, grazing lawn extent increased by 2.86%, 0.77% and 9.82% for Site 1, Site 2 and Site 3 respectively.

**Table 3.5:** Summary of land cover area and changes for all sites. Only land cover classes relevant for grazing lawn dynamics are considered. Changes (%) correspond to proportion of the total area for each study site.

Land Cover	Area Coverage and Percentage Changes					
	2002 (km <sup>2</sup> )	2014 (km <sup>2</sup> )	2019 (km <sup>2</sup> )	2002/2014 Change(%)	2014/2019 Change(%)	2002/2019 Change(%)
Site 1						
Woody	25.35	22.45	20.18	-2.33	-1.83	-4.16
Bunch grass	32.21	30.34	34.15	-1.51	3.07	1.56
Grazing lawn	3.89	6.56	7.44	2.15	0.71	2.86
Bare	0.13	0.09	0.16	-0.03	0.06	0.02
Site 2						
Woody	24.01	24.23	22.54	0.33	-2.53	-2.20
Bunch grass	16.41	14.23	16.48	-3.26	3.36	0.10
Grazing lawn	1.11	2.48	1.63	2.05	-1.27	0.77
Bare	0.09	0.20	1.14	0.16	1.41	1.57
Site 3						
Woody	21.31	22.83	31.51	1.51	8.61	10.11
Bunch grass	65.26	60.75	43.79	-4.47	-16.81	-21.28
Grazing lawn	3.08	5.39	12.99	2.29	7.53	9.82
Bare	0.45	0.75	0.79	0.29	0.04	0.34

Apart from changes in the overall extent of grazing lawns, the specific direction of change between grazing lawns and other land cover types was revealed using a transition matrix (see Section 3.2.5). Based on the ecology of grazing lawn dynamics (Hempson *et al.*, 2015), three main change trajectories can be identified. This includes: (i) grazing lawns taking over areas that are either covered by other vegetation components such as bunch grasses and shrubs or bare soil surface; (ii) grazing lawn patches being colonized by other vegetation types; and (iii) grazing lawns degrading and dying off to create bare soil surfaces. Thus, the analysis of grazing lawn dynamics was limited to and from woody, bunch grass and bare land cover categories. The full transition matrices for all sites is available in Appendix B.

Table 3.6 shows the magnitude of different grazing lawn change trajectories expressed in terms of the proportion of total landscape area for Site 1, Site 2 and Site 3. As expected, the most gains in grazing lawn area coverage were from bunch grass. This pattern was prevalent for all sites and periods, although there were variations between different periods. For example, greater proportions of bunch grass to grazing lawn switches were recorded between 2002 and 2014 than between 2014 and 2019

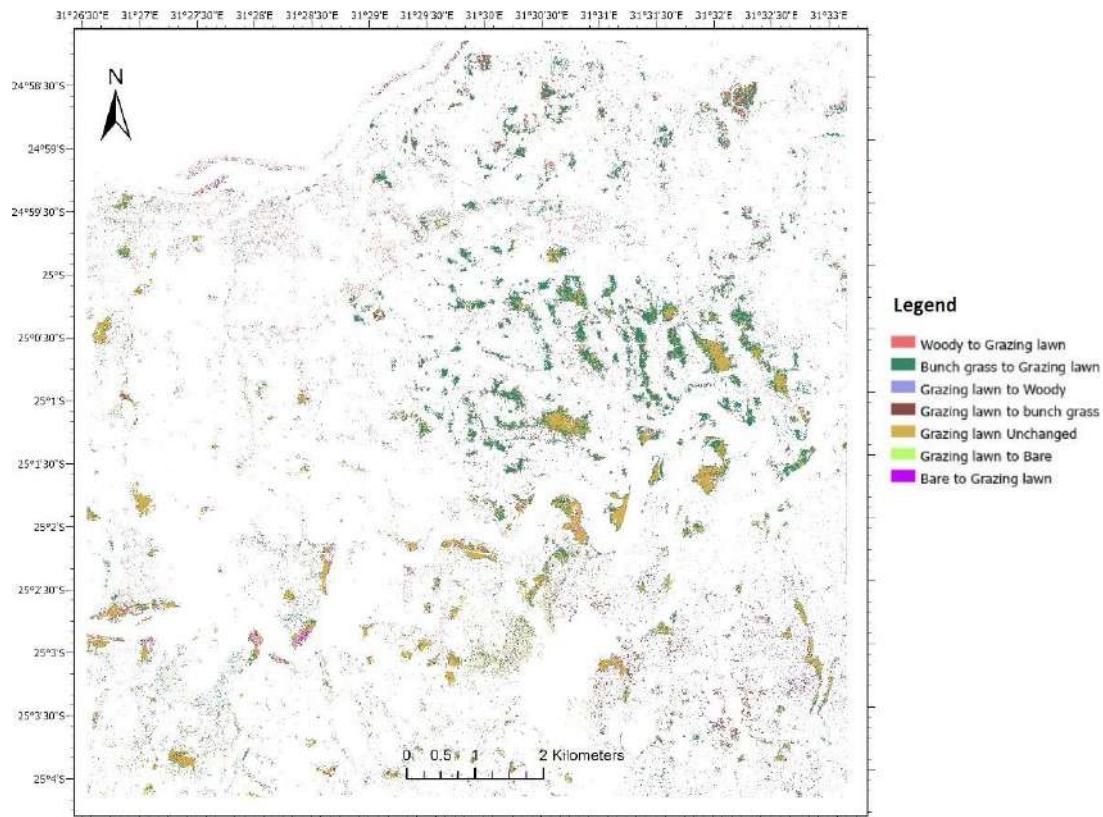
for Site 1 (2002/2014 = 3.51% [4.36 km<sup>2</sup>] and 2014/2019 = 2.01% [2.49 km<sup>2</sup>]) and Site 2 (2002/2014 = 1.89% [1.26 km<sup>2</sup>] and 2014/2019 = 1.25% [0.83 km<sup>2</sup>]). On the contrary, Site 3 recorded more bunch grass to grazing lawn switches between 2014 and 2019 (8.39% [8.46 km<sup>2</sup>]) than between 2002 and 2014 (2.93% [2.95 km<sup>2</sup>]).

**Table 3.6:** Relative change trajectories of grazing lawn cover expressed as proportion of the overall landscape area.

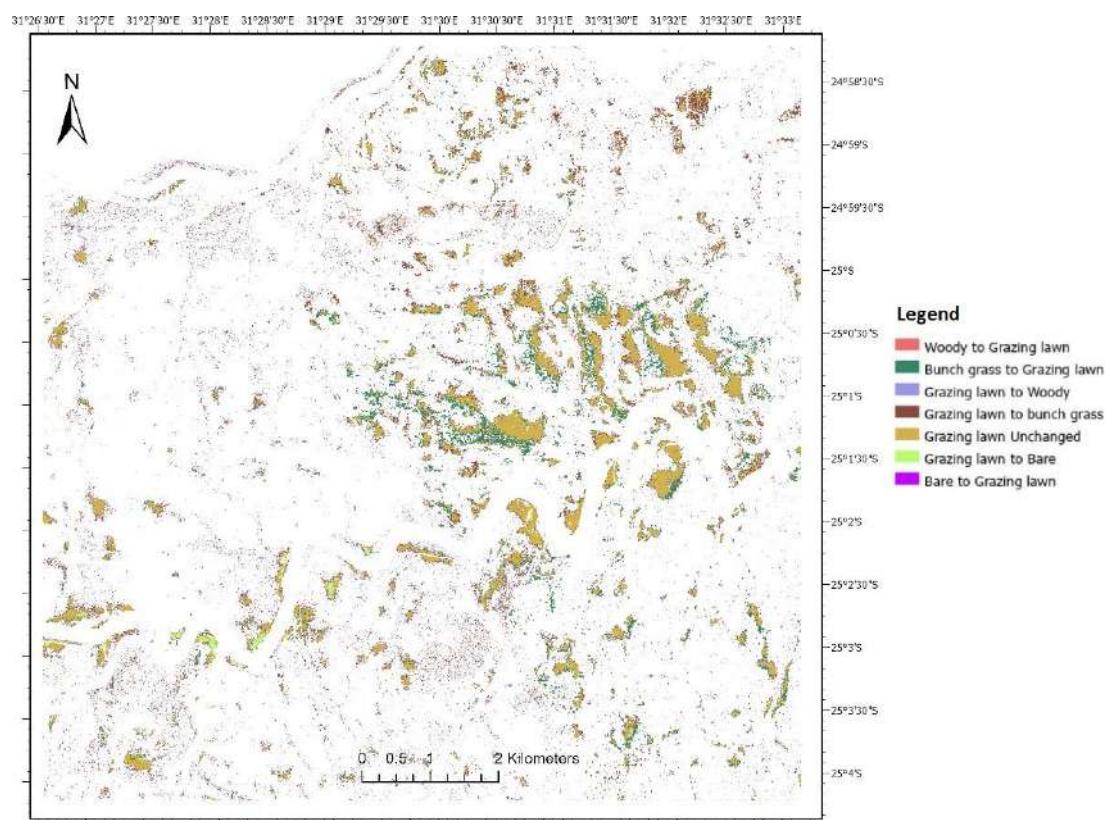
Change Trajectory	Proportion of Landscape Area (%)		
	2002/2014	2014/2019	2002/2019
Site 1			
Woody to Grazing lawn	1.19	0.38	0.75
Bunch grass to Grazing lawn	3.51	2.01	5.51
Grazing lawn to Woody	0.27	0.62	0.42
Grazing lawn to Bunch grass	0.84	2.59	2.55
Grazing lawn Unchanged	1.59	3.57	2.86
Grazing lawn to Bare	0.17	0.09	0.04
Bare to Grazing lawn	0.07	0.26	0.08
Site 2			
Woody to Grazing lawn	1.02	0.88	0.59
Bunch grass to Grazing lawn	1.89	1.25	1.41
Grazing lawn to Woody	0.36	1.53	0.45
Grazing lawn to Bunch grass	0.37	1.77	0.47
Grazing lawn Unchanged	0.66	1.19	0.39
Grazing lawn to Bare	0.05	0.49	0.28
Bare to Grazing lawn	0.02	0.02	0.01
Site 3			
Woody to Grazing lawn	0.87	1.81	2.32
Bunch grass to Grazing lawn	2.93	8.39	9.72
Grazing lawn to Woody	0.47	1.52	0.89
Grazing lawn to Bunch grass	1.33	1.02	0.72
Grazing lawn Unchanged	1.02	2.34	1.19
Grazing lawn to Bare	0.13	0.27	0.16
Bare to Grazing lawn	0.17	0.17	0.12

Transitions from grazing lawn to other land cover followed a similar trajectory as from other land cover to grazing lawns, where majority of the transitions were with bunch grass followed by woody cover (Table 3.6). This was the case particularly for Site 1 and Site 2 in all periods (i.e. 2002/2014 and 2014/2019). However, a different transition pattern from grazing lawn to other land cover types emerged for Site 3 between 2014 and 2019. where a greater proportion of grazing lawns were colonized by woody

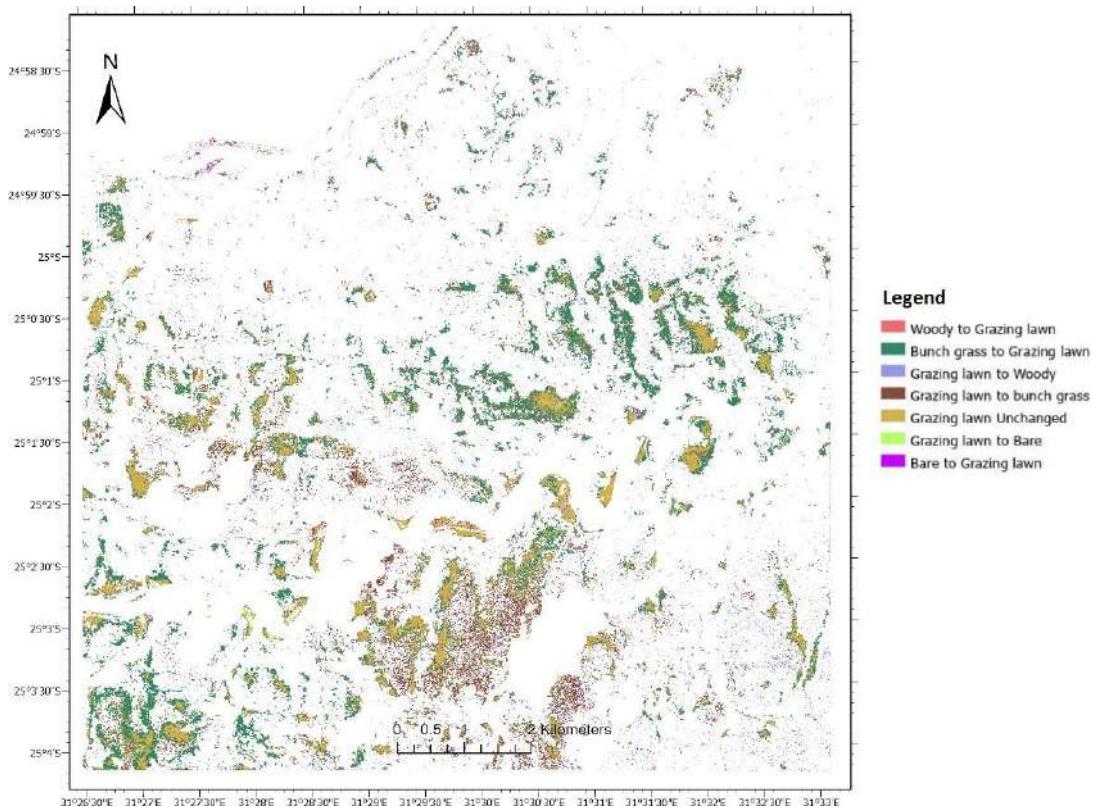
cover (1.52% [1.53 km<sup>2</sup>]) compared to bunch grass (1.02% [1.03 km<sup>2</sup>]). The spatial distributions of grazing lawn transitions are presented in Figures 3.12, 3.13 and 3.14 for Site 1; Figures 3.15, 3.16 and 3.17 for Site 2; and Figures 3.18, 3.19 and 3.20 for Site 3.



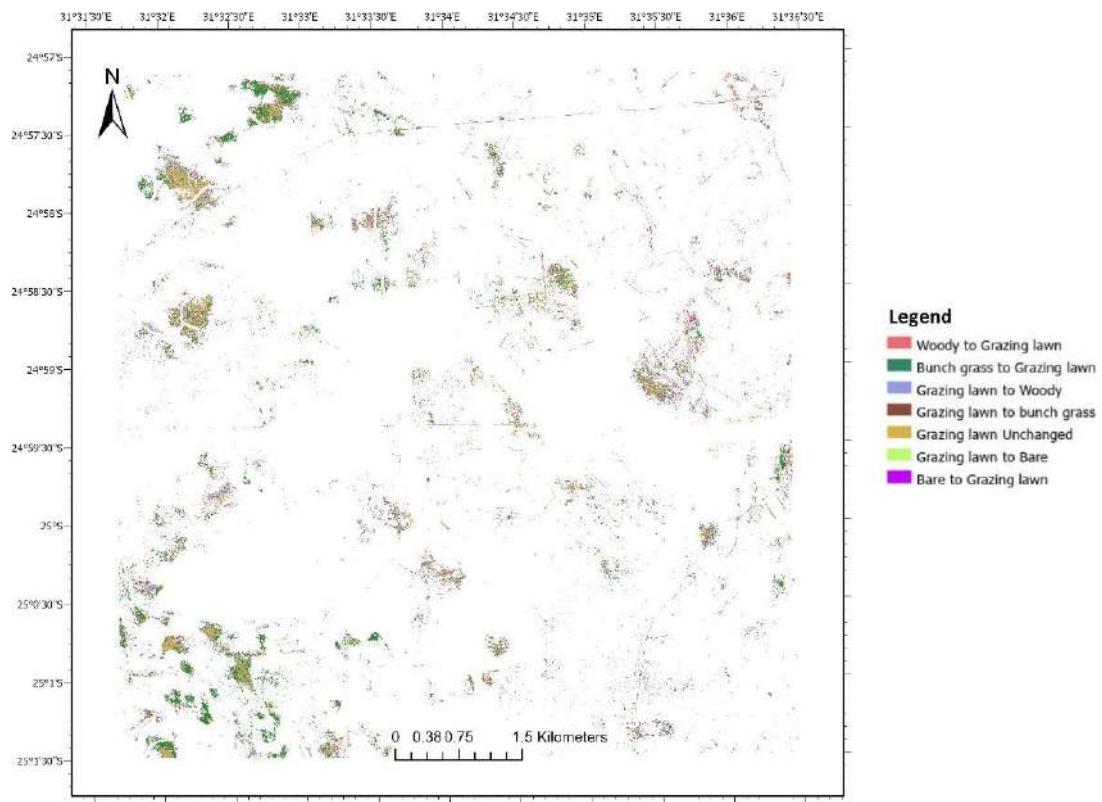
**Figure 3.12:** Grazing lawn change map for study site 1 for the period between 2002 and 2014.



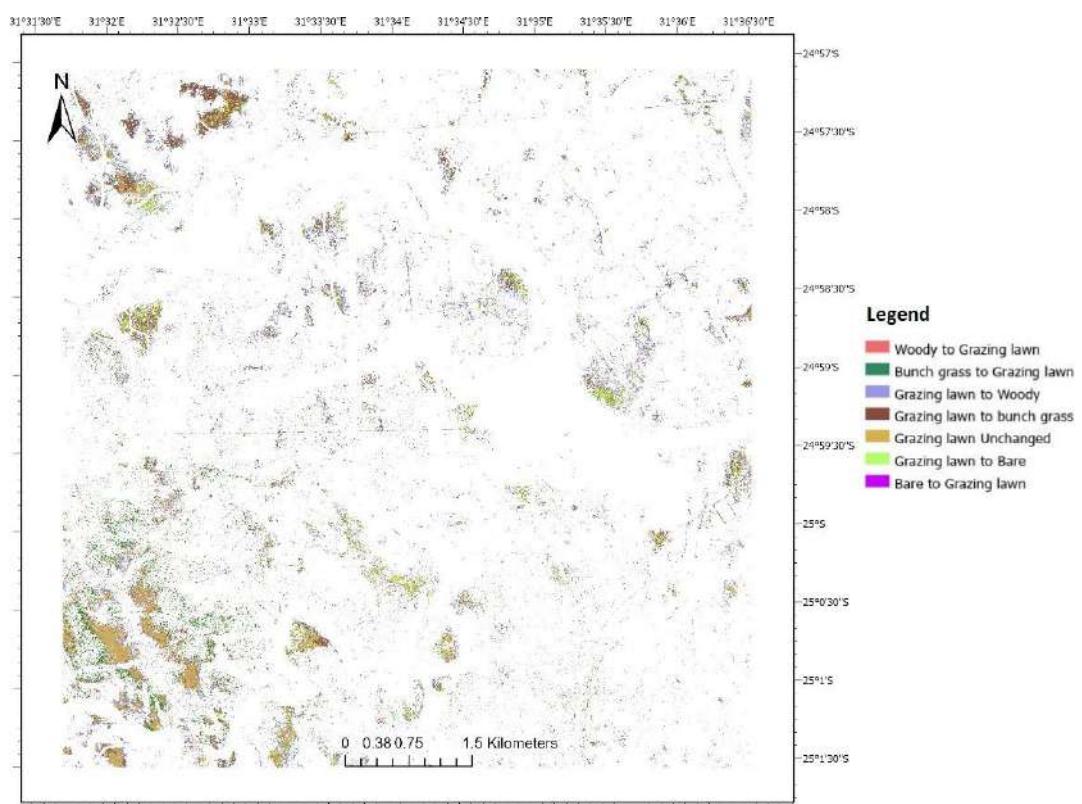
**Figure 3.13:** Grazing lawn change map for study site 1 for the period between 2014 and 2019.



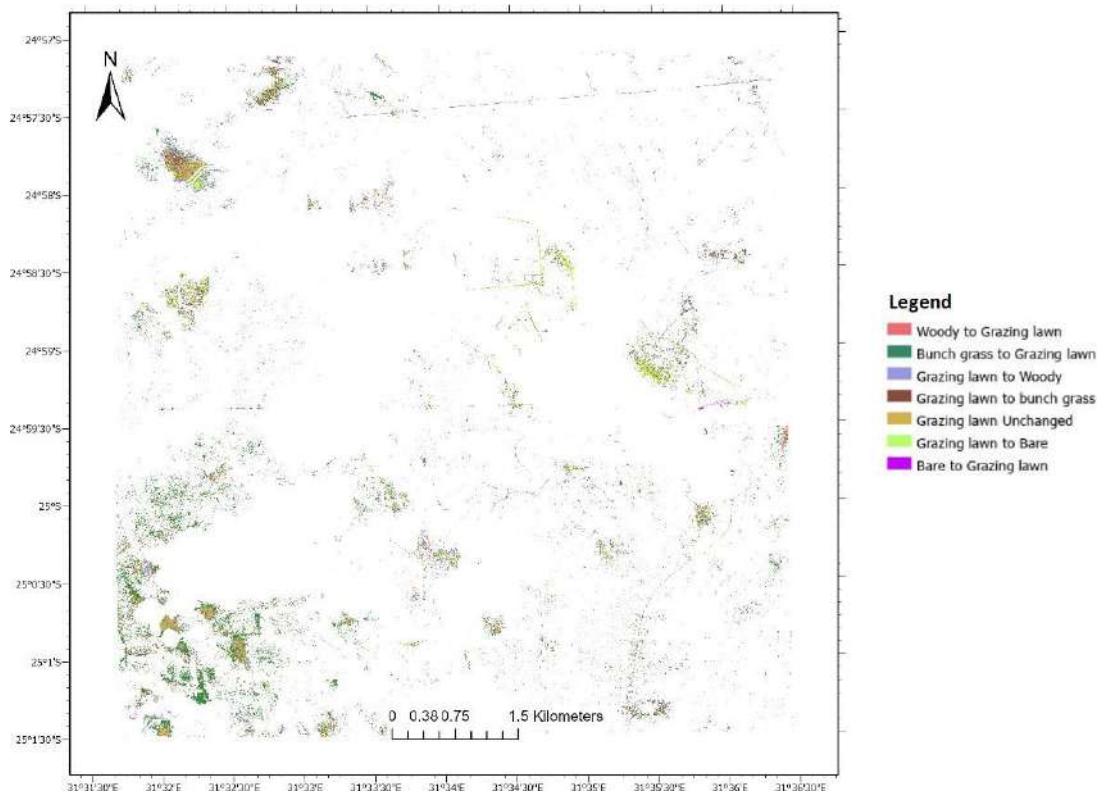
**Figure 3.14:** Grazing lawn change map for study site 1 for the period between 2002 and 2019.



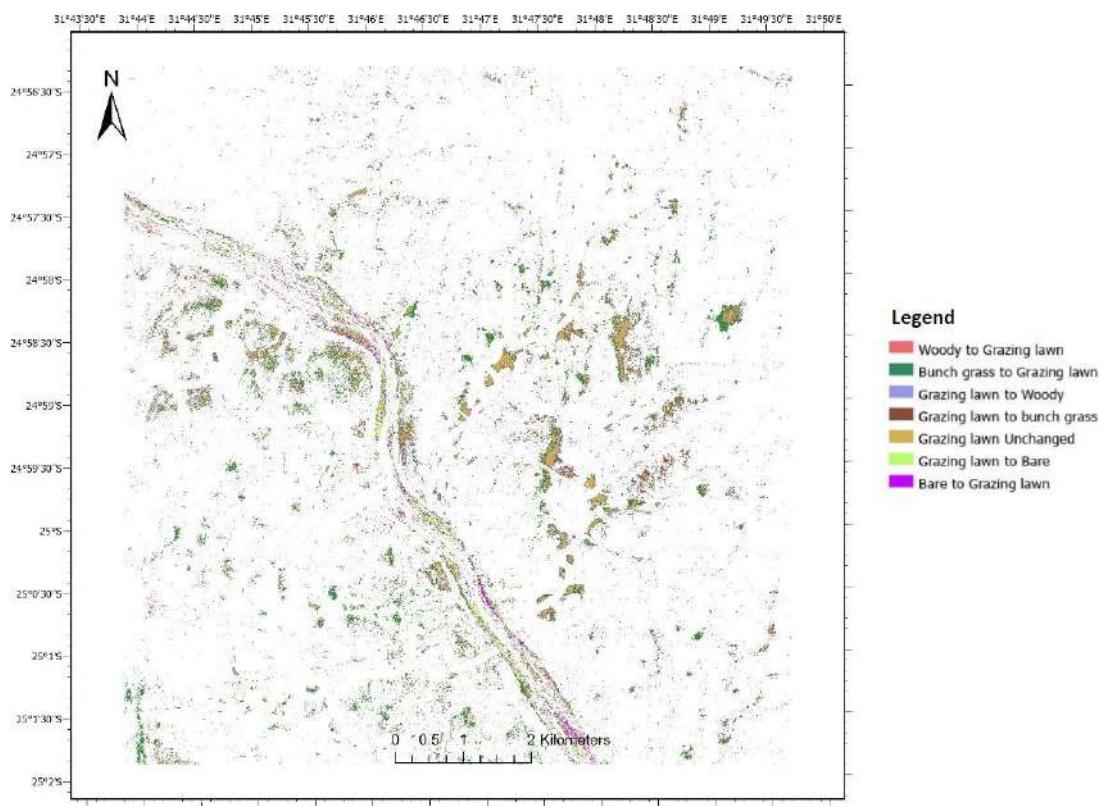
**Figure 3.15:** Grazing lawn change map for study site 2 for the period between 2002 and 2014.



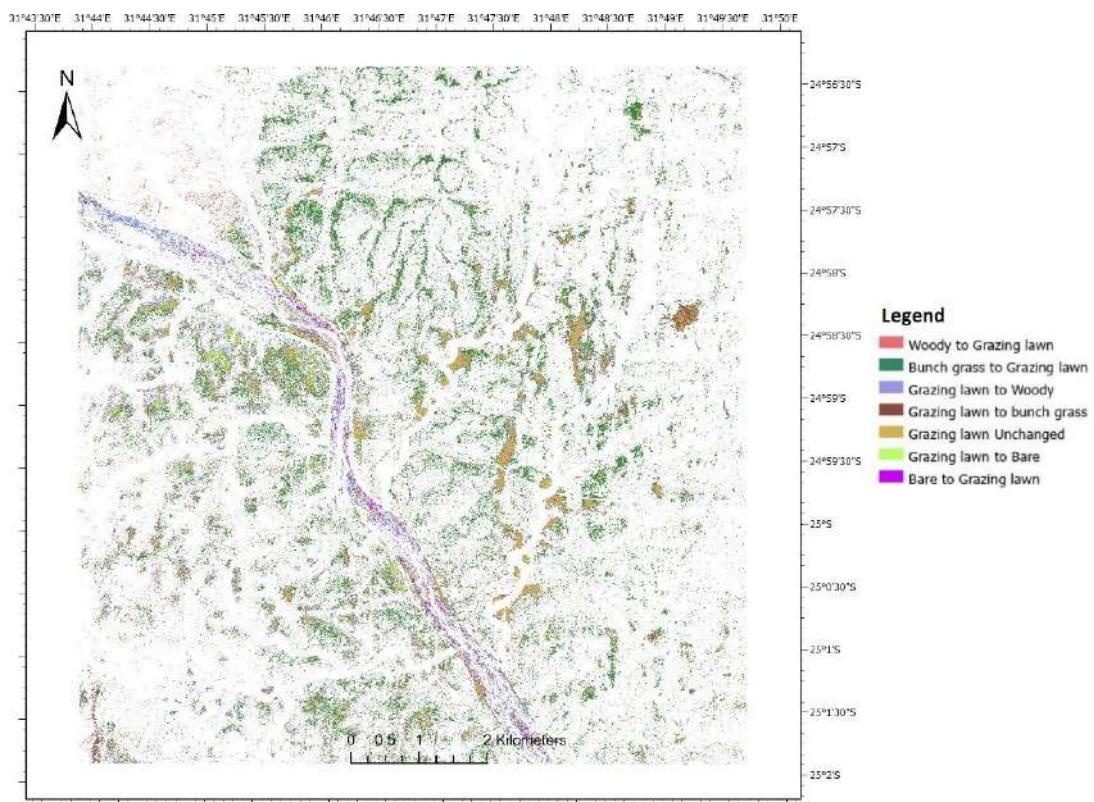
**Figure 3.16:** Grazing lawn change map for study site 2 for the period between 2014 and 2019.



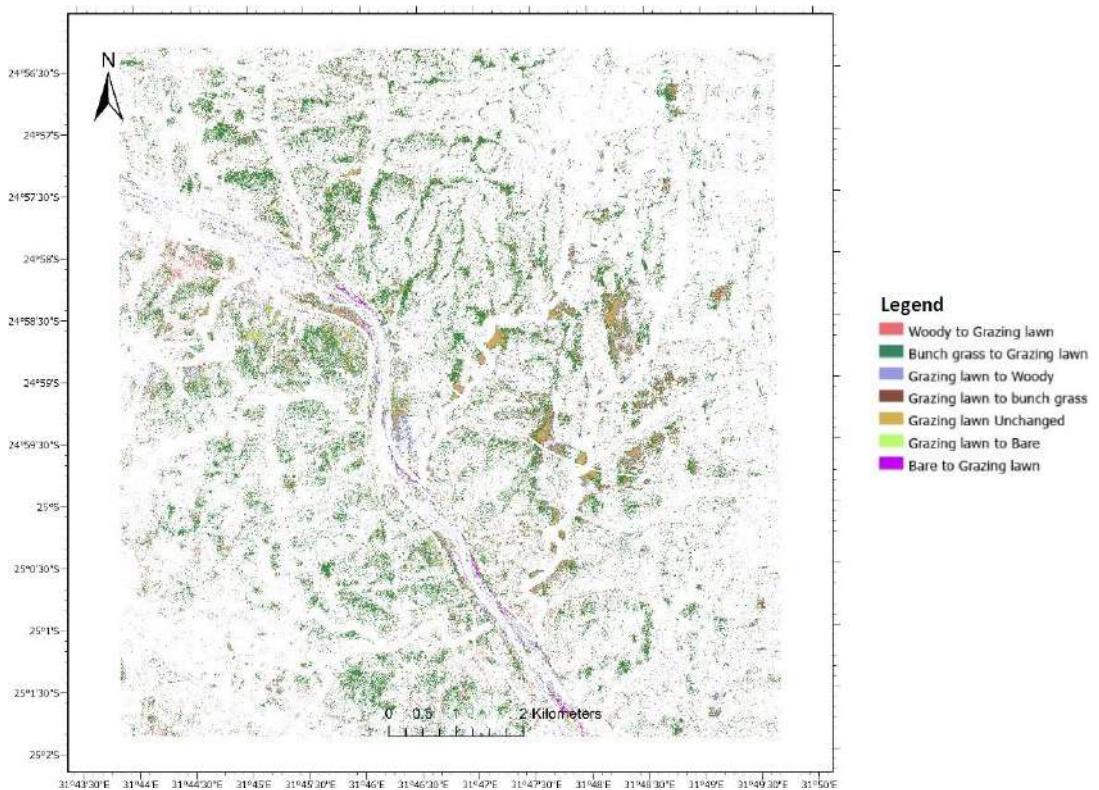
**Figure 3.17:** Grazing lawn change map for study site 2 for the period between 2002 and 2019.



**Figure 3.18:** Grazing lawn change map for study site 3 for the period between 2002 and 2014.

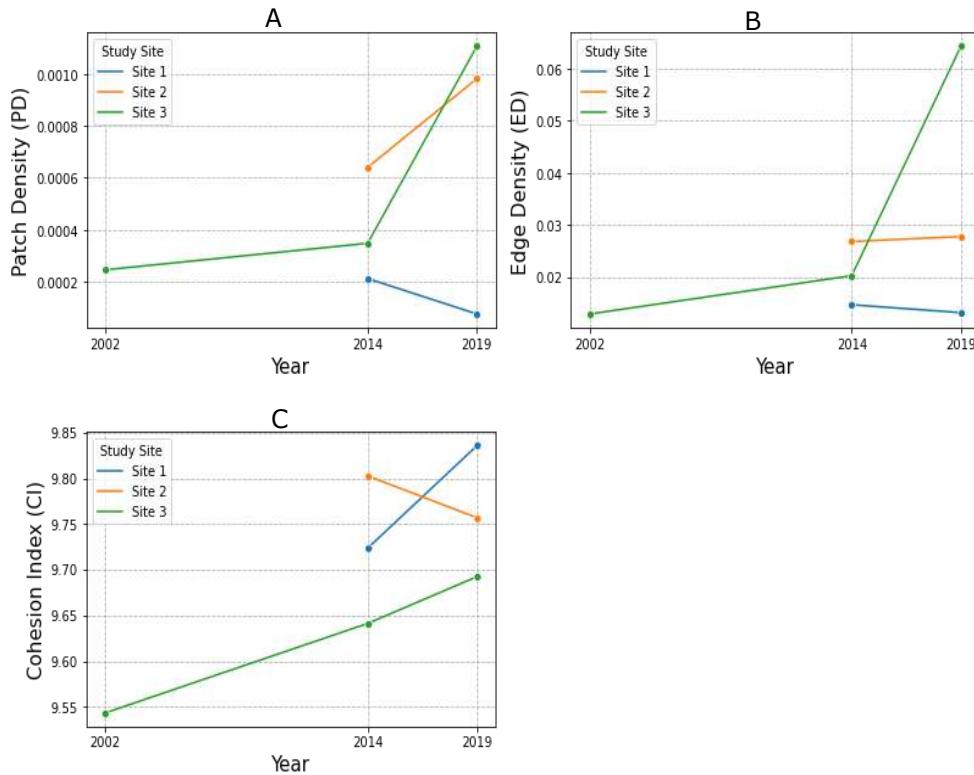


**Figure 3.19:** Grazing lawn change map for study site 3 for the period between 2014 and 2019.



**Figure 3.20:** Grazing lawn change map for study site 3 for the period between 2002 and 2019.

Figure 3.21 shows trends in grazing lawn structure, expressed in terms of the spatial configuration of grazing lawn patches for each study site. Results from 2002 for Site 1 and Site 2 were eliminated due to cloud cover, in order to avoid incorrect comparison and conclusions. For Site 1, both PD and ED decreased between 2014 and 2019, pointing to a decrease in number of grazing lawn patches per unit area, whereas CI showed an increasing trend showing an increasing degree of connectedness of grazing lawn patches within the landscape. Site 2 showed more grazing lawn patches but a decrease in connectivity from 2014 to 2019. Site 3 had results from all years and showed a positive trend for all metrics (Figure 3.21), depicting a general increase in grazing lawn patches and in closer proximity to each other from 2002 to 2014 and from 2014 to 2019.



**Figure 3.21:** Plots of spatial metrics showing dynamics in grazing lawn structure over time: (A) Patch Density (PD); (B) Edge Density (ED); and (C) Cohesion Index (CI).

## 3.4 Discussion

Analysis of changes in grazing lawn cover showed a consistent increase in the area coverage of grazing lawns and a corresponding wider spread of grazing lawns between 2002 and 2014 for all sites (see Table 3.5). This is corroborated by the fact that the combined areas of grazing lawns that remained unchanged and areas that switched from other land cover types to grazing lawns far exceeded transitions from grazing lawns to other land cover categories for the same period (Table 3.6). This pattern of grazing lawn persistence and spread could be explained by their characteristic attractiveness to grazers whose activity plays a significant role in the formation and maintenance of grazing lawns in the savannah landscape (Cromsigt & Beest, 2014; Cromsigt & Olff, 2008; Hempson *et al.*, 2015). A typical pathway for grazing lawn formation has been linked to the interactive effect of fire and herbivory, where fresh grass sprouting from newly burned areas attracts grazers, resulting in continuous grazing on post-burn sites which leads to the establishment of grazing lawns (Donaldson *et al.*, 2018). Although this increases the risk of already established grazing lawns patches

being colonized by tall grass swards due to abandonment by grazers, such fire-herbivore interaction drives the spread of grazing lawns (Archibald, 2008; Donaldson *et al.*, 2018). This may explain the observed wider distribution of grazing lawns, given high frequency of fire events in the savannah landscape (Archibald *et al.*, 2005). Apart from the provision of nutritional benefits for grazers, the low herbaceous cover on grazing lawn patches provides protection to herbivores by significantly reducing the risk of predation due to increased visibility (Hempson *et al.*, 2015; Voysey *et al.*, 2021). It therefore possible that the cumulative activities of herbivores resulted in the widespread formation and persistence of grazing lawns in the 2002/2014 study period.

In contrast with the findings from the 2002/2014 period, a differential pattern in grazing lawn dynamics was observed during the drought-impacted 2014/2019 period (Table 3.5). For example, grazing lawn cover increased marginally for Site 1, decreased for Site 2 and increased substantially for Site 3. This inconsistent trend in grazing lawn dynamics could be explained by drought-induced interruptions in ecosystem processes (Sankaran, 2019) with spatially varying implications. For example, Donaldson *et al.* (2020) reported that droughts decouple grazers from already established grazing lawn patches, which leads to colonization of such areas by tussock grasses during post-drought vegetation recovery. Additionally, droughts can intensify the impacts of grazing and grazer movement which can lead to substantial declines or degradation of grasses both in drought areas and non-drought refugia (Staver *et al.*, 2019). The intensity of drought impacts varies spatially within savannah landscapes, particularly due to spatial variation in factors such as topography and available soil moisture (Sankaran, 2019). Lower lying areas with relatively more soil moisture and surface water availability have been associated with high grazer activity (Redfern *et al.*, 2003; Smit, 2011; Smit *et al.*, 2007). Such landscape-scale distribution of grazers is even more pronounced during dry periods when surface water is spatially limited (Redfern *et al.*, 2003) and moisture content of graze is low (Berry & Louw, 1982). Compared with Site 1 and Site 2, Site 3 is strongly influenced by the Sabie river catchment, which together with the drainage channels, makes up approximately 80% of the landscape area (Figure 3.1). It is therefore conceivable that the impacts of drought stress on grasses would be relatively less intense in areas with relatively more water/moisture availability than the surrounding landscape. This coupled with the strong association of grazers to available moisture (Smit, 2011) could explain the recorded increase in grazing lawn presence during post-drought recovery (see site 3 maps and Table 3.5).

The majority of the transitions between grazing lawns and other land cover types occurred with the bunch grass class. This finding is not surprising given that such change trajectory is known to be the most common in savannah vegetation dynamics (Archibald *et al.*, 2005; Archibald, 2008; Cromsigt & Olff, 2008; Donaldson *et al.*, 2018; Hempson *et al.*, 2015). This suggests that the principal drivers of change in grazing lawn cover are grazers and their interaction with fire. This accords with earlier observations by Donaldson *et al.* (2018) and Hempson *et al.* (2019), which showed grazers, fire and their interaction as the key drivers of alternate grassland states (i.e. fire adapted tall grasses verses grazer adapted short grasses). Between 2014 and 2019, results for Site 3 however deviated from the grazing lawn to bunch grass trajectory, with a greater proportion of grazing lawns switching to woody cover. This is consistent with observed drought impacts on fire and herbivory and how that impacts post-drought woody vegetation recovery (Sankaran, 2019). Balfour & Howison (2002) reported that increasing drought frequency and protracted drought events are expected to decrease frequency and size of individual burns due to reduced grass fuel biomass. Additionally, droughts reduce browsing pressure due to herbivore mortality (Walker *et al.*, 1987). With fire and browsers being the main consumers of woody vegetation biomass, drought-induced reduction of these consuming forces will favour recruitment of woody vegetation, which could potentially explain the observed pattern of grazing lawn transition.

Analysis of grazing lawn structure revealed their spatial configuration for all sites and periods (Figure 3.21). Results are presented for only 2014 and 2019 for site 1 and Site 2 due to cloud cover in the respective 2002 imagery. Site 3 however yielded results from all years (i.e. 2002, 2014 and 2019). The results showed differing patterns of grazing lawn expansion and spatial heterogeneity for all sites. Even though Site 1 experienced an increase in grazing lawn cover between 2014 and 2019 (Table 3.5) the number of grazing lawn patches per unit area decreased during the same period, while the degree of connectedness increased. This shows that most of the observed grazing lawn expansion in Site 1 occurred as an extension of previous grazing lawn patches. This phenomenon is consistent with the maps of grazing lawn transition where most of the gains in grazing lawns occurred adjacent to stable patches (see Site 1 transition maps). In contrast, Site 2 showed a more heterogeneous pattern of grazing lawn spatial configuration between 2014 and 2019, with the observed losses in previously established cover (see Table 3.5) increasing the degree of spatial isolation. Such high heterogeneity in grazing lawn cover could favour a more diverse habitat structure (Hempson *et al.*, 2019) within the landscape. Given the low fuel

biomass on grazing lawns, widespread occurrence of small pockets of grazing lawn patches could effectively alter fire regimes within the landscape, including fire spread, intensity and frequency (Archibald *et al.*, 2005; Donaldson *et al.*, 2018; Hempson *et al.*, 2015). A unique pattern of grazing lawn spatial configuration was observed for Site 3, which showed a progressive increase in grazing lawn patches per unit area with increasing proximity of patches between 2002 and 2014 and from 2014 to 2019 (see Site 3 maps and Figure 3.21). From the perspective of grazing lawn ecology, physical characteristics of Site 3 makes it ideal for increased presence of grazing lawn cover. For example, Site 3 is dotted by a number of sodic sites and is largely covered with drainage channels with their associated catenas. Such landscape characteristics naturally attract and concentrate grazers in the savannah landscape resulting in widespread formation and persistence of grazing lawns (Archibald *et al.*, 2005; Archibald, 2008; Donaldson *et al.*, 2018; Hempson *et al.*, 2015).

### 3.5 Conclusions

This study sought to analyse spatial and temporal dynamics of grazing lawns in southern African savannahs using multi-temporal very high-resolution satellite images. Two temporal windows were used. The first one from 2002 to 2014, represented a period with no drought stress (Figure 3.2), and highlighted what can be considered as the regular dynamics in grazing lawn cover, driven by factors that continuously shape the ecology of grazing lawns. The second period from 2014 to 2019, experienced a major drought event and provided the opportunity to investigate how drought stress impacted dynamics in grazing lawn cover. The results showed an increase in grazing lawn cover within all the sites used for the analysis in the absence of drought stress (i.e. between 2002 and 2014). During this period (2002 to 2014) change trajectory of grazing lawn cover was dominated by transitions to bunch grass and vice versa, indicating that the activities of grazers and fire were the principal drivers of change. On the other hand, differential patterns of grazing lawn cover and structural changes were observed between 2014 and 2019, a period that recorded a major drought event. The impacts of drought on grazers and fire regimes therefore confounded post-drought vegetation recovery with mixed consequences for grazing lawn dynamics. The find-

ings from this study support previous understanding of the origin and persistence of grazing lawns as mainly shaped by the individual and interactive effects of fire and grazers. Additionally, the findings suggest that drought impacts on grazing lawns and overall vegetation recovery is strongly mediated by local landscape characteristics.

---

## Chapter 4

# **Multi-sensor Optical Image Fusion for Land Cover Classification in a Heterogeneous African Savannah: Toward Accurate and Cost Effective Grazing Lawn Monitoring**

---

### **4.1 Introduction**

Short grass grazing lawn patches are significant components of habitat heterogeneity in savannah ecosystems (Cromsigt & Olff, 2008; Hempson *et al.*, 2015,1). Their formation is controlled by the dynamic interactions between the activities of grazers and fire but are maintained by constant grazing (Archibald *et al.*, 2005; Donaldson *et al.*, 2018). Frequent grazing creates a feedback of dense nutrient-rich growth which in turn attracts more grazers. Further, their presence may be relatively stable in different parts of the savannah landscape where localised availability of resources such as nutrient hot-spots (e.g. termite mounds and sodic sites (Hempson *et al.*, 2015)) and water bodies (Awuah *et al.*, 2020; Hempson *et al.*, 2015) concentrate grazers. Compared to bunch grasses, grazing lawns are highly nutritious and the preferred food resource for certain short-grass grazer species due to their low stem to leaf and carbon to nitrogen (C:N) ratios. In southern Africa, the availability of such nutritious forage has been linked to grazer movement (Voysey *et al.*, 2021) and herbivore reproductive phenology with consequences for ungulate population dynamics (Owen-Smith & Ogutu, 2013). Additionally, grazing lawn patches act as natural breaks that curtail the spread of fire due to their limited above ground fuel biomass (Gill *et al.*, 2009; Waldram *et al.*, 2008). By influencing patterns in herbivory and fire regimes, the presence and ex-

tent of grazing lawn cover could potentially have significant cascading impacts on ecosystem processes such as plant community composition, nutrient cycling, habitat structure and biodiversity (Hempson *et al.*, 2019). Accurate knowledge of grazing lawn extent and distribution is thus essential to enhance understanding of the ecology of savannah grassland systems.

Satellite remote sensing technology offers unparalleled opportunities for cost-effective spatio-temporal quantification and analysis of land cover distributions. The availability of open-access satellite imagery - such as those from the Landsat (Wulder *et al.*, 2012) and Sentinel-2 (Drusch *et al.*, 2012) missions - and accompanying image processing frameworks have expedited land cover classification workflows with enhanced accuracy. Some regions and biomes such as forests have been studied extensively, partly due to standardized classification workflows and methodology. Conversely, savannahs in particular have received relatively little attention and remain a major source of uncertainty in land cover products globally (Bastin *et al.*, 2017; Griffith *et al.*, 2017). Recent research has focused on optimizing savannah land cover characterization mainly through the use of multi-source satellite images (Borges *et al.*, 2020; Liu *et al.*, 2017; Symeonakis *et al.*, 2018), improved image resolution and benchmarking of classification methods (Awuah *et al.*, 2020; Kaszta *et al.*, 2016). However, the inherent high spatial heterogeneity in savannah landscapes and spectral similarity between different vegetation components still present considerable challenges for accurate discrimination of vegetation components.

To deal with the spectral confusion among different vegetation components, integration of multi-season image acquisitions (Borges *et al.*, 2020; Symeonakis *et al.*, 2018) and the use of images from the dry season (Bucini *et al.*, 2009,1) have been proven to capitalize on phenological differences and increase discrimination space among spectrally similar vegetation components. The use of radar imagery or LiDAR has also been useful to discriminate between vegetation components by including structural information, e.g. grass vs woody vegetation (Cho *et al.*, 2012; Symeonakis *et al.*, 2018). Yet, classification accuracy is influenced significantly by image spatial resolution (e.g. see Marston *et al.* (2017)), where medium resolution imagery is impaired by pixel mixing under spatially heterogeneous savannah landscape conditions, restricting their success to rather general and continuously distributed target land cover classes. Current research favours very high spatial resolution (VHR,  $\leq 5\text{m}$ ) satellite imagery for dealing successfully with these challenges (e.g. Awuah *et al.* (2020); Kaszta *et al.* (2016); Marston *et al.* (2017)). For example, Marston *et al.* (2017)

showed that VHR IKONOS, Quickbird and WorldView-2 imagery can discriminate more savannah vegetation classes and at higher accuracies (OA: > 87.7%) than medium resolution imagery (OA: 75.2%). In a related study, Awuah *et al.* (2020) demonstrated the potential of WorldView-3 imagery for mapping plant functional types including grazing lawns in mesic and semiarid southern African savannahs with very high accuracies (> 90%).

Generally, pixel-based classification tends to perform better when image spatial resolution is relatively coarse (Gao & Mas, 2008). However, the prevalence of high within-class spectral variance in VHR imagery increases misclassification and can create a 'salt-and-pepper' effect when using pixel-based classification (Kaszta *et al.*, 2016; Van der Sande *et al.*, 2003). Alternatively, object-based image analysis (OBIA) aggregates pixels through the process of segmentation, creating contiguous groups of pixels with high spectral homogeneity known as image objects (Blaschke *et al.*, 2014). OBIA is widely recommended for the classification of VHR imagery (Blaschke *et al.*, 2014; Maxwell & Warner, 2015), particularly in heterogeneous landscapes (Ali *et al.*, 2016; Kaszta *et al.*, 2016; Xu *et al.*, 2019). In a specific African savannah classification using WorldView-2 imagery, Kaszta *et al.* (2016) showed that OBIA had a significantly higher accuracy across all seasons and different machine learning frameworks relative to pixel-based classification.

While VHR satellite images are ideal for heterogeneous savannah landscapes, such high resolution data are costly for operational monitoring of savannah vegetation and other land cover dynamics. A recent partnership involving Norway's International Climate and Forests Initiative (NICFI), Kongsberger Satellite Services (KSAT), Planet Labs and Airbus now grant free access to high-resolution, analysis-ready mosaics of Planet imagery over tropical regions (Planet, 2017). This promises new opportunities for rapid and cost-effective vegetation monitoring at high spatial scales. High spatial resolution Planet imagery have been proven successful for vegetation mapping, particularly in forest (Francini *et al.*, 2020) and agricultural (Laso *et al.*, 2020) landscapes. However, their spectral characteristics are limited and have shown less success in resolving the spectral similarity of different savannah vegetation components (Symeonakis *et al.*, 2019). Alternatively, open-access medium resolution satellite images such as Landsat Operational Land Imager (OLI) and Sentinel-2 sample relatively high numbers of spectral bands, potentially useful for mapping vegetation in African savannah landscapes (Symeonakis *et al.*, 2019).

In this study, we leverage both the high number of spectral bands of Sentinel-2 imagery and high spatial resolution of Planet imagery, integrating these two data sources for discrimination of different vegetation components in a heterogeneous southern African savannah landscape, towards accurate grazing lawn classification.

This study explores pixel-level image fusion approaches for integrating Sentinel-2 and Planet imagery. In comparison with other popular image fusion techniques such as the Intensity-Hue-Saturation transform (Siddiqui, 2003) and Principal Component transform (Yocky, 1995), Gram-Schmidt transformation (Laben & Brower, 2000) has been found to be fast and generates fused images with high spatial detail while preserving spectral integrity (Jenerowicz & Woroszkiewicz, 2016; Laben & Brower, 2000). Rokni *et al.* (2015) compared Modified Intensity-Hue-Saturation, High Pass Filter, Gram Schmidt and Wavelet-PC (Principal Components) techniques for fusion of multi-temporal Landsat Enhanced Thematic Mapper Plus (ETM+) 2000 band 8 (15 m) and Landsat Thematic Mapper (TM) 2010 multispectral (30 m) images towards surface water detection. The study (Rokni *et al.*, 2015) found the Gram-Schmidt image fusion technique to produce better outcomes both in terms of resulting image quality and classification. Further, the potential of Gram-Schmidt transformation for multi-sensor image fusion has been successfully demonstrated, with examples including pixel-level fusion of Unoccupied Aerial Vehicle (UAV) derived VHR aerial imagery with Landsat-8 OLI (Zhao *et al.*, 2019) and Sentinel-2 imagery (Jenerowicz & Woroszkiewicz, 2016) in agricultural landscapes. The performance of the different pan-sharpening approaches has not been demonstrated in heterogeneous savannah landscape. Therefore this study compares the Gram-Schmidt and Principal Component transformations in the fusion of Sentinel-2 and Planet imagery.

We hypothesize that, fusion of Sentinel-2 images with Planet imagery leverages their spectral and spatial advantages and enhances discrimination of savannah vegetation types leading to higher grazing lawn detection accuracies. To test this hypothesis, we (i) compare the performance of Gram-Schmidt and Principal Component methods for image fusion to select the best fusion outcome; and (ii) compare the original Sentinel-2 and Planet imagery with their fused version (i.e. Sentinel-2 + Planet imagery) for discriminating grazing lawns from other heterogeneous savannah land cover types.

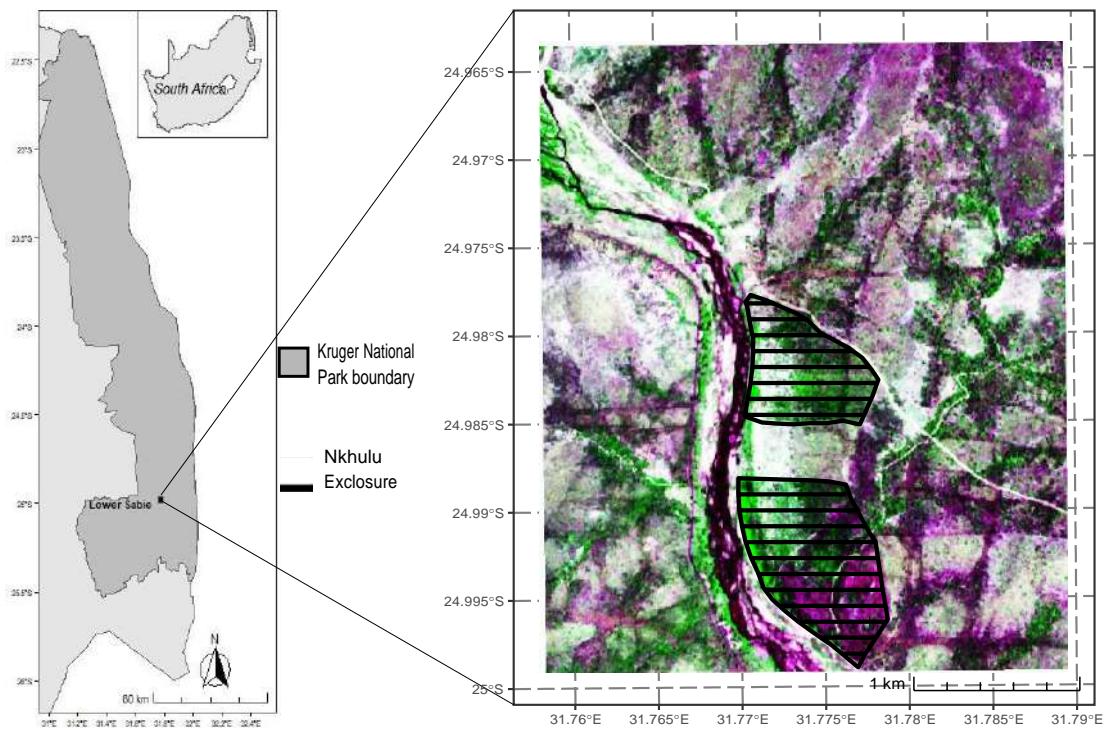
## **4.2 Materials and Methods**

### **4.2.1 Study Site**

The study site is located in Kruger National Park (KNP), South Africa (Figure 4.1). The test area covers 1220 ha extending over portions of the Sabie River catchment and encompassing the Nkhulu exclosure within the Lower Sabie region of KNP. Mesic savannah landscape conditions prevail within the area with mean rainfall ranging between 600 mm and 700 mm annually (Venter *et al.*, 2003). The site has underlying granite geology and is dotted with sodic sites often occurring at the footslopes of catenas, which are known to have high soil sodium content. Sodic sites provide dietary sodium supplement and thus, attract and concentrate grazers on these parts of the savannah landscape, making them significant drivers in grazing lawn formation (Hempson *et al.*, 2015). Vegetation within the study site is predominantly composed of thorn thickets (e.g., *Acacia robusta*), silver clusterleaf (*Terminalia sericea*) and sour grasses (*Hyparrhenia filipendula*) (Venter, 1991). Apart from the abiotic determinants of vegetation type distribution in KNP, the presence of a high herbivore population density such as the African elephant (*Loxodonta africana*) and dominant grass consumers (with >50% of grass in diet) such as impala (*Aepyceros melampus*), blue wildebeest (*Connochaetes taurinus*), zebra (*Equus quagga*), buffalo (*Syncerus caffer*) and white rhino (*Ceratotherium simum*) (Kleynhans *et al.*, 2011) exert significant impact on vegetation structure.

### **4.2.2 Land Cover Classification Nomenclature**

To discriminate grazing lawns from other vegetation types, three vegetation classes were identified based on structure. These included (i) woody; (ii) bunch grass; and (iii) grazing lawns. The target vegetation classes provided for a more objective in-situ and image-based discrimination of vegetation components during reference data collection. Trees and shrubs in savannah landscapes are structurally complex to discriminate due to disturbance adaptations (e.g. multiple stems, toppled, leaning etc.) and height limitations among others (Zizka *et al.*, 2014). These were thus grouped under a single woody category. Grazing lawns were identified as short grass patches with height  $\leq 20$  cm, often with sparse distribution and stoloniferous growth behavior,



**Figure 4.1:** Map of study area showing location of the Lower Sabie field site in Kruger National Park (KNP), with enlarged view of the Planet image scene (False colour: R, NIR, G) from NICFI overlaid with the Nkhulu exclosure boundary (sourced from South African National Parks Scientific Services (Siebert & Eckhardt, 2008)). Inset map shows the location of KNP within South Africa.

while grass patches > 20 cm in height were categorized as bunch grasses. In addition to the vegetation classes, water bodies, bare soil and built-up were considered, totalling six land cover classes used for image classification. A summary of land cover classification nomenclature is provided in Table 4.1.

**Table 4.1:** Description of land cover classification nomenclature and reference data. Reference samples are expressed as number of polygons and total area covered by polygons (in sq km) for each land cover class.

ID	Name	Land cover	Reference samples		
			Description	Number of polygons	Area coverage (km <sup>2</sup> )
1	Woody	Woody vegetation components including trees and shrubs at different phenological stages.		1908	0.071
2	Bunch grass	Tall grass patches often with upright growth form and dense distribution that are >20 cm in height.		679	0.101
3	Grazing lawn	Short grass patches with stoloniferous growth form and <20 cm in height, often sparsely distributed.		457	0.077
4	Water body	Water bodies occurring within the landscapes including rivers, streams and reservoirs. The landscape is mainly drained by the Sabie River.		56	0.03
5	Bare	Patches of exposed soil including dusty trails and rocky outcrops.		262	0.034
6	Built-up	Built artificial structures within the landscape, mostly asphalt and concrete coated surfaces such as roads and bridges and buildings.		36	0.007

### **4.2.3 Reference Data**

Field surveys were conducted in the dry season of 2019 between June 27 and July 21. A total of 111 point locations systematically distributed within a 200 m buffer and beyond 100 m distance from access roads were visited. Site descriptions of surrounding vegetation patches and other land cover were recorded together with cardinal direction photographs (north, east, south, west) to aid reference data generation. Data from in-situ reference locations were supplemented with 497 randomly distributed points within the bounds of the image footprint, with minimum spatial separation distance of 100 m to minimize autocorrelation. Overall, 3398 polygons were manually digitized from spectrally homogeneous areas within 100 m radius of sample point locations based on knowledge from field visits. Digitized polygon objects have been shown to produce the most accurate classification outcomes compared to other approaches for extracting training data such as points, lines and segmentation objects (Corcoran *et al.*, 2015; Ma *et al.*, 2017). Reference polygons were labeled according to the different land cover categories (Table 4.1) via visual interpretation of very high-resolution WorldView-3 imagery and Google Earth satellite scenes with the same temporal window (June/July 2019), augmented by field photos and descriptions of land cover.

### **4.2.4 Remote Sensing Data and Preprocessing**

The remote sensing data used in this study comprise cloud-free Sentinel-2A and Planet images taken during the dry season. Image acquisition dates coincide with the period of in-situ land cover surveys (see Section 4.2.3), to allow consistent classification and accuracy assessment. The Sentinel-2A and Planet images were co-registered to the WorldView-3 image used in reference polygon creation to allow spatial consistency during the extraction of reference pixels and comparison of classification results.

Sentinel-2 is a multi-spectral satellite imaging mission designed and operated by the European Space Agency (ESA) as part of the Global Monitoring for Environment and Security (GMES) initiative (Drusch *et al.*, 2012). It includes a constellation of two polar-orbiting satellites (Sentinel-2A and Sentinel-2B), phased at 180° to each other in a sun-synchronous orbit. The integration of two satellites provides a 5-day revisit frequency at the equator with a 290 km swath (Sadeh *et al.*, 2021). The mission features a Multi Spectral Instrument (MSI) which samples 13 spectral bands ranging

from visible and near infrared to shortwave infrared portions of the electromagnetic spectrum. The spatial resolution varies depending on the spectral bands: four bands at 10 m, six bands at 20 m and three bands at 60 m. Sentinel-2 images are freely available for download from a number of open-access portals, but primarily served on the Copernicus Open Access Hub website (<https://scihub.copernicus.eu/>). In this study, the Sentinel-2A image was acquired through the Google Earth Engine platform (Gorelick *et al.*, 2017). Images were acquired between May and June 2019 and had been pre-processed to surface reflectance using the Sen2Cor module (Louis *et al.*, 2016) as Level-2A products, from which a median composite was created. Overall, 10 spectral bands were used covering the visible and near infrared (NIR) channels at 10 m spatial resolution and the red edge (RE) and shortwave infrared (SWIR) bands at 20 m spatial resolution. The RE and SWIR bands were resampled to 10 m to ensure consistency with the visible and NIR bands.

The Planet image is part of the open-access analysis-ready (ARD) satellite image data over tropical regions. The archive features bi-annual mosaics between December 2015 and August 2020 and monthly mosaics available from September 2020 onward. Access to this VHR ARD satellite imagery was made available through a partnership between NICFI, KSAT, Planet Labs and Airbus (operational on October 22, 2020), intended to support efforts against tropical deforestation. The mosaics are constructed using best pixels from daily Planet imagery processed to surface reflectance either as three-band (Blue, Green and Red) or four-band (Blue, Green, Red and Near Infrared) image composites. In this study, the four-band ARD composite covering June 2019 to December 2019, with spatial resolution of 4.77 m per pixel was used.

#### **4.2.5 Multi-sensor Image Fusion**

Pixel-level image fusion of Sentinel-2A and Planet imagery was achieved via panchromatic sharpening using both Gram-Schmidt transformation (Laben & Brower, 2000) and Principal Component transformation (Yocky, 1995). In this study, the higher spatial resolution panchromatic band was derived from the Planet image as the mean of all four bands. To fuse the derived higher spatial resolution panchromatic band with the lower spatial resolution Sentinel-2A image, first, a lower spatial resolution panchromatic band is simulated from the Sentinel-2A image. Second, Gram-Schmidt and Principal Component transformations are separately performed on the simulated lower spatial resolution panchromatic band and the plurality of the Sentinel-2A bands,

with the lower spatial resolution panchromatic band employed as the first band in the transformation process. Third, the statistics (mean and standard deviation) of the higher spatial resolution panchromatic image are adjusted to match the statistics of the first output band from the Gram-Schmidt and Principal Component transformations using histogram matching. Fourth, the first transform band is replaced with the adjusted higher spatial resolution panchromatic band to produce a new set of Gram-Schmidt and Principal Component transform bands. Finally, an inverse transformation is performed to return the new set of transform bands into a higher spatial resolution multi-spectral band space.

The resulting Fused image, and the original Sentinel-2A and Planet images were tested generally for savannah land cover classification and specifically for grazing lawn detection, and their accuracy scores were compared.

#### **4.2.6 Image Quality Assessment**

The image fusion approach was evaluated to ascertain the improvement of spatial resolution and preservation of image spectral characteristics. First the fused image was visually compared to the original images (Sentinel-2A and Planet imagery) to assess improvement in spatial detail. Second, two statistical metrics were computed to evaluate the preservation of image spectra. This included: (i) correlation coefficient (CC) between the original Sentinel-2A image bands and the equivalent bands from the fused image (CC ranges between -1 and 1, where values closer to 1 or -1 show increasing positive or negative correspondence between band pairs); and (ii) structural similarity index (SSIM) (Wang *et al.*, 2004) comparing the multi-spectral Sentinel-2A and fused images. SSIM combines luminance, contrast and structure, and was calculated with a moving 9 x 9 pixel neighbourhood. SSIM computation using a moving window over both images helps to capture spatial variation and ultimately leads to better estimates of (dis)similarity between images (Avanaki, 2009). The 9 x 9 pixel neighborhood was deemed appropriate to cover enough spatial variation while minimizing computational cost. The SSIM values range between -1 and 1, and only equals 1 if the two images are identical. CC and SSIM were calculated using Equations 4.1 and 4.2

$$CC = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (4.1)$$

where  $\sigma_{xy}$  is covariance for band pairs ( $xy$ ) and  $\sigma_x$  and  $\sigma_y$  are standard deviations for band  $x$  and  $y$  respectively.

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{\mu_x^2 + \mu_y^2 + C_1 \quad \sigma_x^2 + \sigma_y^2 + C_2} \quad (4.2)$$

where  $\mu_x$  and  $\mu_y$  are mean intensities,  $\sigma_{xy}$  is covariance,  $\sigma_x$  and  $\sigma_y$  are standard deviations,  $C_1$  and  $C_2$  are constants for luminance and contrast comparisons respectively, used to stabilize the equation in the presence of weak denominators (Wang *et al.*, 2004). In this case  $x$  and  $y$  represent the original Sentinel-2 image and the fused image respectively. The output from the best performing fusion approach (i.e., between Gram-Schmidt and Principal Component pan-sharpening) was selected for further processing and classification.

#### 4.2.7 Image Segmentation

Image segmentation was performed on all image datasets (i.e. Sentinel-2, Planet and fused images). This involved the computation of a gradient image from the original input image bands using a Sobel edge detection method (Vincent *et al.*, 2009), followed by a watershed transform (Roerdink & Meijster, 2000). The watershed transform algorithm is modeled after the hydrologic watershed concept, with a landscape composed of basins. Basins fill up with water starting at the lowest points until flooded, and dams are built at the meeting point of water coming from different basins. The landscape is thus divided into regions (watersheds) separated by dams (watershed lines). Similarly, the watershed transform algorithm considers the gradient image as a topological surface where dark pixels represent lower elevation (minimum). The algorithm sorts the gradient image pixels by increasing intensity values and 'floods' the image starting with the lowest gradient values (uniform portions of object) to the highest gradient values (object edges). The results is a segmentation image partitioned into regions with similar pixel intensities where each region is assigned the mean intensity value. A Full Lambda algorithm, which merges smaller segments within larger, textured areas, was used to minimize over-segmentation (Jin, 2012).

#### 4.2.8 Feature Extraction

In addition to the original segmentation image bands, two spectral indices and three Haralick textural image features were calculated for all datasets in order to enhance land cover discrimination space. The spectral indices included the Global Environmental Monitoring Index (GEMI) (Pinty & Verstraete, 1992) (Equation 4.3) and Modified Soil Adjusted Vegetation Index-2 (MSAVI2) (Laosuwan & Uttaruk, 2014) (Equation 4.4) while Mean (Equation 4.5), Variance (Equation 4.6) and Sum Average (Equation 4.7) constituted the Haralick texture features (Haralick *et al.*, 1973). The spectral indices and texture features were selected following results from a related study (Awuah *et al.*, 2020), where they ranked amongst the most important features in savannah plant functional types (PFTs) classification across different machine learning models. The Haralick texture features were calculated on the NIR band based on Gray Level Co-occurrence Matrix (GLCM) (Haralick *et al.*, 1973) using a 3x3 moving window, at an offset distance of 1 pixel in all directions ( $0^0$ ,  $45^0$ ,  $90^0$  and  $135^0$ ) (Pratt, 2013; Symeonakis *et al.*, 2018).

$$GEMI = \frac{2(N^2 - R^2) + 1.5N + 0.5R}{N + R + 0.5} \quad (4.3)$$

$$MSAVI2 = \frac{(2N + 1) - \sqrt{(2N + 1)^2 - 8(N - R)}}{2} \quad (4.4)$$

where  $N$  and  $R$  are near infrared and red bands respectively.

$$Mean = \sum_{i,j=0}^{G-1} i P_x(i) \quad (4.5)$$

$$Variance = \sum_{i,j=0}^{G-1} (P_x(i) - \mu_x)^2 \quad (4.6)$$

$$Sumaverage = \sum_{i,j=0}^{2G-2} i P_{x+y}(i) \quad (4.7)$$

where  $G$  is the number of distinct gray levels,  $i$  and  $j$  are co-occurring intensity values,  $P_x(i)$  is the marginal probability of the  $i$ th entry obtained by summing rows of the GLCM,  $\mu_x(i)$  is mean of row sums of the GLCM and  $P_{x+y}(i)$  represents GLCM sum distribution.

#### 4.2.9 Image Classification and Accuracy Assessment

The segmentation image composed of original bands, spectral indices and texture features (for each dataset) was normalized by subtracting the mean and dividing by its standard deviation for each image band to minimize bias during algorithm training and image classification. The normalized segmentation image served as input in a deep neural network classification model using Multilayer Perceptron (MLP) network architecture. MLP is a simple feed-forward deep learning model trained via back-propagation (Goodfellow *et al.*, 2016). MLP neural networks excel at handling complex non-linear classification tasks (Bischof *et al.*, 1992; Venkatesh & Raja, 2003) making them ideal for application in a highly heterogeneous savannah landscape (Awuah *et al.*, 2020). Three hidden layers with 64, 32 and 8 neurons for the first, second and third hidden layers respectively were found to yield optimal accuracy after initial experimentation. Dropout layers (Srivastava *et al.*, 2014) were used between the dense layers, which randomly neglected 20% of weights in order to reduce overfitting and improve generalization error. The model was compiled using a constant learning rate of 0.001 with Adams optimizer (Kingma & Ba, 2014) and a categorical cross-entropy loss function.

Accuracies of classification outputs were assessed using a confusion matrix (Congalton & Green, 2019) from which a combination of aggregate and individual class based metrics were calculated. The accuracy metrics included precision (Equation 4.8), recall (Equation 4.9) and F-score (Equation 4.10).

$$Precision = \frac{tp}{tp + fp} \quad (4.8)$$

$$Recall = \frac{tp}{tp + fn} \quad (4.9)$$

$$F\_score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4.10)$$

where tp, fp and fn represent the number of true positive, false positive, true negative and false negative cases, respectively.

A stratified 10-fold cross-validation approach was adopted during accuracy assessment. Compared to the split-sample approach, cross-validation is more robust due to the availability of multiple metrics and minimizes the risk of chance results while providing for a more efficient use of limited data. Cross-validation accuracy results (i.e. using F-scores) from the individual datasets (i.e. Sentinel-2A, Planet and Fused) were compared using the Friedman test (Friedman, 1937,4) to determine the significance (at  $\alpha = 0.05$ ) of accuracy differences. The Friedman test is a non-parametric test for comparing three or more treatment scenarios (Demšar, 2006) by assigning ranks to treatments separately for each data point. In the case of this study, the Friedman test ranks F-scores from each iteration of cross-validation across all datasets (i.e. Sentinel-2A, Planet and Fused) and compares the average ranks under the null hypothesis that all datasets have equal average ranks. Assuming  $r_i^j$  is the rank for the  $j$ -th of  $k$  datasets on the  $i$ -th out of  $N$  cross-validation scores, the average rank is  $R_j = \frac{1}{N} \sum_i r_i^j$ . The Friedman statistic is calculated using Equation 4.11 (Demšar, 2006).

$$\chi_F^2 = \frac{12N}{k(k+1)} \sum_j R_j^2 - \frac{k(k+1)^2}{4} \quad (4.11)$$

The Friedman test follows a chi-square distribution with  $k-1$  degrees of freedom. We used the Nemenyi test (Nemenyi, 1963) for post-hoc analysis in cases where there were significant differences in accuracy scores among the three datasets used. Similar to the Tukey test for Analysis of Variance (ANOVA), Nemenyi test is used to compare all groups (i.e. datasets in the case of this study) to each other. Two groups are significantly different if the difference between their average ranks is greater than a critical difference (CD) calculated using Equation 4.12.

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}} \quad (4.12)$$

where  $q_\alpha$  represents a critical value from the studentized range statistic divided by  $\sqrt{\frac{2}{N}}$  (Demšar, 2006).

#### **4.2.10 Feature Importance Estimation**

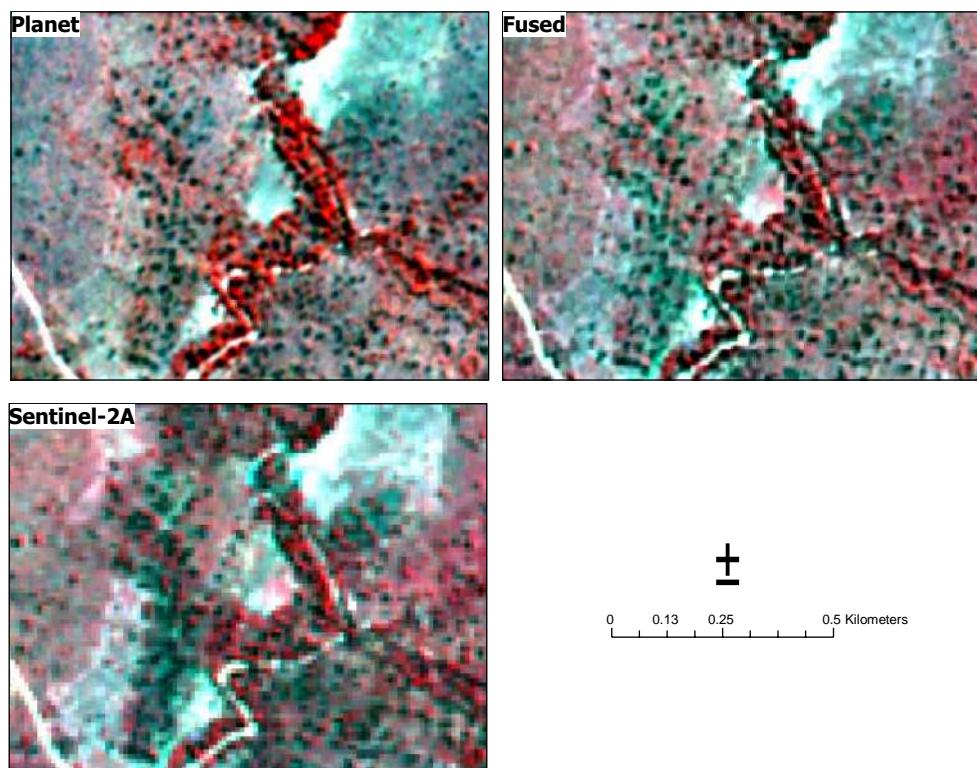
The contribution of individual image features to classification accuracy was assessed through permutation (Pedregosa *et al.*, 2011). This was achieved by randomly shuffling each image feature to determine importance estimates as weights based on degree of reduction in classification accuracy score.

### **4.3 Results**

#### **4.3.1 Image Fusion Accuracy**

Pixel-level image fusion resulted in the creation of a new dataset (Fused image), which retained both the spectral characteristics of Sentinel-2A and the high spatial resolution from Planet data. Figure 4.2 shows an example visual comparison of spatial details resolved by all three datasets (i.e. Sentinel-2A, Planet and Fused images). As expected, the Sentinel-2A image at 10 m spatial resolution has a higher level of pixelation at the same visualization scale as Planet and Fused images due to its relatively coarse resolution. Further, land cover features are less distinctly resolved, with high prevalence of pixel mixing in some instances. For example, the sparsely distributed and smaller-sized woody vegetation components become mixed-up with the background grassy vegetation and are indistinguishable (for crowns < 10 m) in the Sentinel-2A image compared to the Fused and Planet images (Figure 4.2). By contrast, the Fused image is finer and shows similar spatial detail as the high resolution Planet data. At 4.7 m spatial resolution, this enables easy identification of objects that are not recognizable in the Sentinel-2A image while retaining the multi-spectral information from Sentinel-2A. Additionally, the spatial boundaries between different land cover types are better distinguished at the spatial resolution of the Fused and Planet images compared to Sentinel-2A (Figure 4.2), which could have important consequences for accurate grazing lawn and other land cover delineation.

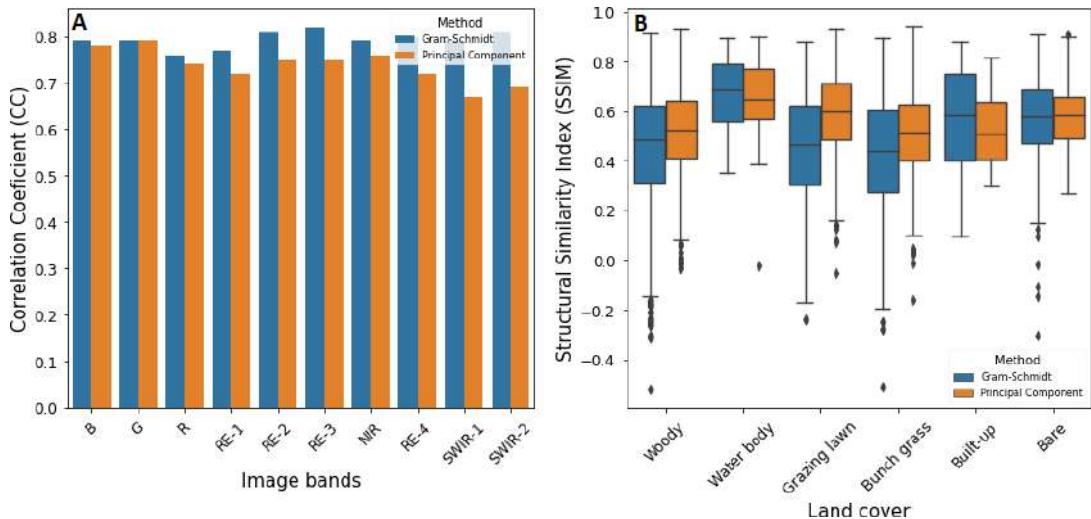
Apart from the enhanced image spatial resolution, image fusion resulted in the fused image having the same number of spectral bands as the Sentinel-2A image, with expected similar spectral characteristics. Figure 4.3 shows results from the assessment of the spectral integrity of the fused image relative to the Sentinel-2A image for both the Gram-Schmidt and Principal Component fusion methods. The correlation between the fused and Sentinel-2A image was generally high for all band pairs. However



**Figure 4.2:** False colour (R: NIR, G: Red, B: Green) display of original and fused images scenes of a sub-area from Gram-Schmidt sharpening. A visual comparison shows higher spatial resolution of Planet and Fused images relative to the Sentinel-2A image.

between the two fusion methods, the Gram-Schmidt approach resulted in higher correlation values than the Principal Component approach, with the differences being relatively more pronounced in the red-edge, near-infrared and short-wave infrared spectra (Figure 4.3B).

The influence of image fusion on different land cover spectral characteristics for both Gram-Schmidt and Principal Component methods, based on the calculated Structural Similarity Index (SSIM) between the multi-spectral fused and Sentinel-2A images is shown in Figure 4.3 B. Overall, a loss of structural information in the Fused images relative to Sentinel-2A was recorded for all savannah land cover types as indicated by SSIM statistics. As expected, the highest SSIM values were observed in the relatively stable and homogeneous land cover types with minimum expected variations in spectra across the savannah landscape, including: bare surfaces, built-up and water bodies. For those land cover types, the Gram-Schmidt fusion approach gave a slightly better outcome.



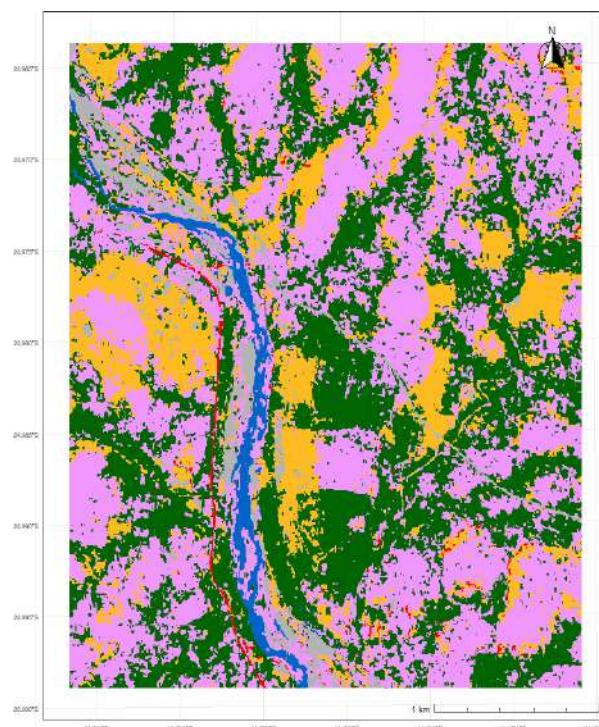
**Figure 4.3:** Spectral quality metrics from pixel-level fusion. The figure shows Correlation Coefficient from Sentinel-2 and fused image band pairs (left) and boxplot showing summary of Structural Similarity Index (SSIM) for different savannah land cover types (right). B = Blue, G = Green, R = Red, RE-1 = Red Edge-1, RE-2 = Red Edge-2, RE-3 = Red Edge-3, NIR = Near Infrared, SWIR-1 = Short Wave Infrared-1 and SWIR-2 = Short Wave Infrared-2.

#### 4.3.2 Land Cover Classification

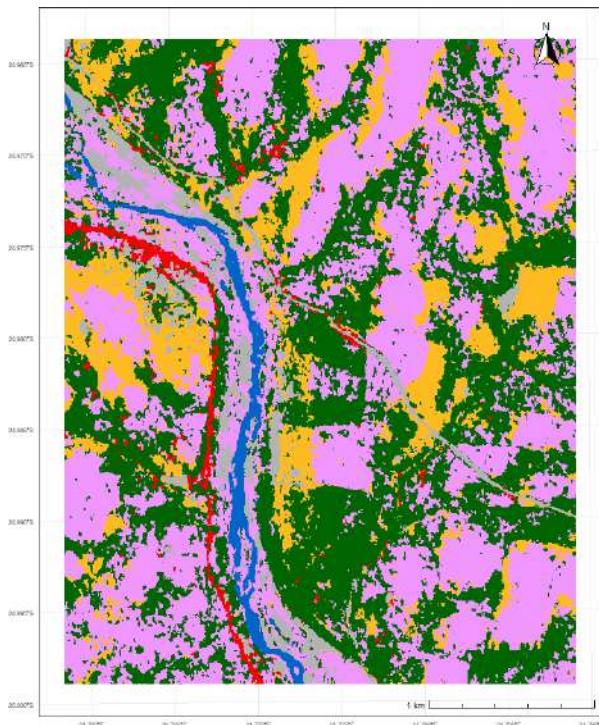
Land cover classifications derived from the three datasets (Planet, Fused images and Sentinel-2A) are presented in Figures 4.4, 4.5 and 4.6. Generally, the thematic representation of different land cover types closely followed visible distribution patterns in the satellite scene shown in Figure 4.1. Apart from the dominance of woody cover in the Nkhulu experimental enclosure 4.1, woody vegetation was mostly confined along drainage channels, with a few isolated occurrences in the midst of continuous bunch grass cover within the landscape. Grazing lawns were mostly present on open sodic sites and adjacent to woody cover along drainage channels or foot-slopes of catenas scattered within the landscape.

Similar thematic details are visible across all maps, however, there is a clear loss of spatial detail in the Sentinel-2A-derived map (Figure 4.6) compared to maps derived from Planet and Fused images (Figure 4.4 and Figure 4.5 respectively). Specifically, boundaries of the different land cover types are less distinctive in the Sentinel-2A map compared to those of Planet and Fused images. For example, built-up (road) appears discontinuous in the Sentinel-2A map while strips of woody cover along drainage channels, particularly in the more open northeastern part of the landscape,

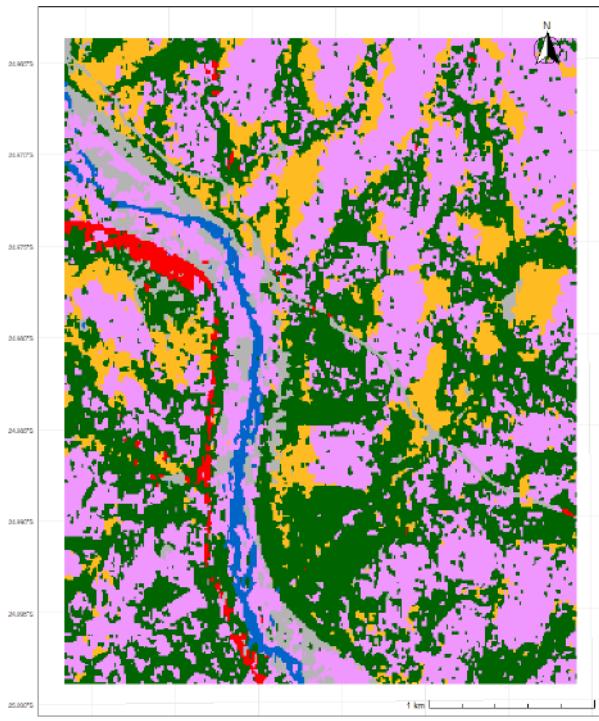
are more distinctly captured in the Fused and Planet maps compared to Sentinel-2A. Additionally, grazing lawn presence and distribution in different parts of the landscape appear congruent in the higher spatial resolution datasets (i.e. Planet and Fused images) and differ slightly from the coarser resolution Sentinel-2A map, particularly in the southeastern portions of the landscape (see Figure 4.4 and Figure 4.5).



**Figure 4.4:** Land cover map derived from Planet image.

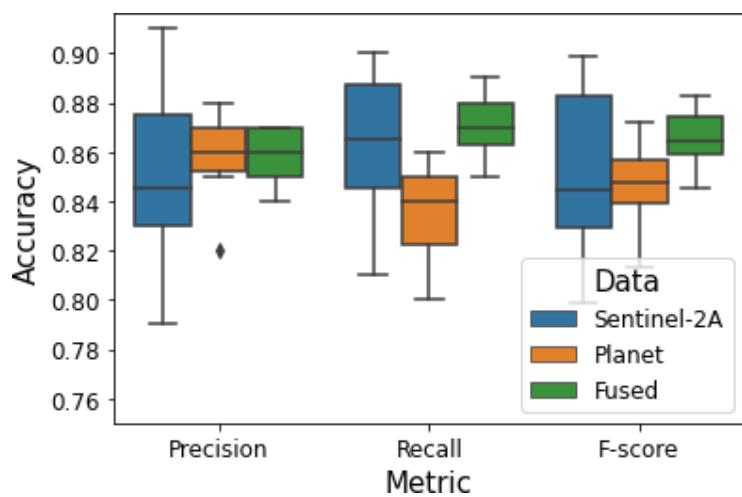


**Figure 4.5:** Land cover map derived from Fused image.



**Figure 4.6:** Land cover map derived from Sentinel-2A image.

Beyond qualitative differences in land cover maps, classification accuracies were quantified using a stratified 10-fold cross-validation approach and compared to further ascertain the utility of the different datasets. Accuracy results from confusion matrices (see Appendix C) are presented in Figure 4.7 and Table 4.2 for the whole landscape and individual land cover types respectively. Overall, the higher spatial resolution images (Fused and Planet images at 4.77m spatial resolution) gave relatively consistent and precise estimates of general land cover classification accuracies compared to the 10 m resolution Sentinel-2A image, which had a wider interquartile range of cross-validation accuracy scores (Figure 4.7). Aggregate accuracy scores for the Fused image were marginally higher, achieving an average F-score of  $0.87 \pm 0.012$ , compared to Planet and Sentinel-2A images with average F-score values of  $0.85 \pm 0.017$  and  $0.85 \pm 0.034$  respectively.



**Figure 4.7:** General classification accuracies from stratified 10-fold cross-validation. Accuracy metrics were calculated as weighted averages from individual land cover class accuracies. Accuracy values represent fractions between 0 and 1.

At the level of individual land cover types, classification accuracy scores obtained from the Fused image were higher than those from Sentinel-2A and Planet images for all land cover classes except woody cover (Table 4.2). Sentinel-2A produced more accurate classification of woody cover, in excess of 4% over the Fused image and 6% over the Planet image in terms of average F-score values. Grazing lawn classification accuracy followed a similar trend as the general map accuracy estimates,

**Table 4.2:** Summary of classification accuracy scores for savannah land cover types showing a comparison across dataset used (Sentinel-2A, Planet, Fused). Accuracy scores represent mean standard deviation from 10-fold stratified cross-validation. Accuracy values represent fractions between 0 and 1.

Data	Land cover	Precision	Recall	F-score
Sentinel-2A	Woody	0.82 ± 0.023	0.77 ± 0.055	0.79 ± 0.029
	Bunch grass	0.79 ± 0.049	0.86 ± 0.053	0.83 ± 0.048
	Grazing lawn	0.88 ± 0.038	0.82 ± 0.068	0.84 ± 0.048
	Water body	0.93 ± 0.057	0.96 ± 0.044	0.94 ± 0.039
	Bare	0.91 ± 0.041	0.89 ± 0.067	0.90 ± 0.042
	Built-up	0.76 ± 0.146	0.88 ± 0.143	0.80 ± 0.113
Planet	Woody	0.76 ± 0.027	0.71 ± 0.039	0.73 ± 0.026
	Bunch grass	0.76 ± 0.027	0.83 ± 0.020	0.79 ± 0.019
	Grazing lawn	0.86 ± 0.024	0.83 ± 0.020	0.84 ± 0.017
	Water body	0.96 ± 0.014	0.95 ± 0.023	0.95 ± 0.016
	Bare	0.91 ± 0.017	0.89 ± 0.013	0.89 ± 0.012
	Built-up	0.91 ± 0.055	0.81 ± 0.072	0.85 ± 0.054
Fused	Woody	0.78 ± 0.024	0.73 ± 0.035	0.75 ± 0.021
	Bunch grass	0.82 ± 0.029	0.86 ± 0.028	0.84 ± 0.021
	Grazing lawn	0.88 ± 0.026	0.85 ± 0.024	0.86 ± 0.020
	Water body	0.96 ± 0.019	0.97 ± 0.011	0.97 ± 0.009
	Bare	0.91 ± 0.02	0.93 ± 0.012	0.92 ± 0.014
	Built-up	0.81 ± 0.066	0.89 ± 0.042	0.85 ± 0.043

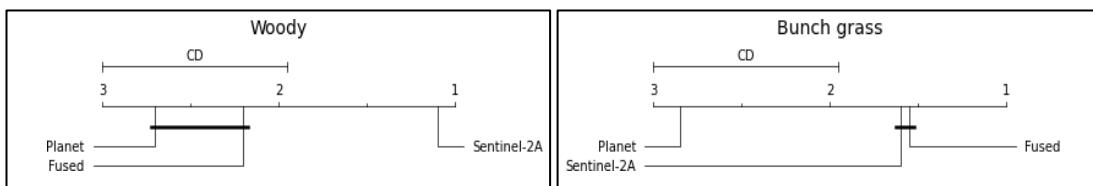
with the Fused image marginally outperforming Planet and Sentinel-2A images by 2% (Table 4.2). All individual land cover classes were successfully discriminated at high accuracies across all datasets. The most obvious inter-class misclassifications occurred among the vegetation classes (see confusion matrices in Appendix ??).

To determine the significance of accuracy differences among the three datasets, the Friedman chi-square ( $\chi^2_F$ ) test for significant differences (at  $\alpha = 0.05$ ) in F-scores for overall map and individual land cover types was used. The test results are presented in Table 4.3 with accompanying post-hoc analysis using Nemenyi test in Figure 4.8 for significantly different class accuracies. Only accuracy results from woody and bunch grass cover classifications showed significant differences in accuracy (Table 4.3).

**Table 4.3:** Summary of Friedman chi-square ( $\chi^2_F$ ) test for significant differences (at  $\alpha = 0.05$ ) in accuracy scores for overall map and individual land cover types. Significance test was conducted on F-scores from stratified 10-fold cross-validation.

Land cover	Friedman statistic ( $\chi^2_F$ )	p-value
Woody	14.11	0.00
Bunch grass	12.05	0.00
Grazing lawn	3.44	0.18
Water body	6.06	0.05
Bare	4.15	0.13
Built-up	2.00	0.37
All	5.60	0.06

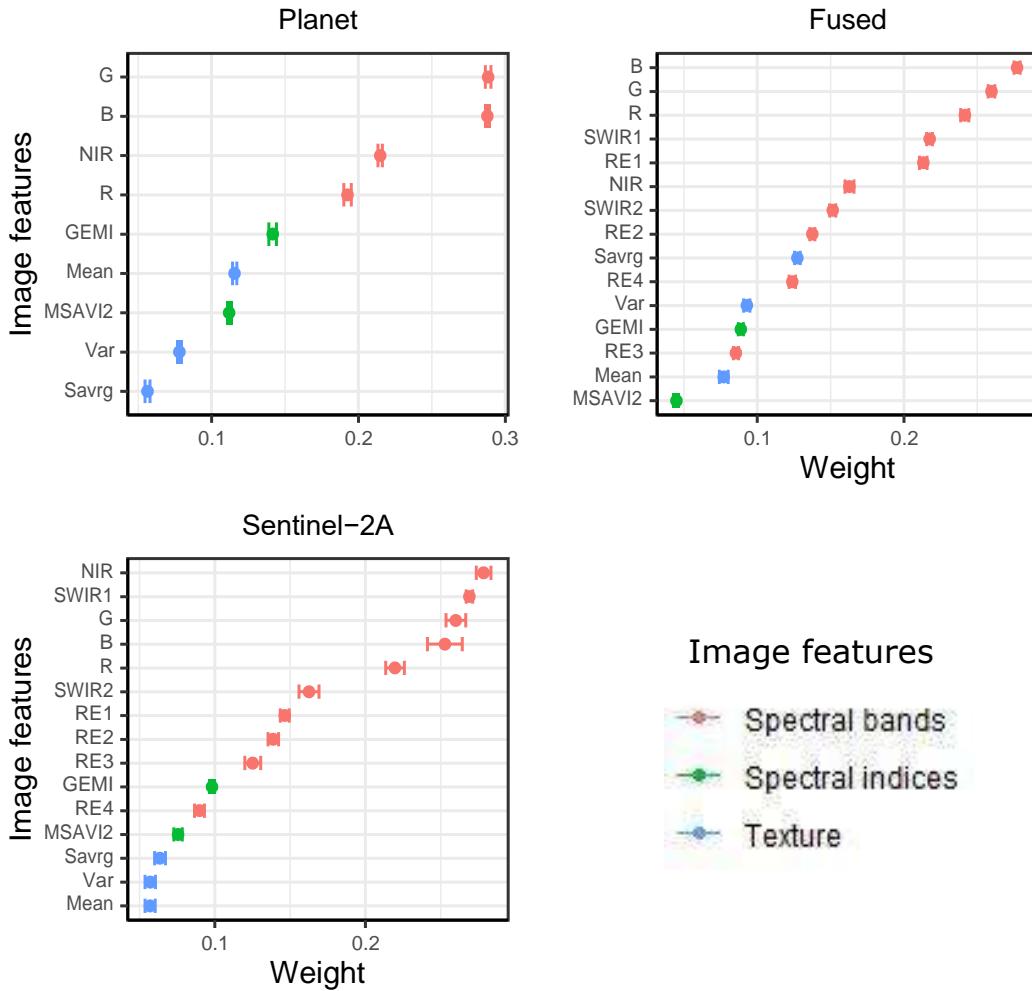
For woody cover, the accuracy estimates from Fused and Planet image pairs were statistically similar, but each paired with results from Sentinel-2A showed significant differences (Figure 4.8). When comparing accuracy scores for bunch grass classification, results from the Planet image was significantly different from results derived from both Fused and Sentinel-2A images. However, comparing results from Fused and Sentinel-2A images showed no significant differences in bunch grass accuracy scores (Figure 4.8).



**Figure 4.8:** Post-hoc comparison of Sentinel-2A, Planet and Fused datasets with Nemenyi test, based on F-scores for Woody and Bunch grass classes. Connected groups are not significantly different at  $\alpha = 0.05$ . Critical Difference (CD) = 1.048

### 4.3.3 Feature Importance

Image feature importance estimates for each classification are presented in Figure 4.9. Generally, the original image bands contributed most to land cover classification accuracy, dominated by the blue, green and red image bands across all datasets. Additional features that comprised the five (5) most important features were near infrared and shortwave infrared-1 for Sentinel-2A, near infrared and Global Environmental Monitoring Index for Planet image and red edge-1 and shortwave infrared-1 for the Fused image (Figure 4.9).

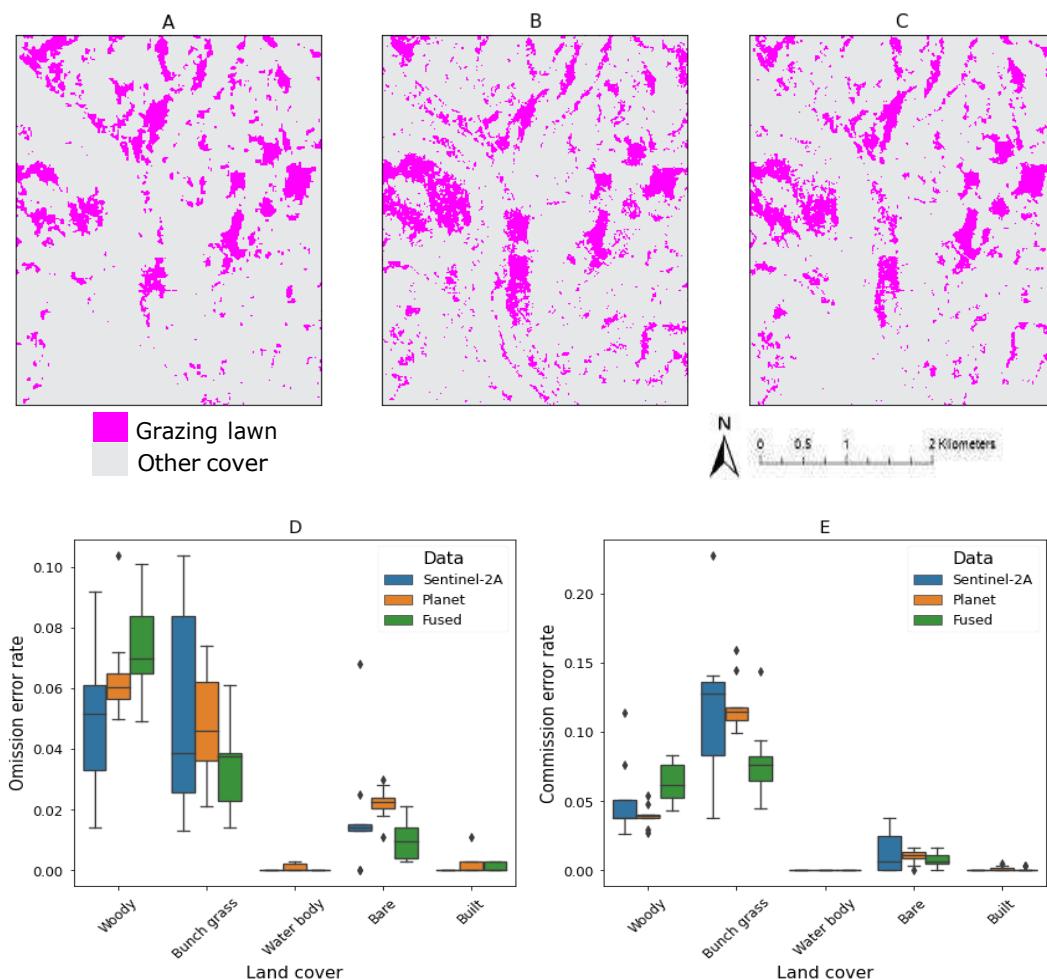


**Figure 4.9:** Image feature importance estimates. Feature weights represent contribution to classification accuracy and are sorted in descending order. For feature acronyms, B = Blue, G = Green, R = Red, RE1 = Red Edge-1, RE2 = Red Edge-2, RE3 = Red Edge-3, RE4 = Red Edge-4, N = Near Infrared, SWIR1 = Shortwave Infrared-1, SWIR2 = Shortwave Infrared-2, GEMI = Global Environmental Monitoring Index, MSAVI2 = Modified Soil Adjusted Vegetation Index-2, Mean = Mean, Var = Variance, Savrg = Sum Average

#### 4.3.4 Grazing Lawn Detection Accuracy

Maps of grazing lawns from each dataset and the distribution of misclassification errors relative to other land cover classes are presented in Figure 4.10. Expectedly, grazing lawn omission and commission errors were dominant among other vegetation classes (i.e. woody and bunch grass), followed by bare surfaces from degraded grazing lawn patches. Woody cover received a greater proportion of misclassified grazing lawn pixels (Figure 4.10D), while relatively more bunch grass pixels were

wrongly classified as grazing lawns (Figure 4.10E) for all datasets. Overall, grazing lawn omission and commission errors in woody cover class were highest for the Fused image and lowest for Sentinel-2A, while the reverse was true for misclassification errors when considering bunch grasses.



**Figure 4.10:** Maps of grazing lawn cover derived from (A) Sentinel-2A, (B) Planet and (C) Fused images; and (D) Omission and (E) Commission error rates for grazing lawn classification against other land cover types. Error rates represent fractions between 0 and 1.

## 4.4 Discussion

### 4.4.1 Image Fusion for high-resolution Savannah Land Cover Classification

This study compared Gram-Schmidt (Laben & Brower, 2000) and Principal Component (Yocky, 1995) pan-sharpening approaches for utilizing the spectral properties of ESA's Sentinel-2A data (Drusch *et al.*, 2012) and the high spatial resolution of the recently released Planet mosaics over the tropics (Planet, 2017), for enhanced land cover classification in a heterogeneous African savannah landscape. Based on results from image quality assessment, the Gram-Schmidt approach produced the most similar image spectral characteristics to the original Sentinel-2 image. This finding corroborates reports of the superiority of Gram-Schmidt transformation for image fusion compared to other techniques (Jenerowicz & Woroszkiewicz, 2016; Rokni *et al.*, 2015; Zhao *et al.*, 2019). For example, Rokni *et al.* (2015) found the Gram-Schmidt image fusion technique to produce better outcomes both in terms of resulting image quality and classification when compared with Modified Intensity-Hue-Saturation, High Pass Filter and Wavelet-PC (Principal Components) techniques for fusion of multi-temporal Landsat Enhanced Thematic Mapper Plus (ETM+) 2000 band 8 (15 m) and Landsat Thematic Mapper (TM) 2010 multispectral (30 m). Based on the results, the Gram-Schmidt fused image was selected for further classification.

While the target spatial resolution of 4.77 m (from the Planet image) was achieved in the resulting Fused images (Figure 4.2) with high correlations between individual bands when compared with the Sentinel-2A data, land cover specific assessment revealed some spectral differences particularly in heterogeneous vegetation classes (see SSIM in Figure 4.3). Traditionally, images are subject to some degree of spectral distortion during pan-sharpening (Pushparaj & Hegde, 2017; Sarp, 2014) which persist even when the panchromatic and multi-spectral bands used are from the same sensor. For example, Sarp (2014) recorded a permanent loss of structural information in pan-sharpened Ikonos and Quickbird images. In the case of this study, the inherent distortion of spectral information from the pan-sharpening process could have been compounded by the use of different sensors with varying spectral response functions, though both Sentinel-2A and Planet images had been pre-processed to surface reflectance. Additionally, the slight differences in temporal window for image acquisition as well as the compositing methods used for data creation likely contributed to the observed spectral difference. For example, the Planet image is a six-

month composite covering June 2019 through to December 2019 (i.e. dry season to early rainy season), and was derived using best available image pixel, whereas the Sentinel-2A image was derived as a median composite of all images between May 2019 and June 2019. Future data interoperability could be enhanced with the release of monthly composites of high resolution Planet imagery from September 2020 onward (Planet, 2017), promising more consistent spectral information with other multi-spectral sensors.

Despite the losses in structural information from pan-sharpening, the spectral integrity of the fused image was most evident in the relatively homogeneous and stable land cover types such as bare patches, built-up and water bodies, with relatively high SSIM statistics (Figure 4.3). This also translated into land cover classification in terms of spatial and thematic details (Figures 4.4, 4.5 and 4.6) as well as accuracy outcomes (Figure 4.7 and Table 4.2). Having more spectral information covering the visible (3 bands), red edge (4 bands), near infrared (1 band) and shortwave infrared (2 bands) at a high spatial scale that can resolve most vegetation patches typical in a heterogeneous savannah landscape, enhanced land cover discrimination space leading to higher classification accuracy for the entire landscape, albeit not statistically significant at  $\alpha = 0.05$  (Table 4.3). Another important finding to emerge from cross-validation accuracy analysis was the markedly low precision in Sentinel-2A accuracy scores, expressed in a much wider inter-quartile range compared to those from Planet and Fused images (Figure 4.7). In general, therefore, it seems that at coarse spatial scales, reference pixels used for training and validation are more prone to high intra-class variation, possibly emanating from the influence of adjacent land cover pixels due to lack of distinct boundaries. In the case of this study, such phenomenon was likely more pronounced due to the high heterogeneity within the studied savannah landscape. This finding suggests the need for extensive reference samples to guarantee consistently high accuracy scores when using Sentinel-2 or data with similar spatial characteristics in heterogeneous landscapes. It should be noted that the Fused image provided the most precise estimates of individual and overall land cover classification accuracies. Contrary to expectations, woody cover classification accuracy deviated from the pattern of dominance by the Fused image out of all individual land cover classes. The Sentinel-2A image recorded significantly higher woody cover classification accuracy compared to the higher spatial resolution

Fused (4 % more) and Planet images (6 % more). Depending on the size distribution of target land cover, medium resolution satellite images can produce similar or even better classification outcomes relative to high spatial resolution imagery (Awuah *et al.*, 2018).

Despite the visible loss of spatial detail in the coarser resolution Sentinel-2A image compared to the Planet and Fused images, all classifications produced very similar thematic representation of land cover (Figures 4.4, 4.5 and 4.6). Marston *et al.* (2017) observed a similar phenomenon in a comparison of very high-resolution and medium resolution satellite image classifications. The marginal differences observed in classification accuracy among the three datasets (see class accuracies in Table 4.2) may be explained by results of feature importance estimates presented in Figure 4.9 which showed similarities in the most important contributing image features to classification accuracy. For example the first five (5) most important image features for classification were dominated by the visible bands (blue, green and red) across all datasets, which highlights the role of similar original image bands in all three classifications. The absence of additional red edge and shortwave infrared bands in the Planet image likely accounted for the differences in classification outcomes in comparison with the Fused image. The additional advantages of red edge bands for land cover mapping in vegetated landscapes have been widely reported in similar savannah applications (Forkuor *et al.*, 2018; Kaszta *et al.*, 2016; Ngadze *et al.*, 2020; Otunga *et al.*, 2019). For example, in a southern African savannah land cover classification context, Kaszta *et al.* (2016) found WorldView-2's red edge band to be crucial in an object-based discrimination of vegetation components. The red edge band captures variations in vegetation pigmentation such as chlorophyll and carotenoids and thus enables successful discrimination of vegetation components even at the level of individual species (Kaszta *et al.*, 2016; Pu & Landry, 2012). Additionally, the shortwave infrared band is highly sensitive to vegetation moisture content due to higher levels of absorption and helps to highlight differences in vegetation states particularly in drier regions such as savannahs (Bueno *et al.*, 2020). When comparing the Fused and Sentinel-2A classifications, similar feature importance estimates were observed for the first five (5) most important image features, with four (4) covering the blue, green, red and shortwave infrared bands. The only difference was the inclusion of near infrared for the Sentinel-2A image and red edge for the Fused image. It can therefore be assumed that the observed accuracy differences between these datasets occurred due to the influence of image spatial resolution. Spatial resolution is widely known to influence classification accuracy, with a general trend favouring higher spatial resolution (Awuah

*et al.*, 2018; Momeni *et al.*, 2016; Suwanprasit & Srichai, 2012). Few studies however report on the accuracy effects of spatial resolution in heterogeneous savannah land cover classification. Our finding is consistent with Marston *et al.* (2017) who recorded higher classification accuracies using VHR imagery than medium resolution images over the same southern African savannah landscape. Similarly, Allan (2007) tested the influence of spatial resolution on savannah land cover classification by comparing SPOT-5, Landsat TM and MODIS sensors and recorded higher accuracy at finer spatial resolution. It is noteworthy that in both studies (Allan, 2007; Marston *et al.*, 2017), the authors used imagery with different spectral properties unlike the similar spectral properties of Fused and Sentinel-2A used in this study. We suggest that optimization of savannah land cover classification would benefit from further robust experimentation of the combinations of spatial resolution and spectral properties that enhance accuracy. Overall, our comparison of classifications derived from the three datasets highlight the combined advantages of high spatial resolution and gain in spectral information from image fusion to enhance land cover classification outcomes.

#### **4.4.2 Grazing Lawn Detection and Monitoring with Open-access Optical Satellite Imagery**

Satellite remote sensing-based monitoring approaches provide vital landscape-scale insights for understanding variations in grazing lawn cover and consequences for broader savannah ecosystem functioning. Apart from overall land cover classification performance, this study set out with the aim of assessing the performance of image fusion for accurate grazing lawn detection. Recent studies have shown that high spatial resolution satellite imagery helps to overcome heterogeneity of savannah landscapes leading to better land cover classification outcomes (Awuah *et al.*, 2020; Kaszta *et al.*, 2016; Marston *et al.*, 2017). Results from this study further emphasizes the utility of both high spatial and spectral resolution imagery by integrating high spatial detail resolved by the Planet image and more useful spectral channels from Sentinel-2A. Image fusion provided the best discrimination space for delineating grazing lawns from other vegetation types (Table 4.2) through the combination of additional red edge and shortwave infrared reflectance (Figure 4.9) and finer spatial scale. This finding supports evidence from previous studies (Underwood *et al.*, 2007; Yu *et al.*, 2020) which reported higher mapping accuracies of different vegetation components using images with high spatial and spectral resolution. Comparing results from the original Planet and Sentinel-2A images however showed similar grazing lawn classification

accuracies. This highlights the particular importance of more varied spectral information in vegetation mapping even at coarser spatial resolutions (Underwood *et al.*, 2007). However, further work is required to develop a full picture of the trade-offs in spectral and spatial resolution in a heterogeneous savannah context.

With similar grazing lawn distribution patterns between the Fused and individual Sentinel-2A and in particular Planet image classifications, this study demonstrates the potential of open-access satellite data for monitoring grazing lawn dynamics. Originally intended for deforestation monitoring over tropical regions, the Planet images come as analysis ready cloud-free composites delivered as both biannual (every six months from December 2015 to August 2020) and monthly collections (from September 2020 onward) (Pandey *et al.*, 2021; Planet, 2017). The free availability of such high resolution analysis ready data (ARD), in addition to the proliferation of cloud-based processing platforms, promises a rapid analysis of grazing lawn attributes at landscape-scale to better understand patterns across different environmental gradients. Of particular interest are the impacts of contemporary threats to savannah ecosystem functioning such as widespread woody vegetation encroachment (Case & Staver, 2017; Mitchard & Flintrop, 2013; Stevens *et al.*, 2017) and frequent, intense and long drought events (Case *et al.*, 2020; Sankaran, 2019), on grazing lawn occurrence and distribution. A substantial extent of grazing lawn cover occurs at the footslopes of catenas adjacent to drainage channels where most woody cover is concentrated. However, the current rate of woody cover intensification within their range of existence (Zhou *et al.*, 2021), coupled with post-drought amelioration of fire (Sankaran, 2019) and dissociation of grazers from grazing lawns (Donaldson *et al.*, 2020), suggest a high likelihood of grazing lawn colonization by adjacent woody vegetation. The combination of historical archives and future acquisitions from Sentinel-2A and high spatial resolution Planet images presents enormous opportunities to investigate such phenomena and recommend spatially targeted management actions.

## 4.5 Conclusions

With the proliferation of different remote sensing datasets comes the opportunity to combine their complementary features for enhanced characterisation of target land cover types, particularly in heterogeneous landscapes. This study presents a multi-sensor image fusion approach for creating images with similar characteristics as the popular high-valued VHR satellite images. The utility of the Fused image compared

with the original Sentinel-2A and Planet images were tested in the context of enhancing land cover classification with the aim of accurately delineating grazing lawns from other vegetation components. The results show that, image fusion yielded the highest and most precise estimates of land cover classification accuracies including grazing lawn detection accuracy in a heterogeneous mesic savannah landscape. The coarser spatial resolution Sentinel-2A image produced a wider range of classification accuracy estimates based on cross-validation analysis, and thus was the least precise. This suggests that, the use of Sentinel-2 images and image data with similar spatial characteristics in heterogeneous landscapes may require more reference samples to ensure consistently high classification accuracy outcomes. The results also show that, depending on factors such as available image quality and degree of multi-sensor image interoperability - mainly due to variation in temporal overlap - as it relates to project objectives, Sentinel-2A or Planet imagery may be individually used for savannah land cover analysis and in particular, grazing lawn monitoring, given the observed marginal differences in classification accuracy and similarities in thematic details. Examining the distribution of grazing lawn classification errors indicates that future work incorporating structural information of different vegetation components could further enhance classification outcomes. This study provides pioneering evidence for the utility of the new NICFI Planet mosaics for land cover classification in a heterogeneous savannah landscape. Ultimately, the results add to the limited literature on remote sensing-based monitoring of grazing lawns and sets a foundation for cost-effective and accurate high resolution analysis of savannah grazing lawn dynamics at a time of contemporary threats to savannah ecosystem functioning such as woody encroachment and frequent drought events.

---

# Chapter 5

## Synthesis and Conclusion

---

### 5.1 Summary

Grazing lawn patches play a central role in maintaining the habitat heterogeneity necessary to support the vast diversity of plants and animals in southern African savannah ecosystems. By influencing patterns in herbivory and fire, which are the main consumers of plant biomass in savannah landscapes, grazing lawns have been linked to important ecosystem processes including nutrient cycling, plant community composition, grazer movement and population dynamics (Archibald *et al.*, 2005; Archibald, 2008; Donaldson *et al.*, 2018; Hempson *et al.*, 2019; Voysey *et al.*, 2021). Further, the characteristic low herbaceous cover of grazing lawn patches makes it easier for tourists to sight iconic African savannah mammal species. As such, the extent and distribution of grazing lawns have a knock-on effect on conservation efforts which depend on revenue from tourism (Gray & Bond, 2013; Lindsey *et al.*, 2007). Eventually, the extent and changes in distribution of grazing lawns have important cascading impacts on the ecological integrity of savannah ecosystems. Yet there is limited understanding of the distribution and spatio-temporal dynamics of grazing lawn cover in savannah ecosystems. Importantly, there is no robust broad scale approach for detecting and monitoring grazing lawns to enable investigation of their distribution and spatio-temporal dynamics. This research addresses these gaps in knowledge using satellite remote sensing. The objectives of this research have been addressed through the analysis presented in Chapter 2, Chapter 3 and Chapter 4. Specifically:

1. The first objective, ***Develop methods for accurate detection and mapping of savannah grazing lawn distribution using very high-resolution satellite imagery and machine learning techniques***, was addressed using analysis in Chapter 2 and Chapter 4. State-of-the-art machine learning algorithms were compared to develop a robust machine learning framework for grazing lawn detection using Worldview-3 images (Chapter 2). Additionally, a fusion approach

- for integrating the spatial and spectral advantages of the recent Planet/NICFI data and Sentinel-2 respectively was developed to demonstrate the utility of open-access high resolution satellite imagery for accurate characterisation of grazing lawns (Chapter 4).
2. The second objective, ***Identify changes in cover and structural distribution of savannah grazing lawns over space and time***, was addressed using analysis in Chapter 3. Post-classification comparison of land cover maps derived from VHR satellite images from 2002, 2014 and 2019, was conducted to highlight patterns of change in grazing lawn cover and structure.
  3. The third objective, ***Identify the dominant drivers of change in grazing lawn cover and structure***, was addressed using analysis in Chapter 2 and Chapter 3. Spatial pattern analysis in Chapter 2 revealed the underlying driver of grazing lawn spatial occurrence under mesic and semi-arid savannah conditions. Further, Chapter 3 highlighted drivers of temporal dynamics in grazing lawn occurrence and the potential impact of drought disturbance on changes in grazing lawn cover.

### **5.1.1 Satellite remote sensing of savannah grazing lawns**

There is substantial evidence of the advantages of remote sensing for vegetation monitoring. While field-based monitoring methods alone offer vital localized details, they are inefficient and costly when dealing with large areas and fail to provide retrospective information necessary for change monitoring. Satellite remote sensing technology, through the provision of synoptic and historical images compliment in-situ field observations and helps to overcome spatio-temporal vegetation monitoring challenges. This is particularly critical from an ecological perspective, which usually involve monitoring at larger spatial and temporal scales to understand underlying processes. This research presents a pioneering attempt to monitor and investigate dynamics in grazing lawn cover and structure using remote sensing. Fundamental to ensuring the success of remote sensing for monitoring grazing lawn dynamics is the ability to detect and characterize grazing lawn patches with remote sensing imagery. Therefore the first objective of this research focused on developing methods for accurate detection and mapping of grazing lawns.

Savannah ecosystems inherently exhibit a high degree of heterogeneity in composition and configuration. From an optical remote sensing perspective, the different composition of vegetation functional types exhibit similar spectral reflectance characteristics. This challenge of spectral similarity between different vegetation components is further compounded by the complex spatial configuration captured in terms of scale and distribution of vegetation patches. In the case of grazing lawns, a spectrum of spatial structure can be observed. Grazing lawns may occur as large continuous patches, typically on sodic sites, or as small patches in a heterogeneous mix of other vegetation components. With a reported average minimum grazing lawn patch size of 2 m in southern African savannahs (Cromsigt & Beest, 2014; Waldram *et al.*, 2008), successful monitoring of grazing lawn extent and distribution require the use of very high-resolution (VHR) data at similar scales. Unlike coarse and medium resolution satellite images which are prone to pixel mixing under spatially heterogeneous savannah landscape conditions, VHR images have been proven to successfully deal with these challenges (Kaszta *et al.*, 2016; Marston *et al.*, 2017). Additionally, multi-spectral VHR images such as WorldView-3 provide the capability of reliably dealing with spectral similarity of different savannah vegetation composition. Instrumental to this is the availability of two near infrared bands and the additional yellow and red-edge bands, which allows detection of photosynthetically active or dying plants and foliar chlorophyll content (Schuster *et al.*, 2015).

The analysis in Chapter 2 therefore involved the use of WorldView-3 imagery for discriminating grazing lawns from other vegetation functional types under mesic and semi arid savannah conditions using machine learning. Application of data science techniques for remote sensing based ecosystem monitoring have been extensively researched (Lawrence & Moran, 2015). However, selecting a fit-for-purpose workflow requires some experimentation due to contextual difference (Foody *et al.*, 2016; Lawrence & Moran, 2015). Four state-of-the-art machine learning models including RF, SVM, CART and MLP were compared. All models achieved very high accuracies, but RF, SVM and MLP models showed similar performance and were generally better than the CART model. The results further highlighted the spectral differences for the same target vegetation types under different savannah landscape conditions. This was apparent in performance reductions when models trained on data from mesic savannah conditions in southern KNP were transferred to the semi-arid savannah landscape towards northern KNP. The KNP landscape is characterized by a rainfall gradient from north to south, with mean annual precipitation increasing along the gradient (Venter *et al.*, 2003). This influences plant moisture stress levels and general

photosynthetic capacity with consequences for spectral characteristics (Sun *et al.*, 2021). Overall, the high spatial resolution of WorldView-3 imagery coupled with a combination of original bands, spectral indices and texture features proved instrumental in dealing with the heterogeneity in savannah vegetation components leading to accurate delineation of grazing lawns.

WorldView-3 imagery provides the ideal spatial and spectral characteristics for vegetation monitoring in heterogeneous savannah landscapes (Awuah *et al.*, 2020; Kaszta *et al.*, 2016; Marston *et al.*, 2017). However, the downside to such VHR imagery is the high cost involved in image acquisition. Using this type of imagery for operational monitoring of grazing lawns would therefore not be economically feasible. Sentinel-2, which is freely available, sample similar spectral bands as WorldView-3 imagery with demonstrated application in savannah vegetation mapping (Symeonakis *et al.*, 2019). However, at the finest spatial resolution of 10 m, Sentinel-2 can only detect relatively large continuous grazing lawn patches, which might limit management decisions. The recent open-access release of analysis-ready Planet image mosaics via a partnership involving Norway's International Climate and Forests Initiative (NICFI), Kongsberger Satellite Services (KSAT), Planet Labs and Airbus (Planet, 2017), represents a major step forward for cost effective high resolution vegetation monitoring. However, in the context of savannah heterogeneity, the spectral characteristics of Planet imagery is limited in resolving the spectral similarities of different vegetation components. Given these challenges, the analysis in Chapter 4 explored options for accurate and cost effective grazing lawn monitoring by leveraging the spatial and spectral advantages of the freely available Planet mosaics and Sentinel-2 satellites images respectively.

Using a Gram-Schmidt pansharpening approach, a new satellite image data set was derived from fusing the high spatial resolution Planet mosaic with the spectral bands from Sentinel-2. This was then tested alongside the original Planet and Sentinel-2 images to assess their utility for savannah grazing lawn detection. The fused image provided the most accurate and precise metrics for grazing lawn and overall savannah vegetation classification. The availability of more spectral information covering the visible, red edge, near infrared and shortwave infrared, at a high spatial scale, enhanced discrimination of different savannah vegetation components for accurate grazing lawn detection. This finding is congruent with other studies which highlight the importance of high spatial and spectral resolution for enhanced classification accuracy of different vegetation types (Underwood *et al.*, 2007; Yu *et al.*, 2020). While the fused image

delivered the most accurate detection of grazing lawn patches, similar distribution of grazing lawns were captured across all datasets. This further highlight the utility of the Planet or Sentinel-2 imagery for rapid assessment of grazing lawn occurrence pattern.

The findings from Chapter 2 and Chapter 4 provide pioneering evidence of the potential of satellite remote sensing for grazing lawn detection. Importantly, this provides a foundation for monitoring grazing lawn dynamics and related impacts on savannah ecosystem functioning, particularly in the context of contemporary threats such as woody encroachment (Case & Staver, 2017; Mitchard & Flintrop, 2013; Stevens *et al.*, 2017) and frequent drought events (Case *et al.*, 2020; Sankaran, 2019)

### **5.1.2 Changes in grazing lawn cover and structure**

Grazing lawn formation is the product of a complex interaction between grazer activity and fire under different landscape productivity conditions (Donaldson *et al.*, 2018; Hempson *et al.*, 2019). Grazing lawn presence in turn strongly influences herbivory and fire regimes, with cascading impacts on biodiversity and savannah ecosystem dynamics (Hempson *et al.*, 2015). By serving as a source of nutrition needed for reproduction and the successful recruitment of juveniles (Gaillard *et al.*, 2000; Parker *et al.*, 2009), while providing refuge from predators (Hempson *et al.*, 2015; Voysey *et al.*, 2021), grazing lawns directly influences herbivore population dynamics and movement. In parallel, grazing lawns effectively curtail fire ignition and propagation within the savannah landscape due to their characteristic low fuel biomass nature (Archibald *et al.*, 2005; Donaldson *et al.*, 2018). From a management and biodiversity conservation perspective, grazing lawns are habitats to a broad diversity of birds and iconic mammals which are of significant tourism interest. This coupled with the ease of sighting animals due to low vegetation cover means that the distribution of grazing lawn patches could have a knock-on effect on much needed revenue for biodiversity conservation (Gray & Bond, 2013; Lindsey *et al.*, 2007). As such, knowledge of changes in the extent, distribution and structure (spatial configuration) of grazing lawns is vital for better understanding savannah ecosystem dynamics and designing effective conservation management decisions, particularly with the observed widespread woody encroachment (Case & Staver, 2017; Marston *et al.*,

2017; Mitchard & Flintrop, 2013; Stevens *et al.*, 2017) and increasing frequency of extreme droughts (Case *et al.*, 2020; Sankaran, 2019). The second objective, which is addressed in Chapter 3, therefore focused on identifying spatio-temporal changes in grazing lawn cover and structure.

Multi-temporal VHR satellite images from IKONOS, QuickBird, WorldView-2, WorldView-3 and Planet sensors covering three sites in southern KNP were used in a post-classification analysis of changes in grazing lawn cover and structure. Post-classification change detection allows thematic quantification of the magnitude and direction of change (Peterson *et al.*, 2004), providing insights into drivers of change over space and time. Images are processed separately without the need for radiometric normalization, which reduces computational cost (Coppin *et al.*, 2004; Tewkesbury *et al.*, 2015). However, the main downside to post-classification change detection is the risk of classification errors propagating into change results (Coppin *et al.*, 2004). In the case of this study, the availability of multi-temporal VHR images helped to minimize classification errors. Further post processing informed by knowledge of grazing lawn ecology helped to focus the analysis on relevant thematic classes and areas within the study sites in order to enhance accuracy of change results. Similar post processing techniques were used by Marston *et al.* (2017) and Peterson *et al.* (2004) to deal with error propagation issues in post-classification change analysis.

Two temporal windows were considered for the change analysis. The period between 2002 to 2014 without any record of significant constraint on landscape productivity, highlighted the regular dynamics in grazing lawn cover and structure. On the other hand, the period between 2014 to 2019 recorded an intense drought event, which provided the opportunity to identify how drought stress alters grazing lawn dynamics. In the absence of drought disturbance, grazing lawn cover increased across all sites by a similar proportion, with most of the gains occurring as an expansion in the extent of already established patches. Where new patches sprung up, they were mostly in areas adjacent to already existing patches. This translated into a higher number of grazing lawn patches per unit area with a higher degree of connectedness. Very little was found in the literature to account empirically for the observed general increase in grazing lawn cover between 2002 and 2014. However, the observed trend could be indicative of an increase in grazer population over the study period. This corroborates the findings of Cromsigt & Beest (2014) in a megaherbivore recolonization experiment, where a higher proportion of short grass cover and 20 times more grazing lawn patches were recorded on high rhino impact than low rhino impact landscapes. For

the period between 2014 and 2019, differential patterns of change in grazing lawn cover and structure were observed within the selected study sites. Marginal changes in grazing lawn extent were observed for Site 1 (increase by 0.7% of landscape area) and Site 2 (decrease by 1.3% of landscape area), whereas Site 3 recorded a significant increase in grazing lawn extent (7.5% of total landscape area). This could be explained by difference in landscape characteristics, with Site 3 largely covered by the Sabie River catchment and adjoining drainage channels, indicative of some surface, soil or leaf moisture availability during drought. It is conceivable, therefore, that Site 3 served as drought refugia, which concentrated grazers in the landscape leading to the creation of short grass patches.

The dominant transitions to and from other habitat types were grazing lawn-to-bunch grass and bunch grass-to-grazing lawns respectively for all study sites. This finding is unsurprising given that bunch grass swards are the most immediate competitors of grazing lawns (Archibald, 2008; Donaldson *et al.*, 2018,2; Hempson *et al.*, 2019). Frequent grazing is required to maintain grazing lawn patches, particularly under highly productive landscape conditions; otherwise facultative bunch grass species out-compete and take over grazing lawn species (Hempson *et al.*, 2015). It is important to note that net gains in grazing lawn cover were recorded for grazing lawn and woody cover transitions between 2002 and 2014 across all sites. This finding is consistent with Voysey *et al.* (2021), who observed that, browsers' use of grazing lawn patches primarily as protective hideouts from predators also prevented woody cover expansion on grazing lawns. However, projected increase in the frequency of intense droughts and their impact on herbivore mortality could disrupt this balance and increase the vulnerability of grazing lawn patches to woody encroachment, as evidenced by the net grazing lawn cover losses to woody cover between 2014 and 2019 for Site 1 and Site 2. In general, drought stress confounded grazing lawn dynamics, the impact of which was strongly mediated by local landscape characteristics.

### **5.1.3 Drivers of change in grazing lawn cover and structure**

The third objective was to identify the dominant drivers of change in grazing lawn cover and structure. This objective draws on findings from Chapter 2 and Chapter 3 to identify the most immediate factors that shape grazing lawn dynamics in southern African savannahs. Results on grazing lawn spatial pattern analysis in Chapter 2 and change detection in Chapter 3 point directly to grazer activity as the most immediate driver of grazing lawn dynamics. Incidentally, continued grazing is among the

most widely reported factors responsible for grazing lawn formation and persistence in southern African savannahs (Archibald, 2008; Cromeis & Olff, 2008; Grant & Scholes, 2006; Hempson *et al.*, 2015, 1). All other factors that act to alter grazer activity such as the availability of resource hot-spots and fire events (Archibald *et al.*, 2005; Archibald, 2008; Donaldson *et al.*, 2018), thus influences grazing lawn dynamics.

One interesting finding is the strong correlation between patterns in grazing lawn structural attributes and distance from water sources, where more and larger grazing lawn patches were found in close proximity to water sources. The availability of reliable water sources is a major limiting factor in dry ecosystems such as savannahs. Water points therefore represent important resource hot-spots which concentrate grazers (Smit *et al.*, 2007; Smith & Buckley, 2011), leading to grazing lawn formation and persistence around water sources. This finding also reflects those of Redfern *et al.* (2003) and Smit *et al.* (2007), who observed that landscape-scale distribution of grazers was generally biased towards areas around reliable rivers and permanent waterholes, particularly in the dry season when moisture content in graze is low. Apart from areas around water points, sodic sites were found to host large contiguous grazing lawn cover. Sodic sites are rich in soil and plant sodium making them very attractive to grazers. Mostly found along drainage lines at the footslope of catenas, the distribution of sodic sites likely contributed to the observed association between grazing lawn presence and water bodies. Other landscape characteristics such as openness has also been found to be attractive to grazers, partly to mitigate the risk of predation (Burkepile *et al.*, 2013). The resulting concentration of grazers in such open savannah landscapes could thus lead to extensive grazing lawn formation, which was evident in the Satara landscape.

Analysis of grazing lawn cover changes showed predominant transitions with tall grasses which points to the interactive influence of fire and grazers on grazing lawn dynamics (Donaldson *et al.*, 2018). Both fire and grazers are consumers of grass biomass (Donaldson *et al.*, 2018; Hempson *et al.*, 2019). Feedback between grass functional traits and consumers shape grassy ecosystem structure into tall grass or short grass communities. Hempson *et al.* (2019) describes this as alternate stable grassland states determined by fire or grazing, where tall grasses dominate fire-driven landscapes and short grass grazing lawns dominate grazer driven landscapes. These positive feedbacks between grass structural states and fire or grazing arise because tall grass and grazing lawn traits that bolster their competitive abilities also attract fire or grazers, making them adaptable under these different consumer re-

gimes (Hempson *et al.*, 2015,1). Tall grasses are relatively unpalatable and highly flammable, with extensive root reserves that allow them to quickly re-sprout after fire. Conversely, grazing lawns are non-flammable, but very palatable and thrive under continuous grazing. Through their stoloniferous growth pattern, grazing lawns successfully hide their reproductive buds from grazers, thus ensuring their persistence under high grazing pressure. Switches between tall grass and grazing lawn states therefore occur with changes in fire regimes and grazer behaviour or population, however, the rate of transition is dependent on landscape productivity (Archibald *et al.*, 2005). Under high productivity conditions, tall grasses shade out grazing lawns, creating flammable conditions through fuel biomass build up, which effectively excludes grazing lawns from the landscape. Frequent grazing is therefore required to maintain grazing lawns under high landscape productivity. Additionally, fire and grazers may interact to shape grazing lawn and tall grass transitions (Donaldson *et al.*, 2018). Post-fire re-sprouts may attract and concentrate grazers on previous tall grass swards, leading to the invasion of grazing lawn species. In parallel, grazer movement away from already established grazing lawn patches could result in their colonization by tall grass species.

## 5.2 Contribution of this research

This research makes significant contributions in two main areas. First, on methodological advancement. The first step to successful monitoring of grazing lawns involves the capability to quantify their occurrence within the savannah landscape. Much of the literature on monitoring savannah grazing lawns focuses on localized field observations and landscape experiments. While such field-based approaches provide detailed insights to advance ecological understanding, they are laborious and costly, particularly in the context of operational management decision making. Additionally, the drivers of grazing lawn dynamics as well as the impact grazing lawns themselves exert on broader ecosystem processes play out at broader spatial and temporal scales, which makes investigations with field-based monitoring inefficient. This research presents the first attempt on the development of a robust broad-scale approach for accurate detection and monitoring of grazing lawns using remote sensing. It provides a novel machine learning-based workflow for grazing lawn detection using very high resolution WorldView-3 imagery. It has also been demonstrated that the spatial and spectral advantages of open-access high and medium resolution

satellite imagery from Planet and Sentinel-2 can be harnessed for accurate and cost effective grazing lawn monitoring at high spatial scales. This particular image fusion approach could have extended applications in different savannah land cover monitoring contexts.

Second, this research contributes to literature on savannah grassland ecology. A greater understanding of grazing lawn spatial distribution and spatio-temporal changes has been achieved, contributing to a better understanding of habitat heterogeneity in southern African savannahs. In particular, the role of water sources in the spatial distribution and persistence of grazing lawns has been demonstrated. Evidence of the main transition pathways between grazing lawns and other habitat types has been provided, which validates previous research focusing on dynamics between grazing lawns and tall grasses. Incidentally, it has been shown that the ongoing phenomenon of woody encroachment is not an immediate threat to grazing lawn patches, but projected increase in frequency and intensity of droughts could hasten invasion of woody cover on grazing lawn patches.

### 5.3 Limitations and future research

The availability of spatially representative and non-autocorrelated in-situ observations (ground-truth) is vital in any remote sensing-based land cover analysis (Congalton & Green, 2019; Foody, 2004). This not only ensures that high quality data are available for calibrating predictive algorithms, but also enables robust validation of outcomes (Foody *et al.*, 2016). In this research most of the field data on the presence of different savannah land cover types were confined to areas around access roads in the study sites due to safety protocols and time limitations during field surveys. As such, field samples were spatially skewed and not representative of the image footprints used for analysis. This may have influenced interpretation of the results. That notwithstanding, a number of steps were taken to ensure high quality data for model calibration and results interpretation. Field locations were sampled beyond 100 m buffer around access roads in order to minimize the possible influence of edge effects. Further, the availability of very high resolution satellite images alongside field photos and in-house experts with extensive experience within the southern African savannah landscape allowed remote selection of spatially extended reference data for analysis. Consequently, reference data from the Lower Sabie region (south KNP), where most of the field survey data were concentrated, were used to calibrate machine learning

models for grazing lawn detection using WorldView-3 imagery (Chapter 2). Higher model performance was achieved within the same landscape. However, transferring the machine learning models to the Satara landscape to the north with relatively drier conditions resulted in a dip in model performance. Future analysis should aim to use reference data that are representative of the spatial scope of application. Further, optimizing savannah vegetation mapping would benefit from a closer analysis of spectral differences in the same vegetation types along precipitation gradients typical of KNP.

This research corroborated evidence of the effectiveness of VHR imagery for vegetation mapping in heterogeneous savannah landscapes. In particular, the WorldView sensor provides vegetation sensitive spectral bands which aided in accurately discriminating grazing lawns from other plant functional types. However, the downside to the use of VHR imagery is the high acquisition cost involved. The closest open-access sensor with similar spectral characteristics is Sentinel-2 MSI (Drusch *et al.*, 2012), whose spatial resolution is relatively coarse for detailed vegetation mapping in heterogeneous savannahs. In this context, the release of high resolution Planet composites (Planet, 2017) with open-access license provides an opportunity to overcome cost limitations of using VHR imagery. This research demonstrated that fusing the high spatial resolution from the Planet data and the spectral bands from Sentinel-2 provides imagery with similar characteristics to that of WorldView, which enhanced accuracy of grazing lawn delineation. A common constraint in the process was the high computing power required for image processing, which is typical for the analysis of VHR images in general. It is envisioned that the increasing availability and access to powerful cloud computing platforms such as Google Earth Engine and Microsoft Planetary Computer would help to offset such computational limitations.

Another important limitation was the access to data on potential drivers of savannah vegetation dynamics such as data on grazer population, fire behaviour and precipitation in the studied landscape. In the absence of such data, this research resorted to reports from literature on grazing lawn ecology, coupled with findings from grazing lawn transition pathways to highlight the dominant drivers of grazing lawn dynamics. Future research should aim to test the relationship between grazing lawn spatial characteristics and data on such biophysical factors to identify empirically the most relevant drivers of grazing lawn dynamics under different landscape conditions.

---

## References

---

- ABBURU, SUNITHA, & GOLLA, SURESH BABU. 2015. Satellite image classification methods and techniques: A review. *International journal of computer applications*, **119**(8).
- ABDI, ABDULHAKIM MOHAMED. 2020. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, **57**(1), 1-20.
- ALI, IFTIKHAR, CAWKWELL, FIONA, DWYER, EDWARD, BARRETT, BRIAN, & GREEN, STUART. 2016. Satellite remote sensing of grasslands: from observation to management. *Journal of Plant Ecology*, **9**(6), 649-671.
- ALLAN, KEAGAN. 2007. *Landcover classification in a heterogenous savanna environment: investigating the performance of an artificial neural network and the effect of image resolution*. Ph.D. thesis.
- ALONSO, MARÍA C, MALPICA, JOSÉ A, & DE AGIRRE, ALEX MARTÍNEZ. 2011. Consequences of the Hughes phenomenon on some classification techniques. *Pages 1–5 of: Proceedings of the ASPRS 2001 annual conference, Milwaukee, Wisconsin, May 1-5.*
- ANDERSON, SAJ, & PEARCE, HG. 2003. Improved methods for the assessment of grassland curing. *Pages 3–6 of: Proceedings: International Wildland Fire Conference.*
- ANGASSA, AYANA, & BAARS, ROBERT MT. 2000. Ecological condition of encroached and non-encroached rangelands in Borana, Ethiopia. *African Journal of Ecology*, **38**(4), 321-328.
- APLIN, PAUL, & SMITH, GEOFFREY M. 2011. Introduction to object-based landscape analysis. *International Journal of Geographical Information Science*, **25**(6), 869-875.
- ARCHIBALD, S, BOND, WJ, STOCK, WD, & FAIRBANKS, DHK. 2005. Shaping the landscape: fire-grazer interactions in an African savanna. *Ecological applications*, **15**(1), 96-109.

- ARCHIBALD, SALLY. 2008. African grazing lawns—how fire, rainfall, and grazer numbers interact to affect grass community states. *The journal of wildlife management*, **72**(2), 492-501.
- ARÉVALO, PAULO, OLOFSSON, PONTUS, & WOODCOCK, CURTIS E. 2020. Continuous monitoring of land change activities and post-disturbance dynamics from Landsat time series: A test methodology for REDD+ reporting. *Remote Sensing of Environment*, **238**, 111051.
- AVANAKI, ALIREZA NASIRI. 2009. Exact global histogram specification optimized for structural similarity. *Optical review*, **16**(6), 613-621.
- AWUAH, KWAME T, NÖLKE, NILS, FREUDENBERG, MAXIMILIAN, DIWAKARA, BN, TEWARI, VP, & KLEINN, CHRISTOPH. 2018. Spatial resolution and landscape structure along an urban-rural gradient: Do they relate to remote sensing classification accuracy?-A case study in the megacity of Bengaluru, India. *Remote Sensing Applications: Society and Environment*, **12**, 89-98.
- AWUAH, KWAME T, APLIN, PAUL, MARSTON, CHRISTOPHER G, POWELL, IAN, & SMIT, IZAK PJ. 2020. Probabilistic mapping and spatial pattern analysis of grazing lawns in Southern African savannahs using WorldView-3 imagery and machine learning techniques. *Remote Sensing*, **12**(20), 3357.
- BALFOUR, DA, & HOWISON, OE. 2002. Spatial and temporal variation in a mesic savanna fire regime: responses to variation in annual rainfall. *African Journal of Range and Forage Science*, **19**(1), 45-53.
- BALZAROLO, MANUELA, VICCA, SARA, NGUY-ROBERTSON, AL, BONAL, DAMIEN, ELBERS, JA, FU, YH, GRÜNWALD, T, HOREMANS, JA, PAPALE, D, PEÑUELAS, JOSEP, et al. 2016. Matching the phenology of Net Ecosystem Exchange and vegetation indices estimated with MODIS and FLUXNET in-situ observations. *Remote Sensing of Environment*, **174**, 290-300.
- BASNET, BIKASH, & VODACEK, ANTHONY. 2015. Tracking land use/land cover dynamics in cloud prone areas using moderate resolution satellite data: A case study in Central Africa. *Remote Sensing*, **7**(6), 6683-6709.
- BASTIN, JEAN-FRANCOIS, BERRAHMOUNI, NORA, GRAINGER, ALAN, MANIATIS, DANAE, MOLLICONE, DANILO, MOORE, REBECCA, PATRIARCA, CHIARA, PICARD, NICOLAS, SPARROW, BEN, ABRAHAM, ELENA MARIA, et al. 2017. The extent of forest in dryland biomes. *Science*, **356**(6338), 635-638.

- BERRY, HH, & Louw, GN. 1982. Nutritional measurements in a population of free-ranging wildebeest in the Etosha National Park. *Madoqua.*, **13**(2), 101-125.
- BERTHRONG, SEAN T, PINERO, GERVASIO, JOBBÁGY, ESTEBAN G, & JACKSON, ROBERT B. 2012. Soil C and N changes with afforestation of grasslands across gradients of precipitation and plantation age. *Ecological Applications*, **22**(1), 76-86.
- BISCHOF, HORST, SCHNEIDER, WERNER, & PINZ, AXEL J. 1992. Multispectral classification of Landsat-images using neural networks. *IEEE transactions on Geoscience and Remote Sensing*, **30**(3), 482-490.
- BITTENCOURT, HELIO RADKE, & CLARKE, ROBIN THOMAS. 2003. Use of classification and regression trees (CART) to classify remotely-sensed digital images. *Pages 3751–3753 of: IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)*, vol. 6. IEEE.
- BLASCHKE, THOMAS, HAY, GEOFFREY J, KELLY, MAGGI, LANG, STEFAN, HOFMANN, PETER, ADDINK, ELISABETH, FEITOSA, RAUL QUEIROZ, VAN DER MEER, FREEK, VAN DER WERFF, HARALD, VAN COILLIE, FRIEKE, et al. 2014. Geographic object-based image analysis-towards a new paradigm. *ISPRS journal of photogrammetry and remote sensing*, **87**, 180-191.
- BOLGER, TP, RIVELLI, AR, & GARDEN, DL. 2005. Drought resistance of native and introduced perennial grasses of south-eastern Australia. *Australian Journal of Agricultural Research*, **56**(11), 1261-1267.
- BOND, WILLIAM J. 2008. What limits trees in C4 grasslands and savannas? *Annual review of ecology, evolution, and systematics*, **39**, 641-659.
- BOND, WILLIAM J, & MIDGLEY, GUY F. 2000. A proposed CO<sub>2</sub>-controlled mechanism of woody plant invasion in grasslands and savannas. *Global Change Biology*, **6**(8), 865-869.
- BORGES, JOANA, HIGGINBOTTOM, THOMAS P, SYMEONAKIS, ELIAS, & JONES, MARTIN. 2020. Sentinel-1 and Sentinel-2 Data for Savannah Land Cover Mapping: Optimising the Combination of Sensors and Seasons. *Remote Sensing*, **12**(23), 3862.

- BOVOLO, FRANCESCA, & BRUZZONE, LORENZO. 2007. A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain. *IEEE Transactions on Geoscience and Remote Sensing*, **45**(1), 218-236.
- BRANDT, MARTIN, TAPPAN, GRAY, DIOUF, ABDOUL AZIZ, BEYE, GORA, MBOW, CHEIKH, & FENSHOLT, RASMUS. 2017. Woody vegetation die off and regeneration in response to rainfall variability in the West African Sahel. *Remote Sensing*, **9**(1), 39.
- BREIMAN, LEO. 2001. Random forests. *Machine learning*, **45**(1), 5-32.
- BREIMAN, LEO. 2017. *Classification and regression trees*. 1st edn. New York, USA: Routledge.
- BUCINI, GABRIELA, SAATCHI, SASSAN, HANAN, NIALL, BOONE, RANDALL B, & SMIT, IZAK. 2009. Woody cover and heterogeneity in the savannas of the Kruger National Park, South Africa. Pages IV-334 of: 2009 IEEE International Geoscience and Remote Sensing Symposium, vol. 4. IEEE.
- BUCINI, GABRIELA, HANAN, NIALL P, BOONE, RANDALL B, SMIT, IZAK PJ, SAATCHI, SASSAN S, LEFSKY, MICHAEL A, & ASNER, GREGORY P. 2010. Woody fractional cover in Kruger National Park, South Africa: remote sensing-based maps and ecological insights. Pages 219–237 of: *Ecosystem function in savannas: measurement and modeling at landscape to global scales*. CRC Press.
- BUENO, INACIO T, McDERMID, GREG J, SILVEIRA, EDUARDA MO, HIRD, JENNIFER N, DOMINGOS, BRENO I, & ACERBI JÚNIOR, FAUSTO W. 2020. Spatial Agreement among Vegetation Disturbance Maps in Tropical Domains Using Landsat Time Series. *Remote Sensing*, **12**(18), 2948.
- BUITENWERF, R, BOND, WJ, STEVENS, N, & TROLLOPE, WSW. 2012. Increased tree densities in South African savannas:> 50 years of data suggests CO<sub>2</sub> as a driver. *Global Change Biology*, **18**(2), 675-684.
- BURKEPILE, DERON E, BURNS, CATHERINE E, TAMBLING, CRAIG J, AMENDOLA, ELIZABETH, BUIS, GREG M, GOVENDER, NAVASHNI, NELSON, VICTORIA, THOMPSON, DAVE I, ZINN, ANDREW D, & SMITH, MELINDA D. 2013. Habitat selection by large herbivores in a southern African savanna: the relative roles of bottom-up and top-down forces. *Ecosphere*, **4**(11), 1-19.

- BUTTERFIELD, HS, & MALMSTRÖM, CM. 2009. The effects of phenology on indirect measures of aboveground biomass in annual grasses. *International Journal of Remote Sensing*, **30**(12), 3133-3146.
- CAMARGO, FLÁVIO F, SANO, EDSON E, ALMEIDA, CLÁUDIA M, MURA, JOSÉ C, & ALMEIDA, TATI. 2019. A comparative assessment of machine-learning techniques for land use and land cover classification of the Brazilian tropical savanna using ALOS-2/PALSAR-2 polarimetric images. *Remote Sensing*, **11**(13), 1600.
- CAMPS-VALLS, GUSTAU, & BRUZZONE, LORENZO. 2009. *Kernel methods for remote sensing data analysis*. 1st edn. Chichester, United Kingdom: John Wiley & Sons.
- CARLSSON, A LISA M, BERGFUR, JENNY, & MILBERG, PER. 2005. Comparison of data from two vegetation monitoring methods in semi-natural grasslands. *Environmental Monitoring and Assessment*, **100**(1-3), 235-248.
- CASADY, GRANT M, VAN LEEUWEN, WILLEM JD, & REED, BRADLEY C. 2013. Estimating winter annual biomass in the Sonoran and Mojave Deserts with satellite-and ground-based observations. *Remote Sensing*, **5**(2), 909-926.
- CASE, MADELON F, & STAVER, A CARLA. 2017. Fire prevents woody encroachment only at higher-than-historical frequencies in a South African savanna. *Journal of Applied Ecology*, **54**(3), 955-962.
- CASE, MADELON F, WIGLEY-COETSEE, CORLI, NZIMA, NOEL, SCOGINGS, PETER F, & STAVER, A CARLA. 2019. Severe drought limits trees in a semi-arid savanna. *Ecology*, **100**(11), e02842.
- CASE, MADELON F, WIGLEY, BENJAMIN J, WIGLEY-COETSEE, CORLI, & CARLA STAVER, A. 2020. Could drought constrain woody encroachers in savannas? *African Journal of Range & Forage Science*, **37**(1), 19-29.
- CAWLEY, GAVIN C, & TALBOT, NICOLA LC. 2010. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, **11**(Jul), 2079-2107.
- CEULEMANS, R, JIANG, XN, & SHAO, BY. 1995. Growth and physiology of one-year old poplar (*Populus*) under elevated atmospheric CO<sub>2</sub> levels. *Annals of Botany*, **75**(6), 609-617.

- CHAVES, ALEXANDRE V, BURKE, JENNIFER L, WAGHORN, GARRY C, & BROOKES, IAN M. 2006. Digestion kinetics of leaf, stem and inflorescence from five species of mature grasses. *Journal of the Science of Food and Agriculture*, **86**(5), 816-825.
- CHEHATA, NESRINE, ORNY, CAMILLE, BOUKIR, SAMIA, & GUYON, DOMINIQUE. 2011. Object-based forest change detection using high resolution satellite images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **38**, 5-7.
- CHO, MOSES AZONG, MATHIEU, RENAUD, ASNER, GREGORY P, NAIDOO, LAVEN, VAN AARDT, JAN, RAMOELLO, ABEL, DEBBA, PRAVESH, WESSELS, KONRAD, MAIN, RUSSELL, SMIT, IZAK PJ, et al. 2012. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment*, **125**, 214-226.
- CONGALTON, RUSSELL G, & GREEN, KASS. 2019. *Assessing the accuracy of remotely sensed data: principles and practices*. 3rd edn. Boca Ranton, USA: CRC press.
- COPPIN, POL, JONCKHEERE, INGE, NACKAERTS, KRISTIAAN, MUYS, BART, & LAMBIN, ERIC. 2004. Review ArticleDigital change detection methods in ecosystem monitoring: a review. *International journal of remote sensing*, **25**(9), 1565-1596.
- CORCORAN, JENNIFER, KNIGHT, JOSEPH, PELLETIER, KEITH, RAMPI, LIAN, & WANG, YAN. 2015. The effects of point or polygon based training data on RandomForest classification accuracy of wetlands. *Remote Sensing*, **7**(4), 4002-4025.
- CORTES, CORINNA, & VAPNIK, VLADIMIR. 1995. Support-vector networks. *Machine learning*, **20**(3), 273-297.
- CRAINE, JOSEPH M, OCHELTREE, TROY W, NIPPERT, JESSE B, TOWNE, E GENE, SKIBBE, ADAM M, KEMBEL, STEVEN W, & FARGIONE, JOSEPH E. 2013. Global diversity of drought tolerance and grassland climate-change resilience. *Nature Climate Change*, **3**(1), 63-67.
- CROMSIGT, JORIS PGM, & BEEST, MARISKA. 2014. Restoration of a megaherbivore: landscape-level impacts of white rhinoceros in Kruger National Park, South Africa. *Journal of Ecology*, **102**(3), 566-575.

- CROMSIGT, JORIS PGM, & KUIJPER, DRIES PJ. 2011. Revisiting the browsing lawn concept: Evolutionary Interactions or pruning herbivores? *Perspectives in Plant Ecology, Evolution and Systematics*, **13**(3), 207-215.
- CROMSIGT, JORIS PGM, & OLFF, HAN. 2008. Dynamics of grazing lawn formation: an experimental test of the role of scale-dependent processes. *Oikos*, **117**(10), 1444-1452.
- DEL FRATE, FABIO, PACIFICI, FABIO, SCHIAVON, GIOVANNI, & SOLIMINI, CHIARA. 2007. Use of neural networks for automatic classification from high-resolution images. *IEEE transactions on geoscience and remote sensing*, **45**(4), 800-809.
- DEL POZO, ALEJANDRO, OVALLE, CARLOS, CASADO, MIGUEL A, ACOSTA, BELÉN, & DE MIGUEL, JOSÉ M. 2006. Effects of grazing intensity in grasslands of the Espinal of central Chile. *Journal of Vegetation Science*, **17**(6), 791-798.
- DEMŠAR, JANEZ. 2006. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research*, **7**, 1-30.
- DENG, JS, WANG, K, DENG, YH, & QI, GJ. 2008. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, **29**(16), 4823-4838.
- DEVELOPMENT TEAM, R CORE. 2016. *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria.
- DEVINE, AISLING P, STOTT, IAIN, McDONALD, ROBBIE A, & MACLEAN, ILYA MD. 2015. Woody cover in wet and dry African savannas after six decades of experimental fires. *Journal of Ecology*, **103**(2), 473-478.
- DONALDSON, JASON E, ARCHIBALD, SALLY, GOVENDER, NAVASHNI, POLLARD, DREW, LUHDO, ZOË, & PARR, CATHERINE L. 2018. Ecological engineering through fire-herbivory feedbacks drives the formation of savanna grazing lawns. *Journal of Applied Ecology*, **55**(1), 225-235.
- DONALDSON, JASON ERNEST, PARR, CATHERINE L, MANGENA, E, & ARCHIBALD, SALLY. 2020. Droughts decouple African savanna grazers from their preferred forage with consequences for grassland productivity. *Ecosystems*, **23**(3), 689-701.

- DORMANN, CARSTEN F, ELITH, JANE, BACHER, SVEN, BUCHMANN, CARSTEN, CARL, GUDRUN, CARRÉ, GABRIEL, MARQUÉZ, JAIME R GARCÍA, GRUBER, BERND, LAFOURCADE, BRUNO, LEITÃO, PEDRO J, et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**(1), 27-46.
- DOXANI, G, KARANTZALOS, KONSTANTINOS, & TSAKIRI-STRATI, MARIA. 2012. Monitoring urban changes based on scale-space filtering and object-oriented classification. *International Journal of Applied Earth Observation and Geoinformation*, **15**, 38-48.
- DRUSCH, MATTHIAS, DEL BELLO, UMBERTO, CARLIER, SÉBASTIEN, COLIN, OLIVIER, FERNANDEZ, VERONICA, GASCON, FERRAN, HOERSCH, BIANCA, ISOLA, CLAUDIA, LABERINTI, PAOLO, MARTIMORT, PHILIPPE, et al. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment*, **120**, 25-36.
- EISAVI, VAHID, HOMAYOUNI, SAEID, YAZDI, AHMAD MALEKNESHAD, & ALIMOHAMMADI, ABBAS. 2015. Land cover mapping based on random forest classification of multitemporal spectral and thermal images. *Environmental monitoring and assessment*, **187**(5), 291.
- ELHAG, MOHAMED. 2016. Evaluation of different soil salinity mapping using remote sensing techniques in arid ecosystems, Saudi Arabia. *Journal of Sensors*, **2016**.
- EVANS, JEFFREY S, & RAM, KARTHIK. 2019. Package ‘spatialEco’. Available at <https://github.com/jeffreyevans/spatialEco>. <https://pbil.univ-lyon1.fr/CRAN/web/packages/spatialEco/spatialEco.pdf>.
- FAJJI, NAHOM GUDETA, PALAMULENI, LOBINA G, & MLAMBO, VICTOR. 2017. Evaluating derived vegetation indices and cover fraction to estimate rangeland aboveground biomass in semi-arid environments. *South African Journal of Geomatics*, **6**(3), 333-348.
- FALCO, NICOLA, DALLA MURA, MAURO, BOVOLO, FRANCESCA, BENEDIKTSSON, JON ATLI, & BRUZZONE, LORENZO. 2010. Study on the capabilities of morphological attribute profiles in change detection on VHR images. *Page 783016 of: Image and Signal Processing for Remote Sensing XVI*, vol. 7830. International Society for Optics and Photonics.

- FEBRUARY, EDMUND C, HIGGINS, STEVEN I, BOND, WILLIAM J, & SWEMMER, LOUISE. 2013. Influence of competition and rainfall manipulation on the growth responses of savanna trees and grasses. *Ecology*, **94**(5), 1155-1164.
- FEHRMANN, LUTZ, KUKUNDA, COLLINS B, NÖLKE, NILS, SCHNELL, SEBASTIAN, SEIDEL, DOMINIK, MAGNUSEN, STEEN, & KLEINN, CHRISTOPH. 2019. A unified framework for land cover monitoring based on a discrete global sampling grid (GSG). *Environmental monitoring and assessment*, **191**(2), 46.
- FENSHAM, RJ, FAIRFAX, RJ, & WARD, DP. 2009. Drought-induced tree death in savanna. *Global Change Biology*, **15**(2), 380-387.
- FERCHICHI, AHLEM, BOULILA, WADII, & FARAH, IMED RIADH. 2018. Reducing uncertainties in land cover change models using sensitivity analysis. *Knowledge and Information Systems*, **55**(3), 719-740.
- FLYNN, E SCOTT, DOUGHERTY, CHARLES T, & WENDROTH, OLE. 2008. Assessment of pasture biomass with the normalized difference vegetation index from active ground-based sensors. *Agronomy Journal*, **100**(1), 114-121.
- FOODY, GILES, PAL, MAHESH, ROCCHINI, DUCCIO, GARZON-LOPEZ, CAROL, & BASTIN, LUCY. 2016. The sensitivity of mapping methods to reference data quality: Training supervised image classifications with imperfect reference data. *ISPRS International Journal of Geo-Information*, **5**(11), 199.
- FOODY, GILES M. 2004. Thematic map comparison. *Photogrammetric Engineering & Remote Sensing*, **70**(5), 627-633.
- FORKUOR, GERALD, DIMOBE, KANGBENI, SERME, IDRIS, & TONDOH, JEROME EBAGNERIN. 2018. Landsat-8 vs. Sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *G/Science & remote sensing*, **55**(3), 331-354.
- FRANCINI, SAVERIO, McROBERTS, RONALD E, GIANNETTI, FRANCESCA, MEN-CUCCI, MARCO, MARCHETTI, MARCO, SCARASCIA MUGNOZZA, GIUSEPPE, & CHIRICI, GHERARDO. 2020. Near-real time forest change detection using PlanetScope imagery. *European Journal of Remote Sensing*, **53**(1), 233-244.
- FRIEDMAN, MILTON. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association*, **32**(200), 675-701.

- FRIEDMAN, MILTON. 1940. A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, **11**(1), 86-92.
- GAILLARD, J-M, FESTA-BIANCHET, M, YOCOZ, NG, LOISON, A, & TOIGO, C. 2000. Temporal variation in fitness components and population dynamics of large herbivores. *Annual Review of Ecology and Systematics*, **31**(1), 367-393.
- GAITÁN, JUAN J, BRAN, DONALDO, OLIVA, GABRIEL, CIARI, GEORGINA, NAKAMATSU, VIVIANA, SALOMONE, JORGE, FERRANTE, DANIELA, BUONO, GUSTAVO, MASSARA, VIRGINIA, HUMANO, GERVASIO, et al. 2013. Evaluating the performance of multiple remote sensing indices to predict the spatial variability of ecosystem structure and functioning in Patagonian steppes. *Ecological indicators*, **34**, 181-191.
- GANDHI, G MEERA, PARTHIBAN, S, THUMMALU, NAGARAJ, & CHRISTY, A. 2015. NDVI: vegetation change detection using remote sensing and GIS-a case study of Vellore District. *Procedia Computer Science*, **57**, 1199-1210.
- GAO, YAN, & MAS, JEAN FRANCOIS. 2008. A Comparison of the Performance of Pixel Based and Object Based Classifications over Images with Various Spatial Resolutions. *Online journal of earth sciences*, **2**(1), 27-35.
- GHANNOUM, OULA. 2009. C4 photosynthesis and water stress. *Annals of botany*, **103**(4), 635-644.
- GHIMIRE, BARDAN, ROGAN, JOHN, GALIANO, VÍCTOR RODRÍGUEZ, PANDAY, PRAJJWAL, & NEETI, NEETI. 2012. An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience & Remote Sensing*, **49**(5), 623-643.
- GHOSH, ANIRUDDHA, & JOSHI, PAWAN KUMAR. 2014. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. *International Journal of Applied Earth Observation and Geoinformation*, **26**, 298-311.
- GILL, JACQUELYN L, WILLIAMS, JOHN W, JACKSON, STEPHEN T, LININGER, KATHERINE B, & ROBINSON, GUY S. 2009. Pleistocene megafaunal collapse, novel plant communities, and enhanced fire regimes in North America. *Science*, **326**(5956), 1100-1103.

GOODFELLOW, IAN, BENGIO, YOSHUA, & COURVILLE, AARON. 2016. *Deep learning*. Cambridge, USA: MIT press. <http://www.deeplearningbook.org>.

GORELICK, NOEL, HANCHER, MATT, DIXON, MIKE, ILYUSHCHENKO, SIMON, THAU, DAVID, & MOORE, REBECCA. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, **202**, 18-27.

GOVENDER, NAVASHNI, TROLLOPE, WINSTON SW, & VAN WILGEN, BRIAN W. 2006. The effect of fire season, fire frequency, rainfall and management on fire intensity in savanna vegetation in South Africa. *Journal of Applied Ecology*, **43**(4), 748-758.

GRANITTO, PABLO M, FURLANELLO, CESARE, BIASIOLI, FRANCO, & GASPERI, FLAVIA. 2006. Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. *Chemometrics and Intelligent Laboratory Systems*, **83**(2), 83-90.

GRANT, CC, & SCHOLES, MC. 2006. The importance of nutrient hot-spots in the conservation and management of large wild mammalian herbivores in semi-arid savannas. *Biological Conservation*, **130**(3), 426-437.

GRAY, EMMA F, & BOND, WILLIAM J. 2013. Will woody plant encroachment impact the visitor experience and economy of conservation areas? *Koedoe: African Protected Area Conservation and Science*, **55**(1), 1-9.

GRIFFITH, DANIEL M, LEHMANN, CAROLINE ER, STRÖMBERG, CAROLINE AE, PARR, CATHERINE L, PENNINGTON, R TOBY, SANKARAN, MAHESH, RATNAM, JAYASHREE, STILL, CHRISTOPHER J, POWELL, REBECCA L, HANAN, NIALL P, et al. 2017. Comment on “The extent of forest in dryland biomes”. *Science*, **358**(6365).

GUERINI FILHO, MARILDO, KUPLICH, TATIANA MORA, & QUADROS, FERNANDO LF DE. 2020. Estimating natural grassland biomass by vegetation indices using Sentinel 2 remote sensing data. *International Journal of Remote Sensing*, **41**(8), 2861-2876.

GUYON, ISABELLE, WESTON, JASON, BARNHILL, STEPHEN, & VAPNIK, VLADIMIR. 2002. Gene selection for cancer classification using support vector machines. *Machine learning*, **46**(1-3), 389-422.

- HAKL, J, HREVUŠOVÁ, Z, HEJCMAN, M, & FUKSA, P. 2012. The use of a rising plate meter to evaluate lucerne (*Medicago sativa* L.) height as an important agronomic trait enabling yield estimation. *Grass and Forage Science*, **67**(4), 589-596.
- HAKLAY, MORDECHAI, & WEBER, PATRICK. 2008. Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, **7**(4), 12-18.
- HARALICK, ROBERT M, SHANMUGAM, KARTHIKEYAN, & DINSTEIN, ITS' HAK. 1973. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, 610-621.
- HARTMANN, HENRIK, MOURA, CATARINA F, ANDERECK, WILLIAM RL, RUEHR, NADINE K, SALMON, YANN, ALLEN, CRAIG D, ARNDT, STEFAN K, BRESHEARS, DAVID D, DAVI, HENDRIK, GALBRAITH, DAVID, *et al.* 2018. Research frontiers for improving our understanding of drought-induced tree and forest mortality. *New Phytologist*, **218**(1), 15-28.
- HE, YUHONG, GUO, XULIN, & WILMSHURST, JF. 2009. Reflectance measures of grassland biophysical structure. *International Journal of Remote Sensing*, **30**(10), 2509-2521.
- HELMAN, DAVID, LENSKY, ITAMAR M, TESSLER, NAAMA, & OSEM, YAGIL. 2015. A phenology-based method for monitoring woody and herbaceous vegetation in Mediterranean forests from NDVI time series. *Remote Sensing*, **7**(9), 12314-12335.
- HEMPSON, GARETH P, ARCHIBALD, SALLY, BOND, WILLIAM J, ELLIS, ROGER P, GRANT, CORNELIA C, KRUGER, FRED J, KRUGER, LAURENCE M, MOXLEY, COURTNEY, OWEN-SMITH, NORMAN, PEEL, MIKE JS, *et al.* 2015. Ecology of grazing lawns in Africa. *Biological Reviews*, **90**(3), 979-994.
- HEMPSON, GARETH P, ARCHIBALD, SALLY, DONALDSON, JASON E, & LEHMANN, CAROLINE ER. 2019. Alternate grassy ecosystem states are determined by palatability-flammability trade-offs. *Trends in ecology & evolution*, **34**(4), 286-290.
- HEROLD, A. 2001. Remote sensing and spatial metrics-a new approach for the description of structures and changes in urban areas. Pages 366–368 of: *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217)*, vol. 1. IEEE.

- HIRAYAMA, HIDETAKE, SHARMA, RAM C, TOMITA, MIZUKI, & HARA, KEITAROU. 2019. Evaluating multiple classifier system for the reduction of salt-and-pepper noise in the classification of very-high-resolution satellite images. *International journal of remote sensing*, **40**(7), 2542-2557.
- HIRD, JENNIFER N, DELANCEY, EVAN R, McDERMID, GREGORY J, & KARIYEVA, JAHAN. 2017. Google Earth Engine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping. *Remote sensing*, **9**(12), 1315.
- HOLDEN, CE, AREVELO, P, & PASQUARELLA, V. 2016. *Yet Another Time Series Model (YATSM): V0. 6.1.*
- HUANG, CHENGQUAN, DAVIS, LS, & TOWNSHEND, JRG. 2002. An assessment of support vector machines for land cover classification. *International Journal of remote sensing*, **23**(4), 725-749.
- HUANG, CHO-YING, & GEIGER, ERIKA L. 2008. Climate anomalies provide opportunities for large-scale mapping of non-native plant abundance in desert grasslands. *Diversity and Distributions*, **14**(5), 875-884.
- HUANG, NI, HE, JIN-SHENG, & NIU, ZHENG. 2013. Estimating the spatial pattern of soil respiration in Tibetan alpine grasslands using Landsat TM images and MODIS data. *Ecological Indicators*, **26**, 117-125.
- HUGHES, GORDON. 1968. On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory*, **14**(1), 55-63.
- HUNT JR, E RAYMOND, DAUGHTRY, CRAIG ST, & LI, LI. 2016. Feasibility of estimating leaf water content using spectral indices from WorldView-3's near-infrared and shortwave infrared bands. *International Journal of Remote Sensing*, **37**(2), 388-402.
- HUNTLEY, BJ. 1982. Southern African savannas. Pages 101–119 of: *Ecology of tropical savannas*. Springer.
- IMMITZER, MARKUS, ATZBERGER, CLEMENT, & KOUKAL, TATJANA. 2012. Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote sensing*, **4**(9), 2661-2693.

- INGLADA, JORDI, & CHRISTOPHE, EMMANUEL. 2009. The Orfeo Toolbox remote sensing image processing software. *Pages IV–733 of: 2009 IEEE International Geoscience and Remote Sensing Symposium*, vol. 4. IEEE.
- JANSSEN, LUCAS LF, & VANDERWEL, FRANS JM. 1994. Accuracy assessment of satellite derived land-cover data: a review. *Photogrammetric engineering and remote sensing;(United States)*, **60**(4).
- JENEROWICZ, AGNIESZKA, & WOROSZKIEWICZ, MALGORZATA. 2016. The pan-sharpening of satellite and UAV imagery for agricultural applications. *Page 99981S of: Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII*, vol. 9998. International Society for Optics and Photonics.
- JENSEN, JOHN R. 2015. *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Boston, USA: Prentice Hall Press.
- JIN, XIAOYING. 2012 (September). *Segmentation-based image processing system*. US Patent 8,260,048.
- JIN, YUNXIANG, YANG, XIUCHUN, QIU, JIANJUN, LI, JINYA, GAO, TIAN, Wu, QIONG, ZHAO, FEN, MA, HAILONG, YU, HAIDA, & XU, BIN. 2014. Remote sensing-based biomass estimation and its spatio-temporal variations in temperate grassland, Northern China. *Remote Sensing*, **6**(2), 1496-1513.
- JOHANSEN, KASPER, & PHINN, STUART. 2004. Mapping indicators of riparian vegetation health using IKONOS and Landsat-7 ETM+ image data in Australian tropical savannas. *Pages 1559–1562 of: IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, vol. 3. IEEE.
- KANELLOPOULOS, I, VARFIS, A, WILKINSON, GG, & MEGIER, J. 1992. Land-cover discrimination in SPOT HRV imagery using an artificial neural network—a 20-class experiment. *International journal of remote sensing*, **13**(5), 917-924.
- KASZTA, ŹANETA, VAN DE KERCHOVE, RUBEN, RAMOELLO, ABEL, CHO, MOSES, MADONSELA, SABELO, MATHIEU, RENAUD, & WOLFF, ELÉONORE. 2016. Seasonal separation of African savanna components using worldview-2 imagery: a comparison of pixel-and object-based approaches and selected classification algorithms. *Remote Sensing*, **8**(9), 763.

- KENNEDY, ROBERT E, YANG, ZHIQIANG, & COHEN, WARREN B. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sensing of Environment*, **114**(12), 2897-2910.
- KHORRAM, SIAMAK, VAN DER WIELE, CYNTHIA F, KOCH, FRANK H, NELSON, STACY AC, & POTTS, MATTHEW D. 2016. Future trends in remote sensing. Pages 277–285 of: *Principles of Applied Remote Sensing*. Springer.
- KINGMA, DIEDERIK P, & BA, JIMMY. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- KLARIC, MATTHEW N, CLAYWELL, BRIAN C, SCOTT, GRANT J, HUDSON, NICHOLAS J, SJAHPUTERA, OZY, LI, YONGHONG, BARRATT, SETH T, KELLER, JAMES M, & DAVIS, CURT H. 2013. GeoCDX: An automated change detection and exploitation system for high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **51**(4), 2067-2086.
- KLEINN, CHRISTOPH, & VILČKO, FRANTI EK. 2006. Design-unbiased estimation for point-to-tree distance sampling. *Canadian Journal of Forest Research*, **36**(6), 1407-1414.
- KLEYNHANS, ELIZABETH J, JOLLES, ANNA E, Bos, MARJAN RE, & OLFF, HAN. 2011. Resource partitioning along multiple niche dimensions in differently sized African savanna grazers. *Oikos*, **120**(4), 591-600.
- KROOK, KRISSIE, BOND, WILLIAM J, & HOCKEY, PHILLIP AR. 2007. The effect of grassland shifts on the avifauna of a South African savanna. *Ostrich-Journal of African Ornithology*, **78**(2), 271-279.
- KUKUNDA, COLLINS B, DUQUE-LAZO, JOAQUÍN, GONZÁLEZ-FERREIRO, EDUARDO, THADEN, HAUKE, & KLEINN, CHRISTOPH. 2018. Ensemble classification of individual Pinus crowns from multispectral satellite imagery and airborne LiDAR. *International journal of applied earth observation and geoinformation*, **65**, 12-23.
- LABEN, CRAIG A, & BROWER, BERNARD V. 2000 ( 4). *Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening*. US Patent 6,011,875.

- LAOSUWAN, TEERAWONG, & UTTARUK, PORNCHAI. 2014. Estimating tree biomass via remote sensing, MSAVI 2, and fractional cover model. *IETE Technical Review*, **31**(5), 362-368.
- LASO, FRANCISCO J, BENÍTEZ, FÁTIMA L, RIVAS-TORRES, GONZALO, SAMPEDRO, CAROLINA, & ARCE-NAZARIO, JAVIER. 2020. Land cover classification of complex agroecosystems in the non-protected highlands of the Galapagos Islands. *Remote Sensing*, **12**(1), 65.
- LAWRENCE, RICK L, & MORAN, CHRISTOPHER J. 2015. The AmericaView classification methods accuracy comparison project: A rigorous approach for model selection. *Remote Sensing of Environment*, **170**, 115-120.
- LEONARD, STEVE, KIRKPATRICK, JAMIE, & MARSDEN-SMEDLEY, JONATHAN. 2010. Variation in the effects of vertebrate grazing on fire potential between grassland structural types. *Journal of Applied Ecology*, **47**(4), 876-883.
- LEVICK, SHAUN R, ASNER, GREGORY P, KENNEDY-BOWDOIN, TY, & KNAPP, DAVID E. 2009. The relative influence of fire and herbivory on savanna three-dimensional vegetation structure. *Biological Conservation*, **142**(8), 1693-1700.
- LI, CONGCONG, WANG, JIE, WANG, LEI, HU, LUANYUN, & GONG, PENG. 2014. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. *Remote sensing*, **6**(2), 964-983.
- LINDSEY, PETER A, ALEXANDER, R, MILLS, MGL, ROMAÑACH, S, & WOODROFFE, R. 2007. Wildlife viewing preferences of visitors to protected areas in South Africa: implications for the role of ecotourism in conservation. *Journal of Ecotourism*, **6**(1), 19-33.
- LIU, BOYU, CHEN, JUN, CHEN, JIAGE, & ZHANG, WEIWEI. 2018. Land Cover Change Detection Using Multiple Shape Parameters of Spectral and NDVI Curves. *Remote Sensing*, **10**(8), 1251.
- LIU, YAN, HILL, MICHAEL J, ZHANG, XIAOYANG, WANG, ZHUOSEN, RICHARDSON, ANDREW D, HUFKENS, KOEN, FILIPPA, GIANLUCA, BALDOCCHI, DENNIS D, MA, SIYAN, VERFAILLIE, JOSEPH, et al. 2017. Using data from Landsat, MODIS, VIIRS and PhenoCams to monitor the phenology of California oak/grass savanna and open grassland across spatial scales. *Agricultural and Forest Meteorology*, **237**, 311-325.

- LOH, WEI-YIN. 2011. Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, **1**(1), 14-23.
- LONG, YU, LI, ZHOU, WEI, LIU, & HUA-KUN, ZHOU. 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.
- LOUIS, JÉRÔME, DEBAECKER, VINCENT, PFLUG, BRINGFRIED, MAIN-KNORN, MAGDALENA, BIENIARZ, JAKUB, MUELLER-WILM, UWE, CADAU, ENRICO, & GASCON, FERRAN. 2016. Sentinel-2 sen2cor: L2a processor for users. *Pages 1–8 of: Proceedings Living Planet Symposium 2016*. Spacebooks Online.
- LU, DENGSHENG. 2005. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *International Journal of Remote Sensing*, **26**(12), 2509-2525.
- LU, DENGSHENG, MAUSEL, P, BRONDIZIO, E, & MORAN, EMILIO. 2004. Change detection techniques. *International journal of remote sensing*, **25**(12), 2365-2401.
- MA, LEI, LI, MANCHUN, MA, XIAOXUE, CHENG, LIANG, DU, PEIJUN, & LIU, YONGXUE. 2017. A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, **130**, 277-293.
- MACFADYEN, SANDRA, ZAMBATIS, NICK, VAN TEEFFELEN, ASTRID JA, & HUI, CANG. 2018. Long-term rainfall regression surfaces for the Kruger National Park, South Africa: A spatio-temporal review of patterns from 1981 to 2015. *International Journal of Climatology*, **38**(5), 2506-2519.
- MACGREGOR, SHAUN D, & O'CONNOR, TIMOTHY G. 2002. Patch dieback of Colophospermum mopane in a dysfunctional semi-arid African savanna. *Austral Ecology*, **27**(4), 385-395.
- MARSTON, CHRISTOPHER G, APLIN, PAUL, WILKINSON, DAVID M, FIELD, RICHARD, & O'REGAN, HANNAH J. 2017. Scrubbing up: multi-scale investigation of woody encroachment in a southern African savannah. *Remote Sensing*, **9**(5), 419.
- MARSTON, CHRISTOPHER G, WILKINSON, DAVID M, REYNOLDS, SALLY C, LOUYS, JULIEN, & O'REGAN, HANNAH J. 2019. Water availability is a principal driver of large-scale land cover spatial heterogeneity in sub-Saharan savannahs. *Landscape ecology*, **34**(1), 131-145.

- MAXWELL, AARON E, & WARNER, TIMOTHY A. 2015. Differentiating mine-reclaimed grasslands from spectrally similar land cover using terrain variables and object-based machine learning classification. *International Journal of Remote Sensing*, **36**(17), 4384-4410.
- MAXWELL, AARON E, WARNER, TIMOTHY A, & FANG, FANG. 2018. Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, **39**(9), 2784-2817.
- MAXWELL, AE, WARNER, TA, STRAGER, MP, CONLEY, JF, & SHARP, AL. 2015. Assessing machine-learning algorithms and image-and lidar-derived variables for GEOBIA classification of mining and mine reclamation. *International Journal of Remote Sensing*, **36**(4), 954-978.
- MCGARIGAL, KEVIN, & MARKS, BARBARA J. 1995. Spatial pattern analysis program for quantifying landscape structure. *Gen. Tech. Rep. PNW-GTR-351*. US Department of Agriculture, Forest Service, Pacific Northwest Research Station, 1-122.
- MCNEMAR, QUINN. 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, **12**(2), 153-157.
- MIDGLEY, GUY F, & BOND, WILLIAM J. 2015. Future of African terrestrial biodiversity and ecosystems under anthropogenic climate change. *Nature Climate Change*, **5**(9), 823-829.
- MILLIE, S, & ADAMS, R. 1999. Measures of Grassland Curing: a comparison of destructive sampling with visual and satellite estimates, Australian Bushfire Conference. *School of Ecology and Environment, Deakin University, Rusden, Albury*, 1-8.
- MITCHARD, EDWARD TA, & FLINTROP, CLARA M. 2013. Woody encroachment and forest degradation in sub-Saharan Africa's woodlands and savannas 1982-2006. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **368**(1625), 20120406.
- MOMENI, RAHMAN, APLIN, PAUL, & BOYD, DOREEN S. 2016. Mapping complex urban land cover from spaceborne imagery: the influence of spatial resolution, spectral band set and classification approach. *Remote Sensing*, **8**(2), 88.

- MUNYATI, C., & MAKGALE, D. 2009. Multitemporal Landsat TM imagery analysis for mapping and quantifying degraded rangeland in the Bahurutshe communal grazing lands, South Africa. *International Journal of Remote Sensing*, **30**(14), 3649-3668.
- MUNYATI, C., & SINTHUMULE, NI. 2016. Change in woody cover at representative sites in the Kruger National Park, South Africa, based on historical imagery. *SpringerPlus*, **5**(1), 1417.
- MUNYATI, C, BALZTER, H, & ECONOMON, E. 2020. Correlating Sentinel-2 MSI-derived vegetation indices with in-situ reflectance and tissue macronutrients in savannah grass. *International Journal of Remote Sensing*, **41**(10), 3820-3844.
- NAIMI, BABAK, HAMM, NICHOLAS AS, GROEN, THOMAS A, SKIDMORE, ANDREW K, & TOXOPEUS, ALBERTUS G. 2014. Where is positional uncertainty a problem for species distribution modelling? *Ecography*, **37**(2), 191-203.
- NEMENYI, PETER BJORN. 1963. *Distribution-free multiple comparisons*. Princeton University.
- NEWNHAM, GJ, GRANT, IF, MARTIN, DN, & ANDERSON, SAJ. 2010. Improved methods for assessment and prediction of grassland curing: satellite based curing methods and mapping. *Bushfire Cooperative Research Centre Project A*, **1**.
- NGADZE, FIONA, MPAKAIRI, KUDZAI SHAUN, KAVHU, BLESSING, NDAIMANI, HENRY, & MAREMBA, MONALISA SHINGIRAYI. 2020. Exploring the utility of Sentinel-2 MSI and Landsat 8 OLI in burned area mapping for a heterogenous savannah landscape. *Plos one*, **15**(5), e0232962.
- NIPPERT, JESSE B, & HOLDO, RICARDO M. 2015. Challenging the maximum rooting depth paradigm in grasslands and savannas. *Functional Ecology*, **29**(6), 739-745.
- O'CONNOR, TG. 1998. Impact of sustained drought on a semi-arid Colophospermum mopane savanna. *African Journal of Range & Forage Science*, **15**(3), 83-91.
- OLOFSSON, PONTUS, FOODY, GILES M, HEROLD, MARTIN, STEHMAN, STEPHEN V, WOODCOCK, CURTIS E, & WULDER, MICHAEL A. 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, **148**, 42-57.

- OMER, GALAL, MUTANGA, ONISIMO, ABDEL-RAHMAN, ELFATIH M, & ADAM, ELHADI. 2015. Performance of support vector machines and artificial neural network for mapping endangered tree species using WorldView-2 data in Dukuduku forest, South Africa. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **8**(10), 4825-4840.
- OSBORNE, COLIN P, CHARLES-DOMINIQUE, TRISTAN, STEVENS, NICOLA, BOND, WILLIAM J, MIDGLEY, GUY, & LEHMANN, CAROLINE ER. 2018. Human impacts in African savannas are mediated by plant functional traits. *New Phytologist*.
- OTUKEI, JOHN RICHARD, & BLASCHKE, THOMAS. 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, **12**, S27-S31.
- OTUNGA, CHARLES, ODINDI, JOHN, MUTANGA, ONISIMO, & ADJORLOLO, CLEMENT. 2019. Evaluating the potential of the red edge channel for C3 (*Festuca* spp.) grass discrimination using Sentinel-2 and Rapid Eye satellite image data. *Geocarto International*, **34**(10), 1123-1143.
- OWEN-SMITH, NORMAN. 1987. Pleistocene extinctions: the pivotal role of megaherbivores. *Paleobiology*, **13**(3), 351-362.
- OWEN-SMITH, NORMAN. 2004. Functional heterogeneity in resources within landscapes and herbivore population dynamics. *Landscape Ecology*, **19**(7), 761-771.
- OWEN-SMITH, NORMAN, & OGUTU, JOSEPH O. 2013. Controls over reproductive phenology among ungulates: allometry and tropical-temperate contrasts. *Eco-ography*, **36**(3), 256-263.
- OWEN-SMITH, R NORMAN. 1992. *Megaherbivores: the influence of very large body size on ecology*. Cambridge university press.
- PAL, MAHESH. 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, **26**(1), 217-222.
- PAL, MAHESH, & FOODY, GILES M. 2010. Feature selection for classification of hyperspectral data by SVM. *IEEE Transactions on Geoscience and Remote Sensing*, **48**(5), 2297-2307.

- PAL, MAHESH, & MATHER, PAUL M. 2003. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote sensing of environment*, **86**(4), 554-565.
- PANDEY, POOJAY, KINGTON, JOE, KANWAR, ANUDEEP, & CURDOGLO, MARINA. 2021. *Addendum to Planet Basemaps Product Specifications: NICFI Basemaps*. Tech. rept. Revision: v02. Planet Labs.
- PANEQUE-GÁLVEZ, JAIME, MAS, JEAN-FRANÇOIS, MORÉ, GERARD, CRISTÓBAL, JORDI, ORTA-MARTÍNEZ, MARTÍ, LUZ, ANA CATARINA, GUÈZE, MAXIMILIEN, MACÍA, MANUEL J, & REYES-GARCÍA, VICTORIA. 2013. Enhanced land use/cover classification of heterogeneous tropical landscapes using support vector machines and textural homogeneity. *International Journal of Applied Earth Observation and Geoinformation*, **23**, 372-383.
- PARKER, KATHERINE L, BARBOZA, PERRY S, & GILLINGHAM, MICHAEL P. 2009. Nutrition integrates environmental responses of ungulates. *Functional ecology*, **23**(1), 57-69.
- PARR, CATHERINE L, LEHMANN, CAROLINE ER, BOND, WILLIAM J, HOFFMANN, WILLIAM A, & ANDERSEN, ALAN N. 2014. Tropical grassy biomes: misunderstood, neglected, and under threat. *Trends in ecology & evolution*, **29**(4), 205-213.
- PATTERSON, PAUL L, & HEALEY, SEAN. 2015. Global ecosystem dynamics investigation (GEDI) LiDAR sampling strategy. In: *In: Stanton, Sharon M.; Christensen, Glenn A., comps. 2015. Pushing boundaries: new directions in inventory techniques and applications: Forest Inventory and Analysis (FIA) symposium 2015. 2015 December 8–10; Portland, Oregon. Gen. Tech. Rep. PNW-GTR-931. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. p. 245.*, vol. 931.
- PEDREGOSA, FABIAN, VAROQUAUX, GAËL, GRAMFORT, ALEXANDRE, MICHEL, VINCENT, THIRION, BERTRAND, GRISEL, OLIVIER, BLONDEL, MATHIEU, PRETTENHOFER, PETER, WEISS, RON, DUBOURG, VINCENT, et al. 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research*, **12**(Oct), 2825-2830.
- PENNINGTON, R TOBY, LEHMANN, CAROLINE ER, & ROWLAND, LUCY M. 2018. Tropical savannas and dry forests. *Current Biology*, **28**(9), R541-R545.

- PETERSON, DANA L, EGBERT, STEPHEN L, PRICE, KEVIN P, & MARTINKO, EDWARD A. 2004. Identifying historical and recent land-cover changes in Kansas using post-classification change detection techniques. *Transactions of the Kansas Academy of Science* (1903), 105-118.
- PINTY, B, & VERSTRAETE, MM. 1992. GEMI: a non-linear index to monitor global vegetation from satellites. *Vegetatio*, **101**(1), 15-20.
- PLANET, TEAM. 2017. Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. <https://api.planet.com>.
- PONTIUS JR, ROBERT G, SHUSAS, EMILY, & MCEACHERN, MENZIE. 2004. Detecting important categorical land changes while accounting for persistence. *Agriculture, Ecosystems & Environment*, **101**(2-3), 251-268.
- PRATT, WILLIAM K. 2013. *Introduction to Digital Image Processing*. 1st edn. Boca Ranton, USA: CRC Press.
- Pu, RUILIANG, & LANDRY, SHAWN. 2012. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment*, **124**, 516-533.
- PUSHPARAJ, JAGALINGAM, & HEGDE, ARKAL VITTAL. 2017. Evaluation of pan-sharpening methods for spatial and spectral quality. *Applied Geomatics*, **9**(1), 1-12.
- QUINLAN, J Ross. 2014. *C4. 5: programs for machine learning*. San Mateo, USA: Morgan Kaufmann Publishers, Inc.
- R CORE TEAM. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- RATAJCZAK, ZAKARY, NIPPERT, JESSE B, & COLLINS, SCOTT L. 2012. Woody encroachment decreases diversity across North American grasslands and savannas. *Ecology*, **93**(4), 697-703.
- REDFERN, JESSICA V, GRANT, RINA, BIGGS, HARRY, & GETZ, WAYNE M. 2003. Surface-water constraints on herbivore foraging in the Kruger National Park, South Africa. *Ecology*, **84**(8), 2092-2107.

- ROBERTO, COLOMBO, LORENZO, BUSETTO, MICHELE, MERONI, MICOL, ROSSINI, & CINZIA, PANIGADA. 2016. 10 Optical Remote Sensing of Vegetation Water Content. *Hyperspectral remote sensing of vegetation*, 227.
- RODRIGUEZ-GALIANO, VICTOR F, & CHICA-RIVAS, MARIO. 2014. Evaluation of different machine learning methods for land cover mapping of a Mediterranean area using multi-seasonal Landsat images and Digital Terrain Models. *International Journal of Digital Earth*, **7**(6), 492-509.
- RODRIGUEZ-GALIANO, VICTOR FRANCISCO, GHIMIRE, BARDAN, ROGAN, JOHN, CHICA-OLMO, MARIO, & RIGOL-SANCHEZ, JUAN PEDRO. 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, **67**, 93-104.
- ROERDINK, Jos BTM, & MEIJSTER, ARNOLD. 2000. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta informaticae*, **41**(1, 2), 187-228.
- ROGGO, Y, DUPONCHEL, L, & HUVENNE, J-P. 2003. Comparison of supervised pattern recognition methods with McNemar's statistical test: Application to qualitative analysis of sugar beet by near-infrared spectroscopy. *Analytica Chimica Acta*, **477**(2), 187-200.
- ROKNI, KOMEIL, AHMAD, ANUAR, SOLAIMANI, KARIM, & HAZINI, SHARIFEH. 2015. A new approach for surface water change detection: Integration of pixel level image fusion and image classification techniques. *International Journal of Applied Earth Observation and Geoinformation*, **34**, 226-234.
- SADEH, YUVAL, ZHU, XUAN, DUNKERLEY, DAVID, WALKER, JEFFREY P, ZHANG, YUXI, ROZENSTEIN, OFFER, MANIVASAGAM, VS, & CHENU, KARINE. 2021. Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. *International Journal of Applied Earth Observation and Geoinformation*, **96**, 102260.
- SAMWAYS, MICHAEL J, & KREUZINGER, KNUT. 2001. Vegetation, ungulate and grasshopper interactions inside vs. outside an African savanna game park. *Biodiversity & Conservation*, **10**(11), 1963-1981.
- SANKARAN, MAHESH. 2019. Droughts and the ecological future of tropical savanna vegetation. *Journal of Ecology*, **107**(4), 1531-1549.

- SANKARAN, MAHESH, & ANDERSON, T MICHAEL. 2009. Management and restoration in African savannas: interactions and feedbacks. *New models for ecosystem dynamics and restoration*, 136-155.
- SANKARAN, MAHESH, & RATNAM, JAYASHREE. 2013. African and Asian Savannas. Encyclopedia of Biodiversity, Volume 1.
- SANKARAN, MAHESH, HANAN, NIALL P, SCHOLES, ROBERT J, RATNAM, JAYASHREE, AUGUSTINE, DAVID J, CADE, BRIAN S, GIGNOUX, JACQUES, HIGGINS, STEVEN I, LE ROUX, XAVIER, LUDWIG, FULCO, *et al.* 2005. Determinants of woody cover in African savannas. *Nature*, **438**(7069), 846.
- SANKARAN, MAHESH, RATNAM, JAYASHREE, & HANAN, NIALL. 2008. Woody cover in African savannas: the role of resources, fire and herbivory. *Global Ecology and Biogeography*, **17**(2), 236-245.
- SARP, GULCAN. 2014. Spectral and spatial quality analysis of pan-sharpening algorithms: A case study in Istanbul. *European Journal of Remote Sensing*, **47**(1), 19-28.
- SCHULP, CATHARINA JE, BURKHARD, BENJAMIN, MAES, JOACHIM, VAN VLIET, JASPER, & VERBURG, PETER H. 2014. Uncertainties in ecosystem service maps: a comparison on the European scale. *PLoS one*, **9**(10).
- SCHUSTER, CHRISTIAN, SCHMIDT, TOBIAS, CONRAD, CHRISTOPHER, KLEINSCHMIT, BIRGIT, & FÖRSTER, MICHAEL. 2015. Grassland habitat mapping by intra-annual time series analysis-Comparison of RapidEye and TerraSAR-X satellite data. *International Journal of Applied Earth Observation and Geoinformation*, **34**, 25-34.
- SHI, DI, & YANG, XIAOJUN. 2016. An assessment of algorithmic parameters affecting image classification accuracy by random forests. *Photogrammetric Engineering & Remote Sensing*, **82**(6), 407-417.
- SHORROCKS, BRYAN, & BATES, WILLIAM. 2015. *The biology of African savannahs*. Oxford University Press, USA.
- SIDDQUI, YUSUF. 2003. The modified IHS method for fusing satellite imagery. Pages 5–9 of: *ASPRS 2003 Annual Conference Proceedings*. Anchorage, Alaska.

- SIEBERT, FRANCES, & ECKHARDT, HOLGER C. 2008. The vegetation and floristics of the Nkhuhlu exclosures, Kruger National Park. *Koedoe*, **50**(1), 126-144.
- SILVA, LEONARDO PEREIRA E, XAVIER, ANA PAULA CAMPOS, DA SILVA, RICHARDE MARQUES, & SANTOS, CELSO AUGUSTO GUIMARÃES. 2020. Modeling land cover change based on an artificial neural network for a semiarid river basin in northeastern Brazil. *Glob. Ecol. Conserv.*, **21**(e00811), e00811.
- SINGH, BIKESH KUMAR, VERMA, KESARI, & THOKE, AS. 2015. Investigations on impact of feature normalization techniques on classifier's performance in breast tumor classification. *International Journal of Computer Applications*, **116**(19).
- SINGH, DALWINDER, & SINGH, BIRMOHAN. 2019. Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 105524. <https://doi.org/10.1016/j.asoc.2019.105524>.
- SMIT, IZAK PJ. 2011. Resources driving landscape-scale distribution patterns of grazers in an African savanna. *Ecography*, **34**(1), 67-74.
- SMIT, IZAK PJ, & ARCHIBALD, SALLY. 2019. Herbivore culling influences spatio-temporal patterns of fire in a semiarid savanna. *Journal of applied ecology*, **56**(3), 711-721.
- SMIT, IZAK PJ, GRANT, CORNELIA C, & DEVEREUX, BERNARD J. 2007. Do artificial waterholes influence the way herbivores use the landscape? Herbivore distribution patterns around rivers and artificial surface water sources in a large African savanna park. *Biological Conservation*, **136**(1), 85-99.
- SMITH, ANNE M, & BUCKLEY, JOSEPH R. 2011. Investigating RADARSAT-2 as a tool for monitoring grassland in western Canada. *Canadian Journal of Remote Sensing*, **37**(1), 93-102.
- SOUTH AFRICAN NATIONAL PARKS. 2021 (Dec). *Kruger climate and Rainfall Data and Information Resources – Scientific Services*. [Online; accessed 11-January-2022].
- SRIVASTAVA, NITISH, HINTON, GEOFFREY, KRIZHEVSKY, ALEX, SUTSKEVER, ILYA, & SALAKHUTDINOV, RUSLAN. 2014. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, **15**(1), 1929-1958.

- STAVER, A CARLA, BOND, WILLIAM J, CRAMER, MICHAEL D, & WAKELING, JULIA L. 2012. Top-down determinants of niche structure and adaptation among African Acacias. *Ecology letters*, **15**(7), 673-679.
- STAVER, A CARLA, WIGLEY-COETSEE, CORLI, & BOTHA, JUDITH. 2019. Grazer movements exacerbate grass declines during drought in an African savanna. *Journal of Ecology*, **107**(3), 1482-1491.
- STEHMAN, STEPHEN V, & OVERTON, W SCOTT. 1994. 9 Environmental sampling and monitoring. *Handbook of statistics*, **12**, 263-306.
- STEVENS, NICOLA, LEHMANN, CAROLINE ER, MURPHY, BRETT P, & DURIGAN, GISELDA. 2017. Savanna woody encroachment is widespread across three continents. *Global change biology*, **23**(1), 235-244.
- SUN, JINHUA, YANG, LIU, YANG, XITIAN, WEI, JIE, LI, LANTAO, GUO, ERHUI, & KONG, YUHUA. 2021. Using spectral reflectance to estimate the leaf chlorophyll content of maize inoculated with arbuscular mycorrhizal fungi under water stress. *Frontiers in Plant Science*, **12**, 1053.
- SUWANPRASIT, CHANIDA, & SRICHAI, NAIYANA. 2012. Impacts of spatial resolution on land cover classification. *Proceedings of the Asia-Pacific Advanced Network*, **33**(0), 39.
- SWEMMER, ANTHONY M, BOND, WILLIAM J, DONALDSON, JASON, HEMPSON, GARETH P, MALHERBE, JOHAN, & SMIT, IZAK PJ. 2018. The ecology of drought-a workshop report. *South African Journal of Science*, **114**(9-10), 1-3.
- SYMEONAKIS, ELIAS, HIGGINBOTTOM, THOMAS P, PETROULAKI, KYRIAKI, & RABE, ANDREAS. 2018. Optimisation of savannah land cover characterisation with optical and SAR data. *Remote Sensing*, **10**(4), 499.
- SYMEONAKIS, ELIAS, VERÓN, SANTIAGO, BALDI, GERMÁN, BANCHERO, SANTIAGO, DE ABELLEYRA, DIEGO, & CASTELLANOS, GEORGE. 2019. Savannah land cover characterisation: a quality assessment using Sentinel 1/2, Landsat, PALSAR and PlanetScope.
- TEWKESBURY, ANDREW P, COMBER, ALEXIS J, TATE, NICHOLAS J, LAMB, ALISTAIR, & FISHER, PETER F. 2015. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, **160**, 1-14.

- UNDERWOOD, EMMA C, USTIN, SUSAN L, & RAMIREZ, CARLOS M. 2007. A comparison of spatial and spectral image resolution for mapping invasive plants in coastal California. *Environmental management*, **39**(1), 63-83.
- VAJSOVA, BLANKA, WALCZYNSKA, AGNIESZKA, BÄRISCH, SAMUEL, ÅSTRAND, PÄR JOHAN, & HAIN, SUSANNE. 2017. *New sensors benchmark report on WorldView-4: Geometric benchmarking over Maussane test site for CAP purposes*. Available at <https://core.ac.uk/download/pdf/93512541.pdf>.
- VAN DER SANDE, CJ, DE JONG, SM, & DE Roo, APJ. 2003. A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment. *International Journal of applied earth observation and geoinformation*, **4**(3), 217-229.
- VAN LANGEVELDE, FRANK, VAN DE VIJVER, CLAUDIOUS ADM, KUMAR, LALIT, VAN DE KOPPEL, JOHAN, DE RIDDER, NICO, VAN ANDEL, JELTE, SKIDMORE, ANDREW K, HEARNE, JOHN W, STROOSNIJDER, LEO, BOND, WILLIAM J, et al. 2003. Effects of fire and herbivory on the stability of savanna ecosystems. *Ecology*, **84**(2), 337-350.
- VAN NIEL, THOMAS G, McVICAR, TIM R, & DATT, BISUN. 2005. On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification. *Remote Sensing of Environment*, **98**(4), 468-480.
- VAN WILGEN, BRIAN W, GOVENDER, NAVASHNI, SMIT, IZAK PJ, & MACFADYEN, SANDRA. 2014. The ongoing development of a pragmatic and adaptive fire management policy in a large African savanna protected area. *Journal of environmental management*, **132**, 358-368.
- VANDENBYGAART, AJ. 2006. Monitoring soil organic carbon stock changes in agricultural landscapes: issues and a proposed approach. *Canadian Journal of Soil Science*, **86**(3), 451-463.
- VELDHUIS, MICHELI P, FAKKERT, HELEEN F, BERG, MATTY P, & OLFF, HAN. 2016. Grassland structural heterogeneity in a savanna is driven more by productivity differences than by consumption differences between lawn and bunch grasses. *Oecologia*, **182**(3), 841-853.

- VENKATESH, YV, & RAJA, S KUMAR. 2003. On the classification of multispectral satellite images using the multilayer perceptron. *Pattern Recognition*, **36**(9), 2161-2175.
- VENTER, FREDERIK JOHANNES. 1990. A classification of land for management planning in the Kruger National Park.
- VENTER, FREDERIK JOHANNES. 1991. *A classification of land for management planning in the Kruger National Park*. Ph.D. thesis, University of South Africa.
- VENTER, FREEK J, SCHOLES, ROBERT J, & ECKHARDT, HOLGER C. 2003. The abiotic template and its associated vegetation pattern. *The Kruger experience: Ecology and management of savanna heterogeneity*, **83**, 129.
- VESCOVO, LORIS, & GIANELLE, DAMIANO. 2008. Using the MIR bands in vegetation indices for the estimation of grassland biophysical parameters from satellite remote sensing in the Alps region of Trentino (Italy). *Advances in Space Research*, **41**(11), 1764-1772.
- VINCENT, O REBECCA, FOLORUNSO, OLUSEGUN, et al. 2009. A descriptive algorithm for sobel image edge detection. Pages 97–107 of: *Proceedings of informing science & IT education conference (InSITE)*, vol. 40. Informing Science Institute California.
- VOYSEY, MICHAEL D, ARCHIBALD, SALLY, BOND, WILLIAM J, DONALDSON, JASON E, CARLA STAVER, A, & GREVE, MICHELLE. 2021. The role of browsers in maintaining the openness of savanna grazing lawns. *Journal of Ecology*, **109**(2), 913-926.
- WAINER, JACQUES, & CAWLEY, GAVIN. 2018. Nested cross-validation when selecting classifiers is overzealous for most practical applications. *arXiv preprint arXiv:1809.09446*.
- WALDRAM, MATTHEW S, BOND, WILLIAM J, & STOCK, WILLIAM D. 2008. Ecological engineering by a mega-grazer: white rhino impacts on a South African savanna. *Ecosystems*, **11**(1), 101-112.
- WALKER, BRIAN H, EMSLIE, RH, OWEN-SMITH, RN, & SCHOLES, RJ. 1987. To cull or not to cull: lessons from a southern African drought. *Journal of Applied Ecology*, 381-401.

- WANG, XIANGYING, & ZHONG, YIXIN. 2003. Statistical learning theory and state of the art in SVM. *Pages 55–59 of: The Second IEEE International Conference on Cognitive Informatics, 2003. Proceedings*. IEEE.
- WANG, ZHOU, BOVIK, ALAN C, SHEIKH, HAMID R, & SIMONCELLI, EERO P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, **13**(4), 600-612.
- WARUI, CHARLES M, VILLET, MARTIN H, YOUNG, TRUMAN P, & JOCQUÉ, RUDY. 2005. Influence of grazing by large mammals on the spider community of a Kenyan savanna biome. *Journal of Arachnology*, 269-279.
- WHITESIDE, TIMOTHY G, BOGGS, GUY S, & MAIER, STEFAN W. 2011. Comparing object-based and pixel-based classifications for mapping savannas. *International Journal of Applied Earth Observation and Geoinformation*, **13**(6), 884-893.
- WOODCOCK, CURTIS E, LOVELAND, THOMAS R, HEROLD, MARTIN, & BAUER, MARVIN E. 2020. Transitioning from change detection to monitoring with remote sensing: A paradigm shift. *Remote Sensing of Environment*, **238**, 111558.
- WULDER, MICHAEL A, MASEK, JEFFREY G, COHEN, WARREN B, LOVELAND, THOMAS R, & WOODCOCK, CURTIS E. 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, **122**, 2-10.
- XIE, YICHUN, SHA, ZONGYAO, & YU, MEI. 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of plant ecology*, **1**(1), 9-23.
- XIE, ZHULI, CHEN, YAOLIANG, LU, DENGSHENG, LI, GUIYING, & CHEN, ERXUE. 2019. Classification of land cover, forest, and tree species classes with ZiYuan-3 multispectral and stereo data. *Remote Sensing*, **11**(2), 164.
- Xu, B, YANG, XC, TAO, WG, QIN, ZH, LIU, HQ, MIAO, JM, & BI, YY. 2008. MODIS-based remote sensing monitoring of grass production in China. *International Journal of Remote Sensing*, **29**(17-18), 5313-5327.
- Xu, BIN, YANG, XIUCHUN, TAO, WEIGUO, QIN, ZHIHAO, LIU, HAIQI, & MIAO, JIANMING. 2007. Remote sensing monitoring upon the grass production in China. *Acta Ecologica Sinica*, **27**(2), 405-413.

- XU, DAEWEI, CHEN, BAORUI, SHEN, BEIBEI, WANG, XU, YAN, YUCHUN, XU, LIJUN, & XIN, XIAOPING. 2018. The Classification of Grassland Types Based on Object-Based Image Analysis with Multisource Data. *Rangeland Ecology & Management*.
- XU, DAEWEI, CHEN, BAORUI, SHEN, BEIBEI, WANG, XU, YAN, YUCHUN, XU, LIJUN, & XIN, XIAOPING. 2019. The classification of grassland types based on object-based image analysis with multisource data. *Rangeland Ecology & Management*, **72**(2), 318-326.
- YIN, XIAOFEI, WANG, CHUNGUANG, ZONG, ZHEYING, WANG, HAICHAO, ZHANG, HAIJUN, & ZHANG, WENXIA. 2018. Biomass estimation of desert steppe based on spectral indices along a precipitation gradient. *Spectroscopy Letters*, **51**(7), 324-331.
- YOCKY, DAVID A. 1995. Image merging and data fusion by means of the discrete two-dimensional wavelet transform. *JOSA A*, **12**(9), 1834-1841.
- YU, XIAOZHI, LU, DENGSHENG, JIANG, XIANDIE, LI, GUIYING, CHEN, YAOLIANG, LI, DENGQIU, & CHEN, ERXUE. 2020. Examining the Roles of Spectral, Spatial, and Topographic Features in Improving Land-Cover and Forest Classifications in a Subtropical Region. *Remote Sensing*, **12**(18), 2907.
- ZHAO, LICHENG, SHI, YUN, LIU, BIN, HOVIS, CIARA, DUAN, YULIN, & SHI, ZHONGCHAO. 2019. Finer Classification of Crops by Fusing UAV Images and Sentinel-2A Data. *Remote Sensing*, **11**(24), 3012.
- ZHOU, YONG, TINGLEY, MORGAN W, CASE, MADELON F, COETSEE, CORLI, KIKER, GREGORY A, SCHOLTZ, RHEINHARDT, VENTER, FREEK J, & STAVER, A CARLA. 2021. Woody encroachment happens via intensification, not extensification, of species ranges in an African savanna. *Ecological Applications*, e02437.
- ZHU, ZHE, & WOODCOCK, CURTIS E. 2014. Continuous change detection and classification of land cover using all available Landsat data. *Remote sensing of Environment*, **144**, 152-171.
- ZIZKA, ALEXANDER, GOVENDER, NAVASHNI, & HIGGINS, STEVEN I. 2014. How to tell a shrub from a tree: A life-history perspective from a South African savanna. *Austral Ecology*, **39**(7), 767-778.

---

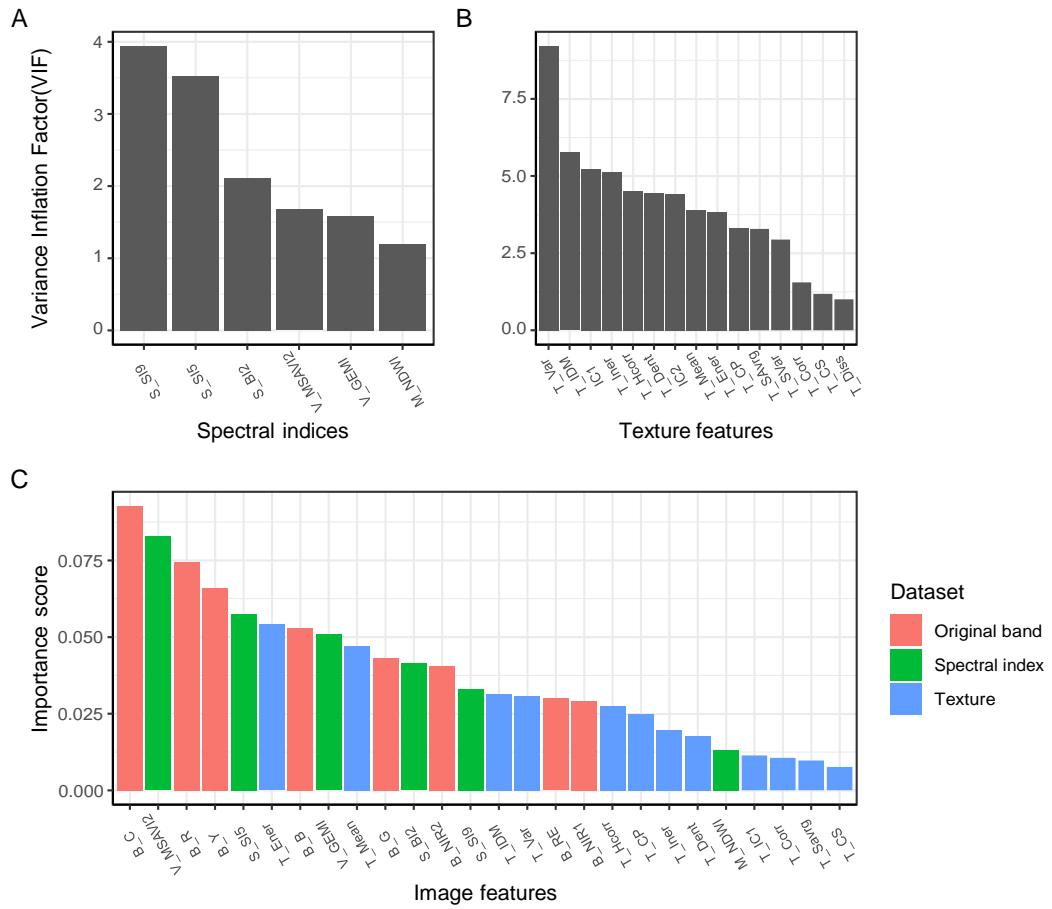
## Appendix A

# Probabilistic Mapping and Spatial Pattern Analysis of Grazing Lawns in Southern African Savannahs Using WorldView-3 Imagery and Machine Learning Techniques

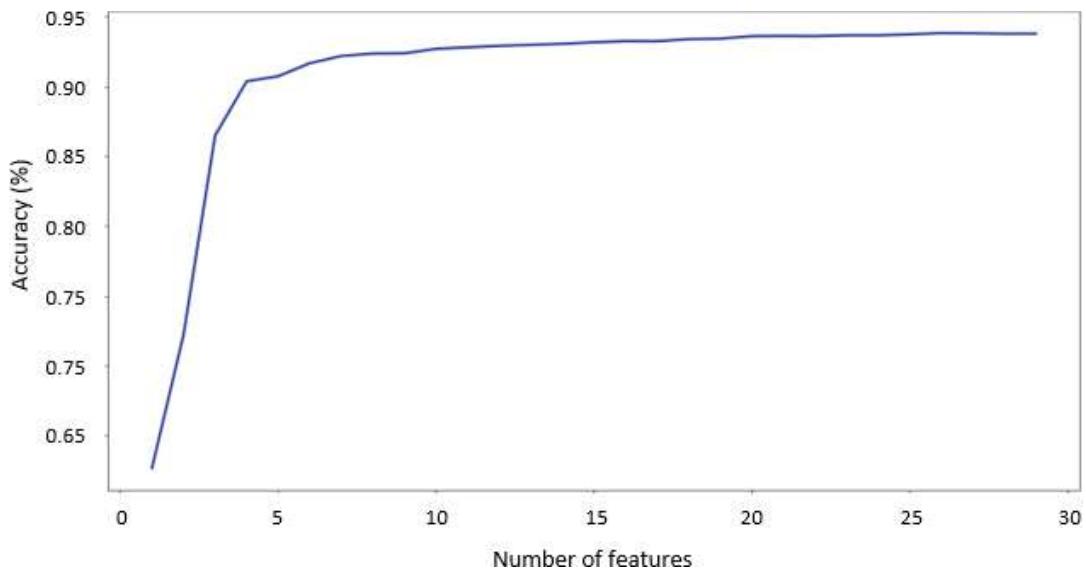
---

### A.1 Multicollinearity and Feature Selection

Multicollinearity analysis among derived image features showed that spectral indices exhibited higher correlation than texture features based on a stepwise elimination threshold of  $VIF \geq 10$ . Twenty-one out of the 27 spectral indices (i.e., vegetation, moisture and soil combined) had collinearity problems. After eliminating the collinear variables, VIF values of retained spectral indices ranged from 1.19 to 3.94 (Figure A.1A) with linear correlation coefficients between  $-0.002$  ( $S\_SI9 \sim M\_NDWI$ ) and  $-0.666$  ( $S\_SI9 \sim S\_SI5$ ). In contrast, only three out of the 18 texture features exhibited collinearity. The retained texture variables had VIF values ranging from 1.00 to 9.22 (Figure A.1B) and linear correlation coefficients ranging between  $-6.14 \times 10^{-5}$  ( $T\_Dent \sim T\_Diss$ ) and 0.79 ( $T\_IDM \sim T\_Ener$ ). Overall, six spectral indices (three soil indices, two vegetation indices and one moisture index) and 15 texture features (seven simple and eight advanced Haralick features) were retained (Figure A.1A and A.1B) and combined with the eight original image bands for final feature selection.



**Figure A.1:** Feature selection results from VIF and RF-RFE analysis. (A) VIF of selected spectral indices. (B) VIF of selected texture features. (C) Importance scores of final selected image features following RF-RFE. Original bands and selected spectral indices and texture features from VIF served as input to RF-RFE. Final selection was based on number of features that retained optimal accuracy



**Figure A.2:** Plot of accuracy versus number of features.

Selection of final input image features was conducted using Random Forest Feature Elimination (RF-RFE). The RF-RFE procedure resulted in 26 image features (Figure A.2) comprising of eight image bands, six spectral indices and twelve texture features. Figure A.1C shows the relative importance scores of selected final input features in differentiating the different land cover categories (Table 2.1). The first 13 (50%) most important features were dominated by nearly equal proportions of spectral bands and spectral indices (six and five, respectively), with two texture features, while the remaining features were largely composed of texture variables (Figure A.1C). Amongst the original bands, variables that exhibited high importance were B\_C, B\_R and B\_Y. V\_MSAVI2, S\_SI5 and V\_GEMI were the spectral indices of high importance, and the most influential texture variables included T\_Ener and T\_Mean.

## A.2 Comparison of Grazing Lawn Area Estimates across Models in Each Landscape.

**Table A.1:** Two-proportions Z-test comparing the proportions of estimated grazing lawn cover. Values in parenthesis represent p-value. Model pairs that show statistically significant difference ( $p < 0.05$ ) in proportion of grazing lawn cover are in bold. CART = Classification and Regression Trees, MLP = Multilayer Perceptron, RF = Random Forest, SVM = Support Vector Machines.

Lower Sabie		Satara	
Model pair	$\chi^2$ -test	Model pair	$\chi^2$ -test
CART v MLP	0.000(1.00)	CART v MLP	<b>5.017(0.025)</b>
CART v RF	0.000(1.00)	CART v RF	<b>8.328(0.003)</b>
CART v SVM	0.000(1.00)	CART v SVM	<b>7.225(0.007)</b>
MLP v RF	0.000(1.00)	MLP v RF	<b>13.146(0.000)</b>
MLP v SVM	0.000(1.00)	MLP v SVM	<b>11.657(0.000)</b>
RF v SVM	0.000(1.00)	RF v SVM	<b>10.083(0.001)</b>

## A.3 Confusion Matrices for the Lower Sabie Landscape

**Table A.2:** Confusion matrix summarising results from Random Forest (RF) model classification of grazing lawn and other cover.

		RF		Commission Error (%)	
Predicted Class	Grazing lawn	Reference Class			
		Grazing lawn	Other		
Grazing lawn	84	16	16.00		
Other	4	693	0.57		
Omission Error (%)	4.55	2.26			

**Table A.3:** Confusion matrix summarising results from Support Vector Machines (SVM) model classification of grazing lawn and other cover.

		SVM		Commission Error (%)	
Predicted Class	Grazing lawn	Reference Class			
		Grazing lawn	Other		
Grazing lawn	93	7	7.00		
Other	7	690	1.00		
Omission Error (%)	7.00	1.00			

**Table A.4:** Confusion matrix summarising results from Classification and Regression Trees (CART) model classification of grazing lawn and other cover.

CART			
	Reference Class		Commission Error (%)
	Grazing lawn	Other	
Predicted Class	Grazing lawn	77	23
	Other	12	685
Omission Error (%)		13.48	3.25

**Table A.5:** Confusion matrix summarising results from Multilayer Perceptron (MLP) model classification of grazing lawn and other cover.

MLP			
	Reference Class		Commission Error (%)
	Grazing lawn	Other	
Predicted Class	Grazing lawn	97	3
	Other	9	688
Omission Error (%)		8.49	0.43

## A.4 Confusion Matrices for the Satara Landscape

**Table A.6:** Confusion matrix summarising results from Random Forest (RF) model classification of grazing lawn and other cover.

RF			
	Reference Class		Commission Error (%)
	Grazing lawn	Other	
Predicted Class	Grazing lawn	90	13
	Other	12	511
Omission Error (%)		11.76	2.48

**Table A.7:** Confusion matrix summarising results from Support Vector Machines (SVM) model classification of grazing lawn and other cover.

SVM			
	Reference Class		Commission Error (%)
	Grazing lawn	Other	
Predicted Class	Grazing lawn	93	10
	Other	14	509
Omission Error (%)		13.08	1.93

**Table A.8:** Confusion matrix summarising results from Classification and Regression Trees (CART) model classification of grazing lawn and other cover.

<b>CART</b>			
Predicted Class	Reference Class		Commission Error (%)
	Grazing lawn	Other	
Grazing lawn	78	25	24.27
Other	25	498	4.78
Omission Error (%)	24.27	4.78	

**Table A.9:** Confusion matrix summarising results from Multilayer Perceptron (MLP) model classification of grazing lawn and other cover.

<b>MLP</b>			
Predicted Class	Reference Class		Commission Error (%)
	Grazing lawn	Other	
Grazing lawn	88	15	14.56
Other	16	507	3.06
Omission Error (%)	15.38	2.87	



---

## Appendix B

# Spatio-temporal Analysis of Grazing Lawn Dynamics in Southern African Savannahs Using Multi-temporal High-resolution Satellite Images.

---

### B.1 Confusion Matrices used for accuracy assessment

**Table B.1:** Site 1 confusion matrix used for classification accuracy assessment for each time period (2002, 2014 and 2019). Columns represent classification labels and rows represent reference labels.

	Woody	Bunch grass	Grazing lawn	Water body	Bare	Built-up
<b>2002</b>						
Woody	449	12	1	1	0	5
Bunch grass	8	309	8	0	0	0
Grazing lawn	0	3	367	0	1	0
Water body	2	0	0	381	0	5
Bare	0	0	0	0	233	1
Built-up	2	0	2	15	2	167
<b>2014</b>						
Woody	136	4	0	0	0	3
Bunch grass	0	45	2	0	0	0
Grazing lawn	0	1	317	0	1	2
Water body	0	0	0	139	0	0
Bare	0	0	2	0	202	0
Built-up	0	1	7	0	1	37
<b>2019</b>						
Woody	47	2	0	1	2	2
Bunch grass	2	29	2	2	0	0
Grazing lawn	0	0	115	1	1	3
Water body	2	1	2	37	1	0
Bare	1	0	3	2	24	1
Built-up	0	1	4	0	1	26

**Table B.2:** Site 2 confusion matrix used for classification accuracy assessment for each time period (2002, 2014 and 2019). Columns represent classification labels and rows represent reference labels.

	Woody	Bunch grass	Grazing lawn	Water body	Bare	Built-up
<b>2002</b>						
Woody	936	11	0	0	0	3
Bunch grass	8	468	8	0	0	0
Grazing lawn	0	2	217	0	2	0
Water body	0	0	0	375	0	0
Bare	0	0	4	0	250	0
Built-up	0	5	3	0	0	683
<b>2014</b>						
Woody	864	11	0	0	0	4
Bunch grass	3	199	1	0	0	1
Grazing lawn	0	6	545	0	1	4
Water body	2	0	0	1105	0	1
Bare	0	0	0	0	389	1
Built-up	1	3	4	2	3	1807
<b>2019</b>						
Woody	1096	26	1	0	0	2
Bunch grass	15	648	62	0	0	0
Grazing lawn	1	52	776	0	10	0
Water body	2	0	0	1088	4	4
Bare	0	0	8	0	716	5
Built-up	0	0	3	4	3	1499

**Table B.3:** Site 3 confusion matrix used for classification accuracy assessment for each time period (2002, 2014 and 2019). Columns represent classification labels and rows represent reference labels.

	Woody	Bunch grass	Grazing lawn	Water body	Bare	Built-up
<b>2002</b>						
Woody	266	7	1	0	0	0
Bunch grass	1	188	0	0	0	0
Grazing lawn	0	0	224	0	0	2
Water body	0	0	0	332	0	1
Bare	0	0	0	0	56	0
Built-up	0	0	5	0	0	8
<b>2014</b>						
Woody	304	6	0	0	0	1
Bunch grass	1	315	1	0	0	1
Grazing lawn	0	2	293	0	2	0
Water body	0	0	0	69	0	0
Bare	0	0	0	0	127	0
Built-up	2	3	0	5	0	58
<b>2019</b>						
Woody	325	3	5	1	0	1
Bunch grass	6	183	10	0	0	0
Grazing lawn	1	11	661	0	4	0
Water body	3	2	1	249	0	0
Bare	0	0	0	0	169	0
Built-up	3	0	1	3	0	164

## B.2 Transition matrices for land cover change analysis.

**Table B.4:** Site 1 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2014. Columns represent 2014 entries and rows represent 2002 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	10.33	10.94	1.48	0.61	0.56	1.35	25.27
Bunch grass	7.06	20.53	4.34	1.04	0.78	2.49	36.24
Grazing lawn	0.33	1.03	1.97	0.26	0.22	0.57	4.37
Water	0.14	0.26	0.16	0.06	0.03	0.05	0.69
Bare	0.02	0.04	0.09	0.03	0.03	0.02	0.23
Built-up	0.01	0.02	0.02	0.01	0.01	0.01	0.08
Totals	17.89	32.82	8.06	2.01	1.63	4.48	66.89

**Table B.5:** Site 1 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2014 and 2019. Columns represent 2019 entries and rows represent 2014 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	10.95	7.00	0.46	0.08	0.06	0.02	18.57
Bunch grass	8.04	23.38	2.48	0.07	0.08	0.03	34.07
Grazing lawn	0.76	3.21	4.41	0.06	0.11	0.00	8.55
Water	0.44	1.17	0.50	0.04	0.02	0.00	2.18
Bare	0.50	0.92	0.32	0.01	0.03	0.00	1.79
Built-up	1.19	2.99	0.77	0.01	0.00	0.00	4.97
Totals	21.87	38.68	8.94	0.27	0.31	0.05	70.13

**Table B.6:** Site 1 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2019. Columns represent 2019 entries and rows represent 2002 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	13.93	16.18	0.92	0.09	0.07	0.02	31.22
Bunch grass	12.53	38.77	6.81	0.08	0.08	0.03	58.30
Grazing lawn	0.52	3.16	3.54	0.04	0.05	0.00	7.31
Water	0.27	0.43	0.19	0.03	0.03	0.00	0.95
Bare	0.04	0.08	0.09	0.01	0.06	0.00	0.29
Built-up	0.02	0.05	0.04	0.00	0.00	0.00	0.12
Totals	27.31	58.67	11.60	0.25	0.30	0.06	98.18

**Table B.7:** Site 2 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2014. Columns represent 2014 entries and rows represent 2002 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	15.73	6.62	0.68	0.02	0.05	0.25	23.35
Bunch grass	7.58	6.88	1.27	0.02	0.06	0.28	16.09
Grazing lawn	0.24	0.25	0.44	0.00	0.03	0.09	1.05
Water	0.25	0.22	0.05	0.04	0.00	0.02	0.58
Bare	0.01	0.01	0.01	0.00	0.01	0.01	0.06
Built-up	0.13	0.10	0.02	0.00	0.00	0.03	0.28
Totals	23.93	14.08	2.47	0.08	0.17	0.68	41.41

**Table B.8:** Site 2 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2014 and 2019. Columns represent 2019 entries and rows represent 2014 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	20.63	10.11	0.59	0.01	0.43	0.06	31.83
Bunch grass	7.42	11.39	0.84	0.00	0.51	0.04	20.20
Grazing lawn	1.02	1.18	0.80	0.00	0.33	0.04	3.37
Water	0.02	0.01	0.00	0.05	0.00	0.00	0.08
Bare	0.07	0.06	0.01	0.00	0.06	0.00	0.21
Built-up	0.37	0.25	0.11	0.00	0.10	0.04	0.87
Totals	29.54	23.01	2.35	0.06	1.43	0.19	56.56

**Table B.9:** Site 2 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2019. Columns represent 2019 entries and rows represent 2002 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	15.04	7.97	0.40	0.01	0.34	0.04	23.80
Bunch grass	6.77	7.95	0.94	0.01	0.53	0.06	16.26
Grazing lawn	0.30	0.31	0.26	0.00	0.19	0.02	1.08
Water	0.24	0.27	0.02	0.04	0.02	0.00	0.59
Bare	0.02	0.02	0.01	0.00	0.05	0.00	0.09
Built-up	0.13	0.10	0.01	0.00	0.01	0.03	0.29
Totals	22.51	16.62	1.64	0.06	1.14	0.16	42.12

**Table B.10:** Site 3 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2014. Columns represent 2014 entries and rows represent 2002 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	8.84	11.02	0.87	0.09	0.09	0.20	21.12
Bunch grass	13.53	48.09	2.95	0.07	0.17	0.40	65.20
Grazing lawn	0.47	1.35	1.03	0.02	0.13	0.05	3.05
Water	0.25	0.12	0.22	0.26	0.11	0.08	1.04
Bare	0.06	0.05	0.17	0.01	0.15	0.02	0.45
Built-up	0.10	0.09	0.15	0.03	0.07	0.03	0.46
Totals	23.24	60.72	5.39	0.49	0.71	0.77	91.32

**Table B.11:** Site 3 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2014 and 2019. Columns represent 2019 entries and rows represent 2014 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	13.25	7.45	1.83	0.57	0.10	0.08	23.28
Bunch grass	15.75	35.00	8.46	1.20	0.23	0.16	60.79
Grazing lawn	1.54	1.03	2.36	0.11	0.27	0.09	5.40
Water	0.19	0.02	0.02	0.24	0.01	0.01	0.49
Bare	0.27	0.07	0.18	0.02	0.16	0.01	0.71
Built-up	0.37	0.20	0.11	0.05	0.02	0.01	0.77
Totals	31.38	43.76	12.96	2.19	0.79	0.37	91.45

**Table B.12:** Site 3 transition matrix ( $\text{km}^2$ ) used for land cover change analysis for the period between 2002 and 2019. Columns represent 2019 entries and rows represent 2002 entries.

Class	Woody	Bunch grass	Grazing lawn	Water	Bare	Built-up	Totals
Woody	11.56	7.91	2.34	0.58	0.13	0.09	22.61
Bunch grass	19.18	37.24	9.79	1.34	0.34	0.20	68.09
Grazing lawn	0.90	0.73	1.20	0.06	0.16	0.03	3.07
Water	0.61	0.04	0.07	0.26	0.04	0.02	1.04
Bare	0.17	0.04	0.12	0.02	0.09	0.01	0.45
Built-up	0.25	0.05	0.08	0.03	0.04	0.02	0.46
Totals	32.65	46.01	13.60	2.28	0.80	0.37	95.72

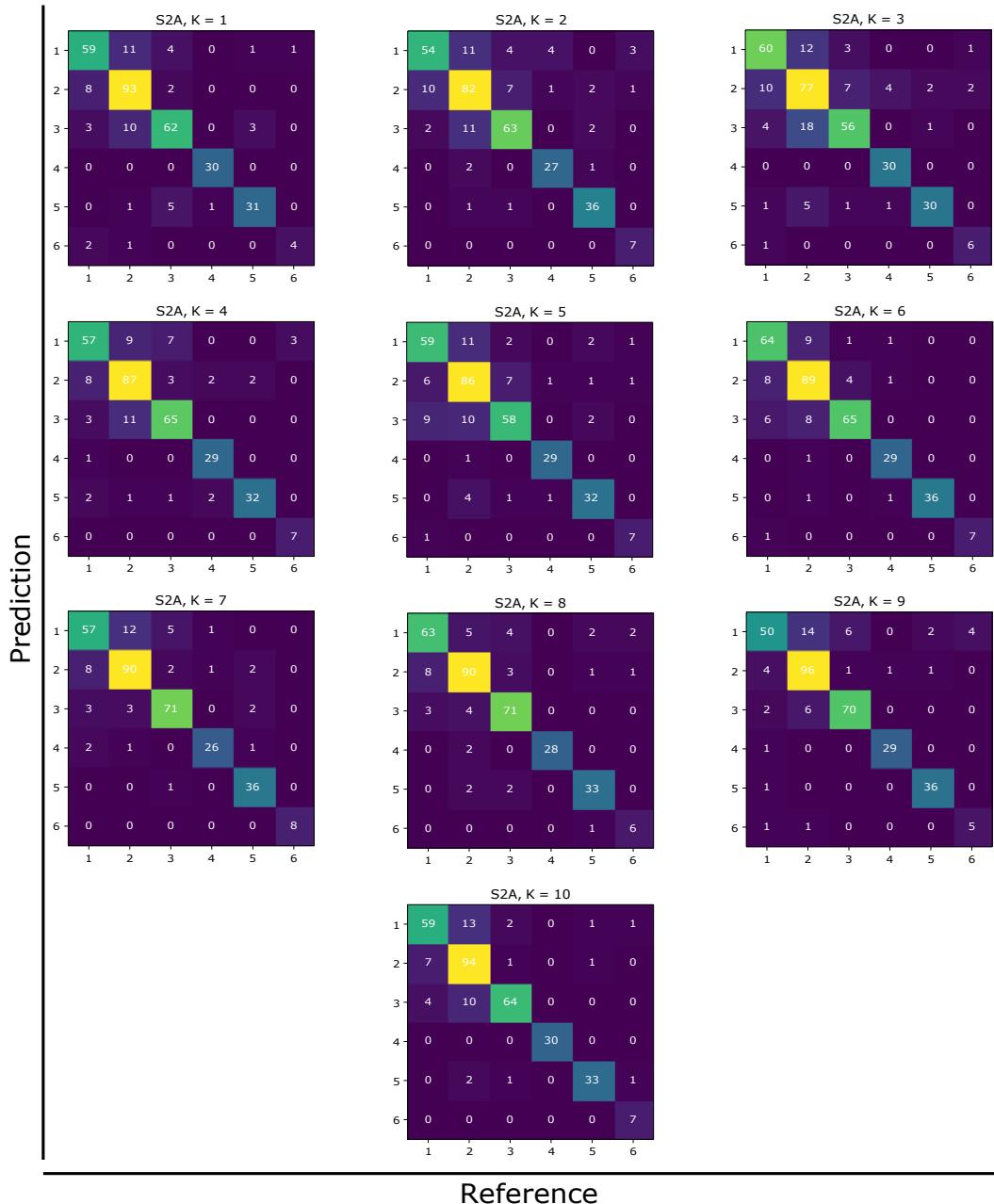
---

---

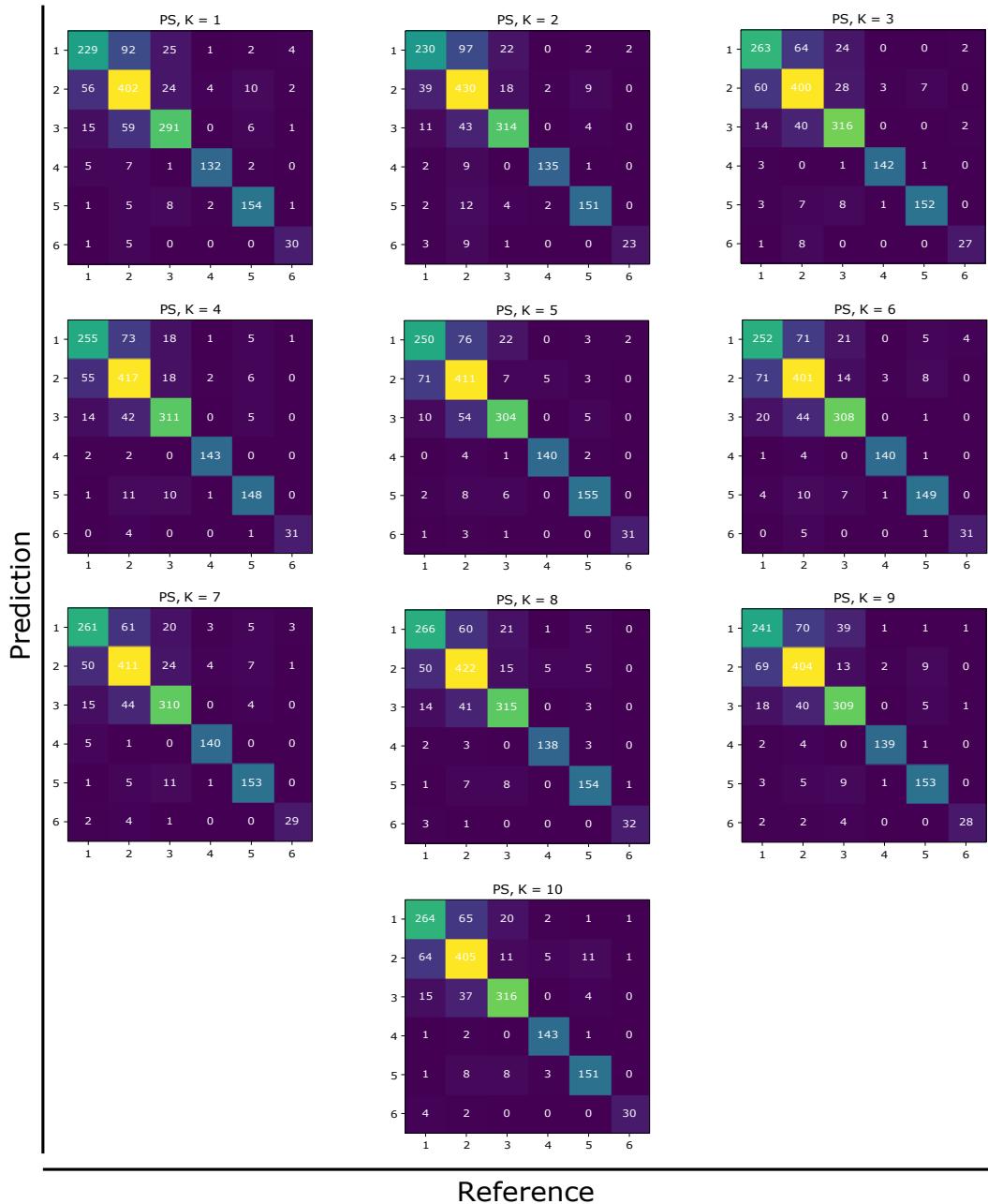
## Appendix C

---

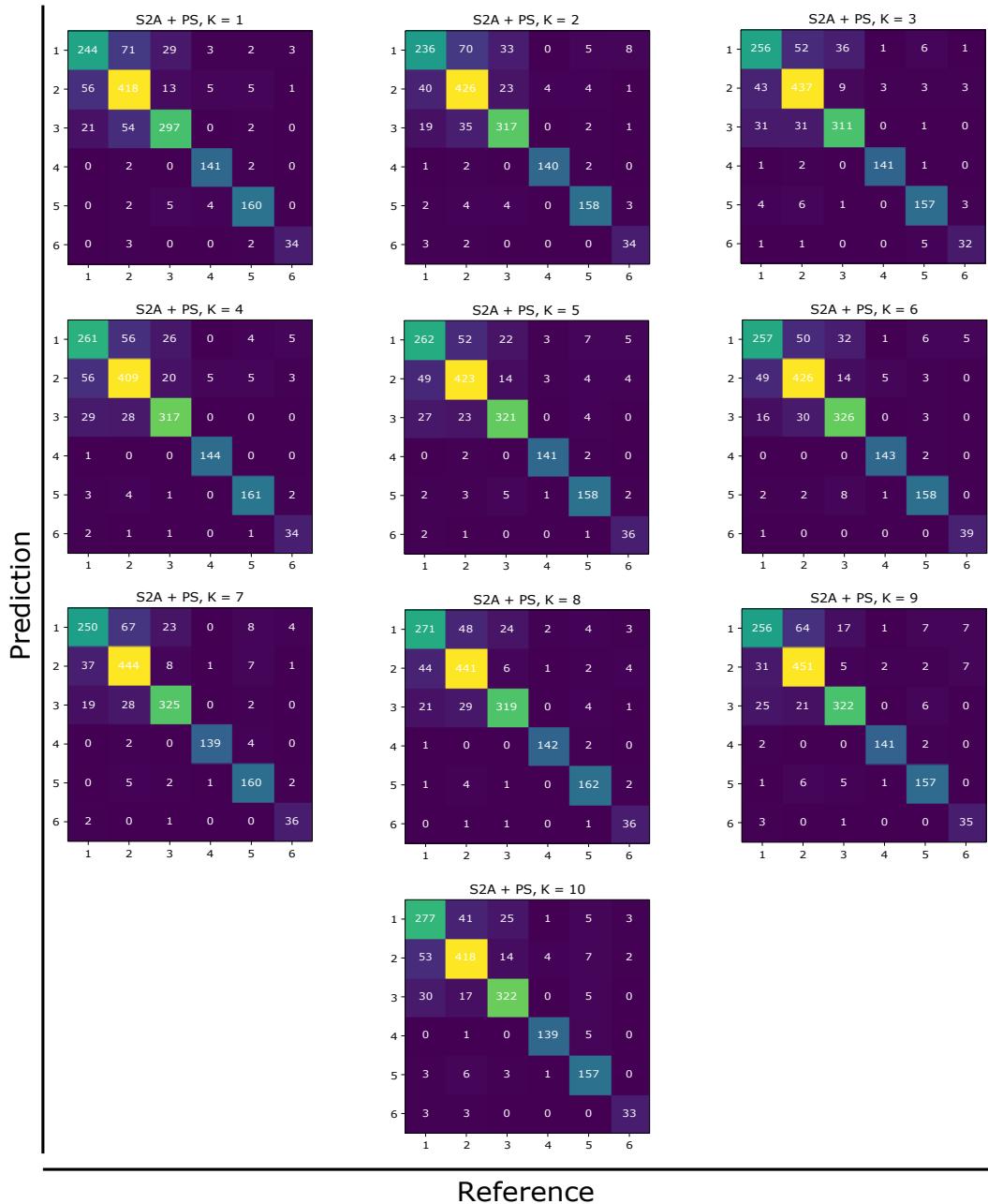
### C.1 Confusion Matrices



**Figure C.1:** Confusion matrices from 10-fold classification and accuracy assessment of the Sentinel-2A image. S2A = Sentinel-2A, K = Iteration. For the land cover classes considered, Woody = 1, Bunch grass = 2, Grazing lawns = 3, Water body = 4, Bare = 5 and Built-up = 6.



**Figure C.2:** Confusion matrices from 10-fold classification and accuracy assessment of the Planet image. PS = Planet, K = Iteration. For the land cover classes considered, Woody = 1, Bunch grass = 2, Grazing lawns = 3, Water body =4, Bare = 5 and Built-up = 6.



**Figure C.3:** Confusion matrices from 10-fold classification and accuracy assessment of the Fused image. S2A+PS = Fused image, K = Iteration. For the land cover classes considered, Woody = 1, Bunch grass = 2, Grazing lawns = 3, Water body = 4, Bare = 5 and Built-up = 6.