Satellite change detection in the Albany Thicket biome

Craig Mahlasi
MHLCRA002

Thesis Presented for the Degree of
DOCTOR OF PHILOSOPHY
Centre for Statistics in Ecology, the Environment and Conservation (SEEC) Department of Statistical Sciences
UNIVERSITY OF CAPE TOWN
February 2024

Supervisors: *Dr Glenn Moncrieff, Dr Vernon Visser, Prof Andreas Altwegg and Dr Jasper Slingsby*

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signed by candidate

Date: 25 July 2024

Acknowledgements

First and foremost I would like to thank my supervisor, Dr. Glenn Moncrieff for all the guidance and knowledge imparted onto me from the inception of the project. Thank you for relentlessly seeking to improve the analysis and the thesis as a whole. I would also like to thank my co-supervisors, Dr. Vernon Visser, for their untiring support and resourcefulness. Thank you for all your constant encouragement. To Dr. Jasper Slingsby and Prof Andreas Altwegg I wish to express my appreciation for your invaluable contributions throughout the journey of this thesis both in the fore with technical guidance and in the background in sourcing funding and other activities.

Thank you to my family for the immeasurable support and beautiful words of encouragement throughout this journey. A special thanks to Ms Kirsty Naiker and Ms Ntokozo Hlatshwayo for always being there to say “you’ve got this”.

I wish to thank Ms. Kyra Lunderstedt for allowing me to access and use data she had collected and maintained. The data formed the basis of this project. Your efforts reduced the overhead that comes with collecting such data. Similarly I wish to thank Google and Google Earth Engine team. The Google Earth Engine was used extensively as a data source and platform for analysis. Developing a dataset of precisely dated land cover changes was made possible by Planet’s Education and Research Program which provides access to daily, high resolution data that permitted the documentation and validation of land cover change.

Finally I want to express appreciation to the South African Environmental Observation Network (SAEON) for funding the project. Vital financial support was obtained from the South African National Space Agency and The Carnegie Carnegie Corporation of New York DEAL2/HUMA fellowship and for this I am thankful.

Dedication

The work embodied in this thesis is dedicated to:

Lorraine Mahlasi (20 August 1968-)
and
Maikeletsi Ziporah Matlhasi (12 March May 1942-)

Abstract

The Albany Thicket Biome has been subject to widespread land cover change, with as much as 63% of the biome being severely degraded. The primary land use activity responsible for much of the land cover change of the biome has been pastoralism and commercial agriculture land expansion. Remote sensing has been long confirmed as being an effective data source for performing land cover change monitoring.There are primarily four traditional remote sensing based change detection frameworks: algebra, transformations, classification, and advanced models. While these methods are able to detect changes in bi-temporal datasets, they are inherently limited in that they are based on the assumption that each pixel's spectral signature is a linear combination of the features on the corresponding physical surface. These methods also suffer from the propensity for false positives resulting from differences in atmospheric conditions, viewing angles and illumination and soil moistures between the two images; another limitation of these methods is the observation interval between the initial and post-change or successive observations are often weeks or up to years apart. which makes the detection of transient changes difficult. Finally they are unable to provide information on changes in land cover that allows for timely intervention by the authorities. Continuous change detection, on the other hand, uses all available and usable observations to detect changes in near-real-time. Continuous change detection classifiers allocate pixels throughout a time series to predefined classes using labelled training data. The majority of tools that seek to perform continuous land cover change detection have been developed for forests. Due to phenological and in turn spectral differences between forests and open such tools are expected to perform poorly in or mixed-canopy ecosystems. This study aims to use multi-temporal satellite imagery to detect land cover change in the Albany Thicket in near-real time. The first data chapter (Chapter 2) seeks to generate a map documenting the changes in the Albany Thicket biome due to land cover change between 2016 and 2019 and to produce an online application to visualise and interpret these changes. Chapter 3 focuses on developing a change detection protocol for identifying vegetation clearing events in the Albany Thicket using Temporal Convolution Neural Networks and comparing it against the Continuous Change Detection and Classification (CCDC), a

state of art algorithm. Finally chapter 4 sets out to develop a Domain adaptive Temporal Convolution Neural Network for continuous change detection in the Albany Thicket biome, the goal of this chapter is to assess whether variations and gradients in biophysical variables affect model performance the chapter presents two potential solutions to this. The study concluded that using medium resolution satellite imagery changes in Albany Thicket vegetation can be reliably detected and discerned. The ability to continuously detect changes using TempCNNs was shown to outperform CCDC. Albany Thicket cover dynamics were shown to be embedded within geographical contexts and that geographical gradients in biophysical variables influence the contextual representations learned by the TempCNNs. The study also showed that fusing TempCNNs with biophysical variables such as surface dryness information can improve their performance. Finally, it was shown that using meta-learning the TempCNN can be adapted to be robust in shifting domains by learning the most optimal parameter initializations that allow for capturing the invariant embeddings that facilitate generalisation across domains possible.

Table of Contents

Declaration.....i
Acknowledgements..... ii
Dedication.....iii
Abstract.....iv
List of Figures..... v
List of Tables..... x
List of abbreviations..... xi
Chapter 1: General Introduction and background..... 1
1.1 Background..... 1
1.2 Albany Thicket.....2
1.3 Land cover.....5
1.4 Remote sensing of land cover.....5
1.5 Methods of remote sensing change detection.....6
1.5.1 Continuous change detection.....9
1.5.2 Deep-Learning and change detection..... 9
1.5.3 Domain Adaptation.....10
1.6 Land cover change mapping in South Africa.....13
1.6.1 Land cover change mapping in Albany Thicket.....14
1.7 Problem statement..... 14
1.8 Aims and Objectives.....16
1.9 Study Area.....17
Chapter 2: Change detection in the Albany Thicket biome..... 22
2.1 Introduction.....22
2.2 Methods.....24
2.2.1 Data.....24
2.2.1.1 Satellite imagery.....24
2.2.1.2 Reference data.....25
2.2.2 Image processing.....27
2.2.3 Differential images.....28
2.2.4 Change classification..... 33
2.2.5 Accuracy assessment..... 34
2.2.6 Variable importance.....36
2.2.7 Extracting and dating clearing events..... 37
2.3 Results..... 37
2.3.1 Accuracy assessment..... 37
2.3.2 Variable Importance.....38
2.3.3 Change classification..... 41
2.3.4 Extracting and dating Albany Thicket loss.....43
2.4 Discussion.....44
2.4.1 Change classification..... 44
2.4.2 Accuracy assessment..... 45

2.4.3 Variable importance....46
2.4.4 Extracting and dating Albany Thicket loss....48
2.4.5 Implications for Albany Thicket disturbance monitoring....48
2.5 Conclusion....48
Chapter 3: Temporal Convolution Neural Networks for continuous change detection in the Albany Thicket biome....50
3.1 Introduction....50
3.2 Methods....56
3.2.1 Data....56
3.2.1.1 Satellite data....56
3.2.1.2 Reference data....57
3.2.2 Models....58
3.2.2.1 Continuous Change Detection and Classification....59
3.3.2.2 Temporal Convolutional Neural Networks....59
3.2.3 Predictions....61
3.2.4 Influence of aridity on performance....61
3.2.5 Influence of training sample size....61
3.2.6 Model Evaluation....62
3.3 Results....63
3.3.1 Change Detection....63
3.3.2 Influence of aridity on performance....65
3.3.3 Influence of training sample size....66
3.4 Discussion....68
3.4.1 Change Detection....68
3.4.2 Influence of aridity on performance....70
3.4.3 Influence of training sample size....71
3.5 Conclusion....72
Chapter 4: Domain adaptive Temporal Convolution Neural Networks for continuous change detection in the Albany Thicket biome....74
4.1 Introduction....74
4.2 Methods....79
4.2.1 Data....79
4.2.1.1 Satellite data....79
4.2.1.2 Surface moisture variables....79
4.2.1.3 Reference data....80
4.2.2. Models....82
4.2.2.1. Surface-dryness fused TempCNN Temporal Convolutional Neural Networks....82
4.2.3.2 MAML-Temporal Convolutional Neural Networks....85
4.2.4 Experiments....89
4.2.4.1 Model Evaluation....90
4.3 Results....91
4.3.1 Change Detection....91
4.3.2 Cross-biome generalisation and adaptation....94
4.4 Discussion....98

4.4.1 Change Detection....98
4.4.2 Cross-biome generalisation and adaptation....100
4.5 Conclusion....102
Chapter 5: Summary and Conclusion....104
5.1 Vegetation loss....104
5.1.1 Findings from Chapter 2....104
5.1.2 Findings from Chapter 3....106
5.2 Environmental gradients and Albany Thicket monitoring....107
5.2.1 Findings from Chapter 4....107
5.3 Implications for Albany Thicket management....109
5.4 Reproducibility....111
5.5 Advances and Future work....112
5.6 General Conclusion....113
6 References....115
7 Appendix....146

List of Figures

Figure 1.1. Photograph showing how Albany Thicket often occurs adjacent to other Vegetation types, in this case Albany Thicket vegetation can be seen adjacent to Forests……3

Figure 1.2. Diagram of the remote sensing process for both passive and active sensors……..……..…….7

Figure 1.3. A map of the vegetation types present in the study area, the vegetation types are based on the National Vegetation Map Project (VEGMAP).....18

Figure 1. 4. a) The mean spectral signatures of the dominant vegetation types (Thicket, forests, and grasses) in Alexandria. b) The mean spectral signatures of the different Albany Thicket subtypes found in Alexandria……..….19

Figure 1.5. Map of moisture availability across the thicket biome. The map shows the gradient of rainfall seasonality, the gradient of mean annual rainfall and and the gradient in the coefficient of variance in mean annual rainfall. ……..20

Figure 1.6. Locality map of the study area in the vicinity of Alexandria, South Africa from OpenStreetMap, below that is a true colour Sentinel-2 composite of the study area (bordered by red dotted line). ……..21

Figure 2.1. Sentinel-2 spectral bands (B1-B12) with corresponding spectral and spatial resolutions. All bands except B10 (cirrus detection) were used in this study. Source: (https://scihub.copernicus.eu/).....25

Figure 2.2. A visual representation of the delineation protocol implemented to create the reference data. Stable Forest (SF), Stable Thicket (ST), Stable

non-Thicket/Forest (STN), Thicket loss (TL), Forest loss (FL), Thicket Gain (TG), Forest Gain (FG)……..……...26

Figure 2.3 A map of the land-cover classes polygons from which the training and testing point data were extracted…..……...27

Figure 2.4. Ranking of the 70 classification features based on Gini importance……39

Figure 2.5. Pairwise Pearson's correlation between the 70 features used in the classification…….40

Figure 2.6. (a) a visualisation of the RF change classification. The map shows the spatial distribution of the classes. b) a barplot of the class-wise areal extents in km2 ……41

Figure 2.7. (a) histogram of the duration (days) in which the Albany Thicket loss events occur. b) the temporal distribution, showing the annual counts of clearing events. c) the temporal distribution of clearing by month. d) map of the spatial distribution of the loss events and the years in which they occurred. ……43

Figure 3.1. One zoomed in example of the time-series labelling protocol. The time-series was subsampled to windows with a fixed length of 20 observations with a five-day update and a step size of three. One of three labels is possible for each pixel's time-series subsample: Stable Thicket, inconclusive and Cleared Thicket. Labels were allocated to the subsamples by inspecting Planet Labs PlanetScope imagery. In this example the pixels within the red bounding box are labelled Stable Thicket up till step 7, at step 8 and 9 it is uncleared (inconclusive) and from step 10 it is evident that some of the pixels within the bounding box are cleared……..57

Figure 3.2. The Temporal Convolutional Neural Network (TempCNN) architecture. The input to the network is a multivariate time series of Sentinel-2 spectral bands and indices from a pixel. Three convolutional filters are applied in succession, a single dense layer, and a Softmax layer…….....59

Figure 3.3. a) confusion matrix of the CCDC prediction, b) confusion matrix of the TempCNN prediction……..63

Figure 3.4. a) RGB image of Thicket-barren land fringes, b) TempCNN classification. The two panels show an example of how the fringes between Albany Thicket vegetation and barren land were misclassified; the reason for this is likely a result of the adjacency effect……..……..64

Figure 3.5. a) RGB image of an area where there are differences between the TempCNN containing the Moisture Stress Index and the TempCNN without the Moisture Stress Index as a feature. b) An example of false positives in the Cleared Thicket class before adding the Moisture Stress Index. c) Example of how adding the MSI reduced the false positives observed in b) however they persist…….65

Figure 3.6. An example of an event-to-flag lag comparison between CCDC and TempCNN. In this example, the RGB column shows Natural Colour Composites of PlanetScope imagery at 10-day steps. The red bounding box is the region where clearing was manually observed. At each of the time steps, the TempCNN is better at detecting clearings while CCDC does not detect some clearings. The event-to-flag lag was between 10 and 20 days for TempCNN. The event-to-flag lag for CCDC was 15 days, provided that the difference in normalised values of all bands of a new observation is greater than one for all observations within a moving window period, typical of 3 observations. ……..……66

Figure 4.1. Spatial variation in total precipitation in Alexandria. b) Map of the study area showing the elevation gradient……76

Figure 4.2. A visual description of the sampling regime implemented in the training of the cross domain adaptation models. Panel a) illustrates the multivariate time-series, The time-series Sentinel-2 data were subsampled to windows with a fixed length of 20 observations with a five-day update and a step size of three (see Fig.3.4). b) depicts the splitting of the time-series by Albany Thicket subtypes. c) The allocation of Albany Thicket subtypes into training and testing regions. d) depicts the regions where Stable Thicket and Cleared Thicket samples were extracted from……..…80

Figure 4.3. The SDF-TempCNN. For the TempCNN branch of the network, the input is a multivariate time series of Sentinel-2 spectral bands and indices. Three convolutional filters are applied in succession, and a single dense layer. For the surface-dryness branch of the network the input is a multivariate of surface-dryness variables which are passed to five dense layers, the two branches are merged by concatenation. Followed by a dense layer and finally a Softmax layer……...84

Figure 4.4. MAML-TempCNN. The TempCNN is initialised, the data are then sampled from the train and test splits. The TempCNN parameters are used in the adaptation (1). During adaptation, the train set is sampled recursively to train the adapter (2); the losses and gradients are computed to update the weights (3). The Meta loss is computed using the update parameters based on the test set. Finally, the meta-gradients, meta-parameters using the meta-loss (4)..……...…….88

Figure 4.5. Grad-CAM++ comparison of saliency between the TempCNN (a) and the SDF-TempCNN (b). Temporally windowed saliency over a single pixel predicted to have undergone clearing. The NDVI values (blue lines) are overlaid to represent the loss of vegetation on the pixel over time. The red vertical lines indicate when the loss event occurred. The darker time regions represent higher saliency, these highlight parts of the input sequence that are most important or carry the most information for detecting changes in vegetarian cover……...……..93

Figure 4.6. a) A comparison of the effect of the number of shots (1,3 & 5 shot) on classification accuracy for the TempCNN, SDF-TempCCNN, MAML-TempCNN. The effect of the number of shots on the MAML-TemPCNN b) 1-shot, c) 3-shot and d) 5-shot on accuracy over 500 iterations. ……..……...95

Figure 4.7. The effect of the number of meta-iterations on the 5-shot MAML-TempCNN……96

Figure 4.8. Qualitative example of the detection results from the MAML-TempCNN, SDF-TempCNN and the baseline TempCNN. The example shows a clearing event in progress between 2019-09-10, 2019-09-20 and 2019-10-01 as well as difference in detection accuracy by the three models. The example shows that the SDF-TempCNN performs better than the other two models.……...……96

Figure 5.1. Resources and services for the development of systems of GEOAI based tools……111

List of Tables

Table.1. The 70 features used in the classification are based on Sentinel-2 (S2) bands. The table lists the position of each band in the stack and their description. Features 1-12 represent the difference between 2016 and 2019 for each Sentinel-2 band, features 13-60 represent the 12 individual surface reflectance bands of each of the four images. Refer to Fig. 2.1 for the wavelength positioning of each band…26

Table. 2**.** Error matrix based on independent testing points. Darker greens indicate a greater number of validation points……38

Table. 3. Performance of the models based on test data. Highlighted in bold is the best performing model. Here st denotes Stable Thicket and ct denotes Cleared Thicket……59

Table. 4. Model performance of varying the reference data sample size. Accuracy is based on test data. Highlighted in bold is the best performing model……..62

Table. 5. The sampling regime used to train the TempCNN and, SDF-TempCNN and MAML-TempCNN for within domain adaptability……75

Table. 6. Performance metrics of the TempCNN and SDF-TempCNN. These results are based on the model trained on the full dataset……88

Table. 7. Performance of the TempCNN and SDF-TempCNN on Albany Thicket subtypes on Albany mesic thicket (AMT), Geluck grassland thicket (GGT), Grahamstown Grassland thicket (GRT), without Nanaga savanna thicket (NST) and Kasouga Dune Thicket (KDT). Here SDF refers to the SDF-TempCNN model and TEM refers to the original TempCNN model……88

List of abbreviations

AMT Albany mesic thicket

BFAST Breaks for Additive Season and Trend Monitor

CD Change Detection

CNN Convolutional neural network

COLD Continuous Monitoring of Land Disturbance

EO Earth Observation

EVI Enhanced Vegetation Index

FG Forest Gain

FL Forest loss

GEE Google Earth Engine

GGT Geluck grassland thicket

GIS Geographic information system

GRAD CAM ++ Gradient-weighted Class Activation Mapping ++

GRT Grahamstown Grassland thicket

IPBES Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem

KDT Kasouga Dune Thicket

LCLUC Land cover and land use change

MAML Model-agnostic meta-learning

MLP multilayer perceptron

MSI Moisture Stress index

NST Nanaga savanna thicket

NDVI Normalised Difference Vegetation Index

NDWI Normalised Difference Water Index

NEMBA National Environmental Management Act

NIR Near infrared

NST Nanaga savanna thicket

RF Random Forest

RGB Red Blue and Green

S-CCD Stochastic Continuous Change Detection

SAR Synthetic Aperture Radar

SAVI Soil Adjusted Vegetation Index

SDF-TempCNN surface-dryness fused TempCNN

SF Stable Forest

SGD stochastic gradient descent

SITS-BERT Satellite image time series Bidirectional Encoder Representations from Transformers

ST Stable Thicket

STN Stable non-Thicket/Forest

SWIR Shortwave Infrared

TempCNN Temporal Convolutional Neural Network

TG Thicket Gain

TL Thicket loss

TOA Top of Atmosphere

VEGMAP National Vegetation Map

VGG Visual Geometry Group

Zero-shot super-resolution (ZSSR)

Chapter 1: General Introduction and background

1.1 Background

The term “biome” is used to describe geographical areas that exhibit similar macroclimate conditions, including average precipitation and temperature ranges, that give rise to distinctive vegetation structure (Woodward et al., 2004; Mucina and Rutherford, 2006; Crisp et al., 2009; Smissen and Rowe, 2018). Biomes can be thought of as large-scale ecosystems within which many ‘’sub-ecosystems’’ exist. Seven general biomes are recognised at a global scale and include Tropical Rainforests, Temperate Forests, Grasslands, Deserts, Taiga, Savannas and the Tundra. Biomes can be further classified into smaller distinctive units (Chitale et al., 2014). Biomes provide important ecosystem services (Parr et al., 2014).

Humans depend on biological resources and therefore their conservation and sustainable use is important for human existence (Iyyappanet al., 2018). According to Ellis et al., (2019), approximately 75% of the Earth’s landmass is under human utilisation to meet the agricultural, forestry and settlement needs of the 7.6 billion people on the Earth (Venter et al., 2016; Ellis et al., 2019). Growing populations coupled with global climate change are creating a plethora of environmental problems that result in ecosystems and in turn, land cover change; this is primarily due to the intensive utilisation of biological resources (Iyyappanet al., 2018; Lang and Song, 2019). According to the WWF, (2022), since 1970 the world has experienced an average decline of 69% in species populations. Due to human-induced alterations in community composition, local terrestrial ecosystems globally have, on average, experienced a minimum 20% reduction in the original abundance of naturally occurring species. Areas rich in endemic species have suffered even greater declines as a result IPBES Global Assessment on Biodiversity and Ecosystem Services 2020).

The main drivers of land cover change and the loss of biodiversity include the encroachment into native vegetation by crop and grazing lands, global climate change, agricultural and forestry practices that are unsustainable, growth of urban

areas and resource extraction, particularly mining (Günay et al., 2018). Human activities and ecological services are intrinsically intertwined, and therefore poor and uncontrolled biological resource usage can compound environmental problems, such as soil erosion, pollution, desertification and salinization (Lang and Song, 2019, IPBES Global Assessment on Biodiversity and Ecosystem Services 2021). According to Günay et al., (2018), the well-being of 3.2 billion people on the planet is compromised by land cover change. The net effects of habitat modification as a result of anthropogenic activities are stressed global carbon budgets and loss of biodiversity (Dafeng et al., 2008).

1.2 Albany Thicket

South Africa hosts nine different biomes in their National Vegetation Map, including Grassland, Savanna, Succulent Karoo, Nama Karoo, Forest, Fynbos, Desert, Indian Ocean Coastal Belt and Albany Thicket (Mucina and Rutherford, 2006). It is also home to three of the world's 36 global biodiversity hotspots (Myers et al., 2000), which are each primarily associated with a particular biomes, including the Cape Floristic Region associated with the Fynbos Biome, the Succulent Karoo, and the Maputaland-Pondoland-Albany hotspot associated with the Albany Thicket biome (Mucina and Rutherford, 2006). The Albany Thicket is believed to have emerged during the Holocene epoch (Cowling et al., 2004). The Albany Thicket biome represents a transitional vegetation category located between Afromontane forest, Subtropical forest, Fynbos, Grassland and Karoo vegetation (Lubke et al., 1986; Mucina and Rutherford, 2006; Parker, 2017) and bears characteristics from these adjacent biomes, for this reason it is thought to be an assemblage of various biomes (Cowling et al., 2004). Fig. 1.1 shows an example of Albany Thicket occurring adjacent to Forests.

Fig. 1.1) Photograph showing how Albany Thicket often occurs adjacent to other Vegetation types, in this case Albany Thicket vegetation can be seen adjacent to Forests. Source: (https://botanicalsociety.org.za/thicket-biome/)

The Albany Thicket biome extends along the valleys of the Fish, Sundays and Gamtoos rivers in the Eastern Cape province and extends westwards along the valleys of the Kouga, Baviaanskloof and Swartberg mountains (Mucina and Rutherford, 2006; Stickler and Shackleton, 2015). Approximately 10-20 % of the plant life in the Albany Thicket biome is understood to be endemic to the region (Stickler and Shackleton, 2015). Albany Thicket is described as low (2-5 m high), dense, semi-succulent, spinescent, evergreen, thorny scrub (Mucina and Rutherford, 2006; Parker, 2017). Vegetation units in the Albany Thicket biome vary greatly between the river catchments (Knight and Cowling, 2003). Along with floral diversity and complexity, the Albany Thicket supports an array of fauna, including 74 mammal species, 3 internationally recognised Important Bird Areas with 307 species of birds, 61 reptile and amphibian species (of which 6 are endemic to the Albany Thicket) and a diverse range of invertebrates (Knight and Cowling, 2003).1.3 Threats to the Albany Thicket biome

Despite the importance of the Albany Thicket for biodiversity, this biome is under threat of on-going land cover change. Approximately 63% of the Albany Thicket biome is believed to be severely changed or degraded and 29% is classified as moderately changed (Powell, 2009; Stickler and Shackleton, 2015). Land cover change in the Albany Thicket biome is described as a shift from a closed, dense canopy system with a restricted a grassy herbaceous layer to an open quasi-savanna (Mills et al., 2005; Powell, 2009; Stickler and Shackleton, 2015). Unsustainable land use practices are listed as the main cause of land cover change of the Albany Thicket biome. The primary land use activity responsible for much of the land cover change in the biome has been pastoralism and commercial agriculture land expansion (Mills et al., 2005; Powell, 2009; Stickler and Shackleton, 2015). Other agents of land cover change in the Albany Thicket biome include alien plant invasions, and altered fire frequency. Reduced fire frequency can create conditions that are conducive to the encroachment of woody vegetation or increases in fire frequency can result in the elimination of fire sensitive native plant species, such as Spekboom (*Portulacaria afra*) (Knight and Cowling, 2003; Stickler and Shackleton, 2015). The Albany Thicket biome has gone through significant land cover change as a result of the growth of settlements in rural and urban areas, the Nelson Mandela Metropolitan Municipality and Buffalo City Metro are examples of the pressure exerted onto the landscape by growing settlements (Knight and Cowling, 2003; Mucina and Rutherford, 2006). Mining currently has little effect on the Albany Thicket landscape, notwithstanding the remnants of past mining activities which exist in the landscape (Mucina and Rutherford, 2006). Land cover change in the Albany Thicket follows one of two main trends dependent on local rainfall: The first is encroachment of bush species, which is characterised by the densification of indigenous woody species in the mesic regions of the Albany Thicket. The second trend is the decline of woody and succulent shrubs in more xeric regions (Hoffman et al., 1987). The degradation of Albany Thicket represents a threat to species richness, functional type diversity, succulent shrub cover and succulent shrub diversity (Schmidt et al., 2019).

The rate of degradation and land cover change is widely regarded as an issue of great concern because of the sensitive nature of the Albany Thicket. The biome is not resilient to degradation and is unable to self-restore (Mills et al., 2005; Stickler

and Shackleton, 2015). The acceptance of the ecological importance of the biome, and that some Albany Thicket vegetation types are endangered (Skowno et al., 2018) has led to the establishment of efforts to halt further degradation and to restore degraded Albany

1.3 Land cover

Land cover refers to the surface type on the ground, which includes classes like urban infrastructure, water, vegetation, bare soil and ice. A concept related to land cover is land use. This refers to the classification of parts of the Earth's surfaces partly into classes representing their socio-economic uses, which are often either a result or driver of their land cover types. Common land use classes include areas utilised for residential, commercial or industrial purposes, for agriculture or forestry, or for conservation or recreational purposes (Ezeomedo and Igbokwe, 2013). Just as land cover, land use can change as a result of human intervention such as changes in zoning. Land cover and land use change (LCLUC) holds direct implications for the climate, biogeochemical cycles, and hydrology (Agidew and Singh, 2017). Land cover change is the principal driver of biodiversity loss both globally and in South Africa (Showalter et al., 2010; IPBES Global Assessment on Biodiversity and Ecosystem Services, 2021; Skowno et al., 2018). Land cover and land use change may have socio-economic ramifications, affecting the resilience and mitigation of societies to global change and threatening sustainable development. Thus, it affects individuals, societies, professionals (such as planners), corporations, governments, and policy makers (Walker et al., 2011).

1.4 Remote sensing of land cover

Remote sensing refers to the collection of information about objects, areas or phenomena on the Earth's surface without being in direct physical contact with them (Lillesand et al., 2015). This is often facilitated by sensors onboard satellites and aircraft that collect imagery of the Earth's surfaces by recording electromagnetic energy reflected by features on the Earth's surface (Heywood et al., 2011). Features on Earth's surface have distinctive spectral characteristics. This is because reflected radiant energy is a function of the structural, physical and chemical properties of the

surface it is incident upon, and the characteristics of the radiant energy, including the angle of incidence, the intensity of the radiant energy and its wavelength (Aggarwal, 2003). These distinctive spectral characteristics allow for the tracking and monitoring of features on Earth's surface through space and time. As a result, remote sensing has been used to provide information on LULC dynamics, mainly land cover change. There are two classes of remote sensing systems, namely passive and active remote sensing. Passive remote sensing systems are designed to record energy naturally reflected or emitted by objects on the target surface, Active remote sensing systems emit energy and measure the energy reflected by objects on the target surface (Lillesand et al., 2015). Fig. 1.2, depicts the remote sensing process for both passive and active sensing systems.

1.5 Methods of remote sensing change detection

There are four common change detection (CD) frameworks: algebra, transformations, classification, and advanced models (van Oort, 2007, Asokan and Anitha, 2019). Algebraic methods assess change by performing mathematical operations on two or more images of different dates (Kennedy et al., 2009, Lu et al., 2010, Kennedy et al., 2014, Asokan and Anitha, 2019). The most common of these methods is image differencing, which involves a pixel-by-pixel subtraction of an initial state image from the final state image and thresholding the result to determine change. The key drawback of algebraic methods is that they are unable to produce a change matrix, without which a comprehensive performance assessment of the CD output is not possible (Macleod and Congalton, 1998, van Oort, 2007).

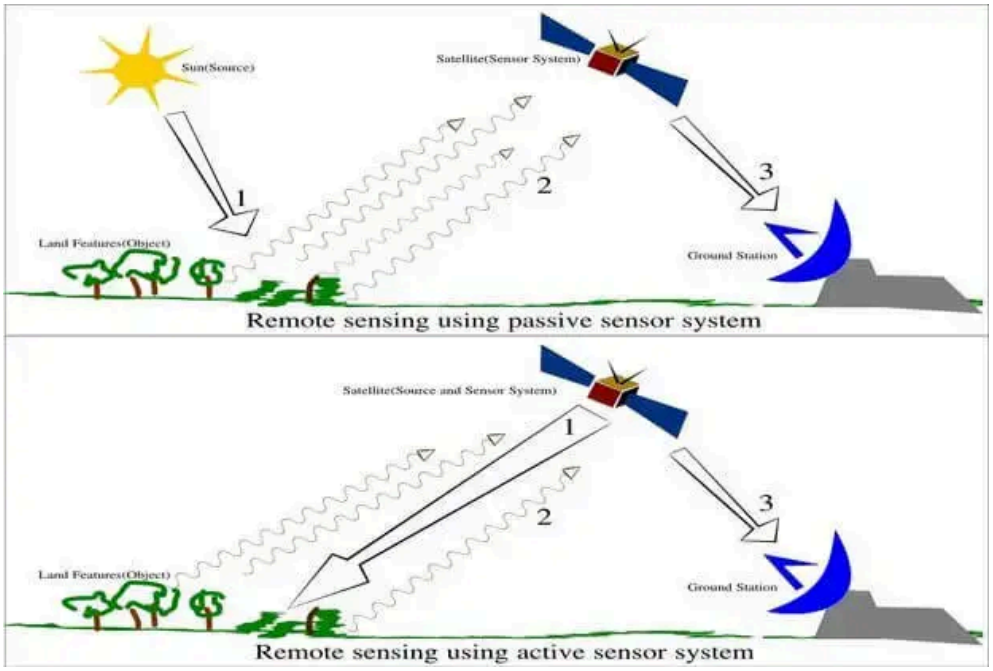


Fig. 1.2) Diagram of the remote sensing process for both passive and active sensors. Source : (https://allaboutcivil.org/components-of-remote-sensing/)

Transform methods of CD apply transformations to pixels. The process commonly involves multiscale decomposition of each image's spectral bands to produce independent components; a transformation is then applied to the image components. A change map is produced through threshold segmentation of the change components. Threshold segmentation involves determining the magnitude or intensity of change between time periods (Lu et al., 2010, Mishra et al., 2017, Asokan and Anitha, 2019). Transform methods include Tasselled Cap Transformation, Principal Component Analysis, Chi-Square Transformation and Gram-Schmidt Transformation. The main limitation to transform methods is the risk of confounding by sensor-noise, differences in atmospheric conditions and illumination. In areas that exhibit considerable degrees of temporal variability of these variables, the sampling frequency must satisfy the Nyquist criterion which states that the ability of an instrument to accurately resolve a signal, it is necessary that the sampling rate is twice the highest frequency of original signal. (Henebry, 2007). This, coupled with the inability to provide "from-to" information on changed pixels and class allocation can become cumbersome (Deng et al., 2008, Lu et al., 2010, Asokan and Anitha, 2019).

Classification methods of CD seek to discern change between two or more images of different dates by classifying the images into a variable number of different classes, e.g., land-cover types (Mas, 1999, Almutairi and Warner, 2010, Mishra et al., 2017). There are two commonly used frameworks of classification for CD: post classification and direct classification. Post classification involves an independent classification of the images followed by a comparison of the classifications. The comparison of the independent classified images involves statistically and visually assessing the spatial and areal extent of each land cover class (Almutairi and Warner, 2010; Lu et al., 2010; Congalton et al., 2014; Mishra et al., 2017; Alonso et al., 2023). The post classification method is handicapped by its sensitivity to the classification accuracies of the separate classifications, the classification inaccuracies are compounded and lead to greater uncertainties in the final map (Deng et al., 2008). In direct classification, instead of carrying out two classifications on separate images, a single classification of an image stack consisting of bands from both dates of interest is produced. This involves a process of annotating pixels that change between the two images and those that do not change (Cohen and Fiorella, 1998, Almutairi and Warner, 2010). Direct classification requires extensive a priori knowledge of the land cover classes present in the area of interest (Deng et al., 2008, Almutairi and Warner, 2010).

Advanced models of CD assess change in images by converting pixel reflectance values to physical parameters. These parameters are then used to extract information on the presence and abundance of surface materials (Sousa and Small, 2019). These models are based on the assumption that each pixel's spectral signature is a linear combination of the features on the corresponding physical surface, mainly land cover times(Salih et al., 2017, Sun et al., 2015). To assess change between two images, scene descriptive pure spectra (endmembers) are first selected and the spectra in each pixel are unmixed using a spectral unmixing model to estimate the abundance of each endmember. Abundance maps of each endmember are produced, and comparison of the abundance maps is done to discern change (Almutairi and Warner, 2010, Salih et al., 2017, Mishra et al., 2017). The limitation of advanced methods of CD is that they rely on in-situ data collection (Lu et al., 2010). Robust understanding of the atmospheric effects at play and the

application of atmospheric compensation are required to mitigate these effects (Jha et al., 2019).

1.5.1 Continuous change detection

The aforementioned attempts at mapping land cover change have relied independently mapping land cover at two periods in time. This is often referred to as “bi-temporal change detection”. While bi-temporal change detection remains the primary method for land cover change detection. Time series analysis of remote sensing data for change detection has gained much attention in the recent past (Woodcock et al., 2020). This is because many researchers are realising the robustness of such analysis, particularly for tracking the trajectory of land cover changes and providing precise information on the timing of land cover changes with improved accuracy. The availability of long-term data archives and planned future co-calibrated sensor missions also makes time series analysis possible. The computational power to perform these analyses has also steadily increased over time. There are examples of the implementation of time series data for continuous change detection in mainly forest ecosystems including (Verbesselt et al., 2012; Woodcock et al., 2020). While the application of these tools for monitoring Albany Thicket may be tempting to researchers, their application may result in discrepancies because they were initially developed for forested systems. Therefore there is a need for tools that are suitable for monitoring the entire Albany Thicket biome or regions that show great variation in aridity between xeric and mesic Albany Thicket.

1.5.2 Deep-Learning and change detection

There is growing adoption and implementation of Deep-Learning algorithms for land cover change detection. Deep-learning algorithms learn the representative and discriminative features from data in a hierarchical nature. Convolutional neural network (CNN) architectures have seen the most implementation, particularly for classification and target recognition exercises on remote sensing data. Shendryk et al. (2019) trained CNNs to perform multi-label classification. Yang et al. (2021) used CNNs to perform consistent classification on land use. Feizizadeh et al. (2021) used CNNs to detect saline flow sources. Chai et al. (2019) used CNNs to detect cloud and cloud shadow in Landsat imagery. These and other studies have reported

superior results compared to traditional supervised classification algorithms. However, the performance and utility of CNNs and other deep-learning architectures for performing continuous remote sensing change detection in open canopy systems such as the Albany Thicket has not been adequately investigated.

1.5.3 Domain Adaptation

The application of deep learning methods to solve land cover classification using remote sensing data has gained much attention and yielded results that confirm the utility of coupling remote sensing data and deep learning models in performing land cover classification tasks in diverse land cover types and geographical regions (Xu et al., 2023). However, the challenge of building a model robust enough to perform change detection in these diverse land cover types and geographical regions remains. This challenge is referred to as domain adaptive remote sensing and results from the inherent low performance of models that experience spectral/distribution shifts that result from images acquired under different surface and atmospheric conditions (Tuia and Bruzzone, 2021; Peng et al., 2022; Xu et al., 2023). These data spectral/distribution shifts tend to lead to poor model performance. The obvious solution to this is to develop and train a model with remote sensing data from all data distributions. However, this is not feasible due to the financial and logistical requirements to collect, store, and maintain the data and to provide the computational resources required to train models with numbers of parameters in excess of hundreds of millions. To make remote sensing models more adaptable to domain gaps/differences, several solutions have been proposed, including: 1) the inclusion and emphasis on cross domain invariant features (Xu et al., 2023); 2) minimising the differences/shifts in data distributions between domains (Marsocci, et al., 2023); and, 3) adaptation of the model towards the distribution of the target domain or through active learning by inclusion of the target domain during model training (Tsai et al., 2020; Tuia and Bruzzone, 2021). The goal of domain adaptation is to mitigate the discrepancies that arise from shifts in distributions and increase the generalisation capacity of models in varying distributions (Ganin and Lempitsky, 2015). Several methods have been proposed to address the issues of data scarcity and domain shift, including data augmentation (Hauberg et al., 2016; Takahashi et al., 2015; Shorten and Khoshgoftaar, 2019; Lalitha and Latha, 2022;),

transfer learning (Villamo et al.,2021; Bouchard et al., 2022; Ma et al., 2024) and meta-learning (Rußwurm et al., 2020; Zeng and Geng, 2022; Zhu et al., 2023; Ma et al., 2024).

Data augmentation has been used as a method of increasing the number of examples in a dataset. This is achieved by performing transformations on the data. Data augmentation transformations include; image rotation, random cropping, mixup, random erasing, and noise introduction (Takahashi et al., 2015; Shorten and Khoshgoftaar, 2019). Transfer learning is another approach that has been proposed to deal with data scarcity in other scientific disciplines.

Transfer learning refers to applying the knowledge learned by a general-purpose model trained on a larger dataset from one domain to a target domain where fewer training examples are available (Xie et al., 2016; Han et al., 2017, Bouchard et al., 2022). Transfer learning has been used in a range of applications including medical imaging (Yu et al., 2019; Rahaman et al., 2020; Kim et al., 2022), astronomy (Vilalta, 2018; kim, 2021), archaeology (Kazimi et al., 2018; Resler, 2021) and many others. Transfer learning has been previously applied in earth observation studies to perform tasks such as object detection (Villamo et al., 2021; Bouchard et al., 2022) and land cover classification (Xie et al., 2016; Li et al., 2020; Pham et al., 2022). In earth observation this can be useful where there are few annotated examples to train models that are able to learn domain specific representations that enable the model to generalise well in unseen data. Therefore the goal of transfer learning is to leverage knowledge learnt from one domain to optimise learning in the target domain (Ma et al., 2024). The advantages of transfer learning is it eliminates the need for having large annotated data which are costly to generate (Zhang et al., 2021; Ma et al., 2024). Another advantage of transfer learning is that it reduces the computational requirements that come with training large domain specific models.

The main shortcoming of transfer learning in remote sensing is that pre-trained models are often trained on three channel (RGB) images such as VGG, AlexNet, Resnet, GoogleNet among others. This widens the domain-gap between the source domain and the target domain (Chen et al., 2022; Wang 2023). Domain-gaps refer to the presence of changes in the features or objects of interest resulting from environmental factors including weather, and seasonal variations and time of day

(Doan et al., 2023). When the target domain is non-RGB such as in multispectral satellite data, the learning features of the pre-trained model do not account for the additional channels, which hold information that is often the basis of spectral signature separation. Another difference between conventional RGB and remote sensing imagery is the angle at which the images are taken. Conventional RGB images in most pre-trained models are at eye-level angles showing the face of the object. Remote sensing images are a high-level angle with variations including nadir, off nadir and oblique. These differences may be a source of error, for example the crown of a tree is different for the face of a tree which includes the trunk of the tree, thus the feature representations extracted from the two are likely to be for different distributions, because of the distinct colour distributions, texture, and context.

While transfer learning from pre-trained networks may be more efficient compared to training a neural network from scratch, the challenge is finding and using a pre-trained model whose source domain is similar to the target domain, to ensure optimal model performance. Spatial variations and heterogeneity of environmental variables make transfer learning a challenging task in earth observation. Secondly, it is not easy to produce a single set of pre-trained model weights that encompasses all satellite imaging platforms due to the inherent differences between these platforms (Stewart et al., 2022). Realising the shortcomings of traditional transfer learning, some researchers have sought to address these issues. Stewart et al., (2022), developed Torchgeo, a Python based package for facilitating deep learning on remote sensing data. The package also provides models pre-trained on multispectral satellite imagery. However, none of the currently available pre-trained models on Torchgeo deal with the temporal domain. Yuan and Lin, (2022), proposed Satellite image time series Bidirectional Encoder Representations from Transformers (SITS-BERT) where a neural network is pre-trained on a large dataset of a source domain. The pre-trained neural network is fine-tuned on small scale task-related labelled samples. A shortcoming of this work is that the pre-trained neural network is fine-tuned on small scale task-related labelled samples, and may fail to extract nuanced features that come from large scale/cross-biome samples.

Another approach to learning transferable representations from the source domain to the target domain, is through meta-learning. Meta-learning refers to “training a model that learns how to learn.” In Meta-learning related tasks are considered and few-shot

learning is implemented to speed-up domain adaptation on unseen tasks. The general idea is to train a model on a set of tasks such that it can learn to adapt to new tasks quickly, without requiring a large amount of training data for each task. Model-agnostic meta-learning (MAML) is a framework for meta-learning. Few-shot learning refers to providing a model a small amount of annotated examples, the goal of few-shot learning is to assess how much data a model needs in order to produce good generalisations when predicting unseen data (Mehra and Hamm, 2021; Zhou and Yu, 2023).

The goal for MAML is to use a small number of gradient steps in learning the most optimal initialization parameters that enable fast adaptation to new tasks. Rußwurm et al. (2020) implemented MAML in remote sensing classification and segmentation tasks using datasets that were distributed across the globe. The experiments showed that meta-learning model optimization may be advantageous for tasks where the data vary across regions. Cha et al., (2023) proposed Zero-shot super-resolution (ZSSR), to generate high resolution remotely sensed data by using the meta-knowledge about ZSSR learned from a compilation of low-resolution images that are treated as ZSSR tasks. The proposed model out-performed current state-of-the-art methods.

1.6 Land cover change mapping in South Africa

The utility of remote sensing in land cover change monitoring has long been recognised by South African researchers. Detailed national land cover assessments of South Africa have been conducted for the years 1994, 2000, 2005, 2013, 2018 and 2020 (https://egis.environment.gov.za/sa_national_land_cover_datasets, von Maltitz et al., 2019). The resulting datasets have been used as the basis for land cover change assessments. Skowno et al. (2021), detailed the Rates and patterns of habitat loss across South Africa's vegetation biomes. Skowno et al. (2017), used Satellite Observations from the period 1990-2013 to document Woodland Expansion in South African Grassy Biomes. Namugize et al.(2018), used these datasets to study effects of land use and land cover changes on water quality in the uMngeni river catchment, South Africa. Gibson et al. (2018), predicted the future land cover change scenarios in South African grasslands using two Landsat derived land cover

datasets for 2000 and 2014. Peter et al. (2018), observed land cover change in Kwelera Nature Reserve and possible implications on water resources using land cover maps generated from Landsat 8 imagery. Nzuza et al. (2021), used Sentinel-2 to predict land degradation in the Lepellane catchment of the Greater Sekhukhune District. While much effort has been put into mapping land cover change in many of South Africa's ecosystems, studies that focus on the application of continuous change detection methods are lacking. A review of the literature revealed three studies that focused on supervised continuous change detection methods including Odindi et al. (2020), Odebiri et al. (2023), Moncrieff (2021).

1.6.1 Land cover change mapping in Albany Thicket

Remote sensing-based change detection has been used to monitor land cover dynamics in the Albany Thicket biome of South Africa. Lloyd et al.(2002), used Iterative Self-Organizing Data clustering on Landsat data to map the patterns of land cover change and degradation in the Albany Thicket biome between 1997 and 2001. Kakembo et al.(2015), assessed elephant induced degradation in the Addo Elephant National Park between 1973 and 2010 by performing post-classification change detection using Maximum Likelihood Classification on Landsat and SPOT data. Carvalho et al. (2021), assessed degradation in the Albany Thicket biome between 2016 and 2018 using Maximum Likelihood Classification.

1.7 Problem statement

Numerous studies have attempted to map land cover change globally and within South Africa. These studies have aided understanding of the global and regional patterns and drivers of land cover change, but they are not sufficiently accurate and are not readily available to inform local monitoring and enforcement activities. For these purposes, detailed local assessments are required. To date, monitoring of the Albany Thicket of South Africa has relied on manual digitisation of high spatial resolution aerial photos or satellite images. While high spatial resolution data are desired for their ability to capture fine detail, such data require substantial financial, computational and human resources input (Showalter et al., 2010; Horning, 201; Giri

et al., 20154). This hampers their implementation for continuous change detection. Several researchers have sought to overcome the subjectivity and labouriousness of manual digitization by implementing supervised classification algorithms (Showalter et al., 2010; Horning, 2014; Hu et al., 2018). These have improved the accuracy and efficiency of change detection assessments. However, these assessments are often performed either as “once-off” exercises, with no repetition assessments, or at very low temporal resolution, typically on an annual basis. Furthermore, these assessments rely on retraining of models from scratch to produce updated land cover maps followed by manual or semi-automated change detection analysis. An additional limitation of these methods is that they are unable to generalise across geographic regions/domains.

This process can be accelerated using freely available satellite imagery and open-source data processing and analysis tools for land cover classification and change detection. While a lot of work has been done on documenting land cover change in the Albany Thicket, not much has been done to aid the rapid detection of Albany Thicket clearing and notification of authorities to allow for timeous intervention. Furthermore, Albany Thicket exhibits phenological traits that vary immensely from those of other vegetation types (Pierce and Cowling, 1984). An example of the reflectance differences can be seen in Fig. 1.4a. The differences are due to the differences in the signals reflected by the two vegetation types which may be explained by differences in Phenological traits between Albany Thicket and Forests. Robust models and tools that consider phenology are required to effectively detect and document changes in Albany Thicket. The mean spectral signatures of the dominant vegetation types in Alexandria are shown in Fig 1.4 a. The difference in the spectral signatures of the dominant vegetation suggest that there may be differences in vegetation structure, species composition, and differences in environmental factors among the vegetation types. Fig. 1.4 b, shows the intra-Thicket spectral variations. This plot suggests that there are structural, species composition, and gradients in environmental factors within Albany Thicket biome. Fig. 1.5. Is a Map of moisture showing the availability across the thicket biome. The gradient of rainfall seasonality shows that in the western regions of the Albany Thicket biome receive rainfall during the winter season, while the eastern regions mainly experience rainfall in the summer months. The gradient of mean annual

rainfall shows that the coastal regions receive the most rainfall and that mean annual rainfall decreases in the direction of the interior along with the coefficient of variance in mean annual rainfall. The three plots show the existence of spectral differences between dominant vegetation types in the study area and within the Albany Thicket as well as the difference in the amount of moisture available across the Albany Thicket biome. The variations in spectral signatures and environmental factors highlight the need for their consideration when performing continuous land cover mapping and change detection.

1.8 Aims and Objectives

Satellite remote sensing is recognized as a robust tool for characterising vegetation dynamics. As mentioned in section 1.12 there is evidence of spectral differences between Albany Thicket and other dominant vegetation types, these differences necessitate the development of protocols and methods for land cover change detection that are suitable for open canopy ecosystems that exhibit variations in environmental factors as is the case in the Albany Thicket. This study aims to use multi-temporal satellite imagery to detect land cover change in the Albany Thicket in near-real time. Degradation is an important phenomenon affecting the Albany Thicket due to the long history of overgrazing. However, the scope of the study was to develop protocols for near real-time change detection in the Albany Thicket. Degradation analysis was not a focus of the study. Furthermore the outputs of the change detection mapping can be used to infer degradation and be used in degradation analysis.

The objectives of the study are:

- To generate an Albany Thicket land cover change map documenting the changes in the Albany Thicket biome between 2016 and 2019 and to produce an online application to visualise and interpret these changes.
- To develop a change detection protocol for identifying clearing events in Albany Thicket using Temporal Convolution Neural Networks and compare it against the Continuous Change Detection and Classification (CCDC) algorithm.

- To develop a Domain adaptive Temporal Convolution Neural Network for continuous change detection in the Albany Thicket biome**.**

1.9 Study Area

In this study I investigated an area of 1339.87 km^2 around Alexandria, a town located 100 km northeast of Port Elizabeth in the Eastern Cape province of South Africa, (33°39.2′S, 26°24.5′E; Fig. 1.3). The average elevation is 195 m above sea-level. The geology of the study area belongs to the Bokkeveld Group, consisting primarily of sandstone and shale (Büttner et al., 2015). Primarily nine types of vegetation are present in the study area, including Albany Mesic Thicket in the northeast, Albany Valley Thicket in the north, and Grahamstown Grassland Thicket between the two. Kasouga Dune Thicket covers the coast from Kenton-on-Sea to Colchester. Nanaga Savanna Thicket covers the western and northwestern parts of the study area. Thorndale Forest Thicket are also found in the western regions (Mucina and Rutherford, 2006) Southern Eastern Coastal Thornveld in the central regions. Southern Coastal Thornveld in the southern regions. Geluk Grassland Thicket covers the region between Nanaga Savanna Thicket and Thorndale Forest Thicket. The differences in the type of vegetation between biomes and within biomes manifest in spectral signatures recorded by optical remote sensing systems, Fig 1.4, shows how the mean spectral signatures vary between biomes and within Albany Thicket subtypes. The spectral differences within the Albany Thicket have the potential of resulting in shifting domains. Rainfall ranged from 323 to 474 mm (Funk et al., 2015) during the training period (2016-2019). A locality map of the study area is depicted in Fig. 1.6.

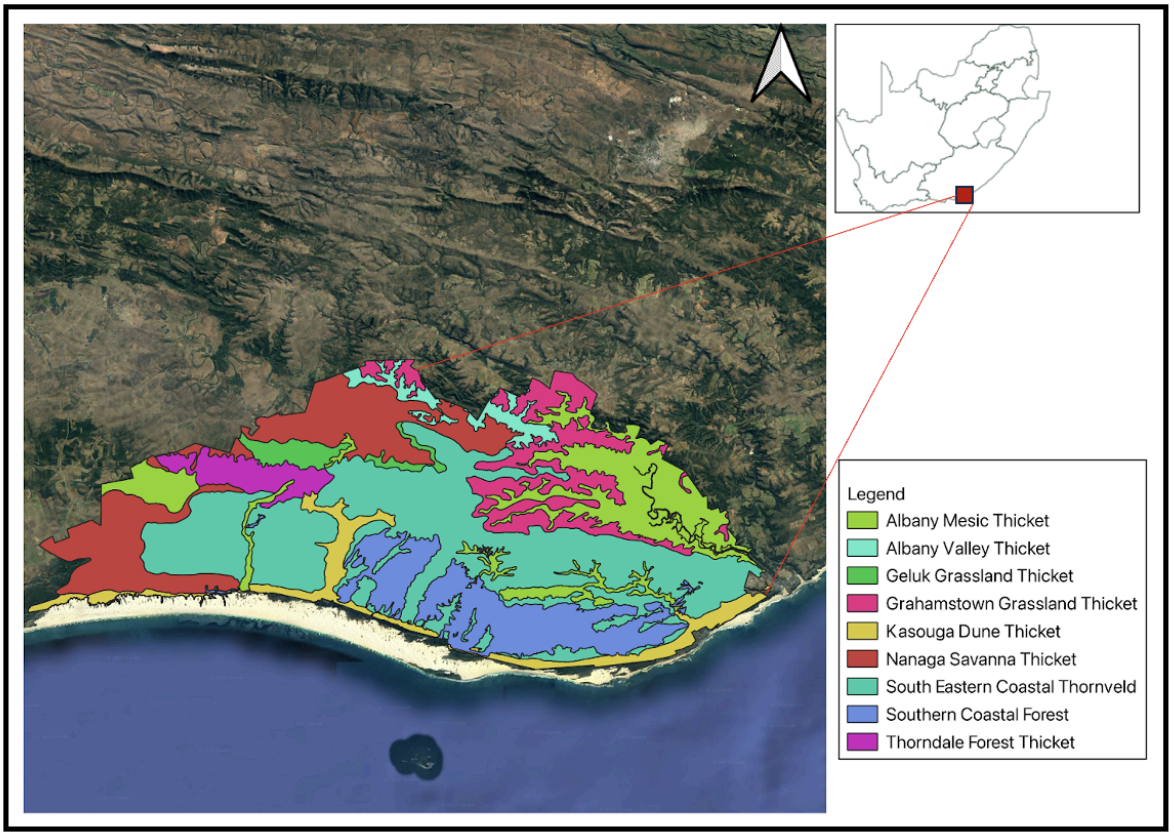


Fig. 1.3) A map of the vegetation types present in the study area, the vegetation types are based on the National Vegetation Map Project (VEGMAP). Source: (http://bgis.sanbi.org/vegmap).

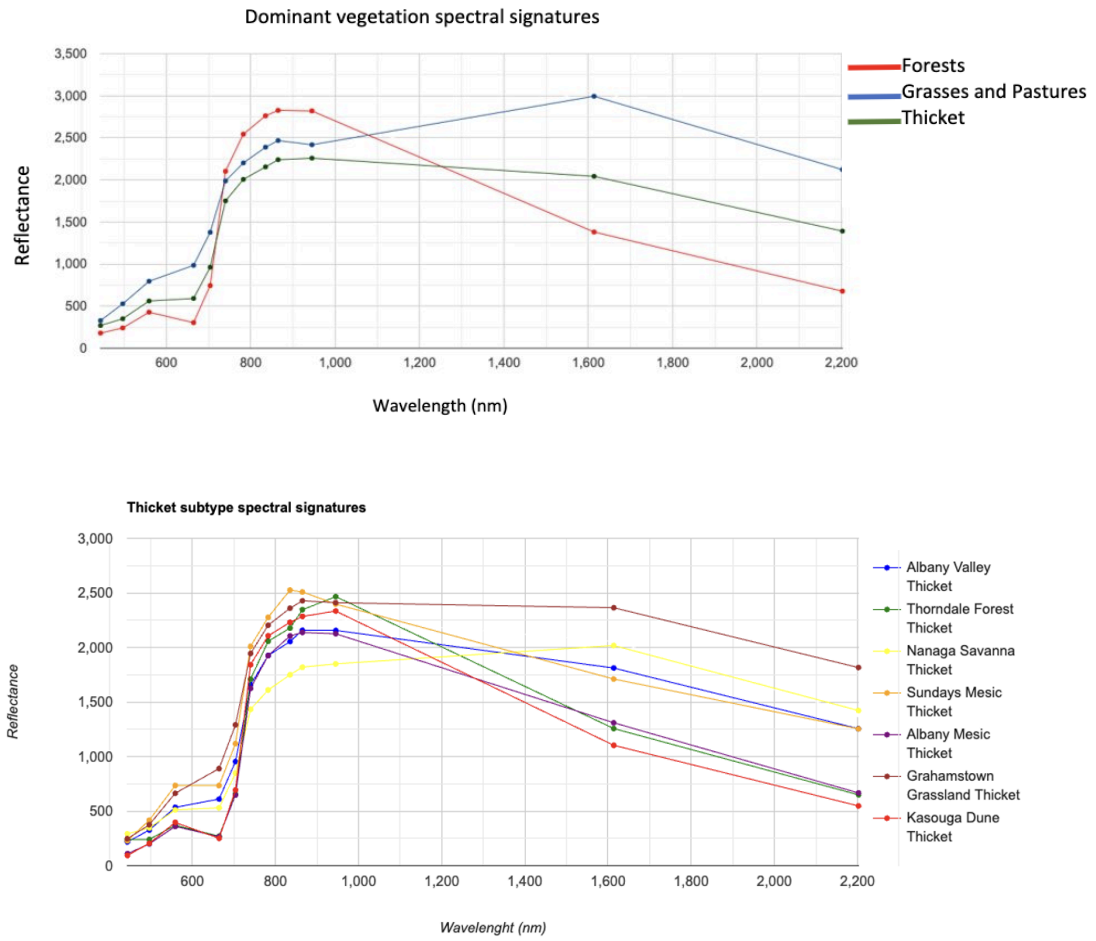


Fig. 1.4) a) The mean spectral signatures of the dominant vegetation types (Thicket, forests, and grasses) in Alexandria. b) The mean spectral signatures of the different Albany Thicket subtypes found in Alexandria.

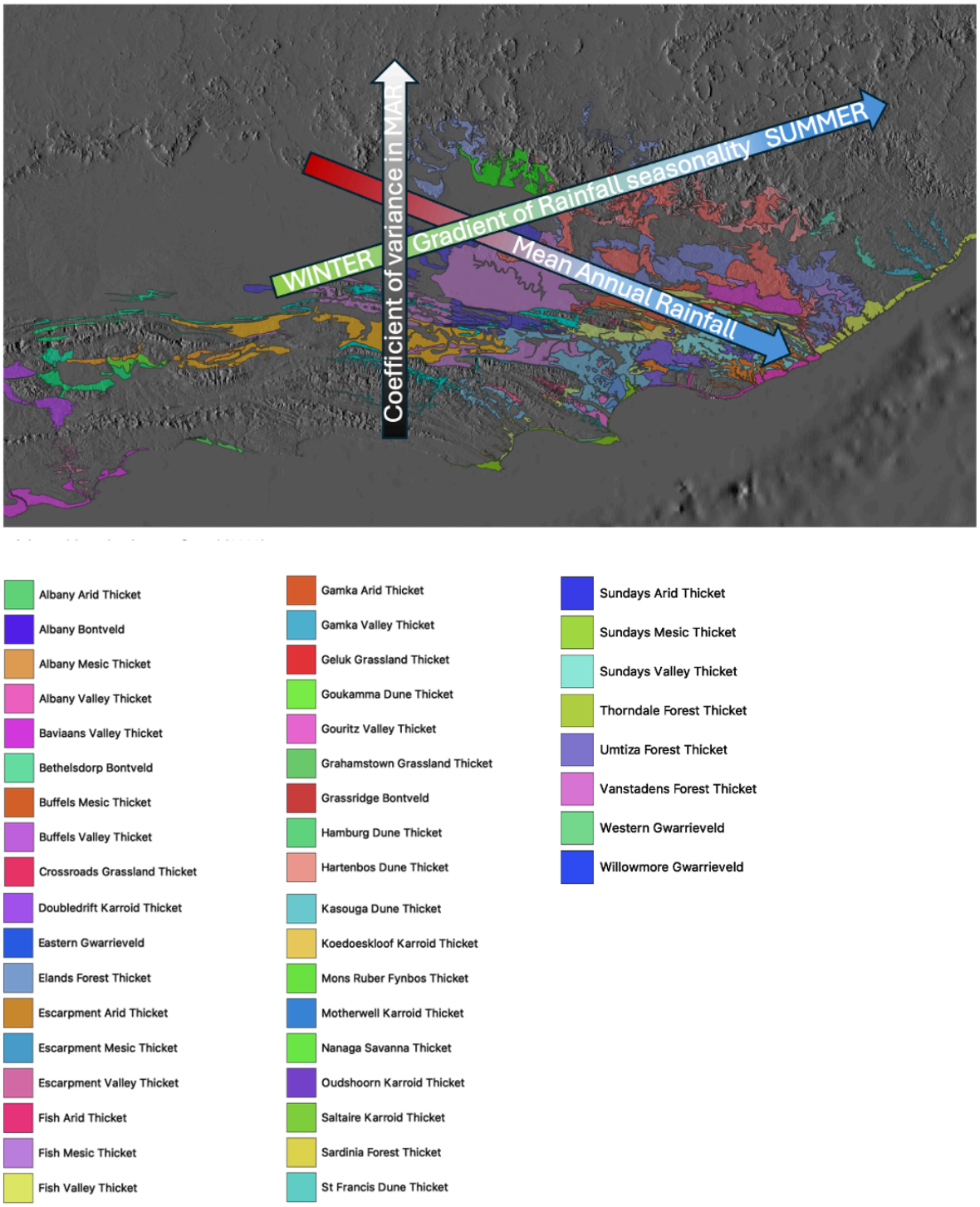


Fig. 1.5) Map of moisture availability across the thicket biome. The map shows the gradient of rainfall seasonality, the gradient of mean annual rainfall and and the gradient in the coefficient of variance in mean annual rainfall.

Fig. 1.6) Locality map of the study area in the vicinity of Alexandria, South Africa from OpenStreetMap, below that is a true colour Sentinel-2 composite of the study area (bordered by red dotted line).

Chapter 2: Change detection in the Albany Thicket biome

2.1 Introduction

Vegetation existing in a natural state is an important proxy for ecosystem health, because it plays important roles in various ecosystem functions that affect atmospheric, biospheric and hydrospheric fluxes (Fairbanks et al., 2000; IPCC, 2013; Chen et al., 2019). Globally It is estimated that 77% of terrestrial habitats across the globe have been changed (IPBES Global Assessment on Biodiversity and Ecosystem Services 2021), primarily as a result of human action. In South Africa, 22% of natural habitat has been changed since the arrival of European settlers (Skowno et al., 2021). The Albany Thicket is one of South Africa's nine nationally recognised biomes along with Desert, Forest, Fynbos, Grassland, Savanna, Indian Ocean Coastal Belt, and the Nama and Succulent Karoo biomes (Lubke et al., 1986; Mucina and Rutherford, 2006; Parker, 2017). It is estimated that only 11% of the Albany Thicket biome remains in a pristine state and approximately 63% is classified as severely degraded (Powell, 2009; Stickler and Shackleton, 2015). The change from Albany Thicket to other land cover types catalyses structural and functional changes and species loss; these changes result in adverse socioeconomic and environmental consequences (Powell, 2009). Agriculture is regarded as the main driver of land cover change in the Albany Thicket biome (Stickler and Shackleton, 2015, Hoffman et al., 1987).

Historically, land cover change detection (CD) has been achieved through manual delineation of land cover classes from greyscale or colour (Red/Green/Blue) aerial photography. The drawback of this method is that it is laborious, time consuming, requires substantial financial inputs, lends itself to subjectivity and is dependent on the availability of updated aerial photography (Showalter et al., 2010; Giri et al., 2015). The advent of Earth observing satellites producing regular images of the Earth's surface at medium to high resolution has ushered in a new era of land cover monitoring (Woodcock et al., 2008; Belgiua and Dragut, 2016; Andualem and Guadie, 2018). Earth observation-based land cover change assessments offer an

efficient alternative to traditional land-cover CD. Satellite imagery collected by Earth observation systems allows for large area coverage and for collection of data in remote areas (Giri et al., 2015). These satellites also typically measure multiple bands across the visible to short-wave infrared range, allowing land cover classes to be distinguished on the basis of differences in their reflectance of the electromagnetic spectrum. Recent developments in the spectral, spatial, and temporal capabilities of satellite imaging systems, as well as advances in computing and storage, have supported the adoption of Earth observation-based CD in various environments. For example, Novo-Fernández et al., (2018) used Landsat Time Series Stacks (LTSS) and the Vegetation Change Tracker (VCT) algorithm to assess forest disturbances in Durango, Mexico. Ye et al. (2020) used a dense Landsat data time series to detect wildfire and pest infestation disturbance across forests in the USA. Eid et al. (2020), mapped three decades of wetland vegetation change in the El-Burullus wetland, Egypt using Landsat data. Yuan and Cohen, (2020) assessed ecosystem conditions in the ridge-slough mosaic of the Everglades (South Florida, USA) between 1996–2007 using Landsat 5 TM derived products.

Land-cover change events can occur both gradually and abruptly (Zanotta and Haertel, 2012). Gradual change means that a parcel of land will often go through several successive spectral bins before reaching a final stable land-cover type. Abrupt change refers to changes that are observed between two consecutive observations, usually occurring within a few days and resulting from events such as fires, floods and landslides. While a lot of work has been done to document abrupt Albany Thicket land cover change, little is understood about the long trajectories of land-cover change (Knight and Cowling, 2003; Powell, 2009; Stickler and Shackleton, 2015). Tracking these trajectories of land-cover change is important in providing insight into the early cues of pending irreversible land cover change (Slingsby et al., 2020).

Despite the confirmed utility of Earth observation in monitoring ecosystem health, most studies have focused on applying remote sensing data in monitoring forest ecosystems and not open or mixed-canopy ecosystems like Albany Thicket. Subtropical ecosystems have received little attention, including the Albany Thicket of South Africa. Lloyd et al. (2002), mapped the patterns of land cover change and degradation in the Albany Thicket biome between 1997 and 2001 using Landsat

data. Carvalho et al. (2021), used Maximum Likelihood Classification to assess degradation in the Albany Thicket biome between 2016 and 2018. Land cover change monitoring in South Africa has generally relied on annual assessments (e.g., Thompson, 1996; Skowno et al., 2017). There is a need for rapid assessments if unauthorised land cover change is to be detected in time for authorities to respond and intervene, rather than simply reporting on historically observed patterns (Moncrieff 2022).

In this study, I present an image differencing-classification hybrid CD protocol for detecting change. The aforementioned CD methods assume that all change is linear and abrupt, which has been addressed in this study by performing trajectory analysis. Central to CD based on trajectory analysis is the use of multi-temporal imagery datasets to assess successive land cover types in a time series (Feng and Liu, 2014). Many of these studies use high temporal frequency trajectory analysis in CD studies (Kennedy et al., 2014, Feng and Liu, 2014). In this study, I use a 4-year set (2016-2019) of Sentinel-2 images to capture the subtle change trends and trajectories. The Random Forest (RF) classifier (Breiman, 2001), is trained to detect and map land cover change in Alexandria between 2016-2019. The resulting land cover change map is used to extract changes in Albany Thicket cover. Central to detecting and mapping change is allocating dates to change events. The resulting dataset of dated change events is used to perform continuous change detection in chapters 3 and 4 of the project.

2.2 Methods

2.2.1 Data

2.2.1.1 Satellite imagery

Multispectral, Level-1C (Top of Atmosphere (TOA) reflectance) Sentinel-2 data were used in this study. Sentinel-2A and B are a pair of multispectral imaging radiometers launched in 2015 and 2017 by the European Space Agency. The spectral and spatial resolution are shown in Fig. 2.1. Data for 23 December 2016, 8 December 2017, 23 November 2018, and 23 December 2019 were downloaded from the Copernicus Open Access Hub (https://scihub.copernicus.eu/). Since the image bands come at 3

different spatial resolutions, they were resample to a 30m grid. The specific scenes were selected to have zero percent cloud cover.

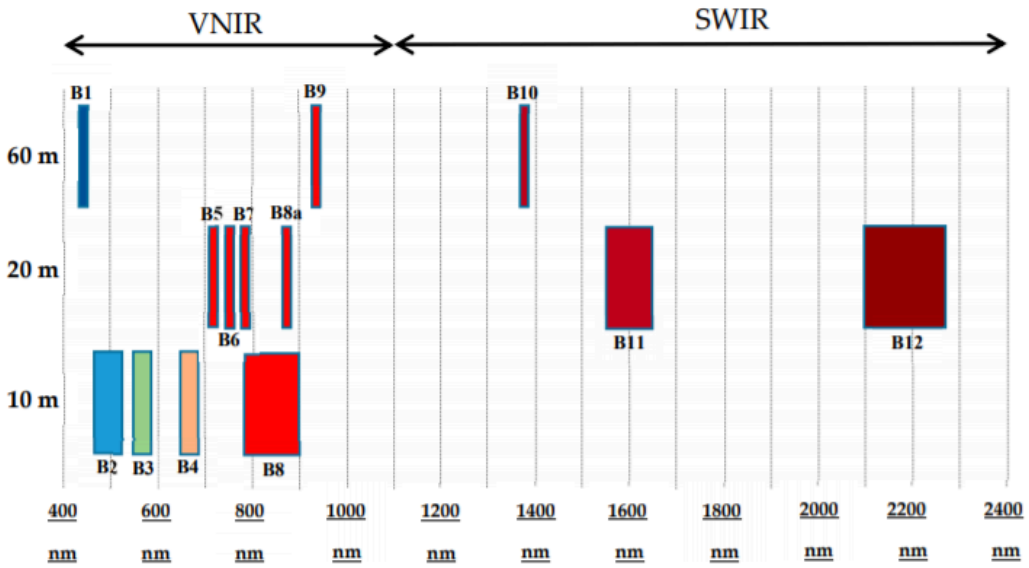


Fig. 2.1) Sentinel-2 spectral bands (B1-B12) with corresponding spectral and spatial resolutions. All bands except B10 (cirrus detection) were used in this study. Source: (https://scihub.copernicus.eu/).

2.2.1.2 Reference data

Training and testing data were adapted from Lunderstedt (2018). The data result from a long-term monitoring program carried out by the Rhodes Restoration Research Group. Using aerial photographs, members of this group manually delineated parcels into classes, including Thicket, Forest, mosaics, fragmented and changed. Here the word “sample” is used to refer to land cover type observed between 2016 and 2019. Samples were in the form of vector polygons. The data were adapted to the following classes: Stable Thicket (743 samples) 129.42km^2, Stable Forest (39 samples) 117.32km^2, Stable non-Thicket/Forest (103 samples) 60.42km^2, Thicket Loss (70 samples) 2.52km^2, Thicket Gain (32 samples) 1.2km^2, Forest Loss (23 samples) 0.89km^2 and Forest Gain (2 samples) 0.05km^2. A sample was considered stable if no apparent change was visually discernible between 2016 and 2019. The label Thicket Loss or Forest Loss were allocated to

samples where Thicket or Forest had changed to another class between 2016 and 2019. It should be noted that gains in Thicket and Forest were represented by densification of existing vegetation rather than the emergence of new vegetation (Fig. 2.2). Degraded areas were considered stable. The reason for this is because the land cover class remained as Albany Thicket. “The class Thicket loss” was allocated only where Albany Thicket was transformed to another class . The polygons were then converted to points/pixels; this resulted in 93000 pixels. The pixels were split 70% (65100 pixels) for training and 30% (27900 pixels) for testing across all classes.The imbalance in the samples across classes reflects their occurrence in reality. Visual confirmation of change/stability was performed by manually interpreting the polygons across the four time stamps. 2016, 2017, 2018 and 2019. For example, if there was no change in a polygon in the four images it was used as a stable sample. If there was change in one of the images that was discernible the sample was deemed to have changed.

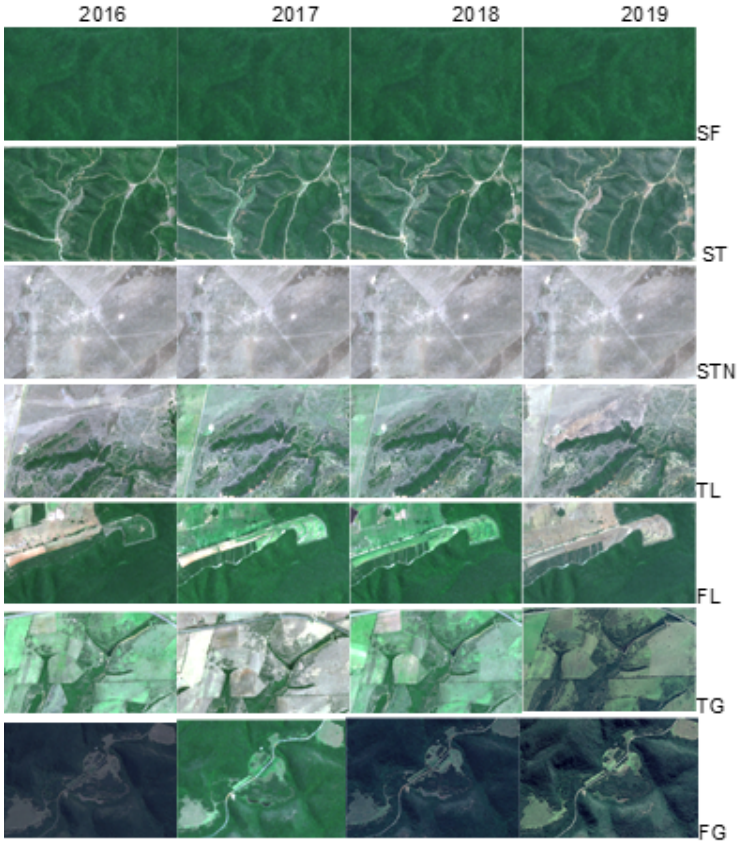


Fig. 2.2) A visual representation of the delineation protocol implemented to create the reference data. Stable Forest (SF), Stable Thicket (ST), Stable

non-Thicket/Forest (STN), Thicket loss (TL), Forest loss (FL), Thicket Gain (TG), Forest Gain (FG).

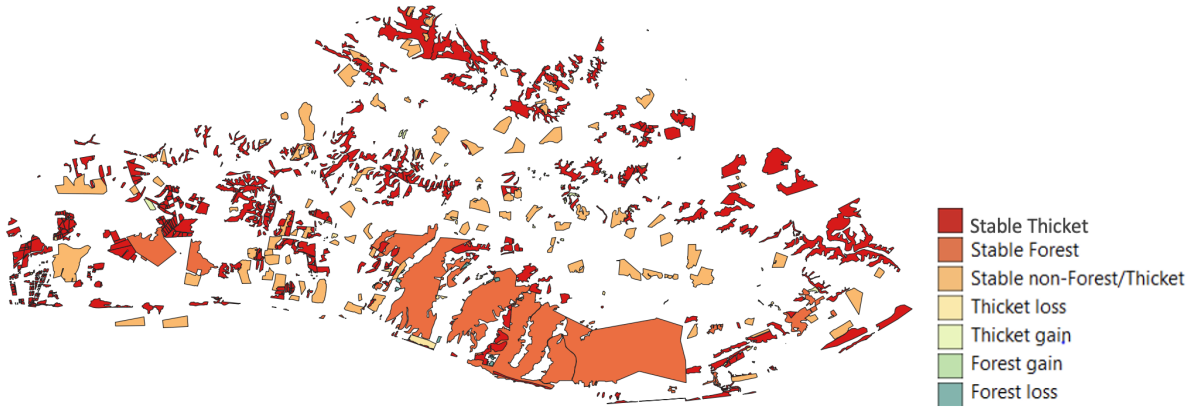


Fig. 2.3) A map of the land-cover classes polygons from which the training and testing data were extracted. The extracted training and testing data represents all the pixels within these polygons.

2.2.2 Image processing

Sen2Cor (version 2.8) atmospheric compensation was performed on each of the four Level-1C images. Sen2Cor is a processing module developed by the European Space Agency (Muller-Wilm, 2017). The module performs atmospheric, terrain and cirrus correction. The outputs of the Sen2Cor algorithm are ortho-corrected, per-pixel surface reference/bottom-of-atmosphere (BOA) values of the individual bands. The Normalised Difference Vegetation Index (NDVI) and Normalised Difference Water Index (NDWI) were also computed for each image.

$$NDVI = \frac{NIR-Red}{NIR+Red} \qquad \text{equation (2.1)}$$

where NIR is the near infrared wavelength band and Red is the red wavelength band.

$$NDWI = \frac{NIR-SWIR}{NIR+SWIR} \qquad \text{equation (2.2)}$$

where NIR is the near infrared wavelength band and SWIR is the shortwave infrared wavelength band

2.2.3 Differential images

Differential images were generated for the 2016 and 2019 images. Image differencing involves calculating the differences between two dates to create a residual image where stable (non-change) pixels are distributed around the mean and the change pixels are distributed on the tails (Mas, 1999). It should be noted that while the image differencing method is critiqued mainly for its inability to produce a change matrix, In this study I use it as a feature engineering technique and not as a change detection method. The rationale of engineering features through the image difference method, is that the random forest classifier is fed with prior guidance of change areas. Differential images of each of the 2016 and 2019 images were generated using the following formula:

$$Dx_{ij} = X_{ij}(T_2) - X_{ij}(T_1) \qquad \text{equation (2.3)}$$

where x is the band of interest, i and j are the row and column pixel values in T_1 (2016 image) and T_2 (2019 image) and D is the resulting differential pixel values. After creating the differential images, the surface reflectance bands from each of the four images and the differential bands were merged, resulting in a 70-band stack to be used as classification features. The classification feature stack is presented in Table 1.

Table.1. The 70 features used in the classification are based on Sentinel-2 (S2) bands. The table lists the position of each band in the stack and their description. Features 1-12 represent the difference between 2016 and 2019 for each Sentinel-2 band, features 13-60 represent the 12 individual surface reflectance bands of each of the four images. Refer to Fig. 2.1 for the wavelength positioning of each band.

Feature	Name	Description
Band 01	DiffB01	Difference between 2016 and 2019 for S2 B01
Band 02	DiffB02	Difference between 2016 and 2019 for S2 B02
Band 03	DiffB03	Difference between 2016 and 2019 for S2 B03
Band 04	DiffB04	Difference between 2016 and 2019 for S2 B04
Band 05	DiffB05	Difference between 2016 and 2019 for S2 B05
Band 06	DiffB06	Difference between 2016 and 2019 for S2 B06
Band 07	DiffB07	Difference between 2016 and 2019 for S2 B07
Band 08	DiffB08	Difference between 2016 and 2019 for S2 B08
Band 09	DiffB08A	Difference between 2016 and 2019 for S2 B08A
Band 10	DiffB09	Difference between 2016 and 2019 for S2 B09
Band 11	DiffB11	Difference between 2016 and 2019 for S2 B11
Band 12	DiffB12	Difference between 2016 and 2019 for S2 B12
Band 13	DiffNDWI	Difference between 2016 and 2019 for NDWI
Band 14	DiffNDVI	Difference between 2016 and 2019 for NDVI
Band 15	2016B01	2016 B01 (Coastal and Aerosol)

Band 16	2016fB02	2016 B02 Blue
Band 17	2016B03	2016 B03 Green
Band 18	2016B04	2016 B04 Red
Band 19	2016B05	2016 B05 Visible and Near Infrared (VNIR)
Band 20	2016B06	2016 B06 Visible and Near Infrared (VNIR)
Band 21	2016B07	2016 B017 Visible and Near Infrared (VNIR)
Band 22	2016B08	2016 B08 Visible and Near Infrared (VNIR)
Band 23	2016B08A	2016 B08A Visible and Near Infrared (VNIR)
Band 24	2016B09	2016 B09 Visible and Near Infrared (VNIR)
Band 25	2016B11	2016 B11 Short Wave Infrared (SWIR)
Band 26	2016B12	2016 B12 Short Wave Infrared (SWIR)
Band 27	NDWI2016	2016 NDWI
Band 28	NDVI2016	2016 NDVI
Band 29	2017B01	2017 B01 (Coastal and Aerosol)
Band 30	2017B02	2017 B02 Blue
Band 31	2017B03	2017 B03 Green

Band 32	2017B04	2017 B04 Red
Band 33	2017B05	2017 B05 Visible and Near Infrared (VNIR)
Band 34	2017B06	2017 B06 Visible and Near Infrared (VNIR)
Band 35	2017B07	2017 B07 Visible and Near Infrared (VNIR)
Band 36	2017B08	2017 B08 Visible and Near Infrared (VNIR)
Band 37	2017B08A	2017 B08A Visible and Near Infrared (VNIR)
Band 38	2017B09	2017 B09 Visible and Near Infrared (VNIR)
Band 39	2017B11	2017 B11 Short Wave Infrared (SWIR)
Band 40	2017B12	2017 B12 Short Wave Infrared (SWIR)
Band 41	NDWI2017	2017 NDWI
Band 42	NDVI2017	2017 NDVI
Band 43	2018B01	2018 B01 (Coastal and Aerosol)
Band 44	2018B02	2018 B02 Blue
Band 45	2018B03	2018 B03 Green
Band 46	2018B04	2018 B04 Red
Band 47	2018B05	2018 B05 Visible and Near Infrared (VNIR)

Band 48	2018B06	2018 B06 Visible and Near Infrared (VNIR)
Band 49	2018B07	2018 B017 Visible and Near Infrared (VNIR)
Band 50	2018B08	2018 B08 Visible and Near Infrared (VNIR)
Band 51	2018B08A	2018 B08A Visible and Near Infrared (VNIR)
Band 52	2018B09	2018 B09 Visible and Near Infrared (VNIR)
Band 53	2018B11	2018 B11 Short Wave Infrared (SWIR)
Band 54	2018B12	2018 B12 Short Wave Infrared (SWIR)
Band 55	NDWI2018	2018 NDWI
Band 56	NDVI2018	2018 NDVI
Band 57	2019B01	2019 B01 (Coastal and Aerosol)
Band 58	2019B02	2019 B02 Blue
Band 59	2019B03	2019 B03 Green
Band 60	2019B04	2019 B04 Red
Band 61	2019B05	2019 B05 Visible and Near Infrared (VNIR)
Band 62	2019B06	2019 B06 Visible and Near Infrared (VNIR)
Band 63	2019B07	2019 B017 Visible and Near Infrared (VNIR)

Band 64	2019B08	2019 B08 Visible and Near Infrared (VNIR)
Band 65	2019B08A	2019 B08A Visible and Near Infrared (VNIR)
Band 66	2019B09	2019 B09 Visible and Near Infrared (VNIR)
Band 67	2019B11	2019 B11 Short Wave Infrared (SWIR)
Band 68	2019B12	2019 B12 Short Wave Infrared (SWIR)
Band 69	NDWI2019	2019 NDWI
Band 70	NDVI2019	2019 NDVI

2.2.4 Change classification

Many studies have demonstrated the utility of the random forest (RF) (Breiman, 2001) algorithm in remote sensing image classification, because of its ability to handle datasets with high dimensionality and multicollinearity (Haas and Ban, 2014, Du et al., 2015, Belgiua and Dragut, 2016, Adjognon et al., 2019). Ease of implementation and general good performance are among other reasons why the RF classifier is favoured over other classifiers (Du et al., 2015). The RF classifier (Breiman, 2001) is a non-parametric, ensemble classifier that employs classification and regression trees (CARTs) to make predictions. In an RF classifier (Breiman, 2001), decision trees are generated by randomly selecting a subset of training samples through bootstrapping aggregation (bagging) (Breiman, 2001). This subset of training samples is referred to as “in-bag” samples. During the tree generation (“growing”) process, nodes are split using a random selection of input variables. The data are classified using a voting system and classes are allocated to each data point through the plurality vote of every single tree in the forest (Rodriguez-Galiano et al., 2012). The remaining samples are known as “out-of-bag” (oob) samples (Breiman, 2001).

The RF classifier (Breiman, 2001) within Python's scikit-learn package was implemented in this study to map change in the study area. The implementation requires the specification of two parameters: the number of decision trees to be grown (*n_estimators*) and the number of variables randomly sampled as candidates for splitting each tree node (*max_features*). The number of trees grown was guided by Belgiua and Dragut, (2016) who reported that errors stabilise before 500 trees are grown, through this the *n_estimators* was set to 500. The total number of classification features (70 bands) was used to set *max_features*, so that all of the features would contribute to node splits (Ghosh et al., 2014). Several studies including Probst, et al, (2019) and Schratz et al, (2019), have proposed hyperparameter tuning for the optimisation of RF classifiers, while hyperparameter tuning is valuable for model optimisation for the reasons above it was not implemented.

2.2.5 Accuracy assessment

To evaluate the performance of the RF classification (Breiman, 2001), several evaluation metrics were produced, including overall accuracy, producer's accuracy, user's accuracy, and Cohen's kappa coefficient (κ). As mentioned above, the RF classifier (Breiman, 2001) performs an internal cross-validation, however complementary external cross-validation by way of the spatial leave-one-out cross-validation method was used to evaluate the accuracy of the classification. The spatial leave-one-out method is a variant of the k-fold cross-validation with a spatial buffer. In k-fold cross-validation the sample set is randomly split into a chain of uniform subsets/folds (Brovelli et al., 2008; Airola et al., 2017; Sharma et al., 2017). The number of folds is denoted by *k*. All but one of the folds are used in the training of the model while the remaining fold is used to test the model. This iterates until every single fold is used as test data. The results from each of the iterations are averaged and then reported. In leave-one-out cross validation, $k = n$, where n is the size of the sample set. The size of each test set is = 1 and the model is therefore trained n times (Brovelli et al., 2008). To implement spatial leave-one-out cross validation, point samples were extracted from the polygon reference data (Fig. 2.3). The points were further split into training and validation sets. A 120 m spatial buffer was established around each of the validation points and training observations falling

within the validation buffers were discarded. The choice of a 120 m spatial buffer was directed by a 6x6 window of the Sentinel-2 20 m pixel size. This was also supported by the fact that the Sentinel-2 bands commonly used to study vegetation dynamics have a native spatial resolution of 20 m (Gascon et al., 2017). The RF was trained using the training set and predictions made for the entire image, with the testing points used to evaluate the predictive performance of the RF classification (Breiman, 2001). In this study n = 27900.

The overall accuracy measures the percentage of correctly classified pixels.

$$\text{Overall accuracy} = \frac{P_a}{t_a} \qquad \text{equation (2.4)}$$

where P_a is the total number of correctly classified pixels, t_a is the total number of pixels in the testing set.

The user's accuracy measures the reliability of the final classification, by measuring the percentage of pixels classified as belonging to class b in all pixels predicted as class b.

$$\text{User's accuracy} = \frac{P_b}{t_b} \qquad \text{equation (2.5)}$$

where, P_b is the total number of pixels correctly classified as belonging to a class, the total number of pixels classified to that class throughout the image is t_b and b are the number of pixels in that class.

The producer's accuracy measures the degree of accuracy in the classification of reference pixels.

$$\text{Producer's accuracy} = \frac{P_b}{t_a} \qquad \text{equation (2.6)}$$

where P_b is the total number of pixels correctly classified in a class and t_a is the total number of pixels belonging to that class in the reference data.

Cohen's Kappa coefficient (κ) measures the degree of agreement between two independent raters. Cohen's Kappa accounts for chance agreement.

$$Kappa = \frac{P_o \; P_c}{1 - P_c} \qquad \text{equation (2.7)}$$

where,

$$Po = \frac{\sum_a P_a}{\sum_a t_a} \qquad \text{equation (2.8)}$$

Where P_a is the total number of correctly classified pixels, t_a is the total number of pixels in the image and t_a is the total number of pixels belonging to that class in the reference data.

$$Pc = \frac{{}_k({}_b P_b \;\; \sum_a P_{ak})}{{}_a t_a \;\; {}_a t_a} \qquad \text{equation (2.9)}$$

where P_a is the total number of correctly classified pixels, t_a is the total number of pixels in the image. P_b is the total number of pixels correctly classified in a class, K is the number of classes and t_a is the total number of pixels belonging to that class in the reference data.

2.2.6 Variable importance

The RF classifier is able to produce measures of variable importance through various ranking protocols. In this study, the Gini feature importance protocol was implemented. In this study, the Gini feature importance was used to assess the spectral sensitivity of each of the bands in detecting and classifying change in Albany Thicket environments. The permuted cases are passed down to sub-node generated trees in the forest. The permuted data are then used to subtract the correctly classified data points from the correctly allocated data points in the non-permuted data. This produces an importance score for each classification

feature (Menze et al., 2009). Pearson's pairwise correlation test was performed on band pairs to test for collinearity between bands to describe the possible impact of correlation between features on the permutation importance.

2.2.7 Extracting and dating clearing events

Pixels that were classified as Thicket loss were extracted from the final classification. After extracting the pixels, boundaries around groups of clearing pixels were manually delineated so as to convert individual pixels to "whole instances". Here the phrase "whole instances" refers to a single clearing event that may take place over a single PlanetScope acquisition date or multiple PlanetScope dates. The utility of allocating dates to clearing instances is that it allows for the inference of information on the date of the occurrence and duration of each clearing event and enables the attribution of trends in clearings to time based events such as events during a farming season. Secondly, in an operational setting, the dating of clearing events allows for the detected clearings to be related with property ownership data in order to find potential perpetrators.

Dates were allocated to the Thicket clearing events. Two dates were assigned to each event: the date when Thicket was last confirmed to be present as well as the date that an entire instance was confirmed to have been cleared (Moncrieff 2021). This exercise was done using daily PlanetScope imagery from Planet (https://www.planet.com/). Using PlanetScope imagery allowed for precise date allocation and boundary reconciliation.

2.3 Results

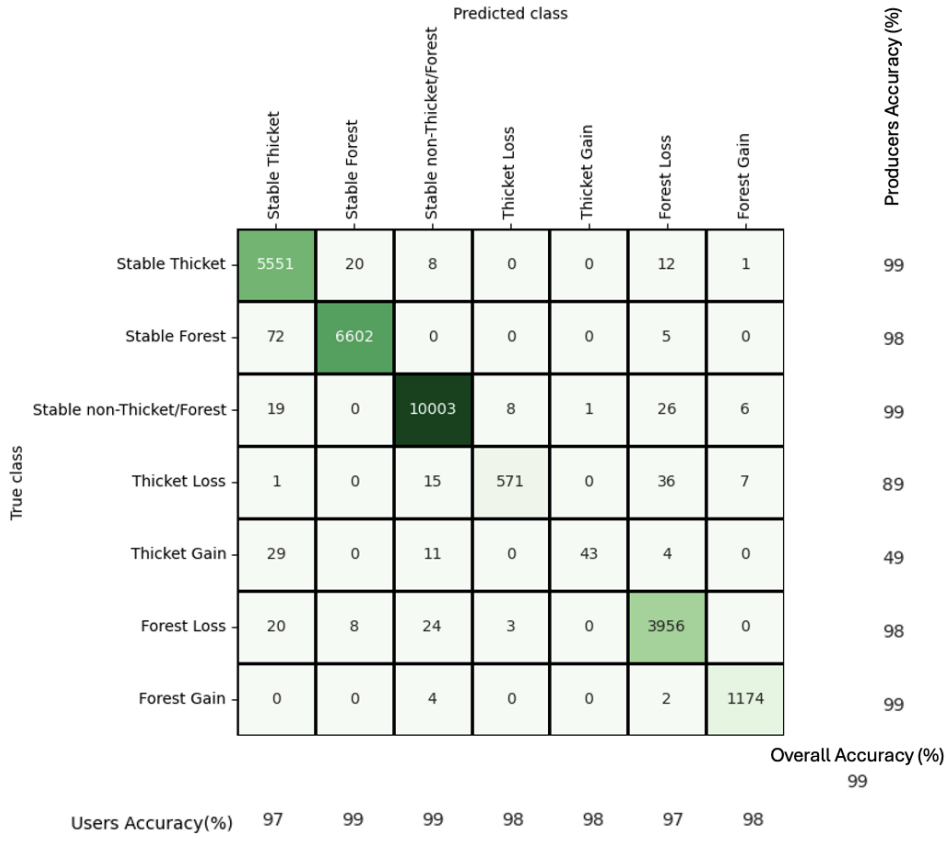
2.3.1 Accuracy assessment

The overall change classification showed high accuracy with an overall accuracy of 99.0% and a Kappa coefficient of 0.98 (Table. 2). All the classes showed high (>96%) User accuracies. Producer accuracies ranged from 49% for the Thicket Gain class to 99% for the Stable non-Thicket/forest and Forest Gain classes. A validation accuracy of 95.0% was observed

2.3.2 Variable Importance

Some of the features were more important than others in classifying change in the study area, The pairwise correlation test are shown in Fig. 2.5. Difference bands were weakly correlated to the surface reflectance bands. Surface reflectance bands were highly positively correlated with each other. The visible bands were weakly correlated to the shortwave infrared (SWIR) bands (B11 and B12). There was relatively high positive correlation between spectral indices and SWIR bands compared to the visible bands. The band 1 difference band showed the highest importance with a score of 0.049 and the 2017 spectral reflectance band 4 ranked as the least important feature with a score of 0.028 in the classification (Fig. 2.4). The 2017 and 2018 red surface reflectance bands had low importance scores. SWIR wavelengths were the most important group of features in classifying change.

Table. 2. Error matrix based on independent testing points. Darker greens indicate a greater number of validation points.



kappa coefficient= 0.98, Validation accuracy = 0.95

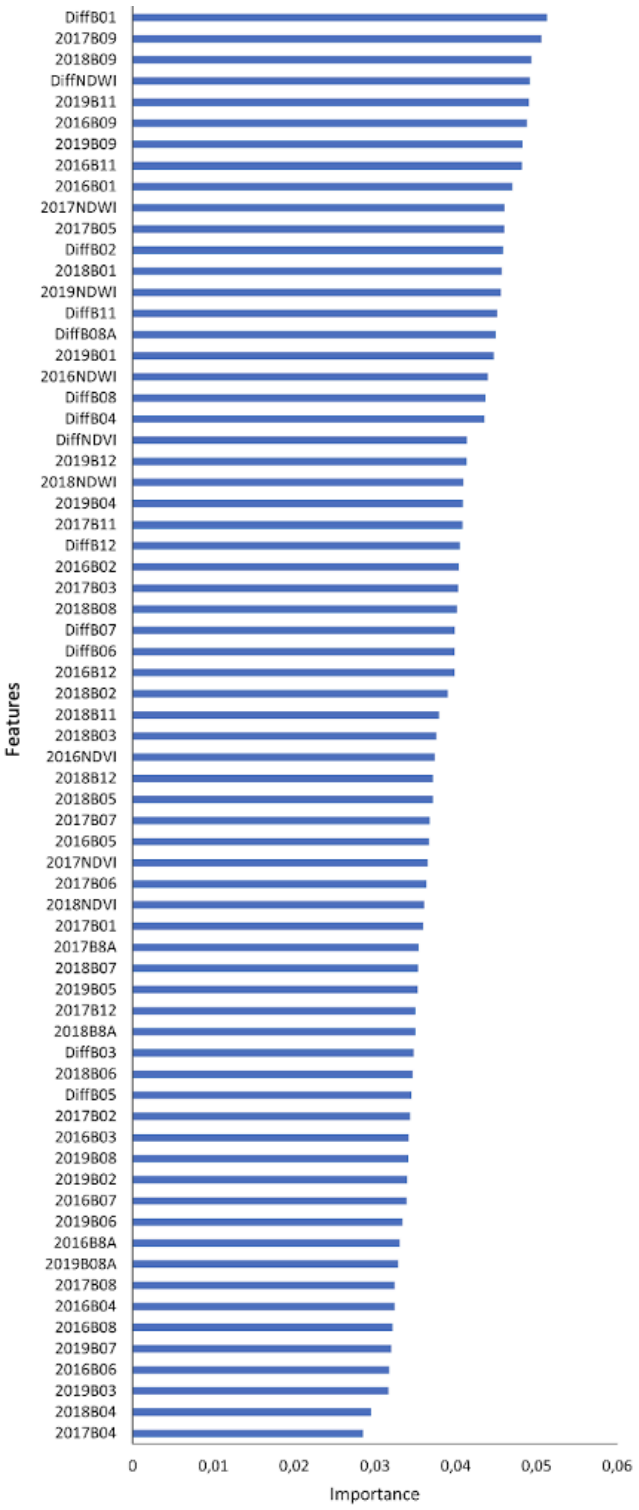


Fig. 2.4) Ranking of the 70 classification features based on Gini importance.

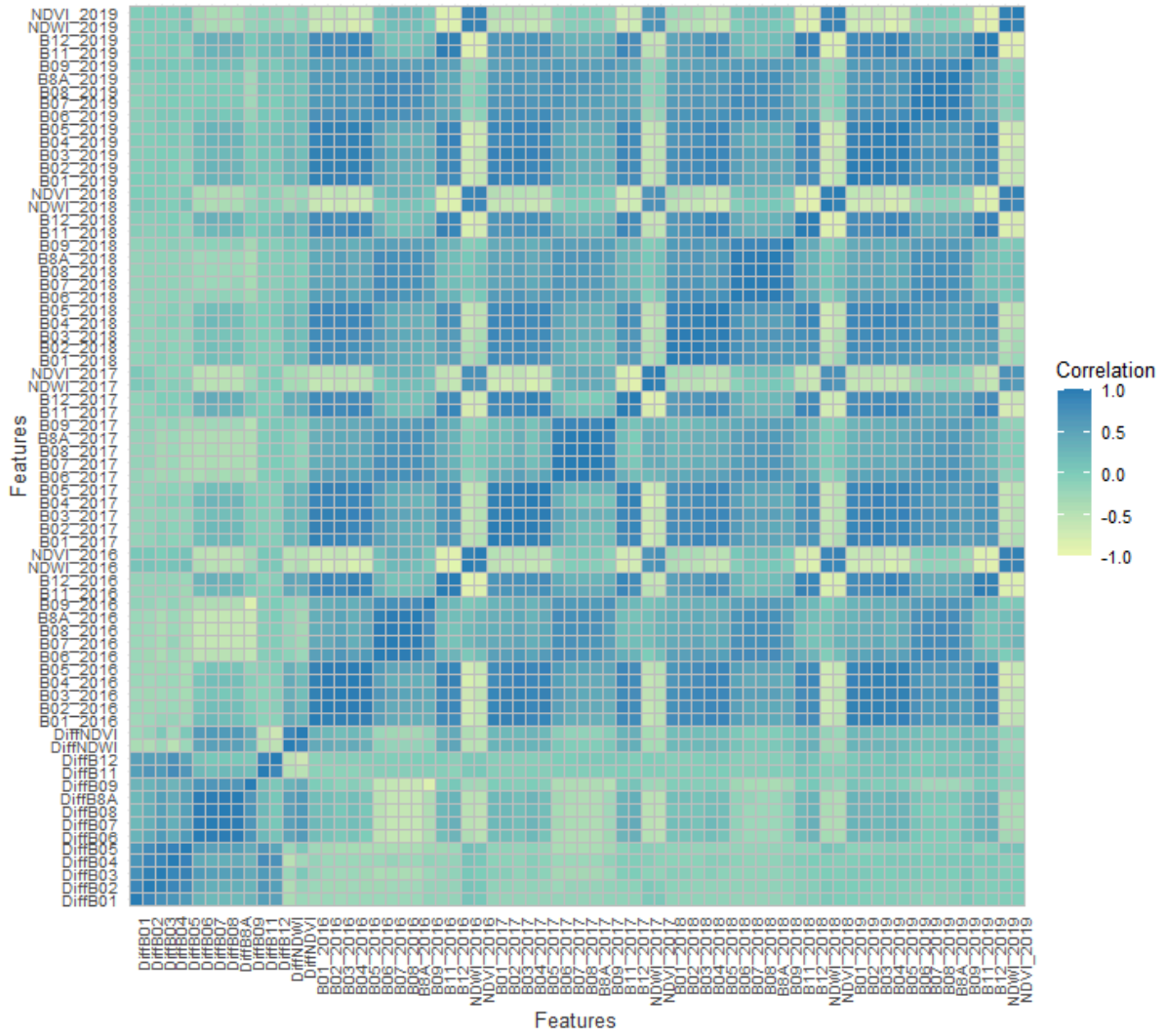


Fig. 2.5) Pairwise Pearson's correlation between the 70 features used in the classification.

2.3.3 Change classification

The spatial distribution of each of the classes is depicted in Fig. 2.6 a, Stable Thicket were distributed across the study area, but were most apparent on the eastern and western portions of the study area. Stable Forests were mostly in the southernmost parts of the study area. The Stable non-Forest/Thicket class was distributed across the study area, as was the Thicket loss class, but obviously over a much smaller extent. The Thicket gain class was most apparent in the valleys, north of Alexandria. Both the forest gain and forest loss classes occurred in the southernmost parts of the study area, where the Stable Forests are found. The Stable Thicket class was the most widespread and abundant, occupying 589.21 km^2 (45%) of the study area (Fig.

2.7 b). This was followed by the Stable non-Forest/Thicket and Stable Forest classes at 44% and 10%, respectively (Fig. 2.7 b). The classes representing gains or losses of Thicket or forest represented the smallest proportions of the study area (a total of 1%), but within these the most abundant was the Thicket Loss class at 0.15% of the study area (Fig. 2.7 b). The Forest Gain class occupied the least surface area (0.11%; Fig. 2.7 b). An online application to visualise and interpret land cover dynamics in Alexandria can be found at this link.

https://craigmahlasi.users.earthengine.app/view/alexandria-2016-2016-final-classification

a

b

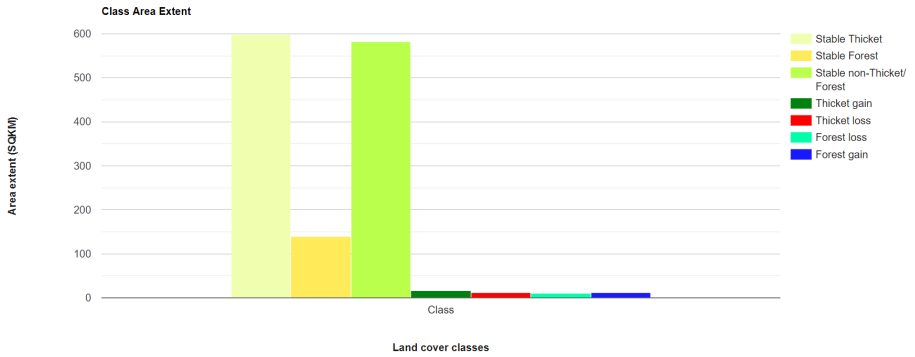


Fig. 2.6) (a) a visualisation of the RF change classification. The map shows the spatial distribution of the classes. b) a barplot of the class-wise areal extents in km^2.

2.3.4 Extracting and dating Albany Thicket loss

Fig. 2.7 d depicts the spatial distribution of Albany Thicket loss events between 2016 and 2019 extracted from the final classification. Of the 6.75 km^2 classified as lost Thicket, 5.6 km^2 (83%) were found to be true positives. The false positives totalled to 1.15 km^2. Clearing events occurred throughout the year with the most clearing taking place between July and November. Only one Albany Thicket loss event was detected in 2016 (Fig. 2.7 a) and most of the loss events occurred in 2018 (67). In 2017, 55 Albany Thicket loss instances were detected. The months between July and November saw the Fig. 2.7 a and b), show the temporal distribution of clearing events. The average duration of clearing was 20 days. Seven of the detected loss events occurred over a period greater than 75 days. There seems to be a spatial concentration of clearing events in the middle of the map in 2017 and 2018.

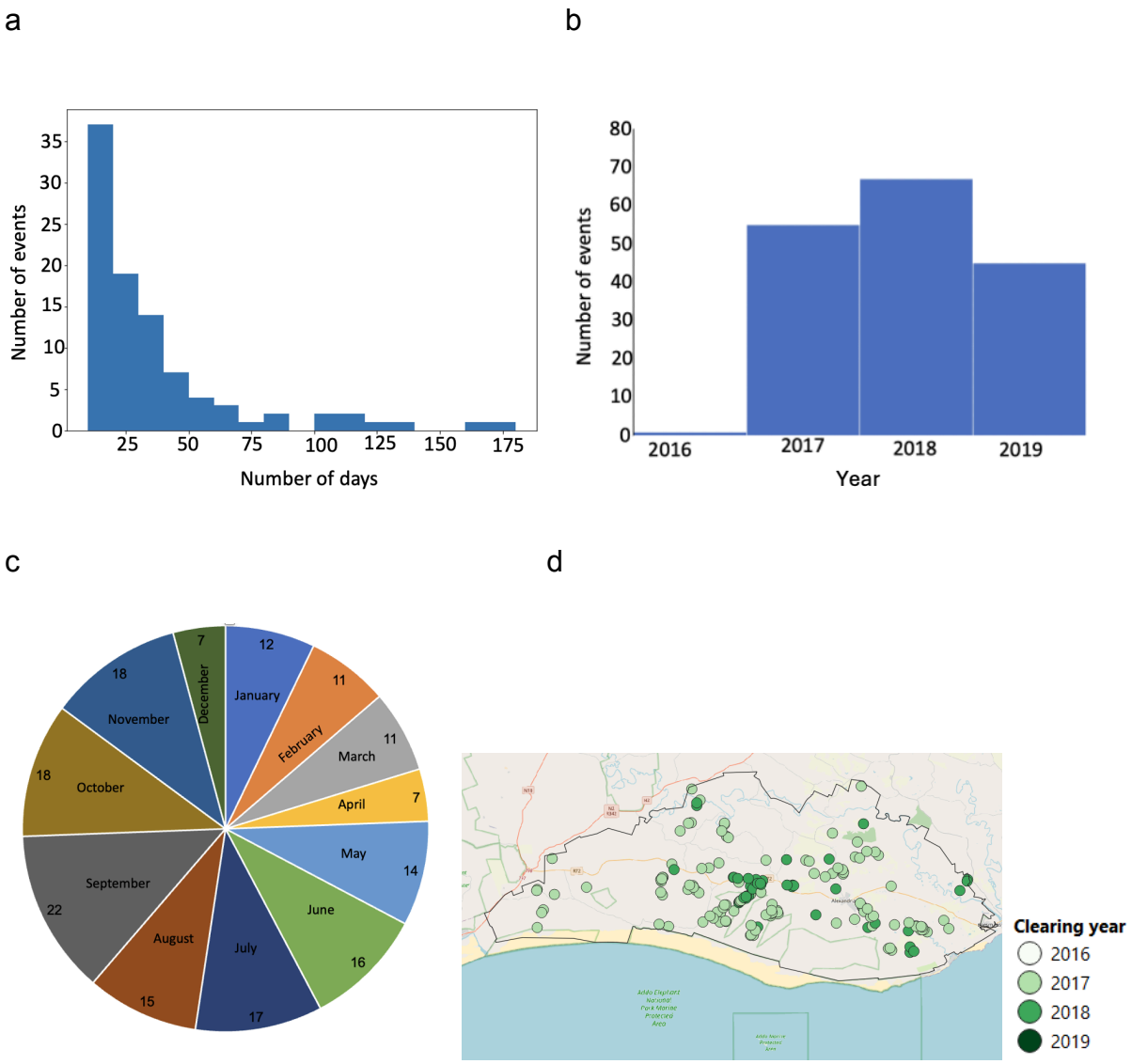


Fig. 2.7) (a) histogram of the duration (days) in which the Albany Thicket loss events occur. b) the temporal distribution, showing the annual counts of clearing events. c) the temporal distribution of clearing by month. d) map of the spatial distribution of the loss events and the years in which they occurred.

2.4 Discussion

2.4.1 Change classification

Temporal variations in atmospheric conditions, surface moisture, and illumination make land cover CD a complex problem in remote sensing. This is exacerbated in

regions that exhibit high vegetation diversity, such as Albany Thicket environments. In this study, we developed a differencing-classification hybrid CD protocol. The results demonstrate that the use of a 4-year set of Sentinel-2 images captures the latent change trends and trajectories. A source of false positives in CD studies based on bi-temporal image sets are pixels where values exceed change thresholds because of temporal differences in sensor-noise, atmospheric conditions, and illumination rather than changes in surface material (Hermosilla et al., 2015). By using a 4-year set (2016-2019) of Sentinel-2 images, temporal noise is resolved through the establishment of an expected range in which stable pixels fluctuate. The adopted method makes the training process efficient and less laborious because only one training set is created, and not multiple sets, as is the case in a post-classification protocol.

The RF classifier can classify changing and stable pixels with high accuracy (Fig. 2.7). The results indicate ongoing change in the Alexandria district. The classification shows that 589.21 km^2 of the study area was occupied by Stable Thicket, 129.56 km^2 by Stable Forests and 572.87 km^2 was classified as Stable non-Thicket/forest. The results show that 6.75 km^2 of Albany Thicket had been lost, gains in Albany Thicket amounted to 1.94 km^2, 0.07 km^2 of forest had been lost and 1.51 km^2 had been lost. The classification results confirm the robustness of the RF classifier against label noise. Label noise refers to discrepancies in the assigning of labels of a dataset. This robustness is a result of using bootstrap samples during the construction of the decision trees (Breiman, 2001; Agjee et al., 2018; Maas et al., 2019). It should be noted that some misclassifications were present. We suspect that the source of these misclassifications was label noise. Some label noise was always expected due to the complexity of Albany Thicket environments. Pixels occupied by mosaics and degraded thicket are very likely to be misclassified. This means that spectrally mixed pixels are much harder to classify by the RF classifier.

2.4.2 Accuracy assessment

The error matrix in Fig. 2.7 shows how robust the RF classifier was in classifying change in Alexandria. Comparable results were reported in other studies (e.g., Sedano et al., 2005; Nyamugama et al., 2015; Baen et al., 2016). Both the producer's and user's accuracies were high for all the classes. The lowest user's

accuracy was observed in the Thicket Gain class. A close look at the error matrix shows confusion between the Stable Thicket and Thicket Gain. This suggests that there is little spectral difference between Stable Thicket and regenerating Thicket.

A likely explanation for the misclassification of Thicket Loss as Stable Thicket or Stable non-Forest/Thicket as forest is the possibility of mixed pixels. This occurs where a pixel is partly cleared and partly vegetated. This is a long-noted limitation of medium and low-resolution imagery (Chen et al., 2003; Suwanprasit and Srichai, 2012). Another explanation may be the adjacency effect where stray (scattered by atmosphere) reflected light comes into the field of view of neighbouring pixels, which results in the fogging of signals and can result in misclassifications (Sterckx et al., 2011).

The spatial leave-one-out cross-validation implemented here yielded good accuracy when applied to assessing the predictive capability of the RF classifier. The rationale of implementing this method of cross-validation was to account for spatial autocorrelation by ensuring spatial-independence between the training and validation datasets. However, caution should be taken in the implementation of spatial leave-one-out cross-validation for two reasons: the first being the effect of small sample size and the rates of prevalence may result in high error estimates (Roberts et al., 2016). An example can be seen in the Thicket Gain class user accuracy. Due to the low prevalence of this class, there was little reference data in this class. Secondly, the spatial leave-one-out may lead to discrepancies in cases where there is an imbalance in sampling (Mahoney et al., 2022), as is the case in this study.

2.4.3 Variable importance

Bands from across the entire Sentinel-2 spectral range were noted to be influential in the classification. Pearson's correlation showed correlation between bands within images from the same year and across years. The high importance scores of the difference bands are attributed to their independence from the surface reflectance bands. Strobl et al. (2006) showed that the interchangeable use of correlated features in classifications results in the ranking of independent features as more important.

The elimination of low-ranking features and re-classification with the remaining features is common practice. This was not done in this study because feature importance scoring can carry bias, especially when features are highly correlated (Zhang and Yang 2020; Strobl et al. 2006). Gregorutti et al. (2014) referred to this as 'correlation bias'. A second limitation of the feature importance ranking is that it does not provide the optimal number of features to use in re-classification (Odindi et al., 2014). As a result, most studies resort to an arbitrary elimination of features.

Notwithstanding the complications that can arise from variable elimination, the 2016-2019 difference Coastal band at 443 nm (band 1) was shown to be most important in the classification. Coastal regions are known to be atmospherically complex and thus require robust aerosol correction. The 443 nm configuration makes the aerosol band sensitive to aerosol optical thickness, the ENVISAT MERIS Global Vegetation Index (MGVI) uses this wavelength to decontaminate the red and near-infrared bands from atmospheric aerosols (Gobron et al., 2004, European Space Agency, 2015). In the context of this study, the importance of the coastal aerosol band suggests that the source of some of the change detected is temporal variation in atmospheric aerosol content. Poursanidis et al. (2019), noted that the coastal aerosol band benefits coastal mapping tasks due to its greater water penetration ability and its sensitivity to atmospheric moisture such as fog. The second most important feature was the 2017 water vapour band at 945 nm (band 9). The water vapour band from 2016, 2018 and 2019 also ranked high. The water vapour band is commonly used for the estimation of atmospheric water vapour absorption and atmospheric correction. SWIR bands have been noted to be sensitive to subtle variations in tree vegetation density and vegetation moisture content. Band 11 was found to be an important feature; the band is located at around 1610 nm (Fig. 2.1), which is synonymous with water absorption. Oldeland et al. (2010) noted the utility of bands located in this region of the electromagnetic spectrum in vegetation discrimination in semi-arid areas. The utility of the NDVI has been confirmed in many studies including; Clark et al. (2005); Chan et al. (2008), Oldeland et al. (2010), Chrysafis et al. (2017); Zhang and Yang (2020); Bhattarai et al. (2021). However in this study lower importance of the NDVI is observed. The red-edge bands (band centre 842 nm) showed relatively weaker importance. The near-infrared (NIR) bands also showed weak importance. Similar performance of the NIR bands was reported

by Chrysafis et al. (2017) and Tasfaye and Awoke (2020), who attributed the weak performance to increased saturation in NIR light in moderate and dense canopies.

2.4.4 Extracting and dating Albany Thicket loss

Most of the clearing occurred in 2018 - this coincides with the driest year in the region noted by Archer et al. (2021). The map of the extracted and dated instances of Albany Thicket loss shows no temporal pattern; the events appear to occur sporadically throughout the year with the months between July and November seeing the most clearing. This suggests that the clearing is done to primarily create pasture lands for livestock rather than for crop cultivation, as cultivation for crops occurs during planting seasons. The seven instances that occurred over a period greater than 75 days were adjacent to plantations; this suggests that the cause of the clearing was for the expansion of plantation forestry.

2.4.5 Implications for Albany Thicket disturbance monitoring

The methods and results presented here demonstrated the utility of satellite imagery and remote sensing tools and products in monitoring changes in Albany Thicket ecosystems. Using freely available satellite imagery and open tools, I was able to map medium-term Albany Thicket disturbances in Alexandria with high accuracy. The methods employed here can be replicated with ease to reproduce comparable results in the study area or biophysically similar regions. The approach used here provides authorities with an accurate tool to help guide attempts to arrest ongoing clearing of Albany Thicket. The detected Albany Thicket disturbances were extracted and manually validated; these will be used as training data in the third and fourth chapters of this thesis which seeks to detect disturbances in Albany Thicket vegetation in near-real time.

2.5 Conclusion

Albany Thicket plays a crucial role in supporting biodiversity and providing ecosystem services. Nonetheless, it faces an ongoing threat of clearing. Efforts to spatially document clearing events have been made. However, existing methods are laborious, time consuming and require substantial financial inputs. In this study we

demonstrated the utility of a differencing-classification hybrid CD protocol for classifying change in Albany Thicket environments. The RF classifier was implemented on a 4-year set (2016-2019) Sentinel-2 images consisting of surface reflectance and difference bands. The results showed that the combination of Sentinel-2 imagery and the RF classifier performed well in detecting and classifying disturbance in Albany Thicket vegetation. The results indicate the presence of spectral differences between Albany Thicket and other cover classes, most importantly non-Albany Thicket vegetation. This study provides a basis for Albany Thicket monitoring using satellite imagery. The robustness of the spatial leave-one-out cross-validation method was also confirmed here. The exploration of variable elimination protocols that account for correlation in variables is recommended. While this study confirms the utility of Earth observation data in detecting and classifying disturbance in Albany Thicket environments, there remains a need to investigate the utility of supervised continuous change detection in the Albany Thicket of South Africa. The next chapter focuses on the application of a time-series supervised classification algorithm to detect changes in the Albany Thicket biome.

Chapter 3: Temporal Convolution Neural Networks for continuous change detection in the Albany Thicket biome

3.1 Introduction

Habitat conversion is the primary driver of biodiversity loss, with 75% of terrestrial ecosystems estimated to be severely modified (IPBES Global Assessment on Biodiversity and Ecosystem Services, 2019; IPCC, 2019). Agriculture is widely regarded as the leading cause of habitat conversion (Dudley and Alexander, 2017; IPBES Global Assessment on Biodiversity and Ecosystem Services, 2019). At present, 40% of the Earth's terrestrial surface area is used for agricultural activities, and this figure. continues to grow (IPBES Global Assessment on Biodiversity and Ecosystem Services, 2019). The continued expansion of agricultural lands will mean further loss of natural vegetation cover (Azeb et al., 2018). The expansion of agricultural activities is driven by increasing populations, which grow markets that require servicing and supplying. This is balanced by advances in agricultural technology including pest and adverse climate resistant seeds which allow for the increased productivity of more lands (IPBES Global Assessment on Biodiversity and Ecosystem Services, 2021). Earth observation tools are important for detecting and providing information on vegetation dynamics, including clearing events. Advances in satellite spatial and temporal resolutions coupled with growing satellite constellations have enabled the development of tools for near-real time retrieval of vegetation clearing events, and include examples such as Global Forest Watch (https://www.globalforestwatch.org/), Terra-i (http://www.terra-i.org/terra-i) and Airbus Starling (https://intelligence.airbus.com/industries/forest-and-environment/starling/). Global Forest Watch provides near real-time forest monitoring and Terra-i provides near real-time detection of land-cover with 16-day updates. Airbus Starling detects forest cover change and flags alerts at quarterly intervals. The system provides information such as the spatial extent of the losses in forest cover.

Because of their role in the sequestration of atmospheric carbon and high biodiversity, forested ecosystems have received most of the research attention in developing Earth observation-based monitoring tools and this has resulted in non-forested ecosystems being largely ignored (Moncrieff, 2021). Existing methods for near real-time change detection have been developed primarily and exclusively for forested ecosystems. In their basic form, near real-time change detection methods timeously ingest, process, and perform change detection on newly available satellite imagery. These methods are collectively termed online change detection (Aminikhanghahi and Cook 2017; Ye et al., 2021; Moncrieff, 2021). Examples of online change detection approaches include the Breaks for Additive Season and Trend Monitor (BFAST Monitor) algorithm (Verbesselt et al., 2012), Continuous Change Detection and Classification (CCDC) (Zhu and Woodcock, 2014), Continuous Monitoring of Land Disturbance (COLD) (Zhu et al., 2020) and Stochastic Continuous Change Detection S-CCD (Ye et al., 2021). These algorithms decompose pixel time series into trend, season and noise elements and detect breaks in trends. The main limitation of these methods is that change is only confirmed when the value of a pixel falls outside the fitted trend after a number of consecutive observations. In the case of CCDC, three consecutive Landsat observations that fall outside the fitted values are needed for change to be confirmed (Zhu et al., 2020). For Landsat, with a revisit rate of 16 days, this means that disturbances are only confirmed at a minimum of 48 days after onset of clearing.

Despite advances in Earth observation and near-real time monitoring tools, cross-biome applicability remains a major limitation, where tools developed for specific biomes tend to perform poorly when applied to other environments. The discrepancy in performance across biomes is a result of differences in plant structural and optical properties, which in turn result in varying spectral signatures (Noda et al., 2021), in addition to variable temporal dynamics due to phenology, disturbance regimes, and sensitivity to rainfall (Slingsby et al. 2020). While Albany Thicket is classified as closed ecosystems like forests, they exhibit far less sub-canopy complexity and greater upper canopy structural complexity (Mucina and Rutherford, 2006; Smith et al., 2017; Cowling and Hoffman, 2021). Change detection in forests is far more straightforward than in non-forested systems because undisturbed forests are more spectrally distinct from the cleared parcels that result

from deforestation (Hu et al., 2022). Change detection in non-forested ecosystems is more difficult due to several factors, including variable reflectance from different plant growth forms, low signal-to-noise ratios for vegetation in these systems, high levels of background reflectance from soils, the existence of photosynthetically active soils, and greater spatial and temporal heterogeneity in evapotranspiration and gross primary productivity (Smith et al., 2019). Natural disturbances, such as fires, storms, frost burn, drought, floods, rockfalls and pest infestations result in either abrupt or gradual disturbances in the structure and health of a habitat (Lingua et al., 2021). The duration these events take to manifest in changes in spectral signatures can make and cover change detection difficult (Coops et al., 2020). These factors can make cross-biome application of change detection tools difficult, because they result in discrepancies in the definition of change thresholds (Rosin and Hervás 2005; Huang et al., 2015; Smith et al., 2019).

The Albany Thicket biome is described as low (2-5 m), impenetrable, semi-succulent, spinescent, evergreen, thorny shrubs (Mucina and Rutherford, 2006; Parker, 2017). In South Africa, Albany Thicket exists on a spectrum of aridity (Smart, 2016), with vegetation units showing great variation between river catchments (Knight and Cowling, 2003). The Albany Thicket is a floristically diverse system which supports a range of fauna including 74 mammal species, 307 species of birds (in three internationally recognised Important Bird Areas), 61 reptile and amphibian species of which six are endemic to the Albany Thicket and a diverse range of invertebrate fauna (Knight and Cowling, 2003). However, this biome has experienced extensive degradation (Carvalho et al., 2021). It is estimated that approximately only 11% of the Albany Thicket Biome remains in a pristine state and 60-63% is classified as severely degraded (Mucina and Rutherford, 2006; Powell, 2009; Stickler and Shackleton, 2015; Mills et al., 2018). The degradation and destruction of Albany Thicket is primarily as a result of pastoralism and agricultural land expansion, with large tracts of Albany Thicket being cleared to make way for crops or degraded by overgrazing (Hoffman et al., 1987; Lloyd et al., 2002; Stickler and Shackleton, 2015; Mills et al., 2018; Carvalho et al., 2021).

The change of Albany Thicket to other land cover types leads to structural and functional changes and species loss, which can result in negative socioeconomic and environmental consequences (Powell, 2009; Hoffman et al., 1987). Albany Thicket can store up to 20 kg/m^2 of carbon (in above- and belowground pools), which is equivalent to mesic forests (Mills and Cowling, 2006). Despite conservation and restoration efforts, Albany Thicket remains under threat of clearing. In South Africa, the National Environmental Management Act (NEMBA) 10 of 2004, states that the disturbance of topsoil and removal of indigenous vegetation occurring over a parcel of land undisturbed for ten years and longer constitutes a crime. There is therefore a need for timeous up-to-date information on clearing and disturbance events if the ongoing clearing of Albany Thicket vegetation is to be halted. Such information could aid the authorities in speedily identifying and documenting changes and enforcing the law. The existence of such information and systems may act as a deterrent to would-be perpetrators.

There is a lack of Earth observation tools that facilitate near-real time change detection in many of South Africa's ecosystems including the Albany Thicket. The current techniques employed are inefficient, laborious, and time-consuming and include visual inspection of bi-temporal aerial and satellite imagery for change followed by manual digitisation (Showalter et al., 2010). Automated classification algorithms have also been used to produce annual land cover maps for multiple years which then become the basis of comparison to assess change (Showalter et al., 2010; Giri el al., 2015). Carvalho et al. (2021), used Maximum Likelihood Classification to assess degradation in the Albany Thicket biome between 2016 and 2018. Kakembo et al. (2015), performed post-classification change detection using Maximum Likelihood Classification to assess elephant induced degradation in the Addo Elephant National Park between 1973 and 2010. While the aforementioned studies were able to identify long-term change, their low temporal resolution hinders the detection of in-progress clearing events. Increasing the frequency of change assessments to satellite revisit times and reducing the lag between successive assessments would improve their utility in timeously detecting changes and thus allowing authorities to respond to change alerts. One approach to addressing this limitation is through supervised time series classification.

Supervised classification methods have been widely used to perform land cover change detection (Khayyam and Waseem, 2021; Magnússon et al., 2021; Moncrieff, 2022; Al-Dousari el al., 2023). Supervised classification methods are trained on a corpus of labelled land cover classes of interest or examples of instances of land cover change (Merry et al., 2023). These models are used to detect changes in succeeding, unlabelled images (Sencaki et al., 2023). Supervised machine learning and deep learning methods are effective in mapping land cover because of their ability to deal with the complex spectral and spatial relationships captured by remote sensing images and their ability to better extract abstract and semantic information.

Supervised land cover change detection methods allow for a model to focus on the change or classes of interest and to be robust against Internal natural variability and natural disturbances. Internal natural variability in factors including differences in background reflectance from soils, variances in photosynthetically active soils, and heterogeneity in evapotranspiration and surface moisture, (Smith et al., 2019). Natural disturbances, including fires, floods, storms, frost burn, pest infestations, rockfalls and pest infestations that can result in either abrupt or gradual changes in the structure and health of a habitat (Lingua et al., 2021). While these factors may manifest in deviations from the class thresholds they are not always changes in land cover (Rosin and Hervás 2005; Huang et al., 2015; Smith et al., 2019).

Most supervised land cover change detection algorithms rely on bi-temporal change detection, which is based on assessing change between two images - one of the initial states and one of the post-change states (Gawlikowski et al., 2022; Nyborg et al., 2022). While the implementation of bi-temporal change detection is straightforward, it suffers from two main limitations: the first is the propensity for false positives resulting from differences in atmospheric conditions, viewing angles and illumination and soil moistures between the two images; the second limitation is that the observation interval between the initial and post-change or successive observations are often weeks or up to years apart (Almutairi and Warner, 2010; Lu et al., 2010; Congalton et al., 2014; Mishra et al., 2017; Alonso et al., 2023). which makes the detection of transient changes difficult.

Continuous change detection uses all available and usable observations to detect changes. Continuous change detection classifiers allocate pixels throughout a time

series to predefined classes using labelled training data. Alonzo et al. (2021) used a Random Forest classifier (Breiman, 2001) on time-series Landsat and Sentinel-2 data to map and quantify land cover dynamics. The use of all available observations allows for the disentanglement of trends and seasonality as well as internal natural variability and natural disturbances. The use of time series reduces biases that result from the omission of seasonal variations when using bi-temporal algorithms (Xian et al., 2022). A few studies have investigated the implementation of neural networks on time series data for continuous change detection (Gondwe et al., 2021, Moncrieff, 2022, Al-Dousari et al., 2023, Chen et al., 2023), although continuous change detection has not been attempted in the Albany Thicket of South Africa.

I postulate that continuous supervised models such as the TempCNN are better suited for performing change detection in open canopy ecosystems because of their ability in dealing with spectral variability and subtle changes in vegetation cover that are characteristic of these environments. Having training data that adequately capture the spectral variability and subtle changes is the main key behind their suitability of detecting change in the Albany Thicket. This study investigates the utility of two methods in detecting changes in The Albany Thicket Biome. The aim was to develop and implement a protocol for detecting clearings in Albany Thicket with low latency. Using an existing dataset of land cover change events developed in Chapter 2 of this thesis, machine learning classifiers were trained to predict near-real time land cover change by taking advantage of new satellite data as it becomes available. The ultimate goal of this study is to test existing methods in a non-forest ecosystem, using a statistical method mostly used for forests (CCDC) and another more flexible machine learning method that has been used in other non-forest ecosystems (TempCNN). To this end The efficacy of TempCNN and CCDC in detecting changes in The Albany Thicket Biome are compared. Directly comparing TempCNN and CCDC in terms of their algorithmic performance would be unfair due to their inherent differences. Instead, the CCDC algorithm is used in this study as a widely recognised baseline state of the art change detection method. This approach provides a benchmark against which TempCNN's novel method can be assessed

3.2 Methods

3.2.1 Data

This study builds on the second chapter of this thesis which generated a land cover change dataset documenting the changes in Alexandria between 2016 and 2019. The contemporary extent of Albany Thicket, representing 589.2 km^2 of the study area, was extracted from the results of Chapter 2 in which a land cover change map of Alexandria was developed for the period 2016 to 2019. The land cover map from Chapter 2 contained 186 instances of Albany Thicket loss, which represent areas where Albany Thicket was changed to another land cover class, typically because of vegetation clearing. Using daily PlanetScope imagery from Planet (https://www.planet.com/), the boundaries of clearing instances were manually defined as polygons. Polygons were assigned a date when Albany Thicket was last confirmed to have been present as well as a date for which the entire polygon was confirmed to have been cleared. Using PlanetScope imagery allowed for the precise allocation of dates and boundary delineation.

3.2.1.1 Satellite data

Multispectral, Level-2A (Top of Atmosphere (TOA) reflectance) Sentinel-2 data were used to generate image time series. Sentinel-2A and B are a pair of multispectral imaging radiometers launched in 2015 and 2017 by the European Space Agency (https://sentinel.esa.int/web/sentinel/missions/sentinel-2). The data were accessed from Google Earth Engine (https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED).The spectral and spatial resolution are shown in Fig. 2. The study area falls within tile 35HMC of the US-Military Grid Reference System. Between 2016 and 2019 a total of 529 images were collected by both Sentinel-2A and B from this tile. The tile revisit time for each satellite is 10 days, i.e, the revisit time is reduced to 5 days due to the orbital path overlap between the two satellites. Preprocessing of the image time series included cloud filtering where images with a total scene cloud cover greater than 60% were excluded from the time series. After the cloud filtering, 360 images remained. The *s2cloudless* algorithm was used to perform cloud and shadow masking within images. Pixels within images with a cloud probability greater than 0.5 were masked. The algorithm also masks

shadows by determining cloud projection intersection with low-reflectance near-infrared (Band 8 pixels) (https://github.com/sentinel-hub/sentinel2-cloud-detector). Temporal gap filling was performed to fill no-data pixels that result from the cloud and shadow masking. This was achieved by replacing masked pixels with the values of a matching unmasked pixel from the temporally nearest preceding image. The temporal gap filling assumes stationarity or no change in the reflectance values across all bands in the no-data pixels, the no-data pixels are replaced with the most recent valid pixel. Gap filling strategy seeks to ensure that an observation for pixels is available throughout the observation period, while not injecting synthetic data.

Normalised Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI), Normalised Difference Water Index NDWI, and Moisture Stress Index (MSI) were computed for each S2 image in the time series. Since the image bands come at 3 different spatial resolutions, they were resample to a 30m grid. Each image was clipped to the extent of currently standing Albany Thicket and pre-identified clearing instances. The full time series of 360 images was then subsampled to windows with a fixed length of 20 images (100 days) this is because the Sentinel-2 constellation has a 5-day latency between two successive images. This equates to 100 samples from each full time series per pixel when using a step size of three, see Fig. 3.

3.2.1.2 Reference data

Reference data were extracted from the Albany Thicket loss instances and Stable Albany Thicket datasets. At the individual pixel level, In total there were 1879 cleared Albany Thicket samples and 1927 Stable Albany Thicket individual pixel samples. The random sampling method was used to extract samples. The time series of each of these pixels was then sampled from pixels within these polygons and subsampled into partially overlapping windows as described above. The sampling was performed in a way that prevents training and validation samples from coming from the same clearing instances/ polygons by subsampling each polygon. These data were then split for training, validation and testing using a 65/5/30 split. Each time series was labelled following a protocol adopted from Moncrieff (2021). The label “Stable Thicket” was allocated to a time series’ subsample if the date when Albany Thicket

was last confirmed to be present occurs later than the end date of the series under consideration, or if no change is reported to occur. Where both the date when Albany Thicket was last confirmed to be present and the date when Albany Thicket was first confirmed to be lost are within the same subsample the label “cleared” is allocated see Fig. 3.1.

3.2.2 Models

A comparison between two continuous change detection models: CCDC (Zhu and Woodcock, 2014) and Temporal Convolution Neural Networks (TempCNN; Pelletier et al., 2019) is made, to assess their performance in detecting changes in Albany Thicket cover.

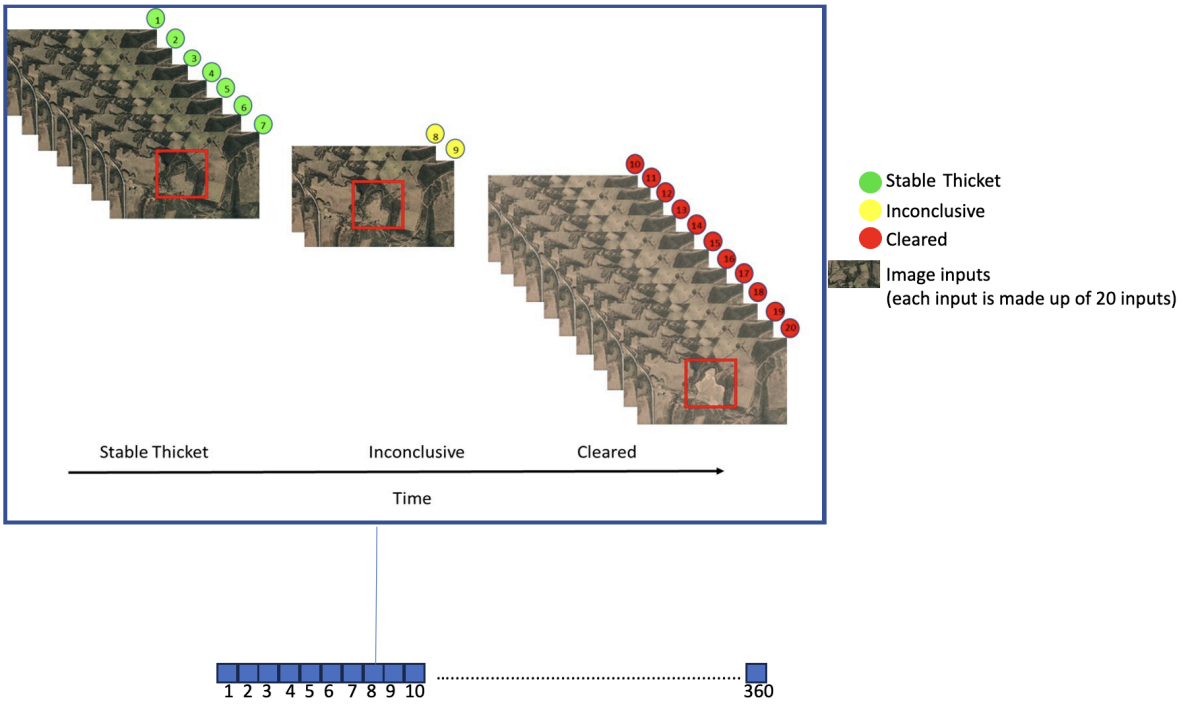


Fig. 3.1) One zoomed in example of the time-series labelling protocol. The time-series was subsampled to windows with a fixed length of 20 observations with a five-day update and a step size of three. One of three labels is possible for each pixel’s time-series subsample: Stable Thicket, inconclusive and Cleared Thicket. Labels were allocated to the subsamples by inspecting Planet Labs PlanetScope imagery. In this example the pixels within the red bounding box are labelled Stable

Thicket up till step 7, at step 8 and 9 it is uncleared (inconclusive) and from step 10 it is evident that some of the pixels within the bounding box are cleared.

3.2.2.1 Continuous Change Detection and Classification

A Google Earth Engine implementation of CCDC developed by Arévalo et al., (2020) is used here as a baseline model to detect Thicket cover change in the Alexandria region. CCDC fits a harmonic regression model to each pixel across the time series. CCDC uses all available images within an archive to detect land cover change. After masking clouds and shadows, Robust Iteratively Reweighted Least Squares (RIRLS) was used to model the per pixel inter- and intra-annual seasonality using a sum of periodic functions. The fitted model was used to compare new observations to predicted observations for each band. The difference is then normalised by three times the RMSE. To detect changes in the times series, newly available observations are added to the time series. Pixels are flagged as have changed If the normalised difference in the newly added observations within a moving window period is greater than one for all bands.

3.3.2.2 Temporal Convolutional Neural Networks

TempCCN is a 1-dimensional CNN used in the temporal domain to extract multivariate temporal structures from multivariate time series satellite imagery (Pelletier et al., 2019; Zhu et al., 2021). The main distinguishing factor between the TempCNN applied here and the generic TempCNN is that Spectro-temporal guidance is implemented in this TempCNN. Spectro-temporal guidance refers to the incorporation of convolutions in both the spectral and temporal dimensions (Hu, et al., 2019, Pelletier et al., 2019). The multivariate input consists of Sentinel-2 spectral bands as well as spectral indices including NDVI, NDWI, SAVI, EVI and MSI. The TempCNN comprises three convolutional layers with convolution filters of the same size. The convolutional layers with 32 units are successively applied. Each convolutional layer is followed by an averaging pooling layer that reduces the dimensions of the feature space. Neurons from the first layer receive a multivariate time series as input and extract both multi-spectral information and temporal dynamics. The three convolutional layers are followed by one dense layer with 256 neurons. Finally, a Softmax layer is derived, which produces the distribution of the

predicting class (Fig. 3.2). The network was trained using Adam optimization with a batch size of 32 and a learning rate of 0.0001. Regularisation mechanisms were implemented including a dropout, L_2-regularisation on the weights were specified as measures against overfitting. Several training iterations were performed on different dropout rates and L_2-regularisation factors. A dropout rate of 0.8, an L_2-regularisation on the weights with a regularisation factor of 1e-8, were finally specified as measures against overfitting. Focal loss was used for the loss function. Model evaluation was performed on the validation dataset. An accuracy assessment was performed on the test dataset. The following metrics were used to evaluate the model: overall accuracy, the F1 score, precision score, and recall score. All preprocessing was performed using Google Earth Engine while TensorFlow 2 (https://www.tensorflow.org/api_docs/python/tf) and eo-flow (https://github.com/sentinel-hub/eo-flow) were used to implement the TempCNN.

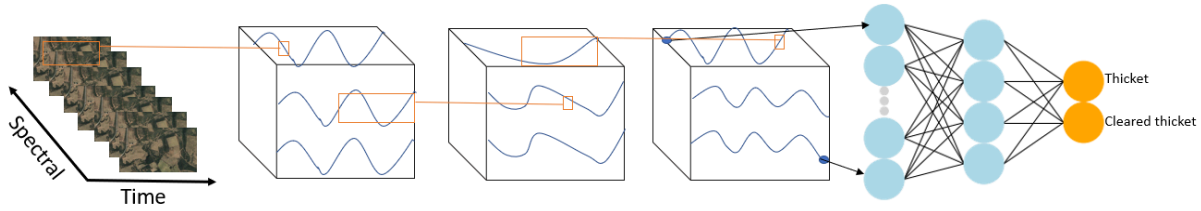


Fig. 3.2) The Temporal Convolutional Neural Network (TempCNN) architecture. The input to the network is a multivariate time series of Sentinel-2 spectral bands and indices from a pixel. Three convolutional filters are applied in succession, a single dense layer, and a Softmax layer.

$$\text{softmax}(x_i) = \frac{exp(x_i)}{\Sigma_j exp(x_j)} \qquad \text{equation(3.1)}$$

where

x_i is the ith element of the input vector

$exp(x_i)$ is the exponential function applied to x_i

$\Sigma_j exp(x_j)$ is the sum of the exponential of all elements in the input vector

3.2.3 Predictions

To test the performance of both the CCDC and TempCNN, the constructed CCDC and TempCNN models perform pixel-wise classification on input time series images. The features learned from the training phase are then used to classify stable Albany Thicket and cleared Albany Thicket on the testing data. A back-looking temporal filter was implemented for the TempCNN workflow. The filter retrieves the present day's date and creates a time series of 20 observations of the most recent Sentinel 2 images that meet the cloud filtering criteria within the past 12 months. Predictions are performed on the stack to produce the latest change detection using the constructed model.

3.2.4 Influence of aridity on performance

Across the study area, rainfall ranged from 323 to 474 mm (Funk et al., 2015) during the training period (2016-2019). Aridity contributes to the spectral complexity described above. This spectral complexity introduces discrepancies in land cover classification exercises and is likely to degrade the performance of any change detection algorithm. To ameliorate the impact of environmental variation, some researchers have incorporated biophysical information as classification features, as has been done in this study through the inclusion of the MSI as a feature (Table 3). The availability of moisture is thought to determine the rate of aridity (Sage and Sage 2013). The effect of the MSI on performance was evaluated by excluding the MSI features from the observations and ran a second time. It should be noted that the Bullock and Arevalo) 2020, incorporate rainfall and temperature data in their implementation of the CCDC.

3.2.5 Influence of training sample size

The availability of sufficient training data often hampers the successful implementation of supervised methods. I examined the influence of training sample size by training and testing models using 100 (3806 samples), 75 (2854 samples) and 50 (1903 samples) percent of the reference data and compared the results.

3.2.6 Model Evaluation

To evaluate the performance of the models, several evaluation metrics were produced, including overall accuracy, precision, recall and F1 score.

The overall accuracy measures the percentage of correctly classified pixels.

Overall accuracy = $\frac{P_a}{t_a}$ equation (3.2)

where

P_a is the total number of correctly classified pixels, t_a is the total number of pixels in the testing set.

F1 Score = $\frac{2*(Precision * Recall)}{(Precision * Recall)}$ equation (3.3)

Precision measures the ratio of correctly predicted pixels to all predicted pixels for specific class:

Precision = $\frac{tp}{tp+fp}$ equation (3.4)

Recall = $\frac{tp}{tp+fn}$ equation (3.5)

Where

tp , fp and fn are true positive, false positive and false negative, respectively.

F1 Score = $\frac{2*(Precision * Recall)}{(Precision * Recall)}$ equation (3.6)

Where

Precision is the ratio of correctly predicted pixels to all predicted pixels for specific class

Recall is the ratio of positively predicted pixels to the total number of positively classified pixels

3.3 Results

3.3.1 Change Detection

Table 3 shows results from the three models. The CCDC outperforms the initial TempCNN, however the TempCNN-MSI showed the best performance when compared to the TempCNN and CCDC, with an accuracy of 0.97, F-score of 0.96, recall of 0.98 and precision of 0.95 when detecting changes in Albany Thicket cover, see Table 3. CCDC performed marginally poorer than TempCNN-MSI with an accuracy of 0.94, F1-score of 0.94, recall of 0.94 and precision of 0.94 when detecting changes in Albany Thicket cover. The higher recall value for cleared Albany Thicket for the TempCNN-MSI model compared to CCDC shows the former is more likely to detect Thicket clearing. What is noteworthy is that the Stable Thicket and Cleared Thicket recall scores are the same for the CCDC model, while the Cleared Thicket recall score was higher than the Stable Thicket recall in the TempCNN-MSI model. Fig. 3.3 shows the confusion matrices of the two models.

The primary objective of continuous change detection is the timeous flagging of change events. Fig. 3.6 illustrates an example of an Albany Thicket clearing event as well as the associated event-to-flag lag comparison between CCDC and TempCNN. The “event-to-flag lag” was calculated as the difference (in days) between the allocated date where thicket was first confirmed to be cleared and the date (Sentinel-2 acquisition date) i.e the date the last image in a prediction input was acquired. While the per-pixel probability of clearing at a specific date is reported, the full boundary and not individual pixels were considered when calculating the “event-to-flag lag”. According to manual interpretation of PlanetScope imagery, the change (loss of Thicket) occurred between 15 and 29 September 2019. Noteworthy is that TempCNN detected Cleared Thicket, while CCDC was less able to do so this can be seen in the 50 versus 49% of true positives observed in both TempCNN and

CCDC respectively. Table 4 shows that the performance of both the CCDC and TempCNN benefit from adding more observations. Overall, the TempCNN achieves better performance in both the recall and precision metrics. However, a plateau in both recall and precision is observed in the CCDC model after 30 days, as the recall and precision do not change when detections are made between 30 and 90 days.

Table. 3. Performance of the models based on test data. Highlighted in bold is the best performing model. Here st denotes Stable Thicket and ct denotes Cleared Thicket.

Model	Accuracy	F1	Recall-st	Recall-ct	Precision-st	Precision-ct
CCDC	0.94	0.94	0.94	0.94	0.95	0.94
TempCNN	0.92	0.91	0.93	0.90	0.93	0.93
TempCNN-MSI	**0.97**	**0.96**	**0.95**	**0.98**	**0.98**	**0.95**

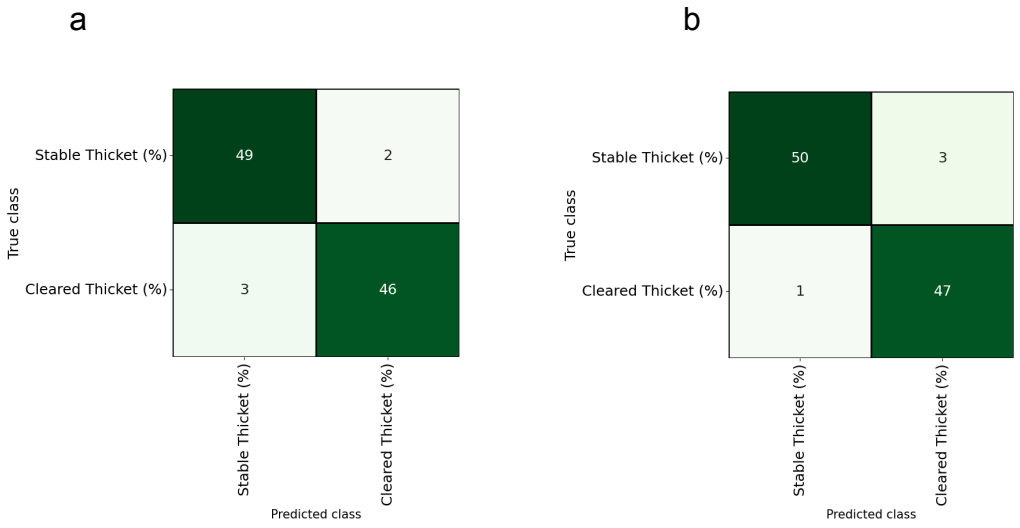


Fig. 3.3 a) confusion matrix of the CCDC prediction, b) confusion matrix of the TempCNN prediction.

3.3.2 Influence of aridity on performance

The TempCNN's ability to detect Albany Thicket clearing declined when the MSI index was removed as a feature in the TempCNN model (Table 3). The model resulted in an accuracy of 0.92, F-score of 0.91, recall of 0.90 and precision of 0.93 on the cleared Albany Thicket class. The TempCNN model is more likely to miss some detections when the MSI feature is excluded from the features for both cleared and Stable Thicket, as indicated by the lower recall (Table 3). The decision to add the MSI as a model feature was based on the presence of an aridity gradient in the study region. A 5% decline in detection accuracy and F1 score was observed when the MSI was removed from the TempCNN features. The recall score of the Cleared Thicket class declined from 0.98 to 0.90. This means that the ability of the TempCNN model to detect clearing events increased with the inclusion of the MSI. This was most apparent in drier Thicket (Fig. 3.5).

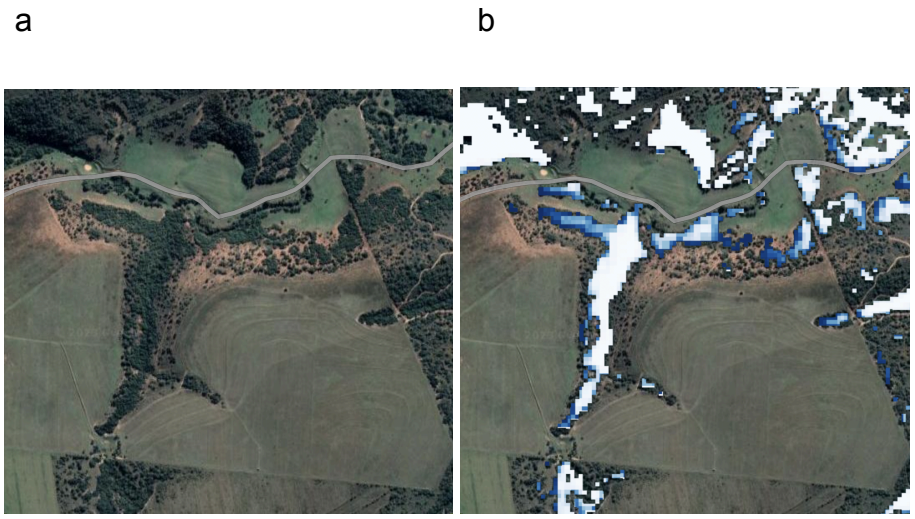


Fig. 3.4 a) RGB image of Thicket-barren land fringes, b) TempCNN classification. The two panels show an example of how the fringes between Albany Thicket vegetation and barren land were misclassified; the reason for this is likely a result of the adjacency effect.

a b c

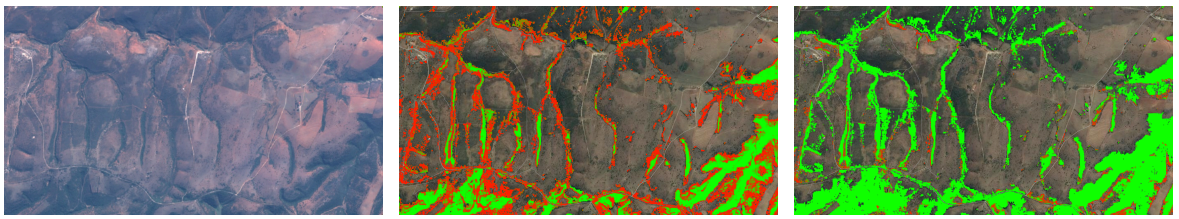


Fig. 3.5 a) RGB image of an area where there are differences between the TempCNN containing the Moisture Stress Index and the TempCNN without the Moisture Stress Index as a feature. b) An example of false positives in the Cleared Thicket class before adding the Moisture Stress Index. c) Example of how adding the MSI reduced the false positives observed in b) however they persist.

3.3.3 Influence of training sample size

Increasing sample size results in increased classification accuracy by TempCNN (Table. 4). The lowest accuracy was observed from the model trained with 50% of reference data in the CCDC model. No difference in accuracy was observed when training the CCDC model using 50% and 75% of the reference data. The TempCNN showed more sensitivity to training data size; this sensitivity was present at all sampling proportions.

Table. 4. Model performance of varying the reference data sample size. Accuracy is based on test data. Highlighted in bold is the best performing model.

Reference Data Sample Size	CCDC Accuracy	TempCNN Accuracy
50%	0.92	0.93
75%	0.92	0.94
100%	**0.94**	**0.97**

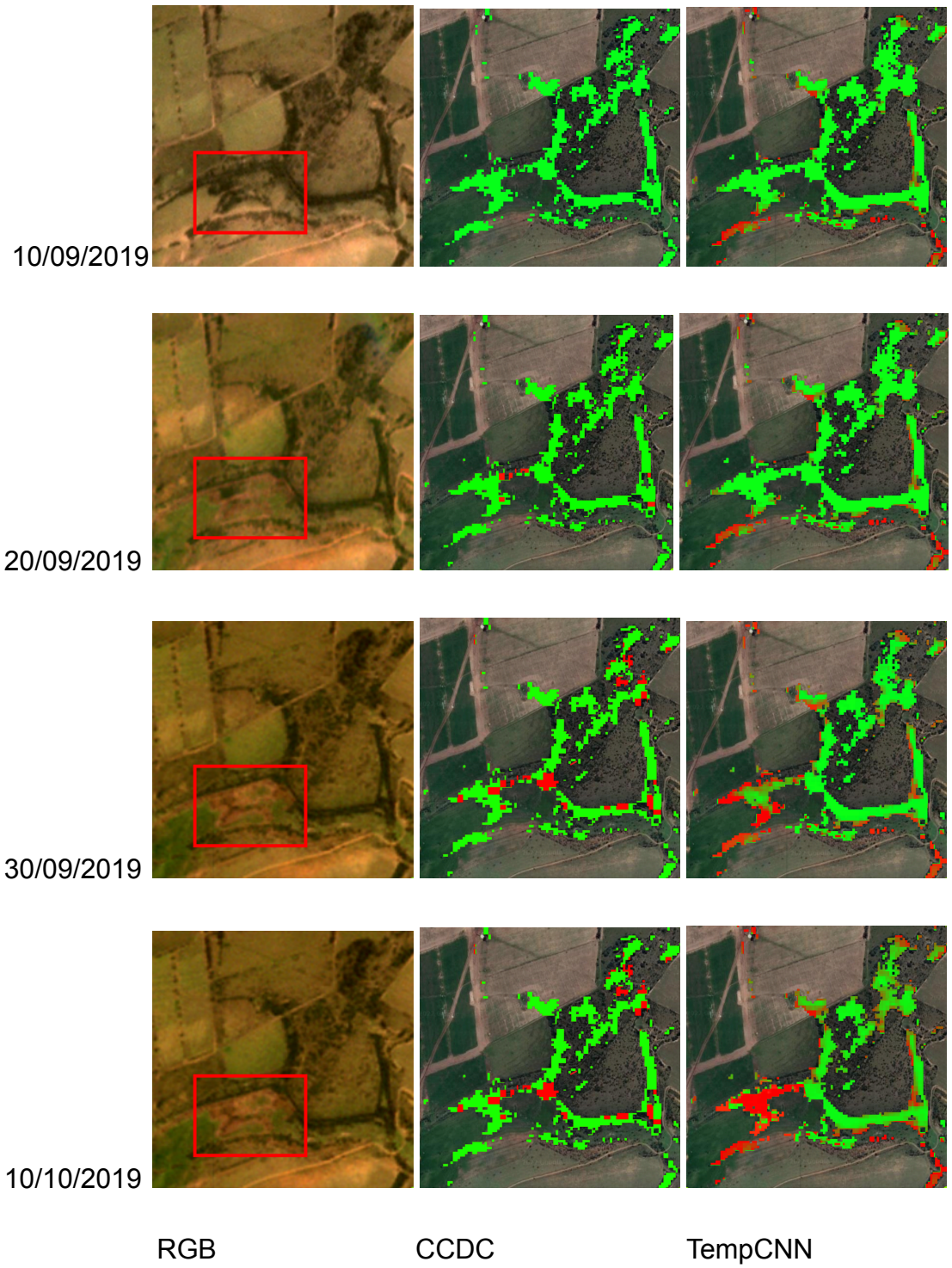


Fig. 3.6) An example of an event-to-flag lag comparison between CCDC and TempCNN. In this example, the RGB column shows Natural Colour Composites of PlanetScope imagery at 10-day steps. The red bounding box is the region where clearing was manually observed. At each of the time steps, the TempCNN is better at detecting clearings while CCDC does not detect some clearings. The event-to-flag

lag was between 10 and 20 days for TempCNN. The event-to-flag lag for CCDC was 15 days, provided that the difference in normalised values of all bands of a new observation is greater than one for all observations within a moving window period, typical of 3 observations.

3.4 Discussion

Despite the important ecosystem services provided by Albany Thickets, they remain under threat of land cover change, primarily for agricultural land expansion. Current estimates suggest that approximately 63% of the original extent of the Albany Thicket biome is severely degraded and continues to undergo change (Mucina and Rutherford, 2006; Powell, 2009; Stickler and Shackleton, 2015; Mills et al., 2018). While several studies have used remote sensing to map the long-term land cover change in Thicket ecosystems, little attention has been afforded to active and continuous monitoring of the land cover dynamics in these ecosystems. The structural and compositional uniqueness of Thicket make remote sensing-based tools designed for monitoring less spectrally diverse ecosystems, such as forests, inefficient. Fig. 1.4a, illustrates the mean spectral signatures of the dominant vegetation types in Alexandria, the spectral distinctness of Thicket from that of forests is evident from these plots. Fig. 1.4b, shows the intra-Thicket spectral variations.

3.4.1 Change Detection

The results of this study show that the TempCNN, TempCNN-MSI and CCDC models were able to detect changes in Albany Thicket cover. Their high performance is attributed to the fact that all three of the models incorporate both the temporal and spectral information in their implementation. The TempCNN-MSI model performed better than the CCDC and TempCNN models. The performance of TempCNN-MSI may be attributed to the fact that TempCNN-MSI allows for far more temporal density in the inputs compared to the CCDC, while also capturing the variance in aridity. The higher temporal density of the TempCNN and TempCNN-MSI results from the temporal gap filling performed in the preparation of their input data. While the CCDC discards cloudy and shadow pixels, the temporal gap filling replaces the gaps with

the previous clear observations, thus allowing for more observations in the input space of both the TempCNN and TempCNN-MSI. Pelletier et al. (2019) found that high observation frequency allows CNNs to extract more temporal information which in turn makes them robust to noise. TempCNNs are able to learn complex patterns in the temporal domain that may be linear or nonlinear, while the CCDC relies on the predefined coefficients of RIRLS and thus is rigid in detecting change.

The CCDC assumes that stable pixels are phenologically consistent. Semi-arid ecosystems might not meet this assumption due to inter-annual variations in selected classes (Zhu and Woodcock, 2014). This phenological inconsistency can lead to temporally dependent species turnover and alterations in species abundance between Albany Thicket species and those from adjacent biomes (Mucina and Rutherford, 2006). The CCDC is far more sensitive to change magnitude, which makes it less likely to make false positive detections but more likely to make false negative detections that result from subtle changes. This means that CCDC may potentially underestimate the extent of changes over time. Healey et al, (2018), found that the CCDC had large commission errors in pixels that had small change magnitudes. Fig. 3.5b, depicts widespread false negatives for the Cleared Thicket class as well as undetected change by CCDC compared to the TempCNN. A probable reason for these false negatives is that the change events may have occurred before the start of the time series; these changes often do not have sufficient flagged/change observations to affect the magnitude of the slope. An event-to-flag lag of between 10-20 days was noted for TempCNN in this study. It should be noted that the 15 days event-to-flag lag of the CCDC is subject to satisfying the condition that successive normalised differences are consecutively equal to one Zhu et al. (2019). Zhu et al. (2019) noted that the detection accuracy of CCDC increases when the number of consecutive anomalous detections is equal to six. In this study, a default number of consecutive anomalous detections of three was specified. This is the reason behind the high number of false negatives in the Cleared Thicket class. Specifying a high number of consecutive anomalous detections could result in commission errors in atmospherically complex areas where high cloud cover is common. Interrogation of the Sentinel 2 archive revealed that of the 404 images acquired over the study area between 2016-2019, 225 images had a cloudy pixel percentage less than 50 percent. A further limitation of the CCDC is that

it requires parameterisation and tuning, which is not straight-forward and can be time-consuming. A strength of the CCDC is that the limitation presented by cloudy imagery can be overcome through the incorporation of harmonised Landsat 8 and Sentinel-2 data Zhou et al. (2019). In theory, harmonised Landsat 8 and Sentinel-2 data can be used with the TempCNN; there are no existing studies confirming its utility.

While this study has confirmed the utility and robustness of the TempCNN, several other researchers have proposed other model architectures for times-series land cover change detection including Rußwurm and Körner, (2018) who used Long Short-Term Memory Network to achieve state-of-the-art accuracy for time-series crop classification. Sun et al. (2020) used Deep Neural Networks to produce reliable in-season maps crop types classification. Xu et al. (2020) used Attention-based Long Short-Term Memory Network for dynamic corn and soybean mapping with improved spatial generalizability.

3.4.2 Influence of aridity on performance

While adding the MSI resolved most of these false positives, it was then apparent that some of the false positives were a result of the adjacency effect: where stray light comes in the field of view of neighbouring pixels this can result in spectral attenuation or greater mixing in heterogeneous surfaces (Lyapustin and Kaufman, 2001), an example of this can be seen in Fig. 3.4, on the classified map the Thicket-barren land fringes within the red box are classified as cleared when in actuality the pixels are Stable Thicket as seen in the RGB image. This phenomenon was observed in Moncrieff (2021), where shrubland pixels lay adjacent to barren or ploughed pixels. The adjacency effect was also observed in non-intact Thicket.

The MSI does improve the performance of the TempCNN, an independent measure of aridity would be more beneficial as it would not be collinear to the other inputs. While the incorporation of the MSI improved the performance of the TempCNN, spectral indices are not without inherent limitations, including the uncertainty born out of mixed spectral curves that manifest from pixels that cover diverse land cover types e.g where vegetation is adjacent to a waterbody. Without the inclusion of spectral unmixing analysis the true contribution of each land cover type to the pixel's

spectral curve is not known. Another limitation of spectral indices is that they are not readily verified and are not direct measurements of the physical properties of interest, this means that they often require additional data to quantify how the spectral index-ground unit of measurement relationships and infer the property's unit of measurement from the unitless spectral index (Zheng et al., 2022). With these limitations in mind, using direct measures of precipitation rather than the MSI should give more insight into how surface moisture data can be used to glean the latent features that differentiate stable and cleared vegetation across the Albany Thicket biome and within Albany Thicket subtypes. Chapter 4 will focus on overcoming this limitation.

3.4.3 Influence of training sample size

This study showed that increasing the training dataset size improves model accuracy (Table 3). Moncrieff (2021) achieved superior performance of TempCNNs even with relatively small amounts of the training data. However, increasing the sample size increases the class distribution and thus the likelihood of label noise due to environmental variations. Label noise refers to discrepancies in the assigning of labels of a dataset. This robustness is a result of using bootstrap samples during the construction of the decision trees (Breiman, 2001; Agjee et al., 2018; Maas et al., 2019). Regularisation mechanisms implemented in the TempCNN, including dropout and weight regularisation, improve the generalisation of TempCNN and minimise the effect of label noise (Pelletier et al., 2019). The benefit of using the CCDC is that unlike the TempCNN it does not rely on having training samples as it can be trained in an unsupervised mode, this voids the need for the collection and creation of labelled samples which is often a laborious task. In this study Planet daily data were used to assign dates of when changes in Albany Thicket cover occurred, Access to such data may not always be possible, this therefore makes the use of CCDC more pragmatic. The limitation of this is that while changes in land cover are detected, information on the class transitions cannot be inferred easily. The primary limitation of the CCDC is the large computing resources required to train the time-series model that produces the coefficients used in the classification. Pasquarella et al. (2022), found that the CCDC algorithm performs well when it is used to detect changes that

result in high-magnitude changes in pixel values. Pasquarella et al. (2022), further found that the CCDC works well in regions that have less diverse land cover types.

While the original CCDC algorithm has shown weaker performance against the TempCNN, several enhancements to the original algorithm have been proposed by researchers including fusing time-series optical and SAR images, Fu et al (2022). By fusing optical and SAR images it was demonstrated that the annual phenological trajectories can be clearly tracked and accurately detect losses and restoration of marsh vegetation. Zhu et al. (2021), found that the incorporation of auxiliary environmental variables into the original CCDC algorithm could result in improved performance.

3.5 Conclusion

The Albany Thicket biome in South Africa has undergone extensive degradation primarily due to agricultural land expansion. The aim of this study was to investigate the utility of supervised continuous change detection in the Albany Thicket of South Africa. CCDC was used as a baseline algorithm to compare the performance of the TempCNN algorithm. The results demonstrate the potential utility of both The CCDC and TempCNNs in detecting changes in Albany Thicket cover within days of the inception of clearing. The performance of TempCNNs is associated with their incorporation of temporal information, in addition to this the TempCNN learns the temporal features of relevance from labelled data. Aridity was shown to be an important biophysical factor influencing the detection accuracy of the TempCNN. Both the CCDC and TempCNN performed well even when training datasets were much smaller in size, but the TempCNN benefited from increased training data size. The TempCNN's performance can also be attributed to the fact that supervised methods such as TempCNN require labelled examples. CCDC and other similar continuous change detection algorithms remain an attractive option when resources or expertise do not allow for the creation of high quality labelled datasets. The TempCNN provides much promise in its ability to detect and map changes in land cover in the Albany Thicket, however it is not free of limitations. The first limitation is that it uses spectral indices to account for variances in environmental conditions. While spectral indices are easy to calculate and integrate into the current inputs of

the TempCNN, they are indirect proxies that are inherently affected by surface and atmospheric conditions that influence the information they generate. Spectral indices often require expert inversion models to convert them to meaningful units of measurement. A second limitation of the TempCNN is that it relies on the availability of large amounts of annotated data to yield good performance. Large amounts of data are often costly and laborious to generate. This shortcoming limits a model's ability to generalise in shifting domains. In chapter 4 I attempt to address these limitations, firstly by incorporating surface moisture variables in the TempCNN to improve the performance of the TempCNN's performance across the Albany Thicket as well as within individual domains with Albany Thicket biome. Secondly I implement a Model Agnostic Machine Learning framework to improve the TempCNN's ability to adapt faster when extending to new domains.

Chapter 4: Domain adaptive Temporal Convolution Neural Networks for continuous change detection in the Albany Thicket biome

4.1 Introduction

Although land use legislation in South Africa has come a long way in ensuring the protection of endangered ecosystems (Botts et al., 2020), clearing and degradation of natural ecosystems persists (Skowno et al., 2021). The Albany Thicket biome continues to experience extensive degradation and loss, with crop production and pastoralism being the primary causes of clearing (Hoffman et al., 1987; Lloyd et al., 2002; Stickler and Shackleton, 2015; Mills et al., 2018; Carvalho et al., 2021). Accurately detecting and documenting clearing events is important for understanding, monitoring and managing land cover change in the Albany Thicket biome. Researchers and practitioners have turned to remote sensing methods that offer greater efficiency compared to traditional methods such as field inspections.

Earth observation data coupled with deep learning algorithms have been recognized as an efficient and accurate approach for detecting and mapping land cover change (Shendryk et al., 2019; Chai et al., 2019; Feizizadeh et al., 2021). Earth observation is an efficient method of gathering information on the surface of the Earth due to the ability of Earth observation platforms to repeatedly acquire data over large and sometimes inaccessible areas. Several studies have demonstrated the ability of deep learning algorithms, particularly Convolutional Neural Networks (CNNs) to accurately classify land cover change using Earth observation data, (Chai et al., 2019; Schiefer et al., 2020; Yang et al. 2021; Solomon and Agnes, 2023). For these reasons remote sensing data and deep learning are implemented to detect land cover changes in Albany Thicket. Previously I have shown that 4-year sets of Sentinel-2 and the Random Forests can be used to map and document changes in land cover in the Albany Thicket biome (Chapter 2). The study has also revealed that Continuous change detection can timeously detect and map clearing of Albany Thicket (Chapter 3).

The successful adoption of these methods for operational use would require that the methods can be implemented through-out the Albany Thicket biome regardless of the Albany Thicket subtype. This means that a model used to detect land cover changes in the Albany Thicket should be able to generalise across different vegetation subtypes within the biome. The importance of such a model is that it would be able to consistently make predictions despite variations in environmental conditions with the biome.

The Albany Thicket exhibits large spatial and temporal variability in vegetation across the biome. This variability is mainly a consequence of a dual climatic system within the biome, this is driven largely by the existence of two rainfall patterns. Summer rainfall is prevalent in the north-east regions, the southwest is dominated by all-year rainfall. Precipitation reliability is a major factor in the Albany Thicket. Rainfall is unreliable, varying the least in coastal regions and escarpment the most variation is observed in the in-land regions of the Gamka Thicket. Months long droughts are a persistent feature of the Albany Thicket (Mucina and Rutherford, 2006; Palmer et al., 2020). The spatial and temporal variation in precipitation is shown in Fig. 4.1. The variation in precipitation along with variances in fire regimes, geology and soils is reflected in species turnover and shifts from Albany Thicket species to those from adjacent biomes (Mucina and Rutherford, 2006). Vegetation subtypes also exhibit differences in timing of leaf emergence, senescence, and flowering. These phenological differences make cross-biome inference challenging (Li et al., 2010; Smith et al., 2019; Kattenborn et al., 2021). The impact of phenological differences on model accuracy grows when models are trained on one or several vegetation biomes and then used to make inference in a vegetation type that was not in the training data (Kattenborn et al., 2021). Differences in structural complexity and species composition among biomes and vegetation subtypes influence light recorded by satellites (Colwell, 1974; Smith et al., 2019). The sum of the light-canopy interactions dictates the rate of spectral mixing. Spectral mixing is the presence of a dilution of a range of land cover types at the sub-pixel level (Ghrefat et al., 2023). Spectral mixing influences the spectral thresholds or the pixel-values at which stable vegetation and disturbed vegetation are separated from each other. This makes change detection in diverse ecosystems as well as in shifting domains more challenging (Chamarthi et al., 2023).

The TempCNN model implemented in the study preceding this one showed good performance, however it used the Normalised Difference Water Index (NDWI) and Moisture Stress Index (MSI) to account for variances in surface moisture across the study area. However, the NDWI and MSI are indirect measures of moisture used as a proxy for precipitation and surface moisture. Spectral indices are sensitive to various contaminants such as atmospheric water and are not independent measures of environmental conditions. Another limitation of spectral indices is that they are not direct measurements of the biophysical properties of interest and therefore require expert knowledge to make associations between indices and the biophysical variable (Zheng et al., 2022). For these reasons, I thought that incorporating independent environmental variables would be more suitable.

Remote sensing based deep learning classification algorithms tend to perform well when trained and used for inference on a single reference image of a single geographical location (Gawlikowski et al., 2022). Challenges arise when inference is performed on images of varying atmospheric, surface moisture, illumination and radiative transfer characteristics or of a different geographical location and of different sensor platforms (Kattenborn et al., 2021; Gawlikowski et al., 2022; Nyborg et al., 2022).

In arid areas such the Albany Thicket, the challenges of cross-domain inference are compounded by the spectral complexity or the range of cover types at the sub-pixel level. Several biophysical factors including mixtures of senescent and green species, photosynthetically active and soil crusts, variations in soil salinity and the presence of litter and dead vegetation as well as shadows influence light recorded by satellites (Tueller, 1987; Smith et al., 2019; Nyborg et al., 2022; Tuia and Bruzzone, 2023). These variations in biophysical conditions can degrade model performance when performing predictions within domains.

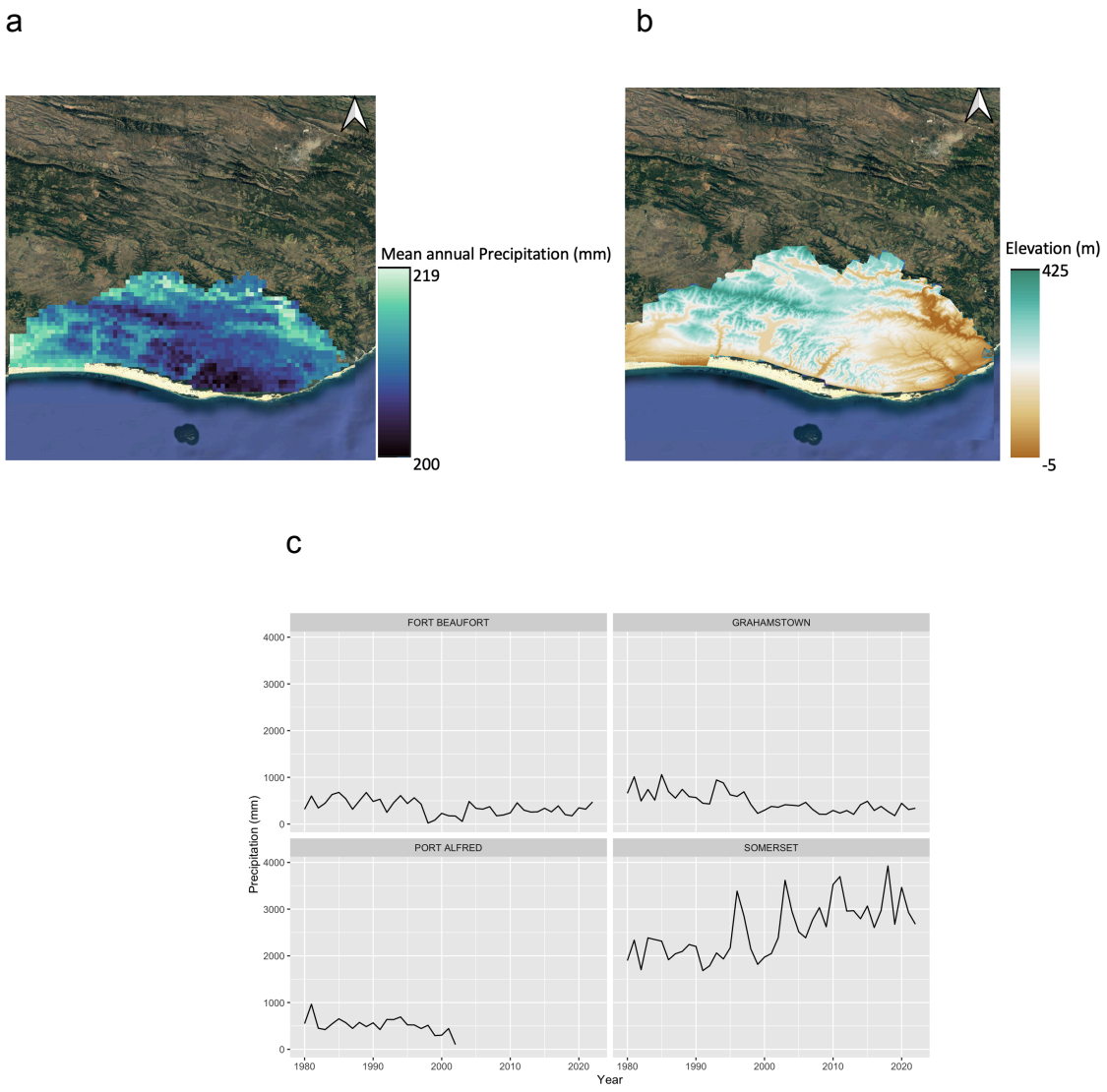


Fig. 4.1 a) Spatial variation in total precipitation in Alexandria. b) Map of the study area showing the elevation gradient. Source (https://search.earthdata.nasa.gov/search?q=SRTM). c) A time-series of the years between 1980-2022 of total precipitation the data were gathered from four weather stations around the Alexandria area. Source (https://www.ndbc.noaa.gov/).

Ideally a model trained on a dataset containing examples of stable and cleared Albany Thicket of all subtypes would learn and extract features that are representative of all stable Albany Thicket subtypes and cleared Albany Thicket. However, the lack of cleared Albany Thicket across a geographically representative subspace of the study region hinders the application of such an approach. Another issue that limits the successful transferability of machine learning models are

deviations in the probability distributions of datasets this is referred to as domain shift (Stacke et al., 2029; Yan et al., 2019; Chamarthi et al., 2024). This happens when data previously unseen by a model is used to make predictions using a model that was trained on data whose distribution is significantly different from the unseen data. The factors listed in the preceding paragraphs are the main sources of distributional shifts in remote sensing data and results in drops in model performance when using traditional machine learning models (Peng et al., 2022). The problem of domain shift is further exacerbated by the focus of continuous/time-series observations (Tuia et al., 2021), as is the case in this study.

This chapter seeks to improve the limitations of the TempCNN implemented in chapter 3. This chapter assesses whether the inclusion of environmental variables can improve the performance of TempCNN. A surface-dryness fused TempCNN (SDF-TempCNN) is proposed for continuous change detection in the Albany Thicket. SDF-TempCNN is a dual-modality convolutional neural network that has two branches, the first extracts features in the temporal domain from a time-series of multispectral images. The extracted Sentinel-2 features are then fused with surface-dryness features extracted from precipitation and evapotranspiration features that make the second branch. By incorporating environmental variables that capture the variations in moisture content, the SDF-TempCNN learns the more latent representations that capture domain contextuality between stable vegetation and cleared vegetation. These latent representations capture the underlying structure in the input data. By learning the invariant representations, the SDF-TempCNN can be used to learn the knowledge that is transferable across domains and unseen datasets (Bhatt et al., 2019). I posit that by introducing domain contextuality through the surface-dryness features, the SDF-TempCNN will improve the accuracy and robustness of change detection in the Albany Thicket biome compared to the baseline TempCNN. The goal of the SDF-TempCNN is to learn the trend in spectral bands and spectral indices in both stable and cleared vegetation while retaining location specific connections of environmental variables. The incorporation of the surface-dryness variables allows the SDF-TempCNN to be robust against distributions/domain shifts that result from the inherent variations in atmospheric, surface and illumination conditions in remote sensing imagery that are inherent in between satellite observations (Tuia et al., 2021), thus optimising the baseline

TempCNN to be domain adaptable. The second goal of this chapter aims to improve the TempCNN by optimising its ability to adapt to the variations within Albany Thicket subtypes. A MAML-TempCNN is implemented to assess cross-biome generalisation and adaptation in instances where data scarcity limits the application of deep learning models. The implementation of the MAML-TempCNN aims to allow the TempCNN to speedily adapt to new and environmentally diverse regions that may have different vegetation types. MAML-TempCNN can effectively adapt to diverse ecological conditions within the Albany Thicket biome, improving the accuracy and robustness of continuous change detection compared to non-meta-learning approaches. By using few-shot learning, the MAML-TempCNN will enable faster optimisation of the model parameters on a few samples during fine-turning in shifting geographical domains.

4.2 Methods

4.2.1 Data

4.2.1.1 Satellite data

Refer to section 3.2.1.1 for a description of the satellite data used in this chapter.

4.2.1.2 Surface moisture variables

Mean total precipitation, and evaporation, for the years 2016-2019 were calculated from The European Centre for Medium-Range Weather Forecasts ERA5 reanalysis data (Hersbach et al., 2020). The ERA5 reanalysis Total precipitation dataset captures the hourly sum water in liquid and frozen on the surface of the Earth. Two datasets were chosen to express evaporation; the first is the ERA5 reanalysis evaporation from bare soil over land surface, this is the hourly sum evaporation from bare soil. The second is ERA5 reanalysis of evaporation from the top of canopy, this is the hourly sum of top of canopy evaporation. These variables were chosen due to the effects of land cover change have on surface water and assumption that this would results in variations up welling spectral signatures and in turn latent representations learned by the models (Chen et al., 2022; Chaemiso et al., 2021; Vandas et al., 2002). ERA5 datasets are available on Google Earth Engine

(https://developers.google.com/earth-engine/datasets/tags/era5-land). Land surface temperature (LST) was calculated from Landsat 8 Thermal Infrared Sensor (TIRS) (https://landsat.gsfc.nasa.gov/satellites/landsat-8/spacecraft-instruments/thermal-infrared-sensor), using the split-window method. These variables were chosen for their association The split-window makes use of the two Thermal Infrared bands on the TIRS sensor onboard the Landsat 8 platform, the algorithm accounts for land surface emissivity, atmospheric water vapour and viewing angle (Wang et al., 2019). changes in land cover have been associated with changes in land surface temperature as such it was deemed that vital representations can be gleaned from Land surface temperature data (Tam et al., 2020; kafy et al., 2021). The surface-dryness variables were extracted and the mean over 2016-2019 were computed for each Stable and cleared Albany Thicket sample.

4.2.1.3 Reference data

Reference data were extracted from the Albany Thicket loss instances and Stable Thicket datasets to create model training and evaluating samples. The time series of each of these was then sampled from pixels within these polygons and subsampled into partially overlapping windows as described above. The pixel sampling was performed in a way that prevents training and validation samples from coming from the same clearing instances. The label “Stable Thicket” was allocated to a time series subsample if the date when Albany Thicket was last confirmed to be present occurs later than the end date of the series under consideration, or if no change had occurred (Fig. 3.1). Where both the date when Albany Thicket was last confirmed to be present and the date when Albany Thicket was first confirmed to be lost are within the same subsample the label “Cleared Thicket” is allocated.

To compare domain adaptability of the models, the reference data were split by the Albany Thicket subtype. The splitting yielded 8 subtypes, 3 of which were used for the training set and 5 in the testing set. A total of 3179 pixels were used to train and evaluate the models, 2203 pixels were placed in the training and validation sets. A total of 976 pixels were drawn from the testing subtypes (Table. 5). A visual description of the sampling regime implemented in the training of the cross domain adaptation models is shown in Fig. 4.2.

Table. 5. The sampling regime used to train the TempCNN and, SDF-TempCNN and MAML-TempCNN for with domain adaptability.

Train	Val	AMT	GGT	GRT	NST	KDT
1323	883	312	102	206	189	164

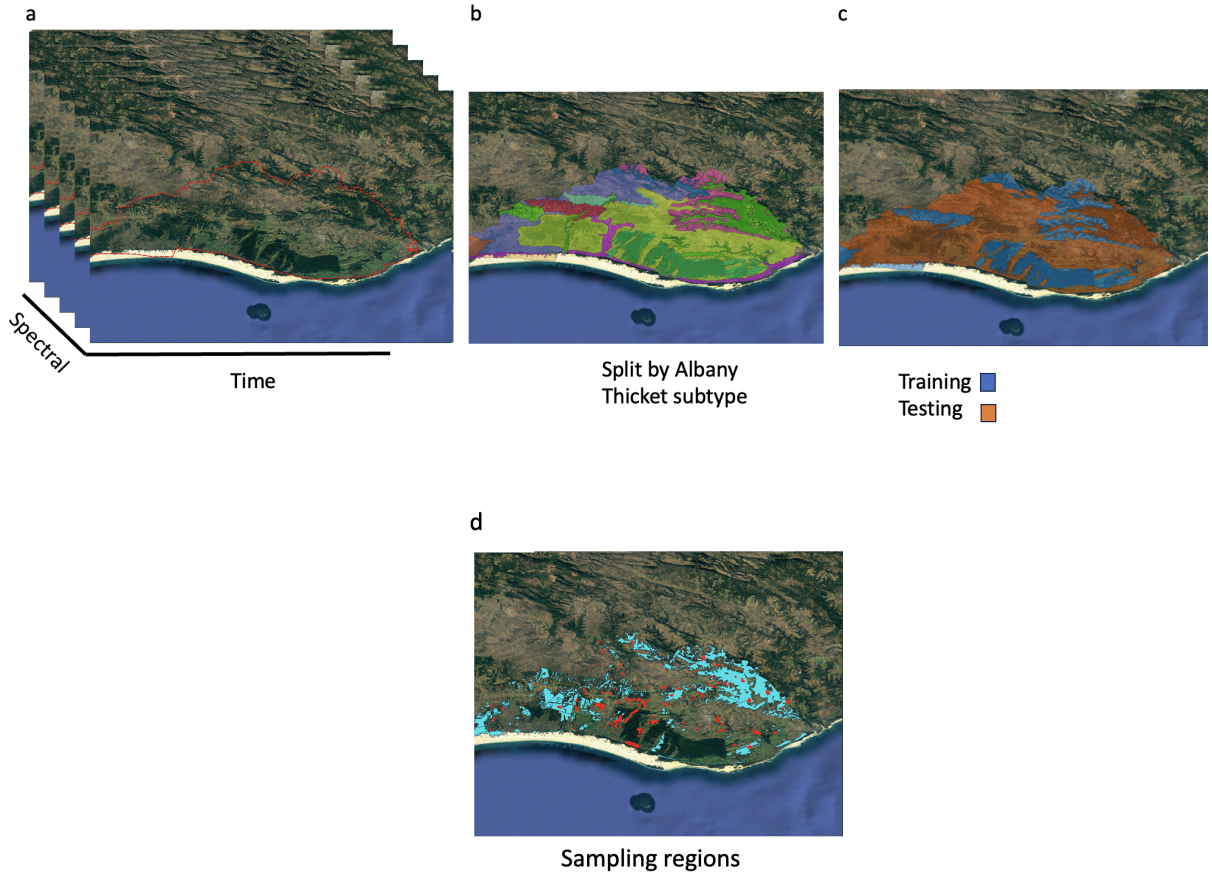


Fig. 4.2) A visual description of the sampling regime implemented in the training of the cross domain adaptation models. Panel a) illustrates the multivariate time-series, The time-series Sentinel-2 data were subsampled to windows with a fixed length of 20 observations with a five-day update and a step size of three (see Fig.3.4). b) depicts the splitting of the time-series by Albany Thicket subtypes. c) The allocation of Albany Thicket subtypes into training and testing regions. d) depicts the regions where Stable Thicket and Cleared Thicket samples were extracted from.

4.2.2. Models

4.2.2.1. Surface-dryness fused TempCNN Temporal Convolutional Neural Networks

Convolutional Neural Networks have received extensive attention in remote sensing-based image classification and object detection. This is primarily due to their ability to extract complex spectral features in remotely sensed imagery (Drönner et al., 2018; Wang et al., 2018). However, most of these algorithms are based on the extraction of spectral features from mono-temporal/snapshot images. As a result, the temporal dynamics are often not accounted for by these algorithms. To overcome this, Pelletier et al. (2019) proposed Temporal Convolutional Neural Networks.

TempCCN is a 1-Dimensional CNN used in the temporal domain to extract multivariate temporal structures from multivariate time series satellite imagery (Pelletier et al., 2019; Zhu et al., 2021). The multivariate input comprises Sentinel-2 surface reflectance bands and spectral indices including NDVI, SAVI, ERA5 total precipitation, ERA5 evaporation from bare soil and top of canopy and Landsat land surface temperature.

The TempCNN used in the third chapter comprises three convolutional layers with convolution filters of the same size. The convolutional layers with 32 units are successively applied. Each convolutional layer is followed by an average pooling layer that reduces the dimensions of the feature space. Neurons from the first layer receive a multivariate time series as input and extract both multi-spectral information and temporal dynamics. The three convolutional layers are followed by one Dense layer with 256 neurons. Finally, a Softmax layer produces the distribution of classes for prediction (Fig. 4).

$$\text{softmax}(x_i) = \frac{exp(x_i)}{\Sigma_j exp(x_j)} \qquad \text{equation (4.1)}$$

where

x_i is the ith element of the input vector

$exp(x_i)$ is the exponential function applied to x_i

$\Sigma_j\, exp(x_j)$ is the sum of the exponential of all elements in the input vector

A surface-dryness fused TempCNN (SDF-TempCNN) is proposed for continuous change detection in the Albany Thicket. SDF-TempCNN is a dual-modality convolutional neural network that has two branches, the first extracts features in the temporal domain from a time-series of multispectral images. The extracted Sentinel-2 features are then fused with surface-dryness features extracted from precipitation and evapotranspiration features that make the second branch. By incorporating environmental variables that capture the variations in moisture content, the SDF-TempCNN learns the more latent representations that capture domain contextuality between stable vegetation and cleared vegetation. These latent representations capture the underlying structure in the input data. By learning the invariant representations, the SDF-TempCNN can be used to learn the knowledge that is transferable across domains and unseen datasets (Bhatt et al., 2019). The goal of the SDF-TempCNN is to learn the trend in spectral bands and spectral indices in both stable and cleared vegetation while retaining location specific connections of environmental variables. The incorporation of the surface-dryness variables allows the SDF-TempCNN to be robust against distributions/domain shifts that result from the inherent variations in atmospheric, surface and illumination conditions in remote sensing imagery that are inherent in between satellite observations (Tuia et al., 2021), thus optimising the baseline TempCNN to be domain adaptable.

To fuse the surface-dryness variables the top layer was removed and replaced with a concatenating layer. The concatenate layer adjoins the TempCNN to the surface-dryness multilayer perceptron (MLP) branch. The MLP comprises five dense layers with a dropout rate of 0.8. The concatenation layer is followed by a single Dense layer and a Softmax layer to output the distribution of the predicting class. An illustration of the architecture can be seen in Fig. 4.3. The method of fusion implemented here is commonly referred to as late fusion. Late fusion refers to independently training several models and fusing them at the decision/classification stage of the model. The late fusion approach was selected because it is more flexible and provides a more robustness against overfitting (Hosseinpour et al.,

2022). The implementation of late fusion is also justified by the fact that the two branches of the SDF-TempCNN use two different data inputs, this allows for flexible learning and extraction of representation.

The network was trained using Adam optimization with a batch size of 32 and a learning rate of 0.0001. To avoid overfitting, regularisation mechanisms were implemented including a dropout rate of 0.8, an L_2-regularisation on the weights with a regularisation factor of 1e-8, was specified. The model was trained for 5 epochs and focal loss was used for the loss function. Unlike cross entropy, focal loss is able to handle imbalances between background and foreground features. It is for this reason it was chosen. All preprocessing was performed using Google Earth Engine while Tensorflow 2 (https://www.tensorflow.org/api_docs/python/tf) and eo-flow (https://github.com/sentinel-hub/eo-flow) were used to implement TempCNN.

$$\text{Focal Loss}\,(p) = -\,(1-p)\,y\ *\ log\,(p) \qquad \text{equation (4.2)}$$

where

Focal Loss (p) is the focal loss for a predicted probability of p

p is the predicted probability that an example belongs to the positive class

y is a hyperparameter that controls the focus of the loss function.

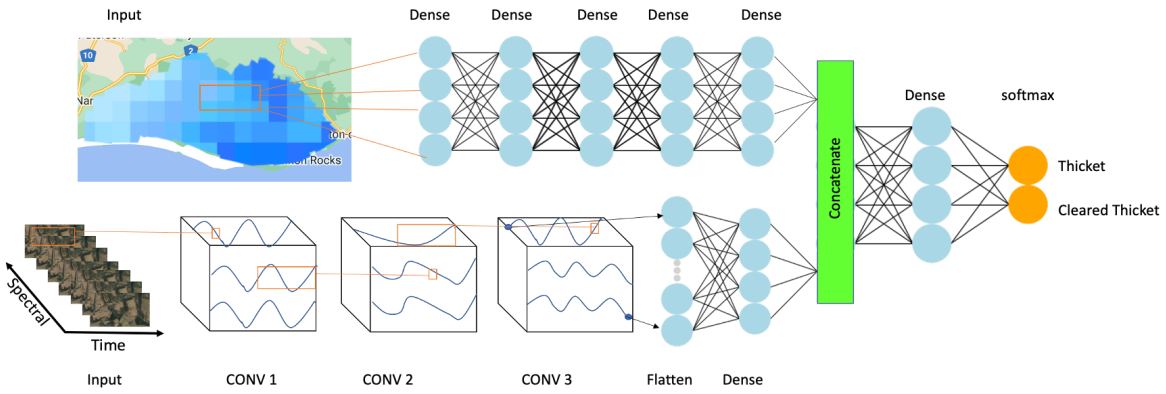


Fig. 4.3) The SDF-TempCNN. For the TempCNN branch of the network, the input is a multivariate time series of Sentinel-2 spectral bands and indices. Three convolutional filters are applied in succession, and a single dense layer. For the surface-dryness branch of the network the input is a multivariate of surface-dryness variables which are passed to five dense layers, the two branches are merged by concatenation. Followed by a dense layer and finally a Softmax layer.

4.2.3.2 MAML-Temporal Convolutional Neural Networks

A MAML-TempCNN was implemented to optimise cross-biome generalisation of the TempCNN. A TempCNN was trained with MAML on 2-way 1-shot, 3-shot and 5-shot datasets from the meta-train regions (5 Albany Thicket subtypes). The data were split into 6 subsets. the first 5 subsets were created to make up the meta-train and validation datasets by iteratively leaving a single Albany Thinket subtype out of that subset, in the end the following subsets were yielded including without Grahamstown Grassland thicket, without Geluk Grassland Thicket, without Nanaga Savanna thicket, without Thorndale Forest Thicket and without Albany Mesic Thicket. During the training and validation phase each subset is iteratively dropped from the training dataset and used for validation. Data containing samples of clearing events and stable vegetation from only the South Eastern Coastal Thornveld regions were used in the meta-test subset.

The MAML was trained to classify Stable Thicket and Cleared Thicket (2-way) given 1 example (1-shot), 3 examples (3-shot) and 5 examples (5-shot). MAML aims to use a small number of gradient steps in learning the optimal initialization parameters

that enable fast adaptation to new tasks. TempCNN was used in the implementation, where a small batch of unseen data is used to train the TempCNN. 500 2-way 1-shot, 3-shot and 5-shot classification tasks were sampled from the test set for the two classes (stable vegetation and lost vegetation) in one of the Albany Thicket subtypes. The three different shot sizes were compared to assess the amount of data needed for optimal generalisation. MAML-TempCNN performs Stochastic Gradient Descent using the difference between weights trained on a small batch of never-before-seen data and the model weights prior to training over 500 meta-iterations. In the meta-learning phase, the learner attempts to quickly adapt to a small number of tasks through training and evaluating on the meta training regions, the hyperparameters are then tuned on the meta-validation (Finn et al., 2017). The model's ability to generalise on new, unseen tasks is assessed on the meta-test set. An illustration of the MAML-TempCNN algorithm can be seen in Fig. 4.4. A step by step of the MAML algorithm is presented below.

1) The algorithm considers a batch of tasks T from a distribution over tasks D, followed by the initialisation of the model parameters theta. For each task Ti in T, the task is split into a few-shot learning set Di and a validation set (Vi,). The loss is then computed on Di set, equation (4.2). Backpropagation is performed to update the model parameters θ_i (equation 4.3).

$$l\,\mathrm{D}\,(\theta_i) = \sum\{x, y(D_i)\}\, l(f_\theta(x_i), y) \qquad \text{equation (4.3)}$$

where

l is the loss function used to compute the difference between the predicted output $l(f_\theta(x), y)$ and the actual output y for each input-output pair (x,y) in the few-shot learning set Di (equation 4.4).

$f_\theta(x_i)$ represents the output of the model with parameters θ_i for input x.

Update f_{θ} = θ - a * gθ l D (θ_i) equation (4.4)

where

θ is the initial set of model parameters that are learned over a distribution of tasks

a is the step size or learning rate, which determines how big of a step is taken in the direction of the negative gradient.

l D(θ) is the gradient of the loss function with respect to the model parameters θ, which measures the direction of steepest ascent for the loss function θ_i is the updated set of model parameters after one or more gradient steps on the i-th task.

2) The updated parameters of θ_i are used to compute loss on the validation set. (equation 4.5).

l V(θ_i)= $\sum\{x, y(V_i)\}\ l(f_{\theta}(x_i), y)$ equation (4.5)

where

l is the loss function used to compute the difference between the predicted output

$l(f_{\theta}(x), y)$ and the actual output y for each input-output pair (x,y) in the validation set Vi

Vi is the validation set for the i-th task, which is used to evaluate the performance of the model on unseen data.

(x,y) A training example from the validation set Vi The input x is a feature vector and y are the corresponding label or target output.

$f_{\theta}(x_i)$ represents the output of the model with parameters θ_i for input x.

3) Compute the loss for θ_i $\text{in the support set}$. The gradients across all tasks are averaged to compute the meta-gradients (equation 4.6).

$\frac{1}{|T|}\ \{T_i T\}\ g\theta\ l\ V(\theta_i)$ equation (4.6)

where

$l\ V(\theta_i)$ is the validation loss for the i-th task, which is computed using the task-specific model parameters θ_i on the validation set Vi for that task.

$g\theta\ l\ V(\theta_i)$ is the gradient of the validation loss $l\ V(\theta)$ with respect to the task-specific model parameters θ_i for the i-th task.

T is the set of few-shot learning tasks used to learn a good initialization of the model parameters θ.

θ is the initial set of model parameters that are learned over the distribution of tasks T.

4) Finally the meta losses of an entire batch are summed up and the parameters are updated using stochastic gradient descent (SGD). The updating is performed on the initial parameters by calculating the gradients of the optimal parameters in the preceding step (equation 4.7).

SGD = $\theta - \alpha\nabla J(\theta,\ x(i),\ y(i))$ equation (4.7)

where

θ is the vector of parameters to be optimised

α is the learning rate, which controls the step size taken in the direction of the gradient

$J\ (\theta, x(i), y(i))$ is the cost function for a particular training example $(x(i), y(i))$

$\nabla J\ (\theta, x(i), y(i))$ is the gradient of the cost function with respect to the parameters θ for a particular training example $(x(i), y(i))$

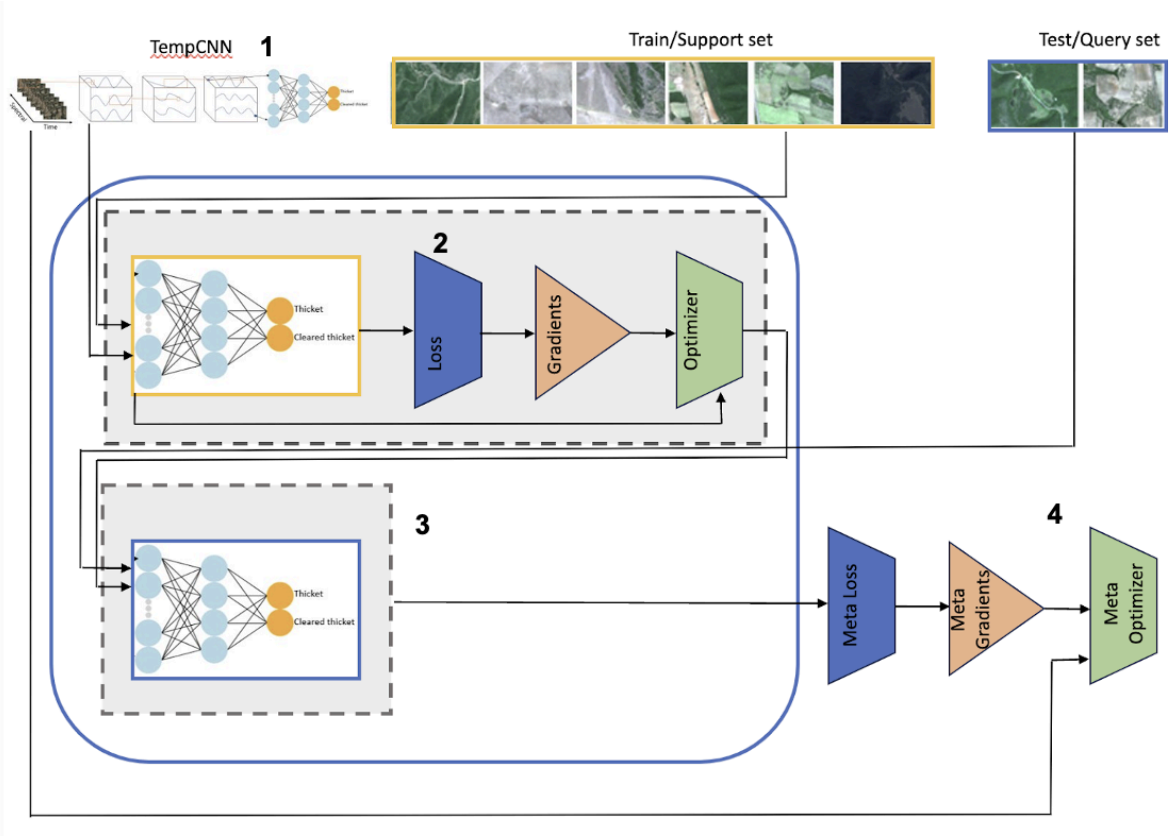


Fig. 4.4) MAML-TempCNN. The TempCNN is initialised, the data are then sampled from the train and test splits. The TempCNN parameters are used in the adaptation (1). During adaptation, the train set is sampled recursively to train the adapter (2); the losses and gradients are computed to update the weights (3). The Meta loss is computed using the update parameters based on the test set. Finally, the meta-gradients, meta-parameters using the meta-loss (4).

4.2.4 Experiments

Two experiments were conducted to determine the utility of considering environmental variables in continuous land cover change detection to account for ecological context. The first experiment assesses whether the TempCNN developed in the third chapter could be extended into other ecological domains by incorporating environmental variables that account for gradients in surface moisture. This was done by training the SDF-TempCNN on the entire Albany Thicket domain and

evaluating the models on a test dataset drawn from the entire domain. This model was compared to the TempCNN developed in chapter 3. TempCNN and SDF-TempCNN were then trained and evaluated on the Albany Thicket subtype dataset described in section 4.2.2.2.

The second experiment focuses on assessing the generalisation and adaptive ability of the TempCNN in shifting domains using MAML. The MAML-TempCNN was trained to classify Stable Thicket and Cleared Thicket on 2-way 1-shot, 3-shot and 5-shot on the tasks/Albany Thicket subtypes described in section 4.2.2.2. The MAML-TempCNN was compared to the TempCNN and SDF-TempCNN which were also trained on the same 2-way 1-shot, 3-shot and 5-shot Albany Thicket subtypes.

4.2.4.1 Model Evaluation

To evaluate the performance of the SDF-TempCNN, several evaluation metrics were produced, including overall accuracy, producer's accuracy, user's accuracy, precision, and recall.

The overall accuracy measures the percentage of correctly classified pixels.

Overall accuracy = $\frac{P_a}{t_a}$ equation (4.8)

where

P_a is the total number of correctly classified pixels, t_a is the total number of pixels in the testing set.

Precision measures the ratio of correctly predicted pixels to all predicted pixels for specific class:

Precision = $\frac{tp}{tp+fp}$ equation (4.9)

Recall measures the ratio of positively predicted pixels to the total number of positively classified pixels

Recall = $\frac{tp}{tp+fn}$ equation (4.10)

Where

$tp,\ fp$ and fn are true positive, false positive and false negative, respectively.

$$\text{F1 Score} = \frac{2*(Precision * Recall)}{(Precision * Recall)} \qquad \text{equation (4.11)}$$

Where

As per equation 4.10 and 4.11

Gradient-weighted Class Activation Mapping ++ (GRAD CAM ++) (Chattopadhyay et al., 2017), was used to generate visual explanations of the localisation of class-discrimination by the two models. GRAD CAM ++ is an effective way to visualise the importance of different temporal regions in the time-series input for both models. These visualisations are often referred to as saliency maps. Saliency maps localises the weights and measures the intensity of their contribution to the classification (Amorim et al., 2023). GRAD CAM ++ calculates the contribution of each temporal window to the final class allocation by taking the weighted average of all positive activations in the last convolutional layer (Chattopadhyay et al., 2017). This therefore gives insight into which temporal windows are important in the decision-making process.

4.3 Results

4.3.1 Change Detection

The results show that by fusing environmental variables onto the TempCNN based on Sentinel 2 spectral bands and indices, the model was able to perform slightly better than the baseline TempCNN that did not possess the SDF-TempCNN features. Table. 6 shows the performance metrics of the SDF-TempCNN as well as those seen from the baseline. A recall of 0.96 and precision of 0.95 when detecting changes in Albany Thicket cover were observed from the SDF-TempCNN. An accuracy of 0.98 was observed from the SDF-TempCNN compared to 0.92 observed from the baseline TempCNN. Both models had similar precision scores for Stable Thicket

(0.93), while the SDF-TempCNN had a precision of 0.95 compared to the 0.93 seen from the baseline TempCNN.

SDF-TempCNN outperforms the original TempCNN in all performance metrics, except for the Cleared Thicket recall score and the Stable Thicket precision score, where it performs equal to it (0.93) when predicting change across the Albany Thicket domain. The original TempCNN had a both precision score and recall score of 0.93 in the Stable Thicket class, this indicates that some changes in this class were not detected by this model, while the SDF-TempCNN scored 0.98 on both precision and recall in both classes

Figure 4.5 shows a GRAD CAM ++ visualisation of the class activation along the NDVI time-series of a single loss event generated from both models. From the two plots it is evident that both models allocated high saliency to periods in the time series where the NDVI value sharply declines. While there seems to be consistency in both models on the regions of importance when making detections, the SDF-TempCNN has higher scores of saliency. Figure 4.8 depicts the qualitative results of the TempCNN and SDF-TempCNN predictions. In the example the SDF-TempCNN appears to detect the clearing event with greater sensitivity.

Table. 7, is a comparison of performance between the TempCNN and SDF-TempCNN Albany Thicket subtypes. While both models showed good performance. On the accuracy metric, the SDF-TempCNN consistently performed better than the TempCNN (> 0.85), the TempCNN ranged between 0.79 and 0.87. The SDF-TempCNN showed the best performance when predicting in the Kasouga Dune Thicket subtype (0.91), while the TempCNN performed best when predicting in the Geluck grassland thicket and Grahamstown Grassland thicket subtypes (0.87).

Table. 6. Performance metrics of the TempCNN and SDF-TempCNN. These results are based on the model trained on the full dataset.

Metric	TempCNN	SDF-TempCNN
Accuracy	0.92	**0.98**
Recall-st	0.93	**0.97**
Recall-ct	0.90	**0.96**
Precision-st	0.93	0.93
Precision-ct	0.93	**0.95**

Table. 7. Performance of the TempCNN and SDF-TempCNN on Albany Thicket subtypes on Albany mesic thicket (AMT), Geluck grassland thicket (GGT), Grahamstown Grassland thicket (GRT), without Nanaga savanna thicket (NST) and Kasouga Dune Thicket (KDT). Here SDF refers to the SDF-TempCNN model and TEM refers to the original TempCNN model.

Metric	**AMT**		**GGT**		**GRT**		**NST**		**KDT**	
	SDF	**TEM**	**SDF**	**TEM**	**SDF**	**TEM**	**SDF**	**TEM**	**SDF**	**TEM**
Accuracy	0.81	0.79	0.88	0.87	0.85	0.87	0.86	0.84	0.91	0.83
Recall-st	0.97	0.79	0.97	0.82	0.98	0.82	0.98	0.90	0.98	0.72
Recall-ct	0.80	0.78	0.96	0.93	0.84	0.80	0.84	0.83	0.92	0.94
Precision-st	0.90	0.78	0.87	0.93	0.89	0.82	0.85	0.84	0.95	0.78
Precision-ct	0.94	0.78	0.88	0.80	0.94	0.93	0.92	0.07	0.90	0.78

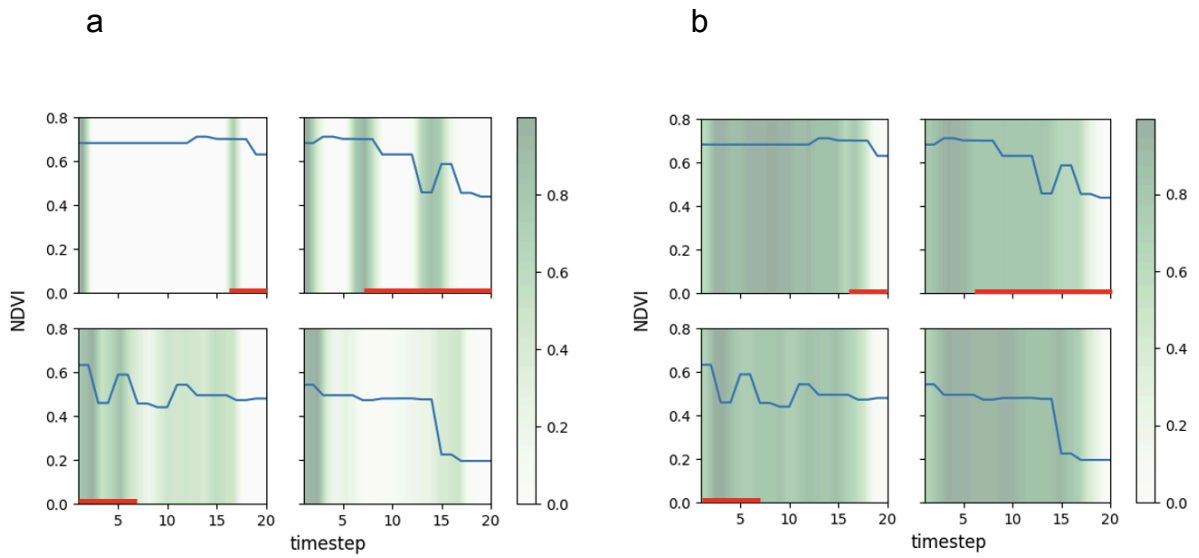


Fig. 4.5) Grad-CAM++ comparison of saliency between the TempCNN (**a**) and the SDF-TempCNN (**b**). Temporally windowed saliency over a single pixel predicted to have undergone clearing. The NDVI values (blue lines) are overlaid to represent the loss of vegetation on the pixel over time. The red vertical lines indicate when the loss event occurred. The darker time regions represent higher saliency, these highlight parts of the input sequence that are most important or carry the most information for detecting changes in vegetarian cover.

4.3.2 Cross-biome generalisation and adaptation

So far TempCNN has been trained and implemented without considering variation within the Albany Thicket Biome. In this section, results of the cross-biome generalisation are presented. MAML's ability to facilitate domain adaptation was assessed on 2-way 1-shot, 3-shot and 5-shot datasets from the meta-train regions (5 Albany Thicket subtypes). MAML-TempCNN was evaluated on one unseen Albany Thicket subtype. The MAML-TempCNN was able to show its ability to reach convergence with a limited amount of data. The 1-shot dataset yielded an accuracy of 0.85, 0.89 was observed from the 3-shot dataset and 0.94 was the accuracy achieved by the 5-shot dataset. Fig. 4.6 a illustrates the performance of the three MAML training shot-sizes. To compare the MAML-TempCNN with the TempCNN and SDF-TempCNN, 1-shot, 3-shot and 5-shot classifiers were also trained on both models. The TempCNN achieved accuracies ranging from 0.47 (1-shot) to 0.53 (5-shot). The SDF-TempCNN yielded accuracies ranging from 0.51 (1-shot) to 0.73

(5-shot). The main observation from this experiment is that MAML frameworks are able to adapt better across the Albany Thicket as well as within Albany Thicket subtypes (Fig. 4.6 a). The MAML-TempCNN was able to maintain consistency in accuracy when trained on 500 iterations (Fig. 4.6 b,c and d). The results show that with one data sample (1-shot) from the meta-testset the MAML-TempCNN can achieve good performance > 0.80 and that increasing the number of shots increases the classification accuracy. To assess the stability of the MAML-TempCNN the best performing shot size (5-shot) was re-trained on 2000 sample iterations. Fig. 4.7 shows that the 5-shot MAML-TempCNN can maintain good performance even over 2000 iterations. The benefit of training on more meta-iterations is that the model may learn richer representations and adapt quicker to various vegetation types.

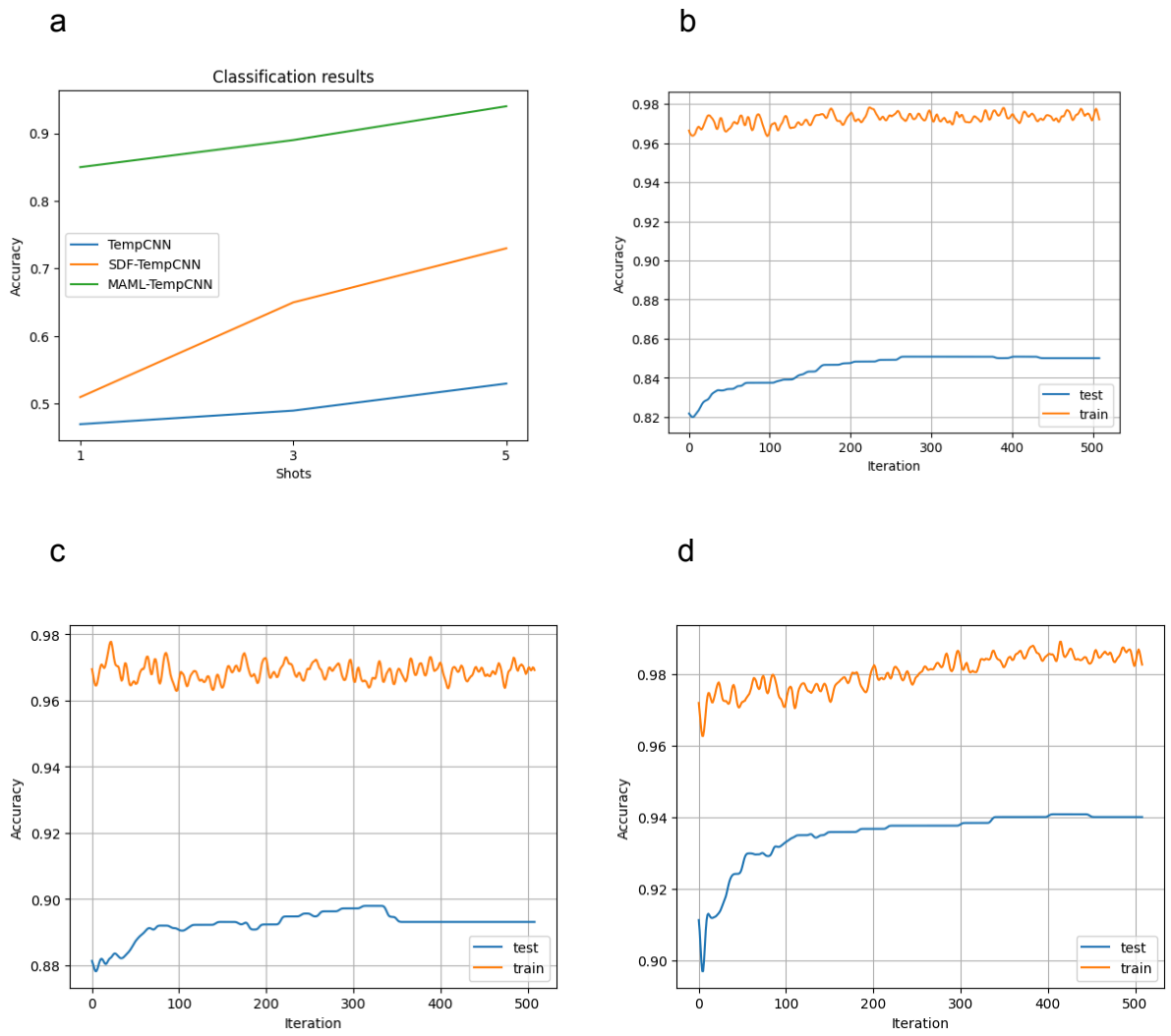


Fig. 4.6) a) A comparison of the effect of the number of shots (1,3 & 5 shot) on classification accuracy for the TempCNN, SDF-TempCCNN, MAML-TempCNN. The effect of the number of shots on the MAML-TemPCNN **b**) 1-shot, **c**) 3-shot and **d**) 5-shot on accuracy over 500 iterations.

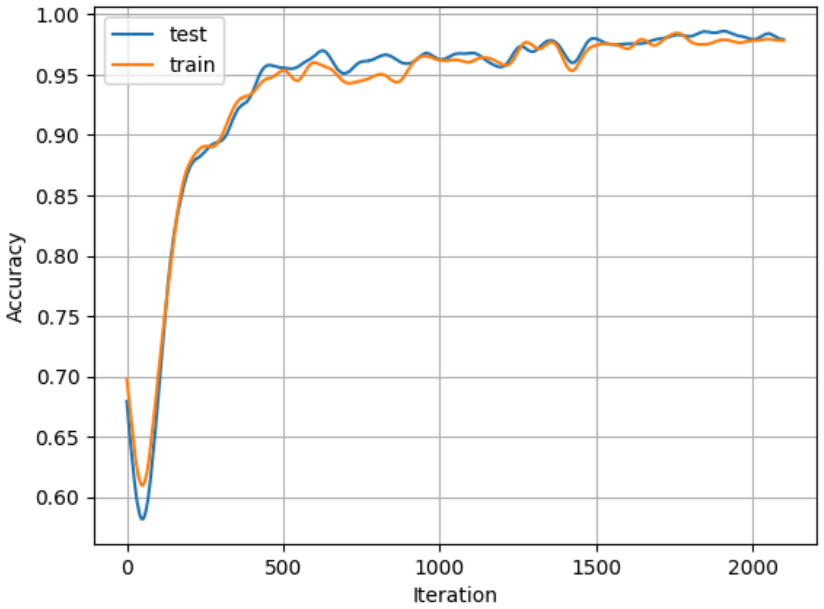


Fig. 4.7) The effect of the number of meta-iterations on the 5-shot MAML-TempCNN over 2000 iterations.

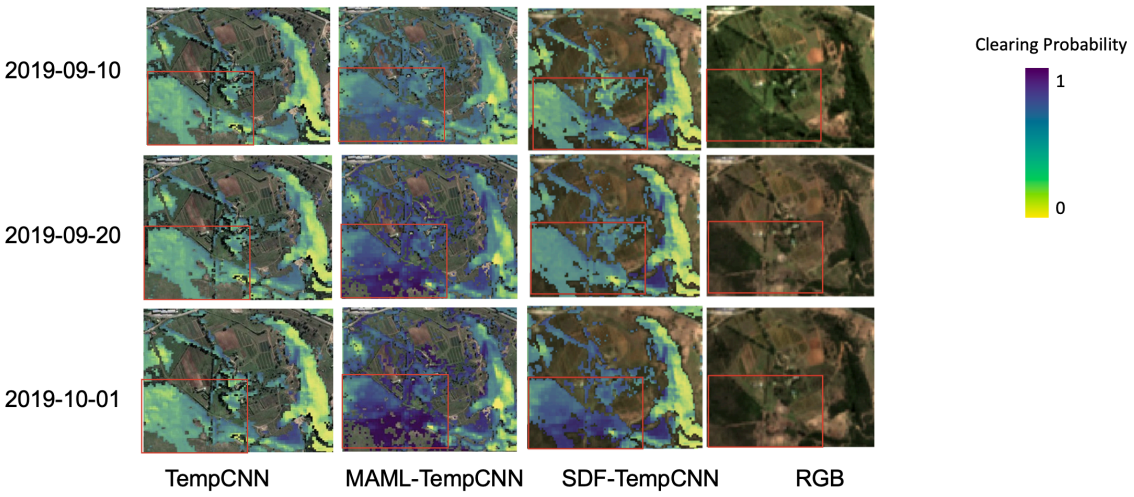


Fig. 4.8) Qualitative example of the detection results from the MAML-TempCNN, SDF-TempCNN and the baseline TempCNN. The example shows a clearing event in progress between 2019-09-10, 2019-09-20 and 2019-10-01 as well as difference in detection accuracy by the three models. The example shows that the SDF-TempCNN performs better than the other two models.

4.4 Discussion

Accurately detecting and documenting clearing events is important for understanding and monitoring the patterns and dynamics of clearing of natural vegetation. Earth observation data coupled with deep learning algorithms have been recognized as efficient in detecting and mapping land cover change. The chapter seeks to overcome the limitations of the TempCNN that was implemented in chapter 3. The aim of this chapter is to develop models that can perform change detection in vegetation cover in changing ecological settings, by developing models that are able to adapt to domain shifts.

Two experiments were set up to evaluate the utility of the SDF-TempCNN in detecting and documenting changes in the Albany Thicket biome, the first was to assess the ability to detect changes in vegetation cover using the SDF-TempCNN. The results showed that the inclusion of surface moisture variables into the TempCNN improved the performance of the model in detecting changes in vegetation cover across the Albany Thicket as well as within individual domains with Albany Thicket biome. The second experiment was to assess the generalisation ability of the TempCNN in shifting domains using MAML. Performance of the MAML-TempCNN varied across the individual Albany Thicket subtypes. The results from this experiment show that when extending to a new domain, using the MAML-TempCNN allowed the model to adapt much faster. The results also highlight the considerations that need to be taken into consideration when implementing the MAML framework, including non-stationarity in vegetation spectral dynamics in earth observation data and environmental conditions across time and within vegetation subtypes. Overall the results from both experiments reveal that continuous change detection can be used to rapidly detect and document changes in vegetation, while accounting for variations in precipitation and surface moisture that are inherent in geographical gradients across different regions and time.

4.4.1 Change Detection

The SDF-TempCNN was compared to a non-SDF-TempCNN as a baseline model that was trained and implemented in the chapter preceding this one. In the previous chapter, adding the Normalised Difference Water Index (NDWI) and Moisture Stress

Index improved performance of the baseline model. However, using a direct measurement of precipitation was deemed more desirable as spectral indices are sensitive to various contaminants such as atmospheric water and are not independent measures of environmental conditions. The same is true for the NDWI and MSI, while it is minimally sensitive to atmospheric water, reflectance from background soils influence water indices values (Gao, 1996). Remote sensing images often violate the assumption of independence because of the correlation and interdependence between spectral bands (Li et al., 2019). This makes the NDWI and MSI susceptible to this interdependence as they are calculated from correlated spectral bands. Land surface temperature along with total precipitation evaporation from bare soil and top of canopy, were used to capture surface-dryness. Table 3 Above shows that by fusing surface-dryness features onto the Sentinel 2 spectral bands and indices, the SDF-TempCNN was able to perform slightly better than the TempCNN. These results indicate that the SDF-TempCNN outperforms the TempCNN in its ability to learn domain contexts and optimise these representations. The SDF-TempCNN also showed improved performances within individual Albany Thicket subtype domains.

The SDF-TempCNN was able to learn domain contextual representations of stable vegetation and cleared vegetation. The benefit of including environmental conditions into machine learning models is that it introduces physical consistency through teaching models to be governed by rules of physics that are at play in Earth systems while conforming to the observational constraints of remote sensing data (García et al., 2013; Reichstein et al., 2019; Shen et al., 2023; Tran et al., 2023). The inclusion of environmental variables further introduces geographical correlation into the model (Li et al., 2018). Zhu et al. (2019), found that the incorporation of both spatial and temporal correlation can enhance the accuracy of prediction of deep learning models. The incorporation of geographical correlation improves accuracy by learning the representations that are shared as a function of geographical proximity. TempCNNs are able to learn temporal dependencies that are present in time series data. The addition of the surface-dryness variables introduces environmental context by leveraging the latent representations that are captured by environmental variables, this allows the SDF-TempCNN to learn complex representations that help the model to make accurate predictions, by utilising the additional information to

activate the temporal regions important for detecting changes in vegetation cover. An example of this can be seen in Fig. 4.5, where the SDF-TempCNN has saliency in temporal regions that coincide with a drop in NDVI values.

These results show the potential of coupling TempCNNs with additional environment variables to improve performance across regions with diverse environmental conditions. Figure 4.8 above, is a qualitative example of the difference between the baseline TempCNN, MAML-TempCNN and the SDF-TempCNN. The example shows an instance where the baseline TempCNN was unable to detect the loss of Albany Thicket cover, the SDF-TempCNN shows better performance when detecting losses in the same region. While the MAML-TempCNN is also able to detect the losses in Albany Thicket coverbIt appears to predict higher probabilities of clearing secondly it detects losses in Albany Thicket that are not present in the RGB composites. Rußwurm et al. (2020), found that for MAML to generalise well on new tasks it requires that it is trained on a large amount of meta-training tasks. Al-Shedivat et al. (2021), found that in addition to training on a large number of tasks, better performance is observed when there are few labels in each task.

The addition of the surface dryness variables acts as modulating or regulating features that force the SDF-TempCNN to be context aware. Textural information is introduced into the SDF branch of the model, this textural information helps the model to consider the biophysical conditions of the area (Nguyen and Pham, 2016). The baseline TempCNN allocated lower probability scores of clearing in most of the pixels compared to the SDF-TempCNN. The accompanying true colour composites images show how the SDF-TempCNN closely approximates the true surface cover; this can be seen in Fig. 4.8.

4.4.2 Cross-biome generalisation and adaptation

After implementing the SDF-TempCNN and observing its improved performance compared to the baseline TempCNN, TempCNN's ability to adapt to new vegetation subtypes was assessed by implementing a MAML-TempCNN. The MAML-TempCNN uses a small number of gradient steps in learning the most optimal initialization parameters that enable domain adaptation of the baseline TempCNN in different vegetation subtypes. The three models were trained on 2-way 1-shot, 3-shot and 5-shot classification tasks sampled from the data in the Albany Thicket subtypes. An

accuracy of 94% observed from the 2way 5-shot on 500 iterations. MAML-TempCNN. While the 5-shot MAML-TempCNN had the best performance, the 1-shot and 3-shot also showed good performances. These results demonstrate that the TempCNN is able to retain information on the difference between Stable and Cleared Thicket when provided at least 1-shot of each land cover the classes of interest. Both the TempCNN and SDF-TempCNN showed greater dependance to the amount of data used to train these models compared to the MAML-TempCNN . The SDF-TempCNN performed decently in the 3-shot and 5-shot training compared to the TempCNN. It was observed that both the TempCNN and MAML-TempCNN were more consistent and stable, this is seen by the marginal increases in accuracies across shot sizes, the SDF-TempCNN on the other hand showed larger increases in accuracy (Fig. 4.8 a). It should be noted that the MAML-TempCNN was trained on the same inputs as the TempCNN.

This experiment assumed that vegetation subtypes are different from each other, however it is possible that the variabilities lay within-distribution of the training set in all three models. Rußwurm et al. (2020), found that the good performance of MAML can be linked to an overlap between the meta-test data distributions. Other researchers including Shen et al. (2021), Gevaert and Belgiu, (2022), Wang et al. (2022), also reported that when both training and testing sets are of a similar distribution, a model is likely to show high performance in the test set. Satrya and Yun (2023), found that the greater the difference in the data distributions the harder domain adaptation becomes. This is referred to as out-of-distribution generalisation, where the test distribution is significantly different from the training distribution. The results of the experiment show that the MAML is a robust alternative to conventional model training especially in few-shot training circumstances. The ability to rapidly adapt parameters means it can be implemented to datasets with diverse land cover classes. Training the 5-shot MAML-TempCNN in 2000 meta iterations shows that it allows the model to robustly learn the initialising parameters that enable fast adaptation and sustained convergence in new domains. This facilitates better generalisation of Albany Thicket clearing across different vegetation subtypes within the biome.

The aim of meta-learning is to train a model to learn the most optimal initialisation parameters that enable fast adaptation to new related tasks. The initialised

parameters from previous tasks in past to new tasks. The transfer of knowledge learned from previous tasks onto new tasks only when the initial tasks and new tasks are related (Garcia et al., 2021). Therefore the implementation of MAML requires that the tasks are homogenous and that there is certainty on the knowledge being shared amongst tasks (Yao et et., 2019). Rußwurm et al, (2020) show that different model weights can result in instances where regions differ or are non-stationary. Here the term non-stationary refers to instances where task distributions evolve over time (Wang et al., 2022), as is the case in many ecological processes (Easdale and Bruzzone, 2022). This means that where distributional shifts exist between tasks or where there are insufficient logical semantic relationships between tasks, MAML models may not learn the most optimal task parameters (Katoch et al., 2020). The concept of non-stationarity applies to natural systems such as the Albany Thicket where over time and space environmental conditions may evolve as a result of natural processes such as shifts in the timing of rainfall seasons. These changes in rainfall timing may not be consistent across regions in the Albany Thicket biome. Such a phenomenon may result in distributional shifts that may cause some Albany Thicket subtypes to fall out of distribution. Figure 3.1, shows minimal differences in the mean spectral signatures of Albany Thicket subtypes, however these spectral signatures may also not be stationary within and across subtypes. The MAML-TempCNN was implemented using Sentinel-2 surface reflectance data using the nanometres (nm) unit of measurement. Inconsistent deviances and processing of these data such as shift to using micrometres (um) as the unit of measurement may also result in some regions being out of distribution. These and other factors may make generalisation across the Albany Thicket subtypes difficult and make the implementation of the MAML-TempCNN challenging. Takimoto et al. (2024) recommend that where task distributions are not similar they should be considered as separate tasks.

4.5 Conclusion

The Albany Thicket biome of South Africa has undergone extensive degradation mainly due to agricultural land expansion. Timeously and accurately detecting and documenting clearing events is important for understanding and monitoring the patterns and dynamics of clearing of Albany Thicket vegetation. Earth observation

data coupled with deep learning algorithms have been recognized as efficient in detecting and mapping land cover change. Therefore, this study proposed a surface-dryness fused TempCNN (SDF-TempCNN) for continuous change detection in the Albany Thicket. The SDF-TempCNN model was compared to a TempCNN model without surface-dryness information. The results highlight the potential of coupling TempCNNs with additional biophysical variables. The TempCNN benefited from the addition of surface-dryness variables which introduce domain contextuality into the model, to facilitate fast adaptation in shifting domains. The learned latent representations are invariant across shifting domains and vegetation subtypes and therefore may be important in discriminating losses in vegetation cover. The speed in the models to adapt to shifting domains when they are provided with limited data was compared. This comparison was facilitated by training 1-shot, 3-shot and 5-shot iterations of each of the three models. The Model Agnostic Machine Learning-Temporal Convolutional Neural Network showed promise in its ability to speed up parameter adaptation for optimal generalisation in few-shot scenarios. This allows models to be expandable in shifting domains or across geographic gradients within the biome. The ability of The Model Agnostic Machine Learning-Temporal Convolutional Neural Network generalise across vegetation types within the biome presents an opportunity for practitioners where data scarcity prevents the implementation of continuous land cover change detection. Training big models that require large amounts of annotated data and expensive computing resources is often not practical, as such the Model Agnostic Machine Learning-Temporal Convolutional Neural Network reduces that need for training on large datasets; this makes Model Agnostic Machine Learning-Temporal Convolutional Neural Network energy and cost efficient for practitioners.

Chapter 5: Summary and Conclusion

5.1 Vegetation loss

Globally 437Mha of tree cover was lost between 2001 to 2021 (www.globalforestwatch.org), and this loss continues. While most of the losses occur in forests, other biomes are also impacted. The Albany Thicket is one of South Africa's nine biomes. The Albany Thicket is recognised for its rich biodiversity, hosting many endemic animal and plant species (Hoffman et al., 1987). Albany Thicket are able to store up to 20 kg/m^2 of carbon (in above- and belowground pools), which is equivalent to mesic forests (Mills and Cowling, 2006). Despite their importance and restoration efforts, Albany Thicket remains under threat of degradation and destruction. Current estimates state that 11% of the Albany Thicket Biome remains pristine with 60-63% being classified as severely degraded (Mucina and Rutherford, 2006; Powell, 2009; Stickler and Shackleton, 2015; Mills et al., 2018).The main drivers of degradation and destruction of the Albany Thicket are pastoralism, the second agent is commercial agriculture land expansion (Mills et al., 2005; Powell, 2009; Stickler and Shackleton, 2015).

The degradation and destruction of Albany Thicket leads to structural and functional changes, causing species loss with negative socioeconomic and environmental impacts (Powell, 2009). The importance of the Albany Thicket biome warrants its protection and conservation. Monitoring natural tree cover plays a central role in this effort. Remote sensing tools have been shown to be effective for monitoring natural resources, including land cover change and object detection (Nagendra et al., 2013; Corbane et al., 2015). However, remote sensing has predominantly focused on forests, with limited application in non-forested ecosystems like the Albany Thicket (Rapinel et al., 2015; Moncrieff, 2022).

5.1.1 Findings from Chapter 2

Several studies have documented the extent and long-term changes in the Albany Thicket biome (Lloyd et al., 2002; Knight and Cowling, 2003; Volk et al., 2003; Carvalho et al., 2021). Except for Carvalho et al. (2021), these studies rely on

manually digitising Albany Thicket from remotely sensed imagery to assess its extent and track changes over time. This process typically involves geographic information system (GIS) software to aid in analysis. However, this approach is time-consuming, expensive, and prone to subjectivity (Showalter et al., 2010; Giri et al., 2015).

There are primarily four automated change detection frameworks including algebra, transformations, classification, and advanced models (van Oort, 2007, Asokan and Anitha, 2019). These are detailed in Chapter one above. While these methods overcome the challenges of manually digitization, they also carry limitations of their own, such as the inability to produce a change matrix by algebraic methods, the inability to provide “from-to” information on changed pixels and class allocation by transformation methods. Most important is that traditional methods assume that all change is linear and abrupt, however changes are often gradual and non-linear.

To address the limitations of traditional change detection frameworks, the first chapter proposed an image differencing-classification hybrid change detection protocol for detecting change. A set of 4-year (2016-2019) Sentinel-2 images representing annual states of the land cover classes present in Alexandria were used to capture their latent change trends and trajectories. The results showed that the combination of Sentinel-2 imagery and the Random Forest classifier have the potential in detecting and classifying changes in Albany Thicket vegetation. The results further indicate the presence of spectral differences between Albany Thicket and other cover types, most important non-Thicket vegetation. The ability to discriminate between Albany Thicket and cover types allows for the use of Sentinel-2 and potentially other medium resolution satellite imagery for monitoring Albany Thicket vegetation dynamics. A total of 186 instances of Albany Thicket that was converted to another land cover type. The detected losses in Albany Thicket amounted to 6.75 km^2 in area extent.

While the methods implemented in the first chapter were successful in detecting and classifying medium-term changes in Albany Thicket vegetation, there is a need for more timely information on the location and date of disturbances and clearing events to allow authorities to intervene in ongoing clearing events. High temporal frequency in change assessments is needed to allow the detection of in-progress change events. This is often performed by independently pre-processing and analysing at

each point in time; this is followed by a comparison of timesteps to determine change.

5.1.2 Findings from Chapter 3

Some scholars have proposed and implemented temporal segmentation as a method for change detection. Temporal segmentation methods include slicing a time-series into multiple segments or portions each segment representing specific time periods. Models are then fitted to each segment and therefore multiple models are fitted to the time series. Chapter three gives a detailed description of temporal segmentation methods. The Continuous Change Detection and Classification (Zhu and Woodcock 2014) algorithm is a temporal segmentation method that is commonly applied. Supervised continuous change detection methods are increasingly being applied. These methods are based on supervised machine learning models that are trained on labelled examples of two or more land cover types and changes in the environment, to identify changes in new, unlabelled remote sensing images. Pelletier et al. (2019), proposed Temporal Convolution Neural Networks, to automatically learn and extract features in the temporal and spectral domain by applying convolutions in the temporal dimension of time-series satellite imagery observations. A detailed description of TempCNNs is given in Chapter three.

A comparison of the Continuous Change Detection and Classification and TempCNN algorithms was conducted in the third chapter. A time-series of Sentinel-2 observations acquired between the years 2016 and 2019 was used to fit the Continuous Change Detection and Classification and train the TempCNN. The results show that TempCNNs can be used to detect changes in Albany Thicket cover within days of the inception of clearing. The performance of TempCCNs is attributed to their incorporation of temporal information from an entire time-series of observations. The primary shortcoming of temporal segmentation methods is that changes are only confirmed when the value of a pixel falls outside the fitted trend after several consecutive observations, for Continuous Change Detection and Classification, a minimum of three consecutive observations falling outside of the fitted pixels values are necessary for a change to be confirmed (Zhu and Woodcock 2014). For example, Continuous Change Detection and Classification fitted with observations from a single Landsat satellite, with a revisit frequency of 16 days,

means that changes are only confirmed at a minimum of 48 days after onset of clearing, assuming that the observations are cloud free.

5.2 Environmental gradients and Albany Thicket monitoring

Aridity and surface moisture were shown to be potentially important biophysical factors influencing the detection accuracy of the TempCNN. Precipitation in the Albany Thicket biome is non-seasonal, however the rate of precipitation varies across the biome. The experiments showed that in regions where the xeric Thicket dominate the landscape the TempCNN tended to make false detections of Cleared Thicket. A reduction of false positives was observed in the TempCNN after including water spectral indices, particularly the Moisture Stress Index. While studying the precipitation and surface moisture dynamics and variability is beyond the scope of this thesis, their role in performance of the models implemented in both the second chapter and third chapter was evident. A key result in the second chapter was in the Random Forest classifier (Breiman, 2001) variable importance assessment. The SWIR-1 bands (features 15, 31, 44 and 53) were shown to be most important in the classification. SWIR bands have been used to study vegetation moisture content (Clark et al., 2005; Chan et al., 2008; Oldeland et al., 2010; Chrysafis et al., 2017; Zhang and Yang 2020; Bhattarai et al; 2021). Another important feature is band 11, which is located at around 1610 nm; which is a region of the electromagnetic spectrum that corresponds with the water absorption feature for vegetation. Oldeland et al. (2010) showed the capacity of these bands in the discrimination of vegetation classes in semi-arid areas.

5.2.1 Findings from Chapter 4

The efficacy of spectral indices in feature extraction has long been recognised (Prasadcet al., 2022). The generation and incorporation of spectral indices is often seamless and efficient. Their efficacy lies in their ability to highlight specific features that are relevant for classification tasks. The addition of the water spectral indices in Chapter 3 resulted in an increase of the TempCNN's accuracy. However, spectral indices like spectral bands are prone to the limitations brought about by the variations in illumination, atmospheric moisture, and gases. Additionally spectral

indices only represent a fraction of the electromagnetic spectrum that is used to calculate them, and some information may be lost or unaccounted for during the index generation. This is particularly evident in the case of multi-spectral images which typically comprise wide bands. Wider spectral bands are prone to spectral confusion as specific absorption peaks are indiscernible (Campbell, 2018; Ge et al., 2022), therefore sources of absorption cannot be attributed with certainty.

To address the shortcomings of spectral indices, some studies have used environmental covariates either directly measured or through robust mathematical modelling. Zhu et al. (2021), found that the addition of auxiliary environmental variables including slope, aspect and elevation improved the estimation of Live fuel moisture content. Zhao et al. (2022), showed that adding directly measured soil evaporative efficiency and other auxiliary environmental variables improved the accuracy of the scaling and mapping of soil moisture. Paisitkriangkrai et al. (2015), showed that a CNN fusing orthophotos and digital surface models outperformed a CNN with orthophotos only in a semantic pixel Labelling task. The primary benefit of including auxiliary environmental variables in remote sensing based deep learning tasks is that they introduce environmental context by leveraging the latent representations that are governed by the natural process (Ge et al., 2022).

These environmental variables are incorporated into machine learning and other modelling frameworks in one of two approaches; namely early fusion and late fusion. Early fusion refers to the approach of merging different datasets at the input level before any processing is performed. Late fusion involves merging datasets at a later stage in the model pipeline, often after features have been extracted from each input. The late fusion approach has been shown to be flexible and provides robustness against overfitting (Hosseinpour et al., 2022).

The fourth chapter proposed a surface-dryness fused TempCNN (SDF- TempCNN) for continuous change detection in the Albany Thicket. The surface-dryness fused TempCNN is a dual-modality convolutional neural network that extracts features in the temporal domain from a time-series of multispectral images, while extracting latent representations in surface dryness variables. The selected environmental variables were total precipitation, evaporation from bare soil and evaporation from top of canopy and land surface temperature. The SDF-TempCNN algorithm was

compared to the TempCNN algorithm implemented in the third chapter. The results demonstrate how TempCNNs can benefit from the inclusion of environmental variables.

A model's ability to generalise is an important metric of evaluation. This refers to how well a model performs a task on new, unseen data that was not included in the training phase. In the case of the surface-dryness fused TempCNN implemented in the fourth chapter, generalisation refers to how well it will detect change in Albany Thicket on newly available or archived data in Alexandria or across the Albany Thicket biome. A surface-dryness fused TempCNN that can generalise is desirable, as it means that in operational use the model would yield reliable detections of land cover change.

To assess cross-biome generalisation of the surface-dryness fused TempCNN, a Meta-Learning approach was adopted through the implementation of Model Agnostic Meta-Learning TempCNN (MAML-TempCNN). MAML-TempCNN is implemented to evaluate the cross-biome generalisation and domain shift adaptation. The framework works by learning the best parameter initialisation which allows for quick adaptation to new tasks. MAML was implemented on five Albany Thicket subtypes and an additional subtype was used to test the performance of MAML. Stochastic Gradient Descent is performed using the difference between weights trained on a small batch of never-before-seen data and the model weights. The results show that the MAML framework can be used to speed up parameter adaptation for optimal generalisation in few-shot scenarios and shifting domains or across geographic gradients. The choice of the source domain is critical in defining the latent representations that are learned from input data which determine the weight matrix that results from the training phase and used to perform Stochastic Gradient Descent.

5.3 Implications for Albany Thicket management

The aim of this thesis was to use multi-temporal satellite imagery to perform continuous detection of land cover change in the Albany Thicket biome. The specific objectives were to 1) generate an Albany Thicket land cover change map documenting the changes in the Albany Thicket biome between 2016 and 2019, 2) To develop a continuous change detection protocol for identifying clearing of Albany

Thicket and 3) to assess the efficacy of incorporating environmental variables in TempCNNs for continuous change detection in the Albany Thicket biome. This is motivated by the observed absence of a system dedicated to monitoring changes in the Albany Thicket biome. The distinctive composition and structure of Albany Thicket necessitate the development of a system that would perform the tasks in part or wholly. The utility of such a system would be to timeously detect and flag possible clearing events. Flagged changes would be made available to authorities to allow for timeous intervention and enforcement of the law and thus becoming a valuable tool in the management of Albany Thicket.

The findings presented in each of the chapters above demonstrate the utility of Earth observation data, products, and methods in the management of Albany Thicket. The management of endangered ecosystems is often multi-faceted, involving protection, restoration, and community engagement. Earth observation can play a role in each aspect of the management of Albany Thicket. Through image analysis, satellite images can be used in the identification of regions that require protection, in restoration programs, image analysis can be used in monitoring the impact restoration activities. Maps generated and other outputs can be used to communicate management plans and goals in community and stakeholder engagements.

This thesis has shown the potential of earth observation data coupled with machine learning methods as tools that can be used in the management of the Albany Thicket biome. The clearing of natural land for agricultural activities and urban development has been identified as the primary cause for habitat loss of terrestrial ecosystems in South Africa (Skowno et al., 2019). Skowno et al. (2019), also note that while the benefits of Earth observation-based products have long been recognised and implemented in the performing of ecosystem assessments in South Africa, the current existing products are outdated and ineffective in timeously detecting fine scale degradation. The methods used in this thesis can be used to produce inputs in conducting future National Biodiversity Assessments, furthermore these methods can be used to develop similar analyses in other ecosystems. These outputs for this thesis can be used to produce fine scale maps of the extent of the Albany Thicket biome. Such maps have the potential for being used in the streamlining of the delineation of the remaining extent of the Albany Thicket biome.

5.4 Reproducibility

An underlying guideline of this project is the repeatability of the methods and analysis. The project was funded as part of an umbrella project termed; “Rapid and Repeatable Tools for monitoring ecosystems (RRetool)”. RRetool aims to develop analyses, methods and protocols that are repeatable by non-expert personnel, through the use of open-source tools, and data and making all analysis and products open and accessible by the public. To this end this project made full use of open, freely available satellite data mainly from the Copernicus Sentinel-2 program and Planetscope data made freely available under the Planet Education and Research Program. Freely available satellite data enables the development of solutions such as those embodied in this project and many others. Open-Source software and resources including Google Earth Engine, Google Colaboratory, Python, TensorFlow and QGIS have been the backbone of this project. Fig 5.1. Shows an example of how the aforementioned tools and services can be integrated to develop and deploy geospatial artificial intelligence solutions such as the ones presented throughout the thesis.

Google Earth Engine is a freely available platform that provides access to geospatial data including Sentinel-2 satellite image (Gorelick et al., 2017). In addition to providing access to data Google Earth Engine provides computing, storage resources and the ability to develop applications for visualisation, interrogation and scientific communication of data and the products of analysis. This project relied on Google Earth Engine for the collection, pre-processing, analysis of satellite imagery and the development of an application for interactively assessing land cover change in Alexandrea between 2016-2019.

Google Earth Engine has a JavaScript and a Python API. The Python API was accessed through Google Colaboratory. Google Colaboratory is a service that provides access to Jupyter Notebooks along with access to computing resources. TensorFlow is an open-source machine learning framework. These tools facilitate the development of the tools and protocols and solutions embodied in this project. Without these resources this project would have been severely hamstrung.

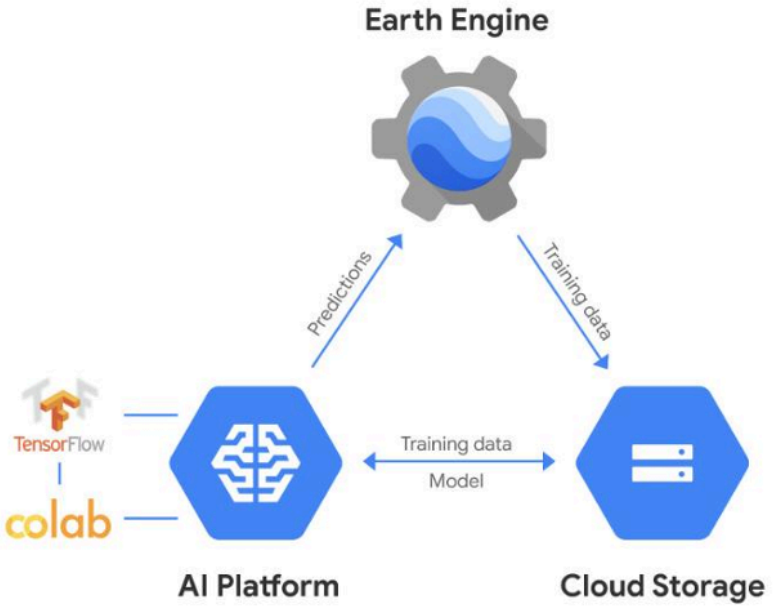


Fig . 5.1) Resources and services for the development of systems of GEOAI based tools. Source: (https://medium.com/google-earth/down-to-earth-with-ai-platform-7bc363abf4fa)

5.5 Advances and Future work

Several advances and alternative methods have been proposed by researchers since the onset of this work presented here. Advances in data collection, storing, and preprocessing have emerged in the form of new software allowing for rapidly intelligently handling and working with remote sensing data. As increased satellite imaging systems are being launched to image the Earth, much work has been done in speeding the downlinking of data through sophisticated mobile terminals that make data available faster than ever before. Such an advance speeds up the time in which land cover change detection is performed. Some researchers have explored the possibility of preprocessing new captured satellite imagery on board satellites and downlinking readily usable data. The rise of cloud computing has accelerated the speed and sheer size of data that researchers have access to and can analyse. While Google Earth Engine is perhaps the most widely used remote sensing big Earth observation data platform, other platforms are increasingly coming online, such as The Microsoft Planetary Computer among others. Such platforms give

researchers access to rapid data discovery, provide computing resources needed to train and deploy big models and analysis and decision support tools. Some systems are open in their make-up allowing for the development of systems and models that rely on federated datasets. Notably advances have been proposed in multi-modality and multi-sensor change detection. These frameworks leverage remotely sensed data from more than one remote sensing platform or from more than one modality. A common example of multi-modality is the fusion of optical and synthetic aperture radar data. An example of multi-sensor change detection is the use of datasets such as the Harmonised Landsat and Sentinel dataset. Both multi-modality and multi-sensor change detection ease the challenges of using earth observation data in cloud prone regions or in cases where short revisit times are necessary for the environment of interest. An exciting and fast-growing field of artificial intelligence is the field of foundation models. Foundation models are large models trained on large sets of unlabelled data. These models are developed with the goal of applying the pretrained models' weights to a wide range of downstream tasks with little labelled data and fine-tuning. Geospatial foundation models are gaining the attention of researchers in the domain of remote sensing. Geospatial foundation models promise to overcome the many challenges of few-shot learning in GeoAI. The advances in remote sensing, computing, and artificial intelligence in general present opportunities for adoption for landcover change detection and other remote sensing applications.

5.6 General Conclusion

Despite their importance and restoration efforts, Albany Thicket ecosystems are subject to severe degradation and destruction. Several researchers have attempted to map and document changes in Albany Thicket cover. While these efforts have yielded valuable information on the geographical location and extent of Albany Thicket, these mapping programs are often performed at an annual or multi-year frequency. Such frequencies are not sufficient for authorities whose roles and activities rely on continuously and rapidly updated information on the extent and changes of Albany Thicket cover. This project aimed at developing protocols for rapidly detecting changes in Albany Thicket cover. To this end it was demonstrated that, using medium resolution satellite imagery, changes in Albany Thicket vegetation can be reliably detected and discerned from changes in other land cover

types. The ability to continuously detect changes using TempCNNs was shown to outperform a state of art algorithm namely, Continuous Change Detection and Classification. Albany Thicket cover dynamics were shown to be embedded within geographical contexts and that geographical gradients in biophysical variables influence the contextual representations learned by the TempCNNs and therefore fusing TempCNNs with biophysical variables including surface dryness information can improve their performance. Finally, it was shown that, using meta-learning, the TempCNN can be adapted to be robust in shifting domains by learning the most optimal parameter initializations that allow for capturing the invariant embeddings that facilitate generalisation in shifting domains and data scarce scenarios. These findings confirm the utility of earth observation data and machine learning models for the monitoring and management of open canopy ecosystems.

6 References

Adjognon, G.S., Rivera-Ballesteros, A., van Soest, D., 2019. Satellite-based tree cover mapping for forest conservation in the drylands of Sub Saharan Africa (SSA): Application to Burkina Faso gazetted forests. Development Engineering 4, 100039. https://doi.org/10.1016/j.deveng.2018.100039\

Agidew, A.A. and Singh, K.N., 2017. 'The implications of land use and land cover changes for rural household food insecurity in the Northeastern highlands of Ethiopia: the case of the Teleyayen sub-watershed', Agriculture & Food Security, 6(1), p. 56. Available at: https://doi.org/10.1186/s40066-017-0134-4

Agjee, N.H., Mutanga, O., Peerbhay, K., Ismail, R., 2018. The Impact of Simulated Spectral Noise on Random Forest and Oblique Random Forest Classification Performance. Journal of Spectroscopy 2018, 1–8. https://doi.org/10.1155/2018/8316918

Airola, A., Pohjankukka, J., Torppa, J., Middleton, M., Nykänen, V., Heikkonen, J., Pahikkala, T., 2019. The spatial leave-pair-out cross-validation method for reliable AUC estimation of spatial classifiers. Data Min Knowl Disc 33, 730–747. https://doi.org/10.1007/s10618-018-00607-x

Aklilu Tesfaye, A., Gessesse Awoke, B., 2020. Evaluation of the saturation property of vegetation indices derived from sentinel-2 in mixed crop-forest ecosystem. Spat. Inf. Res. https://doi.org/10.1007/s41324-020-00339-5

Al-Shedivat, M., Li, L., Xing, E., Talwalkar, A., 2021. On Data Efficiency of Meta-learning. http://arxiv.org/abs/2102.00127

Al-Dousari, A.E., Mishra, A. and Singh, S., 2023. 'Land use land cover change detection and urban sprawl prediction for Kuwait metropolitan region, using multi-layer perceptron neural networks (MLPNN)', The Egyptian Journal of Remote Sensing and Space Science, 26(2), pp. 381–392. Available at: https://doi.org/10.1016/j.ejrs.2023.05.003.

Almutairi, A., Warner, T.A., 2010. Change Detection Accuracy and Image Properties: A Study Using Simulated Data. Remote Sensing 2, 1508–1529. https://doi.org/10.3390/rs2061508

Alonzo, M., Van Den Hoek, J., Murillo-Sandoval, P.J., Steger, C.E., Zinda, J.A., 2021. Mapping and quantifying land cover dynamics using dense remote sensing time series with the user-friendly pyNITA software. Environmental Modelling & Software 145, 105179. https://doi.org/10.1016/j.envsoft.2021.105179

Amorim, J.P., Abreu, P.H., Santos, J., Cortes, M., Vila, V., 2023. Evaluating the faithfulness of saliency maps in explaining deep learning models using realistic perturbations. Information Processing & Management 60, 103225. https://doi.org/10.1016/j.ipm.2022.103225

Andualem, T.G., Belay, G., Guadie, A., 2018. Land Use Change Detection Using Remote Sensing Technology. J Earth Sci Clim Change 9. https://doi.org/10.4172/2157-7617.1000496

Asokan, A., Anitha, J., 2019. Change detection techniques for remote sensing applications: a survey. Earth Sci Inform 12, 143–160. https://doi.org/10.1007/s12145-019-00380-5

Awty-Carroll, K., Bunting, P., Hardy, A., Bell, G., 2019. Using Continuous Change Detection and Classification of Landsat Data to Investigate Long-Term Mangrove Dynamics in the Sundarbans Region. Remote Sensing 11, 2833. https://doi.org/10.3390/rs11232833

Baena, S., Boyd, D.S., Smith, P., Moat, J., Foody, G.M., 2016. Earth observation archives for plant conservation: 50 years monitoring of Itigi-Sumbu thicket. Remote Sens Ecol Conserv 2, 95–106. https://doi.org/10.1002/rse2.18

Banko, G., n.d. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data and of Methods Including Remote Sensing Data in Forest Inventory 43.

Bazi, Y., Bashmal, L., Rahhal, M.M.A., Dayil, R.A., Ajlan, N.A., 2021. Vision Transformers for Remote Sensing Image Classification. Remote Sensing 13, 516. https://doi.org/10.3390/rs13030516

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Bhattarai, R., Rahimzadeh-Bajgiran, P., Weiskittel, A., Meneghini, A., MacLean, D.A., 2021. Spruce budworm tree host species distribution and abundance mapping using multi-temporal Sentinel-1 and Sentinel-2 satellite imagery. ISPRS Journal of Photogrammetry and Remote Sensing 172, 28–40. https://doi.org/10.1016/j.isprsjprs.2020.11.023

Biswas, M., Banerji, S., Mitra, D., 2019. Land-use–land-cover change detection and application of Markov model: A case study of Eastern part of Kolkata. Environ Dev Sustain. https://doi.org/10.1007/s10668-019-00387-4

Botts, E.A., Skowno, A., Driver, A., Holness, S., Maze, K., Smith, T., Daniels, F., Desmet, P., Sink, K., Botha, M., Nel, J., Manuel, J., 2020. More than just a (red) list: Over a decade of using South Africa's threatened ecosystems in policy and practice. Biological Conservation 246, 108559. https://doi.org/10.1016/j.biocon.2020.108559

Bouchard, I., Rancourt, M.E., Aloise, D., Kalaitzis, F., 2022. 'On Transfer Learning for Building Damage Assessment from Satellite Imagery in Emergency Contexts', Remote Sensing, 14(11), p. 2532. Available at: https://doi.org/10.3390/rs14112532.

Butt, A., Shabbir, R., Ahmad, S.S., Aziz, N., 2015. Land use change mapping and analysis using Remote Sensing and GIS: A case study of Simly watershed, Islamabad, Pakistan. The Egyptian Journal of Remote Sensing and Space Science 18, 251–259. https://doi.org/10.1016/j.ejrs.2015.07.003

Carvalho, S.L., Campbell, E.E., Richard du Preez, D., 2022. Degradation of the Albany Thicket Biome. How much of the extent remains 20 years after the initial mapping in 1998? Journal of Arid Environments 196, 104649. https://doi.org/10.1016/j.jaridenv.2021.104649

Chaddad, F., Mello, F.A.O., Tayebi, M., Safanelli, J.L., Campos, L.R., Amorim, M.T.A., Barbosa De Sousa, G.P., Ferreira, T.O., Ruiz, F., Perlatti, F., Greschuk, L.T., Rosin, N.A., Fim Rosas, J.T., Demattê, J.A.M., 2022. Impact of mining-induced deforestation on soil surface temperature and carbon stocks: A case study using remote sensing in the Amazon rainforest. Journal of South American Earth Sciences 119, 103983. https://doi.org/10.1016/j.jsames.2022.103983

Chai, D., Newsam, S., Zhang, H.K., Qiu, Y., Huang, J., 2019. Cloud and cloud shadow detection in Landsat imagery based on deep convolutional neural networks. Remote Sensing of Environment 225, 307–316. https://doi.org/10.1016/j.rse.2019.03.007

Chamarthi, S., Fogelberg, K., Brinker, T.J., Niebling, J., 2024. Mitigating the influence of domain shift in skin lesion classification: A benchmark study of unsupervised domain adaptation methods. Informatics in Medicine Unlocked 44, 101430. https://doi.org/10.1016/j.imu.2023.101430.

Chaemiso, S.E., Kartha, S.A., Pingale, S.M., 2021. Effect of land use/land cover changes on surface water availability in the Omo-Gibe basin, Ethiopia. Hydrological Sciences Journal 66, 1936–1962. https://doi.org/10.1080/02626667.2021.1963442

Chan, J.C.-W., Paelinckx, D., 2008. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Remote Sensing of Environment 112, 2999–3011. https://doi.org/10.1016/j.rse.2008.02.011

Chattopadhyay, A., Sarkar, A., Howlader, P., Balasubramanian, V.N., 2017. Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks. https://doi.org/10.48550/ARXIV.1710.11063

Chen, H., Zhang, H., Chen, K., Zhou, C., Chen, S., Zou, Z., Shi, Z., 2023. Continuous Cross-resolution Remote Sensing Image Change Detection. IEEE

Trans. Geosci. Remote Sensing 61, 1–20. https://doi.org/10.1109/TGRS.2023.3325829

Chen, W., Zhu, D., Ciais, P., Huang, C., Viovy, N., Kageyama, M., 2019. Response of vegetation cover to CO2 and climate changes between Last Glacial Maximum and pre-industrial period in a dynamic global vegetation model. Quaternary Science Reviews 218, 293–305. https://doi.org/10.1016/j.quascirev.2019.06.003

Chen, D., Stow, D.A., Gong, P., 2004. Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case. International Journal of Remote Sensing 25, 2177–2192. https://doi.org/10.1080/01431160310001618464

Cheng, C., Zhang, F., Shi, J., Kung, H.-T., 2022. What is the relationship between land use and surface water quality? A review and prospects from remote sensing perspective. Environ Sci Pollut Res 29, 56887–56907. https://doi.org/10.1007/s11356-022-21348-x

Cho, M.A., Ramoelo, A., 2019. Optimal dates for assessing long-term changes in tree-cover in the semi-arid biomes of South Africa using MODIS NDVI time series (2001–2018). International Journal of Applied Earth Observation and Geoinformation 81, 27–36. https://doi.org/10.1016/j.jag.2019.05.014

Chrysafis, I., Mallinis, G., Gitas, I., Tsakiri-Strati, M., 2017. Estimating Mediterranean forest parameters using multi seasonal Landsat 8 OLI imagery and an ensemble learning method. Remote Sensing of Environment 199, 154–166. https://doi.org/10.1016/j.rse.2017.07.018

Clark, M., Roberts, D., Clark, D., 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sensing of Environment 96, 375–398. https://doi.org/10.1016/j.rse.2005.03.009

Cohen, W.B., Fiorella, M., n.d. Comparison of Methods for Detecting Conifer Forest Change with Thematic Mapper Imagery 16.

Congalton, R., Gu, J., Yadav, K., Thenkabail, P., Ozdogan, M., 2014. Global Land Cover Mapping: A Review and Uncertainty Analysis. Remote Sensing 6, 12070–12093. https://doi.org/10.3390/rs61212070.

Colwell, J.E. 1974. 'Vegetation canopy reflectance', Remote Sensing of Environment, 3(3), pp. 175–183. Available at: https://doi.org/10.1016/0034-4257(74)90003-0.

Coops, N.C., Shang, C., Wulder, M.A., White, J.C., Hermosilla, T., 2020. Change in forest condition: Characterizing non-stand replacing disturbances using time series satellite imagery. Forest Ecology and Management 474, 118370. https://doi.org/10.1016/j.foreco.2020.118370

Corbane, C., Lang, S., Pipkins, K., Alleaume, S., Deshayes, M., García Millán, V. E., Strasser, T., Vanden Borre, J., Toon, S., & Michael, F. 2015. Remote sensing for mapping natural habitats and their conservation status – New opportunities and challenges. *International Journal of Applied Earth Observation and Geoinformation*, *37*, 7–16. https://doi.org/10.1016/j.jag.2014.11.005

Crisóstomo de Castro Filho, H., Abílio de Carvalho Júnior, O., Ferreira de Carvalho, O.L., Pozzobon de Bem, P., dos Santos de Moura, R., Olino de Albuquerque, A., Rosa Silva, C., Guimarães Ferreira, P.H., Fontes Guimarães, R., Trancoso Gomes, R.A., 2020. Rice Crop Detection Using LSTM, Bi-LSTM, and Machine Learning Models from Sentinel-1 Time Series. Remote Sensing 12, 2655. https://doi.org/10.3390/rs12162655

Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P., Linder, H.P., 2009. Phylogenetic biome conservatism on a global scale. Nature 458, 754–756. https://doi.org/10.1038/nature07764

Dean, A.M., Smith, G.M., 2003. An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. International Journal of Remote Sensing 24, 2905–2920. https://doi.org/10.1080/01431160210155910

Deng, J.S., Wang, K., Deng, Y.H., Qi, G.J., n.d. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data 18.

Doan, A.-D., Nguyen, B.L., Gupta, S., Reid, I., Wagner, M., Chin, T.-J., 2024. Assessing domain gap for continual domain adaptation in object detection. Computer Vision and Image Understanding 238, 103885. https://doi.org/10.1016/j.cviu.2023.103885.

Drönner, J., Korfhage, N., Egli, S., Mühling, M., Thies, B., Bendix, J., Freisleben, B., Seeger, B., 2018. Fast Cloud Segmentation Using Convolutional Neural Networks. Remote Sensing 10, 1782. https://doi.org/10.3390/rs10111782

Du, P., Samat, A., Waske, B., Liu, S., Li, Z., 2015. Random Forest and Rotation Forest for fully polarized SAR image classification using polarimetric and spatial features. ISPRS Journal of Photogrammetry and Remote Sensing 105, 38–53. https://doi.org/10.1016/j.isprsjprs.2015.03.002

Easdale, M. H., Perri, D., & Bruzzone, O. A. 2022. Arid and semiarid rangeland responses to non-stationary temporal dynamics of environmental drivers. *Remote Sensing Applications: Society and Environment*, *27*, 100796. https://doi.org/10.1016/j.rsase.2022.100796

Eid, A.N.M., 2020. Coastal wetland vegetation features and digital Change Detection Mapping based on remotely sensed imagery: El-Burullus Lake, Egypt. International Soil and Water Conservation Research 14.

Elagouz, M.H., Abou-Shleel, S.M., Belal, A.A., El-Mohandes, M.A.O., 2019. Detection of land use/cover change in Egyptian Nile Delta using remote sensing. The Egyptian Journal of Remote Sensing and Space Science S1110982317304301. https://doi.org/10.1016/j.ejrs.2018.10.004

Ellis, E.C., Pascual, U., Mertz, O., 2019. Ecosystem services and nature's contribution to people: negotiating diverse values and trade-offs in land systems. Current Opinion in Environmental Sustainability 38, 86–94. https://doi.org/10.1016/j.cosust.2019.05.001

Erpul, G., Huang, Y., Roué, M., Saw, L.G., Mketeni, F.G., n.d. MEMBERS OF THE MANAGEMENT COMMITTEE WHO PROVIDED GUIDANCE FOR THE PRODUCTION OF THIS ASSESSMENT: 48.

Feng, Y., Liu, Y., Liu, D., 2015. Shoreline mapping with cellular automata and the shoreline progradation analysis in Shanghai, China from 1979 to 2008. Arab J Geosci 8, 4337–4351. https://doi.org/10.1007/s12517-014-1515-7

Finn, C., Abbeel, P., Levine, S., 2017. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. http://arxiv.org/abs/1703.03400

Frenay, B., Verleysen, M., 2014. Classification in the Presence of Label Noise: A Survey. IEEE Trans. Neural Netw. Learning Syst. 25, 845–869. https://doi.org/10.1109/TNNLS.2013.2292894

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2, 150066. https://doi.org/10.1038/sdata.2015.66

Ganin, Y. and Lempitsky, V. 2015. 'Unsupervised Domain Adaptation by Backpropagation'. http://proceedings.mlr.press/v37/ganin15.pdf

Gao, B. (1996) 'NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space', Remote Sensing of Environment, 58(3), pp. 257–266. Available at: https://doi.org/10.1016/S0034-4257(96)00067-3.

Gao, Z., Shi, X., Wang, H., Zhu, Y., Wang, Y., Li, M., Yeung, D.-Y., 2023. Earthformer: Exploring Space-Time Transformers for Earth System Forecasting. Available at: http://arxiv.org/abs/2207.05833

Garcia, J.R., Freddi, F., Liao, F.-T., McGowan, J., Nieradzik, T., Shiu, D., Tian, Y., Bernacchia, A., 2021. Meta-Learning with MAML on Trees. http://arxiv.org/abs/2103.04691

García, M., Sandholt, I., Ceccato, P., Ridler, M., Mougin, E., Kergoat, L., Morillas, L., Timouk, F., Fensholt, R., Domingo, F., 2013. Actual evapotranspiration in drylands derived from in-situ and satellite data:

Assessing biophysical constraints. Remote Sensing of Environment 131, 103–118. https://doi.org/10.1016/j.rse.2012.12.016

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., De Bonis, R., Isola, C., Martimort, P., Fernandez, V., 2017. Copernicus Sentinel-2A Calibration and Products Validation Status. Remote Sensing 9, 584. https://doi.org/10.3390/rs9060584

Gawlikowski, J., Saha, S., Kruspe, A., Zhu, X.X., 2022. 'An Advanced Dirichlet Prior Network for Out-of-Distribution Detection in Remote Sensing', IEEE Transactions on Geoscience and Remote Sensing, 60, pp. 1–19. Available at: https://doi.org/10.1109/TGRS.2022.3140324.

Gevaert, C.M. and Belgiu, M. (2022) 'Assessing the generalization capability of deep learning networks for aerial image classification using landscape metrics', International Journal of Applied Earth Observation and Geoinformation, 114, p. 103054. Available at: https://doi.org/10.1016/j.jag.2022.103054.

Ghosh, A., Sharma, R., Joshi, P.K., 2014. Random forest classification of urban landscape using Landsat archive and ancillary data: Combining seasonal maps with decision level fusion. Applied Geography 48, 31–41. https://doi.org/10.1016/j.apgeog.2014.01.003

Ghrefat, H., Awawdeh, M., Howari, F., Al-Rawabdeh, A., 2023. Mineral exploration using multispectral and hyperspectral remote sensing data, in: Geoinformatics for Geosciences. Elsevier, pp. 197–222. https://doi.org/10.1016/B978-0-323-98983-1.00013-2.

Giri, C.P. (Ed.), 2016. Remote Sensing of Land Use and Land Cover: Principles and Applications, 0 ed. CRC Press. https://doi.org/10.1201/b11964

Godet, L., Devictor, V., 2018. What Conservation Does. Trends in Ecology & Evolution 33, 720–730. https://doi.org/10.1016/j.tree.2018.07.004

Gondwe, J.F., Lin, S. and Munthali, R.M. (2021) 'Analysis of Land Use and Land Cover Changes in Urban Areas Using Remote Sensing: Case of Blantyre City', Discrete Dynamics in Nature and Society. Edited by L. Li, 2021, pp. 1–17. Available at: https://doi.org/10.1155/2021/8011565.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Grebner, D.L., Bettinger, P., Siry, J.P., 2013. Forest Products, in: Introduction to Forestry and Natural Resources. Elsevier, pp. 97–124. https://doi.org/10.1016/B978-0-12-386901-2.00004-X

Gregorutti, B., Michel, B., Saint-Pierre, P., 2017. Correlation and variable importance in random forests. Stat Comput 27, 659–678. https://doi.org/10.1007/s11222-016-9646-1

Grigoraș, G., Urițescu, B., 2019. Land Use/Land Cover changes dynamics and their effects on Surface Urban Heat Island in Bucharest, Romania. International Journal of Applied Earth Observation and Geoinformation 80, 115–126. https://doi.org/10.1016/j.jag.2019.03.009

Haas, J., Ban, Y., 2014. Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, River Delta and the Pearl River Delta. International Journal of Applied Earth Observation and Geoinformation 30, 42–55. https://doi.org/10.1016/j.jag.2013.12.012

Han, X., Zhong, Y., Cao, L., Zhang, L., 2017. Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification. Remote Sensing 9, 848. https://doi.org/10.3390/rs9080848.

Haque, Md.I., Basak, R., 2017. Land cover change detection using GIS and remote sensing techniques: A spatio-temporal study on Tanguar Haor, Sunamganj, Bangladesh. The Egyptian Journal of Remote Sensing and Space Science 20, 251–263. https://doi.org/10.1016/j.ejrs.2016.12.003

Hauberg, S., Freifeld, O., Boesen, A., Larsen, L., Fisher,J., & Hansen, L. 2016 Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, PMLR 51:342-350, 2016.

Healey, S.P., Cohen, W.B., Yang, Z., Kenneth Brewer, C., Brooks, E.B., Gorelick, N., Hernandez, A.J., Huang, C., Joseph Hughes, M., Kennedy, R.E., Loveland, T.R., Moisen, G.G., Schroeder, T.A., Stehman, S.V., Vogelmann, J.E., Woodcock, C.E., Yang, L., Zhu, Z., 2018. Mapping forest change using stacked generalization: An ensemble approach. Remote Sensing of Environment 204, 717–728. https://doi.org/10.1016/j.rse.2017.09.029

Henebry, G., 1997. “Advantages of principal components analysis for land cover segmentation from SAR image series.” (1997).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J., 2020. The ERA5 global reanalysis. Quart J Royal Meteoro Soc 146, 1999–2049. https://doi.org/10.1002/qj.3803

Horning, N. 2004. Land cover classification methods, Version 1.0. American Museum of Natural History, Center for Biodiversity and Conservation. http://biodiversityinformatics.amnh.org

Hosseinpour, H., Samadzadegan, F. and Javan, F.D. (2022) ‘CMGFNet: A deep cross-modal gated fusion network for building extraction from very high-resolution remote sensing images’, ISPRS Journal of Photogrammetry

and Remote Sensing, 184, pp. 96–115. Available at: https://doi.org/10.1016/j.isprsjprs.2021.12.007.

Hu, M., Wu, C. and Du, B., 2022. 'Multi-Temporal Spatial-Spectral Comparison Network for Hyperspectral Anomalous Change Detection'. arXiv. Available at: http://arxiv.org/abs/2205.11395

Hu, Q., Sulla-Menashe, D., Xu, B., Yin, H., Tang, H., Yang, P., Wu, W., 2019. A phenology-based spectral and temporal feature selection method for crop mapping from satellite time series. International Journal of Applied Earth Observation and Geoinformation 80, 218–229. https://doi.org/10.1016/j.jag.2019.04.014.

Hu, Y., Dong, Y., Batunacun, S. 2018. An automatic approach for land-change detection and land updates based on integrated NDVI timing analysis and the CVAPS method with GEE support. ISPRS Journal of Photogrammetry and Remote Sensing 146, 347–359. https://doi.org/10.1016/j.isprsjprs.2018.10.008.

Huang, B., Zhao, B., Song, Y., 2018. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. Remote Sensing of Environment 214, 73–86. https://doi.org/10.1016/j.rse.2018.04.050

Hui, D., Biggs, R., Scholes, R.J., Jackson, R.B., 2008. Measuring uncertainty in estimates of biodiversity loss: The example of biodiversity intactness variance. Biological Conservation 141, 1091–1094. https://doi.org/10.1016/j.biocon.2008.02.001

Jaisankar, I., Velmurugan, A., Sivaperuman, C., 2018. Biodiversity Conservation: Issues and Strategies for the Tropical Islands, in: Biodiversity and Climate Change Adaptation in Tropical Islands. Elsevier, pp. 525–552. https://doi.org/10.1016/B978-0-12-813064-3.00019-3

Jha, C.S., Rakesh, Singhal, J., Reddy, C.S., Rajashekar, G., Maity, S., Patnaik, C., Das, A., Misra, A., Singh, C.P., Mohapatra, J., Krishnayya, N.S.R., Kiran, S.,

Townsend, P., Martinez, M.H., 2019. Characterization of Species Diversity and Forest Health using AVIRIS-NG Hyperspectral Remote Sensing Data. Current Science 116, 1124. https://doi.org/10.18520/cs/v116/i7/1124-1135

Kafy, A.-A., Dey, N.N., Al Rakib, A., Rahaman, Z.A., Nasher, N.M.R., Bhatt, A., 2021. Modeling the relationship between land use/land cover and land surface temperature in Dhaka, Bangladesh using CA-ANN algorithm. Environmental Challenges 4, 100190. https://doi.org/10.1016/j.envc.2021.100190

Katoch, S., Thopalli, K., Thiagarajan, J.J., Turaga, P., Spanias, A., 2020. Invenio: Discovering Hidden Relationships Between Tasks/Domains Using Structured Meta Learning. http://arxiv.org/abs/1911.10600

Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 173, 24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010

Kennedy, R.E., Andréfouët, S., Cohen, W.B., Gómez, C., Griffiths, P., Hais, M., Healey, S.P., Helmer, E.H., Hostert, P., Lyons, M.B., Meigs, G.W., Pflugmacher, D., Phinn, S.R., Powell, S.L., Scarth, P., Sen, S., Schroeder, T.A., Schneider, A., Sonnenschein, R., Vogelmann, J.E., Wulder, M.A., Zhu, Z., 2014. Bringing an ecological view of change to Landsat-based remote sensing. Frontiers in Ecology and the Environment 12, 339–346. https://doi.org/10.1890/130066

Kennedy, R.E., Townsend, P.A., Gross, J.E., Cohen, W.B., Bolstad, P., Wang, Y.Q., Adams, P., 2009. Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects. Remote Sensing of Environment 113, 1382–1396. https://doi.org/10.1016/j.rse.2008.07.018

Kerley, G.I.H., Cromsigt, J.P.M.G., 2020. From reindeer to rhino: Reflections on ‘Climate change mitigation and adaptation benefits of wilder rangelands. S. Afr. J. Sci 116. https://doi.org/10.17159/sajs.2020/7604

Khayyam, U., Waseem, S., 2021. Carbon emission as a result of forest land change in Islamabad, Pakistan, in: Climate Change Science. Elsevier, pp. 41–57. https://doi.org/10.1016/B978-0-12-823767-0.00003-3

Kim, D.-W., Yeo, D., Bailer-Jones, C.A.L., Lee, G., 2021. Deep transfer learning for the classification of variable sources. A&A 653, A22. https://doi.org/10.1051/0004-6361/202140369.

Knight, A.T., Cowling, R.M., n.d. Conserving South Africa's 'Lost' Biome 118.

L., A., J., P., J., A., 2023. Automatic forest change detection through a bi-annual time series of satellite imagery: Toward production of an integrated land cover map. International Journal of Applied Earth Observation and Geoinformation 118, 103289. https://doi.org/10.1016/j.jag.2023.103289

Lang, Y., Song, W., 2019. Quantifying and mapping the responses of selected ecosystem services to projected land use changes. Ecological Indicators 102, 186–198. https://doi.org/10.1016/j.ecolind.2019.02.019

Li, J., Huang, X., Gong, J., 2019. Deep neural network for remote-sensing image interpretation: status and perspectives. National Science Review 6, 1082–1086. https://doi.org/10.1093/nsr/nwz058

Li, P., Ren, P., Zhang, X., Wang, Q., Zhu, X., Wang, L., 2018. Region-Wise Deep Feature Representation for Remote Sensing Images. Remote Sensing 10, 871. https://doi.org/10.3390/rs10060871.

Li, T., Shen, H., Yuan, Q., Zhang, X., Zhang, L., 2017. Estimating Ground-Level PM $_{2.5}$ by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach. Geophysical Research Letters 44. https://doi.org/10.1002/2017GL075710

Li, W., Wang, Z., Wang, Y., Wu, J., Wang, J., Jia, Y., Gui, G., 2020. Classification of High-Spatial-Resolution Remote Sensing Scenes Method Using Transfer Learning and Deep Convolutional Neural Network. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 13, 1986–1995. https://doi.org/10.1109/JSTARS.2020.2988477

Li, Y., Shao, Z., Huang, X., Cai, B., Peng, S., 2021. Meta-FSEO: A Meta-Learning Fast Adaptation with Self-Supervised Embedding Optimization for Few-Shot

Remote Sensing Scene Classification. Remote Sensing 13, 2776. https://doi.org/10.3390/rs13142776

Lingua, E., Marchi, N., Bettella, F., Costa, M., Pirotti, F., Piras, M., Garbarino, M., Morresi, D., Marzano, R., 2022. Natural Disturbances and Protection Forests: At the Cutting Edge of Remote Sensing Technologies for the Rapid Assessment of Protective Effects against Rockfall, in: Teich, M., Accastello, C., Perzl, F., Kleemayr, K. (Eds.), Protective Forests as Ecosystem-Based Solution for Disaster Risk Reduction (Eco-DRR). IntechOpen. https://doi.org/10.5772/intechopen.99509

Lu, D., Moran, E., Hetrick, S., 2011. Detection of impervious surface change with multitemporal Landsat images in an urban–rural frontier. ISPRS Journal of Photogrammetry and Remote Sensing 66, 298–306. https://doi.org/10.1016/j.isprsjprs.2010.10.010

Lubke, R.A., Everard, D.A., Jackson, S., 1986. The biomes of the eastern Cape with emphasis on their conservation. Bothalia 16, 251–261. https://doi.org/10.4102/abc.v16i2.1099

Luo, C., Qi, B., Liu, H., Guo, D., Lu, L., Fu, Q., Shao, Y., 2021. Using Time Series Sentinel-1 Images for Object-Oriented Crop Classification in Google Earth Engine. Remote Sensing 13, 561. https://doi.org/10.3390/rs13040561

Lyapustin, A.I. and Kaufman, Y.J. (2001) 'Role of adjacency effect in the remote sensing of aerosol', Journal of Geophysical Research: Atmospheres, 106(D11), pp. 11909–11916. Available at: https://doi.org/10.1029/2000JD900647.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing 152, 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015

Ma, Y., Chen, S., Ermon, S., Lobell, D.B., 2024. Transfer learning in environmental remote sensing. Remote Sensing of Environment 301, 113924. https://doi.org/10.1016/j.rse.2023.113924

Macleod, R.D., Congalton, R.G., 1998. A Quantitative Comparison of Change-Detection Algorithms for Monitoring Eelgrass from Remotely Sensed Data 10.

Magnússon, R.Í, Limpens, J, Kelijn, D, van Huissteden, K, Maximov, T, C, Lobry, S, Heijmans, M,M,P,D, 2021. 'Shrub decline and expansion of wetland vegetation revealed by very high resolution land cover change detection in the Siberian lowland tundra', Science of The Total Environment, 782, p. 146877. Available at: https://doi.org/10.1016/j.scitotenv.2021.146877.

Marsocci, V., Gonthier, N., Garioud, A., Scardapane, S., Mallet, C., 2023. GeoMultiTaskNet: remote sensing unsupervised domain adaptation using geographical coordinates, in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Presented at the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, Vancouver, BC, Canada, pp. 2075–2085. https://doi.org/10.1109/CVPRW59228.2023.00201

Mas, J.-F., 1999. Monitoring land-cover changes: A comparison of change detection techniques. International Journal of Remote Sensing 20, 139–152. https://doi.org/10.1080/014311699213659

Maulana Azad National Institute of Technology, Bhopal, India, Mishra, S., Shrivastava, P., Maulana Azad National Institute of Technology Bhopal, Dhurvey, P., Maulana Azad National Institute of Technology, Bhopal, India, 2017. Change Detection Techniques in Remote Sensing: A Review. IJWMCIS 4, 1–8. https://doi.org/10.21742/ijwmcis.2017.4.1.01

Maxwell, A.E., Warner, T.A., 2020. Thematic Classification Accuracy Assessment with Inherently Uncertain Boundaries: An Argument for Center-Weighted Accuracy Assessment Metrics. Remote Sensing 12, 1905. https://doi.org/10.3390/rs12121905

Menze, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., Hamprecht, F.A., 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and

classification of spectral data. BMC Bioinformatics 10, 213. https://doi.org/10.1186/1471-2105-10-213

Mehra, A., Hamm, J., 2021. Penalty Method for Inversion-Free Deep Bilevel Optimization. http://arxiv.org/abs/1911.03432

Merry, K., Bettinger, P., Crosby, M., Boston, K., 2023. Geographic data processing—raster data, in: Geographic Information System Skills for Foresters and Natural Resource Managers. Elsevier, pp. 231–267. https://doi.org/10.1016/B978-0-323-90519-0.00009-1

Miller, L., Zhu, L., Yebra, M., Rüdiger, C., Webb, G.I., 2022. Multi-modal temporal CNNs for live fuel moisture content estimation. Environmental Modelling & Software 156, 105467. https://doi.org/10.1016/j.envsoft.2022.105467

Mills, A.J., Milton, S.J., Taplin, B.D., Allen, J.L., 2018. Viability of watering Portulacaria afra truncheons to facilitate restoration of subtropical thicket: Results from a nursery experiment and cost model. South African Journal of Botany 115, 58–64. https://doi.org/10.1016/j.sajb.2017.10.017

Moncrieff, G.R., 2021. Locating and Dating Land Cover Change Events in the Renosterveld, a Critically Endangered Shrubland Ecosystem. Remote Sensing 13, 834. https://doi.org/10.3390/rs13050834

Moncrieff, G.R., 2022. Supervised continuous land cover change detection in a critically endangered shrubland ecosystem. Remote Sensing. *14*, 12. https://doi.org/10.3390/rs14122766

Mullissa, A., Vollrath, A., Odongo-Braun, C., Slagter, B., Balling, J., Gou, Y., Gorelick, N., Reiche, J., 2021. Sentinel-1 SAR Backscatter Analysis Ready Data Preparation in Google Earth Engine. Remote Sensing 13, 1954. https://doi.org/10.3390/rs13101954

Myers, N., Mittermeier, R., Mittermeier, C. Biodiversity hotspots for conservation priorities. Nature 403, 853–858 (2000). https://doi.org/10.1038/35002501

Nagendra, H., Lucas, R., Honrado, J. P., Jongman, R. H. G., Tarantino, C., Adamo, M., & Mairota, P. 2013. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators*, *33*, 45–59. https://doi.org/10.1016/j.ecolind.2012.09.014

Namugize, J.N., Jewitt, G., Graham, M., 2018. Effects of land use and land cover changes on water quality in the uMngeni river catchment, South Africa. Physics and Chemistry of the Earth, Parts A/B/C 105, 247–264. https://doi.org/10.1016/j.pce.2018.03.013

Nguyen, T.T.H. and Pham, T.T.T. (2016) 'Incorporating ancillary data into Landsat 8 image classification process: a case study in Hoa Binh, Vietnam', Environmental Earth Sciences, 75(5), p. 430. Available at: https://doi.org/10.1007/s12665-016-5278-1.

Noda, H.M., Muraoka, H., Nasahara, K.N., 2021. Phenology of leaf optical properties and their relationship to mesophyll development in cool-temperate deciduous broad-leaf trees. Agricultural and Forest Meteorology 297, 108236. https://doi.org/10.1016/j.agrformet.2020.108236

Novo-Fernández, A., Franks, S., Wehenkel, C., López-Serrano, P.M., Molinier, M., López-Sánchez, C.A., 2018. Landsat time series analysis for temperate forest cover change detection in the Sierra Madre Occidental, Durango, Mexico. International Journal of Applied Earth Observation and Geoinformation 73, 230–244. https://doi.org/10.1016/j.jag.2018.06.015

Nyamugama, A., Kakembo, V., 2015. Monitoring land Cover Changes and Fragmentation dynamics in the subtropical thicket of the Eastern Cape Province, South Africa. SA J of Geomatics 4, 397. https://doi.org/10.4314/sajg.v4i4.4

Nyborg, J., Pelletier, C., Lefèvre, S., Assent, I., 2022. TimeMatch: Unsupervised cross-region adaptation by temporal shift estimation. ISPRS Journal of Photogrammetry and Remote Sensing 188, 301–313. https://doi.org/10.1016/j.isprsjprs.2022.04.018

Nzuza, P., Ramoelo, A., Odindi, J., Kahinda, J.M., Madonsela, S., 2021. Predicting land degradation using Sentinel-2 and environmental variables in the Lepellane catchment of the Greater Sekhukhune District, South Africa. Physics and Chemistry of the Earth, Parts A/B/C 124, 102931. https://doi.org/10.1016/j.pce.2020.102931

Odebiri, O., Mutanga, O., Odindi, J., Naicker, R., Slotow, R., Mngadi, M., 2023. Evaluation of projected soil organic carbon stocks under future climate and land cover changes in South Africa using a deep learning approach. Journal of Environmental Management 330, 117127. https://doi.org/10.1016/j.jenvman.2022.117127

Odindi, J., Adam, E., Ngubane, Z., Mutanga, O., Slotow, R., 2014. Comparison between WorldView-2 and SPOT-5 images in mapping the bracken fern using the random forest algorithm. Journal of Applied Remote Sensing 8, 18.

Odindi, J.O., Nongebeza, S., Siro, N., 2022. The influence of seasonal land-use-land-cover transformation on thermal characteristics within the city of Pietermaritzburg. SA J of Geomatics 9, 348–364. https://doi.org/10.4314/sajg.v9i2.23

Oldeland, J., Dorigo, W., Lieckfeld, L., Lucieer, A., Jürgens, N., 2010. Combining vegetation indices, constrained ordination and fuzzy classification for mapping semi-natural vegetation units from hyperspectral imagery. Remote Sensing of Environment 12.

Olsen, J., Ceccato, P., Proud, S., Fensholt, R., Grippa, M., Mougin, E., Ardö, J., Sandholt, I., 2013. Relation between Seasonally Detrended Shortwave Infrared Reflectance Data and Land Surface Moisture in Semi-Arid Sahel. Remote Sensing 5, 2898–2927. https://doi.org/10.3390/rs5062898

Palmer, A. R., Ezenne, G. I., Choruma, D. J., Gwate, O., Mantel, S. K., & Tanner, J. L. 2020. A comparison of three models used to determine water fluxes over the Albany Thicket, Eastern Cape, South Africa. *Agricultural and Forest Meteorology*, *288–289*, 107984 https://doi.org/10.1016/j.agrformet.2020.107984

Parker, D.M., 2017. The composition and complexity of the woody and succulent components of Albany thicket with and without elephants. South African Journal of Botany 112, 19–28. https://doi.org/10.1016/j.sajb.2017.05.004

Parr, C.L., Lehmann, C.E.R., Bond, W.J., Hoffmann, W.A., Andersen, A.N., 2014. Tropical grassy biomes: misunderstood, neglected, and under threat. Trends in Ecology & Evolution 29, 205–213. https://doi.org/10.1016/j.tree.2014.02.004

Pelletier, C., Valero, S., Inglada, J., Champion, N., Marais Sicre, C., Dedieu, G., 2017. Effect of Training Class Label Noise on Classification Performances for Land Cover Mapping with Satellite Image Time Series. Remote Sensing 9, 173. https://doi.org/10.3390/rs9020173

Pelletier, C., Webb, G., Petitjean, F., 2019. Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series. Remote Sensing 11, 523. https://doi.org/10.3390/rs11050523

Peng, J., Huang, Y., Sun, W., Chen, N., Ning, Y., Du, Q., 2022. Domain Adaptation in Remote Sensing Image Classification: A Survey. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 15, 9842–9859. https://doi.org/10.1109/JSTARS.2022.3220875

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., Stenseth, N.Chr., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in Ecology & Evolution 20, 503–510. https://doi.org/10.1016/j.tree.2005.05.011

Pham, L., Tran, K., Ngo, D., Lampert, J., Schindler, A., 2022. Remote Sensing Image Classification using Transfer Learning and Attention Based Deep Neural Network. http://arxiv.org/abs/2206.13392

Pierce, S.M., Cowling, R.M., 1984. Phenology of fynbos, renosterveld and subtropical thicket in the south eastern Cape. South African Journal of Botany 3, 1–16. https://doi.org/10.1016/S0022-4618(16)30074-2

Powell, M.J., n.d. 2009. Restoration of degraded subtropical thickets in the Baviaanskloof Megareserve, South Africa 164.

Poursanidis, D., Traganos, D., Reinartz, P., Chrysoulakis, N., 2019. On the use of Sentinel-2 for coastal habitat mapping and satellite-derived bathymetry estimation using downscaled coastal aerosol band. International Journal of Applied Earth Observation and Geoinformation 80, 58–70. https://doi.org/10.1016/j.jag.2019.03.012

Pu, R. .2021. 'Assessing scaling effect in downscaling land surface temperature in a heterogenous urban environment', International Journal of Applied Earth Observation and Geoinformation, 96, p. 102256. Available at: https://doi.org/10.1016/j.jag.2020.102256.

Ramoelo, A., Cho, M.A., Mathieu, R., Madonsela, S., van de Kerchove, R., Kaszta, Z., Wolff, E., 2015. Monitoring grass nutrients and biomass as indicators of rangeland quality and quantity using random forest modelling and WorldView-2 data. International Journal of Applied Earth Observation and Geoinformation 43, 43–54. https://doi.org/10.1016/j.jag.2014.12.010

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566, 195–204. https://doi.org/10.1038/s41586-019-0912-1

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40, 913–929. https://doi.org/10.1111/ecog.02881

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 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. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Roe, D., 2019. Biodiversity loss—more than an environmental emergency. The Lancet Planetary Health 3, e287–e289. https://doi.org/10.1016/S2542-5196(19)30113-5

Rußwurm, M., Körner, M., 2018. Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders. IJGI 7, 129. https://doi.org/10.3390/ijgi7040129

Sage, R.F., Sage, T.L., 2013. C4 Plants, in: Levin, S.A. (Ed.), Encyclopedia of Biodiversity (Second Edition). Academic Press, Waltham, pp. 361–381. https://doi.org/10.1016/B978-0-12-384719-5.00192-1

Salih, A.A.M., Ganawa, E.-T., Elmahl, A.A., 2017. Spectral mixture analysis (SMA) and change vector analysis (CVA) methods for monitoring and mapping land degradation/desertification in arid and semiarid areas (Sudan), using Landsat imagery. The Egyptian Journal of Remote Sensing and Space Science 20, S21–S29. https://doi.org/10.1016/j.ejrs.2016.12.008

Schiefer, F., Kattenborn, T., Frick, A., Frey, J., Schall, P., Koch, B., Schmidtlein, S., 2020. Mapping forest tree species in high resolution UAV-based RGB-imagery by means of convolutional neural networks. ISPRS Journal of Photogrammetry and Remote Sensing 170, 205–215. https://doi.org/10.1016/j.isprsjprs.2020.10.015

Schmidt, A., Kerley, G., Watson, L., 2019. Threshold changes in the structure and composition of bush clumps in piospheres in an arid thicket mosaic, South Africa. J. Arid Environ. 169, 19–28. https://doi.org/10.1016/j.jaridenv.2019.103999

Schramm, M., Pebesma, E., Milenković, M., Foresta, L., Dries, J., Jacob, A., Wagner, W., Mohr, M., Neteler, M., Kadunc, M., Miksa, T., Kempeneers, P., Verbesselt, J., Gößwein, B., Navacchi, C., Lippens, S., Reiche, J., 2021. The

openEO API–Harmonising the Use of Earth Observation Cloud Services Using Virtual Data Cube Functionalities. Remote Sensing 13, 1125. https://doi.org/10.3390/rs13061125

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., Brenning, A., 2019. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. Ecological Modelling 406, 109–120. https://doi.org/10.1016/j.ecolmodel.2019.06.002

Sedano, F., Gong, P., Ferrao, M., 2005. Land cover assessment with MODIS imagery in southern African Miombo ecosystems. Remote Sensing of Environment 98, 429–441. https://doi.org/10.1016/j.rse.2005.08.009

Selman, C., Lingard, S., Gems, D., Partridge, L., Withers, D.J., Taguchi, A., White, M.F., n.d. COMMENT ON "Brain IRS2 Signaling Coordinates Life Span and Nutrient Homeostasis" 2.

Sencaki, D.B., Putri, M.N., Santosa, B.H., Arfah, S., Arifandri, R., Afifuddin, Habibie, M.I., Putra, P.K., Anatoly, N., Permata, Z.D.O., Frederik, M.C.G., Agustan, Sumargana, L., Priyadi, H., 2023. Land cover multiclass classification of wonosobo, Indonesia with time series-based one-dimensional deep learning model. Remote Sensing Applications: Society and Environment 32, 101040. https://doi.org/10.1016/j.rsase.2023.101040

Sharma, R.C., Hara, K., Hirayama, H., 2017. A Machine Learning and Cross-Validation Approach for the Discrimination of Vegetation Physiognomic Types Using Satellite Based Multispectral and Multitemporal Data. Scientifica 2017, 1–8. https://doi.org/10.1155/2017/9806479

Sharma, Y. and Ross, R. (no date) 'Less is More when Applying Transfer Learning to Multi-Spectral Data'.

Shen, C., Appling, A.P., Gentine, P., Bandai, T., Gupta, H., Tartakovsky, A., Baity-Jesi, M., Fenicia, F., Kifer, D., Li, L., Liu, X., Ren, W., Zheng, Y., Harman, C.J., Clark, M., Farthing, M., Feng, D., Kumar, P., Aboelyazeed, D., Rahmani, F., Song, Y., Beck, H.E., Bindas, T., Dwivedi, D., Fang, K., Höge, M., Rackauckas, C., Mohanty, B., Roy, T., Xu, C., Lawson, K., 2023. Differentiable

modelling to unify machine learning and physical models for geosciences. Nat Rev Earth Environ 4, 552–567. https://doi.org/10.1038/s43017-023-00450-9

Shen, Z., He, Y., Zhang, X., Xu, R., Yu, H., Cui, P., 2023. Towards Out-Of-Distribution Generalization: A Survey.http://arxiv.org/abs/2108.13624

Shendryk, Y., Rist, Y., Ticehurst, C., Thorburn, P., 2019. Deep learning for multi-modal classification of cloud, shadow and land cover scenes in PlanetScope and Sentinel-2 imagery. ISPRS Journal of Photogrammetry and Remote Sensing 157, 124–136. https://doi.org/10.1016/j.isprsjprs.2019.08.018

Shorten, C. and Khoshgoftaar, T.M. (2019) 'A survey on Image Data Augmentation for Deep Learning', Journal of Big Data, 6(1), p. 60. Available at: https://doi.org/10.1186/s40537-019-0197-0.

Showalter, P.S., 2021. A Thematic Mapper Analysis of the Prehistoric Hohokam Canal System, Phoenix, Arizona 15.

Slingsby, J.A., Moncrieff, G.R., Wilson, A.M., 2020. Near-real time forecasting and change detection for an open ecosystem with complex natural dynamics. ISPRS Journal of Photogrammetry and Remote Sensing 166, 15–25. https://doi.org/10.1016/j.isprsjprs.2020.05.017

Smart, K.G. 2016. The role of vegetation in regional climate regulation feedback processes, University of Witwatersrand, Johannesburg, http://wiredspace.wits.ac.za/handle/10539/21695

Smissen, P.J., Rowe, K.C., 2018. Repeated biome transitions in the evolution of Australian rodents. Molecular Phylogenetics and Evolution 128, 182–191. https://doi.org/10.1016/j.ympev.2018.07.015

Sohl, T., Sleeter, B., 2012. Role of Remote Sensing for Land-Use and Land-Cover Change Modeling, in: Remote Sensing of Land Use and Land Cover. CRC Press, pp. 225–240. https://doi.org/10.1201/b11964-18

Solomon, A. and Sundaresan, A.A., 2023. 'Land-cover classification with hyperspectral remote sensing image using CNN and spectral band selection', *Remote Sensing Applications: Society and Environment*, 31, p. 100986. Available at: https://doi.org/10.1016/j.rsase.2023.100986.

Sousa, D., Small, C., 2019. Mapping and Monitoring Rice Agriculture with Multisensor Temporal Mixture Models. Remote Sensing 11, 181. https://doi.org/10.3390/rs11020181

South African Government (2008) 'National Environmental Management Act'. https://www.gov.za/documents/national-environmental-management-act

Stacke, K., Eilertsen, G., Unger, J., Lundström, C., 2019. A Closer Look at Domain Shift for Deep Learning in Histopathology. https://arxiv.org/pdf/1909.11575

Sterckx, S., Knaeps, E., Ruddick, K., 2011. Detection and correction of adjacency effects in hyperspectral airborne data of coastal and inland waters: the use of the near infrared similarity spectrum. International Journal of Remote Sensing 32, 6479–6505. https://doi.org/10.1080/01431161.2010.512930

Stickler, M.M., Shackleton, C.M., 2015. Local Wood Demand, Land Cover Change and the State of Albany Thicket on an Urban Commonage in the Eastern Cape, South Africa. Environmental Management 55, 411–422. https://doi.org/10.1007/s00267-014-0396-6

Stocker, T.F., Intergovernmental Panel on Climate Change (Eds.), 2013. Climate change 2013: the physical science basis ; summary for policymakers, a report of Working Group I of the IPCC, technical summary, a report accepted by Working Group I of the IPCC but not approved in detail and frequently asked questions ; part of the Working Group I contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change, New York.

Sun, S., Hu, C., Tunnell, J.W., 2015. Surface oil footprint and trajectory of the Ixtoc-I oil spill determined from Landsat/MSS and CZCS observations. Marine Pollution Bulletin 101, 632–641. https://doi.org/10.1016/j.marpolbul.2015.10.036

Sun, Z., Di, L., Fang, H., Burgess, A., 2020. Deep Learning Classification for Crop Types in North Dakota. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 13, 2200–2213. https://doi.org/10.1109/JSTARS.2020.2990104

Suwanprasit, C., Srichai, N., 2012. Impacts of spatial resolution on land cover classification. APAN Proceedings 33, 39. https://doi.org/10.7125/APAN.33.4

Takahashi, R., Matsubara, T. and Uehara, K. (2020) 'Data Augmentation using Random Image Cropping and Patching for Deep CNNs', IEEE Transactions on Circuits and Systems for Video Technology, 30(9), pp. 2917–2931. Available at: https://doi.org/10.1109/TCSVT.2019.2935128.

Takimoto, Y., Tanaka, Y., Iwata, T., Okawa, M., Kim, H., Toda, H., & Kurashima, T. 2024. *Meta-Learning for neural network-based temporal point processes*. arXiv.Org. https://arxiv.org/abs/2401.15846

Tan, J., Yu, D., Li, Q., Tan, X., Zhou, W., 2020. Spatial relationship between land-use/land-cover change and land surface temperature in the Dongting Lake area, China. Sci Rep 10, 9245. https://doi.org/10.1038/s41598-020-66168-6

Tariq, A., Shu, H., 2020. CA-Markov Chain Analysis of Seasonal Land Surface Temperature and Land Use Land Cover Change Using Optical Multi-Temporal Satellite Data of Faisalabad, Pakistan. Remote Sensing 12, 3402. https://doi.org/10.3390/rs12203402

Tran, B.N., Van Der Kwast, J., Seyoum, S., Uijlenhoet, R., Jewitt, G., Mul, M., 2023. Uncertainty assessment of satellite remote-sensing-based evapotranspiration estimates: a systematic review of methods and gaps. Hydrol. Earth Syst. Sci. 27, 4505–4528. https://doi.org/10.5194/hess-27-4505-2023

Tsai, Y.-H., Hung, W.-C., Schulter, S., Sohn, K., Yang, M.-H., Chandraker, M., 2020. Learning to Adapt Structured Output Space for Semantic Segmentation. http://arxiv.org/abs/1802.10349.

Tueller, P.T. 1987. 'Remote sensing science applications in arid environments', Remote Sensing of Environment, 23(2), pp. 143–154. Available at: https://doi.org/10.1016/0034-4257(87)90034-4.

Tuia, D., Persello, C. and Bruzzone, L. 2016. 'Recent Advances in Domain Adaptation for the Classification of Remote Sensing Data', IEEE Geoscience and Remote Sensing Magazine, 4(2), pp. 41–57. Available at: https://doi.org/10.1109/MGRS.2016.2548504.

van Oort, P.A.J., 2007. Interpreting the change detection error matrix. Remote Sensing of Environment 108, 1–8. https://doi.org/10.1016/j.rse.2006.10.012

Vandas, S.J., Winter, T.C., Battaglin, W.A., 2002. Water and the environment, AGI environmental awareness series. American Geological Institute, Alexandria, VA.

Venter, O., Sanderson, E.W., Magrach, A., Allan, J.R., Beher, J., Jones, K.R., Possingham, H.P., Laurance, W.F., Wood, P., Fekete, B.M., Levy, M.A., Watson, J.E.M., 2016. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. Nat Commun 7, 12558.

Wang, B., Jia, K., Liang, S., Xie, X., Wei, X., Zhao, X., Yao, Y., Zhang, X., 2018. Assessment of Sentinel-2 MSI Spectral Band Reflectances for Estimating Fractional Vegetation Cover. Remote Sensing 10, 1927. https://doi.org/10.3390/rs10121927

Wang, H., Li, J., Wu, H., Hovy, E., Sun, Y., 2023. Pre-Trained Language Models and Their Applications. Engineering 25, 51–65. https://doi.org/10.1016/j.eng.2022.04.024

Wang, M., He, G., Zhang, Z., Wang, G., Wang, Z., Yin, R., Cui, S., Wu, Z., Cao, X., 2019. 'A radiance-based split-window algorithm for land surface temperature retrieval: Theory and application to MODIS data', International Journal of Applied Earth Observation and Geoinformation, 76, pp. 204–217. Available at: https://doi.org/10.1016/j.jag.2018.11.015.

Wang, M., Tan, K., Jia, X., Wang, X., Chen, Y., 2020. A Deep Siamese Network with Hybrid Convolutional Feature Extraction Module for Change Detection Based on Multi-sensor Remote Sensing Images. Remote Sensing 12, 205. https://doi.org/10.3390/rs12020205

Wang, W., Dou, S., Jiang, Z., Sun, L., 2018. A Fast Dense Spectral–Spatial Convolution Network Framework for Hyperspectral Images Classification. Remote Sensing 10, 1068. https://doi.org/10.3390/rs10071068

Wang, Z., Shen, L., Fang, L., Suo, Q., Zhan, D., Duan, T., Gao, M., 2022. Meta-Learning with Less Forgetting on Large-Scale Non-Stationary Task Distributions. http://arxiv.org/abs/2209.01501

Wheeler, K.I. and Dietze, M.C. (2019) 'A Statistical Model for Estimating Midday NDVI from the Geostationary Operational Environmental Satellite (GOES) 16 and 17', Remote Sensing, 11(21), p. 2507. Available at: https://doi.org/10.3390/rs11212507.

Whitehorn, P.R., Navarro, L.M., Schröter, M., Fernandez, M., Rotllan-Puig, X., Marques, A., 2019. Mainstreaming biodiversity: A review of national strategies. Biological Conservation 235, 157–163. https://doi.org/10.1016/j.biocon.2019.04.016

Woodcock, C.E., Loveland, T.R., Herold, M., Bauer, M.E., 2020. Transitioning from change detection to monitoring with remote sensing: A paradigm shift. Remote Sens. Environ. 238, 111558. https://doi.org/10.1016/j.rse.2019.111558

Woodward, F.I., Lomas, M.R., Kelly, C.K., 2004. Global climate and the distribution of plant biomes. Phil. Trans. R. Soc. Lond. B 359, 1465–1476. https://doi.org/10.1098/rstb.2004.1525

WWF, 2022, Living Planet Report 2022 – Building a naturepositive society. Almond, R.E.A., Grooten, M., Juffe Bignoli, D. & Petersen, T. (Eds). WWF, Gland, Switzerland.

Xie, J., Qi, T., Hu, W., Huang, H., Chen, B., Zhang, J., 2022. Retrieval of Live Fuel Moisture Content Based on Multi-Source Remote Sensing Data and Ensemble Deep Learning Model. Remote Sensing 14, 4378. https://doi.org/10.3390/rs14174378.

Xu, Jinfan, Zhu, Y., Zhong, R., Lin, Z., Xu, Jialu, Jiang, H., Huang, J., Li, H., Lin, T., 2020. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. Remote Sensing of Environment 247, 111946

Xu, Q., Shi, Y., Yuan, X., Zhu, X.X., 2023. Universal Domain Adaptation for Remote Sensing Image Scene Classification. IEEE Trans. Geosci. Remote Sensing 61, 1–15. https://doi.org/10.1109/TGRS.2023.3235988

Xue, J., Su, B., 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. Journal of Sensors 2017, 1–17. https://doi.org/10.1155/2017/1353691

Yan, W., Wang, Y., Gu, S., Huang, L., Yan, F., Xia, L., Tao, Q., 2019. The Domain Shift Problem of Medical Image Segmentation and Vendor-Adaptation by Unet-GAN. https://arxiv.org/pdf/1910.13681.pdf

Yang, C., Rottensteiner, F., Heipke, C., 2021. A hierarchical deep learning framework for the consistent classification of land use objects in geospatial databases. ISPRS Journal of Photogrammetry and Remote Sensing 177, 38–56. https://doi.org/10.1016/j.isprsjprs.2021.04.022

Yao, H., Wei, Y., Huang, J., Li, Z., 2019. Hierarchically Structured Meta-learning. http://arxiv.org/abs/1905.05301

Yao, P., Shi, J., Zhao, T., Lu, H., Al-Yaari, A., 2017. Rebuilding Long Time Series Global Soil Moisture Products Using the Neural Network Adopted the

Microwave Vegetation Index. Remote Sens. 2017, 9, 35. Remote Sensing 9, 849. https://doi.org/10.3390/rs9080849

Ye, S., Rogan, J., Zhu, Z., Eastman, J.R., 2021. A near-real-time approach for monitoring forest disturbance using Landsat time series: stochastic continuous change detection. Remote Sensing of Environment 252, 112167. https://doi.org/10.1016/j.rse.2020.112167

Yu, S., Liu, L., Wang, Z., Dai, G., Xie, Y., 2019. Transferring deep neural networks for the differentiation of mammographic breast lesions. Sci. China Technol. Sci. 62, 441–447. https://doi.org/10.1007/s11431-017-9317-3

Yu, W., Li, J., Liu, Q., Zhao, J., Dong, Y., Zhu, X., Lin, S., Zhang, H., Zhang, Z., 2021. Gap Filling for Historical Landsat NDVI Time Series by Integrating Climate Data. Remote Sensing 13, 484. https://doi.org/10.3390/rs13030484

Yuan, J., Cohen, M.J., 2020. Remote detection of ecosystem degradation in the Everglades ridge-slough landscape. Remote Sensing of Environment 247, 111917. https://doi.org/10.1016/j.rse.2020.111917

Zeng, Q., & Geng, J. 2022. Task-specific contrastive learning for few-shot remote sensing image scene classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *191*, 143–154. https://doi.org/10.1016/j.isprsjprs.2022.07.013

Zhang, F., Yang, X., 2020. Improving land cover classification in an urbanized coastal area by random forests: The role of variable selection. Remote Sensing of Environment 251, 112105. https://doi.org/10.1016/j.rse.2020.112105

Zhang, Y., Hui, J., Qin, Q., Sun, Y., Zhang, T., Sun, H., Li, M., 2021. Transfer-learning-based approach for leaf chlorophyll content estimation of winter wheat from hyperspectral data. Remote Sensing of Environment 267, 112724. https://doi.org/10.1016/j.rse.2021.112724

Zhou, Y., Yu, L., 2023. Few-shot learning via weighted prototypes from graph structure. Pattern Recognition Letters 176, 230–235. https://doi.org/10.1016/j.patrec.2023.11.017

Zhu, L., Webb, G.I., Yebra, M., Scortechini, G., Miller, L., Petitjean, F., 2021. Live fuel moisture content estimation from MODIS: A deep learning approach. ISPRS Journal of Photogrammetry and Remote Sensing 179, 81–91. https://doi.org/10.1016/j.isprsjprs.2021.07.010

Zhu, J., Yang, K., Guan, N., Yi, X., & Qiu, C. 2023. HCPNet: Learning discriminative prototypes for few-shot remote sensing image scene classification. *International Journal of Applied Earth Observation and Geoinformation*, *123*, 103447. https://doi.org/10.1016/j.jag.2023.103447

Zhu, Q., Chen, J., Shi, D., Zhu, L., Bai, X., Duan, X., Liu, Y., 2020. Learning Temporal and Spatial Correlations Jointly: A Unified Framework for Wind Speed Prediction. IEEE Trans. Sustain. Energy 11, 509–523. https://doi.org/10.1109/TSTE.2019.2897136

Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote Sensing of Environment 144, 152–171. https://doi.org/10.1016/j.rse.2014.01.011

Zhu, Z., Zhang, J., Yang, Z., Aljaddani, A.H., Cohen, W.B., Qiu, S., Zhou, C., 2020. Continuous monitoring of land disturbance based on Landsat time series. Remote Sensing of Environment 238, 111116. https://doi.org/10.1016/j.rse.2019.03.009

7 Appendix

Code and software will be made available on the github repository below

https://github.com/PixelGSD/Albany-change-detection-