

**Mapping drought stress in commercial *Eucalyptus* forest plantations using
remotely sensed techniques in Southern Africa**

by

Andile Lubanyana

214580702



**Submitted in fulfillment of the academic requirements for the degree of
Master of Science in the Discipline of Geography in the
School of Environmental Sciences, Faculty of Science and Agriculture.
University of KwaZulu-Natal, Pietermaritzburg**

December 2021

Abstract

Drought is one of the least understood and hazardous natural disasters that leave many parts of the world devastated. To improve understanding and detection of drought onset, remote sensing technology is required to map drought affected areas as it covers large geographical areas. The study aimed to evaluate the utility of a cost-effective Landsat 8 imagery in mapping the spatial extent of drought prone *Eucalyptus dunnii* plantations. The first objective was to compare the utility of Landsat spectra with a combination of vegetation indices to detect drought affected plantations using the Stochastic gradient boosting algorithm. The test datasets showed that using Landsat 8 spectra only produced an overall accuracy of 74.70% and a kappa value of 0.59. The integration of Landsat 8 spectra with vegetation indices produced an overall accuracy of 83.13% and a kappa of 0.76. The second objective of this study was to do a trend analysis of vegetation health during drought. The normalized difference vegetation index (NDVI) values fluctuated over the years where 2013 had the highest value of 0.68 and 2015 the lowest NDVI of 0.55 and the normalized difference water index (NDWI) had the lowest value in 2015. Most indices showed a similar trend where 2013 had the highest index value and 2015 the lowest. The third objective was to do a trend analysis of rainfall and temperature during drought. The rainfall trend analysis from 2013 to 2017 indicated that the month of February 2017 received the highest rainfall of 154 mm. In addition, July of 2016 received the highest rainfall compared to 2013, 2014, 2015 and 2017 with rainfalls of 6.4 mm, 0.6 mm, 28 mm, and 1 mm, respectively. The temperature trend analysis from 2013 to 2017 indicated that December 2015 had the highest temperature of 28 ° C compared to December of 2013 2014, 2016 and 2017 with temperatures of 24°C, 25°C, 27°C, 24°C, respectively. Furthermore, it was also noted that June 2017 had the highest temperature of 23°C while June 2015 had the lowest at 20°C. The fourth objective of this study was to compare the utility of topographical variables with a combination of Landsat vegetation indices to detect drought affected plantations using the One class support vector machine algorithm. The multiclass support vector machine using Landsat vegetation indices and topographical variables produced an overall accuracy of 73.86% and a kappa value of 0.71 with user's and producer's accuracies ranging between 61% to 69% for drought damaged trees, while for healthy trees ranged from 84% to 90%. The one class support vector machine using Landsat vegetation indices and topographical variables produced an overall accuracy of 82.35% and a kappa value of 0.73. The one class support

vector machine produced the highest overall accuracy compared to the multiclass SVM and stochastic gradient boosting algorithm. The use of topographical variables further improved the accuracies compared to the combination of Landsat spectra with vegetation indices.

Keywords: Drought, remote sensing, vegetation indices, Multiclass support vector machine, One class support vector machine.

Acknowledgements

I would like to extend my gratitude to the following for their contributions:

Firstly, I would like to thank my supervisors Dr Romano Lottering and Dr Kabir Peerbhay for their feedback, guidance, support, patience and advice throughout the process. I have acquired so much knowledge from our interactions and your contribution is greatly valued.

The National Research Foundation for funding this project.

The South African Weather Services for providing climatic data.

A special thank you to Sappi forests for access to field sites and existing datasets required for the completion of my study

My family for their words of encouragement throughout this journey.

Declaration

This research was undertaken in fulfilment of the requirements within the School of Agriculture, Earth and Environmental Sciences and represents original work of the author. Work or ideas taken from other sources have been duly acknowledged in a form of citation and reference list.


Signature (Student)
Andile Lubanyana

..... 25/08/2022
Date

.....
Signature (Supervisor)
Dr. Romano Lottering

.....
Date

.....
Signature (Co-Supervisor)
Dr. Kabir Peerbhay

.....
Date

Plagiarism Declaration

I, Andile Lubanyana declare that;

1. The research reported in this thesis, except where otherwise indicated is my original work.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - a. Their words have been re-written, and the general information attributed to them has been referenced.
 - b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and reference sections.

Signed:  ..

Declaration for Publications

Details of contribution to publications that form part of and/or include research presented in this thesis (includes publications in progress, submitted and published and give details of the contributions of each author of each publication).

Publication 1: A Lubanyana, K Peerbhay, R Lottering (in preparation), Detecting drought stress on commercial Eucalyptus varieties using satellite remote sensing and stochastic gradient boosting.

Publication 2: A Lubanyana, K Peerbhay, R Lottering (in preparation), Mapping of drought stress in commercial forest plantations using supervised learning approaches and remotely sensed techniques.

The work was done by the first author under the guidance and supervision of the second and third authors.

Signature: Student

Andile Lubanyana

Signature: Supervisor

Dr. Romano Lottering

Date

Date

Table of Contents

Abstract	i
Acknowledgements	iii
Declaration	iv
Plagiarism Declaration	v
Declaration for Publications	vi
List of Tables	xi
List of Figures	xi
1 General Introduction	13
1.1 Introduction	13
1.2 Research Problem.....	17
1.3 Aim and Objectives	17
1.4 Key research questions:	17
1.5 General Structure of the Thesis.....	18
2 Chapter Two: Detecting drought stress in Commercial <i>Eucalyptus</i> plantations using satellite remote sensing and the stochastic gradient boosting algorithm	19
Abstract.....	19
2.1 Introduction	21
2.2 Materials and Methods	24

2.2.1	Study area.....	24
2.2.2	Field data collection.....	26
2.2.3	Landsat-8 OLI image acquisition and pre-processing.....	26
2.2.4	Climatic data.....	28
2.2.5	Statistical analysis.....	28
2.3	Results.....	29
2.3.1	High flats rainfall.....	29
2.3.2	High flats temperature.....	30
2.3.3	Mapping drought-affected compartments using Landsat 8 spectral variables.....	31
2.3.4	Mapping drought-affected compartments using Landsat 8 derived vegetation indices	32
2.3.5	Vegetation indices and trends between drought affected and healthy forest stands	33
2.3.6	Variable importance.....	35
2.3.7	Mapping drought distribution.....	35
2.4	Discussion.....	37
2.4.1	Classification using Landsat 8 and stochastic gradient boosting.....	37
2.4.2	Variable importance and drought classification.....	38
2.4.3	Recommendations for future research.....	39
2.4.4	Conclusion.....	40
3	Chapter Three: Mapping of Drought stress in commercial forest plantations using one-class and multi-class Supervised Learning Approaches and remotely sensed techniques.....	41
3.1	Introduction.....	42
3.2	Methodology.....	45
3.2.1	Study Area.....	45

3.2.2	Field verification data	47
3.2.3	Image Acquisition and Pre-processing	47
3.4	Environmental variables.....	48
3.4.1	Topographic metrics	48
3.4.2	Bioclimatic data	48
3.4.3	Statistical analysis.....	50
3.5	Accuracy assessment.....	51
3.5.1	Confusion matrix	51
3.5.2	Overall accuracy	52
3.6	Results	53
3.6.1	Data plots using multiclass SVM using Land Sat vegetation indices and topographic variables	53
3.6.2	Data plots using one-class SVM using Landsat vegetation indices and topographical variables	54
3.6.3	Classifying drought affected compartments using Landsat vegetation indices and topographical variables using a multiclass SVM.....	55
3.6.4	Mapping drought affected compartments using Landsat vegetation indices and topographical variables using a one class SVM	56
3.6.5	Variable importance.....	56
3.6.6	Drought distribution.....	57
3.7	Discussion	59
3.7.1	Multiclass classification using Landsat vegetation indices and topographical variables	59
3.7.2	One class classification using Landsat vegetation indices and topographical variables	60
3.8	Conclusion.....	61

4	CHAPTER FOUR: SUMMARY OF STUDY FINDINGS	62
4.1	Mapping drought affected forest compartments using multispectral data and Stochastic gradient boosting.....	62
4.2	Mapping drought affected forest compartments using multispectral data combined with topographical variables using a one class classification approach	63
4.3	Implications of this study	64
4.4	Future research recommendations.....	64
5	Conclusion.....	65
7	References	67

List of Tables

Table 2.1 Vegetation indices used in this study.....	27
Table 2.2 Showing Overall accuracy, Kappa and Error rate.	31
Table 2.3 Confusion matrix using stochastic gradient boosting with Landsat 8 spectra for mapping drought affected forest compartments using the best monthly results obtained in August 2015.	31
Table 2.4 Showing Overall accuracy, Kappa and Error rate.	32
Table 2.5 Stochastic gradient boosting using Landsat 8 spectra and drought indices for detecting healthy and drought affected forest compartments.	32
Table 3.1 Topo-graphic variables used to assess the drought impact on commercial forestry.	49
Table 3.2 Confusion matrix with two classes	52
Table 3.3 Multiclass support vector machine with Landsat 8 vegetation indices.	55
Table 3.4 One class support vector machine with Landsat 8 vegetation indices.....	56

List of Figures

Figure 2.1 Study area showing drought affected and unaffected forest compartments in the Sappi high-flats plantations. Where (A) shows South Africa, (B) KwaZulu-Natal and (C) the study area.	25
Figure 2.2 High flats monthly total rainfall from 2013-2017 based on weather station data received from South African weather service.	30
Figure 2.3 High flats monthly average temperature 2013-2017 based on weather station data received from SAWS.	30
Figure 2.4 Vegetation indices highlighting the trends between drought affected and unaffected compartments based on monthly data for August from 2013 to 2017 using Landsat 8 datasets..	34
Figure 2.5 Variable importance showing most effective variables used in final prediction of the model.....	35
Figure 2.6 illustration of classification maps based on Landsat 8 spectra and Landsat 8 derived vegetation indices	36

Figure 3.1 Study area showing drought affected and unaffected forest compartments in the Sappi High-flats plantation. Where (A) shows South Africa, (B) KwaZulu-Natal and (C) the study area.	46
Figure 3.2 Multiclass data plot showing drought affected, unaffected and other classes generated using R software.	53
Figure 3.3 One class data plot showing positive class drought and negative class other generated using R software.	54
Figure 3.4 Variable importance showing contribution of each variable using the One class and multiclass support vector machine.	57
Figure 3.5 illustration of derived classification maps with one class and multiclass support vector machines.	58

1 General Introduction

1.1 Introduction

Eucalyptus is one of the most commonly grown hardwood genus worldwide, covering more than 19 million hectares of land (Albaugh *et al.*, 2013). These species can be grown under varying climatic environments for products that include, among others; pulp and paper, charcoal, fuelwood, and solid wood products. In South Africa, forest plantations cover 1.2% of the countries land area of which 515 000 hectares are planted to *Eucalyptus* (Albaugh *et al.*, 2013; García *et al.*, 2017). These plantations are grown in a wide range of environments particularly those that are situated in the subtropical and humid warmer temperate regions. Commercial plantations are threatened by adverse weather conditions, particularly in developing countries, which are most vulnerable to drought events. During drought, there is generally a significant decline in rainfall over more extended periods of time. About 5% of Africa's total cultivated land is said to be under irrigation therefore, making the continent extremely vulnerable to drought (Belal *et al.*, 2014). Severe drought events negatively impact agriculture, especially in large geographical areas in Africa where people rely on rainfed agriculture (Masih *et al.*, 2014). During the last decade, the frequency and impact of adverse weather in the agricultural community in southern Africa have increased and the most common type of adverse weather is drought. Hydrological, meteorological, agricultural and socioeconomic droughts are the four drought categories characterised in literature. Hydrological drought occurs when there is a deficiency of water in the hydrological system, this is usually experienced when there are extremely reduced stream flows in rivers and reduced levels in lakes, reservoirs, and ground water. This type of drought has a huge impact on agricultural productivity and the economy of many countries in the world (Van loon, 2015). Meteorological droughts occur when there is a lack of rainfall over an extended temporal period and often exacerbated by other meteorological surroundings such as increased temperatures, increased evapotranspiration rates and dehydrating winds (Spinoni *et al.*, 2020). Furthermore, meteorological droughts influence hydrological droughts as they significantly reduce the availability of water. This affects the quality of water in rivers as the assimilative or dilution capacity is significantly reduced (Wolff and Van vliet, 2021). This may result in domestic and industrial waters being rationed to manage the limited supply of water (Mosley, 2015).

Agricultural droughts, on the other hand, are caused by a shortage of rainfall and inadequate soil water supply, this results in crop failure, pasture and economic loss and may expand the incidence of pests and disease that affect crops, livestock and forage (Cao *et al.*, 2019; Meza *et al.*, 2021). Lastly, socioeconomic droughts refer to a situation where water supply from a regional water source system cannot meet water demand (Liu *et al.*, 2020). However, this study will focus on agricultural drought and its impact on valuable commercial forest timber resources.

The recent 2015-2016 drought in South Africa was attributed to the El Niño southern oscillation (ENSO) (Monyela *et al.*, 2017). This was regarded as the worst drought in 23 years, affecting the country's water reserves, which resulted in certain parts of the country being declared as drought disaster areas (Schreiner *et al.*, 2018). A study by Ngaka, (2012) indicated that drought is a major disaster in South Africa in terms of its impact on people and total economic loss. This was also evident during the recent drought as it resulted in South Africa moving from exporting approximately 1 million tons of agricultural products to neighbouring countries to becoming a net importer of crops (Baudoin *et al.*, 2017). Furthermore, this also had serious socioeconomic impacts as it left thousands of people unemployed. One particular sector threatened by drought and its subsequent impact is forestry, as it is heavily reliant on water resources both, above and below ground (Hais *et al.*, 2019; Idris and Mahrup, 2017).

The South African commercial forest industry is highly sensitive to adverse weather as about 1.5% of the country's land is suitable for tree crops under the current climate (García *et al.*, 2017). Forest plantations are concentrated in areas that receive higher rainfall and contribute a large share of total streamflow (García *et al.*, 2017). South Africa is regarded as a semi-arid country with an average yearly precipitation of approximately 460 mm (Dallas *et al.*, 2014). With the country's limited water resources forest plantations are under threat, as drought significantly affects the production of forest plantations. For example, Xulu *et al.*, (2018) found that the recent 2015-2016 drought resulted in the decline of eucalyptus productivity in the east-coast of the South African Zululand forestry region. Furthermore, the study suggested that some commercial forest plantations in this region were projected to be affected by frequent and severe droughts, bringing new challenges to the commercial plantation production. Considering the recent drought events, understanding and monitoring of ENSO related droughts is of great importance as it informs decision making for adapting to drought hazards. In this regard, remote sensing is a useful tool for studying temporal evolution and drivers of drought due to limited access and

inconsistency of drought related in-situ data (Lausch *et al.*, 2018). Conventional methods of assessing and monitoring drought events often depend on readily available rainfall data, which is limited in most parts of the African continent and may also be, inaccurate and challenging to obtain in near real-time (Aghakouchak *et al.*, 2015). In contrast, remote sensing satellite sensors are always accessible and can be utilized for detecting drought onset episodes, its magnitude and duration (Dash *et al.*, 2017). For instance, Nasilowska *et al.*, (2019) used vegetation indices computed from Landsat 8 OLI imagery (0.43-1.38) covering the 2014 and 2015 growing period. The study attempted to establish the most important factors that affect drought resistance by testing the biophysical and physical parameters of trees, including forest habitat characteristics. The study found that mid-infrared based indices such as the Normalized difference moisture index (NDMI) and Moisture stress index were useful for monitoring water shortage in forests. It also found that ground water and rainfall shortfalls affect forest conditions differently and is dependent on the type of soil. Furthermore, this study utilized algorithms such as index differences, PCA analysis and ANOVA statistical analysis. The index difference approach was found to be simple and functions effectively without extensive in situ field data. Similarly, Rousta *et al.*, (2020) utilized MODIS derived Normalized Difference Vegetation Index (NDVI), Tropical Rainfall Measuring Mission (TRMM) data, Vegetation Condition Index (VCI) and Land Surface Temperature (LST) indices with a 1km resolution to determine the impacts of drought on vegetation from 2001 to 2018 in Afghanistan. Based on VCI data, in 2009 and 2010 the area experienced 28% and 21% drought, respectively. In addition, it also found a relatively high correlation between NDVI and VCI, but lower between NDVI and precipitation. Furthermore, LST played an important role in inducing the NDVI values, therefore, LST and precipitation were recommended for effectively capturing the correlation between drought and NDVI. The study successfully used NDVI, VCI, LST and rainfall indices to investigate the impact of drought on vegetation and showed that these indicators are a reliable method of drought monitoring. Lastly, Richman *et al.*, (2016) examined the roles of temperature, precipitation and El Niño in characterizing both current and previous drought events. The logistic regression and primal estimated sub gradient solver were utilized to determine drought. The logistic regression produced an overall accuracy of 81.14%. The use of remote sensing is imperative in regions that lack continuous in-situ data monitoring however, there have been further developments in remote sensing technology involving the utilization of algorithms and remote sensing data to map

drought prone areas. To predict drought, Kuswanto and Naufal (2019) derived a 3-month standardized precipitation index (SPI) from TRMM and the Modern-ERA Retrospective Analysis for Research and Applications (MERRA-2). The classification methods used were Classification and Regression Trees (CART) and Random Forest (RF). However, using multiple predictor analyses, several predictors showed CART to be less predictive than RF. The accuracy of the prediction using RF was 100% with an average Area Under the Curve (AUC) of about 0.8 whereas the AUC for CART was 0.75 with an accuracy of over 80%. The analysis showed that using MERRA-2 dataset predicted by RF is effective in predicting drought. Furthermore, Das *et al.*, (2021) used Landsat OLI/TIRS images (10.6 - 11.19) to measure LST and soil moisture index (SMI) to examine drought conditions in tea plantations. Sentinel-2 satellite images were used to develop maps for NDMI, NDVI, Leaf Area Index (LAI), as well as yield maps. A Pearson's correlation and simple linear regression analyses were run to determine the relationship between soil moisture and vegetation canopy moisture as well as to illustrate the relationship of LAI and time-series information for predicting yield during a drought season. Based on this study, the frequency for the Sylhet station was 38.46% during near-normal conditions, 35.90% during normal conditions, and 25.64 % during moderately dry conditions. Similarly, the Sreemangal station illustrated frequencies of 28.21%, 41.02%, and 30.77% for near-normal, normal, and moderately dry months, respectively. For the drought periods 2018-2019, 2019-2020 and 2020-2021, correlation coefficients between SMI and NDMI were 0.84, 0.77 and 0.79, respectively. This illustrated a robust association between soil and plant canopy moisture. In this study, satellite remote sensing with SPI was shown to be valuable for land use planners, policy makers, and scientists to determine drought stress in tea estates. The use of remote sensing data is very useful in developing countries for planning and undertaking drought risk assessments. This can be used to drive positive change through formulating drought policies that focus on adapting to drought rather than on providing relief to the people affected. The commercial forestry sector can also benefit from this through early detection and to critically assess the impacts before more profound structural changes occur.

1.2 Research Problem

The recent 2015-2016 drought affected many parts of the country leaving many people devastated and the agricultural sector taking a huge knock due to declining agricultural yields. The commercial forestry sector was affected by this phenomenon resulting in reduced *Eucalyptus* growth and productivity. The High flats region in KwaZulu-Natal is predominantly an agricultural area with commercially grown *Eucalyptus dunnii* plantations. Certain forest stands within the plantation were found to be affected by drought with visible tree diebacks. The impacts of the 2015-2016 drought on forests have not been studied in detail and are uncertain. While a few studies by (Xulu *et al.*,2018; Xulu *et al.*,2019) have looked at drought influence on commercial plantations using MODIS. None of the studies conducted in the country have studied the impacts of drought using multiple variables such as Landsat 8 spectra, vegetation indices and climatic data. This has prompted a study to be conducted in the area to better understand the impacts of drought by mapping drought prone parts of the plantation. This has enabled remote sensing technology to be used to allow for supervised and unsupervised classification of drought.

1.3 Aim and Objectives

The study aims to evaluate the utility of a cost-effective Landsat 8 (0.43 – 1.38 μ m) imagery in mapping the spatial extent of drought prone *Eucalyptus dunnii* plantations and to contribute in improving techniques for drought analysis. This will be achieved through the subsequent objectives:

1. Comparing the utility of Landsat spectra with vegetation indices to classify drought affected plantations using the Stochastic gradient boosting.
2. To assess the trend of vegetation health, rainfall and temperature over the duration of this study
3. Compare the utility of topographical variables and vegetation indices as input data sources for one class and multiclass support vector machines analysis.

1.4 Key research questions:

1. Is the use of spectral bands only sufficient as input data for stochastic gradient boosting to classify drought in forest plantations?
2. Will the integration of spectral bands and drought indices improve the overall accuracy of drought classification using the Stochastic gradient boosting?

3. Will the integration of environmental variables, spectral bands and vegetation indices further improve the overall accuracy of drought classification using the one class support vector machine algorithm?

1.5 General Structure of the Thesis

This thesis is made up of four chapters. The first chapter is made up of the general introduction which includes the aim and objectives. The second and third chapters consist of two publishable standalone research papers and Chapter four comprises of a brief summary of the study, conclusion and a review of the aim and objectives. Below is a brief summary of both chapters two and three.

Chapter 2 is the first publishable paper where drought stress on commercial *Eucalyptus* was detected using satellite remote sensing and stochastic gradient boosting. This includes a trend analysis of temperature, rainfall as well as the classification of drought affected compartments using remote sensing algorithms.

Chapter 3 is the second publishable paper and focuses on improving drought classification using local environmental variables and the one class support vector machine algorithm. Landsat 8 spectral information and drought indices were also used as input data sources.

2 Chapter Two: Detecting drought stress in Commercial *Eucalyptus* plantations using satellite remote sensing and the stochastic gradient boosting algorithm

Abstract

Eucalyptus species are one of the most significant forest species that are commercially grown worldwide, covering area of about 20 million ha. This study compared the utility of Landsat spectra with a combination of vegetation indices to detect drought affected plantations using the Stochastic gradient boosting algorithm. Using Landsat spectra only produced an overall accuracy of 74.70% and a kappa statistic of 0.59. The combination of Landsat spectra with vegetation indices produced an overall accuracy of 83.13% and a kappa statistic of 0.59. The study then conducted a trend analysis of vegetation health using key indices over the duration of this study. NDVI values fluctuated over the years where 2013 had the highest NDVI value of 0.7 and 2015 the lowest NDVI value of 0.55. The normalized difference water index (NDWI) in 2013 had the highest value of 0.68 and NDWI in 2015 had the lowest value which was 0.50. Most indices showed a similar trend where 2013 had the highest index value and 2015 the lowest. Finally, a trend analysis of rainfall and temperature during drought was undertaken. The years 2013 to 2017 monthly rainfall trend analysis indicated that February 2017 received the highest rainfall of 154 mm. July of 2016 received the highest rainfall of 55 mm compared to 2013 (6.4 mm), 2014 (0.6 mm), 2015 (28 mm) and 2017 (1 mm). It was also observed that in June 2017 no rainfall was recorded. The monthly temperature trend showed that December 2015 had the highest temperature of 28°C compared to December 2013 (24°C), 2014 (25°C), 2016 (27°C) and 2017 (24°C). June 2017 had the highest temperature of 23°C while June 2015 had the lowest at 20°C. The research undertaken showed that it is possible to classify drought prone plantations using a blend of remote vegetation indices and Landsat spectral bands. The year 2015 was deemed a drought year and this was attributed to below average rainfall and high temperatures for a long period of time. Using remote sensing tools and the stochastic gradient boosting algorithm, early mapping at the onset of drought on forest resources may allow for effective mitigation and protection management strategies.

Keywords: Forest drought, Remote sensing, Stochastic gradient boosting, Vegetation indices, Landsat 8.

2.1 Introduction

Eucalyptus species are one of the most vital forest species that are commercially grown worldwide, covering a global surface of roughly 20 million ha (Barradas *et al.*, 2018). In South Africa, commercial forests cover about 1.27 million ha of which 515 000 ha are planted to *Eucalyptus* (Albaugh *et al.*, 2013; DAFF ,2017). The forest sector contributes approximately 1% to the gross domestic product (GDP) and employs around 165 900 workers and provides about 62 700 jobs directly and 30 000 jobs indirectly (DAFF, 2017). Furthermore, the sector provides biomass for fuel and raw materials for pulp and paper manufacturing, timber mining, pole manufacturing, saw milling and composite board manufacturing (du Toit *et al.*, 2017; Foli *et al.*, 2014). Although forests are commercially exploited, they are vital sources of livelihoods to millions of people particularly those residing in rural areas (Köhl *et al.*, 2015). The majority of commercial forests in South Africa are located adjacent to rural areas where rural communities benefit through employment opportunities, access to grazing and harvesting of non-woody products (Bojang and Ndeso-Atanga, 2011). However, the forestry sector is identified as extremely vulnerable to climatic events, since they directly affect the productivity of forests thereby threatening the livelihoods of its dependents as well as stakeholder profits (Bottero *et al.*, 2017; Law, 2014). Among the threatening natural impacts experienced in South African forestry such as wind, fire, hail and snow damage lies a much more lethal and widespread threat to forest productivity in general. Drought stress on forests, caused by prolonged periods of no rainfall has become a serious concern within the industry due to irregularity in occurrences and intensity of damage to forest resources. More specifically, the impact of drought to forest trees requires immediate response mechanisms to detect, diagnose and quantify affected regions for effective management solutions. However, determining such impacts using conventional approaches may prove challenging. Frequent monitoring of drought across all scales is important for the forest industry as it is used to influence decision making and reduce the risk of damage to forests (Xulu *et al.*, 2018). Such practices may also improve our understanding of what triggers drought, so that useful forest management strategies can be continuously developed.

During severe drought conditions, forest trees experience water stress (Hope *et al.*, 2014). This poses a serious threat to the functioning, structure, and ecological value of forest ecosystems (Archaux and Wolters, 2006). However, forest trees react to water stress through the closure of stomata and a reduction in LAI to prevent further water loss (Assal *et al.*, 2016). A decline in LAI

results in a reduction in photosynthetic ability and alters canopy structure (Zhang *et al.*, 2017). This further leads to a reduced chlorophyll and water content in plant leaves (Clark *et al.*, 2016). The closure of the stomata minimizes the ability of forest trees to take up carbon and this could result in tree mortality and forest dieback (Pasho *et al.*, 2011). Since severe drought reduces forest productivity and tree growth and often weakens trees by decreasing their vigor, the impact of drought on forests may last for several years, which incurs high economic costs to the forest industry (Peltier *et al.*, 2016). Droughts are complex natural hazards, and this is partly because their impacts spread over large geographical areas (Chopra, 2006). Although conventional in situ measurements have been used to monitor drought damage, they may not cover large areas as they require intense field observations which is labour intensive. However, remote sensing techniques provide high density data and improve in situ derived data.

Remote sensing tools are very useful in detecting drought stress in forests, which is challenging to observe using conventional methods (Xie *et al.*, 2008). The two most common multispectral sensors that are used in drought studies are MODIS and Landsat satellites (Assal *et al.*, 2016; Bastos *et al.*, 2014; Hope *et al.*, 2014; Huang and Anderegg, 2012). These remote sensing platforms are both suited for drought studies as they cover large geographical areas where access to ground measurements may be impossible, data is repeatable, cost effective and are reliable (Byer and Jin, 2017). In remote sensing studies, canopy water content can be used as an indicator of drought impact on forests (Ustin *et al.*, 2012). For example, Asner *et al.*, (2016) used a High-fidelity imaging spectroscopy (HIFIS) (980 – 1160 nm) to estimate forest canopy water content during the 2011-2015 drought event in California. The HIFIS data was combined with 3D forest images derived from Light detection and ranging (LIDAR) to exclude non-forest canopy measurements and to allow for estimation of canopy water content. This data infusion method called laser-guided high-fidelity imaging spectroscopy permits establishing the quantity and forecast of forest canopy water content in three dimensions. A statewide supervised machine learning algorithm was used to scale up the aircraft-based canopy water content imagery to the full state level. Both the HIFIS and LIDAR data were acquired utilizing the Carnegie airborne observatory (CAO). The study found that during the drought about 10.6 Ha of forest experienced a significant decline in canopy water content. It was also found that a 30% canopy water loss occurred over 1 million hectares of forest affecting approximately 58 million trees. Although both MODIS and Landsat satellites were not utilized in this study, it is important to point out that using

CAO sensors can be costly as opposed to remote sensing platforms that have long-term data and, are open access. Therefore, this method may not be applicable in southern Africa due to cost constraints. Other methods used to monitor drought impact on forests are the application of vegetation indices (Breshears *et al.*, 2005).

Vegetation indices (VI) are calculated from simple formulas that include two or more reflective wavebands derived from satellite data (Sonobe *et al.*, 2017). VI's are widely used to assess differences in the physiological state and biophysical properties of vegetation (Maselli, 2004; Mohd Razali *et al.*, 2016; Sruthi and Aslam, 2015). Several remote sensing indices have been developed and used in drought studies (Dennison *et al.*, 2005; Dutta *et al.*, 2015; Amalo and Hidayat, 2017; Potter, 2015). For example, Potter, (2015) successfully used Landsat derived NDWI and NDVI to detect changes during the 2013-2014 drought along a 100km transect of the central California coast. Grasslands showed signs of more significant drought stress than forests or shrub lands. A study by Dennison *et al.*, (2005) compared the effectiveness of using MODIS NDVI and NDWI in monitoring fuel moisture content in shrub land ecosystems. These indices were both correlated with live moisture data obtained from the Los Angeles county fire department. Using linear regression models the relationship between indices and live fuel moisture were established. The study found that the NDWI was more significant in monitoring fuel moisture in shrub land ecosystems when compared to NDVI. NDWI could hypothetically be used to determine canopy water content in forests, as it is highly sensitive to canopy water content and spongy mesophyll of vegetation canopies (Ceccato *et al.*, 2001). Although both remote sensing platforms and vegetation indices have the potential to monitor drought in forests, a combination of a suitable sensor and vegetation indices as well as a good algorithm could help in classifying and mapping the spatial extent of drought-affected areas (Lewińska *et al.*, 2016). For instance, Barradas *et al.*, (2021) compared machine learning methods for classifying plant drought stress. Where a combination of artificial intelligence with reflective spectroscopy was tested and an entire dataset of reflective spectra was used as input to machine learning algorithms namely decision trees, random forest and extreme gradient boosting. Random forest, extreme gradient boosting, and decision trees produced accuracies of 94 %, 92 % and 88 %, respectively. These classification methods are a promising tool to detect plant physiological responses to drought. Dao *et al.*, (2021) used RF, support vector machine (SVM) and deep neural networks (DNN) to analyze full spectra

and first order derivative detecting drought impact at various stages for *Bromus inermis* grass. SVM produced the overall accuracy of 97.5% followed by DNN (94.6%) and RF (90.4%).

In summary, drought stress has been causing grave concerns towards the productivity of commercial forestry and the need to monitor drought using remote sensing techniques has the potential to detect drought stress. However, the cost is a major limiting factor and we need to exploit freely available sensors. Therefore, this study seeks to evaluate the utility of a cost-effective Landsat 8 image in mapping the spatial extent of drought-prone *Eucalyptus dunnii* plantations in KwaZulu-Natal, South Africa. Specifically, the study compares the utility of Landsat spectra only with vegetation indices and Landsat spectra to detect drought -affected plantations. The study also adopted a stochastic gradient boosting approach due its low sensitivity to outlier, capacity to utilize inaccurate training and unbalanced datasets and its robustness in dealing with interaction effects between variables (Li *et al.*, 2019).

2.2 Materials and Methods

2.2.1 Study area

Eucalyptus dunnii is one of the fastest-growing wood plants, growing at a mean rate of up to 30 m³/ha. About 20 million hectares have been planted worldwide due to their adaptability and fast growth rates (Delgado and Pukkala, 2011). This study area (Figure 2.1) is located in the High flats forest plantations in KwaZulu-Natal, which is situated along the eastern-half of the country on a crest of a hill with a very gentle slope. The site is managed by SAPPI forests and lies between 30° 16' 00" S; 30° 12' 00" E and has a total area of 10380 hectares. The area experiences cold, dry winters with average monthly temperatures ranging from 6.1°C in July to 25°C in January (Schulze *et al.*, 2007). The area generally falls within the summer rainfall region. The mean annual precipitation ranges between 800-1000mm with much rainfall occurring from October to March (Schulze *et al.*, 2007). The elevation of the area is 950 m above sea level and is underlain by mainly coarse-textured sandstone and tillite which give rise to different soil types (Crous *et al.*, 2013).

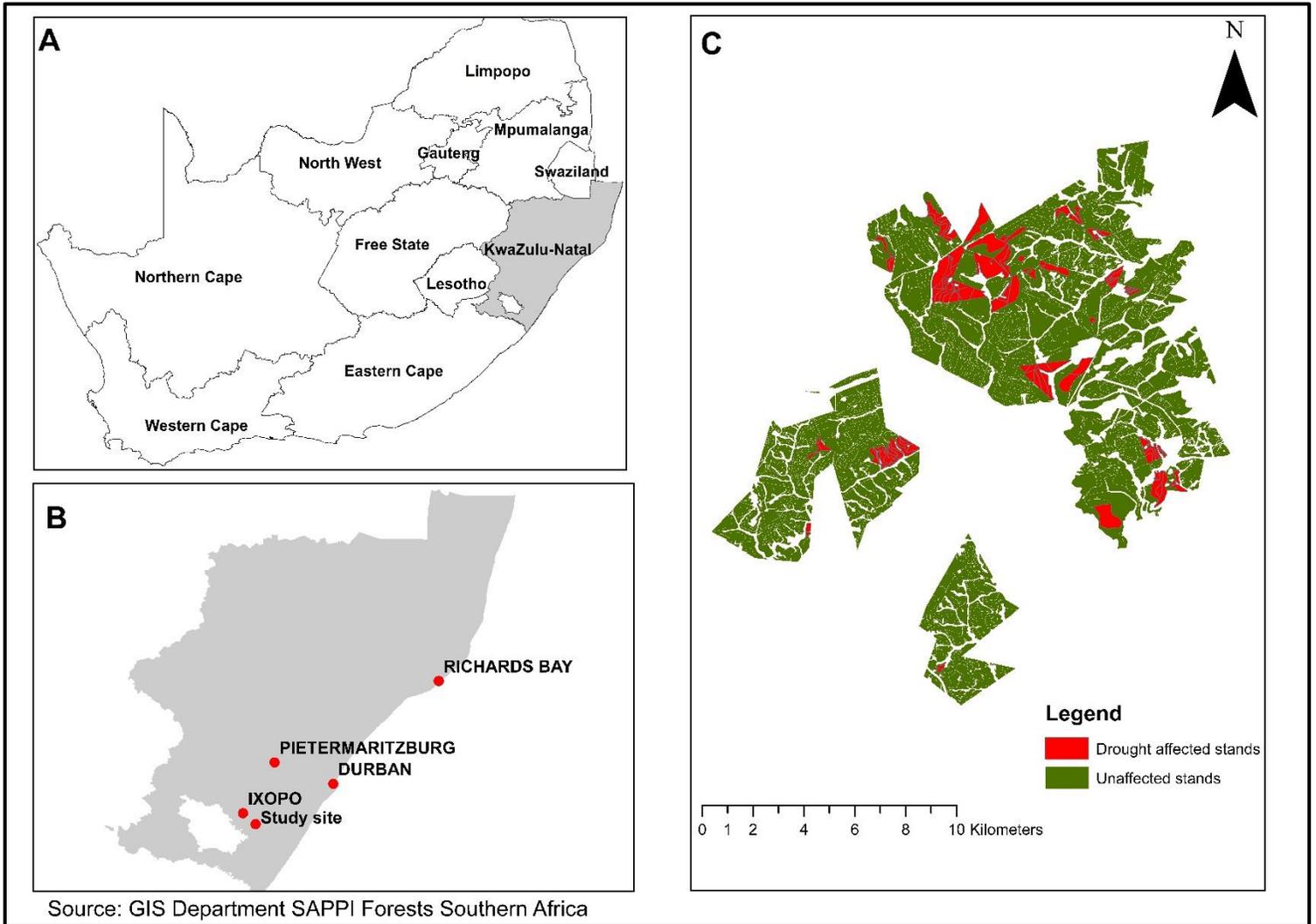


Figure 2.1 Study area showing drought affected and unaffected forest compartments in the Sappi high-flats plantations. Where (A) shows South Africa, (B) KwaZulu-Natal and (C) the study area.

2.2.2 Field data collection

Field data collection and verification was done monthly at the SAPPI (South African Pulp and Paper Industry) plantation in the High flats region of KwaZulu-Natal near Ixopo from the 8th of August 2015 to the 15th of December 2015. Forest compartments were physically observed for drought stress with data collected on the absence and presence of drought damaged trees and the number of trees damaged, which was expressed as the number of drought damaged trees over the total number of assessed trees. Compartments displaying over 50% drought induced stress were categorized as damaged stands, while compartments displaying 0% of drought damage were categorized as undamaged. Stands between 1% and 49% were removed from the analysis to avoid underestimating the occurrence of drought damage. Furthermore, trees that showed physical signs of stunted growth, leaf dropping and discoloration as well as embolism formation were categorized as drought damaged trees. GPS points were collected using a differentially corrected Trimble GeoXT handheld GPS receiver with an accuracy of < 1m. The GPS data were used for ground verification of each forest compartment, which was then later extracted for statistical analysis using each image dataset. This dataset was used to extract image spectra and vegetation indices for statistical and classification analysis as well as for deriving trend analysis for drought affected and unaffected stands. The final dataset contained 450 stands with over 50% drought damage, while the remaining 380 stands displayed no signs of drought stress. These compartment means were extracted to assess trends over time using satellite imagery and rainfall from a nearby plantation weather station.

2.2.3 Landsat-8 OLI image acquisition and pre-processing

Cloud free Landsat-8 OLI (Operation land imager) multispectral images with path/row:168/81 were acquired from August 2015 to December 2015 from the United States Geological Survey (USGS) Earth Resources Observation and Science (<https://earthexplorer.usgs.gov>) and covered the entire study area. The images were acquired during clear skies, sunny conditions with an average azimuth angle of 38.25° and sun elevation angle of 31.66°. The image dates were specifically chosen to match the period from when trees first began to show signs of stress due to drought occurrence in this region. Landsat-8 OLI has two modes of capturing images which are spectral and panchromatic. Landsat 8 OLI consists of seven spectral bands spread across the

visible (433 nm – 680 nm), near-infrared (845 nm - 885 nm) to short wave infrared (1560 nm – 2300 nm) regions with a spatial resolution of 30-m. The Landsat-8 OLI sensor has a temporal resolution of 16 days. Landsat-8 OLI image spectral bands were converted from digital numbers format (DN) to reflectance using ENVI 5.1 software following the approach described on the USGS website (<http://landsat.usgs.gov>). MODTRAN was used to correct the acquired images based on the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercube (FLAASH) radiative transfer algorithm. (Shao and Zhang, 2016; Dube and Mutanga, 2015). Monthly images succeeding December 2015 were unavailable due to heavy cloud cover of the region and was therefore not considered in this study. The Landsat-8 OLI spectral bands from each monthly image were used to calculate vegetation indices shown in Table 2.1. The vegetation indices were selected based on key studies that were successful in monitoring drought stress on vegetation using remotely sensed data (Zhang *et al.*, 2013; Ghaleb *et al.*, 2015, Amalo and Hidayat, 2017). Furthermore, image data from 2013 to 2017 for each year was collected monthly to analyze vegetation index trends along with rainfall data.

Table 2.1 Vegetation indices used in this study.

Vegetation indices	Formula	References
Normalized difference vegetation index (NDVI)	$NDVI = \frac{Nir - Red}{Nir + Red}$	(Rouse Jr <i>et al.</i> 1974)
Normalized difference water index (NDWI)	$NDWI = \frac{Nir - Swir}{Nir + Swir}$	(Gao 1996)
Enhanced vegetation index (EVI)	$EVI = \frac{Nir - Red}{(Nir + 6Red - 7.5Blue + 1)}$	(Huete <i>et al.</i> 2002)
Vegetation condition index (VCI)	$VCI = 100 \frac{Evi_i - Evi_{min}}{Evi_{max} - Evi_{min}}$	(Kogan 1995)
Temperature condition index (TCI)	$TCI = 100 \frac{LST_{max} - LST_i}{LST_{max} - LST_{min}}$	(Kogan 1995)
Vegetation health index (VHI)	$VHI = 0.5Vci + 0.5TCI$	(Kogan 1995)

2.2.4 Climatic data

Monthly rainfall data for High-flats was obtained from the South African Weather Services (SAWS). The data was collected from a nearby weather station (0210099A7) in Ixopo. Monthly temperature data was also obtained from SAWS, and a monthly temperature graph was produced.

2.2.5 Statistical analysis

2.2.5.1 Stochastic gradient boosting

The stochastic gradient boosting algorithm (SGB) is an innovative machine learning technique associated with "bagging" and "boosting" established to enhance overall classification accuracy of analysis (Freeman *et al.*, 2015). Classification trees are derived from boosting and bagging methods where boosting methods are founded on prior classification trees while bagging methods are founded on subsets of the training data (Dube *et al.*, 2015). The SGB algorithm follows a specific procedure where at every phase of the boosting method, a subset of data is randomly selected. Selected subsets are separated using the steepest gradient algorithm, where the gradient is defined by deviance (twice the binary negative log likelihood), which is a replacement for misclassification rates. This results in fairly small trees at each repetition, rather than generating complete classification trees at each phase. The last phase involves stacking all trees together, of which there are usually between 100 and 200, and classifying each observation based on the most prevalent classification among the trees. The SGB uses three user-defined parameters that direct its output which are the learning rate (r_1), the number of regression trees (N_{tree}) and tree complexity (C_t) (Chirici *et al.*, 2013). Two parameters are considered most significant, which are r_1 and c_t where r_1 governs the influence of each tree to the developing model, and C_t is the number of samples utilized in the last node. The SGB model was run in the R statistical software version 3.5.1 where r_1 was set between 0.0001 and 0.1 and c_t was set between 1 and 5. The bagging fraction was set to 0.3 in this study.

2.2.5.2 Variable importance

SGB variable importance uses variable selection to eliminate redundant predictor variables which don't improve the model's accuracy. The model scales the relative contribution of each variable and a high number indicates a robust impact of the response variable. In this study, relative variable importance was determined utilizing the average decrease in accuracy (Han *et al.*, 2016).

2.2.5.3 Accuracy assessment

The study utilized a confusion matrix which summarizes the performance of the stochastic gradient boosting method (Novaković *et al.*, 2017). The final dataset was split into 70% training data and 30% test data. The kappa coefficient was utilized to determine the performance of the SGB method. Where a kappa statistic was calculated and values closer to 1 indicate a perfect classification. The producer's accuracy is a measure of how well the real features on the ground are represented on the classified map. It also includes error of omission which is part of the observed features not represented on the map. High errors of omission results in low producer's accuracy. The user's accuracy is a measure of the reliability of the map and indicates to the user how well the map represents what's on the ground (Patil and Taillie, 2003). This is also inclusive of the error of commission where a high error of commission results in a lower user's accuracy (Patel and Kaushal, 2010).

2.3 Results

2.3.1 High flats rainfall

The monthly total rainfall from the year 2013 to 2017 was seasonally variable and varied across the five years. The high summer rainfall months were generally around October to March while the lowest rainfall months were around April to September (Figure 2.2). Over the period of 2013 to 2017, February 2017 received the highest rainfall of 154 mm compared to February 2013, 2014, 2015, 2016 with rainfalls of 111 mm, 43 mm, 136 mm and 118 mm, respectively. July of 2016 received the highest monthly rainfall of 55 mm compared to 2013, 2014, 2015 and 2017 with rainfalls of 6.4 mm, 0.6 mm, 28 mm and 1 mm. It is also observed that in June 2017 no rainfall was recorded.

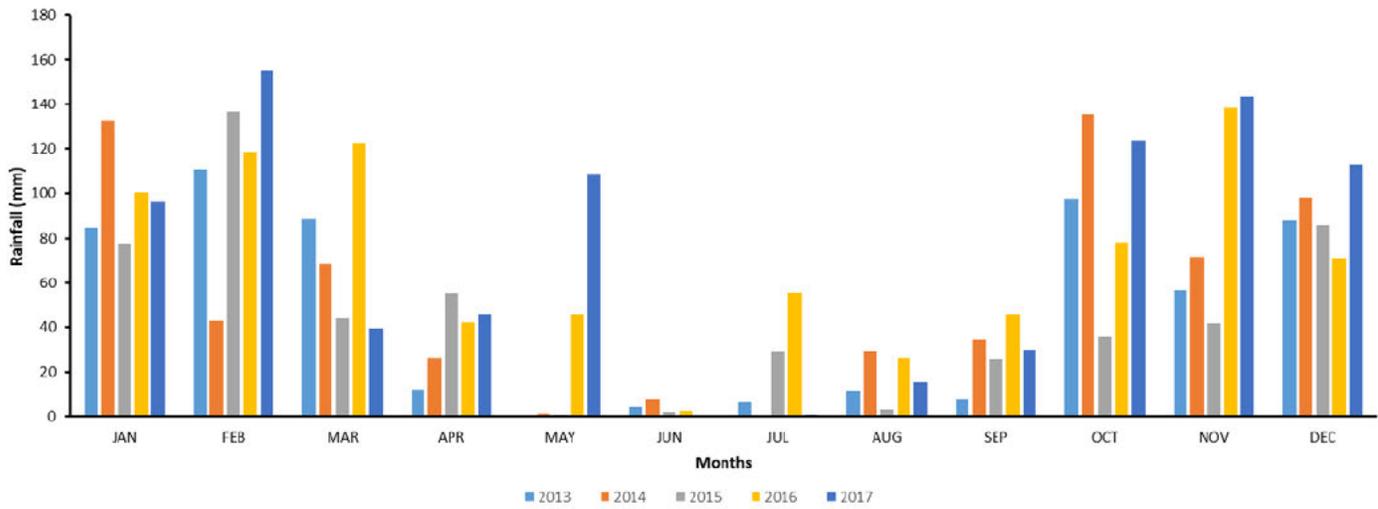


Figure 2.2 High flats monthly total rainfall from 2013-2017 based on weather station data received from South African weather service.

2.3.2 High flats temperature

Figure 2.3 Shows monthly average temperatures from 2013 to 2017. December 2015 had the highest temperature of 28°C compared to December 2013, 2014, 2016 and 2017 with temperatures of 24°C, 25°C, 27°C and 24°C, respectively. June 2017 had the highest temperature of 23 °C while July 2016 had the lowest at 19.8°C. The monthly temperature graph clearly shows that 2015 was one of the hottest years.

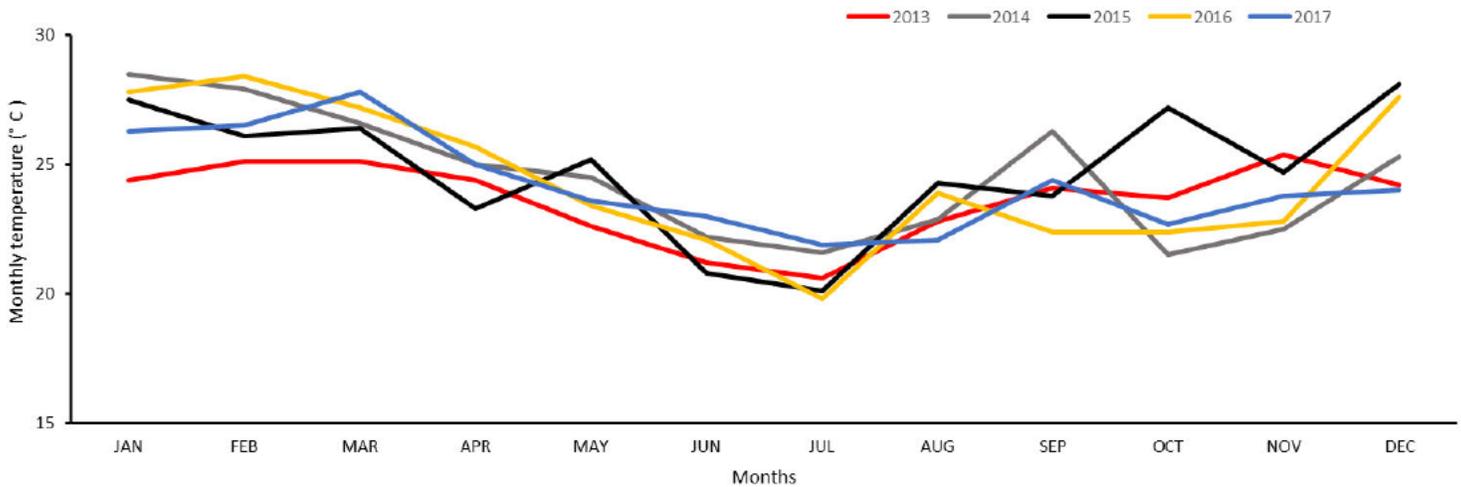


Figure 2.3 High flats monthly average temperature 2013-2017 based on weather station data received from SAWS.

2.3.3 Mapping drought-affected compartments using Landsat 8 spectral variables

Table 2.2 displays the accuracy results when using individual Landsat 8 bands to detect drought affected compartments between August to December 2015. From the table it is noticeable that the highest accuracy was obtained in the month of August.

Table 2.2 Showing Overall accuracy, Kappa and Error rate.

Dataset	Overall accuracy (%)	Kappa (K)	Error rate (%)
August 2015	74.70	0.59	25.3
September 2015	74.30	0.59	25.7
October 2015	70.44	0.58	29.56
November 2015	69.78	0.58	30.22
December 2015	64.45	0.57	35.55

Using the best dataset, the test results for August, indicated that using Landsat 8 spectral bands and the SGB algorithm produced high producer's and user's accuracies ranging from 76% to 71% for healthy trees and 73% to 79% for drought-affected trees. The results show that 290 healthy compartments (non-drought affected) were correctly classified as healthy, and only 90 compartments were incorrectly classified as drought affected. In the drought-affected column 120 compartments were correctly classified as drought-affected, while 330 compartments were incorrectly classified as healthy (Table 2.3).

Table 2.3 Confusion matrix using stochastic gradient boosting with Landsat 8 spectra for mapping drought affected forest compartments using the best monthly results obtained in August 2015.

	Healthy Stands	Drought Affected Stands	Total
Healthy	290	120	410
Drought Affected	90	330	330
Total	380	450	830

2.3.4 Mapping drought-affected compartments using Landsat 8 derived vegetation indices

Table 2.4 displays the accuracy results when using individual Landsat 8 bands combined with the derived vegetation drought indices to detect drought affected compartments between August to December 2015. August produced the best overall accuracy and kappa statistic compared to the rest of the months under investigation.

Table 2.4 Showing Overall accuracy, Kappa and Error rate.

Dataset	Overall accuracy (%)	Kappa (K)	Error rate (%)
August 2015	83.13	0.76	16.87
September 2015	82.14	0.75	17.86
October 2015	79.87	0.73	20.13
November 2015	77.90	0.72	22.1
December 2015	76.56	0.72	23.44

Using the best results of August, test dataset results indicated that using an integration of vegetation indices with Landsat 8 spectra produced high producer's and user's accuracies of 82% for healthy trees and 84 % for drought-affected trees. The confusion matrix shows that 310 healthy (non-drought affected) forest compartments were correctly classified and 70 compartments were incorrectly classified as being drought affected. In the drought affected column, 70 compartments were correctly classified as drought-affected while 380 compartments were incorrectly classified as healthy (Table 2.5).

Table 2.5 Stochastic gradient boosting using Landsat 8 spectra and drought indices for detecting healthy and drought affected forest compartments.

	Healthy Stands	Drought Affected Stands	Total
Healthy stands	310	70	380
Drought affected stands	70	380	450
Total	380	450	830

2.3.5 Vegetation indices and trends between drought affected and healthy forest stands

For comparison purposes, Figure 2.4 shows a five-year vegetation index trend from 2013 to 2017 between healthy and drought affected forest compartments during the month of August. NDVI values steadily declined from 2013 which had the highest NDVI value of 0.7 to the lowest in 2015 with an NDVI value of 0.55. The NDVI value in 2016 slightly increased and then later dropped in 2017, by 0.03. The Normalized difference water index (NDWI) in 2013 had the highest value of 0.68 while in 2015 had the lowest value, which was 0.50. It was observed that most of the indices showed a similar trend where 2013 had the highest index value and 2015 the lowest with a slight distinction between healthy and drought affected stands. Nonetheless, indices such as the vegetation health index (VHI), VCI and the enhanced vegetation index (EVI) displayed this trend with greater distinction between drought affected and those healthy forest compartments.

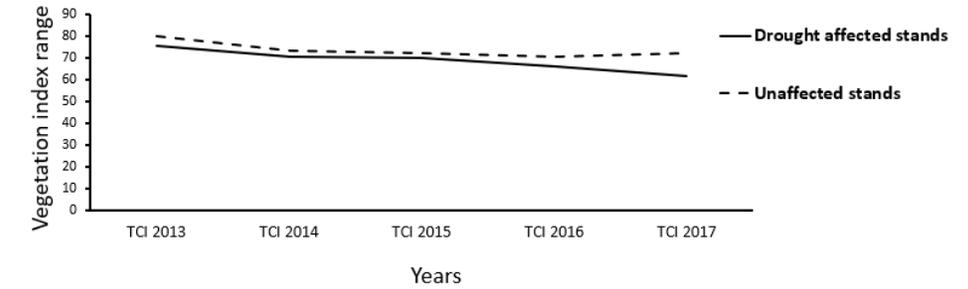
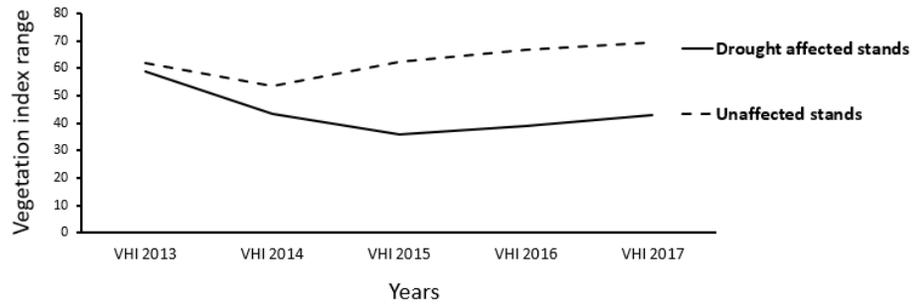
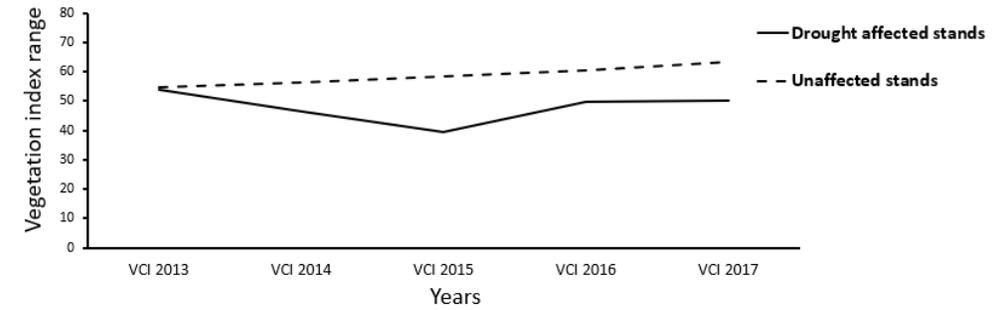
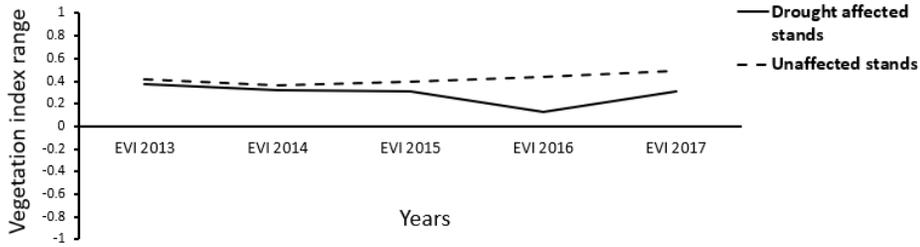
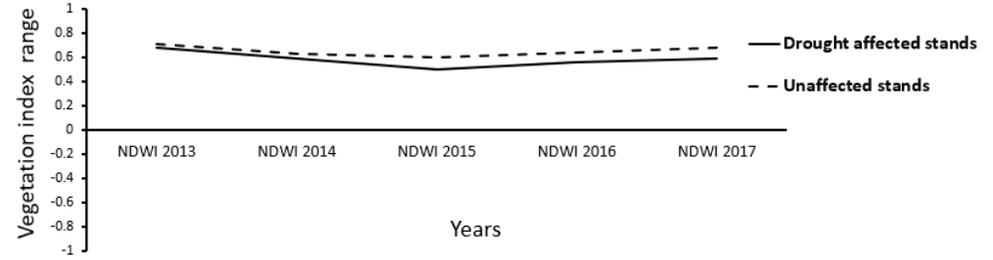
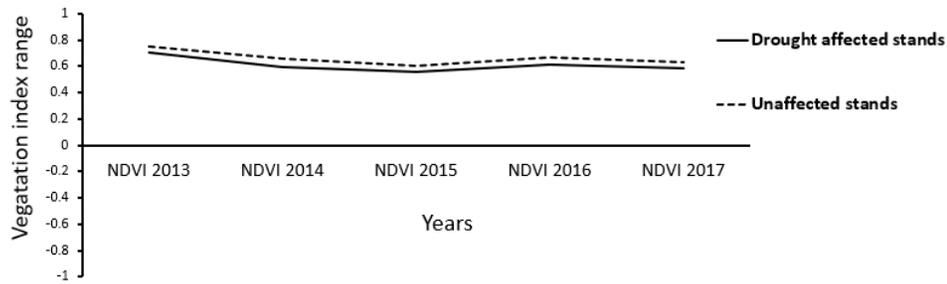


Figure 2.4 Vegetation indices highlighting the trends between drought affected and unaffected compartments based on monthly data for August from 2013 to 2017 using Landsat 8 datasets.

2.3.6 Variable importance

Figure 2.5 shows the most important variables utilized in the final prediction model using Landsat 8 bands combined with derived vegetation indices. The results are shown for the best month of August 2015 where variable importance was measured using the mean decrease in accuracy (MDA).

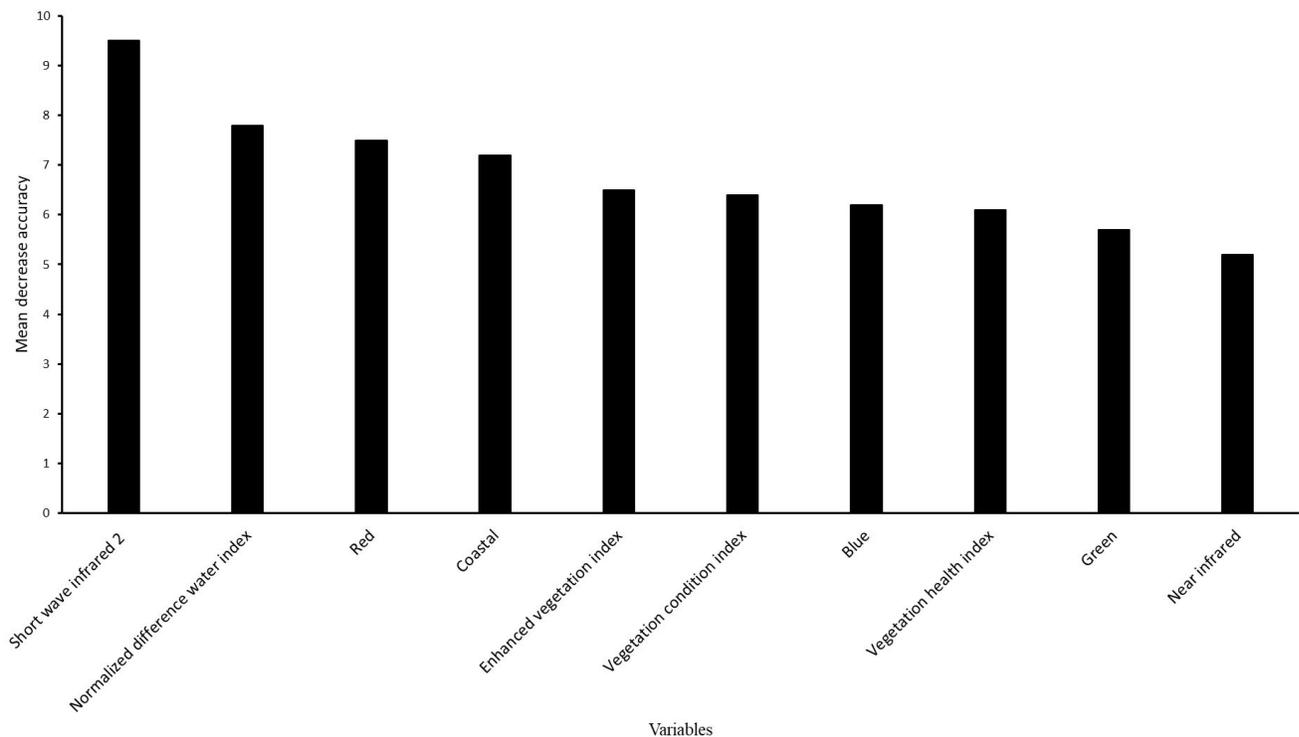
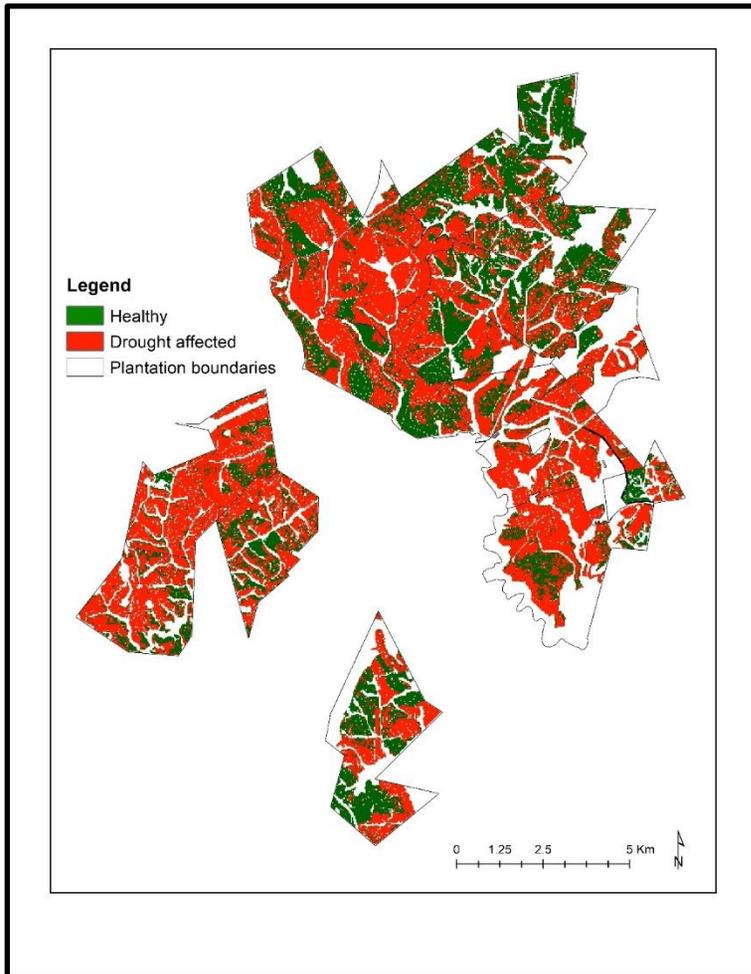


Figure 2.5 Variable importance showing most effective variables used in final prediction of the model.

2.3.7 Mapping drought distribution

There are noticeable differences between the maps classified using Landsat 8 spectra variables (Figure 2.6a) and those classified with Landsat 8 derived vegetation indices (Figure 2.6b). The August 2015 map (Figure 2.6a) contained less trees that were not affected by drought compared to the August 2015 map (Figure 2.6b).

(a)



(b)

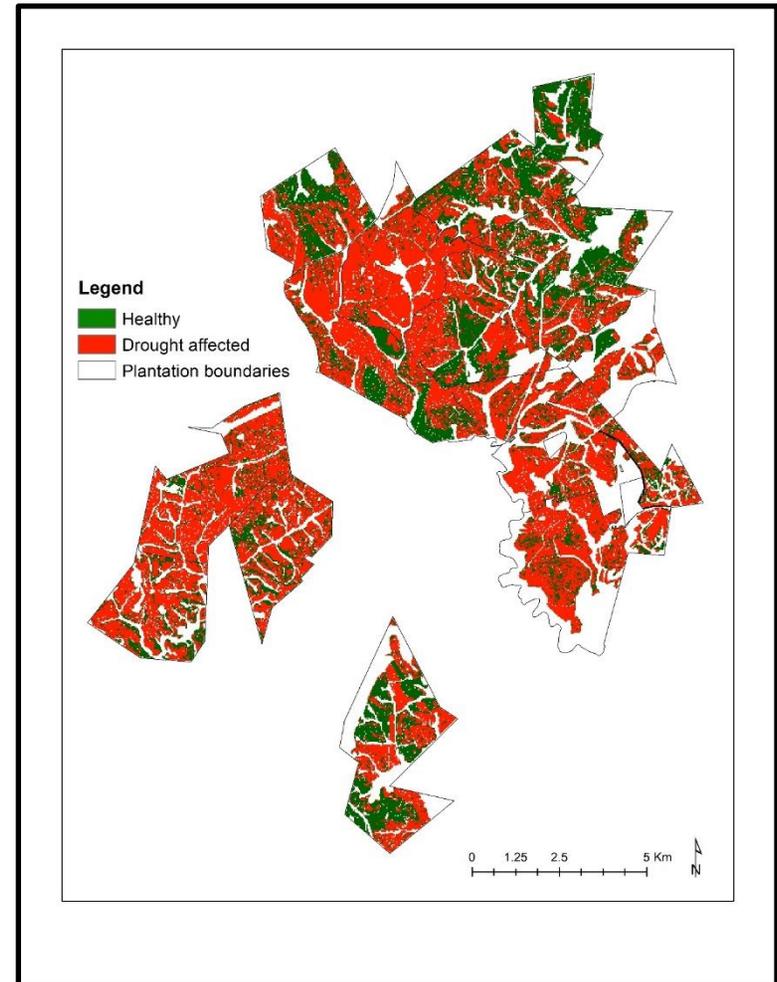


Figure 2.6 illustration of classification maps based on Landsat 8 spectra and Landsat 8 derived vegetation indices .

2.4 Discussion

Drought is considered as one of the most catastrophic natural events that threatens production in the forestry sector. Severe drought impacts forests by reducing their productivity and triggers stomatal closure which may result in tree mortality causing huge and widespread economic losses. The main essence of this study was to evaluate the utility of cost-effective Landsat 8 imagery in mapping the spatial extent of drought-prone *Eucalyptus dunnii* plantations in KwaZulu Natal. This was achieved using a combination of remote sensing vegetation indices and Landsat 8 spectral bands, whereby drought-prone plantations were successfully classified and mapped.

2.4.1 Classification using Landsat 8 and stochastic gradient boosting

The integration of vegetation indices and Landsat spectral bands produced a higher overall accuracy and kappa as well as higher user and producer accuracies compared to Landsat spectral bands only. Vegetation indices improve the accuracy of algorithms and enhances spectral information as well as separability of classes of interest (Fajji *et al.*, 2017). They discriminate landscape features better whereas spectral bands only provide spectral signatures of a particular surface or object. For instance, Godinho *et al.*, (2016) mapped the montado ecosystem using Landsat 8 multispectral data (0.43 -1.38), vegetation indices such as the enhanced vegetation index, shortwave index, carotenoid reflectance index 1, green chlorophyll index, normalized multiband index and soil adjusted total vegetation index as well as the stochastic gradient boosting algorithm. This study found that the integration of vegetation indices resulted in a significant improvement in the overall classification accuracy. This led to a difference of 4.90% overall accuracy and 0.06 in kappa value. A similar trend was also observed in the producer and user accuracies with an increment of 3.64% and 6.26%, respectively. Xie *et al.* (2019) identified suitable variables and algorithms for identifying land cover, forests and tree species based on land cover and forest types. The results indicated that multiple variables, such as vegetation indices, textures and topographic factors enhanced land cover and forest classification accuracy compared to using spectral bands exclusively. A combination of multiple variables improved forest classification by 1% to 12.7 % and land cover classification by 3.7% to 15.5% compared to spectral bands only. The study also analyzed climatic data measured over five-years from 2013 - 2017 indicated that 2015 received the lowest rainfall and highest temperatures, particularly during the

summer rainfall months of between October – March. The year 2015 was therefore called a drought year as below-average rainfall and high temperatures were recorded. According to Botai *et al.*, (2016) the Free State and North West provinces were declared drought disaster areas in the 2015/2016 hydrological year. KwaZulu-Natal was later declared as severely drought-stricken in the same year (Schreiner *et al.*, 2018). The results of this study therefore advocate for regular drought monitoring and for-repeated mapping exercises rather than a one snapshot in time analysis.

One of the major drawbacks of this study was obtaining cloud-free imagery between the summer rainfall months of October to March. These months are very crucial as they positively affect agricultural productivity. Due to difficulties in the availability of cloud free imagery between this period, the months of August 2015 to December 2015 allowed for a fixed window to conduct this study. Another shortfall was experienced when obtaining rainfall data within the study area due to sparse rainfall monitoring networks in the country since there are not many rainfall stations available. A nearby rainfall station in Ixopo was used to represent rainfall patterns of the High-flats region.

2.4.2 Variable importance and drought classification

The SGB algorithm selected a few variables that were useful in classifying drought-prone compartments. One of the variables that were selected was the near infrared band (NIR) (0.85 – 0.88 nm). The NIR bands are very useful in monitoring vegetation health (Wang *et al.*, 2012). Plant leaves generally absorb most of the radiance in the visible band pigments such as chlorophyll and xanthophylls and reflect most radiance in the NIR (Kim *et al.*, 2011). When plants are stressed the reflectance, pattern is altered due to a decrease in photosynthetic absorbance, this results in the increase of reflectance in the visible band and decrease in the NIR (Gerhards *et al.*, 2016). The NIR bands are integrated with Red bands to enhance vegetation health monitoring and provide more detail for stress detection (Mushtaq and Asima, 2016). These bands are used to derive the Normalized difference vegetation index. Zhang *et al.*, (2017) studied the effects of the 2009/2010 drought in south western China by determining the standardized anomalies of NDVI, enhanced vegetation index (EVI), normalized difference water index (NDWI) and LST. The results indicated that the NDVI, EVI and NDWI reduced while LST increased in the 2009/2010 drought stricken vegetated areas. Other bands such as the short-wave infrared (SWIR) are widely used to detect plant water stress. The NIR and SWIR bands are used to derive indices such as the normalized

difference water index. Mashaba *et al.*, (2016) revealed that NDWI better determined water stress in winter when compared to NDMI. This was also consistent with the findings of Gulacsi and Kovacs, (2015). Other indices such as the vegetation health index (VHI) and temperature condition index (TCI) were also selected by the model. The VHI is based on the average of VCI and TCI. The TCI is based on thermal emission to measure surface temperature. However, Andujar et al (2017) suggested that TCI may be best suited for detecting incipient drought than VCI. VHI was also found to have a high correlation with the standardized precipitation evapotranspiration index (SPEI) which suggested that it may be most suitable for monitoring effects of long-term droughts. Most vegetation indices were selected by the model as they enhance spectral detail and improve classification accuracy.

2.4.3 Recommendations for future research

Rainfall monitoring stations in the country are limited, therefore obtaining accurate and reliable data is a challenge. Future studies can utilize the TRMM satellite. This sensor provides important rainfall information covering the tropical and sub-tropical parts of the earth. Future studies can also look into incorporating climatic and environmental variables into algorithms. Future research may also aim to improve accuracies using higher-resolution imagery with spatial resolutions below 5m, for example IKONOS and Quick bird imagery for mapping drought prone areas using spectral bands or strategically derived indices. The methodology developed would be suitable using higher-resolution datasets for detailed analysis of drought stress mapping at a sub-compartment level. Finally, the utilization of one class classifiers such as SVM may be investigated for the detection of drought stress in forest compartments and compared to contemporary multi-class approaches such as when using SGB.

2.4.4 Conclusion

This study aimed to evaluate the utility of cost-effective Landsat 8 imagery in mapping the spatial extent of drought-prone *Eucalyptus dunnii* plantations located in the High-flats region of KwaZulu-Natal. The research undertaken showed that it is possible to classify drought-prone plantations using a combination of vegetation indices and Landsat spectral bands. The vegetation indices were also capable of detecting drought-affected parts of the plantation as lower values indicated a presence of drought. The most important variables for classifying drought affected forests using multispectral image data were prevalent in the NIR (0.85 – 0.88) , red (0.64 – 0.67) and SWIR (1.57 – 1.65) regions of the electromagnetic spectrum. The year 2015 was deemed a drought year and this was attributed to below-average rainfall and high temperatures for a long period of time. Using remote sensing tools and the stochastic gradient boosting algorithm, early mapping of drought impact on forest resources may allow for effective mitigation and protection management strategies. While this study was successful in classifying drought prone plantations in the high-flats region, the next chapter will highlight issues related to sampling of training data and apply techniques such as the one class support vector machine to address this challenge.

3 Chapter Three: Mapping of Drought stress in commercial forest plantations using one-class and multi-class Supervised Learning Approaches and remotely sensed techniques

Abstract

Forest ecosystems are recognized as vital sources of livelihoods to millions of people worldwide. This study seeks to test the utility of both one class and multiclass support vector machines to classify drought prone *Eucalyptus dunnii* forest compartments. Using a combination of Landsat 8 vegetation indices and local topographical variables, the one class SVM produced an overall accuracy of 82.35% and a kappa value of 0.73. The user's and producer's accuracies ranged from 68% - 88% for drought affected compartments while for the other class the producer's and user's accuracies ranged from 79% - 93%. The multiclass SVM produced an overall accuracy of 73.86% and a kappa value of 0.71. The user's and producer's accuracies ranged from 61% to 69% for drought affected compartments and for the other compartment classes, ranged from 84% - 90%. The study showed that using a one class SVM can be utilized to accurately map classes of interest and the use of topographical variables can be utilized to enhance the overall classification accuracy of classifiers.

Keywords: One class support vector machine, Multiclass support vector machine, topographical variables, Landsat 8, vegetation indices.

3.1 Introduction

Forest ecosystems are recognized as vital sources of livelihoods to millions of people around the world (Köhl *et al.*, 2015). They offer a wide range of ecosystem services which include carbon sequestration, contribution to soil formation, water and climate regulation (Foli *et al.*, 2014; Mori *et al.*, 2017; Aznar sanchez *et al.*, 2018). In South Africa, commercial forest plantations cover 1% of the total land area and directly supports the economic, social and environmental needs of communities, particularly in rural poor areas (Bojang and Atanga, 2011). The community at large benefit through employment opportunities, access to grazing and harvesting of non-woody products. Nonetheless, the forestry sector is identified as extremely vulnerable to climatic events such as the impact of drought, since they alter the functioning and structure of forest ecosystems (Law, 2014; Assal *et al.*, 2016).

Drought is defined as a lack of precipitation over a region for an extended period and can be classified into meteorological, economical, agricultural and socioeconomic droughts (Edossa *et al.*, 2014; Wanders and Wada, 2015). Agricultural droughts are most common and occur when there is a significant decline in soil moisture, resulting in crop or forest mortality (Yu *et al.*, 2018, Cao *et al.*, 2019). Severe drought events contribute to tree stress and mortality through the direct impacts of decreased soil moisture and high air temperatures (Reyer *et al.*, 2015). Many studies have found that an increase in drought frequency and intensity may affect forests through impeding tree recruitment, reducing growth, a reduction in leaf area index (LAI) and increasing tree mortality (Archaux and Wolters, 2006; Matusick *et al.*, 2012; Hope *et al.*, 2014; Assal *et al.*, 2016). Severe drought events usually affect large geographical areas and conventional methods for monitoring drought may not be ideal, as they are time consuming and field intensive (Naumann *et al.*, 2018). Remote sensing tools are therefore useful in drought assessments as the technology can detect changes that may not be readily observed, especially when over a wide landscape such as a commercial forest plantation (Anderson *et al.*, 2010). The detection of drought events is further enhanced by the application of remote sensing indices (Tuvdendorj *et al.*, 2019).

Several remote sensing indices have been established and tested in detecting drought. For instance, the normalized difference vegetation index (NDVI) has been widely used to monitor the effect of drought on vegetation by quantifying the amount of greenness or vigor from the radiometric properties of plants (Maselli *et al.*, 2004; Breshears *et al.*, 2005; Mohd Razali *et al.*, 2016) For

example, Sruthi and Aslam, (2015) used MODIS derived NDVI and land surface temperature (LST) to analyze vegetation stress in India. The combination of NDVI and LST provided important data for monitoring agricultural drought and served as an early warning detection for farmers. NDVI was successfully correlated with LST and yielded better results when combined with LST to detect agricultural drought. Furthermore, Sholihah *et al.*, (2016) identified agricultural drought extent using the vegetation health index (VHI). The outputs illustrated that VHI decreased more than 50% from 30.86 (mild) in 2000 to 14.66 (severe) in 2015. This was attributed to a sharp increase in LST from 27°C to 40°C in 2015. Lastly, Dutta *et al.*, (2015) attempted to establish the effectiveness of remote sensing and GIS methods in monitoring the spatio-temporal degree of agricultural drought. The National oceanic and atmospheric administration advanced very high-resolution radiometer (NOAA – AVHRR) data was used through NDVI based vegetation condition index (VCI). The VCI values of normal (2003) and drought (2002) were compared with the meteorological based standardized precipitation index, rainfall anomaly index and yield anomaly index and were found to be in agreement. The correlation between VCI and yield of major rainfed crops ($R > 0.75$) supports the effectiveness of remote sensing computed indices for evaluating agricultural drought. It is evident that vegetation indices have been widely used in drought studies. Recent studies successfully mapped drought sensitivity in areas using various algorithms (Roodposhti *et al.*, 2017).

Remotely sensed data has become an important tool for the derivation of landcover maps. In most cases, users are not interested in a complete map of the landscape, but rather in a subset of the classes found in the area. The problem usually occurs when using multiclass classifiers where there are numerous labels in a category. This is called multiclass classification. One of the drawbacks of this approach is that it is time consuming and costly. For instance, in order to map a region for a user interested in urban landcover, all training data points must be collected not only on the urban class but also on secondary classes not relevant to the user, such as forest, water, and crops, if they are present in the study. If these classes are not included in the training data set this may result in misclassifications where the classifier will assign pixels of untrained classes into trained classes. For example, if the forest class was not incorporated in the training data, pixels of forests may have been classified as a shrub or crop, which will greatly overestimate the real extent of those shrub and crop classes (Silva-Palacios *et al.*, 2017). In such instances, one class classifiers may be promising. One class classifiers, consider only the class of interest (positive class) and the main

difference between these two methods is in their training data and the amount of time and effort needed to produce it. According to Brereton (2011), the fact that one class classifiers consider only the positive class can be a major limitation, since only data about a single class is available, therefore only one side of the discriminative boundary can be established. Furthermore, in a feature space it is challenging to establish how closely a boundary should fit around a class of interest. To deal with this challenge, some one class classifiers such as the one class support vector machine (OC-SVM) assumes that the non-interest classes have a specific distribution around the class of interest (Ao *et al.*, 2017; Chaitra and Kumar, 2018).

Against this background, droughts threaten the productivity of commercial forests and places the economy of many countries at risk. Therefore, there is a need to develop efficient methods to detect the onset, magnitude and to quantify the impacts of drought. This study seeks to test the utility of both one class and multiclass support vector machine algorithms for classifying drought prone *Eucalyptus dunnii* plantations in the Highflats region of KwaZulu-Natal. The occurrence of drought is a one class problem that is widespread and surveying non- drought compartments for mapping purposes may prove unnecessary.

3.2 Methodology

3.2.1 Study Area

Eucalyptus dunnii plantations are valuable and are the most widely cultivated commercial plantation tree species in the world. Approximately 20 million hectares are planted due to its adaptability, rapid growth rate, good shape and excellent wood and fiber properties (Delgado and Pukkala, 2011). *Eucalyptus dunnii* is very good for producing kraft pulp and dissolving wood pulp. They are considered to be amongst the fastest growing wood plants in the world with mean growth rates up to 20 – 30 m³/ha. The study area is located in the Sappi High-flats forest plantations in KwaZulu-Natal situated at a latitude of 30° 16' 00'' S and longitude 30° 12' 00'' E and has a total area of 10380 hectares (Figure 3.1). The area is rural and consists mainly of agricultural plantations, traditional authority land and natural vegetation. The region has an average annual precipitation that ranges between 800-1000 mm and rainfall occurs between October – March. The mean annual temperature is 17°C with a lithology consisting mainly of coarse textured sandstone and tillite resulting in different soils present in the area. The forest plantation is dominated mainly by *Eucalyptus dunnii*.

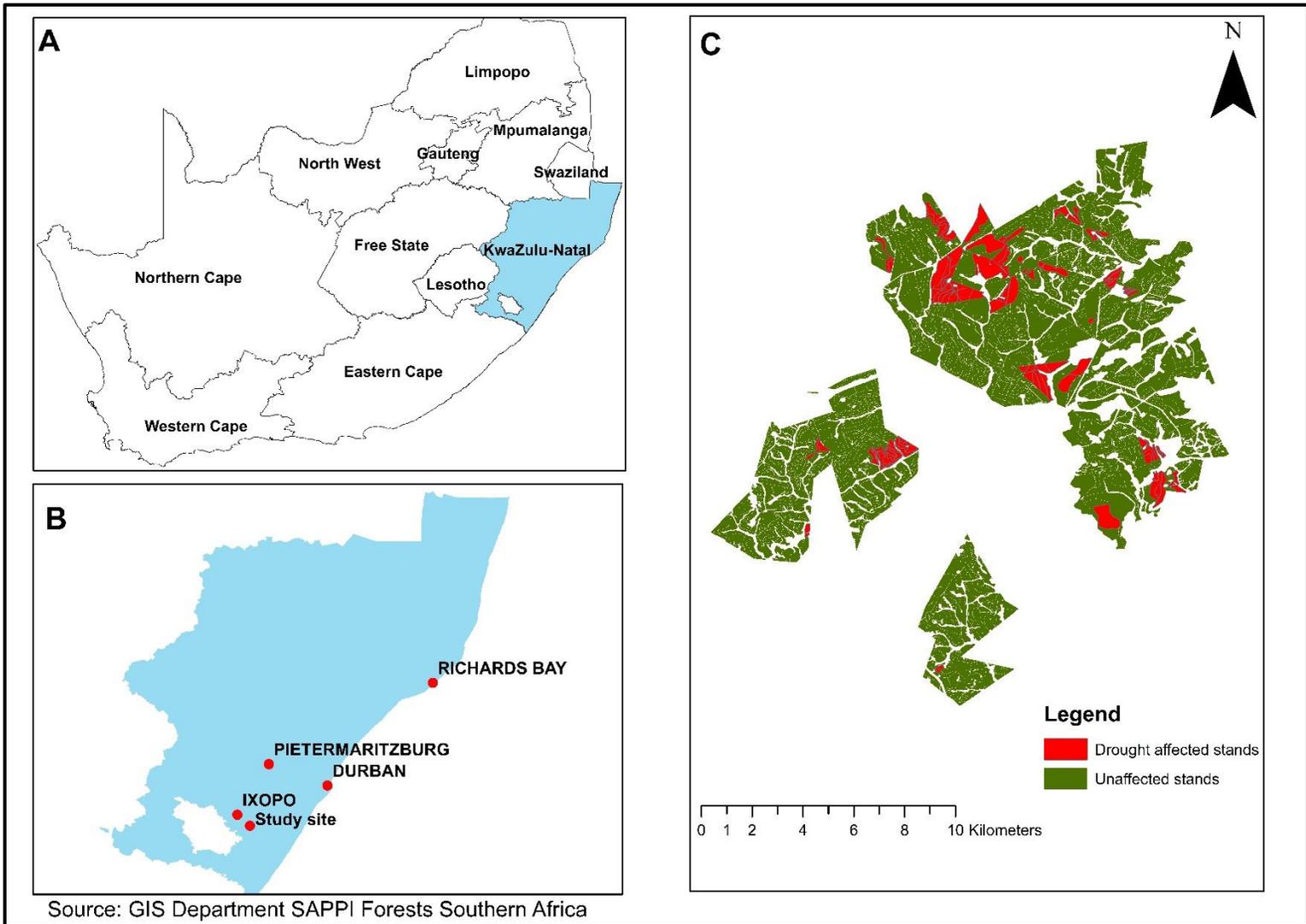


Figure 3.1 Study area showing drought affected and unaffected forest compartments in the Sappi High-flats plantation. Where (A) shows South Africa, (B) KwaZulu-Natal and (C) the study area.

3.2.2 Field verification data

Field verification was done monthly at SAPPI (South African Pulp and Paper Industry) plantation in the High-flats region of KwaZulu-Natal near Ixopo from 8th of August 2015 to 15th of December 2015. Forest compartments were physically observed for drought stress with data collected on the absence and presence of drought damaged trees and the number of trees damaged, which was expressed as the number of trees damaged over the total number of trees assessed. Compartments displaying drought induced stress of more than 50% were categorized as damaged stands, while compartments displaying 0% of damage were categorized as undamaged. Stands between 1% and 49% were removed from the analysis to avoid underestimating the occurrence of drought damage. In addition, trees that showed physical signs of stunted growth, leaf dropping and discoloration as well as embolism formation were categorized as drought damaged trees. GPS points were collected using a differentially corrected Trimble GeoXT handheld GPS receiver with an accuracy of < 1m. The GPS points data was used for ground verification of each forest compartment which was then later extracted for statistical analysis using each image dataset. For multiclass analysis, 51 stands were fire burnt, followed by 51 stands affected by drought and 51 unaffected stands to maintain a balanced dataset i.e. $51 \times 100 = 5100$ (Peerbhay *et al.*, 2013). The dataset was divided into one class and multiclass where for one class it was drought damaged versus all other data while the multiclass dataset consisted of all 3 classes i.e. burnt, drought damaged and unaffected stands.

3.2.3 Image Acquisition and Pre-processing

A cloud free Landsat-8 OLI (Operation land imager) multispectral image with path/row:168/81 was acquired from August 2015 covering the study area and was attained from the United States Geological Survey (USGS) Earth Resources Observation and Science (<https://earthexplorer.usgs.gov>). The image was acquired during a clear sky, sunny condition with azimuth angle of 38.06° and sun elevation angle of 31.92° . The image date was specifically selected to match the date at which drought damage was observed on the ground. Landsat-8 OLI has two modes of capturing images which are spectral and panchromatic. Landsat 8 OLI consists of seven spectral bands spread across the visible (433 nm – 680 nm), near infrared (845 nm - 885 nm) to short wave infrared (1560 nm – 2300 nm) regions with a spatial resolution of 30-m. The Landsat-8 OLI sensor has a temporal resolution of 16 days. Landsat-8 OLI image spectral bands were converted from digital numbers format (DN) to

reflectance using ENVI 5.1 software following the approach described on the USGS website (<http://landsat.usgs.gov>). The acquired images were atmospherically corrected using the MODTRAN method based on the Fast Line-of-sight Atmospheric analysis of Spectral hypercube (FLAASH) radiative transfer algorithm (Shao and Zhang, 2016; Dube and Mutanga, 2015).

3.4 Environmental variables

3.4.1 Topographic metrics

Spatial topographic metrics are categorized into three groups which are; non- local, local and combined topographical variables. Non-local attributes are relative positions of designated points, namely, catchment area, openness, and flow buildup (Li and McCarty, 2018). Local topographical variables are defined as the surface geometry at a particular point on the land surface such as elevation, curvatures and slope, while combined topographical variables are a combination of both local and non-local variables which are namely slope length factor, stream power index and topographical variables (Nunes *et al.*, 2019). These variables were generated from a 30 m resolution Digital Elevation Model (DEM) created from shuttle radar topography mission (SRTM) data in SAGA GIS (2.3.2) and ArcGIS 10.4 packages.

3.4.2 Bioclimatic data

Bioclimatic data such as temperature and rainfall are important indicators of drought. This study used rainfall bio-climatic variables and mean temperature in conjunction with topographic variables to classify drought affected areas (Table 3.1). The bio-climatic variables were acquired from a one square kilometer (1 km²) 30 arc seconds spatial resolution worldclim datasets (<http://www.worldclim.org/>). The worldclim datasets consists of long term mean annual data (30-year period) with grids including temperature, rainfall and other climatic data such as driest, wettest, coldest and hottest periods of the year. The generated rainfall and temperature bioclimatic variables were resampled to match the SRTM derived DEM spatial resolution (30 m).

Table 3.1 Topo-graphic variables used to assess the drought impact on commercial forestry.

N.o	Variables	Description	unit	Reference
1	Aspect	Slope direction	radian	(Vico and Porporato, 2009)
2	Catchment area	Runoff velocity and volume	m ²	(Gericke and Duplesis, 2012)
3	Convergence index	Indicates structure of relief as a set of converging(channels) and diverging areas (ridges)	m	(Adams <i>et al.</i> , 2016)
4	Cross sectional curvature	Morphometric features	degree m ⁻¹	(Ehsani and Queil, 2009)
5	Direct insolation	Potential Incoming insolation	kw/m ²	(Saad <i>et al.</i> , 2014)
6	General curvature	Horizontal and vertical curvature	degree m ⁻¹	(Rosenberg <i>et al.</i> , 2010)
7	Elevation	Ground height	m	(Kellner <i>et al.</i> , 2009)
8	Longitudinal curvature	Morphometric features	degree m ⁻¹	(Ehsani and Malekian, 2011)
9	Mass balance index (MBI)	Mass balance index	m	(Christian <i>et al.</i> , 2016)
10	Maximum curvature	Highest curvature	degree m ⁻¹	(Ehsani, 2008)
11	Minimum curvature	Lowest curvature	degree m ⁻¹	(Rana, 2006)
12	Negative openness	Drainage features, soil water content		(Cavalli <i>et al.</i> , 2013)
13	Normalized height	Relative height and slope position	m	(Adhikari <i>et al.</i> , 2018)
14	Positive openness	Drainage features, soil water content		(Li <i>et al.</i> , 2018)
15	Plan curvature	Horizontal (contour) curvature	degree m ⁻¹	(Nanomura <i>et al.</i> , 2020)

16	Profile curvature	Vertical rate of change of slope	degree m ⁻¹	(Krebs <i>et al.</i> , 2015)
17	Skyview factor	Visibility	kw/m ²	(Bernard <i>et al.</i> , 2018)
18	Slope	Steepness of the ground	radian	(Nakil and Khire, 2016)
19	Standardized height	Relative height and slope position	m	(sovilla <i>et al.</i> , 2010)
20	Terrain ruggedness Index (TRI)	Quantifies topographic heterogeneity		(Skentos and qurania, 2017)
21	Topographic wetness Index (TWI)	Steady state wetness index		(Hojati and Mokarram <i>et al.</i> , 2018)
22	Valley depth	Relative heights	m	(Tsai <i>et al.</i> , 2012)
23	Wind effect	Effect of wind on the surface	m/s	(Li <i>et al.</i> , 2019)
24	Temperature (mean annual)	Temperature	° c	(Taha <i>et al.</i> , 2018)
25	Precipitation (mean annual)	Rainfall	mm	(Yacoob and Taylor, 2009)

3.4.3 Statistical analysis

3.4.3.1 One class support vector machine

The OC-SVM was developed by Scholkopf *et al.*, (2001) and is an addition of the novel two-class algorithm, which allows training of classifiers in the absence of negative data (Senf *et al.*, 2006). Training is conducted by making a certain number of positive data points as if they belong to the negative class (Xu *et al.*, 2018). The aim is to establish a boundary between the bulk of positive data points and outliers. The OC-SVM utilizes the parameter ν to outline the tradeoff between the number of data points treated as positive and negative class. There are two approaches to generate a separating boundary (Pereira *et al.*, 2017). The first approach involves training the OC-SVM in such a way to define a classification function that adapts to a hypersphere boundary among the positive class and the outliers founded on a distribution function. The ν parameter establishes the shape of the boundary. The

second method involves fitting a hyperplane between the origin and the data points. This method has been found to be equal to the decision hypersphere and is utilized in several OC-SVM implementations due to its simplicity (An *et al.*, 2015). The requirements for a separation boundary are represented mathematically:

$$F(z) = I(d(z) < \Theta_d) \text{ or } F(z) = F(z) I(p(z) > \Theta_d) \dots\dots(1)$$

where $F(z)$ is a measure of the distance $d(z)$ to the positive class or of the probability $P(z)$ belonging to the positive class (Chen *et al.*, 2019), and a threshold Θ to describe the difference between the positive class and the outliers. I is a function representing the positive or negative class. One class classifiers learn by optimizing $d(z)$ or $p(z)$. Some approaches may further optimize the parameter Θ . In this study the parameters of the SVM were tuned in R statistical version 3.5.1 software.

3.5 Accuracy assessment

3.5.1 Confusion matrix

The confusion matrix summarizes the performance of the one class support vector machine and consists of the kappa value, error rate, overall accuracy, user’s and producer’s accuracy as well as the number of forest compartments that are correctly and incorrectly classified. The final dataset was split into 70% training data and 30% test data. The producer’s accuracy is a measure of how well the real features on the ground are represented on the classified map. It also includes error of omission which is part of the observed features not represented on the map. High errors of omission results in low producer’s accuracy. The user’s accuracy is a measure of the reliability of the map and indicates to the user how well the map represents what’s on the ground (Mas *et al.*, 2014). This is also inclusive of the error of commission where a high error of commission results in a lower user’s accuracy (Salk *et al.*, 2018). According to Visa *et al.*, (2011) a confusion matrix includes data about actual and predicted classifications. For instance, a confusion matrix of size $n \times n$ related to a classifier showing the predicted and actual classification where n is a number of different classes. Table 3.2 shows a confusion matrix with $n = 2$ and consist of entries with the following meanings:

- A is the amount of correct negative classifications
- B is the amount of incorrect positive classifications
- C is the amount of in correct positive classifications
- D is the amount of correct positive classifications

The classification accuracy can be deduced from this matrix as follows:

$$\text{Accuracy} = \frac{a+d}{a+b+c+d} \dots\dots\dots (2)$$

$$\text{Error} = \frac{b+c}{a+b+c+d} \dots\dots\dots (3)$$

$$D = \int \frac{|b-c|}{\text{MAX}\{B|C\}} \text{ if } B = C= 0 \dots\dots\dots (4)$$

3.5.2 Overall accuracy

The overall accuracy is regarded as one of the most basic metrics and is defined as the amount of pixels or points that are correctly classified (Salk *et al.*, 2018). Which is the total confusion matrix diagonal divided by the total of the entire confusion matrix. In this study the dataset was split to 70% training and 30% test and the Kappa coefficient was used to measure the performance of the OC-SVM approach. A K (KHAT) correlation between predicted and observed values was obtained and values closer to one show a perfect classification. A producer’s accuracy measures errors of omission and how well real-world land cover types can be classified (Rwanga and Ndaambuki, 2017). High errors of omission results in low producer’s accuracy. While user’s accuracy measures errors of commission and informs the user how well the map represents what’s actually on the ground (Pierce, 2015). High errors of commission results in a lower user’s accuracy.

Table 3.2 Confusion matrix with two classes

	Predicted Negative	Predicted Positive
Actual Negative	A	B
Actual positive	C	D

3.6 Results

3.6.1 Data plots using multiclass SVM using Land Sat vegetation indices and topographic variables

The aim of the multiclass support vector machine (MC-SVM) is to discover a hyperplane in an N-dimensional space that clearly classifies the data points. When categories have more than two dissimilar labels the problem is categorized as multiclass classification (Figure 3.2).

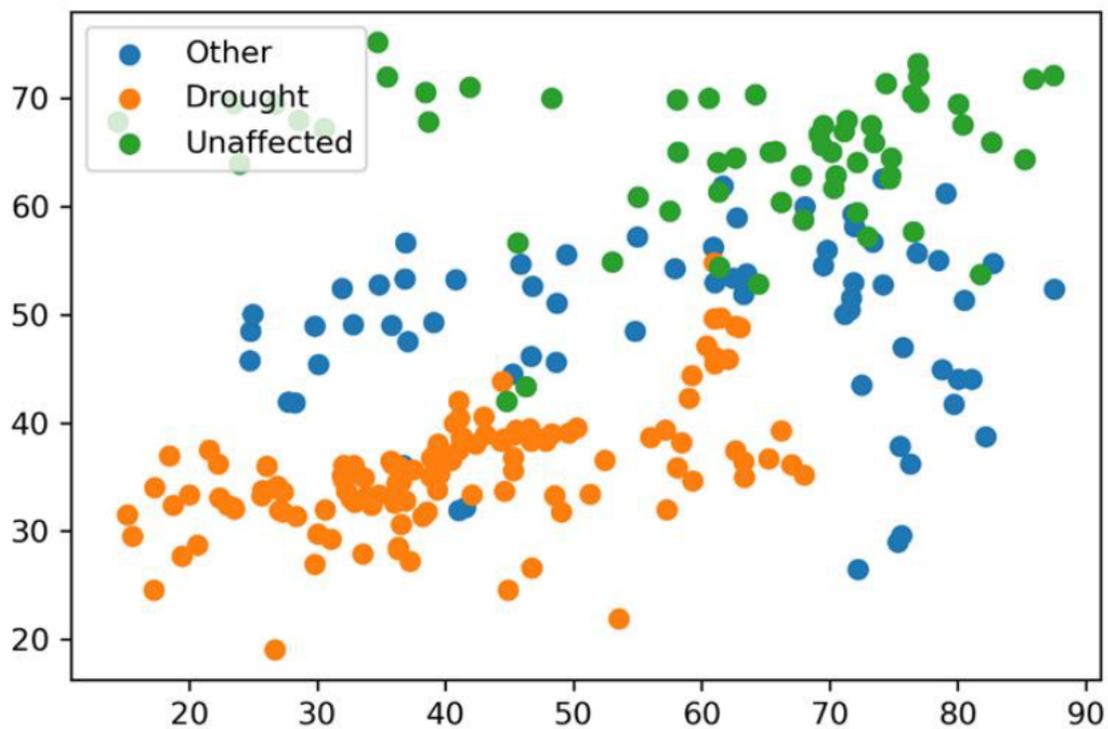


Figure 3.2 Multiclass data plot showing drought affected, unaffected and other classes generated using R software.

3.6.2 Data plots using one-class SVM using Landsat vegetation indices and topographical variables

The OC-SVM attempts to learn a decision boundary (hypersphere) that achieves the maximum separation between the points and the origin. Points belonging to a specific class of interest (drought) are called positive data while those belonging to other classes are referred to as negative data (Figure 3.3).

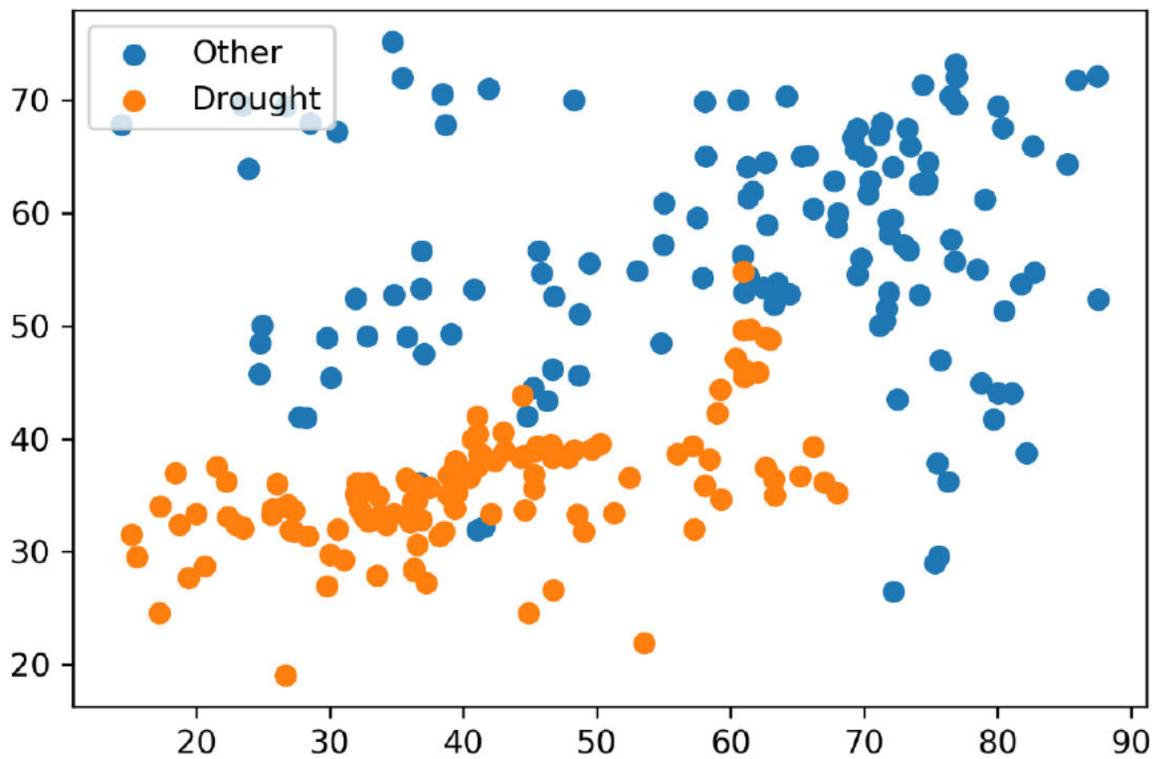


Figure 3.3 One class data plot showing positive class drought and negative class other generated using R software.

3.6.3 Classifying drought affected compartments using Landsat vegetation indices and topographical variables using a multiclass SVM

A multiclass SVM mapper was initially utilized in the study with a combination of Landsat vegetation indices and the selected topographical variables. An overall accuracy of 73.86% was obtained and a kappa value of 0.71. The user's and producer's accuracies ranged from 61% to 69% for drought affected compartments. While for the unaffected, the producers and users accuracy ranged from 84% to 90% and the error rate was 26.14% (Table 3.3).

Table 3.3 Multiclass support vector machine with Landsat 8 vegetation indices.

	Drought affected	Unaffected	Fire damage	Total
Drought affected	3500	700	1500	5700
unaffected	400	4300	100	4800
Fire	1200	100	3500	4800
Total	5100	5100	5100	15300

3.6.4 Mapping drought affected compartments using Landsat vegetation indices and topographical variables using a one class SVM

A one class mapper was utilized in the study with a combination of Landsat vegetation indices and topographical indices. Table 3.4 shows an overall accuracy of 82.35% was obtained and a kappa value of 0.73. The users and producers accuracies range from 68% to 88% for drought-affected compartments. While for the other column the producers and users accuracy range from 79% to 93%.

Table 3.4 One class support vector machine with Landsat 8 vegetation indices.

	Drought affected	Other	Total
Drought affected	4500	2100	6600
Other	600	8100	8700
Total	5100	10200	15300

3.6.5 Variable importance

Figure 3.4 indicates the most significant variables that were selected by the model. Valley depth, VCI, VHI, and longitudinal curvature had the highest Mean decrease accuracy (MDA). While mean rainfall, cross sectional curvature and positive openness had the lowest MDA. The model selected a total of 31 variables that successfully contributed in classifying drought-stricken areas within the forest plantations.

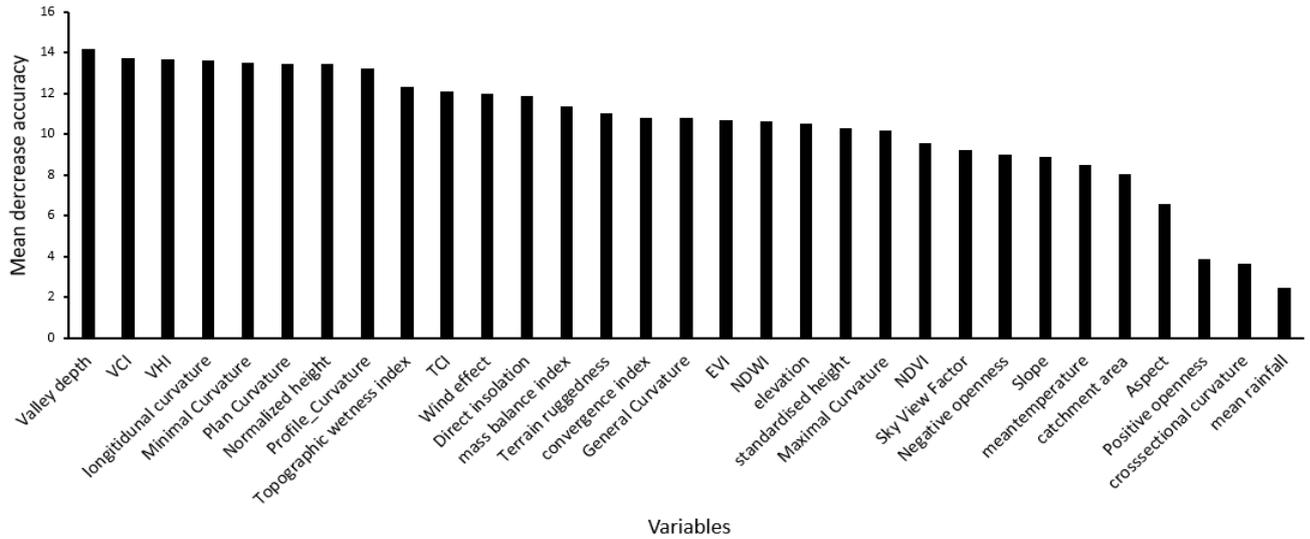
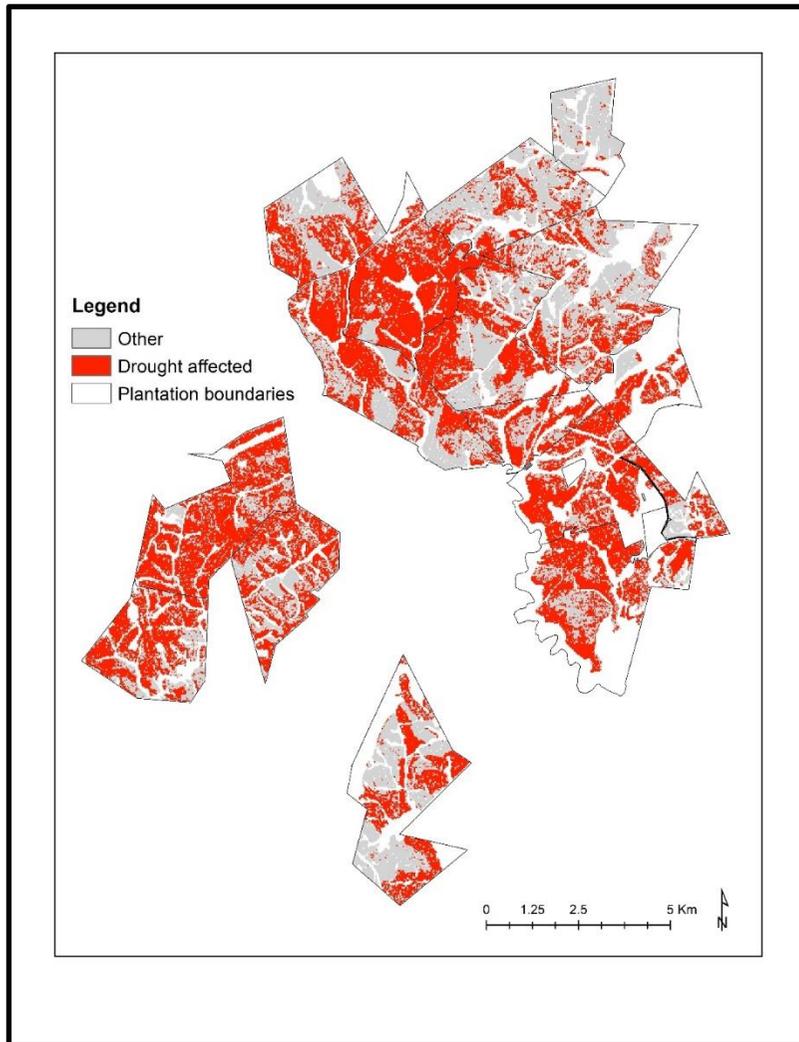


Figure 3.4 Variable importance showing contribution of each variable using the One class and multiclass support vector machine.

3.6.6 Drought distribution

There are detectable changes between the maps classified based on Landsat vegetation indices and topographical variables using one class support vector machine (Figure 3.5a) and those based on Landsat vegetation indices and topographical variables using a multiclass support vector machine (Figure 3.5b). The map in (Figure 3.5a) contained more trees that were affected by drought compared to (Figure 3.5b).

(a)



(b)

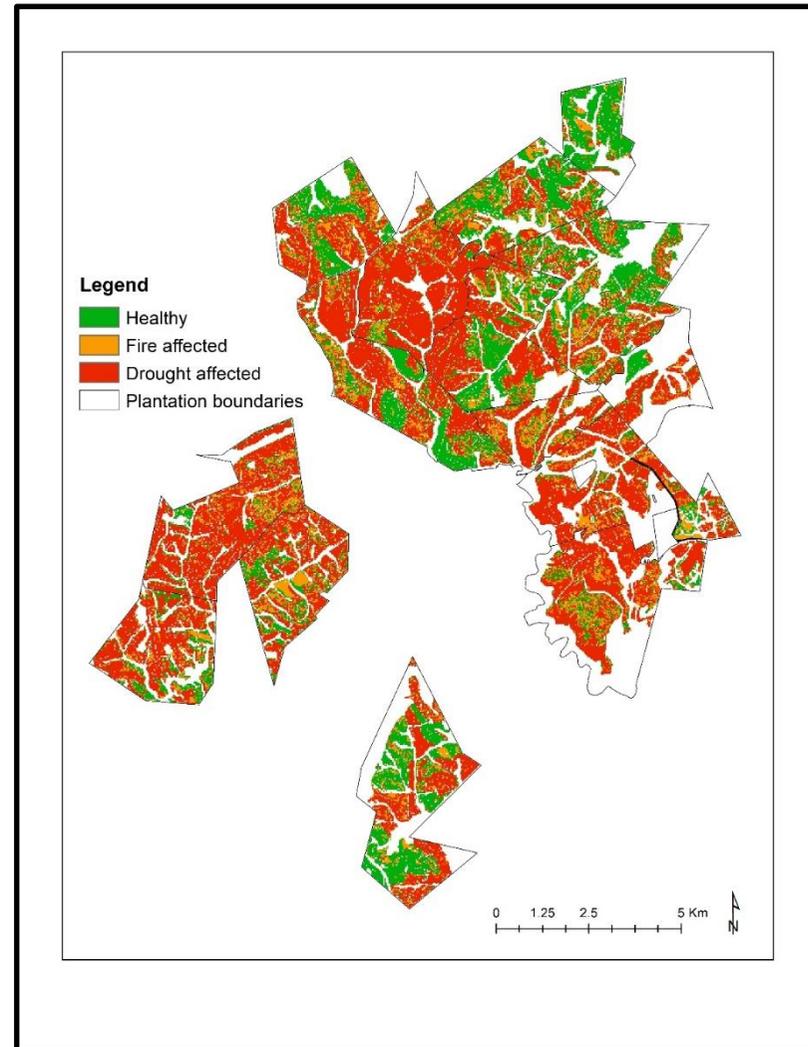


Figure 3.5 illustration of derived classification maps with one class and multiclass support vector machines.

3.7 Discussion

To improve the knowledge and understanding of drought classification, a OC-SVM was used and compared to utilizing a MC-SVM. Both Landsat vegetation indices and environmental variables were also used in this study. These variables were selected as they improved the overall accuracy compared to the previous study where Landsat spectral bands and drought indices were used. This study seeks to test the utility of a one class support vector machine for mapping drought prone *Eucalyptus dunnii* plantations in the KwaZulu-Natal High-flats region.

3.7.1 Multiclass classification using Landsat vegetation indices and topographical variables

The integration of Landsat vegetation indices and topographical variables using a MC-SVM was explored. The overall accuracy and Khat statistic (Kappa) was lower compared to that of the OC-SVM. This may be attributed to the number of training samples which tend to be exhaustive and time consuming in a multiclass classification. For instance, the class of interest for this study are those that are affected by drought and the secondary classes are those compartments that are unaffected and those damaged by fire. This required a collection of many training points including those deemed unnecessary to prevent the classifier from underestimating the class of interest. Silva *et al.*, (2017) pointed out that the MC-SVM classifiers may correctly discriminate secondary classes to the disadvantage of the class of interest, which may lead to a high probability of error. A study by Hasan *et al.*, (2019) utilized machine learning methods including SVM, artificial neural networks (ANN) and convolutional neural network (CNN) to classify thirteen vegetation species. The performance was evaluated based on their overall accuracy and each classifier was tested for the advantage associated with an increase in training samples. The accuracy obtained by CNN, ANN and SVM was 99%, 94% and 91%, respectively. It is important to note that although an increase in training samples improved the performance of the classifiers, SVM obtained a lower overall accuracy. Furthermore, Raczko and Zagajewski, (2017) compared three classification algorithms including Support vector machines (SVM), RF and ANN for tree species classification. The overall accuracies for ANN, SVM and RF was 77%, 68% and 52%, respectively. Similarly, Wang *et al.*, (2019) classified landcover of wetlands using RF, support vector machines (SVM) and K-nearest neighbor (KNN). The classification accuracies of RF, SVM and KNN were 86.61%, 79.96% and 77.23%, respectively. The studies clearly illustrated that a multiclass

SVM may not perform efficiently when a high number of training samples are used and therefore, a one class approach may be preferred, as in the findings of this study. Moughal, (2013) pointed out that as the number of classes increases, the number of parameters to be estimated increases and that in turn affects the classification performance in terms of accuracy.

3.7.2 One class classification using Landsat vegetation indices and topographical variables

The integration of Landsat vegetation indices and topographical variables produced a higher overall accuracy for OC-SVM compared to MC-SVM. This may be attributed to the fact which Deng *et al.*, (2018) points out that the OC-SVM only requires manually labelled samples of the class of interest known as the positive class and therefore it is less likely to suffer from the problem of incomplete training data. A study by Xu *et al.*, (2018) used OC-SVM to extract rice cultivated using Landsat optical land imager. Instead of sampling and training all land cover types as done by multiclass classification methods, the OC-SVM only used training samples of the target class (rice) for rice mapping. The performance of OC-SVM was evaluated in terms of the overall classification accuracy and was compared to the MC-SVM approach, decision tree classification (DTC) and vegetation index-based thresholding (VIT). The overall accuracy of OC-SVM was 91.15% which was comparable to that of DTC with 91.53%. While overall accuracy for MC-SVM was 83.85% and VIT, 57.63%. Zhao *et al.*, (2020) further utilized the biased support vector machine, a one class type- classifier that also requires labelled data for the class of interest and unlabeled data for the other classes to classify fallows in dryland cropping systems. The study found that this classifier allowed in-situ observations to be extrapolated throughout the flowering period to the rest of the growing season to create large training data sets, therefore reducing data collection requirements. The same approach was tested to monitor fallows in the northern grains of Australia and illustrated that the seasonal fallow extent could be mapped with an accuracy of over 92%, both during summer and winter seasons. In addition, Roodposhti *et al.*, (2017) attempted to produce a drought sensitivity map for vegetation cover using two OC-SVM algorithms. To attain promising results, a combination of 30 years statistical data of synoptic stations and 10 years MODIS imagery were used for extraction of SPI and EVI. For drought sensitivity mapping four variables were considered which included elevation, slope, aspect and geomorphic classes. The results of the investigation showed spatio-temporal patterns of drought impacts on vegetation cover. The Area under the curve method was used to evaluate accuracy of the output map which produced a value of 0.8.

This study showed that topographical variables such as slope, aspect elevation to name a few can be used in OC-SVM to derive effective drought sensitivity maps.

3.8 Conclusion

The study aimed to test the utility of both one class (OC-SVM) and multiclass support vector machines (MC-SVM) for classifying drought prone *Eucalyptus dunnii* plantations. Using both vegetation indices and topographical variables and a high amount of training samples for the MC-SVM produced a lower overall accuracy. It was therefore concluded that OC-SVM performs better when using a few training samples that are labelled accurately. Additionally, the results demonstrated that:

- The ability of OC-SVM to perform drought damage classifications accurately
- OC-SVM was more superior in mapping drought compared to MC-SVM
- The onset of 2015-2016 drought was before 2015
- Labelling multiple classes reduces the overall accuracy of the classifier

The results obtained in this study are promising, however, more research is still required to improve these results and to determine how good the OC-SVM can be for drought classification and analysis. Research into drought damage classification shouldn't be limited to forest cover, it should also be aimed at other environmental issues such as mapping invasive species and grasslands to name a few.

4 CHAPTER FOUR: SUMMARY OF STUDY FINDINGS

This study aimed to evaluate the utility of a cost-effective Landsat 8 imagery in mapping the spatial extent of drought prone *Eucalyptus dunnii* plantations while exploring the effectiveness of one class classification approach against multiclass algorithms. These algorithms were investigated at a forest catchment scale using vegetation indices and topographical variables. The obtained results demonstrated the capabilities of classification algorithms and advantages they present in analysing drought. The study is one of the very few studies to utilize OC-SVM and MC-SVM for the classification of drought damage on commercial forests. The study also highlighted the advantages of integrating Landsat spectra with vegetation indices over Landsat spectra only when analysing drought damage on commercial forests. In this chapter the conditions set in chapter 1 will be reviewed against the findings.

4.1 Mapping drought affected forest compartments using multispectral data and Stochastic gradient boosting

In chapter 2 of the study the stochastic gradient boosting classifier was utilized to classify drought prone *Eucalyptus dunnii* plantations. Where Landsat spectra only and an integration of Landsat spectra and vegetation indices were used. The combination of Landsat spectra and vegetation indices produced a higher overall accuracy compared Landsat spectra only. This was attributed to the fact that multiple variables such spectral bands and vegetation indices improves the accuracy of algorithms. This improved the overall accuracy by over 8.43% and this significant improvement highlights the importance of vegetation indices in drought analysis studies. The results are comparable to Godinho *et al.*, (2016) where the study found that an integration of Landsat 8 multispectral data (0.43 – 1.38), vegetation indices and the stochastic gradient boosting resulted in a significant improvement of 4.9 % in the overall accuracy. Furthermore Xie *et al.*, (2019) identified suitable variables for classifying land cover, forest and tree species. The study found that a combination of multiple variables improved forest classification by 1% to 12.7% and land cover classification by 3.7% to 15.5% compared to spectral bands only. These studies highlight the importance of integrating variables to improve the overall accuracy. It also is worth noting that the overall accuracy improvement in this study was significantly higher compared to Godinho *et al.*, (2016) and Xie *et al.*, (2019). One of the objectives of the study was to do a trend analysis of vegetation health using vegetation indices. The key finding related to this objective was the correlation of indices with

drought occurrence in 2015. This finding highlighted the importance of vegetation indices in analysing drought impacts on vegetation. Other vegetation indices such as the EVI, TCI and VHI detected drought occurrence from 2015 and continued into 2016. These indices proved to be useful for drought analysis.

A trend analysis of monthly rainfall and temperature was done for a period of five years. The analysis indicated that monthly rainfall was seasonally variable. Monthly rainfall prior to drought occurrence were higher compared to the drought year 2015 and pre-drought years 2013 and 2014. The analysis for temperature indicated that December 2015 was the highest compared to December 2013, 2014, 2016 and 2017. The highest annual temperature was measured in 2015 and the lowest in 2013. Both temperature and rainfall results indicated that 2015 was a drought year.

4.2 Mapping drought affected forest compartments using multispectral data combined with topographical variables using a one class classification approach

The combination of topographical variables, Landsat spectra and drought indices produced a higher overall accuracy (82.35%) for the OC-SVM compared to MC-SVM (73.86%). This was attributed to a low number of training samples that required labelling. Whereas MC-SVM requires all training samples to be labelled including those that are of no interest to the user. This objective highlighted the importance of one class mapping in studies that are interested mapping a class of interest. OC-SVM allows for this to be done accurately and in a timely manner. The results were comparable to Xu *et al.*, (2018) who used OC-SVM to extract rice cultivated area using Landsat optical land imager. The OC-SVM used training samples of the target class (rice) for rice mapping. The performance of OC-SVM was evaluated and was compared to the MC-SVM approach, decision tree classification (DTC) and vegetation index-based thresholding (VIT). The overall accuracy of OC-SVM was 91.15% which was comparable to that of DTC with 91.53%. While overall accuracy for MC-SVM was 83.85% and VIT, 57.63%. Zhao *et al.*, (2020) utilized the biased support vector machine one-class type classifier that also requires labelled data for the class of interest and unlabeled data for the other classes, to classify fallows in dryland cropping systems. The results showed that the seasonal fallow extent can be mapped with over 92% accuracy, both during summer and winter seasons. This study indicated the effectiveness of one class classifiers in accurately mapping drought affected plantations.

4.3 Implications of this study

The approaches derived in this study should not only be applied to forestry but also extended to other agricultural mapping exercises for key economic crops that may be at risk of being impacted by drought. Other adverse climate conditions such as frost and hail should also be examined as they pose a serious risk on agriculture.

One of the main advantages of this study is the use of a cost-effective Landsat 8 OLI imagery along with a OC-SVM and drought indices which produced good results. However, there is a need to test other one class classifiers such as the presence and background learning algorithm (PBL) and ensemble one class classification (EOCC) algorithm. For example, Ao *et al.*, (2017) successfully investigated the possibility of using a novel one class classification algorithm the PBL to classify LIDAR in an urban scenario. Whereas Liu *et al.*, (2020) used the EOCC in combination with NDVI time series analysis to detect *Spartina alterniflora* encroachment.

Lastly this study could be expanded throughout the forest industry across South Africa or expanded over larger areas using big datasets given the fact that this approach is advantageous over multiclass techniques as it is quick and cost effective.

4.4 Future research recommendations

Future research should focus on improving methods of decomposing multiclass training samples to one or binary class to ensure that all data used for training is complete and accurate. The study utilized supervised classifiers and future studies should utilize more of these classifiers to determine their efficiency compared to the ones used in this study. Also, OC-SVM can be compared to other proposed one class classifiers such as the biased support vector machine (BSVM) and maxent.

The study focused on classifying forested areas and future research can focus on different land covers such as grasslands and areas that grow important agricultural crops. It was also found that *Eucalyptus dunnii* forest species were affected by drought in certain plantations and this should be investigated thoroughly. Future research can also focus on higher-resolution remote sensing imagery to conduct drought analysis on forested catchments. Droughts are complex and affect large geographical areas therefore new methods should be developed in addition to algorithms that can investigate the extent, magnitude and duration of a drought event. In addition, studies could consider real-time analysis using a time series of images for change detection. The rate of change and stand age can help distinguish drought from other damage

types such as frost, snow, fire, pest and disease. Summer and winter baselines for various indices such as NDVI and NDWI can be set and using this information thresholds below which fall into levels of drought impact can be detected. Future studies could also use a soil deficit model to determine a single value of the number of water deficit days using temperature, rainfall and soil data. Furthermore research could look at recovery after drought to understand if, where and at what rate trees are recovering after drought and the type of recovery healthy canopy vs epicormic shoot production. Lastly, the application of Sentinel-2, which has a greater swath width of 290 km, 12 bands and spatial resolution of 10 m, would be suitable to accurately map drought prone areas.

5 Conclusion

Droughts threaten the livelihoods of people and have a devastating impact on developing countries that heavily rely on rainfed agriculture. Therefore, drought analysis is necessary to develop measures to mitigate its impacts. Drought is a complex phenomenon and requires a holistic approach of all its facets which include duration, severity and extent. This study explored the use of classification algorithms to map drought prone *Eucalyptus dunnii* plantations. The results from the study demonstrated that classification algorithms are capable of mapping and visualizing the extent of drought damage on forest plantations.

This conclusion is based on the findings from this thesis and covers research questions established in the first chapter of the thesis:

Are the use of spectral bands only sufficient as input data for stochastic gradient boosting to classify drought damage in forest plantations?

- The use of Landsat 8 spectral bands only produced a fairly good overall accuracy however the results could still be improved by using multiple variables.

Will the integration of spectral bands and vegetation indices improve the overall accuracy of drought classification using stochastic gradient boosting?

- The use of spectral bands and vegetation indices improved the overall accuracy significantly and this was attributed to the fact that vegetation indices enhances spectral information and improve the accuracy of algorithms.

Will the integration of topographical variables, and vegetation indices improve the overall accuracy of OC-SVM and MC-SVM?

- The use of topographical variables and vegetation indices produced a higher overall accuracy for OC-SVM compared to MC-SVM. This was attributed to the low amount of training data samples that are required to be labelled which results in high accuracy and is less time consuming.

Overall, this study provided insight into three classification algorithms namely SGB, OC-SVM and MC-SVM. The OC-SVM and MC-SVM produced lower overall accuracies compared to SGB. However, the results were comparable although both OC-SVM and MC-SVM didn't utilise Landsat 8 spectra as input data. Researchers should utilize this study as a base for future drought analysis using classification algorithms.

7 References

- Adams, B. A., Whipple, K. X., Hodges, K. V., & Heimsath, A. M. (2016). In situ development of high-elevation, low-relief landscapes via duplex deformation in the Eastern Himalayan hinterland, Bhutan. *Journal of Geophysical Research: Earth Surface*, 121(2), 294-319.
- Adhikari, K., Owens, P. R., Ashworth, A. J., Sauer, T. J., Libohova, Z., Richter, J. L., & Miller, D. M. (2018). Topographic controls on soil nutrient variations in a silvopasture system. *Agrosystems, Geosciences & Environment*, 1(1), 1-15.
- AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., & Hain, C. R. (2015). Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*, 53(2), 452-480.
- Albaugh, J. M., Dye, P. J., & King, J. S. (2013). *Eucalyptus and water use in South Africa*. International Journal of Forestry Research, 2013.
- Amalo, L. F., & Hidayat, R. (2017). Comparison between remote-sensing-based drought indices in East Java. In *IOP Conference Series: Earth and Environmental Science* (Vol. 54, No. 1, p. 012009). IOP Publishing
- An, W., Liang, M., & Liu, H. (2015). An improved one-class support vector machine classifier for outlier detection. *Proceedings of the institution of mechanical engineers, part c: Journal of mechanical engineering science*, 229(3), 580-588.
- Anderson, L. O., Malhi, Y., Aragão, L. E., Ladle, R., Arai, E., Barbier, N., & Phillips, O. (2010). Remote sensing detection of droughts in Amazonian forest canopies. *New Phytologist*, 187(3), 733-750.
- Ao, Z., Su, Y., Li, W., Guo, Q., & Zhang, J. (2017). One-class classification of airborne LiDAR data in urban areas using a presence and background learning algorithm. *Remote Sensing*, 9(10), 1001.
- Archaux, F., & Wolters, V. (2006). Impact of summer drought on forest biodiversity: what do we know?. *Annals of Forest Science*, 63(6), 645-652.
- Asner, G. P., P. G. Brodrick, C. B. Anderson, N. Vaughn, D. E. Knapp & R. E. Martin (2016) Progressive forest canopy water loss during the 2012–2015 California drought. *Proceedings of the National Academy of Sciences*, 113, E249-E255.
- Assal, T. J., P. J. Anderson & J. Sibold (2016) Spatial and temporal trends of drought effects in a heterogeneous semi-arid forest ecosystem. *Forest Ecology and Management*, 365, 137-151.
- Aznar-Sánchez, J. A., Belmonte-Ureña, L. J., López-Serrano, M. J., & Velasco-Muñoz, J. F. (2018). Forest ecosystem services: An analysis of worldwide research. *Forests*, 9(8), 453.

- Barradas, A., Correia, P. M., Silva, S., Mariano, P., Pires, M. C., Matos, A. R., ... & Marques da Silva, J. (2021). Comparing machine learning methods for classifying plant drought stress from leaf reflectance spectra in *Arabidopsis thaliana*. *Applied Sciences*, 11(14), 6392.
- Barradas, C., Pinto, G., Correia, B., Castro, B. B., Phillips, A. J. L., & Alves, A. (2018). Drought× disease interaction in *Eucalyptus globulus* under *Neofusicoccum eucalyptorum* infection. *Plant Pathology*, 67(1), 87-96.
- Bastos, A., Gouveia, C. M., Trigo, R. M., & Running, S. W. (2014). Analysing the spatio-temporal impacts of the 2003 and 2010 extreme heatwaves on plant productivity in Europe. *Biogeosciences*, 11(13), 3421-3435.
- Baudoin, M. A., Vogel, C., Nortje, K., & Naik, M. (2017). Living with drought in South Africa: lessons learnt from the recent El Niño drought period. *International journal of disaster risk reduction*, 23, 128-137.
- Belal, A. A., El-Ramady, H. R., Mohamed, E. S., & Saleh, A. M. (2014). Drought risk assessment using remote sensing and GIS techniques. *Arabian Journal of Geosciences*, 7(1), 35-53.
- Bernard, J., Bocher, E., Petit, G., & Palominos, S. (2018). Sky view factor calculation in urban context: computational performance and accuracy analysis of two open and free GIS tools. *Climate*, 6(3), 60.
- Bojang, F. (2011). Economic and social significance of forests for Africa's sustainable development. *Nature and Faune*, 25(2), 1-92.
- Botai, C. M., Botai, J. O., Dlamini, L. C., Zwane, N. S., & Phaduli, E. (2016). Characteristics of droughts in South Africa: a case study of free state and north west provinces. *Water*, 8(10), 439.
- Bottero, A., D'Amato, A. W., Palik, B. J., Bradford, J. B., Fraver, S., Battaglia, M. A., & Asherin, L. A. (2017). Density-dependent vulnerability of forest ecosystems to drought. *Journal of Applied Ecology*, 54(6), 1605-1614.
- Brereton, R. G. (2011). One-class classifiers. *Journal of Chemometrics*, 25(5), 225-246.
- Breshears, D. D., Cobb, N. S., Rich, P. M., Price, K. P., Allen, C. D., Balice, R. G., ... & Meyer, C. W. (2005). Regional vegetation die-off in response to global-change-type drought. *Proceedings of the National Academy of Sciences*, 102(42), 15144-15148.
- Byer, S., & Jin, Y. (2017). Detecting drought-induced tree mortality in Sierra Nevada forests with time series of satellite data. *Remote Sensing*, 9(9), 929.
- Cao, Y., Chen, S., Wang, L., Zhu, B., Lu, T., & Yu, Y. (2019). An agricultural drought index for assessing droughts using a water balance method: A case study in Jilin Province, Northeast China. *Remote Sensing*, 11(9), 1066.

- Cavalli, M., Trevisani, S., Goldin, B., Mion, E., Crema, S., & Valentinotti, R. (2013). Semi-automatic derivation of channel network from a high-resolution DTM: the example of an Italian alpine region. *European Journal of Remote Sensing*, 46(1), 152-174.
- Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., & Grégoire, J. M. (2001). Detecting vegetation leaf water content using reflectance in the optical domain. *Remote sensing of environment*, 77(1), 22-33.
- Chaitra, P. C., & Kumar, R. S. (2018). A Review of Multi-Class Classification Algorithms. *International Journal of Pure and Applied Mathematics*, 118(14), 17-26.
- Chen, Y., Wu, W., & Zhao, Q. (2019). A bat-optimized one-class support vector machine for mineral prospectivity mapping. *Minerals*, 9(5), 317.
- Chirici, G., R. Scotti, A. Montagni, A. Barbati, R. Cartisano, G. Lopez, M. Marchetti, R. E. McRoberts, H. Olsson & P. Corona (2013) Stochastic gradient boosting classification trees for forest fuel types mapping through airborne laser scanning and IRS LISS-III imagery. *International Journal of Applied Earth Observation and Geoinformation*, 25, 87-97.
- Chopra, P. (2006, January). Drought risk assessment using remote sensing and GIS: a case study of Gujarat. Enschede, The Netherlands: ITC.
- Christian, J. E., Siler, N., Koutnik, M., & Roe, G. (2016). Identifying dynamically induced variability in glacier mass-balance records. *Journal of Climate*, 29(24), 8915-8929.
- Clark, J. S., Iverson, L., Woodall, C. W., Allen, C. D., Bell, D. M., Bragg, D. C., ... & Zimmermann, N. E. (2016). The impacts of increasing drought on forest dynamics, structure, and biodiversity in the United States. *Global change biology*, 22(7), 2329-2352.
- Crous, J., Burger, L., & Sale, G. (2013). Growth response at age 10 years of five Eucalyptus genotypes planted at three densities on a drought-prone site in KwaZulu-Natal, South Africa. *Southern Forests: a Journal of Forest Science*, 75(4), 189-198.
- DAFF (2017). Annual Report 2016/2017. In F.a.F. Department of Agriculture (Ed.): Department of Agriculture, Forestry, Fisheries
- Dallas, H. F., & Rivers-Moore, N. (2014). Ecological consequences of global climate change for freshwater ecosystems in South Africa. *South African Journal of Science*, 110(5-6), 01-11.
- Dao, P. D., He, Y., & Proctor, C. (2021). Plant drought impact detection using ultra-high spatial resolution hyperspectral images and machine learning. *International Journal of Applied Earth Observation and Geoinformation*, 102, 102364.
- Das, A. C., Noguchi, R., & Ahamed, T. (2021). An Assessment of Drought Stress in Tea Estates Using Optical and Thermal Remote Sensing. *Remote Sensing*, 13(14), 2730.

- Dash, P. K., Rai, R., Mahato, A. K., Gaikwad, K., & Singh, N. K. (2017). Transcriptome landscape at different developmental stages of a drought tolerant cultivar of flax (*Linum usitatissimum*). *Frontiers in chemistry*, 5, 82.
- Deng, X., Li, W., Liu, X., Guo, Q., & Newsam, S. (2018). One-class remote sensing classification: one-class vs. binary classifiers. *International Journal of Remote Sensing*, 39(6), 1890-1910.
- Delgado-Matas, C., & Pukkala, T. (2011). Comparison of the growth of six Eucalyptus species in Angola. *International Journal of Forestry Research*, 2011.
- Dennison, P. E., Roberts, D. A., Peterson, S. H., & Rechel, J. (2005). Use of normalized difference water index for monitoring live fuel moisture. *International journal of remote sensing*, 26(5), 1035-1042.
- du Toit, B., Malherbe, G. F., Kunneke, A., Seifert, T., & Wessels, C. B. (2017). Survival and long-term growth of eucalypts on semi-arid sites in a Mediterranean climate, South Africa. *Southern Forests: a Journal of Forest Science*, 79(3), 235-249.
- Dube, T. & O. Mutanga (2015) Investigating the robustness of the new Landsat-8 Operational Land Imager derived texture metrics in estimating plantation forest aboveground biomass in resource constrained areas. *ISPRS Journal of Photogrammetry and Remote sensing*, 108, 12-32.
- Dube, T., Mutanga, O., Abdel-Rahman, E. M., Ismail, R., & Slotow, R. (2015). Predicting Eucalyptus spp. stand volume in Zululand, South Africa: an analysis using a stochastic gradient boosting regression ensemble with multi-source data sets. *International Journal of Remote Sensing*, 36(14), 3751-3772.
- Dutta, D., Kundu, A., Patel, N. R., Saha, S. K., & Siddiqui, A. R. (2015). Assessment of agricultural drought in Rajasthan (India) using remote sensing derived Vegetation Condition Index (VCI) and Standardized Precipitation Index (SPI). *The Egyptian Journal of Remote Sensing and Space Science*, 18(1), 53-63.
- Edossa, D. C., Woyessa, Y. E., and Welderufael, W. A. (2014). Analysis of droughts in the central region of South Africa and their association with SST anomalies. *International Journal of Atmospheric Sciences*, 2014.
- Ehsani, A. H. (2008). Morphometric and Landscape Feature Analysis with Artificial Neural Networks and SRTM data: Applications in Humid and Arid Environments (Doctoral dissertation, KTH).
- Ehsani, A. H., & Malekian, A. (2011). Landforms identification using neural network-self organizing map and SRTM data.
- Ehsani, A. H., & Quiel, F. (2009). Self-organizing maps for multi-scale morphometric feature identification using shuttle radar topography mission data. *Geocarto International*, 24(5), 335-355.

- Fajji, N. G., Palamuleni, L. G., & Mlambo, V. (2017). Evaluating derived vegetation indices and cover fraction to estimate rangeland aboveground biomass in semi-arid environments. *South African Journal of Geomatics*, 6(3), 333-348.
- Foli, S., Reed, J., Clendenning, J., Petrokofsky, G., Padoch, C., & Sunderland, T. (2014). To what extent does the presence of forests and trees contribute to food production in humid and dry forest landscapes?: a systematic review protocol. *Environmental Evidence*, 3(1), 1-8.
- Freeman, E. A., Moisen, G. G., Coulston, J. W., & Wilson, B. T. (2016). Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance. *Canadian Journal of Forest Research*, 46(3), 323-339.
- Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58(3), 257-266.
- García Chevesich, P., Neary, D. G., Scott, D. F., Benyon, R. G., & Reyna, T. (2017). Forest management and the impact on water resources: a review of 13 countries.
- Gerhards, M., Rock, G., Schlerf, M., & Udelhoven, T. (2016). Water stress detection in potato plants using leaf temperature, emissivity, and reflectance. *International journal of applied Earth observation and geoinformation*, 53, 27-39.
- Ghaleb, F., Mario, M., & Sandra, A. N. (2015). Regional landsat-based drought monitoring from 1982 to 2014. *Climate*, 3(3), 563-577.
- Godinho, S., Guiomar, N., & Gil, A. (2016). Using a stochastic gradient boosting algorithm to analyse the effectiveness of Landsat 8 data for montado land cover mapping: Application in southern Portugal. *International Journal of Applied Earth Observation and Geoinformation*, 49, 151-162.
- Gulácsi, A., & Kovács, F. (2015). Drought monitoring with spectral indices calculated from MODIS satellite images in Hungary. *Journal of Environmental Geography*, 8(3-4), 11-20.
- Hais, M., Hellebrandová, K. N., & Šrámek, V. (2019). Potential of Landsat spectral indices in regard to the detection of forest health changes due to drought effects. *Journal of Forest Science*, 65(2), 70-78.
- Han, H., Guo, X., & Yu, H. (2016, August). Variable selection using mean decrease accuracy and mean decrease gini based on random forest. In 2016 7th IEEE international conference on software engineering and service science (icse) (pp. 219-224). IEEE.

- Hasan, M., Ullah, S., Khan, M. J., & Khurshid, K. (2019). Comparative analysis of SVM, ANN and CNN for classifying vegetation species using hyperspectral thermal infrared data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 1861-1868.
- Hojati, M., & Mokarram, M. (2016). Determination of a topographic wetness index using high resolution digital elevation models. *European Journal of Geography*, 7(4), 41-52.
- Hope, A., Fouad, G., & Granovskaya, Y. (2014). Evaluating drought response of southern Cape Indigenous Forests, South Africa, using MODIS data. *International journal of remote sensing*, 35(13), 4852-4864.
- Huang, C. Y., & Anderegg, W. R. (2012). Large drought-induced aboveground live biomass losses in southern Rocky Mountain aspen forests. *Global Change Biology*, 18(3), 1016-1027.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, 83(1-2), 195-213.
- Idris, M. H., & Mahrup, M. (2017). Changes in hydrological response of forest conversion to agroforestry and rainfed agriculture in Renggung Watershed, Lombok, Eastern Indonesia. *Jurnal Manajemen Hutan Tropika*, 23(2), 102-110.
- Kellner, J. R., Clark, D. B., & Hofton, M. A. (2009). Canopy height and ground elevation in a mixed-land-use lowland Neotropical rain forest landscape: Ecological Archives E090-233. *Ecology*, 90(11), 3274-3274.
- Kim, Y., Glenn, D. M., Park, J., Ngugi, H. K., & Lehman, B. L. (2011). Hyperspectral image analysis for water stress detection of apple trees. *Computers and Electronics in Agriculture*, 77(2), 155-160.
- Kogan, F. N. (1995). Application of vegetation index and brightness temperature for drought detection. *Advances in space research*, 15(11), 91-100.
- Köhl, M., Lasco, R., Cifuentes, M., Jonsson, Ö., Korhonen, K. T., Mundhenk, P., ... & Stinson, G. (2015). Changes in forest production, biomass and carbon: Results from the 2015 UN FAO Global Forest Resource Assessment. *Forest Ecology and Management*, 352, 21-34.
- Krebs, P., Stocker, M., Pezzatti, G. B., & Conedera, M. (2015). An alternative approach to transverse and profile terrain curvature. *International Journal of Geographical Information Science*, 29(4), 643-666.
- Kuswanto, H., & Naufal, A. (2019). Evaluation of performance of drought prediction in Indonesia based on TRMM and MERRA-2 using machine learning methods. *MethodsX*, 6, 1238-1251.

- Lausch, A., Bastian, O., Klotz, S., Leitão, P. J., Jung, A., Rocchini, D., ... & Knapp, S. (2018). Understanding and assessing vegetation health by in situ species and remote-sensing approaches. *Methods in ecology and evolution*, 9(8), 1799-1809.
- Law, B. E. (2014). Regional analysis of drought and heat impacts on forests: current and future science directions. *Global Change Biology*, 20(12), 3595-3599.
- Lewińska, K. E., Ivits, E., Schardt, M., & Zebisch, M. (2016). Alpine forest drought monitoring in South Tyrol: PCA based synergy between scPDSI data and MODIS derived NDVI and NDII7 time series. *Remote Sensing*, 8(8), 639.
- Li, B., Yu, Q., & Peng, L. (2022). Ensemble of fast learning stochastic gradient boosting. *Communications in Statistics-Simulation and Computation*, 51(1), 40-52.
- Li, X., & McCarty, G. W. (2018). Use of principal components for scaling up topographic models to map soil redistribution and soil organic carbon. *JoVE (Journal of Visualized Experiments)*, (140), e58189.
- Li, X., McCarty, G. W., Lang, M., Ducey, T., Hunt, P., & Miller, J. (2018). Topographic and physicochemical controls on soil denitrification in prior converted croplands located on the Delmarva Peninsula, USA. *Geoderma*, 309, 41-49.
- Li, Y., Li, Y. G., Li, Q. S., & Tee, K. F. (2019). Investigation of wind effect reduction on square high-rise buildings by corner modification. *Advances in Structural Engineering*, 22(6), 1488-1500.
- Liu, X., Liu, H., Datta, P., Frey, J., & Koch, B. (2020). Mapping an Invasive Plant *Spartina alterniflora* by Combining an Ensemble One-Class Classification Algorithm with a Phenological NDVI Time-Series Analysis Approach in Middle Coast of Jiangsu, China. *Remote Sensing*, 12(24), 4010.
- Mas, J. F., Pérez-Vega, A., Ghilardi, A., Martínez, S., Loya-Carrillo, J. O., & Vega, E. (2014). A suite of tools for assessing thematic map accuracy. *Geography Journal*, 2014.
- Maselli, F. (2004). Monitoring forest conditions in a protected Mediterranean coastal area by the analysis of multiyear NDVI data. *Remote sensing of environment*, 89(4), 423-433.
- Mashaba, Z., Chirima, G., Botai, J., Combrinck, L., & Munghemezulu, C. (2016). Evaluating spectral indices for winter wheat health status monitoring in Bloemfontein using Lsat 8 data. *South African Journal of Geomatics*, 5(2), 227-243.
- Masih, I., Maskey, S., Mussá, F. E. F., & Trambauer, P. (2014). A review of droughts on the African continent: a geospatial and long-term perspective. *Hydrology and Earth System Sciences*, 18(9), 3635-3649.
- Matusick, G., Ruthrof, K. X., & Hardy, G. S. J. (2012). Drought and heat triggers sudden and severe dieback in a dominant Mediterranean-type woodland species. *Open Journal of Forestry*, 2(4), 183-186.

- Meza, I., Rezaei, E. E., Siebert, S., Ghazaryan, G., Nouri, H., Dubovyk, O., ... & Hagenlocher, M. (2021). Drought risk for agricultural systems in South Africa: Drivers, spatial patterns, and implications for drought risk management. *Science of the Total Environment*, 799, 149505.
- Mohd Razali, S., Marin Atucha, A. A., Nuruddin, A. A., Abdul Hamid, H., & Mohd Shafri, H. Z. (2016). Monitoring vegetation drought using MODIS remote sensing indices for natural forest and plantation areas. *Journal of Spatial Science*, 61(1), 157-172.
- Monyela, B. M. (2017). A two-year long drought in summer 2014/2015 and 2015/2016 over South Africa (Master's thesis, University of Cape Town).
- Mori, A. S., Lertzman, K. P., & Gustafsson, L. (2017). Biodiversity and ecosystem services in forest ecosystems: a research agenda for applied forest ecology. *Journal of Applied Ecology*, 54(1), 12-27.
- Mosley, L. M. (2015). Drought impacts on the water quality of freshwater systems; review and integration. *Earth-Science Reviews*, 140, 203-214.
- Moughal, T. A. 2013. Hyperspectral image classification using support vector machine. In *Journal of Physics: Conference Series* (Vol. 439, No. 1, p. 012042). IOP Publishing.
- Mushtaq, A. G., & Asima, N. (2016). Determining the Vegetation Indices (NDVI) from landsat 8 satellite data. Article DOI 10.21474/IJAR01/1348. *International Journal of Advance Research*, 4(8), 1459-1463.
- Nakil, M., & Khire, M. 2016. Effect of slope steepness parameter computations on soil loss estimation: review of methods using GIS. *Geocarto international*, 31(10), 1078-1093.
- Nasiłowska, S. A., Kotlarz, J., Kacprzak, M., Rynkiewicz, A., Rotchimmel, K., & Kubiak, K. (2019). The impact of drought in 2015 on the health forest condition determined using Landsat-8 OLI images. *Lesne Prace Badawcze*, 80(1), 55-68.
- Naumann, G., Alfieri, L., Wyser, K., Mentaschi, L., Betts, R. A., Carrao, H., ... & Feyen, L. (2018). Global changes in drought conditions under different levels of warming. *Geophysical Research Letters*, 45(7), 3285-3296.
- Ngaka, M. J. (2012). Drought preparedness, impact and response: A case of the Eastern Cape and Free State provinces of South Africa. *Jàmá: Journal of Disaster Risk Studies*, 4(1), 1-10.
- Nonomura, A., Hasegawa, S., Kanbara, D., Tadono, T., & Chiba, T. (2020). Topographic analysis of landslide distribution using AW3D30 data. *Geosciences*, 10(4), 115.
- Novaković, J. D., Veljović, A., Ilić, S. S., Papić, Ž., & Milica, T. (2017). Evaluation of classification models in machine learning. *Theory and Applications of Mathematics & Computer Science*, 7(1), 39-46.

- Nunes, A., Köbel, M., Pinho, P., Matos, P., Costantini, E. A., Soares, C., ... & Branquinho, C. (2019). Local topographic and edaphic factors largely predict shrub encroachment in Mediterranean drylands. *Science of the Total Environment*, 657, 310-318.
- Pasho, E., Camarero, J. J., de Luis, M., & Vicente-Serrano, S. M. (2011). Impacts of drought at different time scales on forest growth across a wide climatic gradient in north-eastern Spain. *Agricultural and Forest Meteorology*, 151(12), 1800-1811.
- Patel, N., & Kaushal, B. K. (2010). Improvement of user's accuracy through classification of principal component images and stacked temporal images. *Geo-spatial Information Science*, 13(4), 243-248.
- Patil, G. P., & Taillie, C. (2003). Modeling and interpreting the accuracy assessment error matrix for a doubly classified map. *Environmental and Ecological Statistics*, 10(3), 357-373.
- Peerbhay, K. Y., Mutanga, O., & Ismail, R. (2013). Commercial tree species discrimination using airborne AISA Eagle hyperspectral imagery and partial least squares discriminant analysis (PLS-DA) in KwaZulu-Natal, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79, 19-28.
- Peltier, D. M., Fell, M., & Ogle, K. (2016). Legacy effects of drought in the southwestern United States: A multi-species synthesis. *Ecological Monographs*, 86(3), 312-326.
- Pereira, A. A., Pereira, J. M., Libonati, R., Oom, D., Setzer, A. W., Morelli, F., ... & De Carvalho, L. M. T. (2017). Burned area mapping in the Brazilian Savanna using a one-class support vector machine trained by active fires. *Remote Sensing*, 9(11), 1161.
- Pierce, K. B. 2015. Accuracy optimization for high resolution object-based change detection: An example mapping regional urbanization with 1-m aerial imagery. *Remote Sensing*, 7(10), 12654-12679.
- Pontius Jr, R. G., & Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32(15), 4407-4429.
- Potter, C. (2015). Assessment of the immediate impacts of the 2013–2014 drought on ecosystems of the California Central Coast. *Western North American Naturalist*, 75(2), 129-145.
- Raczko, E., & Zagajewski, B. (2017). Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing*, 50(1), 144-154.
- Rana, S. (2006). Use of plan curvature variations for the identification of ridges and channels on DEM. In *Progress in spatial data handling* (pp. 789-804). Springer, Berlin, Heidelberg.
- Reyer, C. P., Rammig, A., Brouwers, N., and Langerwisch, F. (2015). Forest resilience, tipping points and global change processes. *Journal of Ecology*, 103(1), 1-4.

- Roodposhti, M. S., Safarrad, T., & Shahabi, H. (2017). Drought sensitivity mapping using two one-class support vector machine algorithms. *Atmospheric Research*, 193, 73-82.
- Rosenberg, H., Souam, R., & Toubiana, E. (2010). General curvature estimates for stable H-surfaces in 3-manifolds applications. *Journal of Differential Geometry*, 84(3), 623-648.
- Rouse Jr, J. W., Haas, R. H., Deering, D. W., Schell, J. A., & Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation (No. E75-10354).
- Rousta, I., Saberi, M. A., Mahmood, S. A. R., Moghaddam, M. M., Olafsson, H., Krzyszcak, J., & Baranowski, P. (2020). Climate Change impacts on vegetation and agricultural drought in the basin of Panjshir River in Afghanistan. *Climate Change Research*, 1(4), 77-88.
- Rwanga, S. S., & Ndambuki, J. M. (2017). Accuracy assessment of land use/land cover classification using remote sensing and GIS. *International Journal of Geosciences*, 8(04), 611.
- Saad, N. M., Hamid, J. A., & Suldi, A. M. (2014). The reliance of insolation pattern on surface aspect. In *IOP Conference Series: Earth and Environmental Science* (Vol. 18, No. 1, p. 012165). IOP Publishing.
- Salk, C., Fritz, S., See, L., Dresel, C., and McCallum, I. (2018). An exploration of some pitfalls of thematic map assessment using the new map tools resource. *Remote Sensing*, 10(3), 376.
- Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7), 1443-1471.
- Schreiner, B. G., E. D. Mungatana & H. Baleta (2018) Impacts of Drought Induced Water Shortages in South Africa: Sector Policy Briefs.
- Schulze, R. E., Lorentz, S. A., Horan, M. J. C., & Maharaj, M. (2007). Sediment yield. Schulze, RE (Editor).
- Senf, A., Chen, X. W., & Zhang, A. (2006, October). Comparison of one-class SVM and two-class SVM for fold recognition. In *International Conference on Neural Information Processing* (pp. 140-149). Springer, Berlin, Heidelberg.
- Shao, Z., & Zhang, L. (2016). Estimating forest aboveground biomass by combining optical and SAR data: a case study in Genhe, Inner Mongolia, China. *Sensors*, 16(6), 834.
- Sholihah, R. I., Trisasongko, B. H., Shiddiq, D., La Ode, S. I., Kusdaryanto, S., & Panuju, D. R. 2016. Identification of agricultural drought extent based on vegetation health indices of Landsat data: case of Subang and Karawang, Indonesia. *Procedia Environmental Sciences*, 33, 14-20.

- Silva, J., Bacao, F., & Caetano, M. (2017). Specific land cover class mapping by semi-supervised weighted support vector machines. *Remote Sensing*, 9(2), 181.
- Silva-Palacios, D., Ferri, C., & Ramírez-Quintana, M. J. (2017). Improving performance of multiclass classification by inducing class hierarchies. *Procedia Computer Science*, 108, 1692-1701.
- Skentos, A., & Ourania, A. (2017). Landform analysis using terrain attributes. A Gis application on the island of Ikaria (Aegean Sea, Greece). *Annals of Valahia University of Targoviste, Geographical Series*, 17(1), 90-97.
- Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., & Mochizuki, K. I. (2017). Mapping crop cover using multi-temporal Landsat 8 OLI imagery. *International Journal of Remote Sensing*, 38(15), 4348-4361.
- Sovilla, B., McElwaine, J. N., Schaer, M., & Vallet, J. (2010). Variation of deposition depth with slope angle in snow avalanches: Measurements from Vallée de la Sionne. *Journal of Geophysical Research: Earth Surface*, 115(F2).
- Spinoni, J., Barbosa, P., Bucchignani, E., Cassano, J., Cavazos, T., Christensen, J. H., ... & Dosio, A. (2020). Future global meteorological drought hot spots: a study based on CORDEX data. *Journal of Climate*, 33(9), 3635-3661.
- Sruthi, S., & Aslam, M. M. (2015). Agricultural drought analysis using the NDVI and land surface temperature data; a case study of Raichur district. *Aquatic Procedia*, 4, 1258-1264.
- Taha, H., Levinson, R., Mohegh, A., Gilbert, H., Ban-Weiss, G., & Chen, S. (2018). Air-temperature response to neighborhood-scale variations in albedo and canopy cover in the real world: Fine-resolution meteorological modeling and mobile temperature observations in the Los Angeles climate archipelago. *Climate*, 6(2), 53.
- Tsai, Y., Wu, Y., Lai, J., & Geary, G. (2012). Ridge-to-valley depth measured with road profiler to control micromilled pavement textures for super-thin resurfacing on i-95. *Transportation research record*, 2306(1), 144-150.
- Tuvdendorj, B., Wu, B., Zeng, H., Batdelger, G., & Nanzad, L. (2019). Determination of appropriate remote sensing indices for spring wheat yield estimation in Mongolia. *Remote Sensing*, 11(21), 2568.
- Ustin, S. L., Riaño, D., & Hunt, E. R. (2012). Estimating canopy water content from spectroscopy. *Israel Journal of Plant Sciences*, 60(1-2), 9-23.
- Van Loon, A. F. (2015). Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water*, 2(4), 359-392.
- Vico, G., and Porporato, A. (2009). Probabilistic description of topographic slope and aspect. *Journal of Geophysical Research: Earth Surface*, 114(F1).

- Visa, S., Ramsay, B., Ralescu, A. L., & Van Der Knaap, E. (2011). Confusion matrix-based feature selection. *MAICS*, 710(1), 120-127.
- Wanders, N., and Wada, Y. (2015). Human and climate impacts on the 21st century hydrological drought. *Journal of Hydrology*, 526, 208-220.
- Wanders, N., Wada, Y., & Van Lanen, H. A. J. (2015). Global hydrological droughts in the 21st century under a changing hydrological regime. *Earth System Dynamics*, 6(1), 1-15.
- Wang, M., Shi, W., & Jiang, L. (2012). Atmospheric correction using near-infrared bands for satellite ocean color data processing in the turbid western Pacific region. *Optics Express*, 20(2), 741-753.
- Wang, X., Gao, X., Zhang, Y., Fei, X., Chen, Z., Wang, J., ... & Zhao, H. (2019). Land-Cover classification of coastal wetlands using the RF algorithm for Worldview-2 and Landsat 8 images. *Remote Sensing*, 11(16), 1927.
- Warrens, M. J. (2015). Five ways to look at Cohen's kappa. *Journal of Psychology & Psychotherapy*, 5(4), 1.
- Wolff, E., & van Vliet, M. T. (2021). Impact of the 2018 drought on pharmaceutical concentrations and general water quality of the Rhine and Meuse rivers. *Science of the Total Environment*, 778, 146182.
- Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. *Journal of plant ecology*, 1(1), 9-23.
- Xie, Z., Chen, Y., Lu, D., Li, G., & Chen, E. (2019). Classification of land cover, forest, and tree species classes with ZiYuan-3 multispectral and stereo data. *Remote Sensing*, 11(2), 164.
- Xu, X., Ji, X., Jiang, J., Yao, X., Tian, Y., Zhu, Y., ... & Cheng, T. (2018). Evaluation of one-class support vector classification for mapping the paddy rice planting area in Jiangsu Province of China from Landsat 8 OLI imagery. *Remote Sensing*, 10(4), 546.
- Xulu, S., K. Peerbhay, M. Gebreslasie & R. Ismail (2018) Drought Influence on Forest Plantations in Zululand, South Africa, Using MODIS Time Series and Climate Data. *Forests*, 9, 528.
- Xulu, S., Peerbhay, K., Gebreslasie, M., & Ismail, R. (2019). Unsupervised clustering of forest response to drought stress in Zululand region, South Africa. *Forests*, 10(7), 531.
- Yacoub, E., & Tayfur, G. (2019). Trend analysis of temperature and precipitation in Trarza region of Mauritania. *Journal of Water and Climate Change*, 10(3), 484-493.
- Yu, C., Huang, X., Chen, H., Huang, G., Ni, S., Wright, J. S., ... & Yu, L. (2018). Assessing the impacts of extreme agricultural droughts in China under climate and socioeconomic changes. *Earth's Future*, 6(5), 689-703.

- Zhang, Y., Peng, C., Li, W., Fang, X., Zhang, T., Zhu, Q., ... & Zhao, P. (2013). Monitoring and estimating drought-induced impacts on forest structure, growth, function, and ecosystem services using remote-sensing data: recent progress and future challenges. *Environmental Reviews*, 21(2), 103-115.
- Zhang, Q., Shao, M. A., Jia, X., & Wei, X. (2017). Relationship of climatic and forest factors to drought-and heat-induced tree mortality. *PLoS One*, 12(1), e0169770.
- Zhao, L., Waldner, F., Scarth, P., Mack, B., & Hochman, Z. (2020). Combining Fractional Cover Images with One-Class Classifiers Enables Near Real-Time Monitoring of Fallows in the Northern Grains Region of Australia. *Remote Sensing*, 12(8), 1337.