UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Mapping and Monitoring the distribution and invasion of Mesquite (*Prosopies Glandulosa*) along the AL Gash River Kassala, Sudan using remote sensing techniques

By

Dineo Revinwa Mokgehle (1095685)

A research report submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirement for the degree of Master of Science in GIS and Remote Sensing at the School of Geography, Archaeology & Environmental Studies

Supervisors: Dr Cletah Shoko and Dr Elhadi Adam
April 2020
Johannesburg, South Africa

DECLARATION

I, **Dineo Mokgehle** declare that this research report entitled 'Mapping and monitoring the distribution and invasion of Mesquite (*Prosopis Glandulosa)* along the AL Gash River Kassala, Sudan using remote sensing techniques' is my original work. This work is being submitted for the degree of Master of Science in GIS and Remote sensing at the University of the Witwatersrand and has not been submitted for any degree at any other university or institution. The research report does not contain other persons' writing unless specifically acknowledged and referenced accordingly.

Ms Dineo Mokgehle (Student): …………………. Date: ……………......

Dr Cletah Shoko (Supervisor): ……………………. Date: ………………..

Dr Adam Elhadi (Co-Supervisor): …………………Date: …....................

Abstract

The accelerated invasion of mesquite (*Prosopis*) has brought negative socio-economic and ecological impacts, both at local and global scale. According to the International Union for conservation of Nature (2004), mesquite has been identified as one of the top 100 worst invasive species in the world. The lack of information regarding the spatial and temporal variability of mesquite invasion has compromised the implementation of monitoring and control efforts. Hence, the mapping and monitoring of mesquite is vital in order to obtain precise and up to date spatial and temporal data about its invasion dynamics.

This study focused on investigating the ability of Sentinel-2 data in mapping the current spatial distribution of mesquite invasion along the Al Gash River in Kassala, Sudan using Support Vector Machine and Random Forest classifiers. Sentinel-2 image, which covered the study area was obtained during the dry period (March). The utilisation of Random Forest classifier achieved an overall accuracy of 93.44%, whereas the Support vector machine classifier achieved an overall accuracy of 87.57%. Additionally, multitemporal Landsat earth observation data were used to monitor the spatio-temporal dynamics of the mesquite invasion over a period of 30 years from 1989 to 2019, with five years intervals. The change detection statistics depicted that the invasion of mesquite has increased over the years. The lowest mesquite areal coverage was found in 1989, with 13.7% (38 967.78ha) of the study area. However, mesquite rapidly increased to 34.46% (124 365ha) in 2014. The year 2019, witnessed a decline in mesquite coverage, covering 101 214.6ha (26.84% of the study area).

Overall, the study demonstrates the ability of Sentinel-2 to detect and discriminate mesquite from other land use and land cover types. The spatial and temporal variability of the new generation multispectral sensor enables continuous monitoring of the invasive species, through the provision of up-to-date and readily available data, at reasonable resolutions.

ACKNOWLEDGEMENTS

I would like to acknowledge the following people for their participation in the completion of this research report. I would like to thank Dr Cletah Shoko for her guidance as the Main Supervisor and Dr Elhadi Adam for his participation as a Co-Supervisor in the project.

I would like to honour the dedication and time invested in me by both the Main Supervisor and Co-Supervisor. It was through their guidance that I was able to collect and interpret data and give valid recommendations. Without their help throughout the year, it would be more difficult complete the project.

I would also like to acknowledge my Parents and Siblings for their involvement, divine encouragements and motivation to a successful completion of this assignment.

Lastly, I would also like to thank the Almighty for His guidance and being a pillar of strength.

Table of Contents

Declaration ... 2

Abstract ... 3

Acknowledgements ... 4

Table of Contents ... 5

List of Figures ... 8

List of Tables ... 9

CHAPTER 1: ... 11

1.1Introduction ... 11

1.2 Problem statement ... 13

1.3 Aim and objectives ... 15

1.4 Organisation of research ... 16

CHAPTER 2: ... 17

LITERATURE REVIEW ... 17

2. Introduction ... 17

2.1 History and phenology of mesquite ... 17

2.2 Impacts associated with mesquite ... 19

2.2.1 Positive impacts associated with mesquite ... 19

2.2.2 Negative impacts associated with mesquite ... 20

2.3 Mesquite management strategies ... 21

2.4 The application of remote sensing in mapping mesquite ... 24

2.4.1 Mapping invasive species through the use of multispectral data ... 26

2.4.2 Mapping invasive species through the use of hyperspectral data ... 27

2.4.3 Mapping invasive species through the use of new generation multispectral data ... 27

2.4.4 Machine learning algorithms in remote sensing of invasive species ... 28

CHAPTER 3: ... 31

RESEARCH METHODOLOGY 31

3 Materials and methods 31

3.1 Study site...31

3.2 Data acquisition and image pre-processing 32

3.3 Ground data..35

3.4 Image classification 38

3.5 Accuracy assessment 40

3.6 Change detection..40

CHAPTER 4:...41

RESULTS AND DISCUSSION 41

4.1 Mapping the current spatial distribution of mesquite along the AL Gash River using Sentinel-2 and Support vector and Random forest classifiers..41

4.1.1 Spatial distribution of mesquite 41

4.1.2 Areal coverage of Mesquite as obtained through Random Forest and other the LULC classes along the Al Gash River..43

4.1.2 Random Forest and Support vector optimization results..43

4.1.3 Accuracy assessment 46

4.2 Determining the historical extent of mesquite invasion along the AL Gash River, Kassala over the past 30 years (from 1989 to 2019) 49

4.2.1 Areal Coverage of mesquite and other LULCC along the Al Gash River over the past 30 years 51

4.2.2 Optimisation of Random Forest 53

4.2.3 Accuracy assessment 55

4.2.4 Change detection statistics of mesquite as detected by Landsat from 1989 to 2019 57

4.3 Discussion...57

4.3.1 The current spatial distribution of mesquite along the AL Gash River using Sentinel-2 and Support vector and Random forest classifiers...58

4.3.2 Historical spatial distribution of mesquite from 1989 to 2019 as detected by Landsat series data....62

CHAPTER 5:....66

Overall Conclusion....66

REFERENCES....69

List of Figures

Figure 1.1 Visual presentation of mesquite invasion along the Al Gash River (Nzumira, 2014) 14

Figure 2.1 Mesquite trees in Kassala, Sudan (Thorp *et al.*, 2000)....... 19

Figure 2 2 Mechanical control methods employed in Kassala through the use of bulldozers (Department of Natural Resources and Mines, 2003) 23

Figure 3.1 Location of the Al Gash River delta in Kassala, Sudan (Author, 2019). 32

Figure 4.1The spatial distribution of mesquite as mapped by Sentinel-2 and Random Forest (A) and Support Vector machine (B) classifiers 42

Figure 4.2 Areal coverage of mesquite and other LULCC along the Al Gash River in 2019 . 43

Figure 4.3 Random Forest optimization 44

Figure 4.4 Variable importance of the Sentinel-2 spectral bands in the classification of the various land cover classes 45

Figure 4.5 The mean decrease in accuracy for the various Sentinel-2 bands in classifying the different land cover classes. The highest mean decrease in accuracy depicts the most significant band 45

Figure 4.6 The mesquite invasion over the past 30 years by Al Gash River in Kassala, using Landsat imagery and Random Forest classifier from 1989 to 2019 at 5-year intervals 50

Figure 4.7 Area invaded by mesquite in comparison to other LULCC from 1989 to 2019 53

Figure 4.8 Optimization results for Random Forest classifier of parameters (mtry and ntree) for the Landsat imagery from 1989 to 2019 54

List of Tables

Table 2.1 Methods employed for controlling mesquite infestation 21

Table 2.2 Previous studies conducted with regards to the mapping and monitoring of mesquite 25

Table 3.1 Sentinel-2 data characteristics 33

Table 3.2 Landsat MSS, TM and OLI characteristics 35

Table 3.3 LULCC, training and test data for Sentinel-2 data 36

Table 3.4 LULC class characteristics identified for Landsat series data 37

Table 3.5 Training and test dataset for the Landsat series data 38

Table 4.1 Areal coverage of mesquite and other LULCC along the Al Gash River 43

Table 4.2 Confusion matrices generated from the Support Vector Machine classifier accuracy assessment...47

Table 4.3 Confusion matrices generated from the Random Forest classifier 48

Table 4.4 Areal Coverage of LULCC along the Al Gash River over the past 30 years 52

Table 4.5 LULC classification confusion matrices...56

Table 4.6 Change detection statistics for mesquite from 1989 to 2019...................................57

List of Abbreviations

ETM+	Enhanced Thematic Mapper Plus
FLAASH	Fast Line-of-sight Atmospheric Analysis of Spectral Hypercube
IUCN	International Union for Conservation of Nature
LULCC	Land Use Land Cover Class
MSS	Multispectral sensor
NIR	Near-infrared
OLI	Optical Land Imager
OOB	Out-of-bag
RF	Random Forest
RS	Remote Sensing
TM	Thematic Mapper
SWIR	Shortwave infrared
USGS	United States Geological Survey

CHAPTER 1:

1.1 Introduction

Mesquite (*Prosopis glandulosa*), which originated from the arid and semi-arid regions of South and North America, consists of over 44 sub-species worldwide (Nie *et al.*, 2012; Mureriwa *et al.*, 2016). Mesquite is classified as an evergreen invasive tree, with an open canopy and can grow to heights ranging from 5 to 10 m (Kazmi, 2009). It is fast-growing, perennial and a drought-resistant invasive species (Kazmi, 2009). Mesquite can also tolerate saline soils and can be planted within soils that have a high acid content. According to the International Union for conservation of Nature (2004), mesquite is considered to be one of the top 100 worst invasive species in the world. It also has various benefits such as fodder for animals, fuel, wood and shade for the people living in the local community (Alredaisy, 2013).

Mesquite usually spreads rapidly through open grazing, thus this is the most common way in which the seed is dispersed (Kazmi, 2009). Mesquite is an aggressive competitor for natural resources and has negative effects on the environment, affecting natural habitats such as coastal belts, riverbeds and rock mountaintops (Kazmi, 2009). In countries such as Australia, Pakistan and other Middle Eastern countries, it has been declared as a noxious weed (Kazmi, 2009). Over the years, due to factors such as the lack of proper management and monitoring schemes, these tree species have rapidly expanded and have invaded agricultural and residential areas within Sudan (Elhag et al., 2015). Mesquite was first introduced in the arid eastern region of Sudan in 1917 (Hoshino *et al.*, 2012). The species was primarily planted in Sudan to combat desert encroachment, desertification and land degradation. Mesquite has an extensive root system with the potential to grow under various environmental conditions (Hoshino *et al.*, 2012). The species aggressive appetite for water sources makes it an invasive species (Elhag et al., 2015). Over the past years the lack of knowledge and poor management of the species led to its invasion along water courses, irrigation channels and agricultural fields (Elhag *et al.*, 2015). The invasion of the species is difficult to control and manage effectively and has negatively affected Sudan's agricultural productivity (Nzumira, 2014). Over the years, there have been many forms of strategies that have been implemented to curb the spread of mesquite in Sudan. However, these strategies mainly used traditional (field-based) forms of obtaining information on the invasion dynamics of the species in order to monitor its spread (Nzumira, 2014). Field based techniques, include ground-based surveys which has suffered major setbacks. These setbacks include: high cost implications, time consumption, and smaller areal coverage, besides being tedious, whereas remote sensing is time and cost effective (Huang and Asner, 2009; Mureriwa *et al.*, 2016). Previous work shows that the detection, mapping and monitoring of invasive species using remote sensing is more efficient and spatially explicit (Huang and Asner, 2009). Remote sensing allows for the useful extraction of subtle plant information (Huang and Asner, 2009; Mureriwa *et al.*, 2016).

To the best of our knowledge, few studies have focused on the application of remote sensing in mapping and monitoring mesquite in sub-Saharan Africa, to date. For example, one study in Somalia was conducted in mesquite mapping using Landsat 8 OLI data and Random Forest classifier, with an overall accuracy of 85% (Meroni *et al.*, 2016). It was also noted that the invasion of mesquite was dominant along river channels, human settlements and agricultural fields (Meroni *et al.*, 2016). Another study conducted in Garissa, Kenya which focused on mapping the spatial distribution of mesquite using Landsat TM data between 2000 and 2006 and to propose a sustainable strategy to aid its management (Zeila, 2011). The results also showed that the areas most invaded by mesquite were agricultural fields and river courses. This has proven to be problematic due to the fact that most farmers in the region depend mainly on the River for irrigation and domestic use (Zeila, 2011). However, with both these studies, there are prevalent limitations that were noticed; satellites such as Landsat have medium spatial and spectral resolutions, which are often inadequate for the effective detection and mapping of invasive species (Adam *et al.*, 2017). Through the use of new generational multispectral sensors, the detection and mapping of mesquite can be effectively done (Adam *et al.*, 2017). For example, Sentinel-2 has spectral resolutions that ensure more vivid images with more information with regards to species mapping (Ramoelo *et al.*, 2015). The most important feature of Sentinel-2 is the inclusion of the red edge band which has proven to be essential in the classification of vegetation species (Ramoelo *et al.*, 2015).

1.2 Problem statement

Over half of the area along the Al Gash River in Kassala has been invaded by mesquite and is a health risk to the environment and population within the region (Elhag et al., 2015). The invasion of this species has brought about disruption in irrigation channels, as depicted in Figure 1.1, thereby affecting water availability, which in turn led to low crop production within the region. This has resulted in the impact on ecosystem services as well as the displacement of subsistence farmers and families (Alredaisy, 2013).

Figure 1.1 Visual presentation of mesquite invasion along the Al Gash River (Nzumira, 2014)

After most mechanical procedures were conducted to remove mesquite between the periods of 2001 to 2010, some control measures were applied to reduce its rapid expansion. These control measures included the Oxfam programme, which involved the removal of 13 thousand acres of mesquite invasion in 2002 and the New Halfa scheme that involved its eradication on farm lands between 2008 and 2010 (Nzumira, 2014). However, most of these control measures for mesquite invasion lacked up to date spatial and temporal information with regards to the key aspects associated with the dynamics of mesquite, such as the speed of the invasiveness, spatial distribution, environmental preference and the relations with hosting ecosystems (Babiker *et al.*, 2008).

Previous studies that were conducted by El Tayeb *et al.*, (2001) and Nzumira (2014) in the eastern regions of Sudan on mapping and monitoring of mesquite have mainly focused on the use of satellite imagery such as Landsat, which has medium to coarse spatial resolution. The studies also focused on a particular period, hence temporal and historical variability is not accounted (Babiker, 2006). The studies have also focused on the use of ground-based surveys of which cover small spatial areas, are tedious and financially straining. An example includes a study that was conducted by Nzumira (2014) in Kassala, Sudan where mesquite infestation in the Gash spate irrigation system was monitored using only Landsat imagery for the past 30 years as well as ground-based surveys. Results depicted that mesquite infestation has increased substantially and had invaded most agricultural regions over the years (Nzumira, 2014). The application of GIS and remote sensing in mapping and monitoring the spread and invasion within the Kassala region is important. This will enable future predictions of invasions or areas that are prone to future invasion, such that decision makers and environmentalists can develop strategies to resolve the issue (Joshi et al., 2004).

1.3 AIM AND OBJECTIVES

The aim of this research was to map and monitor the current and historical spatial distribution and invasion of mesquite over thirty years (1989 - 2019) along the Al Gash River Kassala, Sudan.

OBJECTIVES

The specific objectives of the research were to:

- To map current spatial distribution of Mesquite along the Al Gash River in Kassala using Sentinel-2 data
- Determine the historical spatial extent of mesquite invasion along the Al Gash River over the past 30 years (1989 - 2019) using multi-temporal Landsat data
- Test the performance of Support Vector Machine (SVM) and Random Forest (RF) in classifying and identifying mesquite species using Sentinel 2 data.

The research will be guided by the following questions:

1. How can the application of remote sensing in mapping the invasion of mesquite improve its monitoring along the AL Gash River?

2. How has the invasion and distribution of mesquite along the Al Gash River changed over space and time?

3. At what accuracy can the occurrence and distribution of mesquite detected and mapped using advanced machine learning and Sentinel-2 data, considering its spatial resolution?

1.4 Organisation of research

This research report consists of five chapters. The first chapter is the general introduction, which include the problem statement, study aim and specific objectives. The second chapter provides a review of literature, including the phenology and impacts of mesquite, as well as the potential of remote sensing and machine learning classifiers in its mapping. The third chapter provides the methodological approach, outlining the study site, datasets used and the classification approach, whereas the fourth chapter provides the results and discussion. Lastly, chapter five provides concluding remarks and further recommendations with regards to the research.

CHAPTER 2:

LITERATURE REVIEW

2. Introduction

Invasive alien species have proven to be a significant problem amongst farmers, governments, ecologists and conservationists in various parts of the world (Joshi *et al.*, 2004; Mureriwa *et al.*, 2016). Invasive plants (namely non-indigenous or non-native species) have undeniably shown that the plants are a threat to livelihoods and biodiversity, besides being considered as a major non-climatic catalyst of global environmental change (Huang and Asner, 2009). Throughout the years, numerous strategies and methods have been implemented to curb the growth and infestation of invasive species within communities (Lowe *et al.*, 2000; Mureriwa, 2017). These actions have all been implemented to keep communities knowledgeable about the dynamics of invasive species and also devise efficient and effective ways to prevent the rate of invasion (Robinson *et al.*, 2016). The detection, mapping and monitoring of invasive species is best conducted through the utilisation of remote sensing (Huang and Asner, 2009). Remote sensing allows for the useful extraction of information without direct physical contact with the invasive species (Huang and Asner, 2009; Mureriwa *et al.*, 2016). Traditional methods of mapping vegetation cover suffer from major setbacks such as time consumption and are very tedious while remote sensing is time efficient and cost effective.

2.1 History and phenology of mesquite

Mesquite is an invasive weed, which originated from Central America has spread onto regions of Panama, Caribbean Islands, Venezuela and the northern parts of Peru (Nie *et al.*, 2012, Mureriwa *et al.*, 2016). Mesquite has spread worldwide over the last 200 years. Currently, it is in many climate zones, as well as in dry and semi dry regions of southern America. Mesquite was first introduced in Senegal in 1822 and in South Africa around 1880 (Mohamed, 2001). Mesquite was first planted in Khartoum (Sudan) in the year 1917. It was primarily planted in Sudan as a way of combating desert encroachment (Elhag et al., 2015). Given its paramount success, the plant was later introduced across Sudan namely in Sinar, Fwar and EL foung. Both residential and vegetated areas were restored in eastern Sudan, during the drought of the 1970s (Mohamed, 2001). Mesquite was later introduced into western and central Sudan namely the White Nile province (Steenbergen, 2014). It was planted in irrigated regions and also in shelter beds around farms along the Nile River. Mesquite has been quite a volatile weed in its spreading and affecting neglected areas, irrigated lands, roads, floodplains and watercourses (Babiker, 2006).

However, over the years, mesquite has proved to be more of a problem than a solution in the eastern and northern regions of Sudan. Mesquite prospers on alluvial soils, which have significantly good water holding capacity; their infestation spread to 230 thousand hectares of land (Nzumira, 2014). Mesquite is also found in Gash delta, Kassala, Wagali and areas of Eritrea in the Atbara River (Steenbergen, 2014). This infestation has spread up to 130 kilometres upstream and was found along the Portsudan highway and the Kassala-Gadarif region (Steenbergen, 2014).

Mesquite (*prosopis glandulosa*) is one of the 44 various species that are part of the Prosopis genus and is known to grow to heights of about 14 meters (Mohamed, 2001). Mesquite is a single-stemmed tree with an extensive root system, growing as deep as 50 cm and an average lifespan of 33-44 years (Mohamed, 2001). The species is classified as an evergreen tree with an open canopy of which can grow to heights ranging from 5 to 10 m as shown in Figure 2.1 (Mwania, 2017).

Figure 2.1 Mesquite trees in Kassala, Sudan (Thorp *et al.*, 2000)

The tree is also known as a multi stemmed acacia-like shrub that is comprised of dark green leaves with reddish-brown branches and axial (Babiker *et al.*, 2008). In Pakistan and most countries in the Middle East, mesquite is usually identified as a tree that has an appearance of a shrub (Kazmi, 2009). Mesquite produces fruit in a form of hanging pods, this usually happens from March to June. These species are well-known to thrive in almost all kinds of weather conditions such as excessive heat and low rainfall conditions. It is fast growing, perennial and drought-resistant (Nzumira, 2014). Mesquite can also tolerate saline soils and can survive in high acidic soils. Hence, Mesquite can thrive in infertile soil and adapts to both arid and semi-arid weather conditions. According to the IUCN (2004) mesquite has been considered to be part of the top 100 worst invasive species in the world.

2.2 Impacts associated with mesquite

2.2.1 Positive impacts associated with mesquite

Mesquite has multiple purposes that are beneficial in providing ecosystem goods and services (Mwangi and Swallow, 2005). According to Kazmi (2009), mesquite is a very important plant for nitrogen-fixation within semi-arid and arid regions of the world. It can also be useful as a source of food, fuel for wood, shade, and a solution to stop or prevent desert encroachment from occurring quite often (Laxen, 2005; Kazmi, 2009). The plant has beneficial pharmacological applications and is a potential source of income to small local communities (Laxen, 2005). Mesquite has been reported as an economic resource in many developing countries such as Pakistan and Kenya and plays a key role in the promotion of afforestation in arid regions (Mwangi and Swallow, 2005).

2.2.2 Negative impacts associated with mesquite

There are negative impacts associated with this species. Over the years, due to factors such as the lack of proper management and monitoring schemes, these tree species have rapidly expanded affecting the environment and the wellbeing of various societies. For example, they have invaded agricultural and residential areas within Kassala (Elhag *et al*., 2015). The invasive is water efficient and very competitive. In countries such as Australia, Pakistan and other Middle Eastern countries, it has been declared as a noxious weed (Kazmi, 2009). It is now estimated that most agricultural regions in Kassala have been invaded by mesquite. Mesquite has posed as a health risk to the population within the region (Elhag et al., 2015). The invasion of mesquite has brought about the prevalence of malaria. It has also caused infection and septicema through injuries by impenetrable thickets formed by the specie in Kenya, Eritrea and Sudan (Hussain *et* al., 2020). The invasion of this specie has brought about the disruption in irrigation channels by invading water courses and consuming water content, thus affecting water availability and has negatively affected crop production within the region. This has resulted in the impact on food security within the region as well as the displacement of families and subsistence farmers (Alredaisy, 2013). Mesquite also has a negative effect on the environment and has also affected habitats such as coastal belts, river beds and rock mountain-tops (Kazmi, 2009).

2.3 Mesquite management strategies

Numerous mesquite eradication and control methods have been applied but with limited success due to the lack of knowledge on the spatial and temporal dynamics of the species within the area (Pasiecznik *et al.*, 2001; Mureriwa *et al.*, 2016). Countries, such as Australia, with more knowledge have been deemed to be more successful in the implementation of various control measures (Van den Berg., 2010). Hence, the best strategy in mesquite eradication is to place focus on attaining knowledge with regards to the dynamics of the propagation of the species (Nzumira, 2014). Control methods are grouped into three categories of which are mechanical, chemical and biological as indicated in table 2.1 (Pasiecznik *et al*., 2001).

Table 2.1 Methods employed for controlling mesquite infestation

Control method	How the method works or its application
Mechanical method	• Mainly focused on the physical removal of mesquite, such as hand grubbing, tree cutting, bulldozing, excavating and the uprooting (Shackleton et al., 2014b). For example, in Sudan the New Halfa and Mesquite tree infestation programmes both employed mechanical methods such as the use of bulldozers, hand cutting and excavators in the removal of mesquite throughout the Gash Agricultural scheme (Steenbergen, 2014). • Has socio-economic benefits such as employment locally and promotes job creation. • Labour intensive, costly and its limited to small spatial coverage
Chemical method	• Involves use of chemicals e.g. spraying specific chemicals on to the invasive species or the use of chemicals after the chopping down of invasive trees (Mureriwa, 2016). • Literature on its success and limitations is limited especially on the specific type of method • Have numerous limitations such as the need for skilled personnel, environmental and health related effects on people (Van den Berg, 2010).
Biological method	• It is organic as it involves the use of other species to get rid of invasive species such as bugs, beetles and other forms of insects. • More sustainable approach and is environmentally friendly. • It does not completely eradicate e.g. has been employed in Kassala through the use of Algarobus prosopis which is an insect that destroys mesquite pods and seeds (Nzumira, 2016) • It is dependent of factors such as climate and the availability of water

Some of the physical, biological and mechanical efforts in eradicating the mesquite have been initiated in various countries. However, in developing countries such as Paraguay, Argentina and South Africa, the efforts have been less intensive (Kazmi, 2009). In Pakistan, the most common method of control is hand-grubbing and is considered as the most effective measure of control (Kazmi, 2009). In North Sudan, the most common control methods are mechanical. There have been projects in the past that have focused on eradicating mesquite, for example, the Oxfam initiatives adopted the use of hand-axes to remove 13 thousand acres of mesquite in Tokar, Sudan (Elhag *et al.*, 2015). The New Halfa project initiated by the Sudanese government and other private firms is also still underway in Kassala, Al Gash which is aimed at controlling mesquite population within the region (Steenbergen, 2014). However, these mechanical methods as depicted in Figure 2.2, have poorly dealt with the problem at hand due to the lack of follow ups conducted during the progress of these programmes, as well as limited spatial coverage, lack of resources and knowledge on the characteristics and spatial dynamics of the invasive species (Babiker *et al.*, 2008).

Figure 2.2 Mechanical control methods employed in Kassala through the use of bulldozers (Department of Natural Resources and Mines (Sudan), 2003)

The evaluation of the specie has been found to be a problem for most decision makers in most sub-Saharan countries, because the measures to control mesquite can only be realised when environmental conditions that favour its success in arid and semi-arid regions are better understood (Nzumira, 2014). Formulating management strategies that are aimed at preserving benefits and minimizing costs associated with controlling the plant need to be properly understood.

2.4 The application of remote sensing in mapping mesquite

Ground based survey is the traditional method for understanding the distribution of mesquite (Adam *et al.*, 2017). However, for the assessment of larger areas, this method is less effective, time consuming, costly, very limiting and labour intensive (Immeitzer *et al.*, 2012). Remote sensing data from multispectral and new generation multispectral sensors can cover large areas and is cost and time efficient and are commonly used for mapping invasive species (Immeitzer e*t al.*, 2012). Remote sensing can provide integrated and detailed information at different resolutions and can allow for the retrieval of information of the spatial dynamics of mesquite (Gomez, 2017). Remote sensing provides the flexibility for monitoring purposes due to its temporal and spatial dynamics. For example, Landsat data can be used for change detection analysis to assess the changes in various infestation rates of invasive species over time which is good for monitoring purposes (Van den Berg *et al*., 2010). In addition, Landsat data can be obtained and utilised freely by the public and can be used for various applications such as change detection analysis (Gomez, 2017). Through the advancements in the field of remote sensing, the utilisation of remote sensing as a tool in the monitoring and mapping of mesquite is gradually increasing globally (Keebine, 2019). There are various developed countries that are utilising remote sensing in mapping mesquite (Babiker, 2006). Table 2.2 shows some of the studies conducted with the application of remote sensing

Table 2.2 Previous studies conducted with regards to the mapping and monitoring of mesquite

Author	Country	Data	Results	Limitations
Nzumira (2014)	Sudan	Landsat	1. An overall accuracy of 76% was achieved 2. Mesquite heavily invaded the inland Gash delta from 1979 to 2013 at a rapid rate, mainly due to animal grazing and anthropogenic processes	1. Low to coarse spectral and spatial resolutions articulated to the data used 2. Misclassifications due to thee spectral variations within the classes
Zeila (2011)	Garissa, Kenya	Landsat TM	1. About 440 square kilometres (33%) of the land within the study region was invaded by mesquite from 2000 to 2006 2. Most dominant area where mesquite was dominant was riverine which proved to be problematic because most farmers depend on the river Tana for irrigation and farming purposes	1. Low to coarse spectral and spatial resolutions articulated to the data used
Van den Berg *et al.* (2010)	South Africa	NOAA, Landsat, MODIS, SPOT-5	1. Achieved an overall accuracy of 72% 2. The prediction of future regions where possible mesquite invasion could occur was achieved 3. The relationship between future mesquite invasion and habitat was highlighted	1. More accurate mapping of the invasive species was required 2. The differentiation between indigenous and alien vegetation within mixed stands
Ansley and Mirik, (2012)	North America	Aerial photography, Landsat	1. Achieved overall accuracies ranging from 95.13% to 97.40% at three different sites	1. Low spatial and spectral resolutions 2. Misclassifications due to thee spectral variations within the classes
Ali (2008)	Sudan	MODIS	1. Achieved an overall survey response rate of 96% with regards to perceptions stemmed from mesquite invasion within communities as well as the control methods 2. Using MODIS Normalized Difference Vegetation Index data from 2000 to 2007, mesquite increased every year by 228.45 hectares per year	1. Low spatial and spectral resolutions
Robinson *et al.*, (2016)	Australia	WorldView-2	1. Achieved an overall accuracy of 88% 2. Red edge and NIR band combinations were successful in the detection of mesquite in the study region	1. The study was only able to distinguish two types of plant and background soil
Adam *et al.*, 2017	South Africa	WorldView-2	3. Random Forest classifier achieved a higher overall accuracy (86.59%) as compared to Support vector machine classifier (85.98%)	1. Mapping vegetation cover only at species level 2.Misclassifications due to thee spectral variations within the classes

However, in the Sub-Saharan region, less research work on mesquite invasion mapping and monitoring through the application of remote sensing technologies has been done (Mureriwa *et al.*, 2016). Adequate management of mesquite infestation requires up-to-date data regarding the invasive species' temporal and spatial variability. Through the use of remote sensing, the provision of effective and economically sustainable methods that have the capabilities to produce timely and accurate information for mapping vegetation species is possible (Adam *et al.*, 2017). There are factors that favour the use of remote sensing in the mapping of vegetation species such as the large aerial coverage, especially with areas that are inaccessible through ground-based surveys (Hoshino *et* al., 2012).

2.4.1 Mapping invasive species through the use of multispectral data

Previous studies have often preferred the use of multispectral satellites such as Satellite Por I'Obersation de la Terre (SPOT) and Landsat to obtain higher accuracies in the mapping and monitoring of invasive species (Shekede *et al., 2008;* Dube et al., 2017; Keebine, 2019). SPOT and Landsat data are freely obtainable and have medium temporal and spatial resolutions (Dube et al., 2017; Kganayago *et al.*, 2018). The outlined characteristics are some of the main reasons that have propagated the use of these sensors. However, attributed to these characteristics are limitations such as the medium spatial and spectral resolutions, which are inadequate for the effective detection and mapping of vegetation (Adam *et al.*, 2017). Mohammed *et al.*, (2013) evaluated the effects of mesquite tree invasion in the Gash Agricultural scheme in Kassala, Sudan from 1979 to 2010. The study used Landsat MSS, TM and ETM+ to compute change detection statistics, as well as to map the spatial distribution in 2010. The highest overall accuracy achieved was 76% from the 2010 image. The change detection results suggested that mesquite tree invasion was evident in the region and was rapidly increasing over the years Mohammed *et al.*, (2013). Through the use of Landsat data, mesquite was mainly detected at river banks, streams and irrigated crop lands. A total of 141 942 hectares of land has been invaded from 1979 to 2010 as reported by Mohammed *et al.*, (2013).

2.4.2 Mapping invasive species through the use of hyperspectral data

The use of hyperspectral sensors has advantages such as the provision of finer results in differentiating spectral features among invasive species (Raczko and Zayajewski, 2017). Hyperspectral data also contains various spectral bands that go beyond the infrared region of the electromagnetic spectrum (Foody, 2002). Hyperspectral data has a finer resolution and can be ideal for mapping and classifying invasive species. However, the disadvantage associated with the hyperspectral system is that it is costly and limited in terms of the availability of data (Asner *et al.*, 2008; Razko and Zayajewski, 2017). This makes it difficult to use them for continuous monitoring of invasive species dynamics over large areas.

2.4.3 Mapping invasive species through the use of new generation multispectral data

According to Adam *et al.* (2017), from the last decade, the emergence of new generation sensors with finer spectral and spatial resolutions has come about. Sensors such as RapidEye, Sentinel-2, Landsat 8 OLI and WorldView-2 are examples of new generation multispectral sensors with spatial resolutions, ranging from 1.5 to 60 m (Adam *et al.*, 2017). These improved spatial and spectral resolutions of new generation sensors ensure more vivid images with more information with regards to land cover mapping. Mureriwa *et al* (2016) conducted a study which focused on mapping mesquite in Northern Cape, South Africa through the use of high spatial resolution WorldView-2 imagery. The main aim of the study was to explore the capability of utilising high spatial resolution imagery in the mapping and distinguishing of mesquite from other co-existing species using two machine learning classifiers. The results indicated that the 8 band WorldView-2 imagery was able to detect and differentiate mesquite from its other co-existing species, with an overall accuracy of 86% (Mureriwa *et al.*, 2016). Higher accuracies facilitate the provision of economically sustainable mapping and monitoring of the spatial distribution of invasive species, as well as aids in the development of more efficient control methods (Mureriwa *et al.*, 2016). Another similar case study was done by Ng *et al* (2017), investigating the potential of Sentinel-2 and Pleiades data in the detection and mapping of mesquite and vachellia in Kenya. The main aim of the study was to test the ability of the random forest classifier to distinguish and discriminate mesquite from other co-existing invasive species using Sentinel-2 imagery (Ng *et al*., 2017). The overall accuracy achieved for the mapping and detection of mesquite was 79%. The results portrayed that it was possible to distinguish and discriminate mesquite from the other LULCC through the use of the Random Forest classifier, high spatial resolution Sentinel-2 imagery and the inclusion of the vegetation red edge band (Ng *et al*., 2017).

2.4.4 Machine learning algorithms in remote sensing of invasive species

In most recent cases, through the advancements in technology, intelligent systems embedded within computers use learning procedures to solve environmental issues with regards to various types of data. These learning procedures imitate some aspects of the human mind in order to resolve complex predicaments (Chauhary *et al.*, 2013). Hence machine learning is based on the enabling of computers to adapt and modify human actions, such that the actions can improve the accuracy and efficiency of the decisions, which are drawn by the algorithms (Chaudhary, *et al.*, 2013; Marsland, 2015). Classification is a problem of machine learning, especially on how to allocate labels to data based mainly on a given set of labels. The procedures of classification involve predicting some outcome based on a given input. In order to predict the outcome, the algorithm processes a training set that contains a set of attributes and the respective outcome, commonly referred to as the objective or prediction attribute. The classification algorithm assigns categories or classes of interest to pixels in the image. There are two types of classification procedures or approaches of which are supervised and unsupervised (Das, 2017). Supervised classification involves the utilisation of spectral bands or signatures obtained from training samples in order to classify an image. In supervised classification, there are various algorithms which include traditional classifiers and new machine learning algorithms such as Maximum Likelihood, Minimum distance, ISO Cluster, Principal Component Analysis, Random and Support Vector (Das, 2017).

The use of machine learning algorithms in mapping invasive species through the use of remote sensing has been applied in many studies. Machine learning algorithms use spectral signatures obtained from ground data (training data) to group pixels with similar digital numbers in order to formulate a single land cover class (Lillesand *et al.*, 2015). The spectral properties of the classes formed from the previous process distinguish the various LULCC from each other. Through this process it easier to detect and map the rate of invasion or the spatial distribution of particular invasive species. One study that compared the two machine learning algorithms (SVM and Random Forest) was conducted in North West, South Africa by Mureriwa (2016). The study aimed to map mesquite and its co-existing species through Worldview-2 imagery, SVM and Random Forest. The results depicted that both algorithms were able to map the spatial distribution of mesquite in the region, however it was observed that the Random Forest classifier was best at detecting mesquite due to the higher accuracy yielded.

2.5 Conclusion

The application of GIS and remote sensing in mapping and monitoring the spread and invasion along the Al Gash River in Kassala remains important. The effective management of mesquite requires updated spatial and temporal information with regards to the spatial distribution and current invasion as well as its impacts of the ecosystem services provided in the region (Joshi *et al.*, 2004). It was evident that the majority of studies done concerning the use of remote sensing have focused on a particular time period and spatial coverage. Others have also been limited to the use of multispectral sensors such as Landsat to map and detect invasive species in small regions whereas other studies have used commercial satellites and aerial photographs which were not readily available for large area mapping. There are however still critical gaps with regards to the lack of knowledge about the spatial and temporal dynamics of mesquite invasion that still exist. Consequently, the use of sensors with finer spatial and spectral resolutions is readily available and should be used in mapping mesquite to obtain higher accuracies. Furthermore, the use and capability of Senitnel-2 in the detection and mapping of mesquite along the Al Gash River in Kassala has not been explored to date. There is therefore a need to embrace new crop of sensors with improved sensing characteristics like Sentinel 2 data in the detection, distinguishing and mapping of mesquite in Kassala, Sudan. Such findings can aid in illustrating the current spatial extent of this invasive species. Additionally, the use of change detection methods in mapping the historical spatial extent of mesquite in the region through the use of Landsat MSS, TM and OLI data remains important given its long history in spatial data acquisition. This will aid local authorities and communities in comprehending the dynamics of what propagates the spread of the invasive species and the major challenges experienced with regards to controlling the management of the invasive species.

CHAPTER 3:

RESEARCH METHODOLOGY

3 Materials and methods

3.1 Study site

Kassala is a town located in the north eastern part of Sudan, near the Eritrean border. The study region lies between latitudes (13° 30' -15° 00') N and longitudes (32° 30'- 36°00') E. The town covers an area of 341 000 square kilometres. The process of desertification, due to the on-going drought has increased rapidly (Nzumira, 2014). Kassala is classified as a region of drought and desertification. The area has a hot dessert climate and receives between 50 and 200 mm of rainfall annually. The average rainfall also hardly exceeds 400 mm. Most rainfall in the area falls between June and September which is classified as the summer period in the North. The greater part of Kassala is semi-arid and its winters are virtually dry and warm with no rainfall (Ali, 2008). (Hellden, 1984; Nzumira, 2014). The region consists of the inland delta where the River Al Gash is. The river Al Gas is considered to be one of Sudan's most prolific alluvial basins (Nzumira, 2014). The river is a sporadic stream that starts in the highlands of Eritrea. It is the area's main agricultural source of irrigation (Seed, 1969). The Al Gash River is a main source for the Gash Agricultural scheme, which constitutes the region's common land use; which is agriculture (Nzumira, 2014). The Gash irrigation scheme covers a total area of 100 800 ha and was initiated by the Sudanese government in the 1920s as a panacea to aid poor farmers (Nzumira, 2014). Kassala, has declined as a centre of cotton production and livestock farming, due to the invasion of mesquite in the scheme from the 1970's. This has exerted pressure on the natural resources in the scheme. Subsistence farmers in the region rely on the small plots of arable land left. One of the main land use activities is agriculture which constitutes subsistence and commercial farmers (Nzumira, 2014). Figure 3.1 shows the map of the study site.

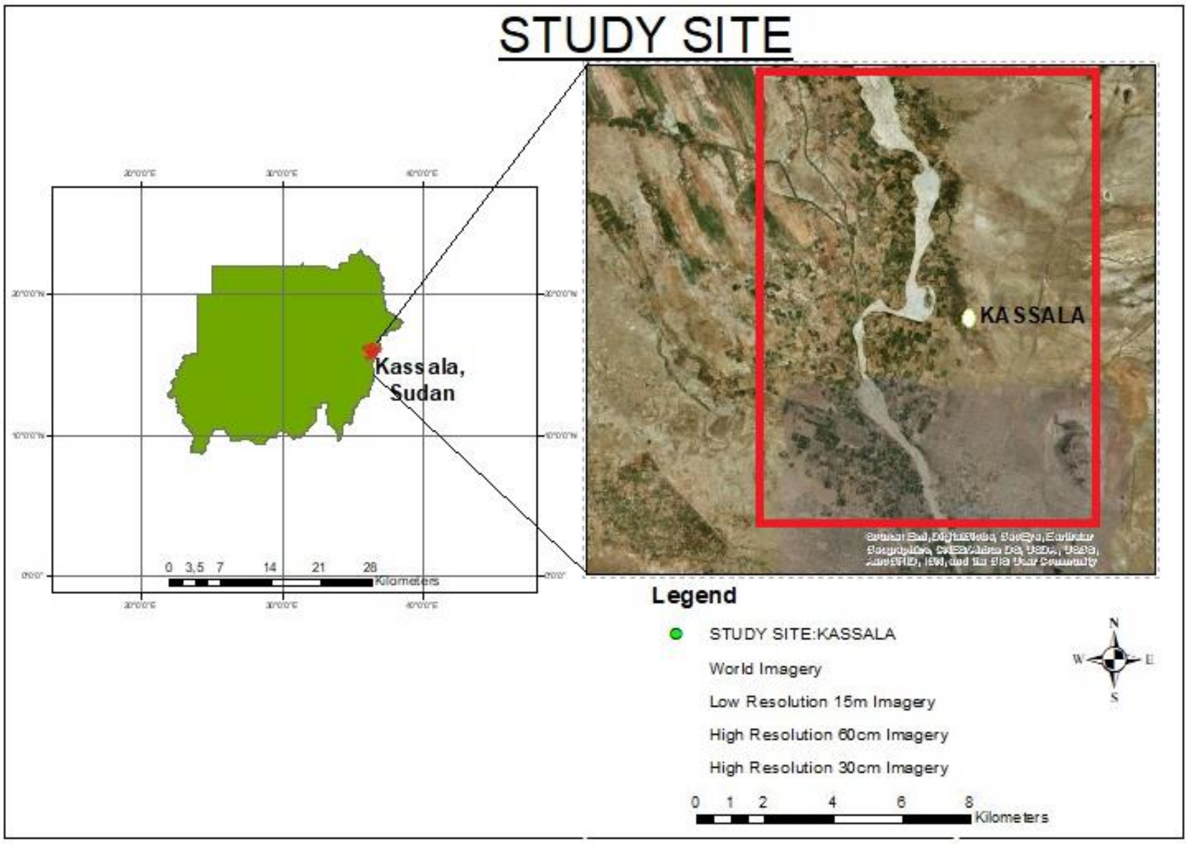

Figure 3.1 Location of the Al Gash River delta in Kassala, Sudan

3.2 Data acquisition and image pre-processing

Sentinel-2 MSI data

Sentinel-2 imagery for March 2019 was obtained from the European Space Agency Copernicus Access Hub (ESA) site to map the current distribution of mesquite. The selection of the dates of acquisition was based on the fact that March lies within the dry season in Sudan, hence it is easier to detect and assess mesquite due to its spectral reflectance (Ali, 2008). This also improves accuracy of the image classification. The Sentinel-2 data was selected based on the suitable processing levels, easy accessibility and low cloud coverage. Sentinel-2A and B have brought a new dimension in terms of obtaining freely accessible remote sensed data (European Space Agency, 2015). The satellite has finer temporal and spatial resolution and has more bands (13) as compared to Landsat. It has a revisit time of 5-10 days and has spatial resolutions of 10, 20 and 60 metres. The sensor also contains a very important part of the electromagnetic spectrum of which is the red-edge band, which is useful for the classification of invasive species (Gomez, 2017). Table 3.1 below shows the various spectral bands and resolution of the sensor.

Table 3.1 Sentinel-2 data characteristics

Spectral Bands	Wavelength(μm)	Spatial Resolution (m)
Band 1- Coastal aerosol	0.443	60
Band 2- Blue	0.490	10
Band 3- Green	0.560	10
Band 4- Red	0.665	10
Band 5- Vegetation red edge	0.705	20
Band 6- Vegetation red edge	0.740	20
Band 7- Vegetation red edge	0.783	20
Band 8- NIR	0.842	10
Band 8A- Vegetation red edge	0.865	20
Band 9- Water vapour	0.945	60
Band 10- SWIR-Cirrus	1.375	60
Band 11- SWIR	1.610	20
Band 12- SWIR	2.190	20

The Sentinel Application platform software was utilised to perform atmospheric correction and geo-referencing of the Sentinel-2 image, through the use of the Sen2Cor plugin found within the software. This plugin reduced all atmospheric effects which were present in the image. Thereafter, the image was resampled to a spatial resolution of 10 m in order for the spatial distribution of mesquite to be captured through the use of the spectral subset tool. This then resulted in the corrected image consisting only of ten spectral bands (bands 2, 3, 4, 5, 6, 7, 8A, 11 and 12) of the original thirteen bands. The image was then converted into a readable ENVI format, which led to the separation of the ten bands. Through the use of the image stacking plugin in the ENVI software, the bands were placed in chronological order into one image.

Landsat series data

The study made use of the Landsat MSS, TM and OLI imagery. Landsat imagery was obtained for the years; 1989, 1994, 1999, 2004, 2014 and 2019 from the USGS earth explorer portal. March was a suitable month for data acquisition. The selection of the dates of acquisition was based on the fact that March lies within the dry season in Sudan, hence it is easier to detect and assess mesquite due to its spectral reflectance (Ali, 2008). This also improves accuracy of the image classification. Landsat has a revisit time of 16 days and a spatial resolution of 30 m (Franklin and Wulder, 2002). Radiometric and atmospheric corrections were then performed using the FLAASH model embedded in the ENVI 5.4 software to reduce atmospheric effects, such as haze and noise (Lu *et al*., 2004). The availability of the historical data from Landsat missions make it more suitable for change detection of mesquite invasion. Table 3.2 below shows the Landsat sensors characteristics

Table 3.2 Landsat MSS, TM and OLI characteristics

Sensor	Spectral Bands	Wavelength(µm)	Resolution(m)
Landsat MSS	Band 1(Green)	0.53-0.60	60
	Band 2(Red)	0.63-0.69	60
	Band 3(NIR)	0.76-0.90	60
	Band 4(NIR)	2.08-2.35	60
Landsat TM	Band 1(Blue)	0.45-0.52	30
	Band 2(Green)	0.52-0.61	30
	Band 3(Red)	0.63-0.69	30
	Band 4(NIR)	0.76-0.90	30
	Band 5(SWIR)	1.55-1.75	30
	Band 6(Thermal infrared)	10.4-12.5	120
	Band 7(SWIR)	2.08-2.35	30
Landsat OLI	Band 1(Coastal aerosol)	0.43-0.45	30
	Band 2(Blue)	0.45-0.51	30
	Band 3(Green)	0.53-0.59	30
	Band 4(Red)	0.64-0.67	30
	Band 5(NIR)	0.85-0.88	30
	Band 6(SWIR-1)	1.57-1.65	30
	Band 7(SWIR-2)	2.11-2.29	30
	Band 8(Panchromatic)	0.50-0.68	15
	Band 9(Cirrus)	1.36-1.38	30
	Band 10(TIRS 1)	10.6-11.19	100
	Band 11(TIRS 2)	11.5-12.51	100

3.3 Ground data

Ground samples were obtained for the various LULCC by randomly creating ground points for the area under study. The ground points that were obtained had to be within a pixel to prevent the confusion of spectral properties from various LULCCs. The various LULCC in the region were defined through the use of the Sentinel-2 and the Landsat series images. A total of five different land cover classes were identified in the area. The identified land cover classes included; water features, bare land, built up area, mesquite and other vegetation as depicted in tables 3.3 and 3.4 Larger samples of points were obtained in order to achieve high classification accuracies (Mesv, 2010; Keebine, 2019). The ground truth collection were then split into 30% test and 70% training data through the use of the R programming software as depicted in tables 3.3 and 3.5. The training data was utilised to classify the images and the accuracy of the whole classification through the test datasets. The tables (3.3 to 3.5) below depict the various LULCC that were identified together with the training and test datasets.

Table 3.3 LULCC, training and test data for Sentinel-2 data

Land cover classes	Code	Characteristics	Training	Test
Water features	WF	Rivers, streams and dams	105	44
Bare land	B	Exposed land and rock with the absence of vegetation	71	71
Built up area	BA	Industrial, residential and commercial structures	252	107
Mesquite	M	Only mesquite	141	60
Other vegetation	V	Orange groves, acacia, agricultural crops, grassland and trees	84	84

Table 3.4 LULC class characteristics identified for Landsat series data

Land cover classes	Code	Characteristics
Water features	WF	Rivers, streams and dams
Bare land	B	Exposed land and rock with the absence of vegetation
Built_up_Area	BA	Industrial, residential and commercial structures
Mesquite	M	Only mesquite
Other Vegetation	V	Orange groves, acacia, agricultural crops, grassland and trees

Table 3.5 Training and test dataset for the Landsat series data

LULCC	1989		1994		1999		2004		2014		2019	
	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
Water features	105	45	131	56	155	67	131	56	210	90	105	44
Bare land	118	51	156	67	214	92	183	79	101	44	166	71
Mesquite	218	94	151	65	170	73	148	64	188	80	252	107
Other Vegetation	140	60	132	56	185	79	113	48	139	59	141	60
Built up area	165	71	187	80	118	50	210	90	176	76	199	84

3.4 Image classification

Supervised classification was performed using the Support vector machine and Random Forest classifiers. Support Vector Machine is a non-parametric supervised machine learning algorithm (Gomez, 2017). This algorithm was developed by Vapnik in 1979 (Mureriwa *et al.*, 2016). The Support Vector machine classifier is a robust classifier. The algorithm is measured by an effective hyper plane probing method where the minimal training region utilised takes less time for processing (Cortes and Vapnik, 1995). It rids problems such as over fitting and needs no notion of a data type. It minimises misclassification and develops boundaries by dividing hyper planes (Vapnik, 1979; Gomez, 2017). This algorithm is a linear classifier and is meant to maximise the distance for each class from the data points in the optimal hyper plane or decision boundary to the training data (Vapnik, 1979; Gomez, 2017). The boundaries are comprised of two support hyper planes that require data points on their edges named support vectors that describe the optimal hyper plane (Mountrakis *et al.*, 2011). It mainly enhances the non-linear procedure through optimisation by the use of a number of methods such as the kernel method of which is a binary linear classifier that utilises radial basis (Vapnik, 1979; Gomez, 2017). The two parameters utilised for tuning are called sigma (*C*) and gamma (γ) (Karatzoglou *et al.*, 2006).

The Random Forest machine algorithm was first introduced by Breiman (2001) as a non-parametric algorithm whereby it improves the accuracy of the classification process and regression trees (Adam *et al*., 2017). The main principle of Random Forest is based on the implementation of a bagging operation whereby several trees are joined and each tree adds a vote to assign a class to a pixel throughout the whole classification process (Breiman, 2001). On the input data bootstrap samples, multiple classification trees (ntree) are created. On a bootstrap model, every single classification tree grows; this is two-thirds of the input data referred to as "in-bag" data (Breiman, 2001). The remaining one-third of the input data omitted from the bootstrap study is known as the OOB samples "out - of-bag." Out-of bag samples are used to calculate the function of each variable in the model's final classification and to estimate the model's misclassification error (Breiman, 2001; Adam et al., 2017). The ensemble then splits the trees randomly into many nodes using random subsets of predictive variables (m*try).* A grid search approach based on the OOB error estimate was used to calculate the optimum combination of m*try* and n*tree* parameters (Breiman, 2001). The study employed the use of the Random Forest and Support Vector machine learning algorithms library of the R studio statistical packages. A tenfold grid approach was utilised based on the OOB error estimate in order to find the optimum combination of n*tree* and m*try* parameters (Breiman, 2001).

3.5 Accuracy assessment

According to Foody (2002), accuracy assessment is defined as the accuracy of a classification or map which provides an impartial representation of the land cover of the region it portrays. Accuracy assessment quantitatively compares satellite images and reference maps of spatial information. The independent test data which constituted 30% of the data was used for the accuracy assessment. Confusion matrices were generated to formulate a comparison between the actual class and the class assigned by the classifier. The Kappa coefficient was also generated, which ranges from 0 to 1 and provides the difference between the actual agreements. Thereafter the overall accuracies were then calculated.

3.6 Change detection

Change detection analysis was performed to determine the historical spatial extent and invasion of mesquite in Kassala for the past 30 years, using the Landsat imagery at 5 year intervals. The change detection statistics were obtained through the use of the post classification algorithm in ENVI 5.4 software (Hegazy and Kaloop, 2015). Hence the change detection statistics were used in this research. The change detection statistics require two input images. The change detection statistics depicts differences in LULC classes over the years. Six image pairs were used which included the following: from 1989 to 1994, from 1995 to 1999, from 1999 to 2004, from 2004 to 2009, from 2009 to 2014 and 2014 to 2019.

CHAPTER 4:

RESULTS AND DISCUSSION

4.1 Mapping the current spatial distribution of mesquite along the AL Gash River using Sentinel-2 and Support vector machine and Random forest classifiers

4.1.1 Spatial distribution of mesquite

The classification of the Sentinel-2 image using the Random Forest and Support vector machine classifiers was conducted to distinguish mesquite from the other LULCC along the Al Gash River. The spatial extent of mesquite in the area was broad and accounted for 20.09% of the total areal coverage which is the second biggest land cover class. Other Vegetation covered a total of 12.42% of the region. The Support Vector and Random Forest machine learning algorithms were able to map the spatial distribution of mesquite and other various LULCC in the region as presented in Figure 4.1. Upon observation from both classifiers, mesquite invasion is prevalent in the study area and is dominant in the north western and western regions. Support vector machine also showed more mesquite coverage, especially in the eastern parts, as compared to random forest.

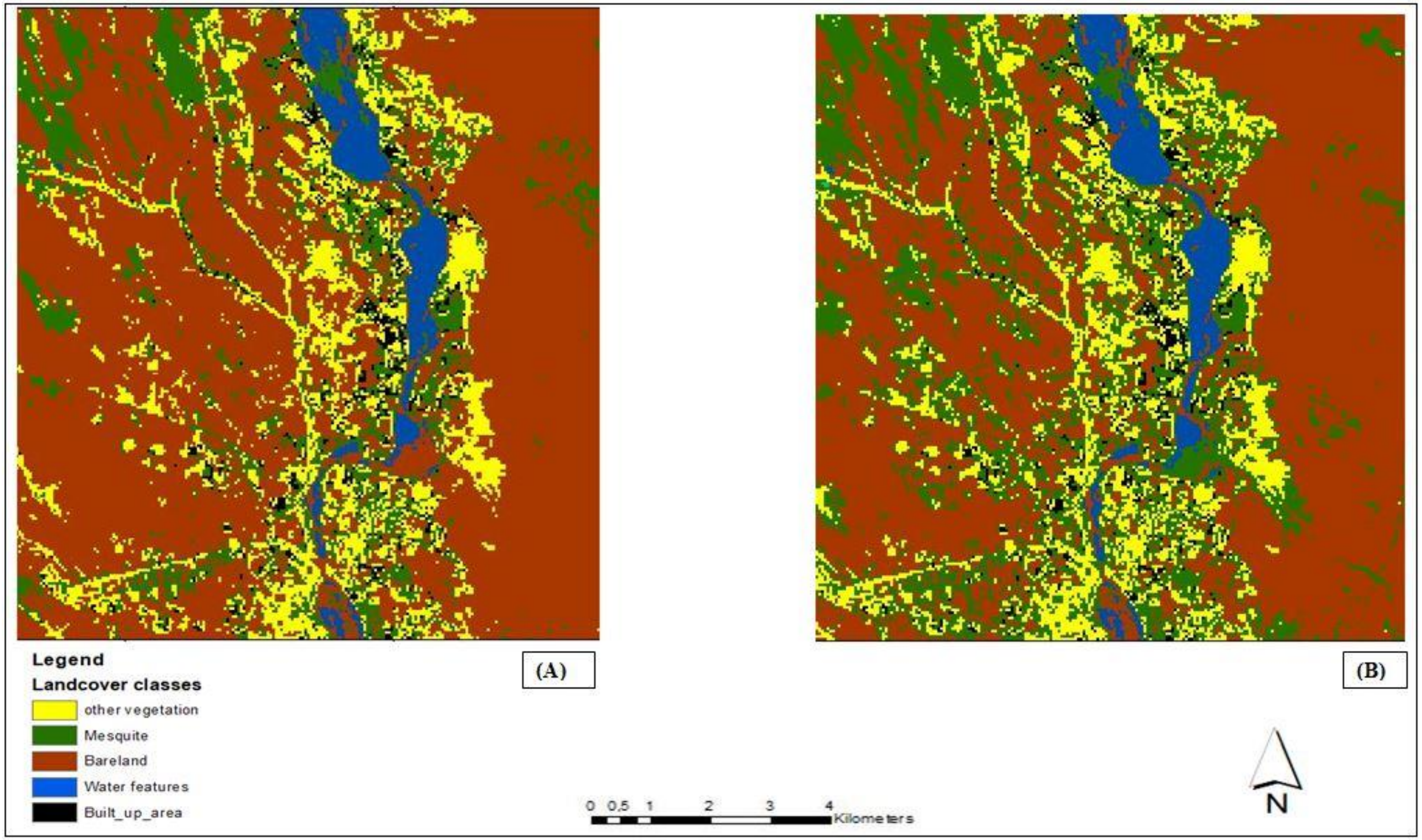

Figure 4.1 The spatial distribution of mesquite as mapped by Sentinel-2 and Random Forest (A) and Support Vector machine (B) classifiers

4.1.2 Areal coverage of Mesquite as obtained through Random Forest and other the LULC classes along the Al Gash River

The areal coverage for mesquite and other LULC is depicted in Table 4.1 and Figure 4.2. The spatial extent of mesquite in the area is broad and accounts for 20.09% of the total aerial coverage of which is the second biggest land cover class. Water features cover a minimal percentage of the area, accounting for 13.7% of the area. Other vegetation covered a total of 12.42% of the region. Bare land was estimated to cover 43.26% of the region.

Table 4.1 Areal coverage of mesquite and other LULCC along the Al Gash River derived using Random Forest

Land cover classes	Area (h)	Area (%)
Mesquite	62 489.70	20.09
Other Vegetation	38 624.31	12.42
Water features	42 689.58	13.70
Built up area	32 684.71	6.88
Bare land	134 584.32	43.26

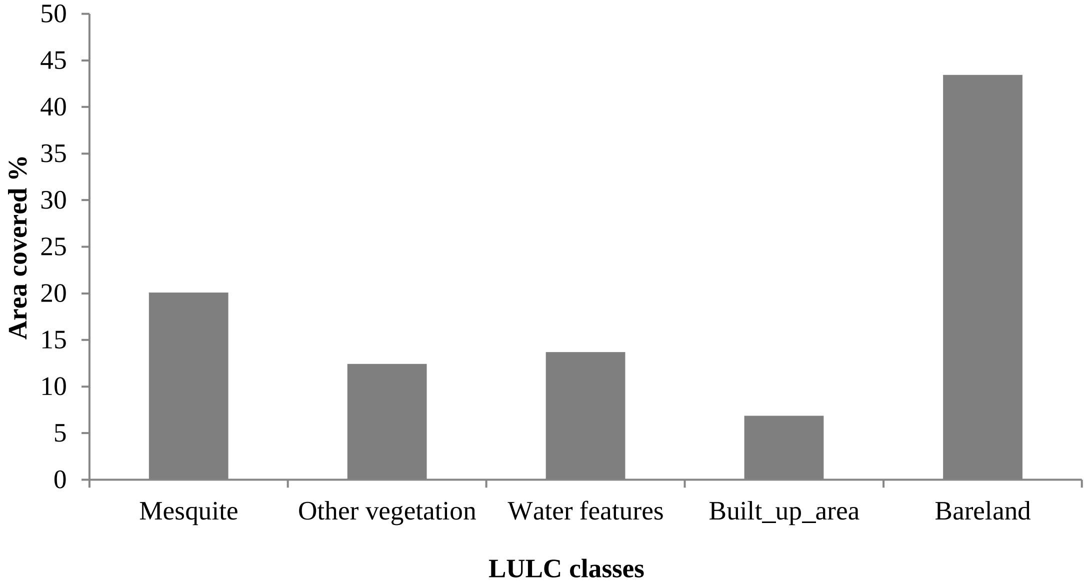

Figure 4.2 Areal coverage of mesquite and other LULCC along the Al Gash River in 2019

4.1.3 Random Forest and Support vector machine optimization results

The various LULCC were classified through the use of the Support vector machine and Random Forest algorithms. Optimising the constraints is vital for determining the best parameter pair to train the RF algorithm for the classification of the identified land cover classes in the region. The optimisation of the RF parameters, n*tree* and m*try*, produced an OOB estimate error degree of 4%. The 4 % error rate was produced from the m*try* and n*try* combination of 9 and 8500 as depicted in Figure 4.3. Variable importance shows that the bands with the highest mean decrease in accuracy are of utmost importance. For this Sentinel-2 image the most important (variables) bands were the NIR, Red and the third vegetation red edge bands.

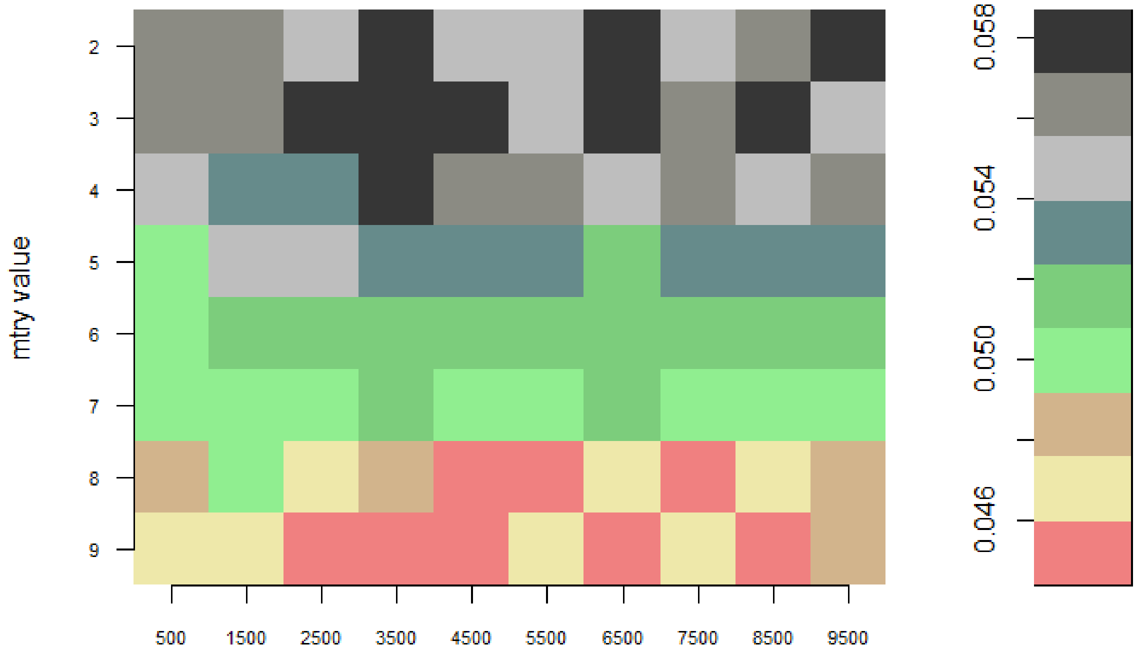

Figure 4.3 Random Forest Optimization

Upon conducting image classification, the error rate as depicted by the gamma and cost parameters of 0.1 and 100 was 14% for the Sentinel-2 image. Furthermore, the efficacy of the separate bands in mapping LULCC was additionally evaluated on the image and the shortwave infrared (2.08-2.32 µm), red (0.646-0.684 µm) and vegetation red edge band (0.783-0.842 µm) were the three most significant bands in the classification of mesquite on the Sentinel-2 image as depicted in Figures 4.4 and 4.5. The Random Forest classifier also provided a mean decrease in accuracy measurement to outline the role that each spectral band plays in the classification procedure as depicted in Figure 4.4.

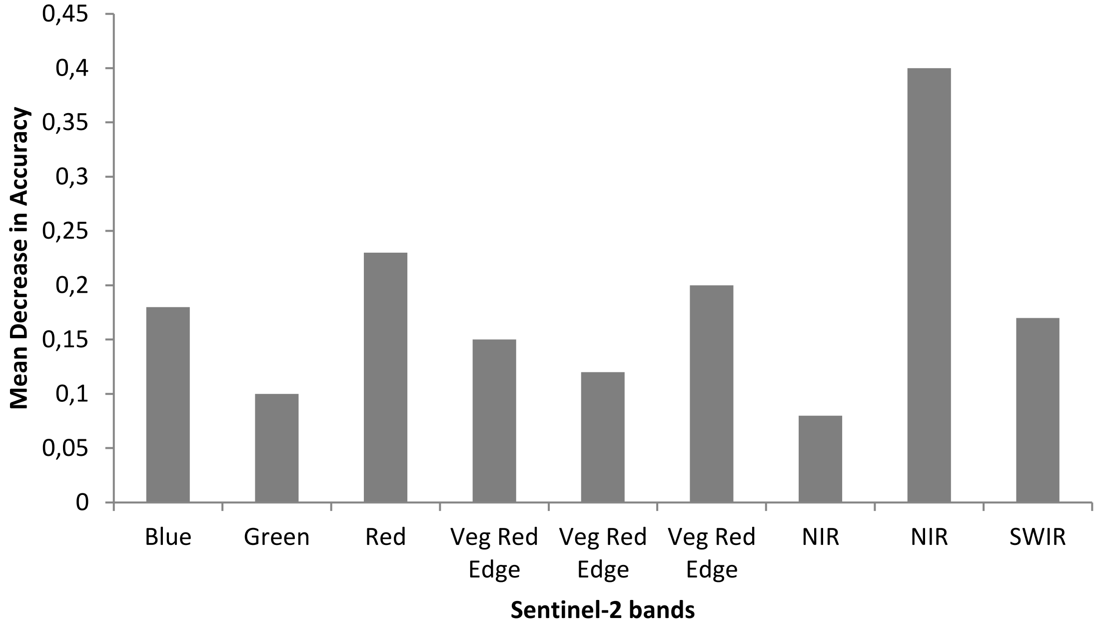

Figure 4.4 Variable importance of the Sentinel-2 spectral bands in the classification of the various land cover classes

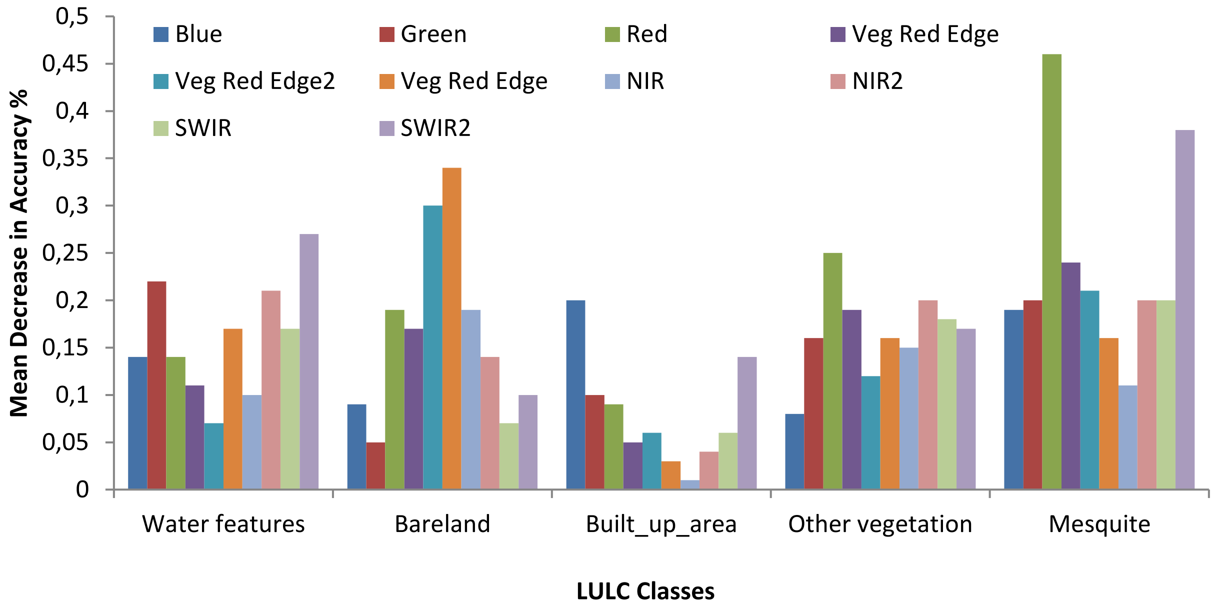

Figure 4.5 The mean decrease in accuracy for the various Sentinel-2 bands in classifying the different land cover classes. The highest mean decrease in accuracy depicts the most significant band

4.1.3 Accuracy assessment

The performance of the support vector machine was validated using the 30% test data set of the ground data. The overall accuracy obtained was 87.57% as shown in table 4.2. The producer accuracies were also high with mesquite (93.94%) being the highest and water features (62.79%) yielding the lowest accuracy. A user's accuracy of 82.30% was achieved due to 12 pixels being misclassified for water features, whereas a producer's accuracy of 93. 44% was obtained due to 3 pixels being classified as water features. A kappa coefficient of 0.82 was produced. The performance of the Random Forest classifier was validated using the test data set which is the 30% of the ground truth collection data. The overall accuracy obtained was 93.44% from the classification process of the image as shown in Table 4.3. The producer accuracies were also high with bare land (100%) and mesquite (96.26%) being the highest and water features (65.91%) yielding the lowest accuracy. A user's accuracy of 99.04% was achieved due to misclassifications for water features, whereas a producer's accuracy of 96.26 % was obtained due to 1 pixel being classified as water features. A kappa coefficient of 0.92 was calculated from the image.

Table 4.2 Confusion matrices generated from the Support Vector Machine classifier accuracy assessment

	Water features	Bare land	Built up area	Other Vegetation	Mesquite	Sum	User's accuracy (%)
Water features	27	2	0	2	3	34	79.41
Bare land	1	65	1	3	0	70	92.86
Built up area	0	0	80	3	1	84	95.24
Other Vegetation	3	2	2	52	2	61	85.25
Mesquite	12	3	3	2	93	113	82.30
Sum	43	72	86	62	99		
Producer's accuracy (%)	62.79	90.28	93.02	83.87	93.94		
Overall accuracy (%)	87.57%						
Kappa coefficient	0.82						

Table 4.3 Confusion matrices generated from the Random Forest classifier

	Water features	Bare land	Built-up area	Other Vegetation	Mesquite	Sum	User's accuracy (%)
Water features	29	0	0	4	3	36	80.56
Bare land	0	71	1	0	0	72	98.61
Built up area	1	0	83	0	0	84	98.81
Other Vegetation	13	0	0	56	1	70	80.00
Mesquite	1	0	0	0	103	104	99.04
Sum	44	71	84	60	106		
Producer's accuracy (%)	65.91	100	98.33	93.33	96.26		
Overall accuracy (%)	93.44						
Kappa coefficient	0.92						

4.2 Determining the historical extent of mesquite invasion along the AL Gash River, Kassala over the past 30 years (from 1989 to 2019)

The spatial distribution maps from Figure 4.6 indicate mesquite invasion and spreading over the years along the Al Gash River. In 1989, mesquite was dominant in the western parts and along the Al Gash River course. The 1994 there was a significant decline in mesquite invasion, especially in the lower western regions. It was also noted that there was a decrease in mesquite along the irrigation channels of the Gash River. In 1999, mesquite expanded greatly in the eastern and western regions as well as along streams feeding into the Al Gash River. Mesquite also covered most agricultural areas close to the Al Gash River. Mesquite drastically declined in 2004 along the eastern parts, agricultural fields and along the Al Gash river course, but expanded in the north-western regions. In 2014, mesquite expansion was more dispersed but was more predominant in the western and south western regions. There was a significant decline in mesquite especially along irrigation channels and along the eastern parts. However, it has expanded gradually in the north western regions.

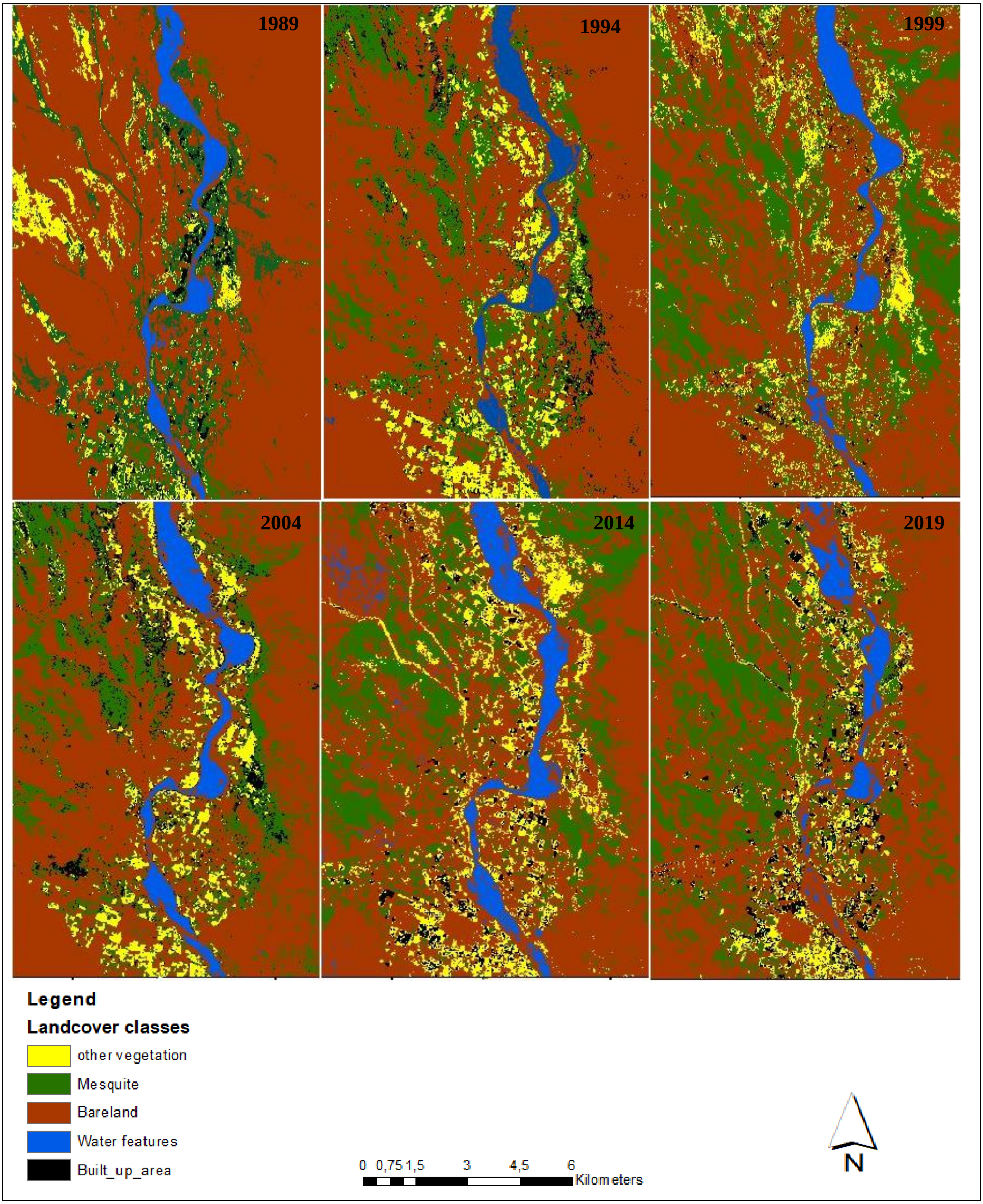

Figure 4.6 The mesquite invasion over the past 30 years by Al Gash River in Kassala, using Landsat imagery and Random Forest classifier from 1989 to 2019 at 5-year intervals

4.2.1 Areal Coverage of mesquite and other LULCC along the Al Gash River over the past 30 years

Table 4.2.1.2 and Figure 4.7 show the areal coverage of Mesquite and other classes along the Al Gash River, over a period of 30 years. In 1989, mesquite covered a total areal coverage of 13.7%, which was 38 679 hectares of land. Bare land was the most dominant LULC class with a total coverage of 51.26%. Mesquite coverage increased in 1994 to 22.51%, amounting to 67 842. 41 hectares of land. The rest of the LULCC decreased in coverage in 1994. In 1999, mesquite rapidly increased to 45.39% of the area, as compared to the other land cover classes which declined. The year 2004 saw a significant decline in mesquite, due to the controls that were in place by then, mainly mechanical methods. Mesquite increased again in the year 2014 to cover 34.4% of the area. In 2019, there was a slight decline in the infestation of mesquite in the area amounting to 29.84% of the area.

Table: 4. 1 Areal Coverage of LULCC along the Al Gash River over the past 30 years

	1989		1994		1999		2004		2014		2019	
	Hectares	%	Hectares	%	Hectares	%	Hectares	%	Hectares	%	Hectares	%
Mesquite	38967.78	13.7	67 842.41	22.51	163 528.6	45.39	74 128.45	26.26	124 365.7	34.46	101 214.6	29.84
Other Vegetation	57678.91	20.28	52 682.52	17.48	62 567.8	17.36	41 587.62	14.73	77 681.92	21.52	87 321.58	25.74
Water features	55 629.34	19.56	48 324	16.03	50 142.41	13.92	52 364.89	18.55	50 281.29	13.93	39 421.6	11.62
Built up area	43 281.67	15.21	46 132.52	15.31	29 684	8.24	37 495.28	13.28	48 921.5	13.55	50 012.8	14.75
Bare land	88 923.41	31.26	86 432.21	28.68	74 341.89	20.64	76 781.4	27.19	59 684.3	16.54	61 218.9	18.05

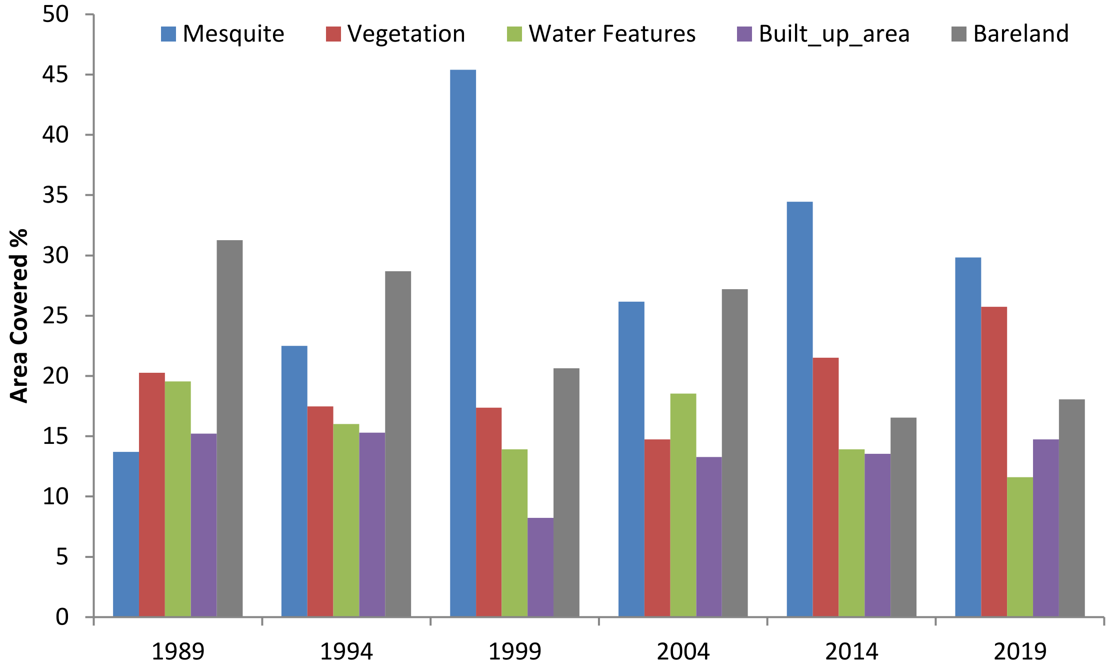

Figure 4.7 Area invaded by mesquite in comparison to other LULCC from 1989 to 2019

4.2.2 Optimisation of Random Forest

The optimisation of the Random Forest parameters was conducted through the use of the 10 fold cross-validation procedure as shown in Figure 4.8. All the variables had to be optimised for each of the Landsat individual images. The OOB estimated errors of the images in chronological order are as follows; 12.6%, 14.03%, 11.15%, 9%, and 4.37% respectively as depicted in Figure 4.8. The image that was captured on the 20th of March 2019 yielded the lowest OOB error rate from n*try* and m*try* combination of 2 and 500. Whereas the image acquired on the 12th of March 1999 achieved an OOB error rate of 14.03% from the Random Forest n*try* and m*try* parameters.

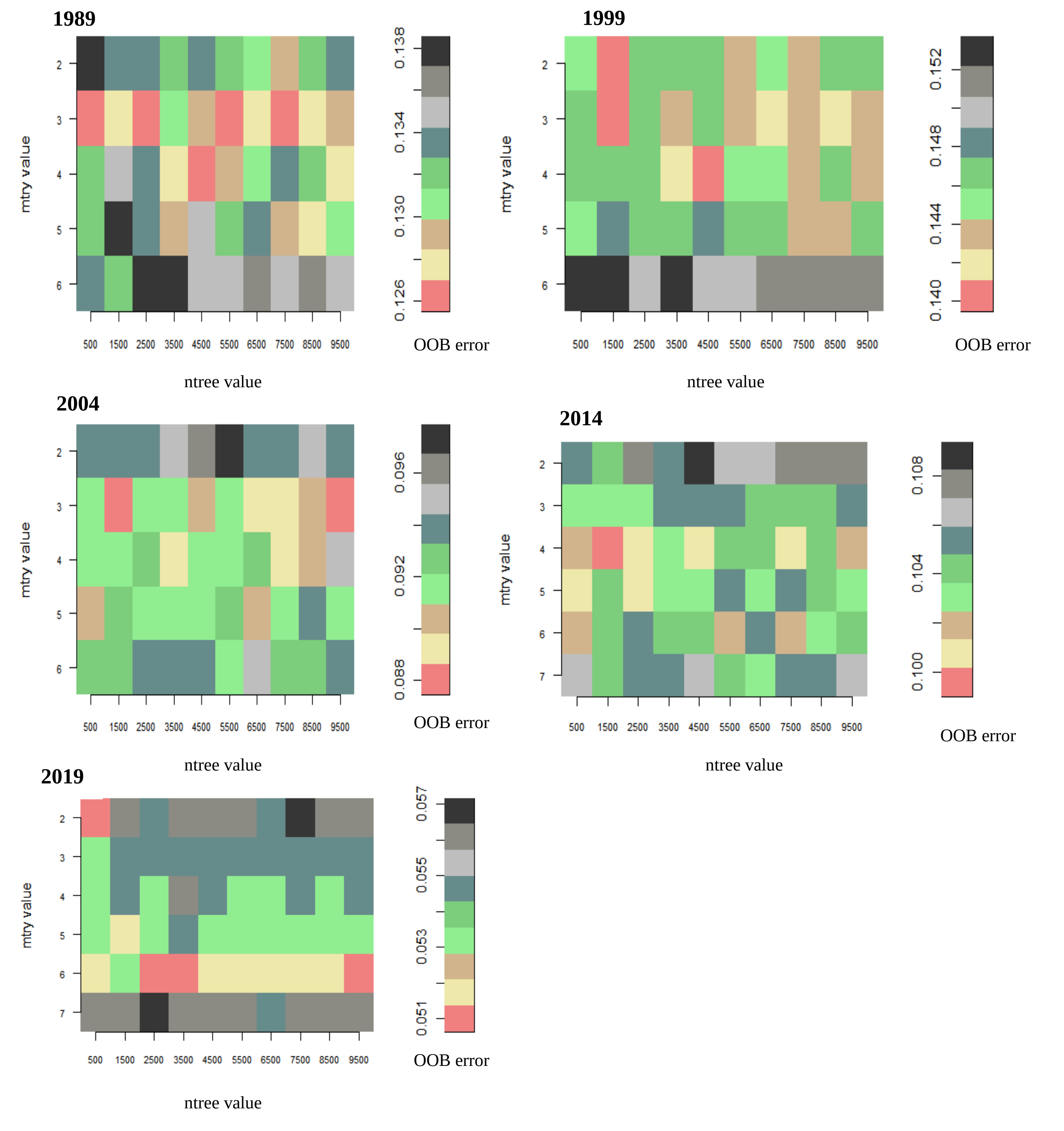

Figure 4.8 Optimization results for Random Forest classifier of parameters (mtry and ntree) for the Landsat imagery from 1989 to 2019

4.2.3 Accuracy assessment

The performance of the Random Forest classifier was validated through using the test data set, which is the 30% of the ground truth collection data for each of the Landsat images. The overall accuracies ranged from 84.7 to 95%. Mesquite had overall high accuracies with the highest (96.43%) from the 2019 image, and the lowest accuracy of 36.84% from the 2004 image. The Kappa coefficient ranged from 0.81 to 0.94 for the images. Table 4.5 shows the confusion matrices generated from the accuracy assessment.

Table 4.5 LULC classification confusion matrices

Year	Overall accuracy (%)	Kappa coefficient	Accuracy type	Water features	Bare land	Mesquite	Other Vegetation	Built up area
1989	88.25	0.85	Producer	75.00	91.55	94.05	89.72	83.33
			User	73.33	91.55	86.81	92.31	90.91
1994	87.42	0.84	Producer	69.2	95.73	75.3	89.5	96.8
			User	88.5	96.6	68.3	97.12	94.38
1999	84.97	0.81	Producer	75.00	88.73	84.52	66.67	97.20
			User	70.21	80.77	84.52	78.43	98.11
2004	88.85	0.85	Producer	68	98.59	41.18	95.00	96.26
			User	89.47	87.47	36.84	98.28	93.64
2014	89.14	0.87	Producer	64.81	91.89	94.74	92.31	94.34
			User	89.74	85.00	86.75	88.24	95.24
2019	95	0.94	Producer	86.36	97.18	96.43	93.33	99.07
			User	95.00	95.83	95.29	91.80	98.15

4.2.4 Change detection statistics of mesquite as detected by Landsat from 1989 to 2019

In the assessment of the control measures that were implemented to combat the invasion of mesquite in the area, change detection statistics were computed between two periods of the classified image pairs. The changes were evaluated by computing the variations of pixels between the image pairs. The change detection variations between 1989 and 1994 depicted a mesquite tree gain of 48 257 as according to table 4.6. Between 1999 and 2004, there was a decrease in the mesquite tree gain and a decrease of 9782 of mesquite trees. However, in the years 2014 to 2019, there was an increase in mesquite gain amounting to 89 142 of which is the biggest change detected in the invasion of mesquite along the Al Gash River in Kassala. Table 4.6 depicts the change detection statistics with regards to mesquite invasion over the past 30 years in Kassala.

Table 4.6 Change detection statistics for mesquite from 1989 to 2019 as detected by Landsat series data

Years	Mesquite trees disappeared (Ha)	No change (Ha)	Mesquite tree gain (Ha)
From 1989 to 1994	1286	301 287	48 257
From 1999 to 2004	9782	214 084	40 789
From 2014 to 2019	10 236	192 485	89142

4.3 Discussion

4.3.1 The current spatial distribution of mesquite along the AL Gash River using Sentinel-2 and Support vector machine and Random forest classifiers

The rapid infestation of mesquite has brought substantial negative effects on various ecosystems and bio-diversity across different landscapes (Nzumira, 2014). To effectively improve the knowledge amongst researchers and decision makers involved in devising effective control measures to be able to deal with the problem, it important that both consistent and precise information is attained on the spatial and temporal dynamics of the invasive species. Remote sensing techniques are increasingly being used for detection of invasive species (Gomez, 2017). The main objectives were to investigate Sentinel-2 in the detection and mapping of the current spatial distribution of mesquite and to test the performance of the Random Forest and Support vector machine algorithms in species classification for identifying mesquite. The spatial distribution of mesquite accounted for 20.09% of study area. From the classification process, mesquite had expanded in the north western regions and was dominant along the Gash river and irrigation channels in Kassala. This invasion is mainly amplified by the movement of animals, anthropogenic factors and the lack of information with regards to the temporal and spatial dynamics of this invasive species (Nzumira, 2014). However, in the eastern regions, mesquite had a low spatial coverage and bare land was the dominant land cover due to less water resources being present in that region, as mesquite has an aggressive appetite for water sources and tends to invade riverine regions predominantly (Nzumira, 2014).

The ability of the sensor to discriminate mesquite from other native species is mainly due to the inclusion of the vegetation red edge band in its spectral properties. This vegetation red edge spectral band holds great capabilities of mapping vegetation of which is attributed to better classification of the image (Drusch, 2012; Ramoelo *et al.*, 2015). The reflective properties of mesquite differ from those of other native species, it reflects wavelengths that are longer and hence is easily detected, attributing to the higher overall accuracies achieved (Ng *et al.*, 2017).

This study is similar to a study by Rajah *et al.*, (2019) based on the utility of Sentinel-2 indices and Sentinel-1(SAR) data for mapping invasive species in uKhahlamba located in Drakensburg South Africa. It was also noted that the derived vegetation indices from the Sentinel-2 imagery achieved an overall accuracy of 80% in detecting and mapping the invasive species in the region (Rajah *et al.*, 2019). It was also observed that through the use of the red edge band, vegetation was effectively detected and mapped (Rajah *et al.*, 2019). The Random Forest algorithm was also effective and produced high accuracies in the detection and mapping of the invasive species.

This research indicated that through the use of the 13-band Sentinel-2 sensor in the detection and mapping of the spatial distribution of mesquite, high accuracies can be produced. Overall accuracies of this calibre have mostly been achieved through the use of hyperspectral data due to the higher spectral and spatial resolution (Mureriwa *et al.*, 2016). However, coupled with these sensors are the limitations of high costs in attaining the data. Conversely, the use of new generational multispectral sensors of which includes the Sentinel- 2 sensor, provide updated spatial and temporal information at no cost and gives high levels of accuracies in the monitoring and mapping of the spatial dynamics of the invasive species (Mureriwa *et al.*, 2016). The overall accuracies were mainly attributed to the high spatial and spectral properties of the sensor. Ng *et al* (2017) achieved an overall accuracy of 79% in mapping the spatial extent of mesquite through the use of Sentinel-2 data in Kenya. The study also mapped the spatial distribution of two invasive species, namely Pleiades and mesquite. It was found that mesquite was the most dominant invasive species within the area of study (Ng *et al.*, 2017). The application of new generational multispectral sensors to map the temporal and spatial distribution patterns of mesquite provides an attractive technique for decision makers and conservationists. This would ensure better management and control of the invasive species.

Currently, there is an increasing demand in utilizing remotely sense data for mapping and monitoring the dynamics of various alien plants. In contrast to traditional classifiers such as Maximum Likelihood, machine learning algorithms have been viewed to yield higher accuracies for image classification due to the fact that they are more robust and advanced in detecting and mapping vegetation species.

Both the classifiers depicted results that validated that mesquite can be distinguished and discriminated accurately from the other land cover classes along the Al Gash River in Kassala. Since the Random Forest and Support Vector Machine algorithms were used on the same training and test data point on the present study, their ability to detect mesquite and other plants was tested. The Random Forest algorithm achieved a relatively higher accuracy as compared to the overall accuracy achieved by the Support vector machine algorithm. These results also agree with other studies performed from literature such as that of Mureriwa *et al.*, (2016) where both Support Vector machine and Random Forest algorithms were applied and high accuracies achieved in mapping of mesquite in the Northern Cape.

Both the random forest and support vector machine algorithms yielded high overall classification accuracies, 93.44% and 87.57%, respectively. The producer and user's accuracies obtained were also generally high with mesquite and bare land being the highest and water features being the lowest for both classifiers. The high overall accuracies were as a result of the fact that in the Random Forest classifier, the several decision trees are joined and each tree adds a vote to assign a class to a pixel and eventually to the overall image classification (Breiman, 2001; Adam et al., 2017). Whereas in the case of the Support Vector machine, it is a non-parametric supervised machine learning algorithm and mainly enhances the non-linear procedure through the use of a number of methods such as kernel of which is a binary linear classifier (Vapnik, 1979; Gomez, 2017). The algorithm of the binary linear classifier is meant to maximise the distance for each class from the data points in the optimal hyper plane or decision boundary to the training data (Vapnik, 1979; Gomez, 2017).

The overall accuracies were also as a result of the higher spectral and spatial resolutions of the Sentinel-2 sensor as well as the advanced machine learning algorithms in the detection and mapping of mesquite amongst the other existing LULCC (Adam *et al.*, 2017). The ability of the sensor to discriminate mesquite from other native species is mainly due to the inclusion of the vegetation red edge band in its spectral properties. This vegetation red edge spectral band holds great capabilities for mapping vegetation, which is attributed to better classification of the image (Drusch, 2012; Ramoelo *et al.*, 2015). The reflective properties of mesquite differ from those of other native species, it reflects wavelengths that are longer and hence is easily detected, attributing to the higher overall accuracies achieved (Ng *et al.*, 2017).

Although, it is seen that through the use of remote sensing it is possible to provide data in relation to the spectral, temporal and spatial variability of the invasion of mesquite, there are however uncertainties with regards to the use of medium spectral and spatial resolutions of commercial satellites such as SPOT and Landsat (Van den Berg *et al*., 2010). Several studies confirmed Sentinel-2 MSI imagery and machine learning algorithms as good for invasive species mapping and discrimination (Dube *et al*., 2015; Shoko and Mutanga, 2017). The overall accuracies obtained by both algorithms were mainly attributed to the spatial and spectral properties of the sensor. A study conducted by Ng *et al* (2017) achieved an overall accuracy of 79% in mapping the spatial extent of mesquite through the use of Sentinel-2 data. The authors utilised Random forest in the classification process, since the main aim of the study was to test the ability of Random Forest in being able distinguish mesquite amongst its other native species (Ng *et al*, 2017). New generational multispectral sensors offer decision makers and conservationists handy tools for efficient mapping of the temporal and spatial distribution patterns of mesquite. This would therefore ensure better management and control of the invasive species.

4.3.2 Historical spatial distribution of mesquite from 1989 to 2019 as detected by Landsat series data

Mechanical control methods were utilised in Kassala, Sudan to deal with mesquite infestation (Nzumira, 2014). The results depicted an overview change in the natural invasion patterns and spatial extent of mesquite over the years along the Al Gash River. The historical spatial extent of mesquite over the last 30 years was observed through change detection analysis and the mapping of mesquite. The LULCC maps generated from the Landsat images were adequate. From the results obtained, it was evident that in 1989, mesquite was dominant in the western parts and along the Gash River course. The year 1994 saw a significant decline in mesquite invasion in the lower western regions of the study area. It was also noted that there was a decrease in mesquite along the irrigation channels of the Gash River. In 1999, mesquite expanded greatly in the eastern and western regions as well as along streams feeding into the Gash River. Mesquite also covered most agricultural areas close to the Gash River. The increase was due to some factors which included the continuous repeated introductions of mesquite from unknown sources (Pasiecznik *et al.*, 2001). It was also due to the species invasive nature (Babiker, 2006). Poor management of the programmes initiated at combating mesquite invasion and spread from 1996 (Babiker, 2006).

Mesquite drastically declined in 2004 along the eastern parts, agricultural fields and along the Gash river course, it however expanded in the north-western regions. This may be attributed to the control methods that were place such as the New Halfa programme which was revised in terms of farmers being told to use mechanical methods such as cutting down mesquite trees and using them for charcoal production. The programme also advised farmers to plant quick growing plants to prevent future mesquite invasion (IFAD, 2004; Nzumira, 2014). In 2014, mesquite expansion was more dispersed but was more predominant in the western and south western regions of the study region. There was a significant decline in mesquite, especially along irrigation channels and along the eastern parts. However, it has expanded gradually in the north western regions. A similar study was conducted by Mohammed *et al.,* (2013) where mesquite was mapped and detected from 1979 to 2010 through Landsat series data. It was found that within the Gash Agricultural scheme, there was an increase mesquite invasion over the years mainly attributed to the lack of on-going monitoring of the programs put in place to curb the spread of mesquite in the region (Mohammed *et al.,* 2013). Mesquite invasion dominated riverine regions within the agricultural scheme. It was also noted that crop production in the region has decreased mainly due to mesquite invasion in the area since the scheme relies mainly on the Al Gash River for irrigation and agricultural purposes (Mohammed *et al.*, 2013).

The introduction of chemical and biological control methods was unable to outperform mechanical and physical methods (Nzumira, 2014). These two major methods have proven to be the most effective and adequate control measures in Kassala (Steenbergen, 2014). However, through the Sudanese government's eradication programmes such as the New Halfa and the Mesquite Trees, management projects have ensured that physical and mechanical control methods remain as the predominant control measures in Kassala (Nzumira, 2014). Such initiatives were explicitly established for the mechanical and physical control of mesquite in Sudan (Babiker, 2006). These initiatives also help the country socially, economically and environmentally by improving the local people's standards of living through employment and promoting the conservation of biodiversity through the eradication of invasive species (Mohamed, 2001).

Change detection results as detected by Landsat series data also depicted that the years from 2014 to 2019 experienced the largest mesquite invasion amounting to 89 142ha. Mohammed *et al.*, (2013) achieved similar results upon conducting change detection throughout the whole

Gash Agricultural scheme whereby the periods with the highest mesquite invasion were 1999 due to some of the eradication programmes poor management and monitoring of mesquite. However, in the year 2010 a decline in the infestation of mesquite was experienced due to the New Halfa programme (Mohammed *et al.*, 2013).

The Sudanese government with the aid of private firms had initiated control methods along with mostly mechanical and biological strategies to avert the infestation of mesquite in the region (Steenbergen, 2014). This was mainly done to avoid future invasions from occurring at great magnitudes (Steenbergen, 2014). Nzumira (2014) noted an increase in mesquite invasion from 1978 to 2013 through the use of Land sat data. Mesquite had evidently invaded the in land delta predominantly over the years at a rapid rate mainly propagated through animal grazing and anthropogenic processes. Mesquite was dominant along river banks and canals and was found to be within agricultural areas and fertile regions (Nzumira, 2014). There was a total areal coverage of 141 942 hectares of land invaded from 1978. Another similar study by Zeila (2011) in Garissa, Kenya mapped the spatial distribution of mesquite over a six year period from 2000 to 2006 by using Landsat data. It was observed that 440 square kilometres (33%) of the land had been invaded by mesquite over the years. The study also showed that the most invaded LULC class was riverine, since mesquite usually invades regions where water sources are abundant (Zeila, 2011). This although has proven to be very problematic due to the fact that most farmers depend on the River Tana in the region for irrigation and domestic purposes.

Although it was not the intention of the study to compare the performance of sensors in classifying and mapping mesquite and other LULC, it was noted that, Sentinel-2 and Landsat 8 OLI achieved considerably high accuracies through the Random Forest and Support Vector algorithms for 2019. The overall accuracies for both the images obtained were 93.44% (Random Forest), 87.57% (Support Vector) and 95% (Random Forest), for Sentinel- 2 and

Landsat 8 OLI, respectively. In addition, in terms of area occupied by mesquite in 2019, Landsat 8 detected more area (29.84%), whereas for Sentinel 2, it was smaller, covering 20.09%.

CHAPTER 5:

OVERALL CONCLUSION

Mesquite is an invasive weed species in Kassala, Sudan and it spread rapidly over land and floodplains. The spatial distribution of mesquite is specifically propagated by both anthropogenic processes and animals through open-grazing. Hence, mesquite should be continuously monitored and evaluated as its invasion can yield negative impacts on the livelihoods, communities and the environment. The phenology of mesquite makes the invasive unique in the sense that it is drought resistant has an extensive root system and generates pods faster than most of the native species in the region, thus making its distribution rapid. The main aim of this study was to map the current spatial distribution and invasion of mesquite along the Al Gash River in Kassala. This aim was achieved through the objectives of the study which were to map the spatial distribution of mesquite through the use of Sentinel-2 data, determine the historical spatial extent of mesquite over the past 30 years through the use of multi-temporal Landsat data and lastly to test the performance of two machine learning classifiers.

The objectives of the study were attained as the spatial distribution of mesquite in Kassala was effectively mapped through the use of remotely sensed data and advanced machine learning classifiers. The study also portrayed the ability of Sentinel-2MSI data in the detection and mapping of mesquite by yielding high accuracies. However, misclassifications attained through the use of the Support Vector machine classifier were as a result of the spectral variations within the same LULCC. It was also found that Sentinel-2 data can be utilised in mapping and monitoring the spatial extent of mesquite, indicating that mesquite covered the majority area of the study site. The spatial extent of mesquite was satisfactorily detected and mapped effectively at a high level of accuracy of 85.57% through the use of the NIR, red edge vegetation and red bands, respectively. Additionally, the LULCC classification of the new

generation multispectral sensor was able to distinguish mesquite spread in the region. The high spatial resolution (10 m) of the sensor allows for the increased detail and accuracy associated with land cover classification. Through the testing of the classifiers, it was observed that both classifiers achieved reasonable overall accuracies. However; the Random Forest classifier was more accurate (93.44%) in the detection of mesquite when compared to the support vector machine classifier. The Random Forest classifier also produced the most important bands for the mapping of mesquite which included the red edge vegetation, the red and SWIR-2 bands. Change detection results depicted that mesquite invasion increased from 2014 to 2019, with a mesquite tree gain of 89142 ha of which was mainly attributed to the lack of adequate management in the initiated programmes in Kassala. The control and management of mesquite in Kassala is challenging and requires economically feasible means and resources.
Formulating management strategies that are aimed at preserving benefits and minimizing costs associated with controlling the plant need to be properly understood. Further, through the application of change detection techniques, the study was able to infer on the invasion of mesquite over the last 30 years and review the success of the eradication programmes initiated at various years. As stated previously the spatial extent of mesquite is mainly due to open grazing by animals and anthropogenic processes as well as through flooding.

Critical gaps with regards to the lack of knowledge about the spatial and temporal dynamics of mesquite invasion still exist. Consequently, the use of sensors with finer spatial and spectral resolutions should be used in mapping mesquite to obtain higher accuracies. Methods such as object-based classification can be additionally explored in the mapping of mesquite. This study employed the use of spectral bands in detecting and mapping mesquite invasion by the AL Gash River in Kassala, it would also be recommended for future studies that the use of indices be applied to yield higher accuracies and for better species detection. The evaluation of the specie has been found to be a problem for most decision makers in most sub-Saharan countries, because the measures to control mesquite can only be realised when environmental conditions that favour its success in arid or semiarid regions are better understood. Formulating management strategies that are aimed at preserving benefits and minimizing costs associated with controlling the plant need to be properly understood.

REFERENCES

Abdulahi, M.M., Ute, J.A., Regasa, T. 2017. Prosopis juliflora L: Distribution, impacts and available control methods in Ethiopia, *Tropical and Subtropical Agroecosystems*, *20*(1). pp75-89.

Abualgasim, M.M.R. 2017. *Mapping and Assessing Impacts of Land Use and Land Cover Change by Means of Advanced Remote Sensing Approach: A Case Study of Gash Agricultural Scheme, Eastern Sudan* (Doctoral dissertation, Technische Universität Dresden).

Adam, E., Mureriwa, N., Newete, S. 2017. Mapping Prosopis glandulosa (mesquite) in the semi-arid environment of South Africa using high-resolution WorldView-2 imagery and machine learning classifiers. *Journal of Arid Environments*, 145, 43-51.

Ali, M.Y.M. 2008. *Use of Remote Sensing and GIS in detecting the Spread of Mesquite (Prosopies chilenses): A case Study of Elgetaina locality (Northern White Nile State)*(Doctoral dissertation, University of Khartoum).

Alredaisy, S.M.A.H. 2013. Ecological Impacts of Mesquite 'Prosopis spp.'Expansion in Delta Toker Agricultural scheme, Northeastern Sudan.

Asner, G.P., Jones, M.O., Martin, R.E., Knapp, D.E. and Hughes, R.F. 2008. Remote sensing of native and invasive species in Hawaiian forests. *Remote Sensing of Environment*, 112, 1912-1926.

Artigas, F., Yang, J. 2005. Hyperspectral remote sensing of marsh species and plant vigour gradient in the New Jersey Meadowlands. *International Journal of Remote Sensing*, *26*, 5209-5220.

Babiker, A. 2006. Mesquite (Prosopis spp) in Sudan: history, distribution and control. *Problems posed by the introduction of Prosopis spp. in selected countries. Food and Agricultural Organization of the United Nations (FAO) Plant Production and Protection Division, Rome, Italy*.

Babiker, A., Hoshino, B., Rakuno, G., Nawata, H., Yoda, K., Ruichen, J. 2008. *Retrieve the soil moisture from radar backscattering coefficient using ALOS/PALSAR polarization (HH/VV) data*.

Bokrezion, H. 2008. The ecological and socio-economic role of Prosopis juliflora in Eritrea. *Academic Dissertation, Johannes Gutenberg-Universität Mainz, Germany*.

Breiman, L. 2001. Random forests. *Machine learning*, 45(1), 5-32.

Burka, A. 2008. Land Use/Land Cover Dynamics in Prosopis juliflora invaded area of Metehara and the Surrounding Districts Using Remote Sensing & GIS Techniques. *Faculty of science, RS & GIS, Department of Earth Science, AAU*.

Castro Gomez, M.G. 2017. Joint use of Sentinel-1 and Sentinel-2 for land cover classification: A machine learning approach. *Lund University GEM thesis series*.

Cortes, C., Vapnik, V. 1995. Support-vector networks. Machine learning, vol. 20, no.3, pp. 273-297.

Das, T. 2017. Machine Learning algorithms for Image Classification of hand digits and face recognition dataset. *Machine Learning*, *4*(12), pp.640-649.

Department of Natural Resources and Mines, Q. 2003. Weed Management Guide. Mesquite-*Prosopis species.*Retrieved from:www.weeds.org.au/WoNS/mesquite/docs/Weed_Management_Guide-Mesquite.pdf.

Drusch, M. *et al*. 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25–36.

Dube, T., Mutanga, O., Sibanda, M., Bangamwabo, V., Shoko, C. 2017. Evaluating the performance of the newly-launched Landsat 8 sensor in detecting and mapping the spatial configuration of water hyacinth (*Eichhornia crassipes*) in inland lakes, Zimbabwe. *Physics and Chemistry of the Earth*, 100, 101-111.

Elhag, A.R., Elsheikh, R.F.A., Abdelradi, S.F.

European Space Agency (ESA). 2015. Sentinel-2 user handbook. Retrieved from https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook

Foody, G. M. 2002. Status of land cover classification accuracy assessment. Remote sensing of environment, 80(1), 185-201.

Forkuo, E.K., Frimpong, A. 2012. Analysis of forest cover change detection. *International Journal of Remote Sensing Applications*, 2(4), 82-92.

Franklin, S.E., Wulder, M.A. 2002. Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. *Progress in Physical Geography*, *26*(2), pp.173-205.

Frenken, K. ed. 2005. *Irrigation in Africa in figures: AQUASTAT Survey, 2005* (Vol. 29). Food & Agriculture Org.

Hegazy, I.R., Kaloop, M. R. 2015. Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt, *International Journal of Sustainable Built Environment*, 39(4), 117-124

Helldén, U. 1984. Drought impact monitoring. A remote sensing study of desertification in Kordofan, Sudan. *Rapporter och Notiser-Lunds Universitets Naturgeografiska Institution*.

HOSHINO, B., KARAMALLA, A., Manayeva, K., Yoda, K., Suliman, M., Elgamri, M., Nawata, H., Yasuda, H. 2012. Evaluating the invasion strategic of Mesquite (Prosopis juliflora) in eastern Sudan using remotely sensed technique.

Huang, C.-y., Asner, G.P. 2009. Applications of remote sensing to alien invasive plant studies. *Sensors*, *9*, 4869-4889

Hussain, M.I., Shackleton, R.T., El-Keblawy, A., Del Mar Trigo Perez, M., Gonzalez, L. 2020. Invasive Mesquite (Prosopis juliflora), an allergy and health challenge. *Plants*, *9*(2), p.141.

Immitzer, M., Vuolo, F., Atzberger, C. 2016. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. Remote Sensing, 8(3), 166-192.

Joshi, C., de Leeuw, J., van Duren, I.C. 2004, July. Remote sensing and GIS applications for mapping and spatial modelling of invasive species. In *Proceedings of ISPRS* (Vol. 35, p. B7).

Karatzoglou, A., Meyer, D., Hornik, K. 2006. Support Vector Machines in R. Journal of Statistical Software 15.

Kazmi,S.J.H., Shaikh, S. Zamir, U.B., Zafar, H., Rasool, A., Tariq, F., Afzal, A., Arif, T. 2009. Ecological and socio-economic evaluation of the use of Prsopies julifloraa for bio-char production in Pakistan. *Pakistan;Drynet* (page numbers)

Keebine, G.L. 2019. *Mapping and monitoring the spatial distribution of eichhornia crasipes(water hyacinth) in Hartebeesport dam, South Africa, using remote sensing datas* .

Kganyago, M., Odindi, J., Adjorlolo, C., Mhangara, P. 2018. Evaluating the capability of Landsat 8 OLI and SPOT 6 for discriminating invasive alien species in the African Savanna landscape. *International Journal of Applied Earth Observation and Geoinformation*, 67, 10-19.

Laxén, J. 2007. Is prosopis a curse or a blessing. An ecological-economic analysis of an invasive alien tree species in Sudan (Dissertation). Helsinki: University of HelsinkiViikki, Tropical Resources Institute. (VTR Institute)

Lowe, S., Browne, M., Boudjelas, S., De Poorter, M., 2000. *100 of the world's worst invasive alien species: a selection from the global invasive species database* (Vol. 12). Auckland: Invasive Species Specialist Group.

Lu, D., Mausel, P., Brondizios, E., Moran, E. 2004. Change detection techniques, *International Journal of Remote Sensing*, 25(12), 2365-2407.

Meroni, M., Ng, W.T., Rembold, F., Leonardi, U., Atzberger, C., Gadin, H., Shaiye, M. 2017. Mapping Prosopis juliflora in west Somiland with Landsat 8 satellite imagery and ground information. *Land Degradation & Development*, *28*(2), pp.494-506.

Mesev, V., 2010. Classification of urban areas: inferring land use from the interpretation of land cover. *Remote sensing of urban and suburban areas*, 141-164.

Mirik, M., Ansley, R.J. 2012. Utility of satellite and aerial images for quantification of canopy cover and infiling rates of the invasive woody species honey mesquite (Prosopis Glandolusa) on rangeland. *Remote Sensing*, *4*(7), pp. 1947-1962.

Mohamed, A. A. 2001. Some Aspects of Germination, Dormancy and Allelopathy of Prosopis juliflora (Mesquite). M.Sc Thesis University of Gezira.

Mohammed, M.R.A., Csaplovics, E., Babatunde. 2013. Mapping and monitoring land-cover/land-use change in the Gash Agricultural scheme (Eastern Sudan) using remote sensing. In *Remote Sensing for Agriculture, Ecosystems and Hydrology XV* (Vol. 8887, p. 88871E). International Society for optics and Photonics.

Mountrakis, G., Im, J., Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 66, no.3, pp. 247-259.

Mureriwa, N., Adam, E., Sahu, A., Tesfamichael, S. 2016. Examining the spectral separability of Prosopis glandulosa from co-existent species using field spectral measurement and guided regularized random forest. *Remote Sensing*, 8(2), 144-159.

Mutanga, O., Adam, E., Cho, M.A. 2012. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *International Journal of Applied Earth Observation and Geoinformation*, 18, 399-406.

Mwangi, E., Swallow, B. 2005. Invasion of Prosopis juliflora and local livelihoods: Case study from the lake Baringo area of Kenya. Nairobi, Kenya: World Agroforestry Centre

Mwania, D.K. 2017. *Distribution and Density of the Invasive Plant Species, Prosopis juliflora, in the Western Turkana Region of Northern Kenya* (Doctoral dissertation, Ohio University).

Ng, W.T., Rima, P., Einzmann, K., Immitzer, M., Atzberger, C., Eckert, S. 2017. Assessing the Potential of Sentinel-2 and Pleidas Data for the detection of Prsopis and Vachellia spp. In Kenya. *Remote Sensing*, 9(11), p.74.

Nzumira, H.S. 2014. *Mesquite trees infestation of the Gash Spate Irrigation system in Kassala state, Sudan* (Doctoral dissertation, UNESCO-IHE).

Pasiecznik, N.M., Felker, P., Association, H.D.R. 2001. *The'Prosopis Juliflora'-'Prosopis Pallida'Complex: A Monograph*. HDRA Coventry

Ramoelo, A., Cho, M., Mathieu, R., Skidmore, A.K. 2015. Potential of Sentinel-2 spectral configuration to assess rangeland quality. *Journal of Applied Remote Sensing*, 9(1), 94-96.

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.

Robinson, T., Wardell-Johnson, G., Pracilio, G., Brown, C., Corner, R., & van Klinken, R. 2016. Testing the discrimination and detection limits of WorldView-2 imagery on a challenging invasive plant target. *International Journal of Applied Earth Observation and Geoinformation*, *44*, 23-30

Shackleton, R.T., Le Maitre, D.C., Pasiecznik, N.M., Richardson, D.M. (2014b). Prosopis: a global assessment of the biogeography, benefits, impacts and management of one of the world's worst woody invasive plant taxa. *AoB Plants, 6*

Shoko, C., Mutanga, O. 2017. Examining the strength of the newly-launched Sentinel 2 MSI sensor in detecting and discriminating subtle differences between C3 and C4 grass species. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129, 32-40.

Singh, A. (1989) Digital Change Detection Techniques using Remotely Sensed Data, *International Journal of Remote Sensing*, 10(6), 989-1003.

Steenbergen, F. v., 2014. *Controlling and /or Using Prosopis Juliflora in Spate Irrigation System.* UNESCO-IHE, Institute for Water Education.

Thorp, J. R., Lynch, R., Trust, N. H., 2000. *The determination of weeds of national significance.* National weeds strategy executive committee.

Van den Berg, E.C. 2010. Detection, quantification and monitoring Prosopis spp. in the Northern Cape Province of South Africa using remote sensing and GIS/EC van den Berg. In: North-West University

Vapnik, V.N. 1995. *The nature of statistical learning theory* New York: Springer-Verlag (Is this a book or journal)

Zeila, A. 2011. *Mapping and managing the spread of Prosopis Juliflora in Garissa Country, Kenya* (Doctoral dissertation, Kenyatta University).