

https://dspace.org/

Mapping the spatiotemporal distribution of the exotic Tamarix species in riparian ecosystem using Multi-temporal remote sensing data

Kekana, Thabiso.

2019

https://hdl.handle.net/10539/29092

Kekana, Thabiso Richard (2019) Mapping the spatiotemporal distribution of the exotic Tamarix species in the riparian ecosystem using multi-temporal remote sensing data, University of the Witwatersrand, Johannesburg, <http://hdl.handle.net/10539/29092>

Downloaded from a DSpace repository. https://dspace.org/



MSc Report

Mapping the spatiotemporal distribution of the exotic *Tamarix* species in riparian ecosystem using Multi-temporal remote sensing data

ThabisoKekana (775338)

Supervisors: Dr Elhadi Adam

Dr Solomon Newete

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

Johannesburg

25 March 2019

Abstract

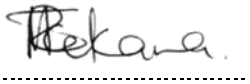
Tamarix spp, commonly known as tamarisk or salt cedar, belong to the family of Tamaricaceae. It is a phreaphytic halophyte with 55 species in the genus *Tamarix*. South Africa has one indigenous (*Tamarix usneoides*) and two exotic (*T. ramosissima* and *T.chinensis*). Not only are the exotic *Tamarix* species becoming infamous invaders, but their hybridisation with the indigenous *T. usneoides* is also complicating morphological discrimination between the different species, and the prospect of potential use of bio-control agents to curb invasion. Thus, lack of spatial information about the current and the past distribution of tamarisk have hampered the effort to control its invasion. This study aimed at investigating the use of multi-temporal remotely sensed data to map the exotic *Tamarix* invasion in the riparian ecosystem of the Western Cape Province of South Africa, where it predominantly occurs. Random Forest (RF) and Support Vector Machine (SVM) were tested to classify *Tamarix* and other land-cover types. Sentinel 2 data and Landsat OLI earth observation data were used to map the current and the temporal exotic *Tamarix* distribution between 2007 and 2018, respectively. This included mapping the current and the multi-temporal *Tamarix* extent of invasion using the multi-spectral sensors Sentinel 2 and Landsat 5 and 8, respectively. Sentinel 2 was able to detect and discriminate the exotic *Tamarix* spp invasion using RF and SVM algorithms. The Random Forest classification achieved an overall accuracy of 87.83% and kappa of 0.85, while SVM achieved an overall accuracy of 86.31% and kappa of 0.83. Multi-temporal Landsat data was able to map the current and previous extent of exotic *Tamarix* invasion for the period between 2007 and 2018. Six land-cover types were classified using SVM. The overall accuracies achieved for 2007, 2014 and 2018 were 87.66%, 91.10%, and 90.62% respectively, and the kappa were 0.85, 0.89, and 0.88, respectively. It was found that the exotic *Tamarix* invasion increased from 284.67 ha to 647.10 ha in De Rust area, 74.70 ha to 97.29 ha in Leeu Gamka and 215.01 ha to 544.41 ha in Prince Albert region in a period of 11 years. Sentinel 2 and Landsat data have shown the potential to be used in *Tamarix* mapping. The results obtained in this study would help in implementation of conservation and rehabilitation plans.

Declaration

I, Thabiso Richard Kekana, declare that:

1. The research reported in this report, except where otherwise indicated, is my original research.
2. This report has not been submitted for any degree or examination at any other university.
3. This report does not contain other persons' data, pictures, graphs, or other information, unless specifically acknowledged as being sourced from other persons.
4. This report does not contain other persons' writing unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted then they are reference in bibliography.

Signed on the 30 day of May 2019 at the University of the Witwatersrand

Signature: 

Acknowledgement

I firstly would like to thank God Almighty for his grace and blessing me throughout my studies. I thank my wife Emelda, my kids Tshego, Remo and Olerato for their motivation, support and patience when I was not there for them due to study commitment. My mother Merriam, brothers Mathews and Robert and my sisters Lebo and Elizabeth I thank you for your support and words of encourage when I felt like quitting. I send a special dedication to my friend Mopeli Makafane and his wife Kgaugelo for their encouragement and motivation for me to pursue my studies.

My gratitude also goes to Dr Elhadi Adam and Dr Solomon Newete for taking me under their wings and mentored me. I appreciate their constructive criticism as this has built me to be an independent researcher, and has encouraged me to work hard and improved my scientific writing.

I would also like to extend my special gratitude to the National Research Foundation (South Africa) for the financial support during the field data collection under the supervision of Dr Solomon Newete (NRF Grant No.,:114345), the Exelis Visual Information Solutions team for providing ENVI software and ESRI Incorporated team for the ArcGIS software where pre- and post-processing were conducted And the R Development Team and for their Rstudio software that was used for image classification in this study.

List of Figures

Figure 1**.** *Tamarix* and its hybrids distribution in Southern Africa. Adopted from Marlin et al.(2017) 3
Figure 2.A Landsat image demonstrating the geographic location of the three study sites (Leeu-Gamka, Prince Albert and De Rust). 12
Figure 3.Random Forest tuning using the 10-cross validation method. Random Forest parameter (*ntree* and *mtry)* optimised andthe OOB sample was used to determine the classification error. 22
Figure 4.Support Vector Machine (SVM)tuning using the cross validation method. The sample error was used to determine the classification error. 23
Figure 5.Mapping current *Tamarix*distribution and other land-cover types using Sentinel 2 data, (a) Random Forest classifier and (b) Support Vector Machine classifier in the Western Cape Province. 24
Figure 6. The measure of the Sentinel 2 (10m spatial resolution) bands variable importance in RF classification process. 25
Figure 7. The relationship between each individual *Tamarix* spp, other Woody vegetation and land cover class and the RF variable importance..... 26
Figure 8.Monitoring *Tamarix* invasion in Olifant River (near De Rust region) using Landsat to classify the land-use and land-cover changes for the period (a) 2007 (b) 2014 and (c) 2018. 32
Figure 9.Monitoring *Tamarix* invasion in Leeu-Gamka using Landsat to classify (a). 2007 land-use and land-cover (b).2014 land-use and land-cover and (c) 2018 land-use and land-cover..... 33
Figure 10.Monitoring *Tamarix* invasion around Prince Albert using Landsat data to classify the land-use and land-cover change for (a) 2007 (b) 2014 and (c) 2018. 34

List of Table

Table 1.Previous successful studies of invasive species mapping using remote sensing 11
Table 2.Sentinel 2 sensor properties and date of Sentinel 2 data acquisition 13
Table 3. Summary of Landsat 8 sensor properties and acquisition dates of Landsat 8 data 14
Table 4.Summary of Landsat 5 TM sensor properties and Landsat 5 data acquisition date. 15
Table 5. Reference data for LULC classes. 17
Table 6. Confusion matrix for validating 2018 classified Sentinel 2 image. SVM accuracies User (UA), Overall (OA), and Producer (PA). 27
Table 7. Confusion matrix for validating 2018 land use and land cover (LULC) map. The RF accuracies: Overall (OA), Producer (PA) and User (UA) were determined from Sentinel 2 image. 28
Table 8.McNemar test showing misclassified pixels and correctly classified pixels by RF and SVM. 28
Table 9.Area and percentages for Land cover class in Prince Albert study site obtained by SVM and RF classifiers. 29
Table 10. Area and percentages for Land cover class in Olifant River (De Rust region) study site obtained by SVM and RF classifiers. 30
Table 11.Area and percentages for Land cover class in Leeu-Gamka study site obtained by SVM and RF classifiers. 30
Table 12. Optimisation of Support Vector Machine (SVM) parameters (*gamma* and *cost*) using radial kernel on Landsat images. 31
Table 13.Confusion matrix for validating 2007 Land-use Land-cover map. The Support Vector Machine accuracies: Overall (OA), Producer (PA) and User (UA) 35
Table 14.Confusion matrix for validating 2014 Land-use Land-cover map. The Support Vector Machine accuracies: Overall (OA), Producer (PA) and User (UA). 35
Table 15.Confusion matrix and for validating 2018 Land-use Land-cover map. The accuracies: Overall (OA), Producer (PA) and User (UA) 36
Table 16.TheArea and percentage of temporal *Tamarix* species and other Land cover classes in De Rust obtained from Support Vector Machine classifiers. 37

Table 17.The Area (ha) and percentage of temporal *Tamarix* species and other Land cover classes in Leeu-Gamka obtained from Support Vector Machine classifiers. 38
Table 18. The Area (ha) and percentage of temporal *Tamarix* species and other Land cover classes in Prince Albert region obtained from Support Vector Machine classifiers..... 38
Table 19.Change detection matrix of Land-use and Land-cover change in Olifants River (De Rust region) between 2007 and 2014 in hectares (ha)..... 39
Table 20. Change detection matrix of Land-use and Land-cover change in Olifants River (De Rust region) between 2007 and 2014 in percentages..... 40
Table 21.Change detection matrix of Land-use and Land-cover change in Olifants River (De Rust region) between 2014 and 2018 in hectares 41
Table 22.Change detection matrix of Land-use and Land-cover change in Olifants River (De Rust region) between 2014 and 2018 in percentages..... 41
Table 23. Change detection matrix of Land-use and Land-cover change in Leeu Gamka region between 2007 and 2014 in hectares. 42
Table 24.Change detection matrix of Land-use and Land-cover change in LeeuGamka region between 2007 and 2014 in percentages 43
Table 25.Change detection matrix of Land-use and Land-cover change in Leeu Gamka region between 2014 and 2018 in hectares. 44
Table 26. Change detection matrix of Land-use and Land-cover change in Leeu Gamka region between 2014 and 2018 in percentages. 44
Table 27.Change detection matrix of Land-use and Land-cover change in Prince Albert region between 2007 and 2014 in hectares 45
Table 28. Change detection matrix of Land-use and Land-cover change in Prince Albert region between 2007 and 2014 in percentages. 46
Table 29. Change detection matrix of Land-use and Land-cover change in Prince Albert region between 2014 and 2018 in hectares. 46

Table of Contents

Content	Page
CHAPTER ONE	2
1. General Introduction	2
1.1 Background	2
1.2 Rationale of the study	4
1.3 Research Aims and Objectives	6
CHAPTER TWO	**7**
2 Literature Review	**7**
2.1 *Tamarix* biology and description	7
2.2 Control of invasive *Tamarix* species	8
2.3 Negative Impact of *Tamarix* on Environment	9
2.4 Remote Sensing for mapping *Tamarix*	10
CHAPTER THREE	12
3 Methodology and Material	12
3.1 Study Area	12
3.2 Data acquisition	13
3.3 Remote Sensing Preprocessing	15
3.4 Reference Data	16
3.4.1 Reference data collected for mapping current Tamarix distribution	16
3.4.2 Reference data collected for multi temporal Landsat classification	17
3.5 Image Classification	17
3.5.1 Random Forest (RF)	18
3.5.2 Support Vector Machine (SVM)	18
3.6 Accuracy Assessment	19
3.7 Change Detection	20
CHAPTER FOUR	22
4 Results	22

4.1 Mapping the current distribution of *Tamarix*and other co-existing species using Sentinel 2 22
4.1.1 Random Forest tuning 22
4.1.2 Support Vector Machine tuning 23
4.1.3 The performance of SVM and RF in mapping *Tamarix*spp 23
4.1.4 Accuracy Assessment 26
4.2 Monitoring exotic *Tamarix* invasion using Landsat multi-temporal data 31
4.2.1 Support Vector Machine parameters tuning 31
4.2.2 Image Classification 31
4.2.3 Accuracy Assessment 34
4.3 Monitoring and quantifying the invasion spread of *Tamarix* using change detection techniques 36
4.3.1 Change Detection matrix 39
4.3.1.1 Change detection matrices of De Rust region (Olifant River) 39
4.1.3.2 Change detection matrices of Leeu Gamka region 42
4.1.3.3 Change detection matrices of Prince Albert region 45
CHAPTER FIVE 48
5 Discussions and conclusion 48
5.1 Discussions 48
5.2 Conclusion 51

CHAPTER ONE

1. General Introduction

1.1 Background

Alien invasive species are plants that occur outside their country of origin and have the ability to invade the native species habitat by outgrowing and replacing them (Milton, 2004). South Africa has a long history of problems with invasive alien species introduction and intervention management (Mcconnachie et al., 2011; Milton 2004). They are a major threat to the ecosystem, biodiversity and the economic livelihoods of the local communities (Richardson and Van Wilgen, 2004). For example, *Tamarix* is one of the most problematic invasive species in South Africa and the United State of America (Marlin et al. 2017).

Tamarix is an old world genus in the family of Tamaricaceae which is native to China, Indian and Mediterranean region, North and Southern Africa (Heywood et al. 2007). There are three *Tamarix* species found in South Africa and these are the two exotic *Tamarix ramosissima* and *Tamarix chinensis;* and the native *Tamarix usneoides* (Mayonde et al. 2015; Newete et al. 2019). *Tamarix ramosissima* is native to central Asia, while *T.chinensis* is native to China and Japan (Glen 2002; Heywood et al. 2007). Both exotic *Tamarix* species are declared invasive category 1 weeds according to the Conversation of Agricultural Resources Act, (ACT 43 of 1983) (Henderson 2001). These exotic *Tamarix* species are also found to hybridize with the native *T. usneoides* (Mayonde et al. 2015). *Tamarix* species are distributed in the semi-desert and Karoo areas of the south-western region of South Africa, in the Northern Cape, Western Cape and Eastern Cape provinces as illustrated in Figure 1.

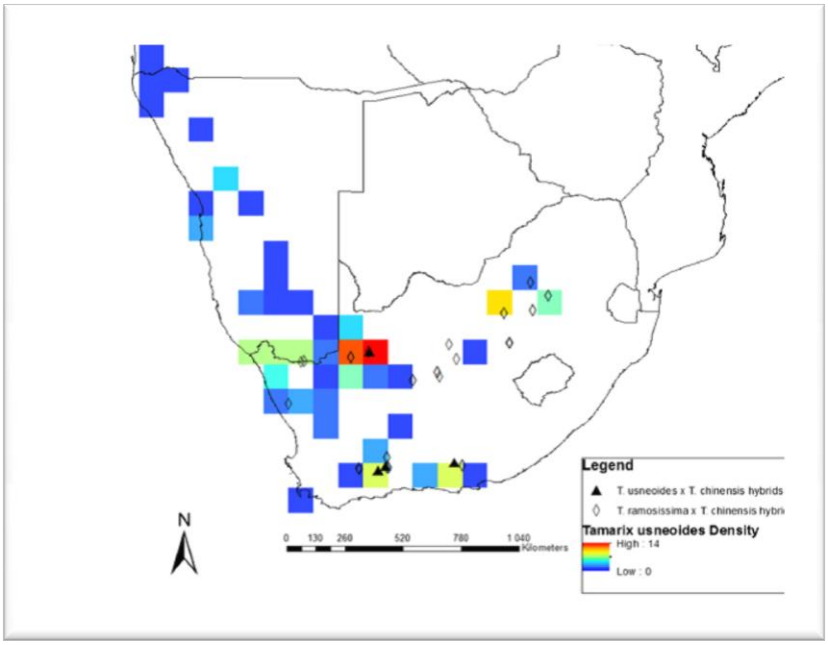


Figure 1. *Tamarix* and its hybrids distribution in Southern Africa. Adopted from Marlin et al. (2017).

Tamarix are used for erosion control, wind break, shade, ornament and for apiculture (DiTomaso 1998; Henderson 2001; Baum 1978). *Tamarix* species are used for phytoremediation in gold mines across South Africa, particularly in the AngloGold Ashanti (Vaal River) gold mine in North West Province, although this has changed strictly to the native *T. usneoides* at least since 2005 (Mayonde et al. 2015). The plant is however, known to alter the physical and chemical condition of the environment and this includes trapping sediments and reducing river channels (Blackburn et al. 1982), reducing ground water due to its excessive water consumption as it has deep roots (McDaniel et al. 2004) and increases soil salinity by excreting salts from salt glands (Brotherson and Field 1987). Furthermore, the salt deposition can limit under canopy species and its flammable litter enhances wildfires (Busch 1995).

Mapping the extent and distribution of *Tamarix* species helps to determine the hotspots for management intervention and prioritize resource allocation. Traditional vegetation mapping is conducted in either quadrant or transect methods where plant information such as plant species, stem diameter and height are recorded in each plot (Gillison 2006). However, these methods are not only time consuming and labour intensive, but also are hampered by inaccessible terrains and if often less feasible for large scale studies (Hoshino et al. 2012). The method might also require

laboratory analysis to predict biochemical and biophysical properties of vegetation, where it is not possible to identify species morphologically (Adam et al. 2012). However, Remote Sensing techniques address such limitations of traditional survey method by giving timely, cost effective and accurate vegetation mapping results. Among these is hyperspectral data, which provides high spectral resolution. However, because of its exorbitant cost it remains highly commercialized with little or no access to data for research. This has promoted the advent of multispectral sensors which offer high spectral, temporal and spatial resolutions at affordable cost or available free of charge and have proven successful in mapping vegetation at species level with accuracy levels ranging from 60% to 90 % (Table 1). Among these are Sentinel 2 and the Landsat series, which are available free of charge at the U.S. Geological Survey website (USGS website). Landsat data offer good temporal resolution which dates back to 1972, while Sentinel 2 with 13 bands of spectral resolution offers high spatial resolution (10mx10m) of up to 10m. The suitability of mapping *Tamarix* species distribution using Sentinel 2 imagery has not been investigated in South Africa and therefore, this study investigated the potential of these multi-spectral sensors in mapping *Tamarix* species accurately.

1.2 **Rationale of the study**

The management of invasive plant species cost South Africa an estimated ZAR 6.5 billion annually (De Lange and van Wilgen, 2010). Exotic *Tamarix* are declared category 1b weed according to the Conversation of Agricultural Resources Act (CARA), (ACT 43 of 1983) in South Africa (Newete et al. 2019). Regulation 15A of the Act states that Category 1 weeds, like exotic *Tamarix* species are strictly prohibited to propagate, or maintain the weed, sell or import (Henderson 2001a).

The two exotic *Tamarix* species (*T. ramosissima* and *T. chinensis*) that occurs in South Africa hybridizes with native *T.usneoides* (Mayonde et al. 2015; Newete et al. 2019). This will have a negative impact on potential bio-control programmes in the country, because of the presence of the native *T.usneoides*, which might be exclusively attacked if the proposed *Tamarix* beetle is released in South Africa (Marlin et al. 2017).

Tamarix trees can grow up to 4 m in a year (Gary 1960) and they can survive floods for 98 days when its root crown is submerged and up to 70 days when it is totally submerged (Warren and Turner 1975; Mortenson et al. 2012). In addition, the seedling can survive inundation for up to a month (Horton et al. 1960).

Tamarix spp alters the physical and chemical conditions of its environment by narrowing river channels and increasing flooding incidences (Mayonde et al. 2016; Brotherson and Field 1987; Blackburn 1982). It has a deep root system that can grow up to 30 meters long and can access underground water (Brotherson and Field 1987). The Western Cape has experienced drought in recent years and since *Tamarix* are phreatophytes and live long, they have the potential to lower ground water (Brotherson and Field 1987; LeMaitre et al. 2013). In addition, invasive *Tamarix* use twice more water than other co-existing species in South Africa (LeMaitre et al. 2013). The ability of *Tamarix* to invade disturbed environments and resprout after physical and mechanical control (Brotherson and Field 1987; Newete et. al 2019) would have negative impacts on the karoo biome and its riparian zones.

Traditional vegetation mapping is laborious, time consuming, expensive and some areas are inaccessible. However, Remote Sensing techniques can be an alternative to addresses the limitations of traditional vegetation mapping by giving timely, cost effective and accurate results. Therefore, remote sensing methods can be used for effective mapping of *Tamarix* spatial distribution at large scale (provincial) to facilitate appropriate management decision making and resource allocation accordingly.

High spatial resolution Quickbird data was used to discriminate between cottonwood and *T. ramosissima* and their finding was that pixel based classification showed that spectral bands alone were not sufficient to discriminate between *Tamarix* and cottonwood trees, because of their similar foliage property. However, semi-object based classification method improved accuracy as unique cottonwood tree crowns was discriminated from *Tamarix* (Ji and Wang, 2015). Hamada et al. (2007) reported object-based parallelpiped classifier with four narrow wavebands of airborne hyperspectral image selected through statistics separability yielded the most accurate and reliable *Tamarix* classification. Yet challenges were that results varied by scene, input data

type, classifier and minimum patch size. Thus, this shows that hyperspectral data not only are expensive and unaffordable for ordinary research programmes, but also are subject to high data dimensionality (Hamada et al. 2007). Few studies have reported on machine learning classification methods in *Tamarix* mapping. Thus, this study explored the use of this classification method to classify between the exotic *Tamarix* species and other co-occurring plants in the Western Cape Province. For instance, Carter et al. (2009) used Maximum likelihood classifier and NDVI in mapping of *Tamarix* at Colorado river (USA) using 40% for training the classifier and 60% for validating classifier and the result achieved were between 80-91%. Evangelista et al. (2009) also tested the Maximum Entropy model (Maxent) and associated vegetation indices in discriminating *Tamarix* from co-existing species using single scene and time series analyses of Landsat 7 data. The best model was time series analysis with all the variables fitted and the overall accuracy achieved was 90%.

1.3 Research Aims and Objectives

The main aim of this study was to map the spatial and temporal extent of the exotic *Tamarix* species in the Western Cape using Sentinel 2 and Landsat earth observation data.

The specific and main objectives of the study were to:

- Investigate the use of Sentinel 2 multispectral sensor for mapping the current distribution of *Tamarix* species and other co-existent land use and land cover types in riparian ecosystem.
- Monitor the extent of *Tamarix* species invasion from 2007-2018 (11 years) using multi-temporal Landsat data.
- Test the performance of Random Forest and Support Vector Machine classifiers in mapping *Tamarix* species effectively.

CHAPTER TWO

2 Literature Review

2.1 *Tamarix* biology and description

Tamarix are evergreen shrubs with feathery branches, the shrub grows 3-6m long and predominantly grows along sandy river banks and river beds (Henderson 2001a). *Tamarix* is a deep-rooted plant capable of growing up to 30 meters deep (Baum 1978) and is tolerant to extreme salt concentrations and soil salinity of up to 50000ppm (Brock 1984). The leaf-salt gland makes it possible for the plant to survive extreme salt conditions by excreting excess salt taken up through the leaf glands (Warren and Turner 1975; Campbell and Strong 1964; Wilson et al. 2017; Zhang et al. 2018). Berry (1970) demonstrated that *Tamarix* can tolerate salty environments and salts and micronutrients are excreted by salt glands. In addition, the composition of the salts is Sodium(Na), Potassium (K), Magnesium (Mg), Calcium (Ca),Chloride (Cl), Nitrates (NO_3), Bicarbonate (HCO_3), and Sulphide (S) and the nutrients are Boron(B), Manganese (Mn), Copper (Cu), Zinc (Zn), and Molybdenum (Mo). However, the salt and nutrient composition excreted by *Tamarix* depend on the root environment (Berry 1970). The three *Tamarix* species that occur in South Africa are the native *Tamarix usneoides* and the two exotic *T. ramosissima* and *T.chinensis*. The native *T.usneoides* is a perennial evergreen shrub with short, congested and scaly branches (Mayonde et al. 2015). It has pale grey-green vaginated leaves, with whitish flowers and the leafy branches usually extend from ground level (Henderson 2001; Mayonde et al. 2015). *Tamarix ramosissima* is commonly known as pink tamarisk. It is an evergreen shrub with reddish-brown bark, and has deep green; bluish-green or greyish sessile leaves which are minute and scale-like (Henderson 2001; Baum 1978). Fruits are 3-4mm papery capsules while flowers are pale to purplish in colour. The flower occurs in racemes 15-70mm long, at the end of thin, long twigs and flowers in late spring and early summer (Baum 1978). Pink tamarisk has obovate petals and unevenly toothed sepals (Henderson 2001a; Baum 1978). *Tamarix chinensis* commonly known as Chinese tamarisk, has a similar characteristics as *T. ramosissima* except that *T. chinensis* has ovate petals and entire sepals which are 0.75-1.25 mm (Henderson 2001a; Baum 1978). It is also an evergreen plant with blackish or dark brown bark.

2.2 Control of invasive *Tamarix* species

The management of *Tamarix* spp depends on three methods which are mechanical, chemical and bio-control methods. Chemical method is not only expensive but also it is not environmentally safe and feasible and is often subjected to regulation restriction. Duncan and McDaniel (1998) have used imazapyr 100 SL (CHOPPER) to control *Tamarix* and the mortality rate was 90%. However, the drawback of imazapyr100 SL (CHOPPER) is that it is a non-selective herbicide that can harm other vegetation species. The efficiency of the method depends on the application of the herbicide on stumps, which necessitate mechanical cutting (DiTomaso et al. 2013; Stevens and Walker 1998). This technique has been used in the South Africa, Australia and USA.

Biocontrol is considered the ideal control method of exotic *Tamarix* species since it is effective, sustainable and economically feasible and the method has already proven great success in the United State of America, where a defoliating *Diorhabda* beetles imported from China were used (Marlin et al. 2017). There are five different *Diorhabda* species known to attack the genus *Tamarix* and these are *D. sublineata, D. carinulata, D. elongata, D. medirionalis and D. carinata* (DiTomaso et al. 2013). Salt cedar leaf beetle (*Diorhabda carinulata*) from China has made significant impact in Nevada, Utah, Colorado and Wyoming regions of the US. *Diorhabda elongata* and *D.carinulata* were successful in suppressing the *Tamarix* invasion in the Great Basin Desert of the USA (DiTomaso et al. 2013) and are, therefore, potential bio-control agents for South Africa, where they are currently under quarantine test (Newete et al., 2019).

Mechanical control technique involves root plowing and stem cutting of the shrub. Root plowing is done by cutting the tamarisk root crowns below the soil surface. *Tamarix* can resprout from the root plowing and cutting of the stem (Brock 1984; Brotherson and Field 1987) and also this can spread the seeds further. Combined methods of stem cutting and chemical control were reported to be successful and up to 100% *Tamarix* mortality was achieved. Furthermore, when Picloram (Grazon) and dichloben (Casoron 56) were applied on the root stumps after cutting (Hollingsworth et al. 1979). This method is effective, but expensive and laborious to conduct. McDaniel and Taylor (2003) found that the cost for chemical control to be about ZAR 4086.49 per hectare and that of physical control to be ZAR 21212.21 per hectare, while seedling disking cost to be about ZAR 707.56 (Smith et al. 2002). Seedling disking involves the removal of the

Tamarix seedling before the root system can fully develop. Fire prescription to control exotic *Tamarix* species not only ineffective, but are also tolerant to fire damage and often burning enhances rapid resprout and plant vigour to spread further (DiTomaso et al. 2013; Busch 1995), while other co-existing species are affected and destroyed by fire.

2.3 Negative Impact of *Tamarix* on Environment

Richardson et al. (2001) described alien invasive species as taxa whose presence in a landscape is either due to accidentally or intentionally introductions by humans as garden ornaments, timber etc. *Tamarix* has a huge reproduction potential, a mature *Tamarix* can produce up to 500000 seeds annually and the seeds are small with a tuft of hair at apex to aid in wind dispersal (McDaniel et al. 2004; Brock 1984; Brotherson and Field 1987). The plant is capable of reproducing sexually either by self or cross pollination of seeds and it is also capable of reproducing vegetatively (asexually) after burning and cutting (Brotherson and Field 1987). Due to this trait the *Tamarix* often has competitive superiority to other co-existing species (Marlin et al. 2017). The exotic *Tamarix* species out-compete the native species for light and water resources owing to its outstanding characteristics which includes deep roots, salt tolerance, high rate of reproduction and seed vigour (Warren and Turner 1975; Baum 1978; Brock 1984; Brotherson and Field 1987). Busch and Smith (1995) confirmed that *Tamarix* had higher concentration of sodium (Na:K ratio = 1.87) than the native co-existing species (*Tessaria sericea*) which had Na:K ratio of 1.56 when they performed foliar elemental analyses. Furthermore, they found *Tamarix* to have low osmotic properties than native co-existing species in Colorado River and previous dominant species (*Salix* and *Populus* spp) were approaching extinction in some parts of riparian ecosystem, due to the saline environment. The root system makes the species thrive in arid and semi-arid conditions (Marlin et al. 2017). It traps alluvial sediments and that was evident on the Brazos River in Texas, where 71% of the river channel's width was reduced over a period of 40 years and flooding resulted due to *Tamarix* invasion (Blackburn et al. 1982). The deposition of salt below the *Tamarix* canopies as a result of direct dew droplets from leaf-salt glands or leaf decomposition inhibits the growth of other species in the vicinity (Marlin et al. 2017).

2.4 Remote Sensing for mapping *Tamarix*

Sabins (1996) defined remote sensing as "the science of acquiring, processing and interpreting images that record the interaction between electromagnetic energy and matter". Remote sensing has the capability to gather data about the earth's features regularly at various scale (local to global) and archived for decades for future use, allowing monitoring of temporal changes in the environment (Xie et al. 2008). Thus, remote sensing techniques has enabled mapping and monitoring of vegetation at various scales (Xie et al. 2008). Moreover, it is an alternative to traditional vegetation mapping which is time consuming, and labour intensive (Hoshino et al. 2012). In addition, remote sensing can effectively be used to map and monitor vegetation in areas that are inaccessible due to political boundaries and rough terrains and has large area coverage (Langley et al. 2001). Herderson (2001) used 'degree quarter method' to survey *Tamarix* and other alien invasive species in South Africa. It is a very coarse method of surveying and estimating tree distribution (Henderson, 2001a).

Tamarix flowers in late spring and summer and they have distinctive white to pink flowers. The panicle inflorescence and thicket makes *Tamarix* to be spectral different from other tree species in the growing season (Hamada et al. 2007 ; Ji and Wang 2016). The plant phenology and foliage is useful in mapping *Tamarix* and it was used by Ji and Wang (2015) and the challenge was that the phenology only lasted few weeks. High spatial resolution images like Quickbird and Rapideye are suitable for mapping *Tamarix*, but the drawback is that they are not freely available. However, there are freely available medium to coarse spatial resolution sensors such as Landsat data and Sentinel 2. Table 1 summarises some studies conducted on invasive species and the remote sensing techniques used.

Table 1.Previous successful studies of invasive species mapping using remote sensing

Plant Species	Country	Imagery	Classification	Accuracy	Reference
Posidonia oceanica and *Cymod oceanodosa*	Greece	Sentinel 2	Support Vector Machine, Random Forest and Maximum likelihood	90 %	(Traganos and Reinartz 2017)
Prosopis and *Vachellia* spp	Kenya	Sentinel 2 and Pléiades	OBIA Random Forest Classier	OOB accuracy 0.79 and OOB accuracy 0.83	(Ng et al. 2017)
Estuarine vegetation types.	St Lucia(South Africa)	Spot 6, Worldview And Rapid Eye	Supervised Maximum likelihood	SPOT-6 =64.3% - 77.9% Worldview =55% RapidEye=66.8%	(Lück-Vogel et al. 2016)
Spartina alterniflora	Sansha Bay (Southeast China)	SPOT6 satellite and UAV	NDVI for fractional vegetation cover(FCV)&aboveground biomass (AGB)	RMSE=0.108 (FCV) and 0.415(AGB), R2 =0.905(FCV) and 0.898(AGB)	(Zhou et al. 2018)
Tamarix	Forgotten River(USA)	Landsat ETM +	linear and nonlinear spectral mixture models	Non Linear Accuracy =73% Linear Accuracy =77%	(Silván-Cárdenas and Wang 2010)
Tamarix	Rio Grande River (USA)	Landsat TM	spectral angle mapper (SAM) algorithm	Overall= 0.93,producer's=0.75 accuracy	(Ji and Wang 2015)
Prosopis species	Northern Cape Province(RSA)	Worldview 2	Random Forest Algorithm	Overall=86. and kappa= 0.84	(Adam et al. 2017)

Remote Sensing has proved to be a reliable tool in mapping invasive species and it can be used for various applications using various readily available data.

CHAPTER THREE

3 Methodology and Material

3.1 Study Area

The study area is located east of the Western Cape Province in South Africa. The Province is the fourth largest province in South Africa (Figure 1), covering 11% of the country's land area. The total size is 129 462 km^2 (van Niekerk et al. 2016). Leeu-Gamka is a town located at 32.7639° S, and 21.9686° E, where the Leeu River enters the Gamka River (Figure 1). Prince Albert is located at 33.2167°S and 22°22.0333°E, and it is situated in Great Karoo 72 km north of Oudtshoorn, while De Rust is 35km west of Oudtshoorn and it is located at 33.4907° S, 22.5305° E. Major routes that passes this areas are N1 and N12 highways (https://www.karoo-information.co.za).

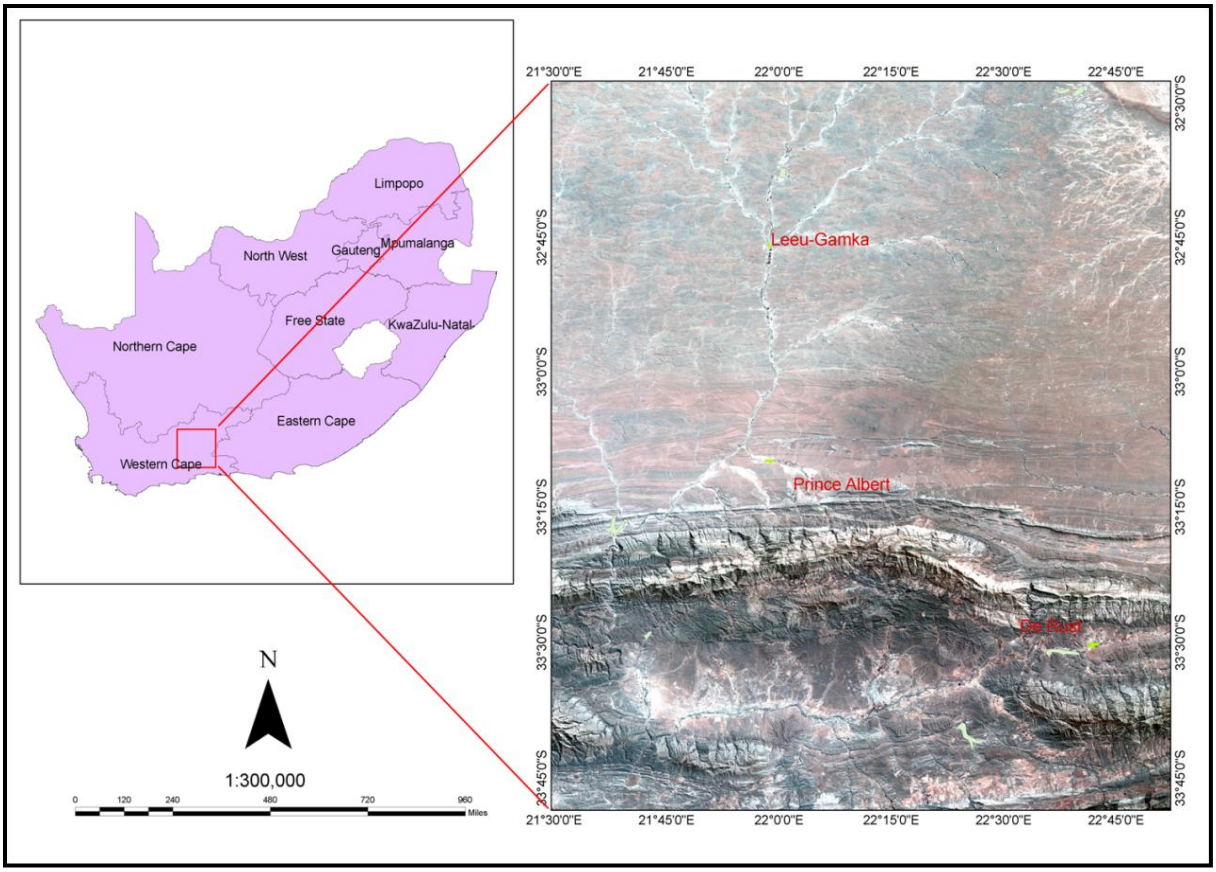


Figure 2. A Landsat image demonstrating the geographic location of the three study sites (Leeu-Gamka, Prince Albert and De Rust).

The exotic *Tamarix* spp are predominately abundant in these regions of the Country (Newete et al. 2019). The climate is arid and continental and also the average rainfall ranges from 50-400mm/year (Dean and Milton 2019). The study areas have a dry summer with maximum and minimum temperatures of 27°C and 15°C respectively and a moist winter with temperature dropping as low as to 3-5º C (Dean and Milton 2019).

3.2 Data acquisition

The two remote sensing data acquired for this study were Landsat satellite data which was used for the multi-temporal analysis and Sentinel 2 data for mapping the current distribution of *Tamarix*. Copernicus Sentinel 2 mission is optical earth observation mission and two satellites constellation obits the Earth at an altitude of 786 km (https://sentinel.esa.int). The sensor has a 290 km swath width, a 5day temporal resolution, and 13 spectral bands. The sensor is characterized by 13 bands which are the blue (442.7 - 492.5 nm), green (559.8 nm), red (664.6 nm), red edge (704.1–782.8 nm), near infrared (832.8 –945.1) and short wave infrared (1373.5 - 2202.4) sections of the electromagnetic spectrum (Malahlela et al. 2019). Sentinel 2 spatial resolutions are 10m, 20m and 60m (Table 2).

Table 2. Sentinel 2 sensor properties and date of Sentinel 2 data acquisition

Date of Acquisition	Bands	Wave length (nm)	Bandwidth (nm)	Spatial resolution (m)
18/11/09	Band 1 Coastal	442.7	21	60
	Band 2 Blue	492.4	66	10
	Band 3 Green	559.8	36	10
	Band 4 Red	664.6	31	10
	Band 5 Vegetation red edge	704.1	15	20
	Band 6 Vegetation red edge	740.5	15	20
	Band 7 Vegetation red edge	782.8	20	20
	Band 8 NIR	832.8	106	10
	Band 8a Narrow NIR	864.7	21	20
	Band 9 Water vapour/NIR	945.1	20	60
	Band 10 SWIR – Cirrus	1373.5	31	60
	Band 11 SWIR 1	1613.7	91	20
	Band 12 SWIR 2	2202.4	175	20

Landsat satellite data is the most common and popular data used to quantify and monitor changes on the earth surface and its efficiency has continuously been improved over the years. The development of Landsat sensors began with Multi-spectral Scanner sensor (MSS) then Thematic Mapper sensors, Enhanced Thematic Mapper Plus (ETM+) sensor, and The Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) (Phiri and Morgenroth 2017). These developments and advancements begun in the early 1970, which now offers high temporal resolution over a long period of time; however, they lack good spatial resolutions. Landsat data has been applied in quantifying various land-cover land-use change across the world and it had helped resource managers and researchers to understand the land-use land-cover dynamics, due to anthropogenic activities or climate change (Gavier-Pizarro et al. 2012; Islam et al. 2018; Mas 1999; Ng et al. 2016; Zhang et al. 2017).

Landsat 8 is an earth observer satellites launched in 11 February 2011 and the sensor orbit the Earth at altitude of 705 km (Phiri and Morgenroth 2017). The sensor has a temporal resolution of 16 days and 11 spectral bands ranging from 0.43 µm to 12.51 µm and spatial resolutions of 20, 60 and 100 meters (Table 3).

Table 3. Summary of Landsat 8 sensor properties and acquisition dates of Landsat 8 data

Acquisition date	Spectral Bands	Spectral range[µm]	Spatial Resolution[m]	Temporal Resolution[days]
14/11/20	Coastal	0.43-0.45	30	16
18/09/29	Blue	0.45-0.51	30	
	Green	0.53-0.59	30	
	Red	0.64-0.67	30	
	Near-Infrared	0.85-0.88	30	
	SWIR1	1.57-1.65	30	
	SWIR 2	2.11-2.29	30	
	Panchromatic	0.50-0.68	15	
	Cirrus	1.36-1.38	30	
	Band 10 TIRS 1	10.6 - 11.19	100	
	Band 11 TIRS 2	11.5 - 12.51	100	

Landsat 5 was launched on 1 March 1984 and carries two push broom Multi Spectral Scanner (MSS) and Thematic Mapper (TM) instruments. The sensor was orbiting the Earth like Landsat 8

at altitude of 705km before it was decommissioned on 25 June 2013 (Phiri and Morgenroth 2017). Thematic Mapper has a spatial resolution of 30m and 120m and 7 spectral bands (Table 4), while Multi Spectral Scanner has a spatial resolution of 57m x 79 m and is characterised by 4 spectral bands.

Table 4. Summary of Landsat 5 TM sensor properties and Landsat 5 data acquisition date.

Date of Acquisition	Landsat 5 Spectral Bands	Spectral range[μm]	Spatial Resolution[m]	Temporal Resolution(day)
02.01.2007	Blue	0.45 - 0.52	30	16
	Green	0.52 - 0.60	30	
	Red	0.63 - 0.69	30	
	Near-Infrared	0.76 - 0.90	30	
	Short wave infrared	1.55 - 1.75	30	
	Thermal	10.40 - 12.50	120	
	Mid-Infrared	2.08 - 2.35	30	

Landsat and Sentinel 2 images were downloaded from the United State Geological Survey (USGS) earth explorer website. Two Landsat 8 and one Landsat 5 images were acquired on the following respective dates: 29/09/2018, 20/11/2014 and 02/01/2007, and Sentinel images were acquired on 09/11/2018. Although, acquisition of multi-temporal data to monitor the changes in the distribution of *Tamarix* invasion was originally proposed at intervals of 5 years starting from 2005 due to data unavailability this was not possible.

3.3 Remote Sensing Preprocessing

The remote sensing data was pre-processed using the radiometric, atmospheric and geometric correction techniques to eliminate distortions and enhance the data acquired. The Environment for Visualizing Images software (ENVI 5.4) was used to perform these corrections on Landsat data. First the Landsat images were radiometric corrected using radiometric calibration and then Fast Line-Of-Sight Atmospheric Analysis of Hyperspectral Cubes (FLAASH) module in ENVI software was used to do atmospheric corrections.

Image registration is the process of overlaying two or more images of the same scene taken at different times by different sensors. Therefore, image registration was applied to Landsat images so that they can overlay on each other. The 2007 and 2014 images were co-registered to 2018 image, which was the base image. Image to image registration function from ENVI classic was used to perform the registration and the Root Mean Square Error (RMSE) of all images was below 0.5 pixels. RMSE of less than or equals to 0.5 pixels in image transformation was acceptable (Jensen, 1995).

Sentinel 2 data was atmospherically corrected using Sen2cor plugin in SNAP 6.0.5 software. Sen2cor is a prototype processor for generating and formatting Sentinel 2 level 2A from level 1C by performing atmospheric, terrain and citrus correction. And therefore it will create Bottom of Atmosphere image and terrain corrected image as a product (Mueller-Wilm et al. unpubl. data.).

3.4 Reference Data

3.4.1 Reference data collected for mapping current Tamarix distribution

Reference data was collected in November 2017 for Sentinel 2 and Landsat 8 images the *Tamarix* flowers mostly in late October and this makes the species distinguishable from other co-existing plants species, because of the colour of the flowers (Henderson, 2001). The samples collected were for Woody vegetation (Acacia karoo and others) and *Tamarix* land-cover types using Global Positioning System (GPS). A total of 86 GPS points were collected from *Tamarix* and other Woody vegetation from the respective sites of each study area. Additional reference data for other land use land cover types (Buildup area, Waterbody and Bareland and Low dense grassland) were obtained by visually interpreting the acquired Sentinel 2 and Landsat 8 images with the help of Google pro software. Therefore, the reference data was randomly divided into training (70%) and validation (30%) data (Table 5). Training data was utilised to train the supervised Random Forest and Support Vector Machine classifiers and validation data was used to construct a confusion matrix table.

3.4.2 Reference data collected for multi temporal Landsat classification

Reference data for historical Landsat were obtained by visually interpreting a 2007, 2014 and 2018 that were observed from Google pro software. Therefore the reference data was randomly split into training data (70%) and validation data (30%) (Table 5). Training data was utilized to teach the supervised Support Vector Machine classifier and validation data was used to create a confusion matrix table. From the constructed confusion matrix producer accuracy, overall accuracy, user accuracy and kappa achieved by the classifier were presented in a tabular form.

Table 5. Reference data for LULC classes.

	Landsat 8(2018)			Landsat 8(2014)			Landsat 5(2007)			Sentinel 2(2018)		
Class	Train	Test	Total	Train	Test	Total	Train	Test	Total	Train	Test	Total
Agriculture	103	44	147	103	44	147	84	35	119	96	40	136
Bareland	113	52	161	113	52	161	101	43	144	115	49	164
Buildup	75	31	106	75	31	106	94	40	134	89	37	126
Tamarix	166	71	237	138	59	197	73	30	103	148	63	211
WoodyVeg	94	94	134	94	94	134	142	60	202	114	48	162
Waterbody	75	31	106	75	31	106	66	27	93	63	26	89

3.5 Image Classification

Random Forest and SVM are supervised image classifiers and were used for image classification, which means the classifier can be trained to classify with the aid of training data and the performance of the classifier can be evaluated by constructing confusion matrix with reference of validating data (Sabins 1996). Reference data is randomly divided into training and validating data before classification process. In mapping the current *Tamarix* distribution Random Forest (RF) and SVM were used to classify Sentinel 2 image classification. SVM was used in mapping the multi-temporal distribution of *Tamarix* and to classify Landsat data. A brief explanation of the SVM and RF algorithms used in the study is provided in the subsections below.

3.5.1 Random Forest (RF)

Breiman (2001) developed Random Forest algorithm to improve the accuracy of regression and classification trees (CART). It can be applied as classifier or regression model and it is based on bagging operation where a number of multiple classification trees are derived from training data from random subset sample (Breiman 2001). This multiple classification trees are combined into a forest and each tree in the forest vote towards the allocation of the most constantly repeated class to the input data. Random Forest (RF) uses two parameters which make the classifier easier when they are optimised, making it better in dealing with outliers and noise associated with the classification algorithm (Breiman 2001). The samples which are not in bagging operation from the subset sample and were not used in unit voting of correct classification are named Out Of Bag (OOB) samples. The OOB sample is used for determining the classification error and determining variable importance. Random Forest uses two parameters which make the classifier easier to use once optimised and they improve accuracy (Cho et al. 2015), which is better in dealing with outliers and noise associated with the classification algorithm (Breiman 2001). The RF parameters are *mtry* is the number of variables to split nodes individual trees and *ntree* is the number of trees to be grown (Díaz-Uriarte and Alvarez de Andrés 2006).

Optimal combination for RF parameter (*mtry and ntree*) is found by using a 10 fold cross validation method based on the OOB error and the parameters were input for classification. The value of *mtry* parameter was mixed from 1 to 10, while *ntree* value was mixed from 500 to 10,000 with an interval of 1500. The 10 fold cross validation method produced a total of 900 combinations of *ntree* and *mtry* values.

3.5.2 Support Vector Machine (SVM)

SVM was employed to classify LULC type in the study area and the advantage of the algorithm is that it can be used as classification or regression models (Cortes and Vapnik, 1995). The algorithm was used to classify Sentinel 2 and multi temporal Landsat data. The other advantage of SVM is that it can successfully work with limited number of training sample dataset (Mountrakis et al. 2011). The limitation is however, the use of SVM for remote sensing image classification is the choice of kernels and the setting of other parameters, which can affect the classification accuracy (Bhavsar and Panchal 2012). The algorithm works by selecting training

sites to create hyperplanes and separate dataset into defined number of classes represented by training areas (Mouuntrakis et al. 2011). Support Vector Machine has been applied efficiently to deal with linear and nonlinear separable classes. Since some of classes are not linear separable, therefore SVM uses kernel function to search for non-linear hyperplane in multidimensional feature space when it is optimised (Mouuntrakis et al. 2011; Karatzoglou et al., 2006). There are four kernel types used in SVM classification, i.e. linear, polynomial, radial basis function (RBF) and sigmoid. So far radial basis function is the most used on remote sensed data (Adam et al. 2014). Two parameters for optimising the radial basis kernel are the kernel width '*gamma* (γ) and *cost* '*sigma (C)*'. In this study radial basis kernel method was applied and *gamma* and *sigma* were optimised using the 10 fold cross validation sampling method. The optimisation and classification were performed in RStudio software. The *cost 'sigma (C)*', controls the agreement between the decision boundary/hyperplane and correctly classifying training point and kernel width *'gamma (γ)'*defines how far the influence of training single site.

3.6 Accuracy Assessment

Validation dataset was used to test the classified maps developed after RF and SVM classification. The Confusion matrix table shows agreement between the classification result and reference imagery (ground pixels). A confusion matrix is adapted and employed to compute and calculate kappa statistics, producer accuracy, overall accuracy and user accuracy (Congalton and Green 2008). User's accuracy (UA) describes the amount of error omission by the user which shows probability that a pixel labelled as specific *Tamarix* and land cover types in the map is the actual class. While producer's accuracy describes the amount of errors of commission which shows that the specific *Tamarix* and land cover types on the ground are accurately classified. User Accuracy (UA) was calculated by dividing the correctly classified samples number of the respective land-cover class by its total number of verified samples belonging to the specific land-cover class. Error of commission was obtained by dividing the correctly classified samples by the total number of reference samples (Richards and Jia 2006). The percentage probability that the random selected pixel is correctly classified in Land use land cover thematic map is referred to as Overall Accuracy (OA). Therefore, the OA was calculated by dividing number of correctly classified pixels with the total number of pixels in specific land cover type. The kappa value is from 0 to 1, where 0 values shows there is negative correspondence between the reference image

and the classified image and positive correspondence between reference image (ground pixels) and classified image (classified pixel) is shown by 1 (Dorn et al. 2015). Mcnemar test was employed to test the performance difference of SVM and RF. Furthermore confusion matrices of the RF and SVM classifiers were used to calculate Mcnemar test which is a non-parametric test (de Leeuw et al. 2006). This test is based on the standardized normal test given as:

$$\mathbf{z} = \frac{f12 - f21}{\sqrt{f12 + f21}} \quad \text{Equation 1}$$

Where $f12$ represents number of pixels correctly by the second classifier, however, misclassified by the first classifier and $f21$ represents the number of pixel misclassified by the first classifier, but correctly classified by the second classifiers. At the significance level of 5% the difference in accuracy is concluded to be statistically significant when the Z value is greater than 1.96.

3.7 Change Detection

Change detection involves quantifying changes of an object or phenomena over time using frequent interval satellite data (Singh 1989). Numerous change detection techniques exist such as principal component analysis, vegetation indexes differencing, image rationing and post classification. However, these techniques depend on the skill and knowledge of analyst as to when and how to apply the optimal technique (Singh 1989). Change detection is a commonly used technique in remote sensing to detect land surface changes and ecosystem dynamics, such as urbanization, desertification and wildfire (Pontius et al. 2017). Different sensors are used to achieve such temporal changes in the environment. For instance, Xiaohan et al.(2015) used MODIS sensor to evaluate changes occurred in the spread of aquatic vegetation in Lake Taihu (China) during his 11-year study. Identification and understanding the long term nature of change in the use of land resources can provide insight to managers and scientist to evaluate the drivers and stressors of change, allowing proper management and planning (Kennedy et al. 2009). Information relating to land cover change can be vital to resource managers to monitor ecosystem service flow (Islam et al. 2018)

Image registration was applied to Landsat images so they can accurately quantify change in LULC. Images of 2007 and 2014 images were co-registered to 2018 image, which was the base image. Image to image registration function from ENVI classic was used to perform the registration and the Root Mean Square Error (RMSE) of all images was below 0.5 pixels. RMSE of less than or equals to 0.5 pixels in image transformation is acceptable (Jensen 1995).

Post classification technique was employed to detect the changes after employing SVM image classification techniques. Furthermore, the technique was employed due to its capability to provide a tabular matrix of change information between the multi-temporal images (Lu et al. 2004). *Change detection statistics* module in ENVI 5.4 software was applied to detect change of classes. In addition, the module provide matrix of changes between two classified images and reports changes between initial state class and final state class as pixel count, percentage and areas (www.harrisgeospatial.com). Change detection statistics matrices were produced between Landsat data of 2007-2014 and 2014-2018. This post classification technique denote changes that took place between the respective images and multi temporal land use land cover change can be noted and assessed.

CHAPTER FOUR

4 Results

4.1 Mapping the current distribution of *Tamarix*and other co-existing species using Sentinel 2

4.1.1 Random Forest tuning

Random Forest was optimised to accurately determine the best input parameters for classification of *Tamarix* and five land-cover classes using 10 cross validation. Results showed that merger of *mtry* value of 2 and *ntree* value of 8500 produced the lowest Out-of-Bag (OOB) error of 15.51%. On the other hand, the highest OOB error produced was 17.5% which resulted from the combination of *mtry* value of 10 and *ntree* value of 4500 (Figure 3). Thus, the combination of *mtry* value of 2 and *ntree* value of 8500 was selected and used.

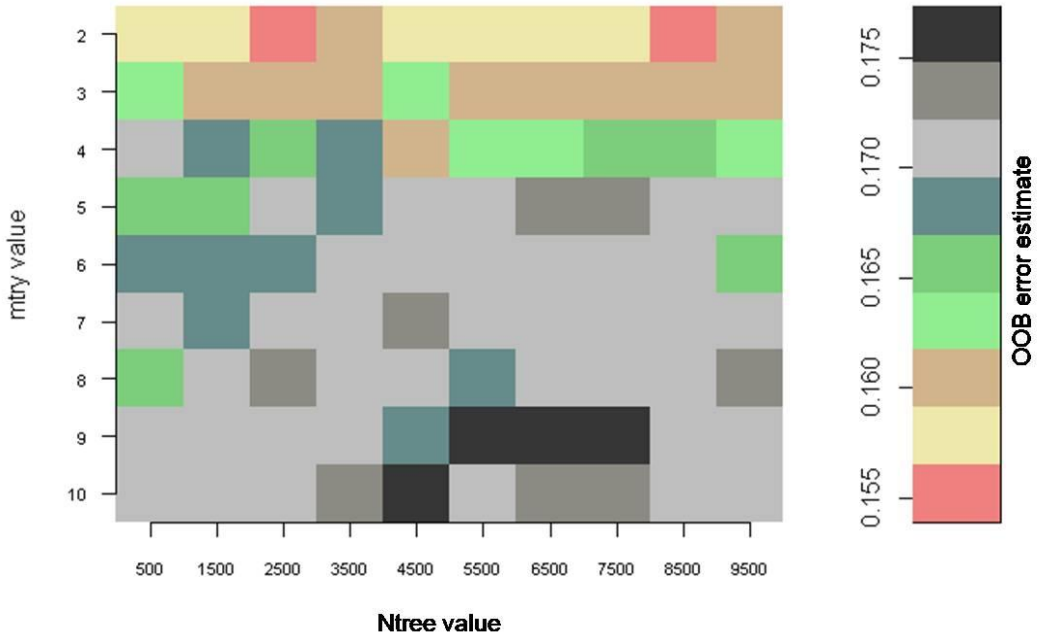


Figure 3. Random Forest tuning using the 10-cross validation method. Random Forest parameter (*ntree* and *mtry)* optimised andthe OOB sample was used to determine the classification error.

4.1.2 Support Vector Machine tuning

The radial kernel method was used and two parameters (*gamma* and *cost*) were optimised to accurately determine the best combination with the lowest error. The lowest error was produced by optimising *gamma* (γ) and *cost* (C). The combination of *gamma* and *cost* of values of 0.1 and 10 respectively, produced the lowest classification error of 15.35% (Figure 4).

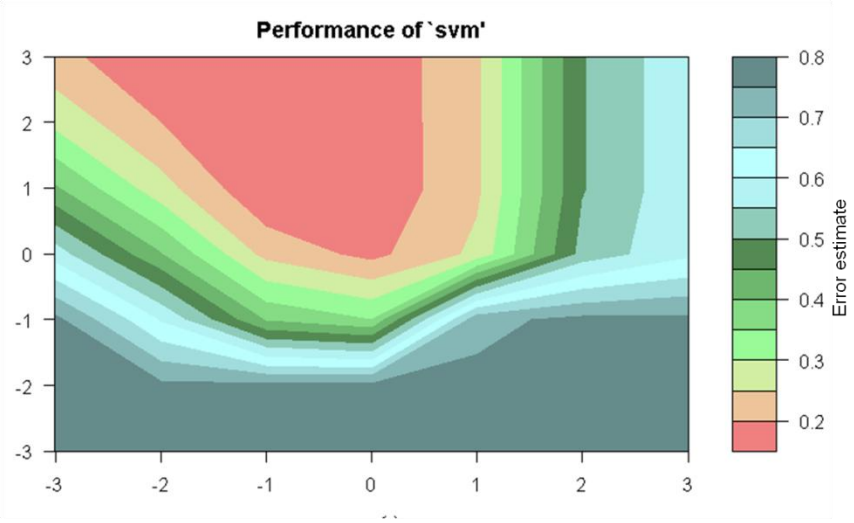


Figure 4. Support Vector Machine (SVM)tuning using the cross validation method. The sample error was used to determine the classification error.

4.1.3 The performance of SVM and RF in mapping *Tamarix*spp

The SVM and RF managed to classify *Tamarix* spp and the other land-cover and land use types (Figure 5). The dominant land-cover was Bareland followed by low-dense grassland and Woody vegetation in both Sentinel 2 and Landsat satellite images classified. *Tamarix* was dominant in low lying areas along the river channels, while other Woody vegetation, mainly dominated by *Acacia karoo,* was found on both low lying and high mountainous areas. A clear ecotone exists between Woody vegetation, *Tamarix* and Bareland and low-dense grassland (Figure 5). When Random Forest classifier was used Woody Vegetation, Agriculture and *Tamarix* land-cover (Figure 5a) where more notable than the same land-cover types when SVM classifier was used (Figure 5b). Nonetheless, the Buildup area was more notable when SVM classifier (Figure 5b) was used than when RF classifier was used (Figure 5a).

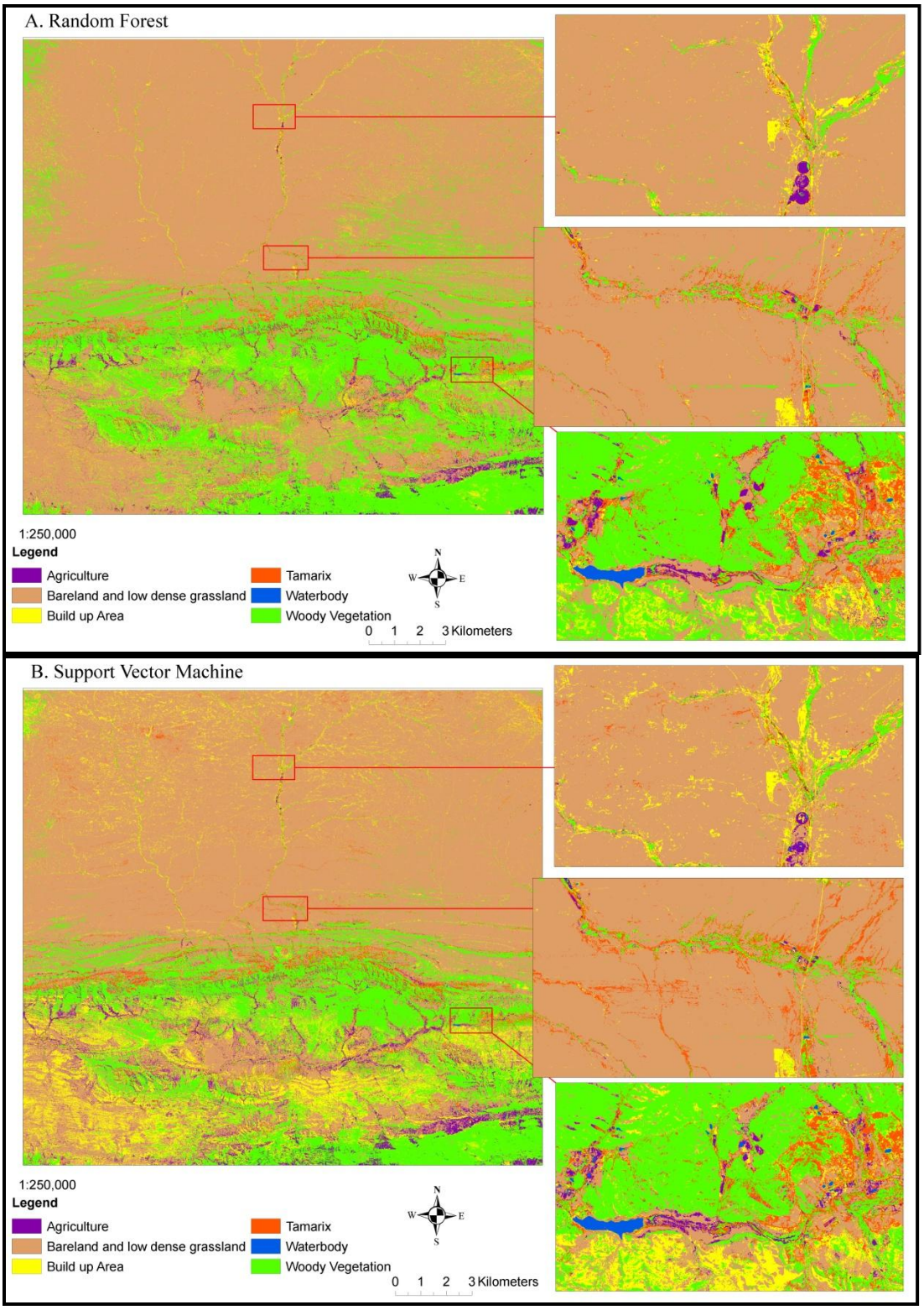


Figure 5. Mapping current *Tamarix*distribution and other land-cover types using Sentinel 2 data, (a) Random Forest classifier and (b) Support Vector Machine classifier in the Western Cape Province.

During the classification process RF provided variable importance which indicates the character of each band in image classification. The most significant bands were red, blue, green, SWIR 2 and red edge bands and they had the highest mean decrease in accuracy (Figure 6).

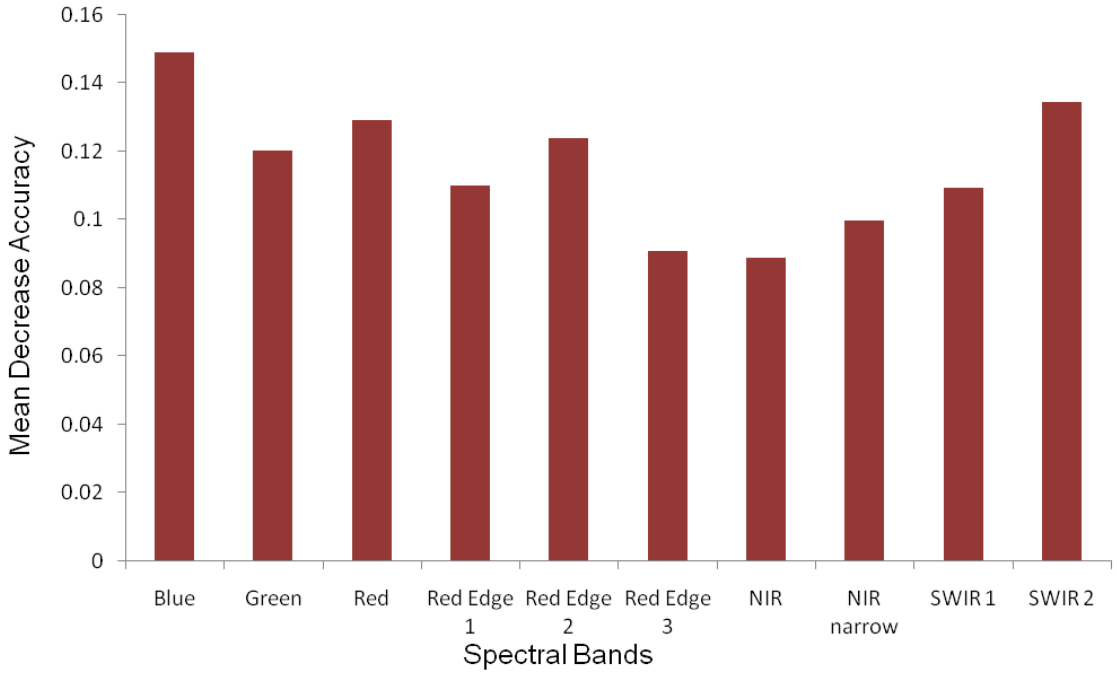


Figure 6. The measure of the Sentinel 2 (10m spatial resolution) bands variable importance in RF classification process.

Red, blue, red edge 1, red edge 2 and SWIR 2 bands were the important bands in classification of *Tamarix* and other vegetated land-cover types and SWIR 2, green and blue bands were important bands in classification of Waterbody (Figure 7). NIR short wave, Red edge 1, red edge 2 and green bands were important in classification of Bareland and low-dense grassland, while the important band in classification of Build-up area was Red edge 2 and red edge 3 bands (Figure 7).

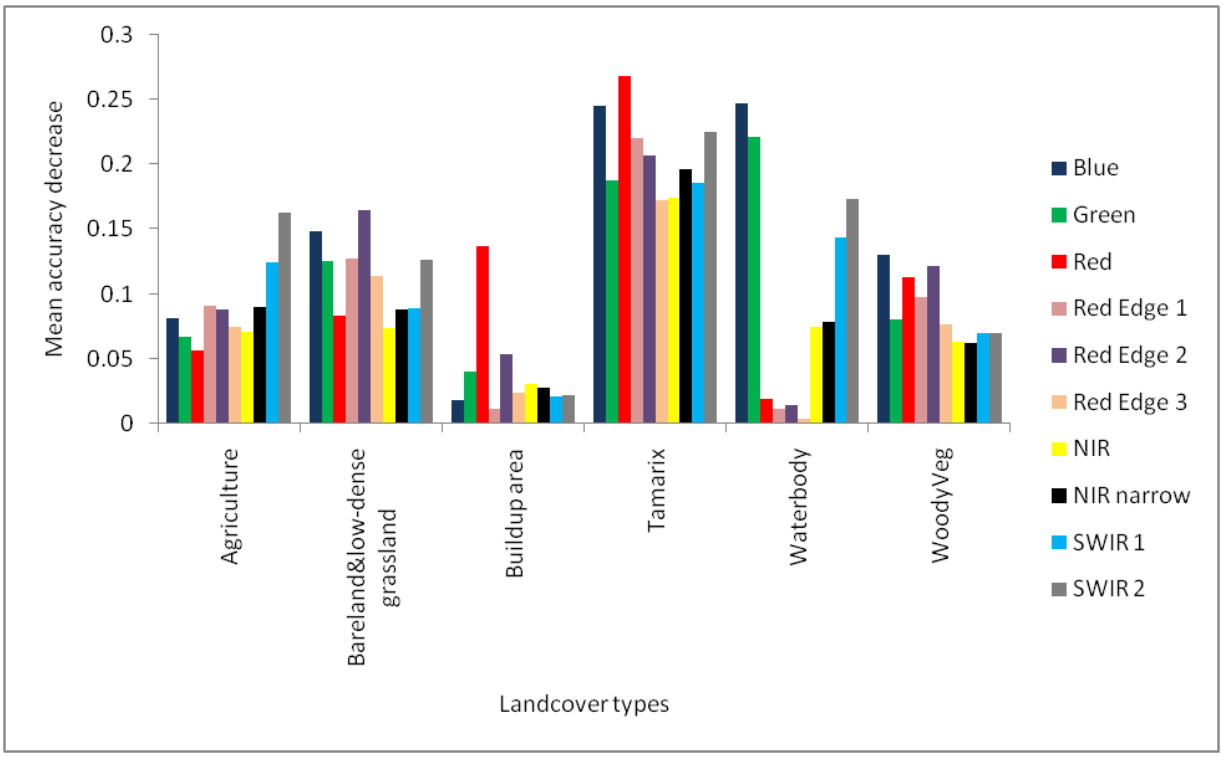


Figure 7. The relationship between each individual *Tamarix* spp, other Woody vegetation and land cover class and the RF variable importance.

4.1.4 Accuracy Assessment

The confusion matrix of SVM produced an Overall Accuracy (OA) of 86.31% and Kappa statistic value of 0.833 (Table 6). Spectral confusion was noted between Build-up area, Woody vegetation and Agriculture when Build-up area was classified, hence low producer and user accuracy of 75.68% and 82.35% were produced, respectively. Also confusion between *Tamarix* and Woody vegetation was noted when *Tamarix* was classified hence user accuracy of 81.2% was produced (Table 6).

Table 6. Confusion matrix for validating 2018 classified Sentinel 2 image. SVM accuracies User (UA), Overall (OA), and Producer (PA).

Class	A	BL	B	T	W	WV	Row total	UA
A	35	2	2	0	0	2	41	85.57
BL	3	40	2	0	0	0	45	88.89
B	1	3	28	2	0	0	34	82.35
T	0	2	3	60	0	8	73	81.2
W	1	0	0	0	26	0	27	96.3
WV	0	2	2	1	0	38	43	88.37
Total Column	40	49	37	63	26	48	263	
PA	87.5	81.63	75.68	95.24	100	79.17		
OA	86.31							
Kappa	0.833							

Note that A = Agriculture, BL = Bareland & low dense grassland, B = Build-up area, T = *Tamarix*, W = Waterbody and WV = Woody Vegetation

Random Forest produced a higher OA of 87.83% and kappa value of 0.8514 compared to SVM (Table 7). However, spectral confusion was also noted between Bareland and low-dense grassland, *Tamarix* and build-up area when build-up area was classified which produced lower PA of 67.57%. Spectral confusion in classifying Waterbody was not noted as it produced high PA and UA 100% for both UA and PA. The other spectral confusion was observed between Bareland and low-dense grassland, Agriculture and build-up area when Bareland and low dense grassland were classified and this produced low UA of 80.77%.

Table 7. Confusion matrix for validating 2018 land use and land cover (LULC) map. The RF accuracies: Overall (OA), Producer (PA) and User (UA) were determined from Sentinel 2 image**.**

Class	A	BL	B	T	W	WV	Row Total	UA
A	33	1	1	2	0	0	37	89.19
BL	4	42	6	0	0	0	52	80.77
B	1	2	25	0	0	1	29	86.21
T	0	2	3	61	0	3	69	88.41
W	0	0	0	0	26	0	26	100
WV	2	2	2	0	0	44	50	88
Column Total	40	49	37	63	26	48	263	
PA	82.5	85.71	67.57	96.83	100	91.67		
OA	87.83							
Kappa	0.8514							

Note that A = Agriculture, BL = Bareland& low dense grassland, B = Build-up area, T = *Tamarix*,W = Waterbody and WV = Woody Vegetation

Although both RF and SVM had different classification accuracies Mcnemar test showed that there was no significant difference at 5% significant level as the Z value was 0.75 which was less than 1.96 (Table 8).

Table 8. McNemar test showing misclassified pixels and correctly classified pixels by RF and SVM.

	Random Forest		
Support Vector Machine	Correctly classified	Misclassified	Total
Correct classified	215	12	227
Misclassified	16	20	36
Total	231	32	263
Z value	0.75		

Table 9 shows the area of the different land-cover type's in Prince Albert region calculated using SVM and RF classification algorithm. The dominated land-cover type was Bareland and low-dense grassland, while Agriculture and Waterbody were less dominant when SVM was used. *Tamarix* was found to cover 6.66% when the SVM classifier was used and 3.91% when RF was

used and the difference between the percentage areas was 2.75% (Table 9). Bareland and low-dense grassland was dominant, while Agriculture and Waterbody were less dominant when RF classifier was applied. All the land-cover types were classified similarly by both classifiers and showed about 0.1- 2% difference among them.

Table 9. Area and percentages for Land cover class in Prince Albert study site obtained by SVM and RF classifiers**.**

Class	SVM Area(ha)	%	RF Area(ha)	RF %	+/-(%)
Agriculture	96.48	0.90%	86.77	0.81%	0.09%
Bareland and low-dense grassland	9332.03	86.92%	9515.69	88.63%	-1.71%
Buildup area	240.58	2.24%	276.2	2.57%	-0.33%
Tamarix	715.46	6.66%	420.06	3.91%	2.75%
Waterbody	6.77	0.06%	2.29	0.02%	0.04%
Woody vegetation	344.56	3.21%	434.87	4.05%	-0.84%

The dominant land-cover types were Woody vegetation, Bareland and low-dense grassland, and *Tamarix* when both SVM and RF were used. However, when SVM was applied on Sentinel 2 data the dominant, Woody vegetation, Bareland and low-dense grassland, and *Tamarix* land-type covered the average area of 2345.51 ha (41.19%), 1507.95 (26.48%), 857.74 (15.06%), respectively. When Random Forest was used Woody Vegetation, Bareland and low-dense grassland, and *Tamarix* land-cover type they covered average area of 2419.01 ha (42.48%), 1488.4ha (26.14%) and 913.39 ha (16.04%) (Table 10). The less dominant land-cover types were Buildup area, Agriculture and Waterbody when both the SVM and RF were used. When SVM was used Agriculture covered the area of 288.56 ha (5.07%), Buildup area covered 461.29 ha (8.10%) and Waterbody covered 123.5 ha (2.17%) (Table10). The land-cover types were similarly classified by both classifiers and showed about 0.2% - 1.4 % difference in their performance. *Tamarix* was found to cover 15.06% of the Olifants River (near De Rust) area when SVM was used and 16.04% when RF was used; in addition the difference was 0.98%.

Table 10. Area and percentages for Land cover class in Olifant River (De Rust region) study site obtained by SVM and RF classifiers.

Class	SVM Area(ha)	%	RF Area(ha)	%	+/-(%)
Agriculture	308.34	5.42%	288.56	5.07%	0.35%
Bareland and low-dense grassland	1507.95	26.48%	1488.4	26.14%	0.34%
Buildup area	539.94	9.48%	461.29	8.10%	1.38%
Tamarix	857.74	15.06%	913.39	16.04%	-0.98%
Waterbody	134.67	2.37%	123.5	2.17%	0.20%
Woody vegetation	2345.51	41.19%	2419.01	42.48%	-1.29%

Agriculture land-cover type covers the average area of 50.1 ha (1.43%) and 69.07 ha (1.97%) when SVM and RF were used and they had a difference of 0.54%. Bareland and low-dense grassland covered an area of 2778.73 ha, Buildup area covered 384.23 ha and *Tamarix* covered 113.5 ha when SVM was used and the same land-cover class covered 2920.53 ha, 259.89 ha and 74.27 ha respectively when RF was used (Table 11). When SVM was used Waterbody covered the average area of 3.3 ha and 0.41 ha when RF was used. Woody vegetation covered 168 ha when SVM was used and 173.71 ha when RF was used. Table 11 showed that the dominant land-cover types were Bareland and low-dense grassland, Buildup area and less dominant land-cover types were Agriculture, *Tamarix* and Waterbody.

Table 11. Area and percentages for Land cover class in Leeu-Gamka study site obtained by SVM and RF classifiers**.**

Class	SVM Area(ha)	%	RF Area(ha)	%	+/-(%)
Agriculture	50.1	1.43%	69.07	1.97%	-0.54%
Bareland and low-dense grassland	2778.73	79.44%	2920.53	83.49%	-4.05%
Buildup area	384.23	10.98%	259.89	7.43%	3.55%
Tamarix	113.52	3.25%	74.27	2.12%	1.12%
Waterbody	3.3	0.09%	0.41	0.01%	0.08%
Woody vegetation	168	4.80%	173.71	4.97%	-0.16%

4.2 Monitoring exotic *Tamarix* invasion using Landsat multi-temporal data

The SVM classifier was used to classify 2007, 2014 and 2018 Landsat images (path 173, row 83) and later they were used for change detection. This algorithm was used because it performed and produced overall accuracy better than RF when it was applied on Landsat images.

4.2.1 Support Vector Machine parameters tuning

The two parameters were optimised in Rstudio software and radial kernel was used during SVM classification. For best performance *gamma* and *cost* were optimised for all Landsat images and the best combined parameters are shown in Table 12. The combination of *gamma* of the value 1 and *cost* of the value 10 resulted in lowest classification error of 0.12 when parameters where optimised for Landsat data of 2007. For 2014 and 2018 Landsat data the combination of *gamma* value of 1 and *cost* value of 1000 yielded the lowest classification error of 0.07 and 0.11, respectively.

Table 12. Optimisation of Support Vector Machine (SVM) parameters (*gamma* and *cost*) using radial kernel on Landsat images.

Parameter	2007	2014	2018
gamma	1	1	1
Cost	10	1000	1000
Lowest Classification Error	0.12	0.07	0.11

4.2.2 Image Classification

The dominant class in the De Rust region was the Bareland and low-dense grassland and *Tamarix* occurred with other Woody vegetation along the flooded riparian zones (Figure 8). High accuracy was achieved when the images were classified using SVM. The classification results showed that the distribution of *Tamarix* changed over the period of the study, although this was also true for the other land-cover class types. Thus, the changes are provided in Table 16 and will be discussed further in the post classification statistic subsection. The classified images were later used for post classification change detection and statistics were computed. In 2007, the dominant land-cover was Woody vegetation and Bareland and low-dense grassland and *Tamarix* and Agriculture covers were less dominant (Figure 8a). The decrease in Agriculture and Bareland and low-dense grassland land-cover type and increase in *Tamarix* and Woody

vegetation was noted in 2014 (Figure 8b). Waterbody and Woody Vegetation showed significant decrease, while *Tamarix* increased significantly in 2018 at Olifants River (near De Rust) (Figure 8c). Therefore, *Tamarix* land-cover has been increasing since the last 11 years (Figure 8).

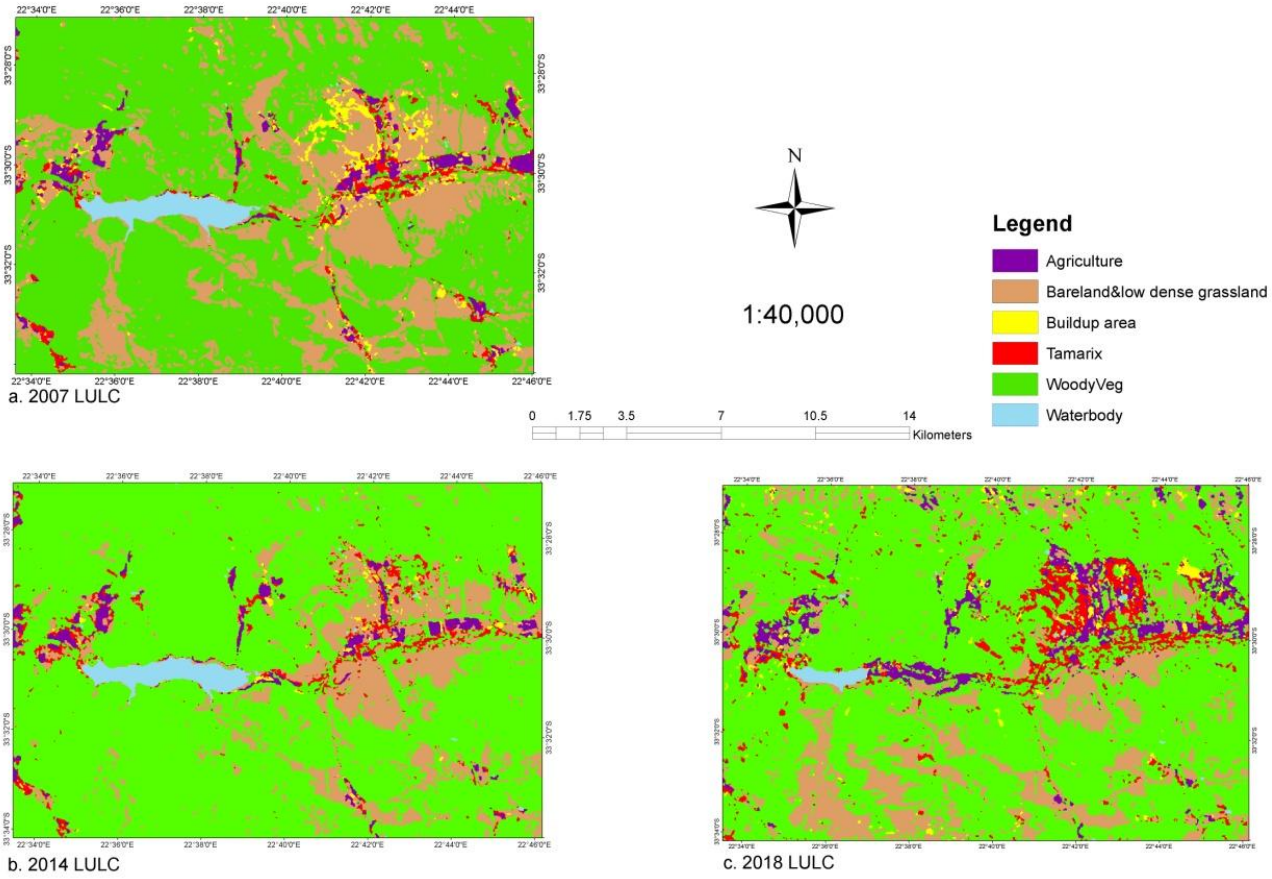


Figure 8. Monitoring *Tamarix* invasion in Olifant River (near De Rust region) using Landsat to classify the land-use and land-cover changes for the period (a) 2007 (b) 2014 and (c) 2018.

Leeu-Gamka region had Bareland and low-dense grassland as the most dominant land-cover and *Tamarix* as a less dominant in 2007 (Figure 9a). However, the *Tamarix* as well as the Woody vegetation land-cover increased significantly in 2014 (Figure 9b) and in 2018 (Figure 9c). The Bareland and low-dense grassland remained dominant throughout the study period, even though *Tamarix* and Woody vegetation was recorded as increasing.

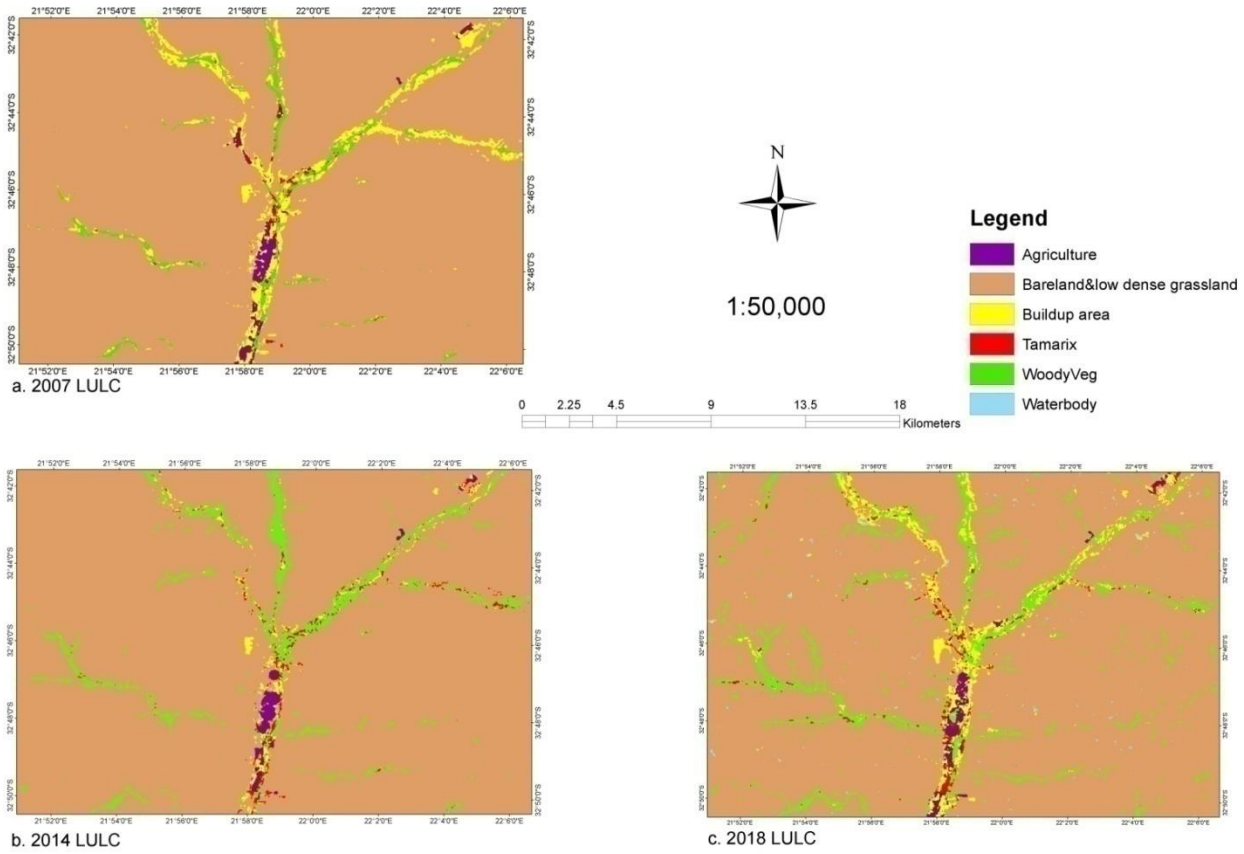


Figure 9. Monitoring *Tamarix* invasion in Leeu-Gamka using Landsat to classify (a). 2007 land-use and land-cover (b). 2014 land-use and land-cover and (c) 2018 land-use and land-cover.

Waterbody and Agriculture were less dominant land-cover and Bareland and low-dense grassland was the most dominant in Swart River near the Prince Albert town region in 2007 (Figure 10a). In 2014, however, *Tamarix* and Woody Vegetation increased, while the Waterbody and Build-up area land-cover decreased (Figure 10b). The increase in Woody Vegetation and Buildup can be noted in 2018 and slight decrease in *Tamarix* (Figure 10c). From these results it can be concluded that *Tamarix* invasion has increased over 11 years.

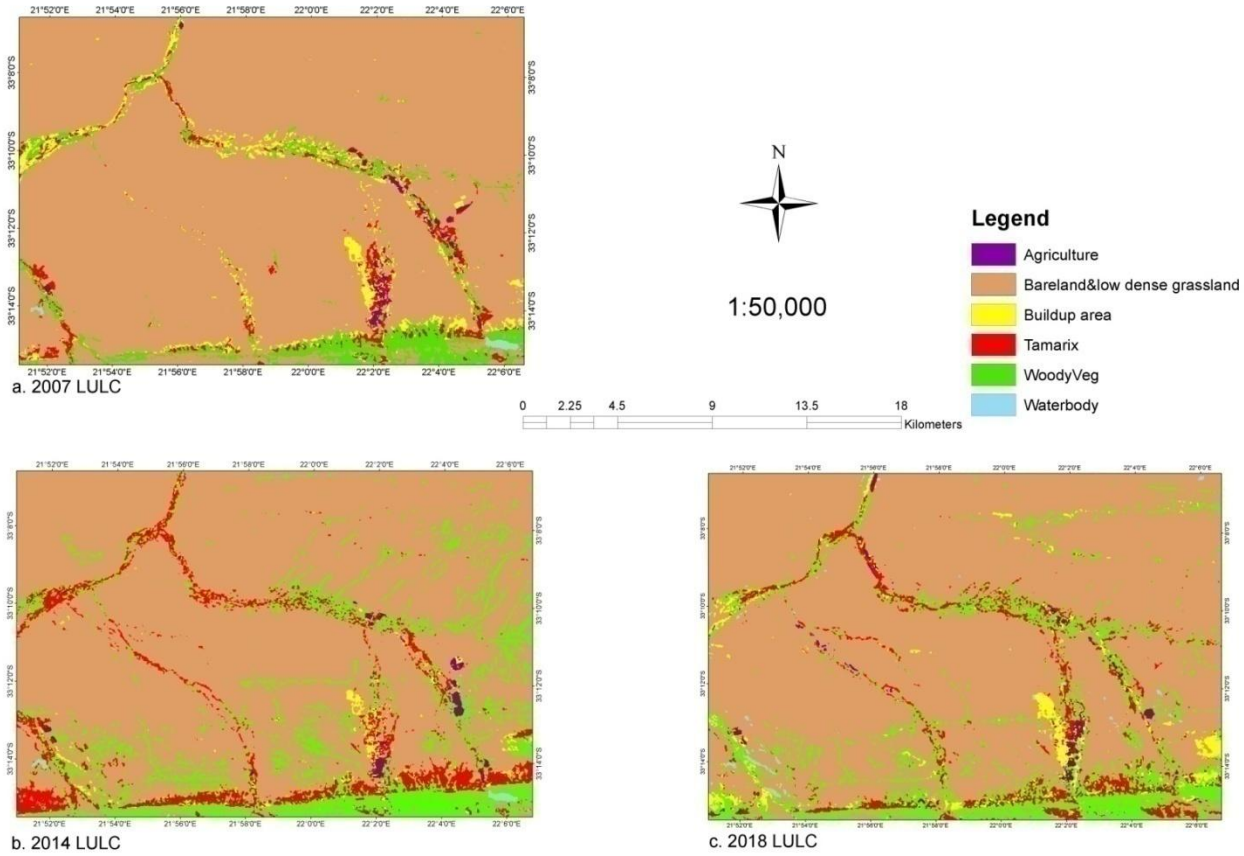


Figure 10. Monitoring *Tamarix* invasion around Prince Albert using Landsat data to classify the land-use and land-cover change for (a) 2007 (b) 2014 and (c) 2018.

4.2.3 Accuracy Assessment

The performance of SVM classifier was tested and using test datasets to construct confusion matrix (Table 13). Table 13, 14 and 15 show the confusion matrix for the 2007 Landsat 5 image, 2014 Landsat 8 image and 2018 Landsat 8 image, respectively.

The SVM classifier produced OA of 87.66% and kappa statistic of 0.849 (Table 13). Spectral confusion was noted between Bareland and low-dense grassland and Build-up areas; hence the classifier had the lowest user accuracy of 75% and 86.84% of producer's accuracy, respectively.

Table 13. Confusion matrix for validating 2007 Land-use Land-cover map. The Support Vector Machine accuracies: Overall (OA), Producer (PA) and User (UA)

LULC	A	BL	B	T	WV	W	Row total	UA
A	31	1	0	0	3	0	35	88.57
BL	1	39	4	3	4	1	52	75
B	1	3	33	0	1	0	38	86.84
T	0	0	3	26	0	0	29	89.65
WV	2	0	0	1	52	1	56	92.86
W	0	0	0	0	0	25	25	100
Column total	35	43	40	30	60	27		
PA	88.57	90.7	82.5	86.67	86.67	92.6		
OA	87.66							
Kappa	0.8497							

NB: Note that A = Agriculture, BL = Bareland& low dense grassland, B = Build-up area, T = *Tamarix*, W = Waterbody and WV = Woody Vegetation

Support Vector Machine produced OA of 91.10% and kappa value of 0.891 (Table 14). Spectral confusion was noted however, between Build-up, Bareland and low-dense grassland when Build-up was classified and had achieved the lowest producer accuracy of 67.74% and user accuracy of 88.24, whilst Bareland and low-dense grassland had UA of 97.67% and PA of 93.75% and *Tamarix* had UA of 82.1% and PA of 93.22%.

Table 14. Confusion matrix for validating 2014 Land-use Land-cover map. The Support Vector Machine accuracies: Overall (OA), Producer (PA) and User (UA).

LULC class	A	BL	T	B	W	WV	**Row total**	UA
A	42	1	0	0	0	0	**43**	97.67
BL	0	45	0	3	1	2	**51**	88.24
T	2	0	55	7	0	3	**67**	82.1
B	0	1	0	21	0	0	**22**	95.45
W	0	0	0	0	30	0	**30**	100
WV	0	1	4	0	0	63	**68**	92.65
Column	**44**	**48**	**59**	**31**	**31**	**68**	**281**	
PA	95.45	93.75	93.22	67.74	96.77	92.64		
OA	91.1							
Kappa	0.891							

Note that A = Agriculture, BL = Bareland& low dense grassland, B = Build-up area, T = *Tamarix*, W = Waterbody and WV = Woody Vegetation

Support Vector Machine classifier produced OA of 90.62% and kappa value of 0.884(Table 15). Spectral confusion was noted between Bareland and low-dense grassland, Agriculture and Woody vegetation and this produced low UA of 82.92% and PA of 85.00% when Agriculture land-cover was classified. Bareland and low dense grassland was misclassified with Buildup area and Agriculture and therefore low PA of 83.33% was achieved.

Table 15. Confusion matrix and for validating 2018 Land-use Land-cover map. The accuracies: Overall (OA), Producer (PA) and User (UA)

LULC Class	A	B	B	T	WV	W	Row total	UA
A	34	3	1	0	3	0	41	82.92
BL	1	40	0	0	1	0	42	95.23
B	1	4	40	0	2	0	47	85.11
T	1	0	0	71	8	0	80	88.75
WV	3	1	0	0	78	0	82	95.12
W	0	0	0	0	1	27	28	96.43
Column	40	48	41	71	93	27		
PA	85	83.33	97.56	100	83.87	100		
OA	90.62							
Kappa	0.884							

Note that A = Agriculture, BL = Bareland& low dense grassland, B = Build-up area, T = *Tamarix*,W = Waterbody and WV = Woody Vegetation

Although Support Vector Machine archived different classification accuracies McNemar test had a Z value of 1.08, at 5% significant level which is less than 1.96. And therefore it can be concluded that there was no significant difference between 2014 and 2018 images classified.

4.3 Monitoring and quantifying the invasion spread of *Tamarix* using change detection techniques

Post classification method was applied on multi-temporal classified images. From 2007 to 2018, LULC had changed over this period (Figure 9). *Tamarix* land-cover class was covering Olifant River area of 284.67 ha in 2007, 315.54 ha in 2014 and 647.10 ha in 2018. It can be noted that

there has been a significant increase in *Tamarix* invasion over the study period of 11 years the change was 5.46% between 2007 and 2014 and 5.73% between 2014 and 2018 (Table 16).

In 2007, Bareland and low-dense grassland land-cover area was 34.84% in 2007 it decreased to 25.61% in 2014 and again decreased to 13.11% in 2018. Agriculture land cover decreased from 5.06% in 2007; to 4.78% in 2014 and increased to 9.48% in 2018 (Table 16). Woody vegetation has shown significant increase from 2007 to 2018 with increase from 41.92% in 2007 to 54.96% in 2014 and finally the increase to 62.47% in 2018.

Table 16. The Area and percentage of temporal *Tamarix* species and other Land cover classes in De Rust obtained from Support Vector Machine classifiers.

Class	**2007 Area**	**%**	**2014 Area**	**%**	**+/-(%)**	**2018 Area**	**%**	**+/-(%)**
Agriculture	289.44	5.06%	276.21	4.78%	-0.28%	548.01	9.48%	4.70%
Bareland and low-dense grassland	1994.67	34.84%	1480.95	25.61%	-9.23%	758.43	13.11%	-12.49%
Buildup area	256.41	4.48%	64.8	1.12%	-3.36%	74.7	1.29%	0.17%
Tamarix	284.67	4.97%	315.54	5.46%	0.48%	647.1	11.19%	5.73%
Waterbody	500.31	8.74%	467.28	8.08%	-0.66%	141.93	2.45%	-5.63%
Woody vegetation	2399.94	41.92%	3178.26	54.96%	13.04%	3612.87	62.47%	7.52%

Agriculture land-cover has decreased in 2014 by 1.62% from 2.64% and again increased by 2.10% in 2018. The decrease was with 1.02% in the period of 7 years (2007-2014). Thereafter, in 2018 agriculture in Leeu-Gamka increased by 2.10% and the change was 0.49% (Table 17). Leeu-Gamka region lacks Waterbody as it covers around 0.01% throughout 11 year period. Bareland and low-dense grassland has increased by 3.74% from 2007 to 2014 and decrease by 7.97% from 2014 to 2018. In 2007, 2014 and 2018, *Tamarix* covered the average area of 74.7 ha, 88.65 ha and 97.29 ha, respectively and has increased by 0.30% from 2.13% in 2007 to 2.43% in 2014 and thereafter increased by 0.24% to 2.67% in 2018. Woody vegetation has also increased in the 11 year period, it started by covering 3.46% in 2007 and increased by 3.70% to 7.16% in 2014 and then it increased by 0.12% to 7.28% in 2018 (Table 17). From Table 17 it can be concluded that *Tamarix* and Woody vegetation had increased in the 11 year period in the Leeu-Gamka region.

Table 17. The Area (ha) and percentage of temporal *Tamarix* species and other Land cover classes in Leeu-Gamka obtained from Support Vector Machine classifiers.

Class	2007 Area	%	2014 Area	%	+/-(%)	2018 Area	%	+/-(%)
Agriculture	92.43	2.64%	59.04	1.62%	-1.02%	76.77	2.10%	0.49%
Bareland and low-dense grassland	2872.08	81.96%	3126.06	85.70%	3.74%	2835.36	77.73%	-7.97%
Buildup	341.73	9.75%	112.86	3.09%	-6.66%	366.57	10.05%	6.96%
Tamarix	74.7	2.13%	88.65	2.43%	0.30%	97.29	2.67%	0.24%
Waterbody	1.98	0.06%	0	0.00%	-0.06%	6.12	0.17%	0.17%
WoodyVeg	121.14	3.46%	261.09	7.16%	3.70%	265.59	7.28%	0.12%

Waterbody has increased from 0.27 ha to 12.15 ha from 2007 to 2018 in Prince Albert region. Woody vegetation has increased by 4.47% in 2014 from 2.61% in 2007 and then decreased by 0.33% in 2018 to 6.75% (Table 18). *Tamarix* in Prince Albert region has increased from 2.00% to 4.17% by 2.17% in 2014 and then increased by 0.70% in 2018. The slight increase between 2014 and 2018 might be attributed by the management and control measures that took place in October 2014 in Wolwekraal Nature Reserve (Milton 2015). Agriculture increased by 0.32% in 2014 and increased by 0.61% in 2018 and Buildup area has decreased from 431.28 ha in 2007 to 115.92 ha in 2014 and increased by 225.72 ha in 2018.

Table 18. The Area (ha) and percentage of temporal *Tamarix* species and other Land cover classes in Prince Albert region obtained from Support Vector Machine classifiers.

Class	2007 Area	%	2014 Area	%	+/-(%)	2018 Area	%	+/-(%)
Agriculture	70.02	0.65%	35.55	0.32%	-0.33%	103.68	0.93%	0.61%
Bareland and low-dense grassland	9773.01	90.73%	9767.79	87.38%	-3.35%	9537.93	85.32%	-2.06%
Buildup	431.28	4.00%	115.92	1.04%	-2.97%	225.72	2.02%	0.98%
Tamarix	215.01	2.00%	465.66	4.17%	2.17%	544.41	4.87%	0.70%
Waterbody	0.27	0.00%	1.62	0.01%	0.01%	12.15	0.11%	0.09%
WoodyVeg	281.61	2.61%	792	7.09%	4.47%	754.65	6.75%	-0.33%

4.3.1 Change Detection matrix

Change detection was compute using ENVI 5.4 software to determine the changes between two classified images and to report the changes between initial and final states of the classes as pixel count, percentage or area in hectares. The column total shows the total area in each final state class and class total shows total area in each final state class in columns. Class changes rows indicate total area, which has changed and image difference row indicates whether there was an increase or decrease in each classes (Table 19). This section focused on the *Tamarix* land-cover transformation in particular since the aim of this study to quantify *Tamarix* distribution dynamics. The subsection below will discuss change matrices of each study site from 2007 - 2014.

4.3.1.1 Change detection matrices of De Rust region (Olifant River)

Table 19 and Table 20 show the change of land-cover class between the year 2007 and 2014. Agriculture, Bareland and low-dense grassland, Buildup area and Waterbody had decreased by 8.46 ha, 1400.4 ha, 349.56 ha, and 32.4 ha, respectively (Table 19). Woody vegetation and *Tamarix* has however increased by 1710.63 ha and 80.19 ha, respectively.

Table 19. Change detection matrix of Land-use and Land-cover change in Olifants River (De Rust region) between 2007 and 2014 in hectares (ha)

	Initial State (2007)							
	Class Name	A	BL	T	B	W	WV	Row Total
Final State (2014)	A	**214.29**	57.87	28.26	26.01	0.36	35.82	362.61
	BL	87.66	**2022.3**	67.41	248.58	24.3	152.55	2602.8
	T	34.11	130.41	**117.18**	59.31	0.9	97.65	439.56
	B	15.84	25.29	14.67	**27.81**	0.9	10.53	95.04
	W	0	3.78	0	0	**469.8**	4.77	478.35
	WV	19.17	1763.55	131.85	82.89	14.49	**6917.85**	8929.8
	Class Total	371.07	4003.2	359.37	444.6	510.75	7219.17	0
	Class Changes	156.78	1980.9	242.19	416.79	40.95	301.32	0
	Image Difference	-8.46	-1400.4	80.19	-349.56	-32.4	1710.63	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

Agriculture, Bareland and low-dense grassland, Buildup area and Waterbody land-cover had decreased increased by 2.28%, 34.982%, 78.623% and 6.344%, respectively. Woody vegetation and *Tamarix* increased by 23.696% and 22.314%, respectively. Table 19 and 20 shows that *Tamarix* land-cover of 32.607 % (117.18 ha) remained unchanged and the total change of other land-cover types were 67.393 % (242.19 ha) with Woody vegetation transformed by 36.689% (131.85 ha). However 95.826% (6917.85 ha) of Woody vegetation remained unchanged and it has transformed into other land-cover type by 4.174% (301.32 ha). The notable transformation was into Bareland and low-dense grassland and *Tamarix* by 2.113% (152.55 ha) and 1.353% (97.65 ha), respectively.

Table 20. Change detection matrix of Land-use and Land-cover change in Olifants River (De Rust region) between 2007 and 2014 in percentages.

	Initial State (2007)							
Final State (2014)	Class name	A	BL	T	B	W	WV	Row Total
	A	**57.749**	1.446	7.864	5.85	0.07	0.496	100
	BL	23.624	**50.517**	18.758	55.911	4.758	2.113	100
	T	9.192	3.258	**32.607**	13.34	0.176	1.353	100
	B	4.269	0.632	4.082	**6.255**	0.176	0.146	100
	W	0	0.094	0	0	**91.982**	0.066	100
	WV	5.166	44.054	36.689	18.644	2.837	**95.826**	100
	Class Total	100	100	100	100	100	100	0
	Class Changes	42.251	49.483	67.393	93.745	8.018	4.174	0
	Image Difference	-2.28	-34.982	22.314	-78.623	-6.344	23.696	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

Land-use and land-cover types for the period 2014 and 2018 were evaluated to determine the change dynamism in Olifants River (De Rust region). Agriculture, *Tamarix* and Build-up area land-covers increased by 401.94 ha, 558.27 ha and72.18 ha respectively while Bareland and low-dense grassland, Woody vegetation and Waterbody decreased by 486.36 ha, 222.21 ha and 323.82 ha (Table 21). *Tamarix* land-cover area that remained unchanged is 153.63 ha and the total changes to other land-cover classes was 303.93 ha with significant transformation was to Woody vegetation by 149.76 ha and Agriculture with 112.68 ha (Table 21).

Table 21. Change detection matrix of Land-use and Land-cover change in Olifants River (De Rust region) between 2014 and 2018 in hectares

		Initial State (2014)						
Final State (2018)	Class Name	A	BL	B	T	WV	W	Row Total
	A	**183.06**	259.56	42.39	112.68	64.35	114.93	776.97
	BL	54.72	**1249.11**	7.56	18.63	805.05	77.85	2212.92
	B	21.96	32.94	**13.95**	21.33	72.81	4.77	167.76
	T	9.9	443.07	17.55	**153.63**	360.99	30.69	1015.83
	WV	105.12	712.62	13.41	149.76	**8457.39**	105.21	9543.51
	W	0.27	1.98	0.72	1.53	5.13	**144.9**	154.53
	Class Total	375.03	2699.28	95.58	457.56	9765.72	478.35	0
	Class Changes	191.97	1450.17	81.63	303.93	1308.33	333.45	0
	Image Difference	401.94	-486.36	72.18	558.27	-222.21	-323.82	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

Table 22 depicts the change in Land-cover types from 2014 to 2018 in percentages. Agriculture, Build-up area, and *Tamarix* land-covers increased by 107.175%, 75.518%, 122.01% respectively while Bareland and low-dense grassland, Woody vegetation and Waterbody decreased by 18.018%, 2.275% and 67.695% (Table 22)

Table 22. Change detection matrix of Land-use and Land-cover change in Olifants River (De Rust region) between 2014 and 2018 in percentages

	Initial State (2014)							
	Class Name	A	BL	B	T	WV	W	Row Total
Final State (2018)	A	48.812	9.616	44.35	24.626	0.659	24.026	100
	BL	14.591	46.276	7.91	4.072	8.244	16.275	100
	BL	5.856	1.22	14.595	4.662	0.746	0.997	100
	T	2.64	16.414	18.362	33.576	3.697	6.416	100
	WV	28.03	26.4	14.03	32.73	86.603	21.994	100
	W	0.072	0.073	0.753	0.334	0.053	30.292	100
	Class Total	100	100	100	100	100	100	0
	Class Changes	51.188	53.724	85.405	66.424	13.397	69.708	0
	Image Difference	107.175	-18.018	75.518	122.01	-2.275	-67.695	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

4.1.3.2 Change detection matrices of Leeu Gamka region

Table 23 and Table 24 shows change detection matrices in Leeu Gamka region from 2007 to 2014, while Table 25 and Table 26 show the change detection matrices from 2014 to 2018. Agriculture has decreased by 38.43 ha and Waterbody has decreased by 1.98 ha. Bareland and low-dense grassland, *Tamarix*, Build up area and Woody vegetation has increased by 125.01 ha, 11.88 ha, 230.85 ha and 134.37 ha, respectively (Table 23). 16.02 ha of *Tamarix* remained unchanged from the initial 86.58 ha in 2007 and it has transformed to other land-cover by 74.7 ha. The change was to Woody vegetation, Bareland and low-dense grassland, Buildup area and Agriculture with the values of 33.75 ha, 14.31 ha, 6.84 ha and 3.78 ha, respectively (Table 23).

Table 23. Change detection matrix of Land-use and Land-cover change in Leeu Gamka region between 2007 and 2014 in hectares.

	Initial State (2007)							
Final State (2014)	Class Name	A	BL	T	B	W	WV	Row Total
	A	38.88	1.17	3.78	8.19	0	1.98	54
	BL	37.44	2791.26	14.31	149.94	0.36	3.78	2997.09
	T	5.31	6.93	16.02	38.25	0.36	19.71	86.58
	B	9.63	18.18	6.84	72.09	1.26	2.88	110.88
	W	0	0	0	0	0	0	0
	WV	1.17	54.54	33.75	73.26	0	92.79	255.51
	Class Total	92.43	2872.08	74.7	341.73	1.98	121.14	0
	Class Changes	53.55	80.82	58.68	269.64	1.98	28.35	0
	Image Difference	-38.43	125.01	11.88	-230.85	-1.98	134.37	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

Agriculture has decreased by 41.577% and Waterbody has decreased by 100%. Bareland and low-dense grassland, *Tamarix*, Build up area and Woody vegetation has increased by 4.353%, 15.904%, 67.553% and 110.921%, respectively (Table 24). Agriculture remained unchanged with 42.064%, Bareland and low-dense grassland remained unchanged with 97.186 % and *Tamarix* remained unchanged with 21.446%. Also, Woody vegetation remained unchanged with 76.597% and 21.096% remained unchanged in Buildup area.

Table 24. Change detection matrix of Land-use and Land-cover change in LeeuGamka region between 2007 and 2014 in percentages

	Initial State (2007)							
Final State (2014)	Class Name	A	BL	T	B	W	WV	Row Total
	A	**42.064**	0.041	5.06	2.397	0	1.634	100
	BL	40.506	**97.186**	19.157	43.877	18.182	3.12	100
	T	5.745	0.241	**21.446**	11.193	18.182	16.27	100
	B	10.419	0.633	9.157	**21.096**	63.636	2.377	100
	W	0	0	0	0	**0**	0	0
	WV	1.266	1.899	45.181	21.438	0	**76.597**	100
	Class Total	100	100	100	100	100	100	0
	Class Changes	57.936	2.814	78.554	78.904	100	23.403	0
	Image Difference	-41.577	4.353	15.904	-67.553	-100	110.921	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

Table 25 and Table 26 show the change detection matrices from 2014 to 2018 in Leeu Gamka region. Bareland and low-dense grassland is the only land-cover type that has decreased in this period. The decrease was by 290 ha however Agriculture, Buildup area, *Tamarix* and Woody vegetation has increased with 17.73 ha, 253.71 ha, 8.64 ha and 4.50 ha, respectively (Table 25). The initial *Tamarix* land-cover in 2014 was 97.29 ha and 18.27 ha remained unchanged. The total area of *Tamarix* that has changed is 70.38 ha and it has transformed to Woody vegetation and Buildup area by 29.34 ha and 33.93 ha, respectively (Table 25).

Table 25. Change detection matrix of Land-use and Land-cover change in Leeu Gamka region between 2014 and 2018 in hectares.

		Initial State (2014)						
		A	BL	B	T	WV	W	Row Total
Final State (2018)	A	38.7	24.75	8.82	3.51	0.99	0	76.77
	BL	0.54	2806.38	9.45	3.6	15.39	0	2835.36
	B	16.29	162.99	83.43	33.93	69.93	0	366.57
	T	0.81	44.1	8.1	18.27	26.01	0	97.29
	WV	2.7	81.72	3.06	29.34	148.77	0	265.59
	W	0	6.12	0	0	0	0	6.12
	Class Total	59.04	3126.06	112.86	88.65	261.09	0	0
	Class Changes	20.34	319.68	29.43	70.38	112.32	0	0
	Image Difference	17.73	-290.7	253.71	8.64	4.5	6.12	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = *Tamarix*,W = Waterbody and WV = Woody Vegetation

The percentage increase for each of these classes is shown in Table 26 where Agriculture, Buildup area, *Tamarix* and Woody vegetation has increased by 30.030%, 224.801%, 9.746% and 1.724% respectively, while Bareland and low-dense grassland decreased by 9.299% and Waterbody has no change (Table 26). *Tamarix* land-cover that remained unchanged was 20.609%.

Table 26. Change detection matrix of Land-use and Land-cover change in Leeu Gamka region between 2014 and 2018 in percentages.

		Initial State (2014)						
		A	BL	B	T	WV	W	Row Total
Final State (2018)	A	**65.549**	0.792	7.815	3.959	0.379	0	100
	BL	0.915	**89.774**	8.373	4.061	5.895	0	100
	B	27.591	5.214	**73.923**	38.274	26.784	0	100
	T	1.372	1.411	7.177	**20.609**	9.962	0	100
	WV	4.573	2.614	2.711	33.096	**56.98**	0	100
	W	0	0.196	0	0	0	**0**	100
	Class Total	100	100	100	100	100	0	0
	Class Changes	34.451	10.226	26.077	79.391	43.02	0	0
	Image Difference	30.03	-9.299	224.801	9.746	1.724	0	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

4.1.3.3 Change detection matrices of Prince Albert region

Table 27 and Table 28 show the change of land-cover class between the year 2007 and 2014, while Table 29 and Table 30 show the change in land-cover between the year 2014 and 2018. Land-use and land-cover types for the period 2007 and 2014 were evaluated to determine the change dynamism in Prince Albert. *Tamarix*, Waterbody and Woody vegetation land-covers increased by 232.20 ha, 1.35 ha and 480.96 ha respectively while Bareland and low-dense grassland, Build-up area and Agriculture decreased by 357.84 ha, 320.85 ha and 35.82 ha (Table 27). *Tamarix* land-cover area that remained unchanged was 87.39 ha and the total changes to other land-cover classes was 127.62 ha with significant transformation was to Bareland and low-dense grassland by 56.25 ha and Agriculture by 53.46 ha (Table 27).

Table 27. Change detection matrix of Land-use and Land-cover change in Prince Albert region between 2007 and 2014 in hectares

	Initial State (2007)							
	Class Name	A	BL	T	B	W	WV	Row Total
Final State (2014)	A	**13.86**	4.14	3.96	2.79	0	9.45	34.2
	BL	34.38	**9121.23**	56.25	179.82	0.09	23.4	9415.17
	T	12.06	174.06	**87.39**	116.28	0	57.42	447.21
	B	7.38	13.68	13.77	**73.44**	0	2.16	110.43
	W	0	0.99	0.18	0	**0.18**	0.27	1.62
	WV	2.34	458.91	53.46	58.95	0	**188.91**	762.57
	Class Total	70.02	9773.01	215.01	431.28	0.27	281.61	0
	Class Changes	56.16	651.78	127.62	357.84	0.09	92.7	0
	Image Difference	-35.82	-357.84	232.2	-320.85	1.35	480.96	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

Tamarix, Waterbody and Woody vegetation land-covers increased by 107.995%, 500.00% and 170.789%, respectively while Bareland and low-dense grassland, Build-up area and Agriculture decreased by 74.395%, 3.662% and 51.157%, respectively (Table 28). *Tamarix* land-cover that remained unchanged was 40.645% and it has transformed to other land-cover classes by 59.355%.

Table 28. Change detection matrix of Land-use and Land-cover change in Prince Albert region between 2007 and 2014 in percentages.

	Initial State (2007)							
	Class Name	A	BL	T	B	W	WV	Row Total
Final State (2014)	A	**19.794**	0.042	1.842	0.647	0	3.356	100
	BL	49.1	**93.331**	26.162	41.694	33.333	8.309	100
	T	17.224	1.781	**40.645**	26.962	0	20.39	100
	B	10.54	0.14	6.404	**17.028**	0	0.767	100
	W	0	0.01	0.084	0	**66.667**	0.096	100
	WV	3.342	4.696	24.864	13.669	0	**67.082**	100
	Class Total	100	100	100	100	100	100	0
	Class Changes	80.206	6.669	59.355	82.972	33.333	32.918	0
	Image Difference	-51.157	-3.662	107.995	-74.395	500	170.789	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

Table 29 and Table 30 show the change of land-cover class between the year 2014 and 2018. Agriculture, Buildup area, *Tamarix* and Waterbody had increased by 68.13 ha, 109.8 ha, 78.75 ha, and 10.53 ha, respectively (Table 29). Bareland and low-dense grassland and Woody vegetation has, however; decreased by 229.86 ha and 37.35 ha, respectively

Table 29. Change detection matrix of Land-use and Land-cover change in Prince Albert region between 2014 and 2018 in hectares.

	Initial State (2014)							
	Class Name	A	BL	B	T	WV	W	Row Total
Final State (2018)	A	15.66	67.05	6.75	12.42	1.8	0	103.68
	BL	0.45	9101.34	7.56	102.42	326.16	0	9537.93
	B	10.08	71.82	78.3	45.99	19.35	0.18	225.72
	T	3.96	227.34	16.83	170.37	125.91	0	544.41
	WV	5.4	290.88	6.21	134.46	317.7	0	754.65
	W	0	9.36	0.27	0	1.08	1.44	12.15
	Class Total	35.55	9767.79	115.92	465.66	792	1.62	0
	Class Changes	19.89	666.45	37.62	295.29	474.3	0.18	0
	Image Difference	68.13	-229.86	109.8	78.75	-37.35	10.53	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

Agriculture, Buildup area, *Tamarix* and Waterbody had increased by 191.646%, 94.72%, 16.911%, and 650%, respectively. Bareland and low-dense grassland and Woody vegetation has, however; decreased by 229.86 ha and 37.35 ha, respectively. Table 29 and 30 show that *Tamarix* land-cover of 36.587% (170.37 ha) remained unchanged and the total change to other land-cover types were 63.413% (295.29ha) with *Tamarix* transforming to Woody vegetation, Buildup area, Agriculture and Bareland and low-dense grassland with 28.875% (134.46 ha), 9.876% (45.99 ha), 2.667% (12.42 ha) and 21.995% (102.42 ha), respectively.

	Initial State (2014)							
	Class Name	A	BL	B	T	WV	W	Row Total
Final State (2018)	A	44.051	0.686	5.823	2.667	0.227	0	100
	BL	1.266	93.177	6.522	21.995	41.182	0	100
	B	28.354	0.735	67.547	9.876	2.443	11.111	100
	T	11.139	2.327	14.519	36.587	15.898	0	100
	WV	15.19	2.978	5.357	28.875	40.114	0	100
	W	0	0.096	0.233	0	0.136	88.889	100
	Class Total	100	100	100	100	100	100	0
	Class Changes	55.949	6.823	32.453	63.413	59.886	11.111	0
	Image Difference	191.646	-2.353	94.72	16.911	-4.716	650	0

Note that A = Agriculture, BL = bareland& low dense grassland, B = Build-up area, T = Tamarix,W = Waterbody and WV = Woody Vegetation

CHAPTER FIVE

5 Discussions and conclusion

5.1 Discussions

Alien invasive *Tamarix* species has negative impacts on riparian ecosystems and biodiversity. The management of alien plants depends on reliable and accurate information on their extent and distribution to facilitate decision making process to undertake an appropriate and effective control measures. Remote sensing is capable of being that tool to provide information about the invasion dynamism in the environment and over the years it has earned enormous popularity as a suitable tool for mapping invasive species. The advent of high temporal and spectral resolution remote sensing data and their availability at more affordable price or free of charge provide a growing demand as a tool of choice for mapping and monitoring invasive plants at species level accurately. Many studies have already shown its success in mapping invasive alien species and other LULC in recent years (Ng et al. 2016 ; Gavier-Pizarro et al. 2012; Adam et al. 2017). Ji and Wang (2015) were able to discriminate *Tamarix* from Cottonwood and other co-existing species using Landsat 5 data in Rio Grande River (USA) and the overall accuracy achieved was 93%. Furthermore, there is little or no information reported on mapping *Tamarix* species using Sentinel 2. However, other invasive species were mapped from their co-existing species using remote sensing techniques (Calleja et al. 2019; Traganos and Reinartz 2017).

The first objective of this study was to investigate the efficiency of Sentinel 2 data in mapping the spatial distribution of *Tamarix*, using the RF and SVM machine learning algorithms for classification. The relatively high accuracies of 87.83% and 86.31% achieved by RF and SVM respectively, demonstrated that high resolution Sentinel 2 is capable of discriminating *Tamarix* from other vegetation land cover. The results agrees with similar study where overall accuracy of 77% was achieved when Sentinel 2 data was used to map invasive *Baccharis halimifolia* in Spain (Calleja et al. 2019). Other previous studies on mapping *Tamarix* demonstrated the capability of hyperspectral and medium resolution multispectral data in discriminating *Tamarix* from other land-cover types or co-existing species (Ji and Wang 2015; Ji and Wang 2016; Silván-Cárdenas and Wang 2010; Hamada et al. 2007).

Overall accuracies of 80-90% on discriminating *Tamarix* spp from other co-existing species have been associated with hyperspectral data (Hamada et al. 2007) and high resolution multispectral data such as Quickbird (Ji and Wang 2015). However, this study demonstrated that Sentinel 2 at 10m spatial resolution was capable of discriminating *Tamarix* from other land cover types with RF and SVM achieving the overall accuracy of 87.83% and 86.31%, respectively. In addition, the study also demonstrated the potential of Landsat data in discriminating *Tamarix* from its co-existing species with overall accuracy ranging from 87-91%. The fact that Sentinel 2 and Landsat imageries are freely available it is an added advantage as potentially effective satellite data for mapping and monitoring of invasive species accurately. As an alternative monitoring tool, it has the potential to resolves some of the unique properties of the high resolution hyperspectral data, which is often limited by its availability due to its enormous cost as well as its data dimensionality and redundancy (Cho et al. 2015). The results from this study are correspond to those found in recent study where the potential of Sentinel 2 data was employed to map invasive sea grass in the Mediterranean Sea in Greece (Traganos and Reinartz 2017). It was also used to map invasive *Baccharis halimifolia* in Bay of Biscay located on Northern coast of Spain (Calleja et al. 2019).

Random Forest and SVM were used to classify the Sentinel 2 data and the performance of both machine learning algorithms were tested and their accuracy compared. The overall accuracies were 86.31 and 87.83% for SVM and RF, respectively. However, Random Forest achieved a slightly higher accuracy level of about 1 % than SVM. The advantage of Random Forest is that it can provide variable importance, which determines the importance of each band in the Sentinel 2 image (see Figure 4 above), while different kernels can be used in SVM classification. Radial kernel was used in the study because it is an efficient method that deals with vegetation species class inseparability issues (Adam et al., 2017). The good accuracies levels attained in this study using Random Forest and Support Vector Machine agree with other studies where RF and SVM were used with multispectral data (Traganos and Reinartz, 2017; Pal, 2005; Adam et al. 2017; Adam et al. 2012) and hyperspectral (Hamada et al. 2007; Burai et al. 2015) data for vegetation mapping. Previous studies showed that *Tamarix* prefer to invade the riparian zones and flood plains and this was confirmed in this study. *Tamarix* preferably invade the riparian zone where its negative impact is more pronounced due to the sensitivity of such ecosystems (Figure 5).

Tamarix also occurred in agricultural lands in De Rust region (Western Cape Province) and this has a direct bearing on the socio-economic activities of the farmers in the region by subjecting the crop land to increased soil salinity and lower groundwater table and reducing the biodiversity of the agricultural ecosystem (Gary 1960). The *Tamarix* invasion in the riparian zone could be attributed to wind and water dispersal ability of the *Tamarix* seeds in addition to its ability to reproduce vegetatively (Newete et al. 2019). The current *Tamarix* distribution cover is approximately 2-16% of the study sites when Sentinel 2 data and Both the RF and SVM classifiers were used and it is 2-.11% when SVM was used on Landsat 8 data. This can be due to the difference in spectral and spatial resolutions between the two satellite data used. For instance, unlike Landsat data, Sentinel 2 has a red edge band which is useful in vegetation classification (Adam et al. 2014).

Remote sensing satellite data has large area coverage and it has a potential to provide temporal data of specific area. Temporal resolution of remote sensing data provides resource managers, decision makers and scientist the opportunity to quantify change in *Tamarix* invasion over a long period of time. Although, the high resolution remote sensing data Sentinel 2, lack long historical data and or temporal resolution, the Landsat data has a potential to map *Tamarix* and other invasive plant species from as early as 1974 to the current date.

The second objective of this study was to investigate the mapping of spatiotemporal distribution of *Tamarix* using Landsat data and SVM classification algorithm in Western Cape Province in South Africa from 2007 to 2018. Support Vector Machine was employed to map the temporal distribution of *Tamarix*, because it was discovered that SVM performed better than RF during classification of Landsat data. Three Landsat images were used to quantify the change and data were acquired for 2007, 2014 and 2018. The overall accuracies achieved for 2007, 2014 and 2018 were 87.66%, 91.1%, and 90.62%, respectively. These high accuracies obtained demonstrated the potential use of Landsat data in discriminating *Tamarix* from other land cover types. Previous studies that mapped *Tamarix* using Landsat images achieved good accuracies and they demonstrated the potential of Landsat in discriminating *Tamarix* from its co-existing species (Ji and Wang 2015; Silván-Cárdenas and Wang 2010). Thus, high accuracies achieved made quantifying temporal change in *Tamarix* possible. In Olifant River (De Rust region)

Tamarix covered 4.97% of land-cover in 2007, 5.46% in 2014 and 11.19% in 2018. Between the periods 2007-2014 *Tamarix* increased by 0.48% and between 2014 and 2018 by5.73% (Table 16). The increase was also noted in Leeu Gamka region where *Tamarix* covered 2.13% of land-cover in 2007, 2.43% in 2014 and 2.67% in 2018 (Table 17). Invasion increase by 0.30% between 2007 and 2014 and also increase between 2014 and 2018 by 0.24%. *Tamarix* land-cover in Prince Albert region was 2.00% in 2007, then increased to 4.17% in 2014 by 2.17% and then again increased to 4.87% in 2018 by 0.70%. The low increase in the period of 2014-2018 can be contributed by the *Tamarix* clearing operation in Prince Albert area that occurred in 2014 (Milton 2015). This information provides useful information about the extent of *Tamarix* invasion and can be used as basis for evaluating environmental degradation and ecosystem function.

5.2 Conclusion

The aim of this study was to map the spatial and temporal extent of the exotic *Tamarix* species in Western Cape using Sentinel 2 and Landsat earth observation data. The results showed that *Tamarix* can be spatially and spectrally discriminated from its co-existing species. Sentinel 2 and Landsat data were able to successfully discriminate between *Tamarix* and other land-cover types, which were Bareland and low-dense grassland, Build-up, Woody vegetation, Waterbody and Agriculture.

High accuracy achieved by Random Forest and SVM illustrated the potential of these classifiers in discriminating *Tamarix* from other land-covers. Two parameters namely *ntree* and *mtry* were optimised in order to determine the best combination using 10 fold cross validation and best combination was *ntree* of 8500 and *mtry* of 2 (Figure 3). Random Forest achieved an overall accuracy of 87.83% and Kappa value of 0.85 and it outperformed the SVM classifier by 1%. However, McNemar test showed that even though RF outperformed SVM there was no significant difference ($Z \leq 1.96$) in the way the classifiers performed as the Z value was 0.75 (Table 8). Variable importance which indicates the most important bands in Sentinel 2 for discriminating *Tamarix* from other land-cover types was provided by RF as red, blue, red edge, and SWIR 2bands.

Support Vector Machine (SVM) also showed the potential of discriminating *Tamarix* from other land-cover type. Radial kernel was employed and two parameters were optimized and 10 fold cross validation yielded best combination of *gamma* and *cost* values of 10 and 0.1 respectively. The overall accuracy achieved was 86.31% and Kappa value of 0.833 (Table 6).

Support Vector Machine was applied on multi-temporal Landsat data of 2007, 2014 and 2018 and overall accuracies achieved were 87.66%, 91.1%, and 90.62%, respectively. Landsat data that was used to map the current *Tamarix* achieved high OA of 90.62%, which was greater than the overall accuracies achieved by Sentinel 2 (87.83% and 86.31%). However, Sentinel 2 data is preferred data to be utilised to discriminate *Tamarix* from co-existing species due to its better spectral and spatial resolution. Multi temporal Landsat data has shown the potential to discriminate previous and current distribution of *Tamarix.* The distribution of *Tamarix* in De Rust region was found to be 284.67 ha, 315.54 ha, 647.10 ha for 2007, 2014 and 2018 respectively. In Leeu Gamka region the distribution was found to be 74.7 ha in 2007, 88.65 ha in 2014 and 97.29 ha in 2018.The distribution of *Tamarix* in Prince Albert region was found to be 215.01 ha, 465.66 ha, 544.41 ha for 2007, 2014 and 2018, respectively. The results show that the *Tamarix* invasion has increased and urged management is needed especially in the Prince Albert region and Olifant River which had the most *Tamarix* land-cover.

Tamarix invasion in Western Cape of South Africa is rapidly increasing and has a negative impact on the environment and biodiversity. This *Tamarix* threats need urgent and effective management plan and river restoration programme. Urgent intervention from decision makers is recommended to avoid massive invasion and alteration of the riparian ecosystem. The overall accuracies of 86.31% and 87.83% achieved in this study showed the potential use of multispectral images such as Sentinel 2 for mapping *Tamarix* at species level. Further studies should be taken by employing and exploring the use of field spectroscopy data to discriminate *Tamarix* species. Since *Tamarix* spp secrete salts via its leaf glands and increases soil salinity, vegetation indices like salinity index can be used to detect the *Tamarix* and also soil conductivity data can be incorporated into auxiliary data for classification. Alternatively the use of object-based classification have to be investigated using multispectral and hyperspectral images which will considers canopy structure, colour and texture of *Tamarix*.

References

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. https://doi.org/10.1016/j.jaridenv.2017.05.001

Adam, E., Mutanga, O., Odindi, J., Abdel-Rahman, E.M., 2014. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. International Journal of Remote Sensing 35, 3440–3458. https://doi.org/10.1080/01431161.2014.903435

Adam, E., Mutanga, O., Rugege, D., Ismail, R., 2012. Discriminating the papyrus vegetation (Cyperus papyrus L.) and its co-existent species using random forest and Hyperspectral data resampled to HYMAP. https://doi.org/10.1080/01431161.2010.543182

Baum, B.B., 1978. The genus Tamarix. Israel Academy of Sciences and Humanities., Jerusalem.

Berry, W.L., 1970. Characteristics of Salts Secreted by Tamarix aphylla. American Journal of Botany 57, 1226–1230. https://doi.org/10.2307/2441362

Bhavsar, H., Panchal, M.H., 2012. A Review on Support Vector Machine for Data Classification 1, 5.

Blackburn, W.H., Knight, R.W., Schuster, J.L., 1982. Saltcedar influence on sedimentation in the Brazos River. Journal of Soil and Water Conservation 37, 298–301.

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brock, J.H., 1984. 4 Tamarix spp. (Salt Cedar), an Invasive Exotic Woody Plant in Arid and Semi-arid Riparian Habitats of Western USA 18.

Brotherson, J., Field, D., 1987. Tamarix: Impacts of a Successful Weed. Rangelands 9.

Burai, P., Tomor, T., Bekő, L., Deák, B., 2015. AIRBORNE HYPERSPECTRAL REMOTE SENSING FOR IDENTIFICATION GRASSLAND VEGETATION. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-3/W3, 427–431. https://doi.org/10.5194/isprsarchives-XL-3-W3-427-2015

Busch, D.E., Smith, S.D., 1995. Mechanisms Associated With Decline of Woody Species in Riparian Ecosystems of the Southwestern U.S. Ecological Monographs 65, 347–370. https://doi.org/10.2307/2937064

Calleja, F., Ondiviela, B., Galván, C., Recio, M., Juanes, J.A., 2019. Mapping estuarine vegetation using satellite imagery: The case of the invasive species Baccharis halimifolia at a Natura 2000 site. Continental Shelf Research 174, 35–47. https://doi.org/10.1016/j.csr.2019.01.002

Campbell, C.J., Strong, J.E., 1964. Salt Gland Anatomy in Tamarix Pentandra (Tamaricaceae). The Southwestern Naturalist 9, 232–238. https://doi.org/10.2307/3669691

Carter, G., Lucas, K., Blossom, G., Lassitter, C., Holiday, D., Mooneyhan, D., Fastring, D., Holcombe, T., Griffith, J., Carter, G.A., Lucas, K.L., Blossom, G.A., Lassitter, C.L., Holiday, D.M., Mooneyhan, D.S., Fastring, D.R., Holcombe, T.R., Griffith, J.A., 2009. Remote Sensing and Mapping of Tamarisk along the Colorado River, USA: A Comparative Use of Summer-Acquired Hyperion, Thematic Mapper and QuickBird Data. Remote Sensing 1, 318–329. https://doi.org/10.3390/rs1030318

Cho, M.A., Malahlela, O., Ramoelo, A., 2015. Assessing the utility WorldView-2 imagery for tree species mapping in South African subtropical humid forest and the conservation

implications: Dukuduku forest patch as case study. International Journal of Applied Earth Observation and Geoinformation 38, 349–357. https://doi.org/10.1016/j.jag.2015.01.015
Congalton, R.G., Green, K., 2008. Assessing the accuracy of remotely sensed data: principles and practices. CRC press.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learning 20, 273–297. https://doi.org/10.1007/BF00994018
de Leeuw, J., Jia, H., Yang, L., Liu, X., Schmidt, K., Skidmore, A.K., 2006. Comparing accuracy assessments to infer superiority of image classification methods. International Journal of Remote Sensing 27, 223–232. https://doi.org/10.1080/01431160500275762
Dean, W.R.J., Milton, S.J., 2019. The dispersal and spread of invasive alien Myrtillocactus geometrizans in the southern Karoo, South Africa. South African Journal of Botany 121, 210–215. https://doi.org/10.1016/j.sajb.2018.11.005
Díaz-Uriarte, R., Alvarez de Andrés, S., 2006. Gene selection and classification of microarray data using random forest. BMC Bioinformatics 7, 3. https://doi.org/10.1186/1471-2105-7-3
DiTomaso, J., Kyser, G., Oneto, S., Wilson, R., Orloff, S., Anderson, L., Wright, S., Roncoroni, J., Miller, T., Prather, T., Ransom, C., Beck, G., Duncan, C., Wilson, K., Mann, J., 2013. {Weed control in natural areas in the Western United States}. Weed Research and Information Center, University of California.
Dorn, H., Törnros, T., Zipf, A., 2015. Quality Evaluation of VGI Using Authoritative Data—A Comparison with Land Use Data in Southern Germany. ISPRS International Journal of Geo-Information 4, 1657–1671. https://doi.org/10.3390/ijgi4031657
Duncan, K.W., McDaniel, K.C., 1998. Saltcedar (Tamarix spp.) Management with Imazapyr. Weed Technology 12, 337–344.
Evangelista, P., Stohlgren, T., Morisette, J., Kumar, S., 2009. Mapping Invasive Tamarisk (Tamarix): A Comparison of Single-Scene and Time-Series Analyses of Remotely Sensed Data. Remote Sensing 1, 519–533. https://doi.org/10.3390/rs1030519
Gary, H.L., 1960. Utilization of five-stamen tamarisk by cattle. Research Notes. Rocky Mountain Forest and Range Experiment Station 51, 4 pp.
Gavier-Pizarro, G.I., Kuemmerle, T., Hoyos, L.E., Stewart, S.I., Huebner, C.D., Keuler, N.S., Radeloff, V.C., 2012. Monitoring the invasion of an exotic tree (Ligustrum lucidum) from 1983 to 2006 with Landsat TM/ETM+ satellite data and Support Vector Machines in Córdoba, Argentina. Remote Sensing of Environment 122, 134–145. https://doi.org/10.1016/j.rse.2011.09.023
Gillison, A.N., 2006. A Field Manual for Rapid Vegetation Classification and Survey for general purposes 85.
Hamada, Y., Stow, D.A., Coulter, L.L., Jafolla, J.C., Hendricks, L.W., 2007. Detecting Tamarisk species (Tamarix spp.) in riparian habitats of Southern California using high spatial resolution hyperspectral imagery. Remote Sensing of Environment 109, 237–248. https://doi.org/10.1016/j.rse.2007.01.003
Henderson, L., 2001a. Alien weeds and invasive plants: a complete guide to declared weeds and invader in South Africa, including another 36 species invasive in that region, Plant Protection Research Institute Handbook. Plant Protection Research Institute, Pretoria.
Henderson, L., 2001b. Alien weeds and invasive plants: a complete guide to declared weeds and invader in South Africa, including another 36 species invasive in that region, Plant Protection Research Institute Handbook. Plant Protection Research Institute, Pretoria.

Heywood, V., Brummitt, R., Culham, A., Seberg, O., 2007. Flowering plant families of the world. Kew Publishing.
Hollingsworth, E.B., P. C. Quimby, J., Jaramillo, D.C., 1979. Control of Saltcedar by Subsurface Placement of Herbicides. Journal of Range Management 32, 288–291. https://doi.org/10.2307/3897833
Horton, J.S., Mounts, F.C., Kraft, J.M., Rocky Mountain Forest and Range Experiment Station (Fort Collins, C.), University of Arizona, Arizona State University, 1960. Seed germination and seedling establishment of phreatophyte species. Fort Collins, Colo. : Rocky Mountain Forest and Range Experiment Station, Forest Service, U.S. Dept. of Agriculture.
Hoshino, B., Karamalla, A., Abd, M.A.M., Suliman, M., Elgamri, M., NAwata, H., Yasuda, H., 2012. Evaluating the Invasion Strategic of Mesquite (Prosopis juliflora) in Eastern Sudan Using Remotely Sensed Technique 4.
Islam, K., Jashimuddin, M., Nath, B., Nath, T.K., 2018. Land use classification and change detection by using multi-temporal remotely sensed imagery: The case of Chunati wildlife sanctuary, Bangladesh. The Egyptian Journal of Remote Sensing and Space Science 21, 37–47. https://doi.org/10.1016/j.ejrs.2016.12.005
Jensen, J.R., 1995. Introductory Digital Image Processing: A Remote Sensing Perspective, 2nd ed. Prentice Hall PTR, Upper Saddle River, NJ, USA.
Ji, W., Wang, L., 2016. Phenology-guided saltcedar (Tamarix spp.) mapping using Landsat TM images in western U.S. Remote Sensing of Environment 173, 29–38. https://doi.org/10.1016/j.rse.2015.11.017
Ji, W., Wang, L., 2015. Discriminating Saltcedar (Tamarix ramosissima) from Sparsely Distributed Cottonwood (Populus euphratica) Using a Summer Season Satellite Image. Photogrammetric Engineering & Remote Sensing 81, 795–806. https://doi.org/10.14358/PERS.81.10.795
Karatzoglou, A., Meyer, D., Hornik, K., 2006. Support Vector Machines in R. Journal of Statistical Software; Vol 1, Issue 9 (2006). https://doi.org/10.18637/jss.v015.i09
Kennedy, R.E., Townsend, P.A., Gross, J.E., Cohen, W.B., Bolstad, P., Wang, Y.Q., Adams, P., 2009. Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects. Remote Sensing of Environment 113, 1382–1396. https://doi.org/10.1016/j.rse.2008.07.018
Langley, S.K., Cheshire, H.M., Humes, K.S., 2001. A comparison of single date and multitemporal satellite image classifications in a semi-arid grassland. Journal of Arid Environments 49, 401–411. https://doi.org/10.1006/jare.2000.0771
Lu, D., Mausel, P., Brondízio, E., Moran, E., 2004. Change detection techniques. International Journal of Remote Sensing 25, 2365–2401. https://doi.org/10.1080/0143116031000139863
Lück-Vogel, M., Mbolambi, C., Rautenbach, K., Adams, J., van Niekerk, L., 2016. Vegetation mapping in the St Lucia estuary using very high-resolution multispectral imagery and LiDAR. South African Journal of Botany 107, 188–199. https://doi.org/10.1016/j.sajb.2016.04.010
Malahlela, O.E., Adjorlolo, C., Olwoch, J.M., 2019. Mapping the spatial distribution of Lippia javanica (Burm. f.) Spreng using Sentinel-2 and SRTM-derived topographic data in

malaria endemic environment. Ecological Modelling 392, 147–158. https://doi.org/10.1016/j.ecolmodel.2018.11.020
Marlin, D., Newete, S.W., Mayonde, S.G., Smit, E.R., Byrne, M.J., 2017. Invasive Tamarix (Tamaricaceae) in South Africa: current research and the potential for biological control. Biological Invasions 19, 2971–2992. https://doi.org/10.1007/s10530-017-1501-6
Mas, J.-F., 1999. Monitoring land-cover changes: A comparison of change detection techniques. International Journal of Remote Sensing 20, 139–152. https://doi.org/10.1080/014311699213659
Mayonde, S., Cron, G., Byrne, M., Gaskin, J., 2016. Tamarix (Tamaricaceae) hybrids: the dominant invasive genotype in southern Africa. Biological Invasions 18, 3575–3594. https://doi.org/10.1007/s10530-016-1249-4
Mayonde, S.G., Cron, G.V., Gaskin, J.F., Byrne, M.J., 2015. Evidence of Tamarix hybrids in South Africa, as inferred by nuclear ITS and plastid trnS–trnG DNA sequences. South African Journal of Botany 96, 122–131. https://doi.org/10.1016/j.sajb.2014.10.011
Mcconnachie, A.J., Strathie, L.W., Mersie, W., Gebrehiwot, L., Zewdie, K., Abdurehim, A., Abrha, B., Araya, T., Asaregew, F., Assefa, F., Gebre-Tsadik, R., Nigatu, L., Tadesse, B., Tana, T., 2011. Current and potential geographical distribution of the invasive plant Parthenium hysterophorus (Asteraceae) in eastern and southern Africa. Weed Research 51, 71–84. https://doi.org/10.1111/j.1365-3180.2010.00820.x
McDaniel, K.C., DiTomaso, J.M., Duncan, C.A., 2004. Common Names and Synonyms 25.
McDaniel, K.C., Taylor, J.P., 2003. Saltcedar Recovery after Herbicide-Burn and Mechanical Clearing Practices. Journal of Range Management 56, 439–445. https://doi.org/10.2307/4003834
Milton, S.J., 2015. Report on clearing of pink tamarisk (Tamarix ramosissima) on Wolwekraal Nature Reserve (Farm 211, part 2), Prince Albert 2014-2015 12.
Milton, S.J., 2004a. Grasses as invasive alien plants in South Africa. South African Journal of Science 7.
Milton, S.J., 2004b. Grasses as invasive alien plants in South Africa. South African Journal of Science 7.
Mortenson, S.G., Weisberg, P.J., Stevens, L.E., 2012. The influence of floods and precipitation on Tamarix establishment in Grand Canyon, Arizona: consequences for flow regime restoration. Biological Invasions 14, 1061–1076. https://doi.org/10.1007/s10530-011-0139-z
Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66, 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001
Mueller-Wilm, U., Louis, J., Richter, R., Gascon, F., Niezette, M., n.d. SEN2COR: Atmospheric and Topographic Correction of Simulated Level-1C Test Datasets 25.
Newete, S.W., Mayonde, S., Byrne, M.J., 2019. Distribution and abundance of invasive *Tamarix* genotypes in South Africa. Weed Research. https://doi.org/10.1111/wre.12356
Ng, W.-T., Meroni, M., Immitzer, M., Böck, S., Leonardi, U., Rembold, F., Gadain, H., Atzberger, C., 2016. Mapping Prosopis spp. with Landsat 8 data in arid environments: Evaluating effectiveness of different methods and temporal imagery selection for Hargeisa, Somaliland. International Journal of Applied Earth Observation and Geoinformation 53, 76–89. https://doi.org/10.1016/j.jag.2016.07.019

Ng, W.-T., Rima, P., Einzmann, K., Immitzer, M., Atzberger, C., Eckert, S., 2017. Assessing the Potential of Sentinel-2 and Pléiades Data for the Detection of Prosopis and Vachellia spp. in Kenya. Remote Sensing 9, 74. https://doi.org/10.3390/rs9010074

Pal, Moo, 2005. Random forest classifier for remote sensing classification. International Journal of Remote Sensing 26, 217–222. https://doi.org/10.1080/01431160412331269698

Pontius, J., Hanavan, R.P., Hallett, R.A., Cook, B.D., Corp, L.A., 2017. High spatial resolution spectral unmixing for mapping ash species across a complex urban environment. Remote Sensing of Environment 199, 360–369. https://doi.org/10.1016/j.rse.2017.07.027

Richards, J.A., Jia, X., 2006. Remote sensing digital image analysis: an introduction, 4th ed. ed. Springer, Berlin.

Richardson, D.M., Van Wilgen, B.W., 2004. Invasive alien plants in South Africa: how well do we understand the ecological impacts?

Sabins, F.F., 1996. Remote sensing : principles and interpretation (3rd ed)., 3rd ed. W.H. Freeman, New York.

Silván-Cárdenas, J.L., Wang, L., 2010. Retrieval of subpixel Tamarix canopy cover from Landsat data along the Forgotten River using linear and nonlinear spectral mixture models. Remote Sensing of Environment 114, 1777–1790. https://doi.org/10.1016/j.rse.2010.04.003

Singh, A., 1989. Review Article Digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing 10, 989–1003. https://doi.org/10.1080/01431168908903939

Smith, L.M., Sprenger, M.D., Taylor, J.P., 2002. Effects of Discing Saltcedar Seedlings during Riparian Restoration Efforts. The Southwestern Naturalist 47, 598–601. https://doi.org/10.2307/3672663

Stevens, R., Walker, S.C., 1998. Saltcedar Control. Rangelands 20, 9–12.

Traganos, D., Reinartz, P., 2017. Mapping Mediterranean seagrasses with Sentinel-2 imagery. Marine Pollution Bulletin. https://doi.org/10.1016/j.marpolbul.2017.06.075

van Niekerk, A., du Plessis, D., Boonzaaier, I., Spocter, M., Ferreira, S., Loots, L., Donaldson, R., 2016. Development of a multi-criteria spatial planning support system for growth potential modelling in the Western Cape, South Africa. Land Use Policy 50, 179–193. https://doi.org/10.1016/j.landusepol.2015.09.014

Warren, D.K., Turner, R.M., 1975a. Saltcedar (Tamarix chinensis) Seed Production, Seedling Establishment, and Response to Inundation. Journal of the Arizona Academy of Science 10, 135–144. https://doi.org/10.2307/40021795

Warren, D.K., Turner, R.M., 1975b. Saltcedar (Tamarix chinensis) Seed Production, Seedling Establishment, and Response to Inundation. Journal of the Arizona Academy of Science 10, 135–144. https://doi.org/10.2307/40021795

Warren, D.K., Turner, R.M., 1975c. Saltcedar (Tamarix chinensis) Seed Production, Seedling Establishment, and Response to Inundation. Journal of the Arizona Academy of Science 10, 135–144. https://doi.org/10.2307/40021795

Wilson, H., Mycock, D., Weiersbye, I.M., 2017. The salt glands of Tamarix usneoides E. Mey. ex Bunge (South African Salt Cedar). International Journal Of Phytoremediation 19, 587–595. https://doi.org/10.1080/15226514.2016.1244163

Xiaohan, Liu., Yunlin Zhang, Kun Shi, Yongqiang Zhou, Xiangming Tang, Guangwei Zhu, Boqiang Qin, 2015. Mapping Aquatic Vegetation in a Large, Shallow Eutrophic Lake: A

Frequency-Based Approach Using Multiple Years of MODIS Data. Remote Sensing, Vol 7, Iss 8, Pp 10295-10320 (2015) 10295. https://doi.org/10.3390/rs70810295

Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. Journal of Plant Ecology 1, 9–23. https://doi.org/10.1093/jpe/rtm005

Zhang, C., Smith, M., Lv, J., Fang, C., 2017. Applying time series Landsat data for vegetation change analysis in the Florida Everglades Water Conservation Area 2A during 1996–2016. International Journal of Applied Earth Observation and Geoinformation 57, 214–223. https://doi.org/10.1016/j.jag.2017.01.007

Zhang, J.-W., D'Rozario, A., Duan, S.-M., Wang, X.-Y., Liang, X.-Q., Pan, B.-R., 2018. Epidermal characters of Tamarix L. (Tamaricaceae) from Northwest China and their taxonomic and palaeogeographic implications. Journal of Palaeogeography 7, 179–196. https://doi.org/10.1016/j.jop.2018.01.003

Zhou, Z., Yang, Y., Chen, B., 2018. Estimating Spartina alterniflora fractional vegetation cover and aboveground biomass in a coastal wetland using SPOT6 satellite and UAV data. Aquatic Botany 144, 38–45. https://doi.org/10.1016/j.aquabot.2017.10.004

(https://en.climate-data.org) accessed on 04 June 2018

(https://sentinel.esa.int/web/sentinel/missions/sentinel-2) accessed on 19 February 2019

(https://www.usgs.gov/land-resources/nli/landsat/landsat-8) accessed on 19 February 2019

(www.harrisgeospatial.com) accessed on 19 December 2018