Faculty of Science

School of Geography, Archaeology and Environmental Studies

An integrated approach for detecting and monitoring the Campuloclinium macrocephalum (Less) DC using the MaxEnt and machine learning models in the Cradle Nature Reserve, South Africa

By

Benjamin Makobe (1685275)

The research report is submitted to the Faculty of Science, University of Witwatersrand Johannesburg in fulfilment of the Degree in Master of Science (Geographical Information System and Remote Sensing) at the School of Geography, Archaeology & Environmental Studies.

Supervisor: Professor Paidamoyo Mhangara

25 March 2024
Johannesburg, South Africa

DECLARATIONS

I, Benjamin Makobe, in my sober senses declare that this research project is my independent work, and it is submitted to meet the requirements for the degree of Master of Sciences in Geographical Information Systems (GIS) and Remote Sensing (RS) at the university of the Witwatersrand, Johannesburg. I hereby declare that this work has never been submitted to any university for examination.

List of published manuscript in peer-reviewed journals

1. **Makobe, B**., Mhangara, P., Gidey, E., Kganyago, M. (2024). Monitoring the invasion of Campuloclinium macrocephalum (less) DC plants using a novel MaxEnt and machine learning ensemble in the Cradle Nature Reserve, South Africa, *Environmental Systems Research, Springer* (***In press***).

.

11/06/2024

DEDICATIONS

This master’s research project is a special dedication to Mama, special dedication to my church and my loving partner.

“*I can do all things through Christ who strengthens me*” **Philippians 4: 13**

ACKNOWLEDMENTS

I extend my heartfelt gratitude to the sovereign God “Yahweh” who carried me through the completion of this project. I am grateful for the unwavering support received from my supervisor Professor Paida Mhangara, providing with necessary tools to complete the project. I am thankful for Gauteng provincial government bursary scheme for funding my university tuitions. Dr Eskinder Gidey, I owe you my life, I would have not completed this project without your assistance, from field data collection, data processing, producing outputs and editing. Dr Mahlatse Kganyago I am grateful for assistance provided during field data collection and inviting the team from South African National Space Agency (SANSA).

Ms Wendy Madungwa, the Cradle nature reserve manager, thank you for allowing us to conduct the field study at the nature reserve. Thank you for providing us with the vehicle and your team in answering all the questions during our field visit. Dr Mncedisi Siteleki I thank you so much for mentoring me. Mrs Litshani Makhuvha, the Environmental Officer at Harmony Gold mining company, Kusasalethu operations. Thank youfor granting me leave (Off-work) to finish my research report. Lastly, I would like to thank my family for their unwavering support and believing in me. To my beautiful partner Rorisang Mathabatha, I thank you for your unconditional love and support. It was through your motivations that I pushed this far; I thank you for believing in my dreams.

ABSTRACT

The invasion of ecosystems by invasive plants is considered as one of the major human-induced global environmental change. The uncontrolled expansion of invasive alien plants is gaining international attention, and remote sensing technology is adopted to accurately detect and monitor the spread of invasive plants locally and globally. The Greater Cradle nature reserve is a world heritage site and intense research site for archaeology and paleontology.It was accorded the world status by the United Nations Educational, Scientific and Cultural Organizations (UNESCO) in 1991 due to its variety of biodiversity present and carries information of significance about the evolution of mankind. The invasion of Campuloclinium macrocephalum (pompom weed) at the Cradle nature reserve is downgrading the world status accorded to the site, lowers the grazing capacity for game animals and replaces the native vegetation. This research study explored the capability of Sentinel-2A multispectral imagery in mapping the spatial distribution of pompom weed at the nature reserve between 2019 and 2024. The non-parametric classification models, support vector machine (SVM) and random forests (RF) were evaluated to accurately detect, and discriminate pompom weed against the co-existing land cover types. Additionally, the species distribution modelling MaxEnt Entropy was incorporated to model spatial distribution and pompom weed habitat suitability. The findings indicates that SVM yielded 44% and 50.7% spatial coverage of pompom weed at the nature reserve in 2019 and 2024, respectively. Whereas, the RF model indicates that the spatial coverage of pompom weed was 31.1% and 39.3% in 2019 and 2024, respectively. The MaxEnt model identified both soil and rainfall as the most important environmental factors in fostering the aggressive proliferation of pompom weed at nature reserves. The MaxEnt predictive model obtained an area under curve score of 0.94, indicating outstanding prediction model performance. SVM and RF models had classification accuracy above 75%, indicating that they could distinguish pompom weeds from existing land cover types. The preliminary results of this study call for attention in using predictive models in predictingcurrent and future spatial distribution of invasive weeds, for effective eradication control and environmental management.

Key words: Campuloclinium macrocephalum; Sentinel-2A; MaxEnt model, Machine learning models, Cradle nature reserve, South Africa.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ASCII	American Standard Code for Information Interchange
AUC	Area Under Curve
COH-WHS	Cradle of Humankind-World Heritage Site
CNN	Convolution Neural Network
EIA	Environmental Impact Assessments
GCP	Ground Control Points
IAPs	Invasive Alien Plants
LAF	Leaf Area Index
NEMBA	National Environmental Management: Biodiversity Act
RF	Random Forests
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
SDM	Species Distribution Modelling
SANBI	South African National Biodiversity Institute
UNESCO	United Nations Educational, Scientific and Cultural Organization
UAV	Unmanned Aerial Vehicle
WHC	World Heritage Convention

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Background

The invasive alien plants (IAPs) are non-native plants introduced accidentally or deliberately into a new geographical environment (Peerbhay *et al.*, 2016). Preston et al. (2018) notes that their introduction has economic and ecological effects. The spread of IAPs alters the functioning of the indigenous ecosystem, because they compete with the indigenous vegetation for water, space and theycan grow across large geographical areas unaided (Peerbhay *et al*., 2016; Sibiya *et al*., 2021). The invasion of IAPs affect the quality and quantity of plant properties such as soil and water retention, consequently IAPs are a threat to biodiversity, human welfare, and agricultural productivity (Kganyago *et al*., 2018; de Villers *et al*., 2020). The South African government has joined the international community in reporting the invasive alien plant status and dispersion at a country level (Kganyago *et al*., 2021, Mafanya *et al*., *2022*). This is done to set funds in exploring eradication methods to minimize their negative impact on the biodiversity and agricultural productivity (de Villers *et al*., 2020; Mafanya *et al*., 2022).

The *Campuloclinium macrocephalun (Less) DC* alternatively known as pompom weed is of *Asteraceae* family, is a foreign hemicryptohytic herb that invades disturbed rangelands (McConnachie *et al*., 2011). The perennial herb, erect up to 1.5m high, has green to purple stems with light purple to pink flowerheads, flowering during summer rainfall (Goodall *et al*., 2011). The weed is originating in Argentina, Brazil and Mexico (SANBI, 2023). In South Africa, pompom weed is becoming one of the major IAPs growing alongside roads, disturbed land, and threatening Savannah rangelands. The weed is becoming a major problem in the country, and it has been categorized as category 1b in Gauteng, Limpopo, Northwest and Mpumalanga provinces and the remaining provinces of the country is proposed to be included in category 1b (SANBI, 2023). In KwaZulu-Natal coast, pompom weed has been reported, and currently is dominant in the central interior Highveld of Gauteng province. Its ferocity has increased since it hosted the invasion in the 1960s (McConnachie *et al.,* 2011). The weed is wind-driven, and human motion has dispersed the powdery seeds through the mud on their wheels and shoes (Goodall *et al*., 2011; SANBI, 2023). The Highveld conditions are more favorable for pompom weed to successfully thrive because frost is prevalent (Goodall *et al*, 2010). The environmental and economic impacts associated with pompom weed it threaten the survival ofgrassland and wetlands (Agricultural Research Council, 2014). It is observed that pompom weed is growing at the UNESCO declared world class heritage site the Greater Cradle nature reserve in Krugersdorp town, Gauteng province. The invasion of pompom weed is threateningthe biodiversity and lowering the world heritage status accorded to the nature reserve.

The general assembly of the United Nations Educational, Scientific and Cultural Organization (UNESCO) in 1972 endorsed a treaty (the World Heritage Convention 1972) with the principle aim of protecting world cultural and heritage sites (UNEP World Conservation Monitoring Centre, 2000). The general assembly ensures protection to the designated sites with outstanding significance and biodiversity within inscribes sites for all humanity and promotes international cooperation (WHC, 2008). The ecosystem and sustainability of inscribed world heritage sites depend on land cover and biodiversity protection. Again, protecting heritage places and monitoring biodiversity can help prevent species extinction and IAP invasion. (Durand *et al*., 2010). Vegetation mapping is critical to understand how different vegetation species alter the environmental conditions of a particular region (Saini and Ghosh, 2012). In South Africa, precise data on the spatial patterns of IAPs is lacking, and there is a need for that spatial data for management and complete eradication of IAPs (Kganyago *et al*., 2018).

The traditional methods of terrestrial mapping, commonly known as site survey are expensive andlaborious and some areas of interest are remotely and difficult to access (Matongera *et al*., 2016; Aldoski *et al*., 2020). The advancement of remote sensing has offered new opportunities to counteract the traditional methods (Ghosh *et al*., 2017; Al-doski *et al*., 2020). Remote sensing is a useful technology for gathering IAP spatial data with comprehensive coverage, temporal observations, and affordability (Dube *et al*., 2020; Mafanya, 2022). The IAP biochemical and biophysical properties can be assessed using electromagnetic energy at different wavelengths (Jensen, 1983). In vegetation mapping studies, the spectral properties form basis of the study and vegetation indices are used for leaf area index (LAF) and biomass estimations (Jensen, 1983). The absorption, reflectance, or transmittance of the electromagnetic radiation within plant's canopy depends on the biophysical properties such as leaf size, structure, and leaf tissue density and the bio-chemical properties such as water, chlorophyll, cellulose (Jensen, 1983).

Remote sensing satellite images range from low to high resolutions. A low-resolution sensor has an advantage of being repeatedly used on short cycles, but low spatial resolution sensors are not suitable for monitoring purposes (JoonGu *et al*., 2018). A high-resolution sensor offers periodic monitoring but hindered by costs and time periods (JoonGu *et al*., 2018). The emergence of unmanned aerial vehicles (UAVs) is seen as an alternative to low and high spatial resolution sensors (Katternboorn *et al*., 2019). The UAVs can record remote data with high temporal and geographical resolution and provide real-time and exact spectrum information (Zhang *et al*., 2021; Narmilan, 2022).

Although a considerable progress has been made in using remote sensing technologies in identification and mapping of IAPs across the globe. However, there is limited literature explaining how different environmental conditions, topographical and biological variables play a crucial role in the infestations of IAPs. These environmental factors are documented as enablers in assisting IAPs to successfully thrive or hinder the adaptability in different areas (Ndlovu and Shoko, 2023). The species distribution models (SDMs) has been an instrumental tool in quantitatively assessing how environmental conditions can aid IAPs to thrive, and predict the possible landscape invasions (Miller, 2010). The SDMs are machine learning algorithms that creates predictive functions based on the observed data of species proliferation and corresponding environmental variables (Valavi *et al*., 2020).

Imperatively, the SDMs assists in understanding the environmental-species interactions, predicting future landscape invasions, and assisting in resource allocations for conservation planning and environmental protection (Mkungo *et al*., 2023). There are several SDMs techniques that are employed in modelling IAP landscape invasions these includes but not limited to maximum entropy (MaxEnt), generalized linear model (GLM), bioclimatic envelope model (BEM), and Logistic Regression are examples. MaxEnt's computing efficiency, robustness, and capacity to interpret partial data made it popular and widely published (Mtengwana *et al*., 2021; Mkungo *et al*., 2023). A study conducted by Ndlovu & Shoko (2023) accounted for the influence of environmental variables such as rainfall and temperature in mapping and future projecting the spread of *L. camara* in Inkomati catchment in the province of Mpumalanga.

Due to the variety of data sources, training data sizes, and difficulties to solve, choosing an image classification algorithm remains difficult (Saini and Ghosh, 2021). Thus far, there is many publications in assessing the performance of various parametric and non- parametric image classifiers in land cover and mapping of IAPs to meet certain scientific and environmental objectives. The parametric image classifiers are distance based and probabilistic, where such an object or feature in an image can be separated or grouped based on the similarity thresholds (Basheer *et al*., 2022). Examples of parametric image classifiers are maximum likelihoodclassifier (MLC), minimum to distance mean, iterative self-organizing data analysis (ISODATA).

In contrast, the non-parametric image classifiers use the analyst created training samples to implement an optimal boundary between objects and examples of this classifier include the robust support vector machine, (SVM), random forest (RF), artificial neural network (ANN) and convolutional neural network (CNN). Due to “big data” the homogenization of machine learning (ML) and deep learning (DL) in remote sensing is being explored to deal with issues of data redundancy and image classification accuracy of land cover (LC) and land use (LU) problems. Thus, this study assesses the efficacy of support vector machine (SVM) and random forests (RF) to accurately model the spatial distribution of invasive forb *Campuloclinium macrocephalun (*Less*)* DC at the Greater Cradle nature reserve, South Africa. Employing Sentinel-2A multispectral bands. The study further sought to use the MaxEnt species distribution model to understand the dynamics and spatial patterns of pompom weed in the research area. In summary, this research project explores the potentiality of using multispectral remote sensing sensor in generating large data in mapping the spatial pattern of invasive weeds on the South African landscape. Additionally, the study will determine how environmental factors influence the adaptability and spatial distributions of invasive weeds.

1.2 Problem statement and justification

The Greater Cradle nature reserve at the Cradle of Humankind situated northwest of Gauteng province is a private game reserve and archaeological site in South Africa. In the year 1999 the Cradle of Humankind was accorded the world historic site for its cultural and environmental outstanding of significance. (Durand *et al*., 2010). The Cradle of Humankind World Heritage Site (COH-WHS) is a major tourism destination and an intense-research site of interest for Paleo-scientists and Archaeologists (Durand *et al*., 2010). The UNESCO assembly of 1972 emphasised that all world historic sites are protected as per the 1972 Convention on Cultural and Natural Heritage treaty adopted (Durand *et al*, 2010). The COH-WHS is protected in terms of the World Heritage Convention Act 1999 (Act no.49 of 1999); National Heritage Resources Act 1999 (Act no. 25 of 1999); the National Environmental Protected Areas Act 2003 (Act no 57 of 2003); the National Environmental Management Biodiversity Act (Act no 10 of 2004); and the Physical Planning Act 1967 (Act no. 88 0f 1967). The above-mentioned legislations serve as a mandatory command to prohibit any mining activities or developments within world heritage sites. In any case of developments environmental impact assessment (EIA) must be performed (UNESCO, 1972). According to the article 11.4 of the convention the appointed heritage committee representing all member states need to report all the inscribed world heritage sites that are threatened by rapid urban developments, illegal land invasions, outbreaks of military, natural or man-made environmental degradation (UNEP World Conservation Monitoring Centre, 2000).

The Greater Cradle nature reserve is currently affected by typical unprecedented growth of highly invasive rangeland weed commonly known as pompom weed. The plant has its origin traceable in South America and Central America. However, in South Africa the exotic weed according to Regulation 15 of the Conservation of Agricultural Resources (Act 43 of 1983), and section 97(1) of Alien and Invasive Species Regulations (2014) of the National Environmental Management: Biodiversity Act (Act 10 of 2004) has been declared invasive/exotic weed and listed as invasive species category 1b plant (SANBI, 2023). The environmental impacts associated with pompom weed, it invades grasslands and wetlands, masking green landscape to pink during Spring-Autum seasons.

1.2.1 Problem statement

The Greater Cradle nature reserve is a preserved site with a significant cultural value and variety of biodiversity. The pompom weed is a declared category 1b “exotic weed” in South Africa. it is threatening the conservation of grassland and savannah biomes in the highveld. The Pompom weed infestation at the UNESCO declared world heritage site is causing environmental degradation and threatening the biodiversity. The uncontrolled expansion of invasive pompom weed lowers the grazing capacity for the game animals, forcing the game animals to migrate. In such the nature reserve loses its significance and the accorded world status. Thus, the Gauteng tourism is compromised subsequently leading to loss of employment due to less revenue generated.

1.2.2 Research justification

The National Environmental Management: Biodiversity Act (Act 10 of 2004) declared to be illegal to plant or germinate invasive weeds. As a result, it is propelled to monitor and eradicate pompom weed to prevent further spreading to other proportions of South Africa and downgrading the heritage status of the Greater Cradle nature reserve. The eradication measures that are currently used to control IAPs include the use of registered herbicides (Plenum, Access, and Climax), uprooting and burning of the weeds (SANBI, 2023). However, such eradication measures remain insufficient in controlling IAPs, they offer temporary solutions. pompom weed grows in spring and it has roots that retain water, and it is one of the IAPs that are resistant to fire and herbicides. The uprooting to some complex, distant areas might be difficult to access.

Therefore, a precise exotic weed detection and management strategies should be effectively implemented. Remote sensing technology has proved successful in monitoring and mapping spatial extent of IAPs in areas that are investigated. Remote sensing provides with precise information and early detection flowering phase of invasive weeds. The information obtained can aid to optimize control of invasive weeds. Additionally, understanding the proliferation and geographical shifting of IAP’s under different environmental conditions, could assists to minimize their impacts on environment and implementing best possible environmental management. The significance of this research study is to provide, where necessary assistance in best environmental management practices and control of pompom weed at the nature reserve employing remote sensing technology.

The following research problems are noted in this study.

I. Several studies have exploited number of different machine learning and deep learning models, to our knowledge never of the machine learning models exploited proved to be superior of the other. Therefore, this gives an opportunity to explore.

II. Mapping of IAPs has offered advancement in environmental management of IAPs, however mapping spatial distribution of pompom weed is not sufficient in addressing the complexes and the rapid spread of pompom weed. Therefore, environmental factors such as soil moisture, topography and rainfall must be investigated in how they influence the distribution of pompom weed and predict landscape susceptible to pompom weed and future infestation.

1.3 Aim and Objectives

1.3.1 Aim

The aim of this study is to compare the performance of support vector machine (SVM) and random forests (RF) models in mapping the spatial extent of *Campuloclinium macrocephalum (*Less.*)* DC at the Greater Cradle Nature Reserve, using multispectral Sentinel-2 imagery.

1.3.2 Objectives

The objectives of the study are as follows:

(i) To assess the efficacy of support vector machines (SVM) and random forests (RF)to accurately model *Campuloclinium macrocephalum* (Less) DC against the existing species using Sentinel-2 multispectral data from 2019 to 2024.

(ii) To model *Campuloclinium macrocephalum* (Less) DC using MaxEnt species distribution model to strengthen the machine learning findings.

(iii) To recommend effective invasive plants eradication measures.

CHAPTER TWO

LITERATURE REVIEW

2. The environmental impacts of invasive alien plants

Given the fastest rate at which global change is happening, it has now gained an international attention to understand on how human-induced activities alter the functioning of biodiversity and ecosystem (Vitousek *et al*., 1997; Tylianakis *et al* 2008). One of the documented key contributors of anthropogenic activities related to global environmental change is the invasion of ecosystems by invasive alien plants (Jensen *et al*., 2020). The Invasive Alien Plants abbreviated as "IAPs" have gained international attention due to their adverse impacts on the biodiversity, altering native ecosystems and associated with high economic costs in weed management (de Villers *et al*, 2020; Mkungo *et al*., 2023). These non-native plants thrive, reproduce, and spread unaided covering large geographic areas subsequently displacing native flora (de Villers *et al*., 2020; Royimani *et al*., 2019). The invasion of IAPs is a threat to the environment especially agricultural land, sensitive protected nature reserves or wetland because they degrade biodiversity and reduce natural ecosystem (Jensen *et al*., 2020). The IAPs outcompete the indigenous vegetation for space, soil nutrients, water retention and consequently it poses a threat to agricultural production and human livelihoods (Kganyago *et al*., 2018; Jensen *et al*., 20220). The South African government is spending an estimated 6.5 billion of rands in eradication of invasive plants (Newete *et al*., 2023).

The infestation of IAPs reduces rangeland that is used by livestock and game animals for grazing. Land-use can also facilitate the infestations of invasive alien plants, as the land or environment is altered by construction of roads and mining activities. It has been observed that

IAPs almost follow an identical trend of infestation, aggressive encroachment of IAPs is on disturbed land due to mining, alongside road constructions and fallow pastural land (Goodall*et al*., 2011; McConnachie *et al*., 2011). The IAPs take advantage of such environments as thesoil is left bare and reduced native flora, therefore less competition (Le Maitre *et al*., 2002). Astudy conducted by (McConnachie *et al*., 2011) revealed that fallow agricultural land, railwaytracks and sideroads are a suitable niche for the survival of *Parthenium hysterophorus* another spreading and problematic invasive plant worldwide. As well as the alien *Compuloclinium macrocephalum* (Less) *DC* follow the disturbance-mediated invasion strategy, invading rangelands with <19% poor vegetation conditions, barren agricultural fields, and wetlands (Goodall *et al*., 2011; McConnachie *et al*., 2011). Several studies indicated that climate change influence the infestations of IAPs in different geographical (Mtengwana *et al*., 2021; Mkungo *et al*., 2023; Ndlovu and Shoko, 2023).

It is also noted that changes in environmental conditions such as increased nitrogen levels, CO_2 concentrations, variability in precipitation may also influence the proliferation of invasiveweeds (Huang *et al*., 2009). Under climate change it is expected that some plant species will extinct, while other plant species will survive and reproduce effectively under the new climatic conditions. Globally, accurate and detailed information on the spatial distribution of IAPs is needed to maximize effective control mechanisms and minimize their impacts on the environment (Kganyago *et al*, 2018). On a country level, South Africa is also conducting studies to understand the status and spatial distribution of IAPs. As that geospatial information might assist in implementation of mitigations and rehabilitation of the affected areas (Mafanya *et al*., 2022).

2.1 The introduction and environmental facilitators of *Campuloclinium macrocephalum* (Less) DC invasion of South African Savannah-rangeland

The earliest record of pompom weed, *Campuloclinium macrocephalum* (Less) DC in South Africa was in 1960s, however it is unascertained of how pompom weed was introduced (McConnachie *et al.*, 2011). The alien pompom weed has its origin in South America and Central America, between 1990s and 2000s it entered a dramatic expansion in South Africa invading grasslands and wetlands particularly in the interior Highveld (McConnachie *et al.*, 2011). The prevalent frost conditions of Highveld regions favoured pompom weed to successfully thrive unaided. The Gauteng province of South Africa was the first to be invaded by pompom weed in the 1960s, before being spread to other proportions of South Africa (Goodall *et al.*, 2011). The KwaZulu Natal, Mpumalanga, Northwest, Limpopo, and Free State provinces the presence of pompom weed is noted. In the Outeniqua mountains near the Western Cape province the small patch of pompom weed has been eradicated with monitoring in place (Agricultural Research Council, 2014). The absence of natural vegetation is cited as the primary facilitator in rapid invasion of invasive alien plants, particularly pompom weed (Goodall *et al.*, 2011). pompom weed highly invades rangeland that is impacted by unsustainable commercial grazing methods, degraded cultivated lands and wetlands (Goodall *et al.*, 2011).

The absence of natural vegetation and high fire frequency increases the severity of pompom weed density, providing opportunity for pompom weed to germinate and establish themselves (Goodall *et al.*, 2011). pompom weed invasion into rangeland is through human activities and wind dispersal, weed dispersal is a dominant factor. About 90% of pompom weed infestations occurs along roadsides (Goodall *et al.*, 2011). Therefore, the roadsides are another favourable niche of pompom weed infestations as they feed on the carbon from the vehicles. The areas that receive >600mm of rainfall per annum are affected by invasion of pompom weed, even though the weed exhibits disturbance-mediated invasion strategy (Goodall *et al.,* 2011).

2.2 Overview of different remote sensing sensors

Remote sensing technology provides with the opportunity of collecting detailed spatial data without physical contact with the object of interest. The spatial information is sensed and collected from the earth surface in the form of the reflected electromagnetic radiation (Rudrapal and Subbhedar, 2015). Satellite sensors range from coarse spatial resolution (CSRS),medium spatial resolution (MSRS) to high resolution (HSRS). It is imperative to understand the pros and cons of different sensors with regards to spatial and temporal resolutions, acquisition time and costs (Al-dowski *et al*., 2020). The user’s experience with different sensors is as well an advantage to know which sensor is suitable for land cover mapping (Al- dowski *et al*., 2020). In an example the Landsat TM images have few spectral bands with widewavelengths that make it difficult to differentiate land use and land cover (LULC) classes, and hyperspectral images consists of significant number of spectral bands with narrow wavelengths which can help to enhance classification accuracy (Al-dowski *et al*., 2020).

Coarse spatial resolutions such as MODIS have one to two days temporal resolution and spatial resolution of 250m in (Bands 1-2), 500m in (Bands 3-7) and 1000m in (8-36 Bands). These sensors are open-source data covering about 36 bands than any other satellites (Chen *et al*., 2018). Medium spatial resolutions (15-120m) examples of this sensor are the Advanced Spaceborne Thermal Emission and Reflection Radiometer abbreviated as ASTER, Advanced Land Imager (ALM), Landsat-TM and ESA Sentinel-2. The Landsat and the ESA Sentinel-2 are the twocommonly open-source and user-friendly sensors. They are used for land cover mapping at regional and national scale (Al-dowski *et al* 2020).

The Sentinel-2 imagery consists of the visible, near-infrared, and red edge bands. The red-edge bands of Sentinel-2 provide the opportunity to derive biophysical variables like Leaf Area Index. The enhanced spatial, spectral, and temporal resolutionof Sentinel-2 sensor gained popularity in vegetation mapping (Adagbasa *et al.,* 2022). The highest sensitivity of the red-edge band in vegetation detection, has made it easy to distinguishbetween vegetation and other land use types (Ndlovu and Shoko, 2023).

High spatial resolution sensors include IKONOS, Quickbird-2 and SPOT 4 and 5 the elevated space resolution is 0.60-20m and the temporal 1-3.5 days (Al-dowski *et al*., 2020). These high spatial resolution sensors cannot have heat bands and they are incapable of showing surface temperature mapping when using energy equilibrium techniques (Al-dowski *et al.,* 2020). In addition, the archived data for a particular region from these sensors might not be easily accessible or quickly retrievable if demanded (Al-dowski *et al*., 2020). High spatial resolution sensors are used for land cover mapping at local level (Jensen, 2015).

2.3 The application of remote sensing in detection and mapping of IAPs

Traditionally, field-based surveys were used to collect information about the spatial distribution of IAPs, however such methods are not precise accurate, observer bias, time-consuming and labour intensive (Royimani *et al.*, 2019; Mafanya *et al*., 2022). Remote sensing technology is proving to be a viable tool for monitoring earth dynamics changes, for example environmental change due to alien plants invasion. Kganyago et al (2018) argues that it is advantageous to use remote sensing because of its comprehensive areal coverage. The ability of remote sensing to capture IAPs electromagnetic energy in the various wavelengths, allows the detection and assessment of IAPs biochemical and biophysical properties (Jensen, 1983).

The use of remote sensing data offers a great opportunity in mapping the spatial distribution of IAPs at different landscapes and different phenological seasons (Royimani *et al*., 2019). The capturing of remote sensing imagery at different growing seasons of plants can facilitate reliable monitoring of the phenological changes of the plants, particularly in the issue of IAPs management. For precise detection and accurate mapping of IAPs can be done using imagery taken at different seasons during the non-flowering and flowering of IAPs (Royimani *et al*., 2019). The IAPs have distinct biochemical and biophysical features that are distinct to native vegetation. These can be detected and differentiated from co-existing vegetation with remote sensing imagery captured at different phenological seasons or different periods (Jensen *et al*., 2020).

2.4 Multispectral remote sensing sensors in IAPs mapping

The multispectral satellite remote sensing sensors provides with large swath-width data and affordable datasets that can be used for various earth observations (Royimani *et al*., 2019). The multispectral satellite sensors typically measure the emitted and reflected energy of an object of interest within the multiple broadbands of 5 to 12 or going up to 36 bands of the electromagnetic spectrum. Multispectral sensors have the following bands the visible green, visible red and near-infrared, examples of multispectral band sensors are the MODIS, Landsat Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+) and ASTER. The revisiting period of the multispectral sensors allows for monitoring of AIPs and the affordability of these datasets accommodates the developing countries which are still battling with IAPs management. There is a vast published literature that used multispectral imagery in mapping of IAPs. For instance, Wilfong et al (2009) mapped *Lonicera maackii* (Rupr) herder (Amur honeysuckle) in the northeast of United States using Landsat TM and Landsat ETM+ imagery for the month of November 2005 and June 2007.

Kganyago et al (2018) evaluated the potential of Landsat 8 OLI and SPOT-6 in mapping the distribution and patch sizes of *Parthenium hysterophorus* northern KwaZulu Natal savannah landscapes, South Africa. The results from their study indicated that SPOT-6 had 86% overall accuracy as compared to 83% overall accuracy of Landsat 8 OLI (Kganyago *et al.*, 2018). The high spatial resolution multispectral sensors such as ESA Sentinel-2, WorldView-2, IKONOS and Quickbird were strategically designed to improve the discrimination of vegetation types as compared to when using low spatial resolution multispectral band sensors in vegetation monitoring studies (Royimani *et al*, 2019). A study by Ngubane et al (2014) utilized high spatial resolution WorldView-2 and medium spatial resolution SPOT-5 in detection of Bracken fern in Durban, South Africa.

The reported overall classification accuracy shows high spatial resolution WorldView-2 outperformed the medium resolution SPOT-5 with 91.67% and 72.22% respectively. Sentinel-2 data is widely cited in modelling, classification, and prediction studies. A study by Mtengwana et al (2021) explored Sentinel-2 MSI and species distribution modelling (SDM) techniques: Random Forests (RF), Maximum Entropy (MaxEnt) and Boosted Regression Trees (BRT) to predict areas that may be at risk of being invaded by IAP’s in the Heuningnes, South Africa. The study provided a baseline in which SDM predictive models can be used to predict landscape susceptible to IAP’s invasions under the current and future climatic events. Despite the improvements with the bands of multispectral satellite remote sensors in discrimination of earth features, the multispectral sensors are still impeded by poor spatial and spectral resolution in vegetation monitoring (Ngubane *et al.*, 2014). Performing image classification on multispectral satellite imagery issues of mixed spectral and mixed pixel confusion are commonly encountered.

2.5 Hyperspectral remote sensing sensors in IAPs mapping

The employment of hyperspectral is added to overcome the spectral resolution challenges encountered when using multispectral sensors. The hyperspectral sensor is made up of hundreds and continuous narrower spectral bands and collecting datasets at 2-16nm bandwidth (Lass *et al*., 2002). For vegetation studies purposes, the narrower and hundreds of bands of hyperspectral imagery offers a great potential to delineate intra and inter species spectral variations (Kganyago *et al*., 2017; Royimani *et al*., 2019). The advantage of using hyperspectral imagery it can discern different tree species within a forested area and produce superior classified imagery even on heterogenous landscape (Royimani *et al*., 2019). This is a good potential of hyperspectral imagery unlike working with multispectral imagery which can only help with mapping the forested area. While the design of multispectral imagery has a great potential in differentiating land and water features. A study by Hunt et al (2002) demonstrated the potential use of hyperspectral sensor in mapping of IAPs, in their study they compared the performance of Visible/Infrared Imaging Spectrometer (AVIRIS) and Landsat ETM+ and SPOT-4 in mapping the *Euporbia esulal* (Leafy spurge) in Devils Tower National Monument. Crook County in the USA. The findings showed that hyperspectral AVIRIS yielded an overall accuracy of 74% when compared to Landsat ETM+ and SPOT-4 which yielded overall accuracy of 49% and 48% respectively. Mafanya et al., (2022) also demonstrated the utility of hyperspectral imagery in IAPs mapping, they used 30m DESIS hyperspectral for mapping invasive *Campuloclinium macrocephalum* (Less.) DC (pompom weed) in Pretoria city heterogenous landscape, South Africa. The hyperspectral remote sensing data is effective in mapping IAPs, the improved spectral resolution allows the discrimination of IAPs from co-existing vegetation based on their biochemical and biophysical properties. The launch of Hyperspectral sensor marked a new era and essential development in vegetation studies; however, the acquisition cost and computational tasks associated with hyperspectral remains a challenge. The hyperspectral narrow and contiguous bands tend to correlate with each other, often resulting in over generalisation error of the classifier.

2.6 The integration of different remote sensing sensors in IAPs mapping

The integration of various data sources or multisource is adopted to improve the classification accuracy. The fusion of two or more remote sensing dataset for the same classification purposes is increasingly evaluated in vegetation monitoring, to counteract the limitations offered by other remote sensing datasets (Peerbhay *et al.*, 2016). Even though there is a gap in data integration for optimal detection and mapping of IAPs, a quite progression has been made. For instance, Peerbhay et al (2016) integrated Eagle airborne hyperspectral AISA with LiDAR imagery and high spatial resolution WorldView-2 with LiDAR. The results showed that AISA with LiDAR earned higher overall accuracy of 78% than 74% for WorldView-2 with LiDAR (Peerbhay *et al.*, 2016). The fusion of the hyperspectral and high spatial resolution with LiDAR demonstrated a significant improvement of image classification accuracies. The integration of AISA with LiDAR yielded higher overall classification accuracy as compared when AISA (68%,) WorldView-2 (63%) and LiDAR (64%) used alone (Peerbhay *et al.*, 2016). Their study is a good push in integration of different remote sensing sensors to map the geospatial distributions of invasive plants.

2.7 The potential use of UAV in mapping of alien invasive plants (IAPs)

The Unmanned Aerial Vehicles (UAV) are emerging as a potential technological tool for monitoring land surface activities. The characteristic of UAV includes its capability of using single band, multispectral and hyperspectral sensors operating within the visible, infrared or microwave of the spectrum (Chen *et al*., 2016). The cost benefits of UAVs over satellites are that they are flexible, mission specific and versatile (Wyard *et al*., 2022). A study by Wyard *et al*., (2022) developed guidelines in acquiring data for land cover mapping using UAV multispectral imagery for a landfill site Wallonia in Belgium. The findings of the study reveal that object-based image analysis (OBIA) supervised classification is suitable for very high-resolution dataset such as UAV for land cover mapping (Wyard *et al*., 2022). The OBIA is suited for land cover mapping because of its characteristics considering shape, texture, and spectral properties (Wyard *et al*., 2022). The UAV at spatial resolution of 10cm and low-cost red-green-blue sensors are enough for the purpose of land cover mapping (Wyard *et al*., 2022).

A study by Kattenborn et al (2019) assessed the potential employment of UAV for semi-automatic sampling acquisition on coverage of three woody invasive species *Pinus radiata*, *Ulex europaeus* and *Acacia dealbata* in Chile. Their study involved training regression models in mapping the cover fractions of the above-mentioned woody species using UAV andSentinel-1 and Sentinel-2 reference data. The UAV produced accurate and highest accuracy when compared to the two ESA missions, therefore demonstrating the potential of using UAVdata as an alternative to traditional field surveys and other remote sensing sensors in delineating and mapping Invasive plants (Kattenborn *et al*., 2019).

2.8 The use of spectral reflectance indices in detection and mapping of IAPs

Since the inception of vegetation monitoring studies, it has heavily relied on the observation of spectral reflectance properties of plant canopy conditions to estimate crop yield using remote sensing techniques (Lass *et al*., 2005; Adagbasa *et al*., 2022). Spectral indices techniques involve the combination of two or more spectral bands that accentuates the relative abundance of features of interests, in such an analyst can distinguish between soil and vegetation on the imagery (Royimani *et al*., 2019). Several studies revealed that different vegetation exhibit different spectral reflectance within the regions of electromagnetic spectrum, enabling the ability to assess vegetation or plant's biochemical and biophysical properties (Jensen, 1983). As well it has been revealed that the native plants have different biophysical and biochemical properties from non-native plants (Jensen *et al*., 2020). The invasive plants distinct biophysical and biochemical features facilitate the superior infestations taking advantage of the ecological niche of native flora (Rajah *et al*., 2020). The satellite derived spectral vegetation indices such as the Normalized Difference Vegetation Index (NDVI), Principal Component Analysis (PCA) and Enhanced Vegetation Index (EVI) have been exploited to aid in accurately mapping the landscapes or heterogenous landscapes invaded by invasive alien plants.

Matongera et al (2017) showed the potential of NDVI in classification of Braken fern computed on Landsat 8 and WorldView-2. The results yielded are 82,93% and 87,8% for Landsat-8 and WorldView-2 respectively. Rajah et al (2020) conducted a study to assess the efficacy of using vegetation indices and the optical bands of Sentinel-2 satellite imagery in mapping of American Bramble (*Rubus cuneifolius*) which invaded the Ukhahlamba Drakensberg Park in KwaZulu Natal province. The support vector machine (SVM) algorithm was used to determine the optimal season for mapping *Rubus cuneifolius*. The outcome of the study showed that highest classification accuracy was obtained using optical bands, therefore the study pushes the implementation of using vegetation indices and optical bands in determining the distribution of invasive plants using cost-free remote sensing datasets. Despite the successful use of vegetation indices in IAPs mapping, their disadvantage is that they are not reliable in different sites with different crop types and NDVI indices have saturation problem especially in high canopy densities (Royimani *et al*., 2019). The vegetation indices are not viable in mapping large areas.

2.9 Application of MaxEnt species distribution model in IAPs monitoring

The detection and constant monitoring of IAPs using remote sensing technologies has offered an opportunity in environmental management. Spatial mapping of IAPs and considering the influence of fluctuating environmental conditions such as temperature, rainfall, soil moisture and aspect could explain the infestations of IAPs in new ecosystems (Mtengwana *et al*., 2021; Ndlovu and Shoko, 2023). Understanding the environmental factors influencing the dynamics of IAPs and constant monitoring of vulnerable ecosystems could minimize the risks of possible land invasions and implementing effective land management. The successful infestations of IAPsinto new ecosystems is driven by changing environmental conditions, anthropogenic activitiesinfluence on climate change (Mtengwana *et al*., 2021). Climate change in Africa is imminent, and it is projected that temperatures may increase between 3°C and 6°C (Serdeczny *et al*.,2016). The regions of Southern Africa may experience high temperatures and frequentdroughts. These extreme climatic events may cause mass extinctions of native vegetation asthe new environmental conditions are not favourable for native vegetation, this would allowopportunistic uncontrolled growth of IAPs due to less competition (Mtengwana *et al*., 2021). In the light of environmental conditions fostering the geographic spread and proliferation ofIAPs, there is a need of seeking holistic and practical methods to predict the current and futureinfestations of IAP's. To empirically minimize areas that may be susceptible to invasion and to ensure that biodiversity is not compromised. The Species Distribution Modeling (SDM) andmachine learning (ML) techniques proved successful to model the spatial distribution anddynamics of IAPs under different environmental conditions (Ndlovu and Shoko 2023).

The Maximum Entropy (MaxEnt) is an ecological model widely used in ecological applications, precisely modelling species distribution (Mtengwana *et al.*, 2021) The MaxEnt proved to be effective in prediction and being able to work with “presence only” data and handling autocorrelations. A study by Mtengwana et al (2021) explored using Sentinel-2 sensor, the bioclimatic and topographic data in predicting the spatial distribution of IAPs under changing climatic conditions at Heuningnes catchment. The Boosted Regression Tree (BRT), MaxEnt and Random Forest (RF) were used to predict the spatial distribution of IAPs within the catchment, the results yielded Area Under Curve (AUC) of 0.89,0.92 and 0.94 respectively. Their study successfully proved that SDM and machine learning algorithms could be used to predict the potential landscape that is vulnerable to IAPs invasion.

A study by Mkungo et al (2023) modelled the future distribution of Bracken fern (*Pteridium aquilinum*) at eThekwini metropolitan municipality. The study further aimed at identifying vulnerable landscape using bio-physical, climatic, and remote sensing techniques using MaxEnt species distribution model. The study formulated four modelling scenarios: model 1 (topographic variables), model 2 (topographic variables and Sentinel-2 spectral bands), model 3 (topographic variables and bio-climate variables) and model 4 (topographic variables, Sentinel-2 spectral bands and bio-climate variables). The four models produced differing results, model 1 and model 2 predicted the central part of eThekwini municipality susceptible to the Braken fern invasion, while model 3 and 4 predicted the central and western parts of the municipality with the potential of Bracken fern invasion. Their findings are consistent with other reported studies, where SDMs and remote sensing techniques are used to model the spatial distributions of IAPs and identify potential landscape that is at risks of IAPs (Mtengwana *et al*., 2021; Ndlovu and Shoko, 2023).

2.10 The Supervised, Unsupervised classifications and integration of machine and deep learning in remote sensing

Not only using remote sensing imagery is a challenge, but also selecting the suitable image classification algorithms for mapping that is able to produce reliable and accurate outputs (Royimani *et al*., 2019). Thus far, to the best of our knowledge little information is known on the synergies of the type of remote sensing dataset that is coherently in conjunction with the ideal image classification algorithms. In remote sensing, image classification has been done through supervised or unsupervised classification processes. The supervised image classification is explained as a human-computer aided classification process in which training samples are created to classify the imagery based on the created training samples. On the other hand, the unsupervised image classification process is carried out without prior information known or creating training samples, the process involves generating classes or grouping of classes based on their reflectance values without being aided by an analyst.

The classification processes are subdivided into parametric and non-parametric image classifiers, the parametric image classifiers include but not limited to, maximum likelihood classification (MLC), Minimum Distance to Mean (MDM) spectral angle mapper (SAM) and iterative self-organizing data analyst technique (ISODATA). These parametric image classifiers are well applied and documented; however, their success in remote sensing image classification is losing significance because they are scene-dependent, *priori* assumptions about data being normally distributed (Royimani *et al*., 2019). The inability to overcome mixedpixels classification in heterogenous landscape (Civico, 1993; Royimani *et al*., 2019). The image classification output is in pixel level form and pixels deviating from training class data are often misclassified or unclassified (Civico, 1993). The output is in the form of pixel level that hinder the classification accuracy of course and medium spatial resolution multispectral sensors. The parametric image classifiers assume normal distribution, that the chosen signature for classification process represents the entire feature or surface being studied.

Furthermore, parametric classifiers are tedious to use because they require large training data often resulting in poor classification. The non-parametric image classifiers examples include, support vector machine (SVM), random forest (RF) and deep learning convolutional neural networks (CNN). The advantage of non-parametric image classifiers is that no prior assumptions about the data is made, they are flexible and adaptable whether working with small or large training samples. The non-parametric image classifiers emerged to improve the capabilities in retrieval of biophysical features of vegetations and the use of statistics could assist in extracting vegetation information. The integration of machine learning (ML) and deep learning (DL) is being experimented in remotely sensed image classification. During the image classification process the machine learning algorithms allows training data migration to sites not visited, while the deep learning like CNN algorithm is able to counteract problems with intra-class spectral variability and spatial heterogeneity (Mafanya *et al.*, 2022). Therefore, the application or working with machine learning and deep learning algorithms could enhance the mapping of IAPs covering large areas that were not previously observed, and further detect IAPs on heterogenous landscape.

2.11 Image classifications models in mapping of IAPs

The assessment of using parametric and non-parametric image classifiers performed on multispectral, hyperspectral, or multisource data is continuously improving, with the aim of meeting certain objectives in mapping IAPs and as well determining the best classifier (Mafanya *et al.*, 2020). There is a vast publication studies that assessed the performance of various image classifiers. A study by Mafanya et al (2022) assessed the potentiality of using machine-learning algorithms in mapping the IAPs using the recently launched DESIS hyperspectral imagery. In their study they trained three image classification algorithms namely: maximum likelihood classifier (MLC), support vector machine (SVM) and spectral angle mapper (SAM) to map the invasive *Campuloclinium macrocephalum (Less.) DC* in a heterogeneous Pretoria landscape of South Africa (Mafanya *et al*., 2022). The results from their study showed that SAM outcompeted SVM and MLC yielding classification accuracy of 87%, 67% and 73% respectively. The overall study provided a potential solution in using spaceborne hyperspectral imagery in generating large number of training samples that can detect and map the distribution of IAPs in heterogeneous landscape.

Another study was by Kganyago et al (2018) they evaluated the capability of spatial and spectral configuration of Landsat 8 OLI andSPOT 6 imagery in mapping the spatial distribution and patch sizes of alien plant *Partheniumhysterophorus*. The robust SVM classifier was computed on both datasets. Their results showedthat in terms of mapping *P. hysterophorus* in the savannah landscape of Kwazulu-natal, SPOT6 outcompeted Landsat OLI gaining an overall accuracy of 86% and 83% respectively. SPOT 6 proved to be capable of mapping small patches of *P.hysterophorus* and this can assist in control and eradication to prevent further spreading in other locations (Kganyago *et al.*, 2018).

De Villers et al (2020) evaluated the performance of two machine learning algorithms, the support vector machine (SVM) and random forest (RF) in mapping the spatial invasion extent of *Prosopis glandulosa* alien plant and land cover change in Prieska, Northern Cape using multi-temporal Landsat data from 1990 to 2018. The performance of the two algorithms almost indicated similar overall accuracies for 1990 to 2018 period. The SVM overall accuracy ranged from 61% to 89% and RF 57% and 83% overall accuracy (de Villers *et al*., 2020). Their study showed the importance of employing Landsat imagery in mapping historical and current land cover dynamics of IAPs (de Villers *et al*., 2020).

CHAPTER THREE

MATERIALS AND METHODOLOGY

3.1 Study area

To fulfil the objectives of this project the Greater Cradle Nature Reserve is chosen as a study area, which is situated between Johannesburg and Pretoria, two of the country's largest cities in the Gauteng Province of South Africa. It is located between 27°43'30ˈˈE and 27°540ˈ E and 27°57'0ˈˈ S to 25°50'0ˈˈ S on the Kromdraai Road (Figure 1). The Greater Cradle Nature Reserve is a privately owned entity covering 3000 to 9000 hectaresof land made up of pristine dolomite grassland on Muldersdrift farm near the town ofKrugersdorp in the Mogale city municipality. The Blaaubank river valley runs on the northernside and is of exceptional palaeoanthropological value as it accommodates the sites (Kroomdraai, Swartkrans, Coppers cave and Bolts farm) that document over three and half million years of landscape, faunal, environmental, and human evolution (Stradford *et al.*, 2016). For over decades the Cradle of Humankind site, which cover the Greater Cradle naturereserve became an extensive research hub for archaeologists and anthropologists owing to discovery of fossilised remains and the site is linked to the origin of the modern race of humanbeings. Due to extensive fossils discovered and the variety of plant and wildlife species preserved in the nature reserve representing diversity of South Africa, the international forumthe United Nations Educational and Scientific Council Organisation (UNESCO) accorded theGreater Cradle Nature Reserve a world heritage site.

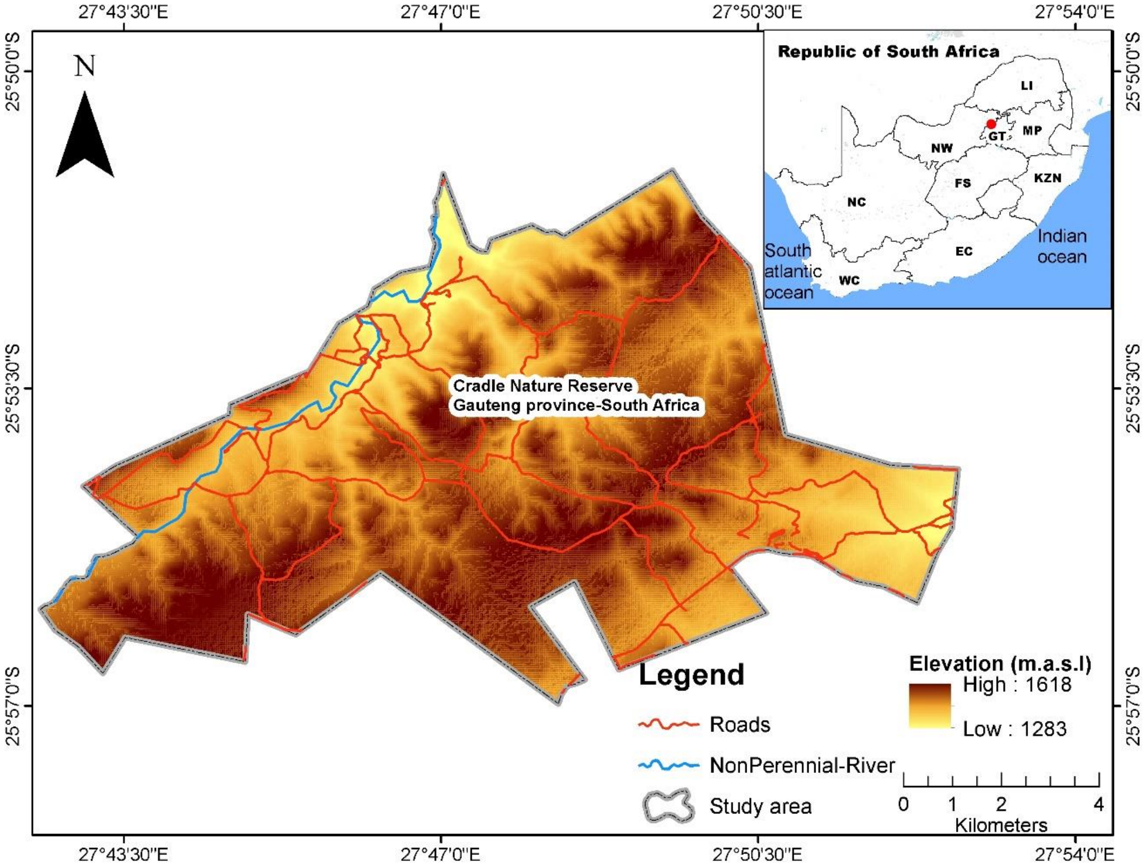

Figure 1. Study area location: The Greater Cradle Nature Reserve, Krugersdorp. South Africa

3.2 Target species description

The pompom weed, *Campuloclinium macrocephalum (Less) DC* is a native plant to Central and South America belonging to *Asteraceae* family (McConnachie *et al*., 2011). Pompom weed is a 1.5-metre-high perennial, erect herb flowering mostly in Spring-Autum season (McConnachie *et al*., 2011; Mafanya *et al*., 2022). The plant has fluffy pink flowerheads, with light green leaves that are scattered around the full length of the green stem and clustered at the base of the plant to form like a rosette (McConnachie *et al*., 2011). The plant starts flowering in spring, early October and die in autumn in April (Gooddall *et al*., 2011). In South Africa, pompom weed was introduced in the 1960s for ornamental purposes, however between the 1990s and 2000s the weed entered a dramatic expansion becoming a noxious weed (McConnachie *et al*, 2011). Currently, 7 out of 9 South African provinces, the pompom weed has been recorded with high infestations in Gauteng province. The invasive weed invades grassland and savannah biomes where it has detrimental impact on the biodiversity, it degrades the rangeland and reducing the grazing capacity for large herbivores (Goodall *et al*., 2022).

3.3 Methods of data acquisitions

3.3.1 Field data collection

On 25th and 26th January 2024, when pompom weeds had aggressively encroached the Cradle Nature Reserve, field data collection took place. It is important to note that the initial field survey was conducted in winter of June 2023, when the pompom weed had completely died. Followed by second survey in November 2023 during the flowering stage of the pompom weed. However, due to erratic rainfall received in 2023 caused the weed to flower late, resulting in sparse availability in the nature reserve. The field survey of 25th and 26th January 2024 was conducted following the full flowering and encroachment of pompom weed, which transformed the nature reserve into a vibrant pink-purple colour. A 10 m × 10 m plot was constructed which was heterogenous, composed of pompom weed and other co-existing land cover types. This was done to ensure intra-species variability and well representation of pompom weed and the co-existing land cover types. A total of 141 ground control points (GCPs) containing pompom weed and varying land cover types were randomly collected within the study area using a handheld Garmin Etrex 10 Global Positioning System. Another 328 GCPs containing only pompom weed were collected randomly within the nature reserve on 27th February 2024, this reference data was used for species modelling on MaxEnt software. A total of 469 samples were utilized.

A priori knowledge of the study area was incorporated and additional land cover types (bushland, riparian zones, water bodies, bareland and cropland) were generated on satellite imagery by pixel classification. The GCPs were tabulated on Microsoft Excel spreadsheets, saved them as comma-separated values (CSV), imported them into Google Earth Engine Pro, and superimposed them on the shapefile of the study area. The CSV format is compatible for MaxEnt software.

The ground truth data was splinted into 70:30 ratio training and testing respectively, for the purpose of classification models (SVM and RF) and MaxEnt species modelling.

3.3.2 Remote sensing data acquisition

The multispectral Sentinel-2A data was acquired from Copernicus Open Access Hub (https://scihub.copernicus.eu/). The 10m spatial resolution Sentinel-2A acquired coincide with the same date the field data was collected (i.e., 26 January 2024) and the cloud cover was below 10%. The European Space Agency (ESA) launched two identical satellites Sentinel 2A and 2B in June 2015 and March 2017, respectively (Noi *et al*., 2017). The Sentinel-2 was designed for land and coastal applications and the satellite gained attention due to its free access and global coverage (Noi *et al*., 2017). The state-of-the-art Sentinel-2 carries a Multispectral Imager (MSI) at a swath width of 290km that allows for land monitoring purposes (ESA, 2016). The 13 spectral bands ranges from 10m to 60m per pixel sizes and captures the data within the visible, near-infrared (NIR) and shortwave infrared (ESA, 2016). The temporal resolution is 10 days when performed with single satellite, and when combined the satellites the deliver the data at a spatial resolution of 10m to 60m in 5 days revisit time (Topaloglua *et al.*, 2016 in Miranda *et al*., 2018). The sensor's 13 bands are as follows Table 1:

Table 1: Sentinel-2 A image characteristics

Band	Resolution	Central Wavelength	Description
B1	60m	443 nm	Ultra blue
B2	10m	490 nm	Blue
B3	10m	560 nm	Green
B4	10m	665 nm	Red
B5	20m	705 nm	Visible & NIR
B6	20m	740 nm	Visible & NIR
B7	20m	783 nm	Visible & NIR
B8	10m	842 nm	Visible & NIR
B8A	20m	865 nm	Visible & NIR
B9	60m	940 nm	SWIR
B10	60m	1375 nm	SWIR
B11	20m	1610nm	SWIR
B12	20m	2190nm	SWIR

The Sentinel-2 multispectral placed as a good sensor over other widely used sensors when working with vegetation monitoring studies. The Sentinel-2 MSI red edge bands between 705nm and 750nm offers an improved retrieval accuracy of crop biophysical variables such as chlorophyll content in the leaf, and leaf area index (Xie et al., 2019). The additional vegetation sensitive bands of Sentinel-2 MSI combined with environmental factors, offers the capability of enhancing species discrimination and enhancing the accuracy of prediction and mapping processes (Mtengwana *et al.*, 2021).

3.3.3 Image pre-processing and analysis techniques

Image pre-processing is a prerequisite for geometric and radiometric errors to be corrected, as well as atmospheric correction to get a cloud free study area (Wong and Sarker, 2014). Sen2Cor is an external plugin in SNAP software that processes Sentinel-2A satellite data that generates Level 2A products and formats outputs. It can be downloaded freely from http://step.esa.int/main/download/. The downloaded Sentinel-2 MSI for 26 January 2024 had 10% cloud coverage, whereas the historical satellite imagery for 26 January 2024 was cloud free (0%). The Sentinel-2 MSI Band 4(Green); Band 3(Green) and Band 2 (Blue) which reflect true-natural colour composite were selected.

A classification approach in which a user oversees the pixel-classification process is known as supervised classification (Miranda *et al*. 2018). For land cover mapping, the user selects various pixel values or spectral signatures that represent a specific class (Miranda *et al*. 2018). The supervised classification process begins by identifying sample locations for the various types of land cover training sites. Following that, the computer algorithm codes the training site's spectral signature and sorts the whole remote sensing image into groups based on the pixel values or the spectral reflectance of the different types of land cover and land use (Civico 1993; Miranda *et al*. 2018; Akalu *et al*. 2019). Ideally, classes should not overlap or overlap slightly with other classes (Miranda *et al*. 2018). This study utilized ESRI ArcGIS Pro to perform supervised image classification. The SVM and RF classification models were employed to perform supervised image classification.

3.4 Supervised image classifications models

Supervised classification is a classification approach in which the user supervises the process of pixel classification (Miranda *et al*., 2018). In the case of land cover mapping the user will select various pixel values or spectral signatures that will represent a specific class (Miranda *et al*., 2018). The supervised classification process will begin by extracting sample locations of land cover types "training area/site" and the computer algorithm will code the spectral signature of the training site, and then classify the entire remote sensed image based on the chosen pixel values or spectral reflectance parameter of LULC classes (Civico, 1993; Miranda *et al.*, 2018). Ideally, it is encouraged that class should not overlap or slightly overlapwith other classes (Miranda *et al*, 2018). This study adopted supervised image classification and was performed on ESRI ArcGIS Pro, the license was acquired through the University of the Witwatersrand. To perform supervised image classification, support vector machine (SVM)and random forests (RF) classification models are employed. The evaluation performance of SVM and RF models was performed on STATA statistical data software version 14.

3.4.1 Support Vector Machine

The support vector machine (SVM) is a widely cited supervised machine learning model usedfor classification and regression modelling (Vapnik 1999; Kganyago *et al.*, 2018). The SVM works by separating different classes by fitting a hyperplane between the training datasets. Thedistance between the classifier and the training datasets is indicated by the margin, and the hyperplane with the maximum margin is selected. The optimal hyperplane with maximum margin reduces the generalization error of the overall classifier (Vapnik 1999). The principle of SVM is adopted from the concept of Kernel approach, where data transformation into higherdimensional space is done using a non-linear transformation. The strength of SVM lies in its ability to overcome high dimensionality and able to perform well with small number of trainingsamples. Several publications have shown that SVM provides high accuracy for land cover classification and alien species distribution mapping (Kganyago *et al.,* 2018; Mafanya *et al.*, 2022). The equation of SVM is as follows:

$$L = \frac{1}{2} \| w \|^2 - \sum_{i=1}^{n} \leftthreetimes i \ (y_i((w.x_i) + \mathrm{b}) - 1)$$

L: Denotes the loss function, it considers the effectiveness of SVM training procedure.

w: SVM linear vector, the hyperplane of SVM is determined by this vector.

||w||: The weight vector is represented by $'w''s$ Euclidean norm. It serves as a regularization of SVM, which gives $'w'$ lower values to avoid overfitting.

n: The sum of data points of training dataset.

3.4.2 Random Forests (RF)

Breiman (2001) defined random forests (RF) as “a tree-based algorithm that depends on the value of an independent random vector sampled for all trees in the forest”. This simply means that RF uses many tree classifications, and a new input vector is classified with several of each tree in the forest. Then each tree is given a classification, which means the tree “votes” for the class that has the most frequent input data, and during the classification process the forest will favour the classification that has the most class “votes” over the forest tree (Adelabu *et al.*, 2014). The tree regression-based model when trained with sufficient field data plots representative of the vegetation variability at national scale produces good results.

Random forests use bootstrap sampling with other trees combined with ensemble regression and tree classification to construct binary classification (Kulkarni and Lowe, 2016). Random forest just like any other algorithms have advantages and disadvantages, to mention few advantages. Random forest is efficient in implementation of large datasets and easily saved structure for future use of pre-generated trees (Kulkarni and Lowe, 2016). The algorithm is not sensitive to noise, it avoids overfitting and high accuracy (Kulkarni and Lowe, 2016). It incorporates spectral bands and feature selection layers such as (soil index, water index, NDVI), and texture features for classification such as (entropy, variance, morphology, line feature) are incorporated (Chaturvedi and de Vries, 2021). The RF algorithm shows a great potential to solve environmental problems such as water resource and natural hazard management (Talukdar *et al.*, 2020). Random forest is integrated with Decision tree algorithm and uses classification regression and regression trees (Breiman, 2001).

3.5 Accuracy assessment

An accuracy assessment was conducted using SVM and RF models to validate the classified Sentinel-2A imagery. The traditional confusion matrix shows the degree of agreement between the classified image and reference ground data for overall accuracy, user accuracy, producer accuracy, and kappa coefficient. Overall accuracy (OA) measures the proportion of accurately classified LULC classes (Petropoulos *et al*. 2012). The overall kappa coefficient measures the agreement between training and validation datasets (Foody, 2009). However, many studies have not used the kappa coefficient due to conceptual flaws. However, this study added a kappa coefficient to the assessment of accuracy. Additionally, precision can be evaluated using various machine learning measures, including the F-score (Gidey and Mhangara 2023). The F-score was used to gain a thorough understanding on the performance of both the SVM and RF models. This is achieved by combining precision and recall into a single metric. Equations 1–5 demonstrate the consideration of the overall accuracy (OA), user accuracy (UA), and producer accuracy (PA).

$$UA = {}^{x_{ii}}/_{x_{i+}} \times 100 \qquad \text{(Eq. 1)}$$

$$PA = {}^{x_{ii}}/_{x_{+i}} \times 100 \qquad \text{(Eq. 2)}$$

$$OA = {}^{D}/_{V} \times 100 \qquad \text{(Eq. 3)}$$

$$\text{F} - \text{score} = 2 \times (PA \times UA)/(PA + UA) \qquad \text{(Eq. 4)}$$

$$\hat{K} = \frac{N\sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k}(x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{k}(x_{i+} \times x_{+i})} \qquad \text{(Eq. 5)}$$

Where $\hat{K}$ = **K**-coefficient, $\boldsymbol{x_{ii}}$ refers to the total number of observations in both the row$_i$ and the column$_i$, $\boldsymbol{x_{i+}}$ and $\boldsymbol{x_{+i}}$ refers to the respective marginal totals, **N** refers to the total number of observations, and **D** refers to the total number of correct pixels in the diagonal, which is the same as the total number of pixels in the **V** error matrix. The **F-score** is calculated as the harmonic mean of the producer's accuracy (**PA**) and user's accuracy (**UA**) (Gidey and Mhangara, 2023).

The Pearson correlation coefficients and regression models were adopted to assess the effectiveness of the SVM and RF models using the STATA software version 14.

3.6 MaxEnt ecological niche modelling for pompom weed distribution

Species distribution models (SDMs) are used in ecological studies to model the geographical distribution of alien plants (Mtengwana *et al.*, 2021; Dai *et al.*, 2022). The SDMs models are effective in modelling the spatial distribution alien plants forecasting their current and future land distributions of IAPs (Ndlovu and Shoko, 2023). If the environmental conditions of the invaded ecosystem by IAPs is the same as of their native, then the conditions necessitate the survival of IAPs and influencing further invasion into other areas.

The open access Maximum Entropy Species Distribution Modeling (MaxEnt), version 3.4.4 is used in this study, due to computer efficiency and ability to compute environmental variables such as rainfall, temperature, and topography. The software utilizes a user-defined landscape divided into grid cells and incorporating the presence-only data, which is the sample of locations where the species are captured (Mtengwana *et al.*, 2022; Mkungo *et al.*, 2023). It produces the habitant suitability of alien species ranging from high to low distributions (Ndlovu & Shoko 2023).

For the purposes of this study, MaxEnt species distribution model was employed to explain the spatial distribution of pompom weed, along with the selected environmental variables such as (i) Elevation; (ii) Land cover; (iii) Rainfall; (iv) Temperature and (v) Soil. The environmental variables (Rainfall, Temperature, Soil and Elevation) were downloaded from NASA Power Data https://power.larc.nasa.gov/data-access-viewer/ .It is important to note that the climatic data is historical from 2019 to 2023 and were monthly aggregated. The Land cover variable was the 2024 SVM classified imagery because it showed highest and accurate classification. ArcGIS (ArcMap) version 10.8.2 was used to process the environmental variables data. The imported environmental variables on ArcMap were converted from raster to ASCII, the American Standard Code for Information Interchange which is compatible for MaxEnt. Additionally, the 469-pompom weed ground control points (GCPs) were converted to comma separated values (csv) to be added on MaxEnt sample file. The MaxEnt model divided the pompom weed GCPs (samples) into 70% training and 30% validation. The environmental layer on MaxEnt was used to import the environmental variables set on continuous. The model was set at 1 with 500 iterations to run ten-fold cross-validation and the output was set “Clog-log” format since it is highly cited and recommended (Phillips and Dudik, 2008; Ndlovu *et al*., 2018; Ndlovu and Shoko 2023).

3.6.1 MaxEnt Species distribution model evaluation

The area under curve (AUC) is used to assess the performance of MaxEnt SDM. The AUC measures the ability of the classifier to correctly predict the species presence only data (sensitivity) against the absence (specificity) by comparing the actual and predicted species distribution (Mkungo *et al*., 2023). The AUC rates the model with values as between 0,5 to 0,6 as poor, 0.7 to 0,8 as decent and 0,9 to 1 as excellent model performance prediction. Therefore, the AUC was used to evaluate the performance of MaxEnt predictive model. The MaxEnt uses a jack-knife test to evaluate the efficacy of predictor variable in predicting landscape vulnerability to invasion by IAPs, and as well producing distinct information on the species distribution.

CHAPTER FOUR

RESULTS

4.1 Spatial distribution of *Campuloclinium macrocephalum* (Less) DC in the Greater Cradle nature reserve for 2019 to 2024

Figures 2 to Figure 4 shows the spatio-temporal trends of pompom weed distributions from 2019 to 2014. Figure 2 presents the changes in areal coverage by pompom weed at the Cradle nature reserve between 2019 and 2024. The Cradle nature reserve total areal coverage is 92.38 km², in January 2019 the pompom weed areal coverage was 31.1 km² and 29.7.42 km² using SVM and RF, respectively. The SVM reported 46.84km² and RF 12.0km² areal coverage for January2024. The comparison between SVM and RF models for the year 2019 and 2024 yielded resultsthat varies significantly, however the temporal dynamics of pompom weed distributions acrossthe nature reserve using SVM model is in consistent with the visual analysis of pompom weeddistribution on the satellite imagery and on the field. Indicating a rapid increase of pompom weed in the nature reserve between 2019 and 2024, thus making the SVM model a powerful supervised machine learning technique.

Figure 3 and Figure 4 displays the performance of SVM and RF models in mapping the spatial distribution ofpompom weed against the co-existing land cover types at the Cradle nature reserve between January 2019 and January 2024. The invasive pompom weed is detected to be present across the nature reserve for both years. The SVM model indicates the increment of pompom weed spatial distribution is 50.7% from 44% in 2024 and 2019 respectively. In contrast, the RF modelindicates the spatial distribution of pompom weed at the nature reserve was 31.14% in 2019 and further increased to 39.3% in 2024. It can be observed that the presence of pompom weed is patchy and heterogeneous as it invades different land cover types or vegetation species, however pompom weed highly invaded the bushland and riparian zones. The SVM and RF models were found to be effective in accurately discriminating and showingthe spatial extent of pompom weed at the Cradle nature reserve. The current observations indicate that 50% of the nature reserve is highly invaded by uncontrolled exotic pompom weed, masking the nature reserve in pink-purple colour during the spring and summer seasons. The map outputs from both years shows a very little indication decrease of pompom weed,rather an increase of pompom weed infestations in areas that were not previously invaded, suchas the north-western side of the nature reserve. This encroachment of pompom weed is causingenvironmental degradations at the nature reserve, posing a high-risk migration of game animals, reducing biodiversity, and lowering the world accorded status of the nature reserve.

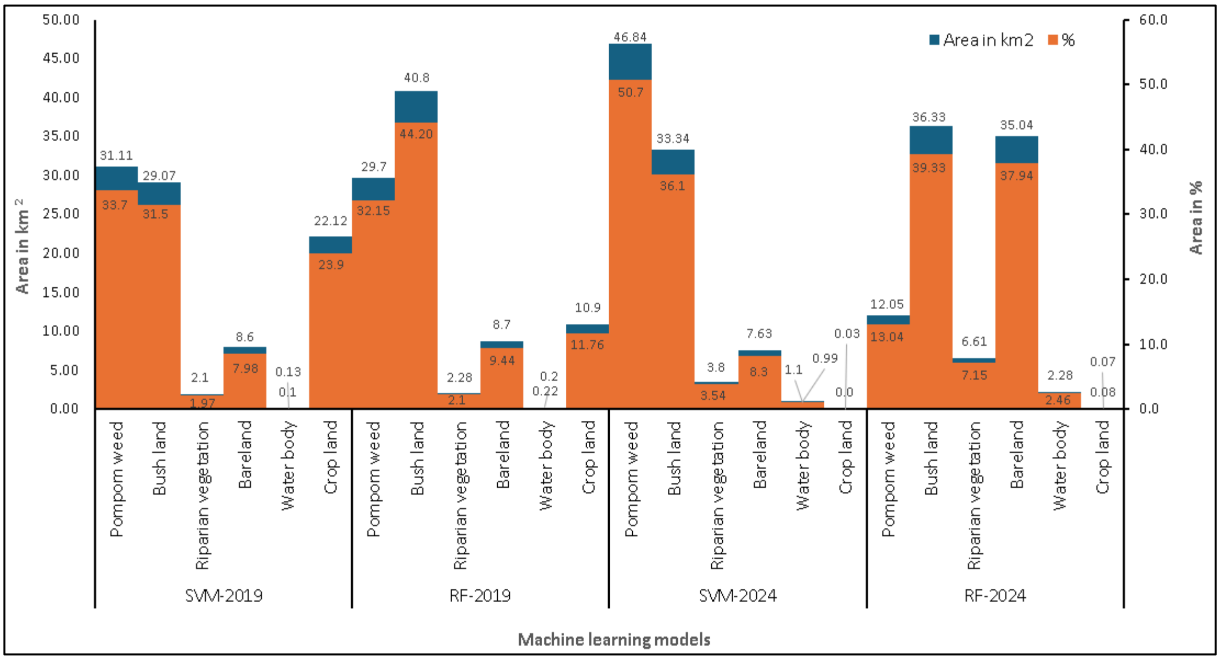

Figure 2: The areal coverage of pompom weed against other land cover types in the Greater Nature Reserve.

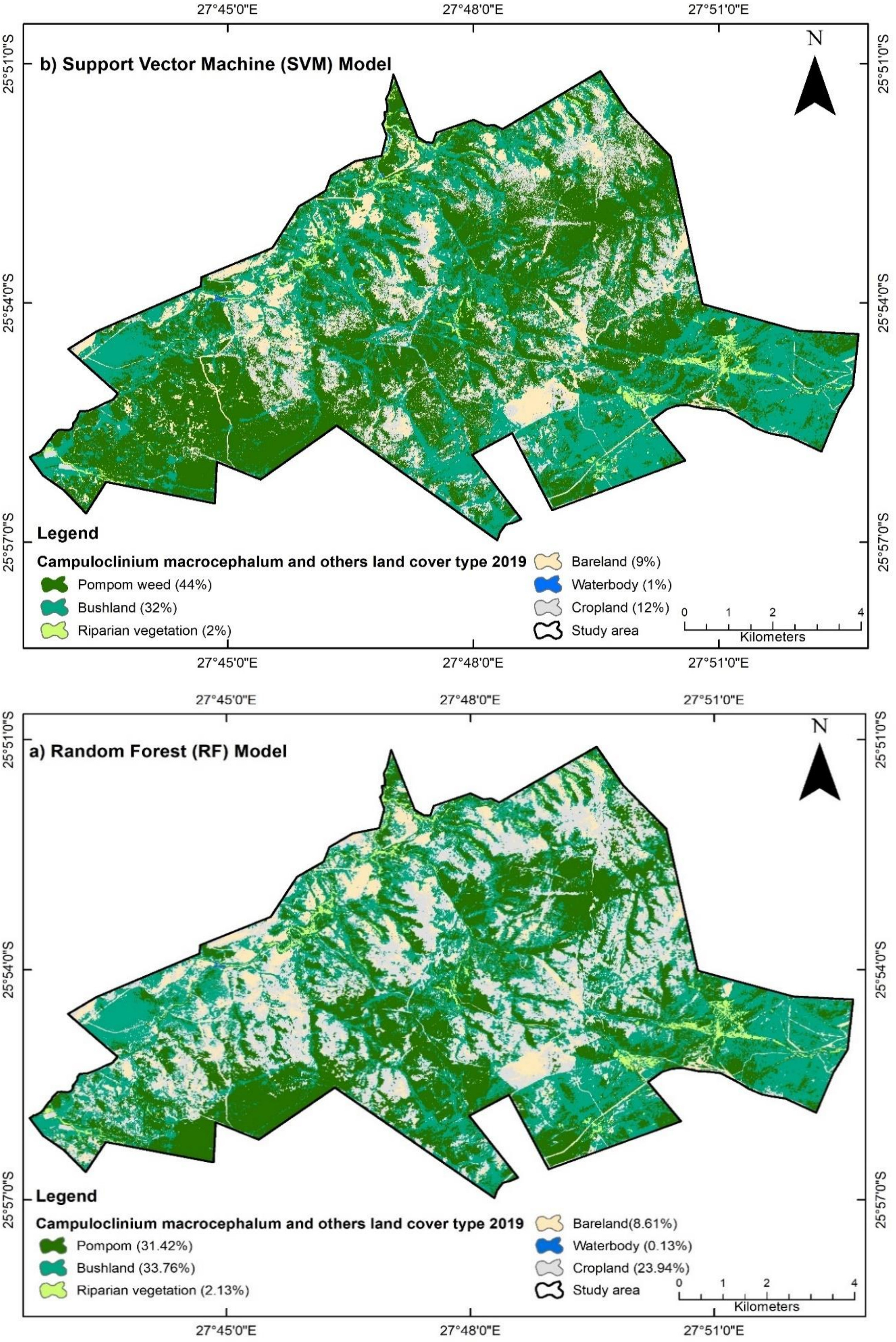

Figure 3. SVM and RF model based spatial distribution of pompom weed at Cradle nature reserve for 2019.

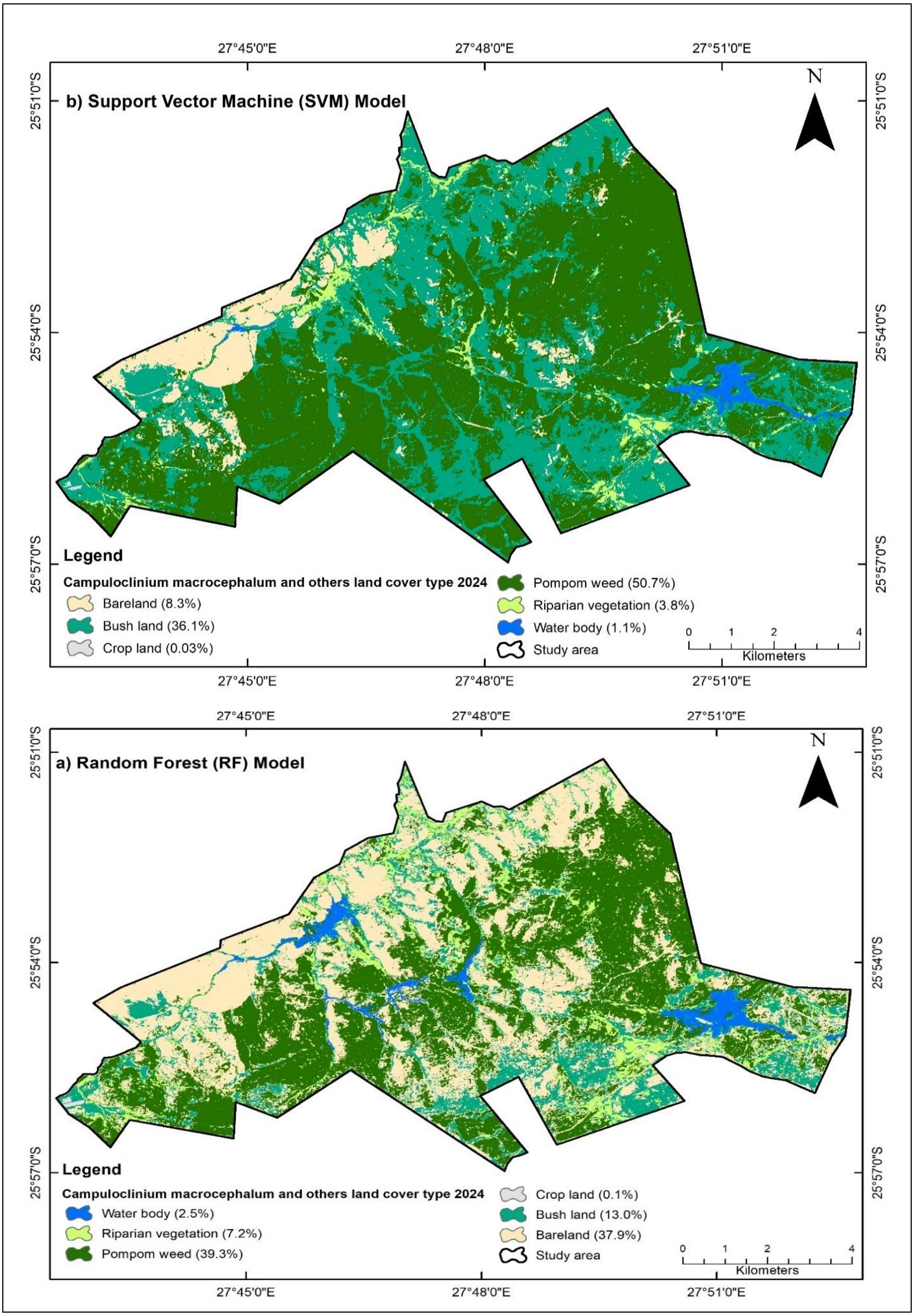

Figure 4. SVM and RF model based spatial distribution of pompom weed at Cradle nature reserve for 2024.

The SVM and RF models were applied to show the dynamics in spatial distribution of pompom weed (Figure 5). The SVM model estimated that from 2019 to 2024 there has been an increase of pompom weed presence at the nature reserve, the land cover that is dominant is pompom weed. In contrast, the RF model estimated that there has been a decline in pompom weed. This explains the discrepancies of working with the two models and as well highlighting the weaknesses of each model. However, the SVM model results are in consistent with the current visual assessments on the field indicating a rapid encroachment of pompom weed at the study area. In relation to other land cover types, the RF model indicates a high decline of Bushland over the years and the SVM is as well indicating a gradual decline in Bushland. The decline in Bushland can be explained as a results of land disturbances and invasion by pompom weed. The cropland is decreasing in both models for the same periods, with SVM indicating a high decline in cropland. The Bareland is a dominant land cover with RF model, the opposite is observed with SVM model.

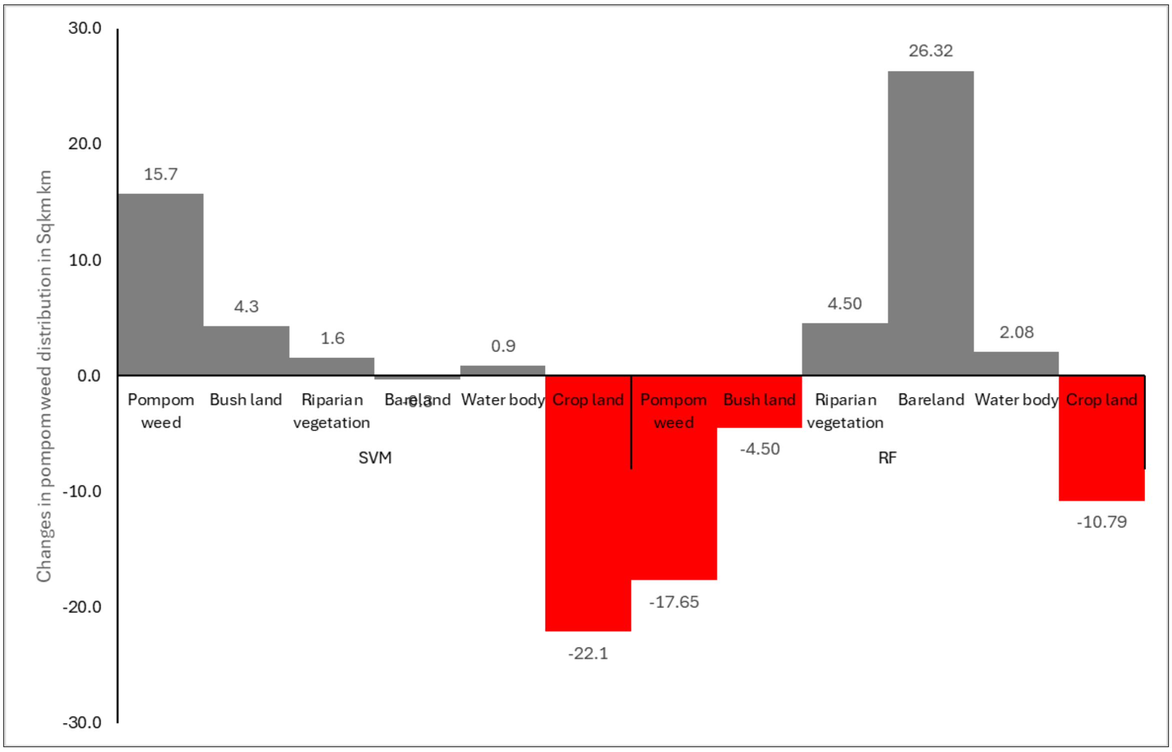

Figure 5: Showing temporal changes in pompom weed distribution from 2019 to 2024.

4.2 Analysis of pompom weed distributions along various land cover types

The Cradle nature reserve land mass is 92.29 km$^{2.}$ figure 6 indicates the changes in land cover types that were invaded by pompom weed between 2019 and 2024. It is noted that pompom weed invaded the bushland and riparian zones from the period 2019 to 2024 by 0.164 km^2 (58%) and 0.086 km^2 (30%), respectively. In such a case, this explains that Bushland are susceptible to pompom weed invasion, due tovarious factors such as bush thickening, land disturbances in terms of road construction during the last five years. The infectious bush thickening is caused by high encroachment of invasive plants, and it isconsidered a serious threat to savannah rangelands. The riparian vegetation grows along theriver streams, in its nature pompom weed also invade wetlands (Goodall *et al*., 2011, SANBI,2023), therefore riparian vegetation is also susceptible to pompom weed invasion. However, the Cropland and Bareland showed a reduction in their spatial areal coverage by -0.0001 km^2 (0%) and -0.010 km^2 (-3%), respectively. This indicated that the distribution of pompom weed in these two land cover types is negative.

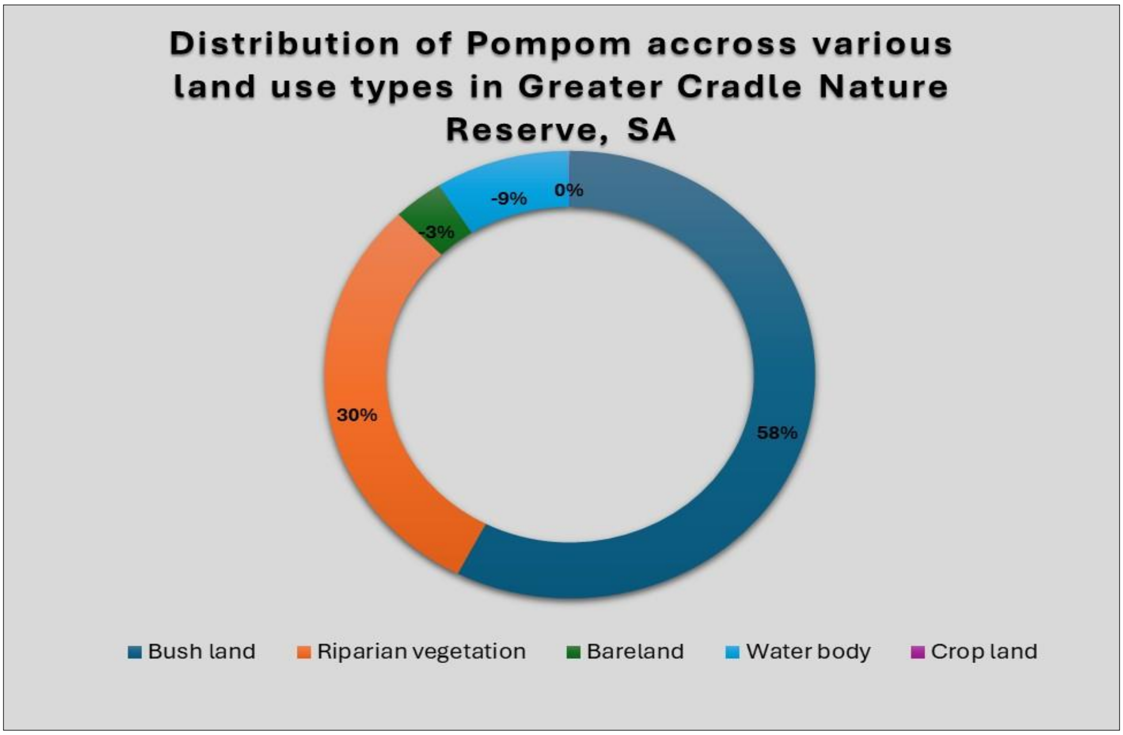

Figure 6: Distribution of pompom weed across land cover types in the nature reserve.

4.3 Accuracy assessment and confusion matrix

The accuracy assessments were performed for both years 2019 and 2024 Sentinel-2 classified imagery using the popular confusion matrix, the results are presented in Tables 2 to 5. The overall accuracy was manually calculated by taking the sum of correctly classified values anddividing it with the total number of values. The user's and producer's accuracy are included, as well as kappa coefficient despite its mentioned critics in redundancy and misleading in remote sensing applications (Pontius and Millones, 2011). The Kappa coefficient is ranked asfollows any value from (0 to 0.4) is regarded as moderate, (0.4 to 0.8) is regarded as substantialagreement and whereas above 0.8 is regarded as excellent agreement. In addition, the F-score values for the same period and model indicate that they exceed the minimum limit thresholds, i.e., 0.5 or 50%. All the findings in this case exceed the minimum standards. The F-score values for the pompom and waterbody range from 0.92 to 0.98, respectively.

Table 2: Confusion matrix for SVM 2024

Land use type	Pompom weed	Bush land	Riparian vegetation	Bare land	Water body	Crop land	Total	U-Accuracy	F-score	Kappa
Pompom weed	**30**	0	0	0	0	0	30	0.97	0.92	-
Bushland	1	**29**	0	0	0	0	30	0.96	0.96	-
Riparian	1	0	**29**	0	0	0	30	0.97	0.94	-
Bareland	0	1	1	**28**	0	0	30	0.93	0.96	-
Waterbody	0	0	1	0	**29**	0	30	0.97	0.98	-
Cropland	2	0	1	0	0	**27**	30	0.90	0.95	-
Total	34	30	32	28	29	27	180	-	-	-
P-Accuracy	0.88	0.97	0.91	1	1	1	0	0.96	-	-
Kappa **OA (95%)**	-	-	-	-	-	-	-	-	-	0.94

The classification accuracy for 2024 classified imagery using SVM model obtained kappa coefficient of 0.94 which is regarded as excellent agreement of classification Table 2. In relation to other land cover types, the classification of pompom weed was successfully detected with user's and producer's accuracy of 0.96 and 0.88, respectively. The Bushland and Riparian vegetations obtained high accuracies that are above 70%, the overall classification accuracy of detecting pompom weed using SVM model is 95%.

Table 3: Confusion matrix for RF 2024

Land use type	Pompom weed	Bush land	Riparian vegetation	Bare land	Water body	Crop land	Total	U-Accuracy	F-score	Kappa
Pompom weed	**28**	1	0	0	0	0	30	0.96	0.85	-
Bushland	1	**29**	0	0	0	0	30	0.95	0.94	-
Riparian vegetation	0	1	**27**	0	2	0	30	1	1.00	-
Bareland	8	2	0	**20**	0	0	30	0.65	0.79	-
Waterbody	0	0	5	0	**25**	0	30	1.00	1.00	-
Cropland	0	0	0	0	0	**30**	30	1.00	1.00	-
Total	39	31	30	20	30	30	180	-	-	-
P-Accuracy	0.76	0.93	1	1	1	1	0	0.94	-	-
Kappa **OA (93%)**	-	-	-	-	-	-	-	-	-	0.92

The overall classification accuracy of using RF model in detecting pompom weed in 2024 classified imagery is 93%. The Kappa coefficient is 0.92, user accuracy is 0.94 and producer accuracy 0.76 as shown in Table 3. In relation to other land cover types, the waterbody was accurately classified with no spectral confusions in 2024 imagery using RF. The classification accuracy for the year 2024 proved successful in obtaining high accuracies, additionally the SVM outperformed the RF.

Table 4: Confusion matrix for SVM 2019

Land use type	Pompom weed	Bush land	Riparian vegetation	Bare land	Water body	Crop land	Total	U-Accuracy	F-Score	Kappa
Pompom weed	**26**	2	1	0	0	1	30	0.94	0.83	-
Bushland	2	**28**	0	0	0	0	30	1.00	0.82	-
Riparian	0	1	**24**	0	5	0	30	1.00	1.00	-
Bareland	0	2	1	**26**	0	1	30	1.00	1.00	-
Waterbody	0	2	5	0	**23**	0	30	1.00	1.00	-
Cropland	0	7	0	5	0	**18**	30	0.58	0.73	-
Total	28	42	31	31	28	20	180	-	-	-
P-Accuracy	0.75	0.7	1	1	1	1	0	0.90	-	-
Kappa **OA (80%)**	-	-	-	-	-	-	-	-	-	0.91

The classification accuracy for 2019 using the SVM model obtained 94% overall accuracy, with kappa coefficient of 0,91 Table 4. The user's accuracy and producer's accuracy obtained0,90 and 0.75 respectively. The classification accuracy for 2019 using RF model obtained overall accuracy of 80% and 0.90 kappa coefficient, 0.92 and 0.75 for user accuracy and producer accuracy, respectively Table 5. In summary, the SVM model performed better than RF model in accurately detecting pompom weed against the co-existing land cover types in 2019 and 2024.

Table 5: Confusion matrix for RF 2019

Land use type	Pompom weed	Bush land	Riparian vegetation	Bare land	Water body	Crop land	Total	U-accuracy	F-score	Kappa
Pompom weed	26	2	1	0	0	1	30	0.94	0.83	-
Bushland	**2**	**28**	0	0	0	0	30	1.00	0.81	-
Riparian	0	1	**24**	0	5	0	30	1.00	1.00	-
Bareland	0	2	1	**26**	0	1	30	1.00	1.00	-
Waterbody	0	2	5	0	**23**	0	30	1.00	1.00	-
Cropland	0	7	0	5	0	**18**	30	0.58	0.73	-
Total	28	42	31	31	28	20	180	-	-	-
P-Accuracy	0.75	0.68	1	1	1	1	0	0.92	-	-
Kappa	-	-	-	-	-	-			-	0.90
OA (80%)										

4.4 SVM and RF models performance evaluation

This study evaluated the performance of SVM and RF models using 120 659 samples containing pompom weed. In this study, the 120 659 samples containing pompom weed extracted at 2 km radius from the Centre of nature reserve, considering the computation processing capability of the computer. The samples were used to test the identification capabilities of each model. Theresults indicated that pompom weed was identified, and the models correlates to each other by 0.68 using Pearson correlation coefficient, which is statistically significant P = 0.00. Additionally, the models were tested using regression model that indicates that both RF and SVM are statistically significant with R-squared of = 0.45 and P-value = 0.00 which indicates a robust positive linear correlation between RF and SVM in the regression model, see (Table 6).

Table 6: Model validation results outputs using STATA v.14 software

Source	***SS***	***df***	***MS***	***Number of Obs*** = 120, 659		
Model	17362.9622	1	17362.9622	F (1, 120656) = 99495.39		
Residual	21055.8799	120,657	.174510222	Prob > F = 0.0000		
				R-squared = 0.4519		
				Adj R-squared = 0.4519		
Total	38418.8421	120,658	.318411063	Root MSE = .41774		
rf	Coef.	Std. Err.	t	P> |t|	[95% Conf.	Interval]
svm	.5232784	.0016589	315.43	0.000	.5200269	.5265299
_cons	.0508302	.0012891	39.43	0.000	.0483036	.0533568

4.5 MaxEnt model-based analysis of pompom along various environmental variables

The results of the jackknife test of variable significance are shown on Figure 7 to Figure 8. Soil is the environmental variable with the highest training gain when ran in isolation Figure 7. Which means it provides useful information in explaining the study area's vulnerability to pompom weed invasion. The environmental variables with highest gain after soil, is rainfall and temperature. However, temperature when it is omitted it reduces the gain which means it has the most information that is not present in other environmental variables. The environmental variables with the lowest training gain are elevation and land cover.

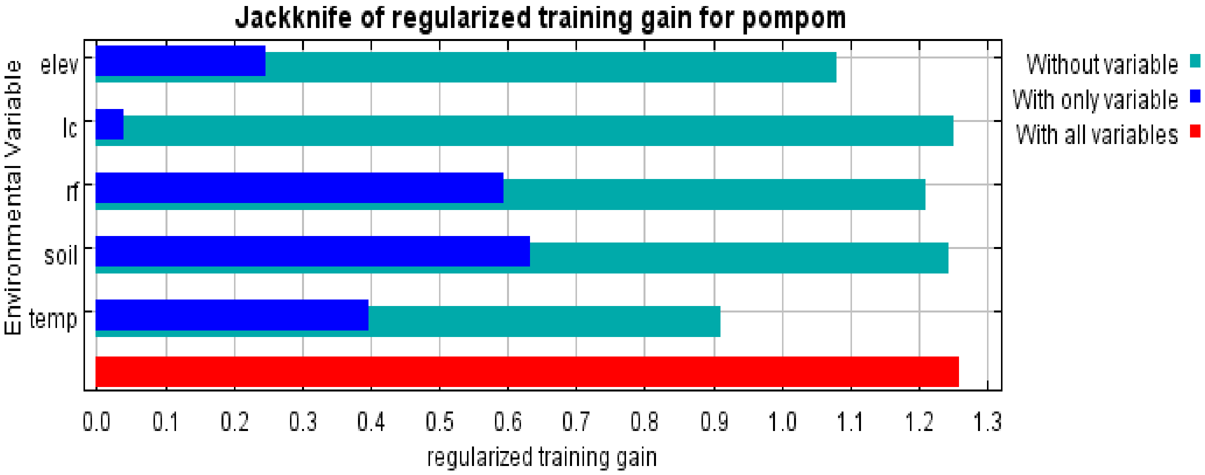

Figure 7: Jacknife results training gain for pompom weed.
**Note: Elevation (ele); Land cover (lc); Rainfall (rf); soil and Temperature (temp)*

The jackknife of test gain results Figure 8 shows that soil and rainfall are still the environmental variables with the highest test gain, which holds the most important information in determining the area's susceptibility to pompom weed invasion. As opposite to Jackknife regularized training gain results, the Jackknife test gain selects elevation as the third contributing environmental variable. Land cover remains the least contributing environmental variable.

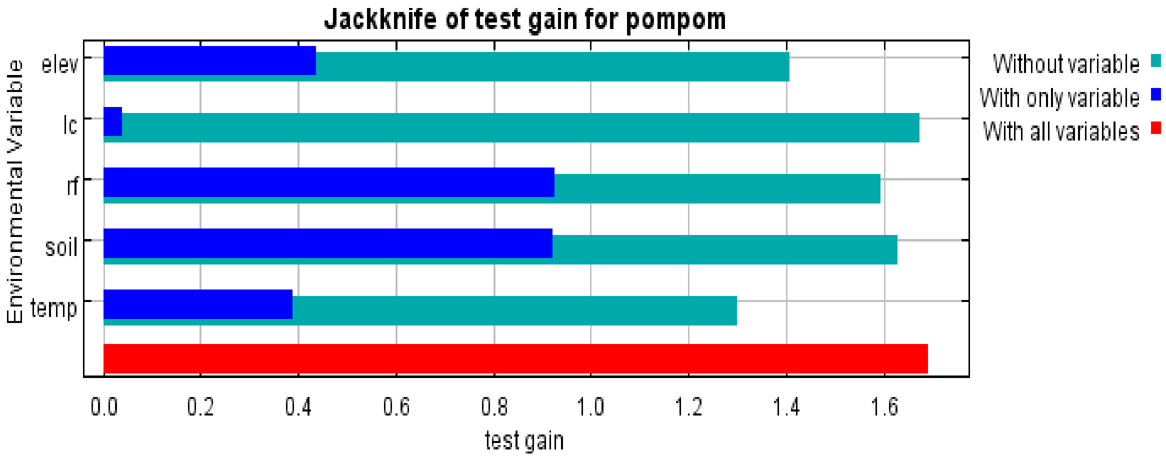

Figure 8: Jackknife results test gain for pompom weed.

Lastly, Figure 9 shows the Jackknife test results using AUC for pompom. Soil, rainfall, and temperature are selected as the environmental variables with most gains. Generally, the result of our analysis indicates that soil, rainfall, and temperature are the most influential environmental variables affecting the spatial distribution of pompom weed in the nature reserve. In particular, Soil holds the most important information in explaining pompom weed invasion. It can be explained that the soil nutrient composition of the nature reserve is favorable for pompom weed to thrive successfully. The elevation and land cover have less contribution to pompom weed infestation.

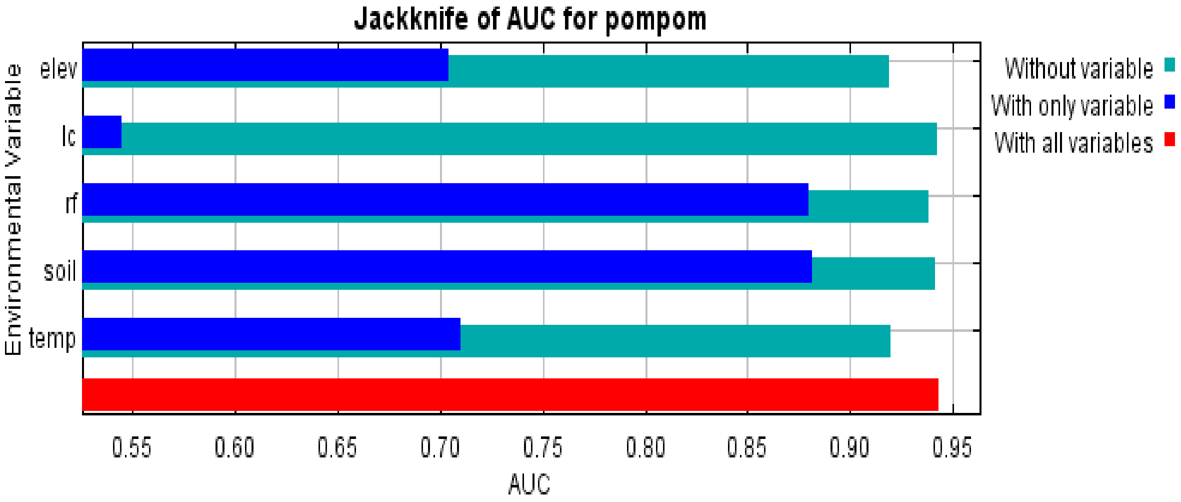

Figure 9: Jackknife test using AUC on test data.

4.6 Analysis of each environmental variable contributions to pompom

Table 7 indicates the estimates of relative contributions of the selected environmental variablestrained on MaxEnt model, to give estimations of how each of the environmental variables contributed to pompom habitat suitability. Soil is the leading environmental variable with 44.8%, this explains that the soil and nutrients composition of the study area is favourable for pompom weed to grow. Temperature and rainfall also show high percent contributions, 29.8 and 14.3 respectively. The elevation and land cover have less contributions, therefore they areless significant.

Table 7: Estimates relative contributions of environmental variables.

Environmental variables	Percent contribution	Permutation importance
Soil	44.8	39.3
Temp	29.8	30.3
Rf	14.3	21.2
Elev	10.7	8.7
Lc	0.4	0.5

4.7 MaxEnt species distribution model evaluation

The performance of MaxEnt model was evaluated by calculating the area under cover (AUC). In respect to AUC model rates 0.5 to 0.6 is regarded as (Poor) 0.7 to 0.8 (Decent) and 0.9 to 1 (Excellent). Figure 10 shows the area under receiver operating characteristic (AUC) results. The selected environmental variables and presence-only data (pompom weed samples) were used in predicting pompom weed distribution and habitat suitability. The model prediction accuracy achieved is AUC score of 0.94 with training and test data, in predicting the spatial distribution of pompom weed in the study area for the year 2024. The 0.94 achieved in this study is regarded as excellent model performance prediction.

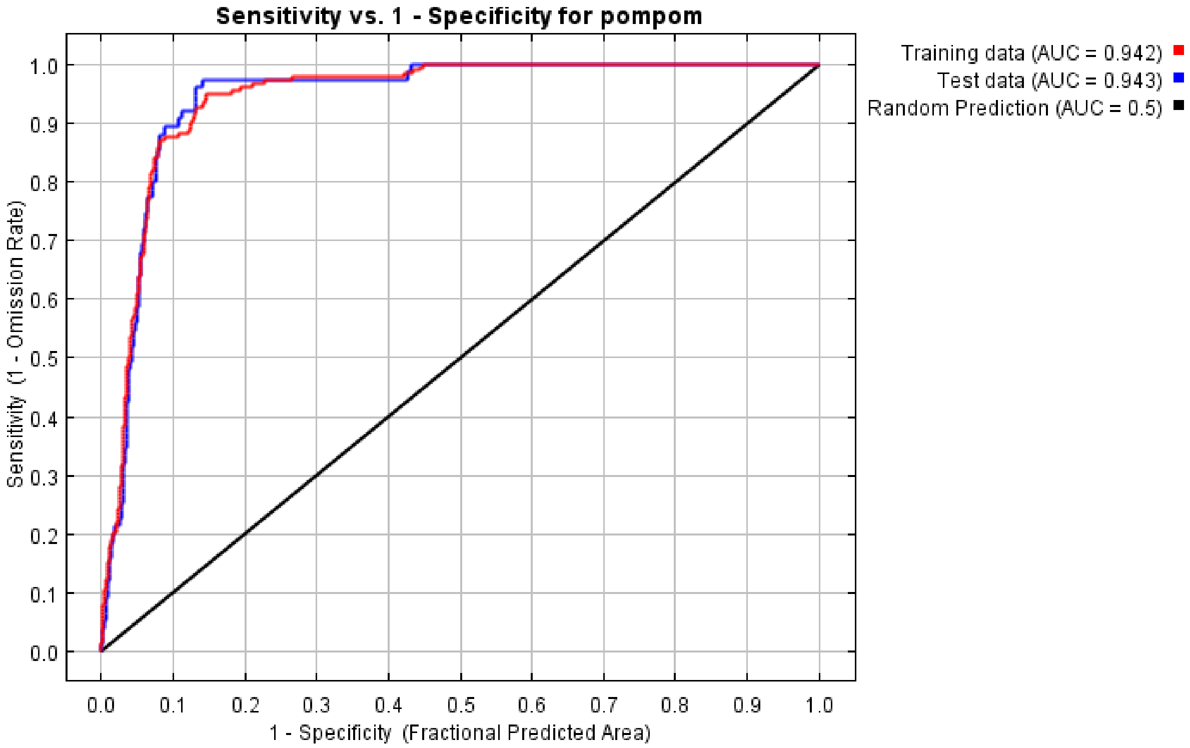

Figure 10: The AUC curve results in predicting pompom weed habitat suitability.

4.8 Spatial distribution of pompom weed based on MaxEnt predictive model

Figure 11 displays the projections of the MaxEnt model prediction of pompom weed distributions based on the selected environmental variables. The central part of the nature reserve is dominated by the presence of pompom weed, the lower southern part of the nature reserve is highly characterized by pompom weed presence, and this is where the locations of pompom weed were collected during the field survey. This explains that MaxEnt model predicted the central, southern, and eastern tip of the nature reserve as high habitat suitability for pompom weed. The Western part depicts low habitat suitability. However, the parts that show low suitability and they are adjacent to central part of the nature reserve are at risk of possible future invasion. Based on the results from MaxEnt predictive model, the parts of the nature reserve that are predicted as high of pompom weed should be prioritized for effective weed control and environmental management.

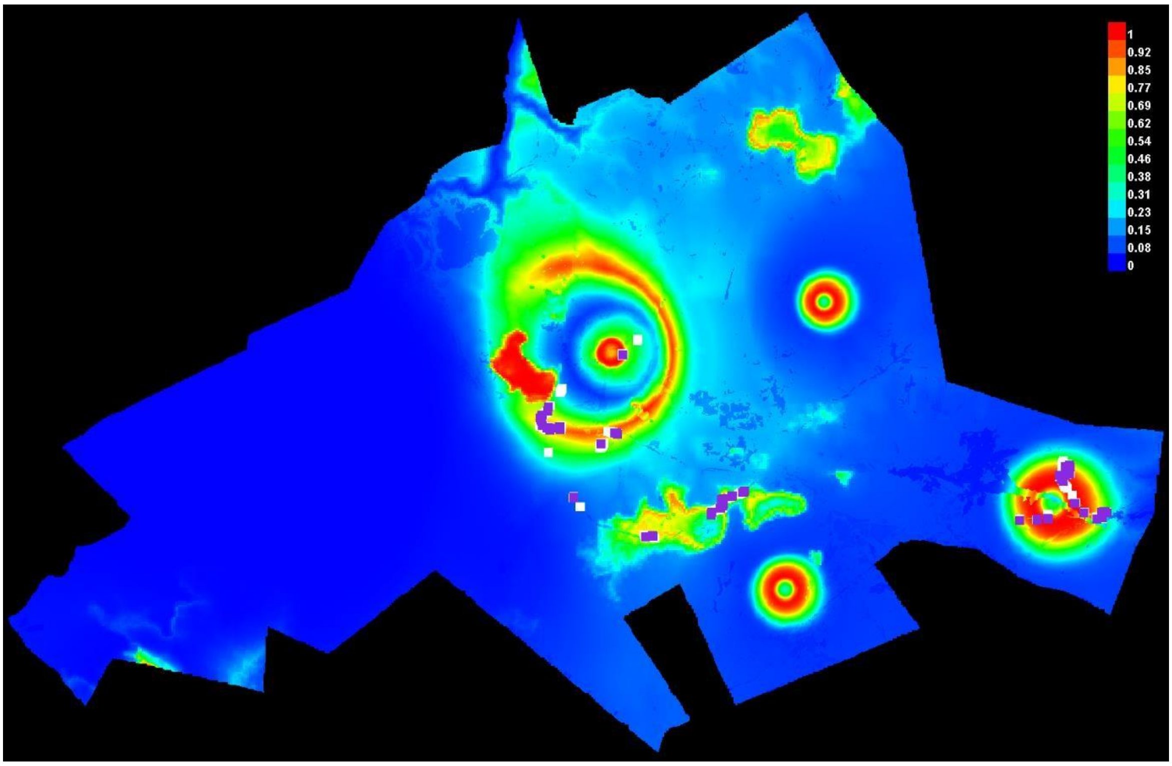

Figure 11: Spatial distribution of pompom weed using MaxEnt predictive model.

4.9 Responsive curves

The responsive curves graphs indicate how individual environmental variable affects the MaxEnt prediction Figure 12. The curves show how the estimated probability of adjusted environmental variables changes, while all other environmental variables are kept constant at their average sample value.

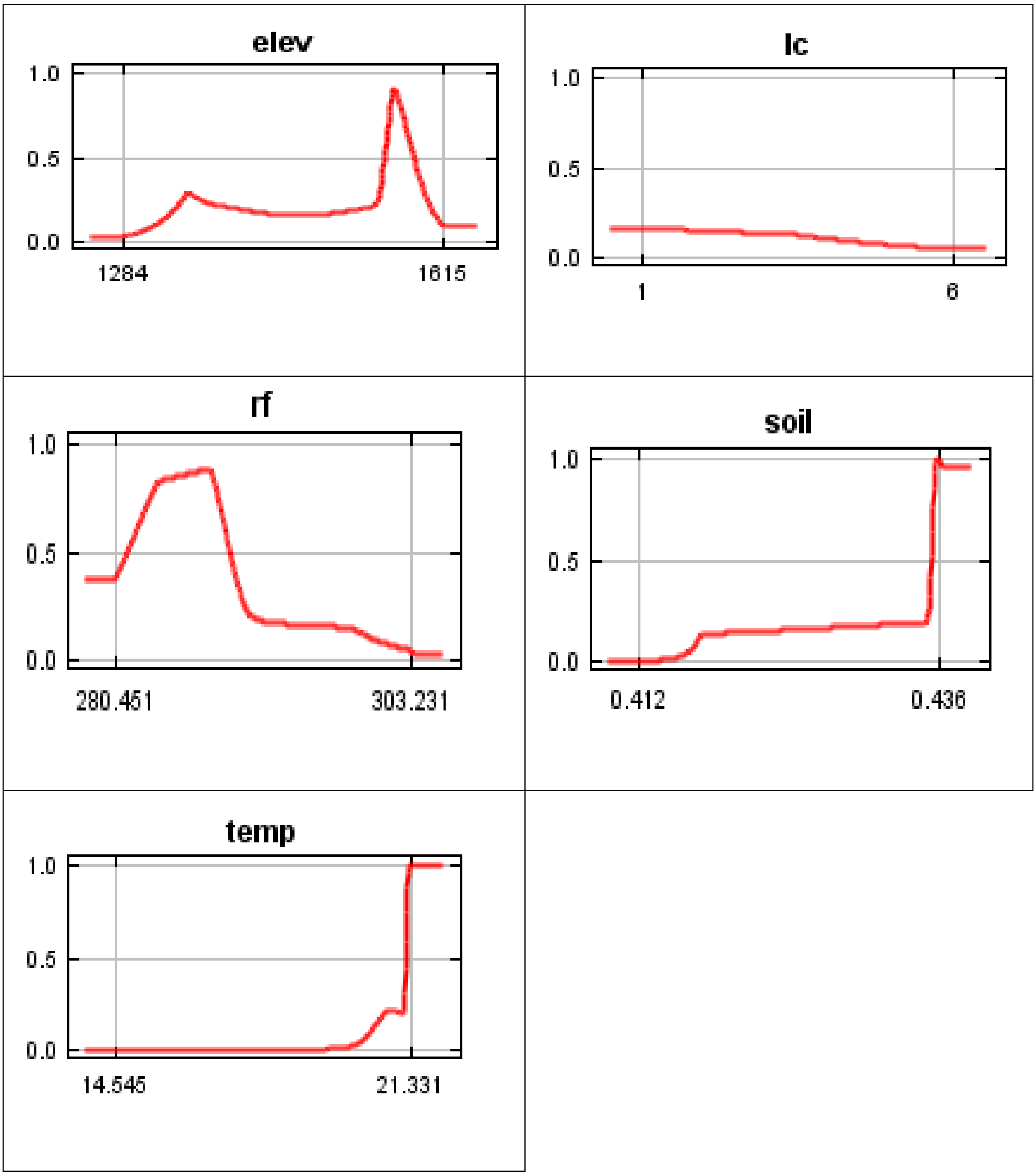

Figure 12: Responsive curve of various environmental variables.

In contrast to figure 13 response curve graphs, each of the following curves indicates a MaxEnt model created using only the using individual environmental variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables.

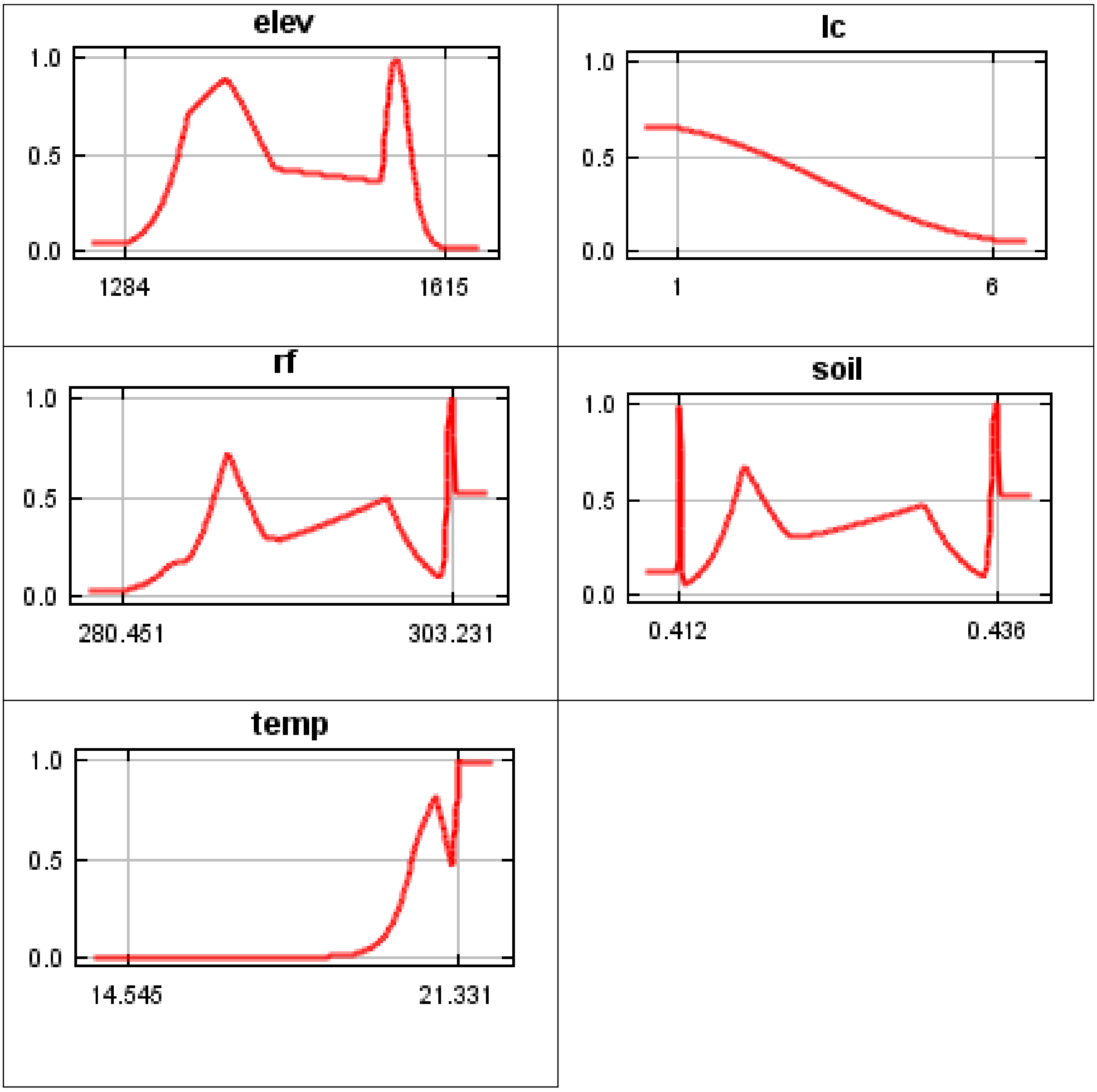

Figure 13: MaxEnt model responsive curve of each environmental variable

CHAPTER FIVE

DISCUSSIONS

5.1 Application of SVM and RF models in mapping spatial distribution of pompom weed using Sentinel-2A imagery

The invasion alien plants are considered the second largest threat to biodiversity (Newete *et al*., 2023). They compete with native vegetation for space, nutrients, and water retention (Mafanya *et al*., 2022). There is still a lack of precise data on the spatial distribution of IAPs, that data is needed for implementing control measures and minimizing their impact on the environment (Kganyago *et al*., 2018). Remote sensing technology is adopted to accurately monitor and map the spatial distribution of invasive weed. The temporal resolution of remote sensing technology gives the user an advantage of obtaining historical information. In terms of IAPs monitoring, a historical data can be obtained to determine the past distribution of IAPs and forecast future distributions and develop effective eradication methods.

In the current study, the Sentinel-2 MSI product proved to be successful in providing the representation of invasive pompom weed spatial distribution in the Cradle nature reserve from 2019 to 2024. This is possible with Band 4: Red (665 nm); Band 3: Green(560nm) and B2: Blue(490nm) that comes with the product. The Red band is effective in vegetation mapping and the 10m spatial resolution of Sentinel-2 is well detailed. The red edge band of Sentinel-2 was instrumental in pompom weed detection. The adoption of SVM and RF models was significant in discriminating pompom weed against the existing land cover types. There are slight differences between the two models, however the overall accuracy obtained are statistically significant. Even though SVM produced higher overall accuracy than RF in classifying the spatial distribution of pompom weed. Therefore, the results obtained from this study are in consistent with other publications in testing the efficacy of machine learning models in invasive alien plants mapping (Mafanya *et al*., 2022, Kganyago *et al*., 2018, Ndlovuand Shoko, 2023).

The findings from this study indicate that the nature reserve is highly invaded by the invasive pompom weed. The encroachment of pompom weed in the nature reserve has been increasingfor the periods chosen for this study. The 2019 imagery showed that pompom weed was presentin the nature reserve, the presence of pompom weed was 44% using SVM model and 31.1% with RF model. For the current year 2024, the SVM indicates 50.7% encroachment of pompomweed in the nature reserve which is an increment of 44% in 2019. The RF model indicates 39.3% for 2024. It is evident that the invasion of pompom weed has been increasing in the nature reserve, and with a possibility of increasing in the future. Therefore, effective seasonal monitoring and eradication management should be implemented.

In relation to the existing land cover types in the nature reserve, pompom weed highly invaded the Bushland. pompom weed accounted 58% of Bushland, making Bushland susceptible to pompom weed. Thus, this has environmental and economic issues, lowering the grazing capacity for game animals resulting in migration of game animals and compromising tourism. The construction of roads in the nature reserve for game vehicles is a form of land disturbance,and this results in bush thickening. The infectious bush thickening is a serious environmental problem caused by the densification of alien plants (Kellner, 2020). In the case of nature reserve, the encroachment of bush thickening replaces the native vegetation, in turn pompom weed takes advantage of that and grow rapidly with less competition. Hence, bushland was highly invaded by pompom weed due to bush thickening. The riparian zones (30%) were otherland cover type highly invaded by pompom weed. Other land cover types such as bareland, waterbody and cropland were less affected by pompom weed. The reason that Cropland is not invaded by pompom weed could be of herbicides or chemicals used for agriculture.

The employment of ESA Sentinel-2 imagery in this study was successful in discriminating the presence of pompom weed from the existing land cover types. The study yielded overall accuracies that are above the 75% for both years. The results from this study can be supplemented by previous studies where Sentinel-2 was used in mapping spatial distributions of alien plants. A study by Ndlovu and Shoko (2023) employed Sentinel-2 imagery to map the spatial distribution and discriminate *L camara* from other LULC types. They obtained 90.27% overall classification accuracy using the RF model on Sentinel-2 data. Another study by Newete et al (2023) used Sentinel-2 imagery and applied RF and SVM algorithms to estimatethe spatial distribution of invasive *Tamarix* genotypes and classify other land cover types in the riparian zones of Leeu, Swart and Olifants in the Western Cape province. The Sentinel-2 imagery was effective in mapping invasive *Tamaix*, achieved 85% overall classificationaccuracy.

5.2 Modelling potential distributions of pompom weed using environmental variables on MaxEnt species distribution model

In this study, the MaxEnt predictive model was found to be effective in working with small, sample data to provide with robust and accurate estimations. The results of this study are in agreement with other published studies where MaxEnt model was used to predict species distribution (Mtengwana *et al.*, 2021; Dai *et al.*, 2022; Ndlovu and Shoko, 2023 and Mkungo *et al.*, 2023).

The historical environmental variables (temperature, rainfall, soil, and elevation) of the study area were selected, together with the land cover from 2024 classified Sentinel-2 imagery. These five environmental variables were selected to investigate their influence on fostering pompom weed distribution at the Cradle nature reserve. It has been well researched that environmental or climatic variables play a significant role in invasive species distribution (Ncube *et al.*, 2020; Ndlovu *et al.*, 2018 and Ndlovu and Shoko 2023). The study predicted areas that are suitable for pompom weed in the nature reserve using the above-mentioned environmental variable on MaxEnt species distribution model. High predictive accuracy AUC of 0.94 score was achieved in this study using presence only data, and the results of this study are in consistent with other published literature in Southern Africa where MaxEnt species distribution model was employed (Mtengwana *et al.*, 2021. Mkungo *et al.,* 2023. Ndlovu and Shoko, 2023). The model predicted the distribution of pompom weed were concentrated in the central and southern parts, as well the eastern tip of the nature reserve. These parts of the nature reserve are regarded as high suitable habitat for pompom weed.

The Jackknife results of MaxEnt model indicates that soil, rainfall, and temperature influence the spatial distribution of pompom weed. These three environmental variables hold a significant information on the establishment and rapid proliferation of pompom weed in our study area of interest. Soil is singled out as the environmental variable with the highest influence on the establishment of pompom weed. It is likely that the soil chemistry of the naturereserve favours pompom weed to thrive successfully. It is recommended that the soil composition of the nature reserve to be researched. Rainfall as well is contributing to the infestation of pompom weed. It is well documented that rainfall is one of the significant climatic variables in fostering growth and spatial distribution of pompom weed (Mtengwana *et al*. 2021).

The Highveld moist conditions are suitable for pompom weed, and the Cradle nature reserve which is in the Highveld region receives rainfall during the spring and summer seasons. The wettest seasons when the nature reserve receives rainfall provides pompom weed with moisture that is important for germination and growth of pompom weed. The availability of water or moisture in the soil during rainfall seasons supports the germination and establishment of pompom weed. It has been observed that pompom weed grows along roadside (Goodall *et al*., 2011). The infestations of pompom weed along roadside can be explained as the result of land disturbance during road constructions and feeding on the carbon from vehicles. It is recommended that a study should be conducted on the effects of carbon on pompom weed.

5.3 The environmental impacts of pompom weed at the Cradle nature reserve

The Cradle nature reserve is accorded a world status by UNESCO, and it is a protected, sensitive nature reserve. The invasion of pompom weed in the nature reserve has detrimental effects on environment, social and economy of Gauteng province. The biodiversity present on the nature reserve is affected by the proliferation of pompom weed. The encroachment of pompom weed reduces the grazing capacity and lowers the water content. The invasive pompom weed grows unaided and replaces the natural vegetation. Considering the results from the predictive model, it is projected that the invasion of pompom weed in nature reserve will increase even in future. Subject to if preventive measures are put into place. Integrated environmental management approaches should be implemented to curb the rapid spread of pompom weed in the nature reserve. The current methods of eradication and weed control taking place in the nature reserve are mechanical which include uprooting and chemical spraying. The currently used eradication methods should be prioritized to areas that are characterized as high suitability of pompom weed by the predictive model to prevent further spreading. Community engagement and raising awareness about invasive species could enlighten community members about the impacts of invasive species. It is highly possible that people in communities are not educated enough about the declared invasive plants.

CHAPTER SIX

CONCLUSIONS

This study examined the competency of using Sentinel-2A imagery in mapping the spatial distributions of pompom weed at the Cradle nature reserve. The SVM and RF were effective in accurately detecting pompom weed against the existing land cover. The findings showed that the spatial distribution of invasive pompom weed has been increasing between 2019 and 2024. The current observations and estimations from models indicate that the presence of pompom weed is worse as compared to previous years. Under the current environmental conditions, it is possible that pompom weed will increase in the future. This was investigated by using MaxEnt species distribution model to assess which environmental variables are significant in supporting the germination and flowering of pompom weed in the study area. The model indicated that soil is the significant variable in influencing the distribution of pompom weed. The findings from the study revealed the robustness and capability of MaxEnt SDM in predicting habitat suitability. The environmental practitioners can use the findings of this study to implement effective eradication methods, monitoring areas where there is high distribution of pompom weed and areas that are at risk of possible invasion.

RECOMMENDATIONS & LIMITATIONS

The study recommends future research should explore if carbon from vehicles triggers Pompom weed infestations along roadsides. We should investigate the effectiveness of eradication methods at nature reserves, starting with a five-year change detection period. However, the study's limitations include uncertainties in prediction models and the selection of important environmental variables.

REFERENCES

Adagbasa, E. G., & Mukwada, G. (2022). Mapping vegetation species succession in a mountainous grassland ecosystem using Landsat, ASTER MI, and Sentinel-2 data. *Plos one*, *17*(1).

Al-dowski, J., Mansor, S.B., San, H. and Khuzaimah, Z. (2020). Land cover mapping using remote sensing data. *American journal of Geographical Information Sytem*, 9(1), pp. 33-45.

Adelabu, S., Mutanga, O. and Adam, E. (2015). Testing the reliability and stability of the internal accuracy assessment of random forest for classifying tree defoliation levels using different validation methods. *Geocarto International*, *30*(7), pp.810-821.

Basheer, S., Wang, X., Farooque, A.A., Nawaz, R.A., Liu, K., Adekanmbi, T. and Liu, S., (2022). Comparison of land use land cover classifiers using different satellite imagery and machine learning techniques. *Remote Sensing*, *14*(19), pp. 4978.

Bartz, R., & Kowarik, I. (2019). Assessing the environmental impacts of invasive alien plants: a review of assessment approaches. *NeoBiota*, *43*, pp. 69-99.

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

Durand, J.F., Meeuvis, J. and Fourie, M. (2010). The threat of mine effluent to the UNESCO status of the Cradle of Humankind World Heritage Site. *TD: The Journal for Transdisciplinary Research in Southern Africa*, *6*(1), pp.73-92.

Civco, D. L. (1993). Artificial neural networks for land-cover classification and mapping. *International journal of geographical information science*, *7*(2), 173-186.

Chaturvedi, V., & de Vries, W. T. (2021). Machine learning algorithms for urban land use planning: A review. *Urban Science*, *5*(3), 68.

Dai, X., Wu, W., Ji, L., Tian, S., Yang, B., Guan, B., & Wu, D. (2022). MaxEnt model-based prediction of potential distributions of Parnassiawightiana (Celastraceae) in China. *Biodiversity Data Journal*, *10*.

Goodall, J., Witkowski, E. T., Morris, C. D., & Henderson, L. (2011). Are environmental factors important facilitators of pompom weed (Campuloclinium macrocephalum) invasion in South African rangelands?. *Biological invasions*, *13*, 2217-2231.

Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A. and Hasanlou, M. (2020). Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. *ISPRS Journal of Photogrammetry and Remote Sensing*, *167*, pp.276-288.

Gidey, E., & Mhangara, P. (2023). An Application of Machine-Learning Model for Analyzing the Impact of Land-Use Change on Surface Water Resources in Gauteng Province, South Africa. *Remote Sensing*, *15*(16), 4092.

Gidey, E., Dikinya, O., Sebego, R., Segosebe, E., Zenebe, A., Mussa, S., ... & Birhane, E. (2023). Land use and land cover change determinants in Raya Valley, Tigray, Northern Ethiopian Highlands. *Agriculture*, *13*(2), 507.

Huang, C. Y., & Asner, G. P. (2009). Applications of remote sensing to alien invasive plant studies. *Sensors*, *9*(6), 4869-4889.

Hütt, C., Koppe, W., Miao, Y. and Bareth, G. (2016). Best accuracy land use/land cover (LULC) classification to derive crop types using multitemporal, multisensor, and multi-polarization SAR satellite images. *Remote sensing*, *8*(8), p.684.

Jensen, J.R., (1983). Biophysical remote sensing. *Annals of the Association of AmericanGeographers*, *73*(1), pp.111-132.

Jensen, T., Seerup Hass, F., Seam Akbar, M., Holm Petersen, P., & Jokar Arsanjani, J. (2020). Employing Machine Learning for Detection of Invasive Species using Sentinel-2 and AVIRIS Data: The Case of Kudzu in the United States. Sustainability, 12 (9), 3544.

Kganyago, M., Odindi, J., Adjorlolo, C. and Mhangara, P. (2017). Selecting a subset of spectral bands for mapping invasive alien plants: a case of discriminating Parthenium hysterophorus using field spectroscopy data. *International Journal of Remote Sensing*, *38*(20), pp.5608-5625.

Kganyago, M., Odindi, J., Adjorlolo, C. and Mhangara, P. (2018). Evaluating the capability ofLandsat 8 OLI and SPOT 6 for discriminating invasive alien species in the African Savannah landscapes. *International journal of applied earth observations and geoinformation , 67 , pp 10-19.*

Lass, L. W., Thill, D. C., Shafii, B., & Prather, T. S. (2002). Detecting spotted knapweed (Centaurea maculosa) with hyperspectral remote sensing technology. *WeedTechnology*, *16*(2), 426-432.

Mafanya, M., Tsele, P., Zengeya, T. and Ramoelo, A. (2022). An assessment of image classifiers for generating machine-learning training samples for mapping the invasive Campuloclinium macrocephalum (Less.) DC (pompom weed) using DESIS hyperspectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *185*, pp.188-200.

Mkungo, L., Odindi, J., Mutanga, O., & Matongera, T. N. (2023). Modelling landscape vulnerability to the Bracken fern (Pteridium aquilinum) invasion in a remnant urban Sandstone Sourveldt grassland ecosystem. *Scientific African*, *22*, e01947.

McConnachie, A.J., Retief, E., Henderson, L. and Mc Kay, F. (2011). The initiation of a biological control programme against pompom weed, Campuloclinium macrocephalum (Less.) DC.(Asteraceae), in South Africa. *African Entomology*, *19*(1), pp.258-268.

Mtengwana, B., Dube, T., Mudereri, B. T., & Shoko, C. (2021). Modeling the geographic spread and proliferation of invasive alien plants (IAPs) into new ecosystems using multi-source data and multiple predictive models in the Heuningnes catchment, South Africa. *GIScience & Remote Sensing*, *58*(4), pp.483-500.

Mhangara, P. and Odindi, J. (2013). Potential of texture-based classification in urban landscapesusing multispectral aerial photos. *South African Journal of Science*, *109*(3), pp.1-8.

Miranda, E., Mutiara, A.B. and Wibowo, W.C. (2018) September. Classification of land cover from Sentinel-2 imagery using supervised classification technique (preliminary study). In *2018 International Conference on Information Management and Technology (ICIMTech)* pp. 69-74.

Ndlovu, H., & Shoko, C. (2023). Assessing Lantana camara spatial distribution in response to rainfall and temperature variations in Inkomati catchment in Mpumalanga. *Physics and Chemistry of the Earth, Parts A/B/C*, *132*, 103470.

Newete, S. W., Mayonde, S., Kekana, T., & Adam, E. (2023). A rapid and accurate method of mapping invasive Tamarix genotypes using Sentinel-2 images. *PeerJ*, *11*, e15027.

Ncube, B., Shekede, M. D., Gwitira, I., & Dube, T. (2020). Spatial modelling the effects of climate change on the distribution of Lantana camara in Southern Zimbabwe. *Applied geography*, *117*, 102172.

Ndlovu, P., Mutanga, O., Sibanda, M., Odindi, J., & Rushworth, I. (2018). Modelling potential distribution of bramble (rubus cuneifolius) using topographic, bioclimatic and remotely sensed data in the KwaZulu-Natal Drakensberg, South Africa. *Applied geography*, *99*, pp.54-62

Odini, J., & Kakembo, V. (2009). Comparison of pixel and sub-pixel-based techniques to separate Pteronia incana invaded areas using multi-temporal high resolution imagery. *Journal of Applied Remote Sensing*, *3*(1), 033545.

Odindi, J. O., & Kakembo, V. (2011). The hydrological response of Pteronia incana-invaded areas in the Eastern Cape Province, South Africa. *Ecohydrology*, *4*(6), pp. 832-840.

Peerbhay, K., Mutanga, O., Lottering, R., Bangamwabo, V., & Ismail, R. (2016). Detecting bugweed (Solanum mauritianum) abundance in plantation forestry using multisource remote sensing. *ISPRS journal of photogrammetry and remote sensing*, *121*, pp 167-176.

Preston, I.R., Le Maitre, D.C., Blignaut, J.N., Louw, L. and Palmer, C.G. (2018). The impacts of Invasive Alien Plants on Water Provision in Selected Catchments. *Water SA*, *(44)*.

Thanh Noi, P, and Kappas, M. (2017). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors* 18, 1,18.

Royimani, L., Mutanga, O., Odindi, J., Dube, T., & Matongera, T. N. (2019). Advancements in satellite remote sensing for mapping and monitoring of alien invasive plant species (AIPs). *Physics and Chemistry of the Earth, Parts A/B/C*, *112*, 237-245.

Rebelo, A. J., Gokool, S., Holden, P. B., & New, M. G. (2021). Can Sentinel-2 be used to detect invasive alien trees and shrubs in Savanna and Grassland Biomes?. *Remote Sensing Applications: Society and Environment*, *23*, 100600.

Rudrapal, D. and Subhedar, M. (2015). Land cover classification using support vector machine. *Int. J. Eng. Res*, *4*(09), pp.584-588.

Rawat, U.S. and Agarwal, N.K. (2015). Biodiversity: Concept, threats and conservation. *Environment Conservation Journal*, *16*(3), pp.19-28.

Vitousek, P. M., D'antonio, C. M., Loope, L. L., Rejmanek, M., & Westbrooks, R. (1997). Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, pp 1-16.

Sibiya, B., Lottering, R. and Odindi, J. (2021). Discriminating commercial forest species using image texture computed from a worldview-2 pan-sharpened image and partial least squares discriminant analysis. *Remote Sensing Applications: Society and Environment*, *23*, 100605.

Saini, R. and Ghosh, S.K. (2021). Crop classification in a heterogeneous agricultural environment using ensemble classifiers and single-date Sentinel-2A imagery. *Geocarto International*, *36*(19), pp.2141-2159.

Shen, G. and Sarris, A. (2008). Application of texture analysis in land cover classification of high resolution image. In *2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, 3, pp. 513-517.

Sothe, C., Almeida, C. M. d., Liesenberg, V. & Schimalski, M. B. (2017). Evaluating sentinel-2and Landsat-8 data to map sucessional forest stages in a subtropical forest in Southern Brazil. Remote Sensing, 9, 838.

Sites, N. (2000). A global overview of protected areas on the World Heritage list of particular importance for biodiversity.

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.A. and Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sensing*, *12*(7), pp.1135.

Vitousek, P.M., D'antonio, C.M., Loope, L.L., Rejmanek, M. and Westbrooks, R. (1997). Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, pp.1-16.

Wyard, C., Beaumont, B., Grippa, T. and Hallot, E. (2022). UAV-Based Landfill Land Cover Mapping: Optimizing Data Acquisition and Open-Source Processing Protocols. *Drones*, *6*(5), pp.123.

Wang, Q., Putri, N.A., Gan, Y. and Song, G. (2022). Combining both spectral and textural indices for alleviating saturation problem in forest LAI estimation using Sentinel-2 data. *Geocarto International*, 37(25), pp.10511-10531.

Wong, S.N. and Sarker, M.L.R. (2014) February. Land use/land cover mapping using multi-scale texture processing of high-resolution data. In *IOP Conference Series: Earth and Environmental Science* 18 (1), pp. 012185. IOP Publishing.

Xu, B. (2021). Improved convolutional neural network in remote sensing image classification. *Neural Computing and Applications*, *33*(14), pp.8169-8180.