Modelling the Potential Distribution of Bramble *(Rubus Cuneifolius)* in the KwaZulu-Natal Drakensberg, South Africa

Phindile P.P Ndlovu
212559529

A thesis submitted in the fulfilment for the degree of Master of Science in Environmental Sciences, in the School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg.

Supervisor:
Professor Onisimo Mutanga

February 2018

Abstract

Invasive Alien Plant (IAP) invasions have been attracting increasing attention as a result of their substantial effects on native ecosystems. Hence, tools for explaining and predicting IAP distributions have been increasingly promoted for proactive ecological management, and Spatial Distribution Models (SDMs) are one such tool. The main aim of this study was to explore the application of SDMs in modelling the potential distribution of invasive American bramble (*Rubus cuneifolius*) in the Ukhahlamba Drakensberg Park, South Africa. The rapid proliferation of this alien plant has had significant adverse impacts on native plants and the stability of grassland ecosystems. However, there is lack of adequate data on its distribution and factors potentially influencing its present-day habitat range expansions. In that regard, the first objective provides a review of the application of SDMs in modelling the distribution of IAPs and associated challenges and opportunities. As a result of the limitations in traditional methods such as ground surveys, SDMs have demonstrated potential in providing relatively quick and feasible means of predicting IAP distributions, ecological niches and suitability of areas not yet invaded. Literature has shown growth in the use of SDMs for predicting biological invasions with presence-only methods gaining popularity than traditional analyses requiring both presence and absence data. Comparative analyses of model performance found contemporary methods such as Maximum Entropy (Maxent) to have better statistical performance compared to well established modelling approaches. Recent studies also demonstrated that remotely sensed data offers opportunities to explore underlying ecological relationships of species beyond climatic factors and improve the performance of SDMs. The second objective was to model the potential distribution of American bramble using topographic, bioclimatic and remotely sensed data using the Maxent modelling approach. Specifically, this study tested whether variable selection affected model accuracy and the spatial distribution of the species. Model performance was evaluated using the Area Under the curve (AUC), True Skill Statistic (TSS) and Kappa statistic. A quantitative comparison of all models showed that the model built with a composite of all variables yielded the highest AUC score of 0.957. The inclusion of spectral reflectance values improved model accuracy from 0.896 to 0.949. Elevation and rainfall of driest quarter were the most influential variables in modelling bramble distribution. Results of this study showed that bramble are species characteristic of warmer areas with sufficient rainfall and low elevation ranges. In addition, this study demonstrated that the Maxent approach based on topographic, bioclimatic and spectral reflectance values effectively predicted areas susceptible to bramble invasion. Overall,

identification of these areas would assist to guide appropriate management measures and control further incursions.

Preface

This research was undertaken in the School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa, from February 2016 to February 2018, under the supervision of Professor Onisimo Mutanga.

I, the author declare that the research work presented in this research represents my original work, except where due acknowledgements are made by means of referencing. This work has never been submitted to any other academic institution.

Phindile P.P Ndlovu Signed ………………… Date…………………….

As the candidate’s supervisor, I certify the above-mentioned statement and therefore approve this thesis for submission.

Prof. Onisimo Mutanga Signed……………….. Date………………………..

Declaration

I Phindile P.P Ndlovu, declare that:

1. The research reported in this thesis, except where otherwise indicated is my original research.

2. This thesis has not been submitted for any degree or examination at any other institution.

3. This thesis does not contain other person's data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted:

    a. Their words have been re-written and the general information attributed to them has been referenced.

    b. Where their exact words have been used, their writing has been placed in italics inside quotation marks and referenced

5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references section.

Signed:………………………

Dedication

For my family

"Those who trust in the LORD for help will find their strength renewed. They will rise on wings like eagles; they will run and not get weary; they will walk and not grow weak."
(Isaiah 40:31)

Acknowledgements

I would like to express my immense gratitude to the following for their contributions in completing this research:

Firstly, I would like to thank my supervisor Professor Onisimo Mutanga for motivating me to pursue my postgraduate studies. I am grateful for your tireless guidance, understanding and support throughout my academic journey; it is always a pleasure working with you.

My deepest appreciation to Dr. John Odindi for all the time and effort you have put in helping me complete this research project. I am indebted to you for your insights, feedback and advice during my writing process.

To Mzwandile Makhanya, I will be forever grateful for your love and unfailing support. I thank you for the *fairytale* moments lived despite the pressures of life and research. May God continually bless and keep you steadfast on His unwavering love.

I would also like to acknowledge with gratitude Mr Ian Rushworth from Ezemvelo KwaZulu-Natal Wildlife (EKZNW) for your assistance and guidance.

Dr Mbulisi Sibanda and Dr Timothy Dube, thank you for your indispensable advice since the beginning of this journey till the very end.

Trylee Matongera (Shamwari), Sizwe Hlatshwayo, Shenelle Sewel and Eddie Mutsaa, you guys have been my pillar of strength. You kept me going even when I felt like giving up, I am forever grateful. Perushan Rajah, thank you so much for your assistance in the field and beyond.

"*There are friends, there is family. Then there are friends that become family*". Thank you to Sane Mkhize, Ayanda Shange, Sharon Khowa and Zipho Shangase. You have made me grow, kept me grounded and lifted my spirit at all times. Amanda Khuzwayo, Nosipho Makaya, Andisiwe Mseswa and Ntombifuthi Nzimande, you guys brought life and joy in the midst of all challenges throughout this journey. Zinhle Mkhabela, our friendship knows no distance for every time we meet we still speak the same language. Thank you for your unconditional love, encouragement and financial support.

I would also like to express my deepest gratitude to Mrs Shani Ramroop and Mr Donavon Devos.

Last but certainly not least, my heart goes out to my family. I bless the Lord for keeping us united despite all odds. Hlengiwe (Lil'sis), you have truly been my rock for the past years and may God continually keep and groom you into the wonderful woman you becoming. Amile, I pray that one day you will get to enjoy the fruits of all my sacrifices and perseverance.

Bhuda Mtho, we have always had each other's backs no matter what because we know our story, and I pray to God the bond continues to grow and solidify through all our challenges. Slindile and Madu, nothing could ever change my love for you, you took it upon yourselves to become our parents. A shout out to my troublesome nephew and niece (Lindokuhle and Londeka), you guys bring joy to my life. A special thank you to Bab'Vela Mngwengwe, this Masters research wouldn't have been a reality without your support and fatherly guidance.

Table of Contents

List of Tables

List of Figures

Abbreviations

IAP (s)	: Invasive Alien Plants
SDM (s)	: Spatial Distribution Models
UDP	: Ukhahlamba Drakensberg Park
GARP	: Genetic Algorithm for Rule set Production
MARS	: Multivariate Adaptive Regression Splines
GAM	: Generalized Additive Model
GLM	: Generalized Linear Model
GLMM	: Generalized Linear Mixed Model
MAXENT	: Maximum Entropy
BRT	: Boosted Regression Trees
RF	: Random Forest
GIS	: Geographic Information Systems
RS	: Remote Sensing
DEM	: Digital Elevation Model
TWI	: Topographic Wetness Index
TPI	: Topographic Position Index
MI	: Moisture Index
NIR	: Near Infrared
NDVI	: Normalized Difference Vegetation Index
GVI	: Green Vegetation Index
EVI	: Enhanced Vegetation Index
MODIS	: Moderate Resolution Imaging Spectroradiometer
CHIRPS	: Climate Hazards group Infrared Precipitation from Station data
LIDAR	: Light Detection and Ranging
ROC	: Receiver Operating Characteristic
AUC	: Area Under the Curve

Chapter One

General Introduction

1.1 Introduction

Invasive Alien Plants (IAPs) have been identified as one of the most contemporary ecological problems as a result of their recurrent and ubiquitous threat to natural ecosystems, including protected areas where protection of biodiversity is a fundamental goal (Lonsdale 1999; Liu *et al.* 2005; Poona 2008; Pyšek *et al.* 2012). They have received a great deal of attention as they have been observed to alter the ecosystem structure and threaten native biological diversity (van Wilgen *et al.* 2008; Vilà *et al.* 2011). *Rubus cuneifolius* also known as American bramble is a woody perennial shrub which has been considered to be one of the top 10 severe invasive plants prominent in grassland landscapes in South Africa (Hansen *et al.* 2018). Literature shows that bramble was introduced in South Africa around 1900 (Erasmus 1984). The species is rapidly spreading as a result of environmental conditions and the lack of natural enemies and competitors (Erasmus 1984; Hansen *et al.* 2018). For example, in KwaZulu-Natal, bramble has been reported to infest 163 475 Ha threatening native plants and preventing seed production through shading effects (Henderson 2007; Shezi and Poona 2006-2010). Its encroachment threatens specialist grassland taxa having negative effects on local biodiversity within the areas it invades (van Wilgen *et al.* 2008). Due to their negative environmental and economic impacts, bramble species are under category 1 of invasive plants according to the Conservation of Agricultural Resources Act of 1983 (Hansen 2015) thus requiring removal. However, as a result of its efficient reproductive system and rapid response to disturbance, the invasion of bramble makes it difficult, time consuming and expensive to control (Hansen *et al.* 2018). Hence, understanding the spatial distribution of IAPs such as bramble is still a pending critical ecological issue.

Early detection and mapping of IAPs has become essential for taking rapid response and formulating control strategies. However, in South Africa, to our knowledge, assessments of the areas invaded by bramble and its potential distribution have not been conducted. Traditional methods such as field surveys are costly, time-consuming (Evangelista *et al.* 2009; Wakie *et al.* 2014) and challenged by accessibility for large remote areas (Matongera *et al.* 2017). Spatial Distribution Models (SDMs) have been applied as feasible and integrative approaches that employ advanced GIS, remote sensing and modelling algorithms to identify areas that are susceptible to invasion as well as the spatial distribution and spread of IAPs (Thuiller *et al.* 2005; Gallien *et al.* 2010; Barbosa *et al.* 2012). SDMs statistically relate the observed distribution of a species (presence/absence) with environmental variables (Austin 2007; Martins et al, 2016). Numerous methods for developing SDMs exist and have been increasingly applied in modelling IAPs. For example, Zhu *et al.* (2007) used Genetic Algorithm for Rule set Production (GARP) to predict and map the potential distribution of invasive Crofton weed in China. Their results showed that Crofton weed was restricted to warmer regions with tropical climate. On the other hand, Ramírez-Albores *et al.* (2016) used Generalized Linear Model (GLM) to model the distribution of *Shinos molle* in Mexico, its presence was predicted in locations beyond climate thresholds that naturally established species can tolerate. However, the challenge with more established models such as GLM is that they require large sample size as well as presence and absence data which is not always available. In comparative studies of statistical performance, the novel presence-only method, Maxent, was found to have better prediction capabilities and produced useful results with sample size as small as five (Hernandez *et al.* 2006; Elith and Leathwick 2009; Elith *et al.* 2011; Gastón and García-Viñas 2011).

Research has shown that many factors are important in determining invasion success. Whereas climate and topographic data have been commonly used as predictor variables for modelling the potential distribution of IAPs, other studies have provided evidence that the inclusion of remotely sensed predictors improve modelling efforts (Zimmermann *et al.* 2007; Cord *et al.* 2010; Truong *et al.* 2017). Contemporary remotely sensed data has availed opportunities of further exploring ecological factors influencing the distributions of IAPs beyond climatic variables. For instance, Cord *et al.* (2010) noted that incorporating remotely sensed data improved the predictive performance of Maxent for modelling tamarisk species distribution in the United States and Mexico. On the other hand, Truong *et al.* (2017) established that the addition of remotely sensed environmental predictors refined modelled species distributions. Although it has been extensively used for mapping the current distribution of IAPs, its application to SDMs is relatively new and not widely explored.

Both the current and potential distribution of invasive bramble has not been quantified in South Africa and estimations of its spatial distribution can assist in guiding effective management efforts. As a result of the paucity of studies relating the distribution of bramble to environmental variables, there is lack of knowledge on its distribution and its ecological niche requirements. Hence, the lack of knowledge about the habitat requirements of bramble limits the ability to develop efficient eradication strategies. Addressing the infestation of IAPs requires integrated efforts that will overcome processes that lead to species establishment and proliferation (Malahlela *et al.* 2015). Maps of potential distribution of IAPs can be important contributions to better understanding the ecological requirements of the species and guide where key management efforts should be focused. Therefore, this study presents an integrative spatial modelling approach using Maxent to model the potential distribution of bramble in the Ukhahlamba Drakensberg Park (UDP), South Africa using topographic, climatic and spectral reflectance values.

1.2 Aims and objectives

The overall aim of this study was to model the potential distribution of invasive American Bramble (*Rubus cuneifolius*) in the Ukhahlamba Drakensberg Park, South Africa.

To achieve this overarching aim the following objectives were tested:

- We reviewed the application of Spatial Distribution Models (SDMs) in modelling the distribution of Invasive Alien Plants (IAPs), and
- Explored the Maxent approach for modelling the potential distribution of American bramble using topographic, climatic and spectral reflectance values.

1.3 General structure of the thesis

This thesis is presented in four chapters. The first chapter gives an overview of the study, highlighting impacts of IAPs on native ecosystems and the role of spatial modelling in investigating their distribution patterns. More so, this chapter outlines the main aim and objectives of the study. The main structure of this thesis is presented in two core chapters (two and three) in the form of publishable papers. Chapter two, which will be submitted to a peer reviewed journal, reviews the application of spatial distribution models for predicting the spatial patterns of IAPs. The chapter highlights the modelling approaches used and their shortcomings in mapping IAPs. Challenges on the suitability of SDMs for modelling IAPs and improvements of the modelling framework currently applied to build reliable models were also discussed. Chapter three (**accepted for publication**) explores the Maximum Entropy (Maxent) approach to model the potential distribution of American bramble in the Ukhahlamba Drakensberg Park. The study investigated environmental variables influencing bramble invasion and how variable selection affected model accuracy. Chapter four evaluates research objectives and provides a synthesis of all major findings.

Chapter Two

Reviewing the application of Spatial Distribution Models (SDMs) in predicting the spatial distribution of invasive alien plants

Abstract

Spatial Distribution Models (SDMs) that relate statistical relationships between occurrence data and environmental conditions have become important tools for predicting spatial patterns of Invasive Alien Plants (IAPs). Therefore, this study reviewed literature on the application of SDMs in modelling the distribution of IAPs and highlights associated challenges and opportunities. There has been a considerable increase on the use of SDMs for predicting biological invasions across landscapes at different spatial scales. Traditional methods requiring both presence and absence data have been challenged by the unavailability of absence data and difficulty of identifying 'true' absences. As a result, there has been an exploration in the application of presence-only models. GARP, CLIMEX, Maximum Entropy and logistic regression were the models most applied for predicting the spatial distribution of IAPs. However, the accuracy and utility of SDMs has been limited by the lack of balance between model equilibrium assumption and niche stability due to the continued spread of IAPs. Hence, modelling frameworks are continually being developed to improve the applicability of SDMs for predicting IAP distributions. Although not yet widespread, common shortfalls from individual models led to the ensemble approach in order to combine the strengths of different models and proved promising for modelling IAPs. In addition, studies have provided evidence that the inclusion of remotely sensed data as ancillary variables improves modelling efforts and further exploration of underlying drivers of alien plants invasion dynamics. This review concluded that integrative modelling approaches that combine correlative models (SDMs) and advanced remote sensing provide possibilities for improving the capacity of SDMs in modelling IAPs. Such approaches should be explored in high biodiversity regions such as South Africa.

Key words: Invasive Alien Plants (IAPs), Spatial Distribution Models (SDMs), remote sensing, environmental variables, ensemble approach

2.1 Introduction

Invasive Alien Plants (IAPs) are now recognized as a problem of global significance, representing recurrent and pervasive threats to natural ecosystems (Lemke and Brown 2012; Hulme 2012; Steyn *et al.* 2017). These invasions have received a great deal of attention as they are known to compromise ecosystem stability and function (Lonsdale 1999; Pyšek *et al.* 2012), reduce native species richness and abundance (Vilà *et al.* 2011; Barbosa *et al.* 2012), alter fire regimes (Hejda *et al.* 2009) and disassemble natural plant communities (Adhikari *et al.* 2015) of invaded landscapes. Global implications of IAPs on resident species, communities and ecosystems are summarized in Vilà *et al.* (2011) and Pyšek *et al.* (2012). Additionally, in many parts of the world such as South Africa, the problem of IAPs is growing in severity and geographic extent (Richardson and Van Wilgen 2004; Henderson and Wilson 2017). Impacts of plant invasions are well illustrated by the example of significant decreases in native species richness and abundance with increasing *Lantana camara* cover in Australia and India (Gooden *et al.* 2009; Sundaram and Hiremath 2012). The invasion of palatable rangelands by unpalatable invasive alien plants such as bracken fern (*Pteridium*) and mesquite (*Prosopis species*) resulted in reduced grazing potential (Richardson and Van Wilgen 2004; Matongera *et al.* 2016). Similarly, remnant grassland patches in South Africa resulted in communities more characteristic of woodlands due to the invasion of American bramble (*Rubus Cuneifolius*) (Henderson 2007; Hansen *et al.* 2018). As a result, tools that accurately assess suitability of areas and potential spread of IAPs are becoming essential for identifying areas where eradication and control efforts should be focused.

Both ecological and economic implications of IAPs have prompted the urgent call for better methods to provide broad-scale assessments to understand their distribution and spread processes. Traditional methods of delineating habitats of IAPs and providing insights of factors driving species abundance through ground surveys have been argued to be laborious and costly (Evangelista *et al.* 2009; Adhikari *et al.* 2015). Malahlela *et al.* (2015) also noted field surveys as not feasible for large, remote and inaccessible areas as a result of terrain. Spatial Distribution Models (SDMs) have become important tools increasingly employed as rapid and cost-effective alternatives to summarize landscape suitability of areas vulnerable to invasion by IAPs. They have afforded the opportunities to predict the likely distribution of IAPs and explore potential factors that support species presence across landscapes and un-surveyed sites (Liu *et al.* 2005; Elith and Leathwick 2009).

Numerous studies have applied SDMs to predict the distribution of IAPs (e.g Peterson *et al.* 2003; Rouget *et al.* 2004; Le Maitre *et al.* 2008; Stohlgren *et al.* 2010; Simpson and Prots 2013; Ramírez-Albores *et al.* 2016; Truong *et al.* 2017; Bjarnason *et al.* 2017). For example, Peterson *et al.* (2003) used GARP to evaluate the invasion of four invasive plants in North America. Bjarnason *et al.* (2017) used the Generalized Additive Model (GAM) to understand the spatial pattern and spread of IAPs in Greece. Related studies have also attempted to improve the prediction of IAPs distributions. For instance, Stohlgren *et al.* (2010) used the ensemble approach to combine the strengths of several modelling approaches to map the distribution of invasive plants in four sites across the United States. Developing reliable and broadly applicable SDMs requires knowledge of the underlying processes influencing biological invasion patterns and this information lacks for most IAPs (Uden *et al.* 2015). Consequently, most SDMs have been built using climatic variables alone, and recently, studies are now exploring a wide range of variables believed to influence IAP distributions. For instance, Truong *et al.* (2017) assessed the contributions of remotely sensed predictors to SDMs. Although climate has been seen as major driver of species distributions, a variety of abiotic and biotic factors are now being incorporated into SDMs.

A wide range of modelling approaches exists for modelling spatial distribution of species and a growing body of literature describes them. These include general framework for SDMs (Franklin 2010; Guisan and Thuiller 2005; Elith and Leathwick 2009), linking SDMs to ecological theory (Austin 2002; Austin 2007), different modelling approaches (Guisan *et al.* 2002) and application to conservation (Rodríguez *et al.* 2007). Higgins and Richardson (1996) provided a review of models for alien plant spread, however, since then, new statistical techniques for predicting the distribution of IAPs have emerged. On the other hand, Barbosa *et al.* (2012) used a quantitative analysis to review the trends, patterns and gaps on the use of SDMs to predict distribution of invasive species, both terrestrial and aquatic. They focused on investigating whether publications on application of SDMs on invasive species have increased over time, impact factors of journals in which papers are published and countries that had most publications. A recent review by Bradley (2014) explored how different choice of predictors, input data and models influenced conclusions in the habitat suitability framework. Their review focused on the conceptual framework of SDMs in modelling and forecasting IAPs distributions.

There is paucity of literature that reviews the application of the various SDM techniques in predicting the distribution of IAPs, and the frameworks developed overtime to improve and produce meaningful conclusions of habitats vulnerable to invasions. In this regard, the aim of this study is to review the various SDM tools used to model the distribution of IAPs. The study provides an overview of environmental predictor variables that influence IAP establishment and spread. Secondly, this study explores the application of different modelling methods used to predict the distribution of IAPs as well as discuss their shortcomings and statistical performances. The use of different approaches to improve SDMs such as ensemble models and the integration of remotely sensed predictor variables are also discussed. Lastly, we summarize important challenges of SDMs in modelling IAPs as well as directions for future research. Figure 2.1 shows a visual textual summary of single terms frequently appearing in this review.

Figure 2.1: A textual summary of the review

2.2 Factors influencing the spatial distribution of Invasive Alien Plants (IAPs)

A variety of environmental factors determine where a species can and cannot maintain populations (Higgins *et al.* 1999; Mujuni 2014). These factors have presented numerous opportunities to investigate features underpinning species invasions through quantifying species-environment relationships. Therefore, selecting ecologically relevant environmental predictors is crucial when explaining invasion-driving processes. Although disturbance is known to be a critical factor for an invasion process to occur, the success of many plant invaders is attributed to climate and other physical factors (Dark 2004). For terrestrial IAPs, a hierarchical scheme of environmental variables either exert direct or indirect effects on species distributions (Guisan and Thuiller 2005; Austin 2007; Bradley 2014). They act at different spatial scales wherein climate influences the global scale while local to regional processes are influenced by topography, geomorphology, land use and biotic interactions (Ohlemüller *et al.* 2006; Barbosa *et al.* 2012). The earliest found example of species distribution modelling is of Johnston 1924 (cited in Guisan and Thuiller 2005) who predicted the invasive spread of a cactus species in Australia through correlating its distribution and climate. Climatic variables such as temperature and rainfall affect plant growth and competitive ability thus defining boundaries where species can establish, persist and potentially become problematic (Collingham *et al.* 2000; Bradley 2014; Truong *et al.* 2017).

Consistently, many studies have commonly used the climatic approach to predict potential distributions of IAPs (e.g Kriticos *et al.* 2003; Robertson *et al.* 2004; Taylor *et al.* 2012; Adhikari *et al.* 2015; Wan *et al.* 2017). For instance, Campos *et al.* (2016) determined the drivers of the spatial patterns of invasive species in Spain and found them to be strongly sensitive to climatic descriptors with mean annual temperature being the most influential predictor. Ohlemüller *et al.* (2006) and Marini *et al.* (2009) also found a strong positive relationship between mean annual temperature and IAP richness. Cheng and Xu (2015) and Padalia *et al.* (2014) demonstrated the value of other climatic variables such as precipitation of the wettest period for strongly influencing the distribution of invasive *S. Vulgaris* in China and *Hyptis Sauveolens* in India, respectively. A review by Barbosa *et al.* (2012) found climatic variables as predictor variables mostly used to predict IAPs (Figure 2.2). Figure 2.2 shows that climatic variables were mostly used across all spatial scales except the local scale where topographic variables were predominantly used. This pattern was also evident in Austin and

Van Niel (2011) where climatic variables such as temperature and water-related predictors were commonly used factors in studies that they reviewed.

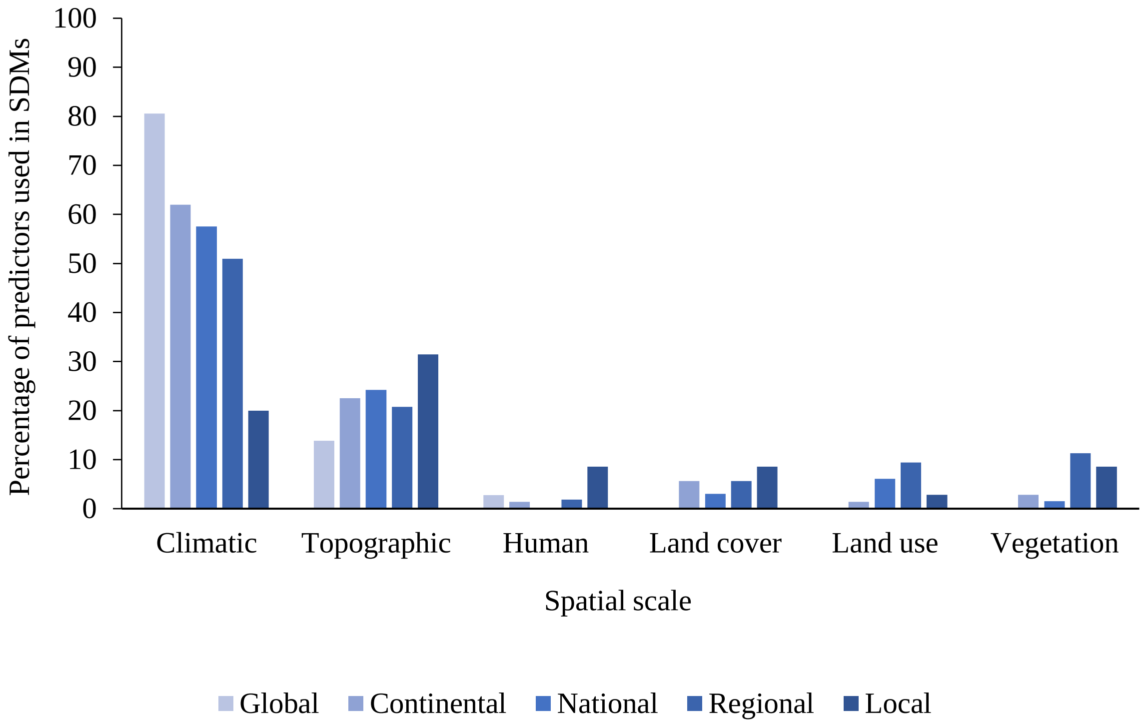

Figure 2.2: Percentage of predictor variables used to predict the distribution of IAPs; adapted from a scientometric analysis by Barbosa *et al.* (2012)

However, studies have questioned the validity of the climatic envelope approach with factors such as topography, disturbance (natural or anthropogenic), hydrological proximity and land use/land cover being also cited as critical factors influencing alien plan invasions (Lonsdale 1999; Dark 2004; Pauchard and Alaback 2004; Ohlemüller *et al.* 2006; Vanderhoof *et al.* 2009; Campos *et al.* 2016). For example, elevation indirectly affects the distribution of species by modulating the micro-climate and soil moisture, while disturbance (roads) act as primary pathways facilitating the introduction of IAPs (Pauchard and Alaback 2004; Dimitrakopoulos *et al.* 2017). Hoffman *et al.* (2008) reported elevation and distance to rivers as the driving factors for the invasion five IAPs in Nebraska. Findings by Benedetti and Morelli (2017) indicated a significant spatial association between IAPs hotspots and roads as well as railways in Germany and Austria. The global potential range and dispersion of 308 IAPs was found to be strongly affected by the human footprint which encompassed human population pressure, land use and infrastructure (Wang and Xu 2016). These studies underscore the importance of other environmental factors in determining invasion success beyond climatic factors alone.

There has also been a considerable debate as to whether biotic interactions influence species distribution patterns (Araújo and Luoto 2007; Zimmermann *et al.* 2010), as a consequence, SDMs have often been derived using abiotic predictors alone. Biotic interactions such as resource opportunity and competition are known to shape the realized environmental niches of IAP species and affect the explanatory power of estimation models (Araújo and Luoto 2007). Contemporary species distribution studies have started to incorporate biotic interactions demonstrating how these reduce the biases in species distributions (e.g Fraterrigo *et al.* 2014; Gallien *et al.* 2015). For example, Gallien *et al.* (2015) identified the importance of including proxies of biotic interactions in detecting IAP occurrence. In their study, they noted that competitive interactions influenced IAP occurrences as most IAPs were found to co-occur with native herbaceous species of similar height. They also found that biotic interaction indices improved model performance with AUC scores ranging from 0.86 (without biotic indices) to 0.96 (with biotic indices). More so, the inclusion of a wide range of variables influencing IAPs distributions has been observed to provide accurate and meaningful predictions of their distributions. For example, models built from both climatic and topographic data were found to be more robust than models based on only climatic factors (Zimmermann *et al.* 2007). Likewise, the addition of remotely sensed data to models built from topo-climatic factors has been observed to increase model accuracy (Buermann *et al.* 2008). These studies indicate that the combination of many factors improves the prediction of areas that are susceptible to invasion by IAPs.

The number of variables used to build SDMs was observed to vary across studies. Austin (2007) noted that the number of predictor variables used to predict IAP distributions ranged between 5 and 38. For example, Wan *et al.* (2017) assessed divergence between native and IAP ranges using 8 bioclimatic variables representing temperature and moisture descriptors. Lemke and Brown (2012) used 28 predictor variables to model the distribution of IAPs at varying levels of occupancy, these included topographic, climatic, disturbance and landcover variables. The number and types of variables used to model the spatial distribution of IAPs has been argued to be greatly influenced by availability as opposed to the consideration of principal/fundamental variables known to determine a species' ecological niche (Duque-Lazo *et al.* 2016). The review by Austin and Van Niel (2011) showed that there are inconsistencies in the choice of variables used for linking known ecological processes, environmental data and SDMs. Yet, studies by Wenger and Olden (2012) and Duque-Lazo *et al.* (2016) demonstrated that model complexity and the number of explanatory variables used are also an important issue in SDMs. Although early studies of species distribution were often built using only climatic variables, the strong interaction between wide ranging factors has been demonstrated to be critical in influencing IAPs distributions. Climatic variables act as large scale determinants while other factors such as topography and disturbance act at smaller resolutions. Consequently, the choice of variables used requires consideration of both the fundamental link to the species' ecological requirements and appropriate scales relative to the species being studied.

2.3 Modelling invasive plant species

Spatial Distribution Models (SDMs) correlate observed distributions with environmental variables to define the spatial habitats of species (Elith *et al.* 2006; Peterson 2001; Bradley 2014). A wide variety of methodological approaches for modelling species distribution exist of which many have been individually or collectively applied to IAPs (Table 2.1). These include BIOCLIM, DOMAIN, General Additive Models (GAMs), Generalized Linear Models (GLMs), logistic regression, Random Forest (RF), Boosted Regression Trees (BRT) GARP and Maximum Entropy (Maxent) (Guisan and Thuiller 2005; Elith *et al.* 2006; Jiménez-Valverde *et al.* 2011; Barbosa *et al.* 2012; Chiou *et al.* 2013). A quantitative analysis by Barbosa *et al.* (2012) showed an increase in the number of publications on spatial modelling of IAPs which is attributed to the increasing interest to curb IAP invasions. They have met many objectives including predicting the spatial distribution and ecological niches of IAPs, identifying habitats vulnerable to species invasions, assessing potential spread and simulating range shifts due to climate change (Lemke and Brown 2012; Young *et al.* 2012; Mujuni 2014; Fernández and Hamilton 2015; Rouget *et al.* 2004). For example, Chiou *et al.* (2013) used BRT modelled the potential range expansion of *Leucaena leucocephala* in Taiwan and found its invasion declined with decreasing average annual temperature and increasing annual precipitation. A Generalized Linear Mixed Model (GLMM) was used to assess the relationship between 106 vascular IAPs attributes and the extent of their range size (Gassó *et al.* 2009). On the other hand, recent studies such as that of Simpson and Prots (2013) demonstrated model extrapolation to forecast invasion range in the face of climate change and their results showed that all species will gain suitable habitats at higher altitudes and spread their ranges.

These modelling approaches differ in their underlying assumptions, algorithms and type of species data they require. One set of methods (for example regression models) use both presence and absence locations to model species locations. Early SDM applications include traditional approaches such as GLM and GAM which predict species presence by finding the relationship between the response and multiple predictor variables (Guisan and Thuiller 2005; Barbosa *et al.* 2012; Elith *et al*. 2015). Ohlemüller *et al.* (2006), for instance, applied the presence-absence GLM algorithm to assess the role played by local and regional predictor variables to alien species richness. Their results found species richness to be highly affected by land cover and climate variables while local soil variables had smaller effects. Likewise, Thuiller *et al.* (2005) applied both GLM and GAM to model 96 IAPs native to South Africa and projected them globally to define high-risk regions susceptible to invasion. Other presence-absence models include Artificial Neural Network (ANN) and decision trees were found to be more flexible and data driven as they were more iterative and appropriate for studies with hierarchical effects on environmental variables (Miller 2010). Related techniques use a splitting strategy through building multiple decision trees, such as Random Forest. Although RF takes advantage of boosting and bootstrap aggregating of the Classification and Regression Tree (CART), it is also challenged by biases in predictors used and interpretability like other data-driven methods (Daliakopoulos *et al*. 2017). However, IAPs absence data has been argued to be problematic because in reality, species may be absent in unsuitable sites for several reasons including failure to disperse in that area (Tsoar *et al.* 2007; Miller 2010). As a result, models such as GLM and GAMs are problematic as they require larger size of sample size due to sensitivity to outliers (Guisan *et al.* 2002). Consequently, models that require both IAPs presence and absence data are said to be suitable for landscape or local scales modelling since the likelihood of 'false' absence is greater at regional scales (Bradley 2014). This also explains the rarity of presence-absence modelling approaches in forecasting species range shifts in response to climate change.

In light of this limitation, presence-only and presence-pseudo absence models have become more popular with GARP, Maxent and CLIMEX being the mostly used algorithms to predict the distribution of IAPs (Elith *et al.* 2006; Barbosa *et al.* 2012; Cheng and Xu 2015). Profile methods models use only presence locations to model species environments without contrasting to absence locations (Elith 2015). For instance, Taylor and Kumar (2013) used presence-only data in developing CLIMEX models for estimating the potential distribution of *Lantana Camara* under current and future climate scenarios in Australia. Their forecasting illustrated that many areas will remain at high risk of invasion. On the other hand, machine learning methods such as Maxent consider the range of the broader landscape through generating background data as absence locations. The model was applied to model suitable areas for pompom weed and two of its biological controls based on the locations of where it had previously been observed (Trethowan *et al.* 2011). The popularity of these aforementioned models is attributed to the wide availability of presence records from historical records and museums. More so, they are software packages that are accessible and easy to implement as compared to other methods that require expertise to use them (Cheng and Xu 2015). This finding is supported by Barbosa *et al.* (2012) who also noted that a large number of studies that modelled AIPs distributions were conducted by presence-only data. Presence-only modelling methods such as GARP and Maxent have also been cautiously applied using extremely small data samples. In that regard, they have been noted to be useful, especially in the cases where IAP are just beginning to spread (Uden *et al.* 2015). However, these modelling approaches also have their individual shortcomings. For instance, presence-only methods have been criticized for always predicting larger areas of suitability as they have no direct mechanism for excluding certain subsets of environmental space. On the other hand, there has been a debate upon how to define ‘background’ from which presence-pseudo absence models choose their random pseudo-absence points (Bradley 2014). Lobo *et al.* (2008) noted that defining the ‘background’ has the likelihood of creating biased models as some IAPs are not in geographic equilibrium. Therefore, presence-only methods are prone to under-prediction while presence-pseudo absence methods suffer from over-prediction. In light of these limitations, the ensemble approach was established to balance and minimize errors (Araújo and Luoto 2007).

Comprehensive comparative analyses such as of Elith *et al.* (2006) that evaluate the statistical performance of a wide range of modelling methods are lacking in the case of IAPs. However, a number of studies have demonstrated variations in performance of models using the current standard of measure Area Under the Curve (AUC). Evident trends reveal that latest methods such as Maxent consistently outperformed more established methods such as GARP, GLM, GAM, DOMAIN and BIOCLIM (Elith *et al.* 2006, Pearson *et al.* 2007, Gastón and García-Viñas 2011). For instance, in a study by Padalia *et al.* (2014), Maxent showed accurate prediction capabilities than GARP for modelling bushmint invasive plants in India with AUC scores 0.86 and 0.75, respectively. On the other hand, a comparative study by Stohlgren *et al.* (2010) found that model performance also varied as a result of species data and study site. For example, in the case of *Carduus nutans* species, AUC scores were high for Random Forest (0.975) followed by logistic regression (0.796) while Maxent performed poorly (0.742). However, Maxent performed better than logistic regression for *Linaria Dalmatica* species with a 0.006 AUC score difference.

Table 2.1: Modelling approaches used to predict the distribution of invasive alien plants

Model	**Data input**	**Model class**	**References**
BIOCLIM	Presence-only	climate envelope	Honig *et al.* 1992; Booth *et al.* 2014; Curtis and Bradley 2015; Wang and Xu 2016
Logistic regression	Presence-absence	regression-based	Vanderhoof *et al.* 2009; Lemke and Brown 2012
Generalized Linear Model (GLM)	Presence-absence	regression-based	Kühn *et al.* 2004; Lloret *et al.* 2005; Mujuni 2014; Ramírez-Albores *et al.* 2016
Generalized Additive Model (GAM)	Presence-absence	regression-based	Goslee *et al.* 2003; Thuiller *et al.* 2005; Campos *et al.* 2016
CLIMEX	presence-only	climate envelope	Dukes and Mooney 1999; Kriticos *et al.* 2003; Shabani and Kumar 2014
DOMAIN	presence-only	multivariate distance	Barbosa *et al.* 2012
MARS	Presence-absence	Non-parametric regression	Friedman 1991; Stohlgren *et al.* 2010; Cunze *et al.* 2013
GARP	presence-background	Machine-learning	Peterson *et al.* 2003; Zhu *et al.* 2007; Padalia *et al.* 2014
Maximum entropy (Maxent)	Presence-background	Machine-learning	Evangelista *et al.* 2009; Simpson and Prots 2013; Truong *et al.* 2017
Artificial Neural Network (ANN)	Presence-absence	Machine-learning	Wang and Xu 2016
Boosted Regression Tree (BRT)	Presence-absence	Decision-tree	Stohlgren *et al.* 2010; Chiou *et al.* 2013

This demonstrated that differences in model performance do not only arise from different algorithm principles but also the selection of variables and input data requirements considered in the model. Models that have demonstrated intermediate predictive performance included GLM, GAM and GARP while BIOCLIM and DOMAIN have poor performance (Elith *et al.* 2006). The above studies indicate that modern modelling techniques demonstrate greater modelling performances as a result of their ability to consider more recent ecological findings and incorporating improved mathematical techniques (Elith *et al.* 2006; Duan *et al.* 2014). However, despite a wide range of comparative studies evaluating the performance of models, no single method was found to be consistently superior across different study sites and characteristics.

Common shortfalls from individual models led to current applications of SDMs demonstrating the use of collective models in an ensemble approach in order to combine strengths of different models. Studies have shown that ensemble models minimize the weakness of any one model and provide a more robust and broad perspective to model results than individual models (Young *et al.* 2012; Fernandes *et al.* 2014). For example, Lemke and Brown 2012 adopted an ensemble approach using Maxent and logistic regression to assess the potential distribution of Japanese Honeysuckle. The ensemble model derived the best fitting model (AUC> 0.8). Conversely, results by Stohlgren *et al.* (2010) indicated that ensemble models do not always improve model accuracy in their study of modelling four harmful IAPs across the United States using five techniques and the ensemble approach. For instance, the ensemble model for *Carduus nutans* (AUC 0.894) was outperformed by RF and BRT with AUC scores 0.975 and 0.931, respectively. Martins *et al.* (2016) identified that the ensemble approach combines the predictions from individual models into a consensus prediction. They noted that this enhanced the agreement of suitable areas in modelling *Hakea sericea* spatial distribution. In this regard, ensemble models are more robust and useful in modelling newly invading IAPs as it is difficult to determine species-environment relationships when species have not spread to all suitable habitats. However, there is still paucity of their application to IAPs.

2.4 Integration of remote sensing to improve SDMs

The inclusion of remotely sensed data as ancillary spatial variables has also been noted to offer opportunities for improving the capacity of spatial distribution models both spatially and temporally. The contribution of integrating remote sensing in SDMs is in two-fold. Firstly, contemporary remotely sensed data enhances predictions of SDMs by providing synoptic, spatial and ecologically relevant variables beyond exclusive climatic suitability (Zimmermann *et al.* 2007; Rocchini *et al.* 2015). To date, a growing number of studies are incorporating remotely sensed data with a range of variables being explored including topography, land cover, spectral and textural indices (Andrew and Ustin 2009; Evangelista *et al.* 2009; Malahlela *et al.* 2015; Truong *et al.* 2017). Predictor variables derived from high resolution Light Detection and Ranging (LIDAR) Digital Elevation Model (DEM) were used to develop a habitat suitability model for *Lepidium Latifolium* in United States of America (USA) (Andrew and Ustin 2009). Studies such as Lemke and Brown (2012) explored the contributions of land cover data in modelling IAPs distributions. However, the use of land cover has been argued to be subject to classification errors such as difficulty in discriminating between vegetation units with similar structure or spectral response (Morán-Ordóñez *et al.* 2012).

Consequently, the accuracy and quality of distribution models may be inadequate for ecological applications. To avoid these limitations, recent studies such as Wakie *et al.* (2014) and Malahlela *et al.* (2015) have focused on the use of spectral variables in invasion ecology to exploit the spectral response from IAPs and analyze underlying ecological relationships of species. For instance, Wakie *et al.* (2014) used monthly Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation indices Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) as input for predicting the current and potential distribution of *Prosopis Juliflora* in Maxent. On the other hand, Malahlela *et al.* (2015) used 11 vegetation indices as ancillary data and found NDVI green vegetation index (NDVIgr) to have a positive correlation to the presence/absence of *Chromoleona Odorata* in South Africa. Alternatively, the use of spectral indices including all channels individually (reflectance values) allow the model to identify the most relevant wavelength channel appropriate for the target species (Moron-Ordonaz *et al*. 2012). Evangelista *et al.* (2009) used Landsat 7 ETM+ satellite scenes, vegetation indices and individual bands to model invasive Tamarisk species. Their study found band 3 as the most influential predictors; however, their study only included 3 individual bands. To the best of our knowledge, this approach remains unexplored in the context of IAPs modelling.

Secondly, the inclusion of remotely sensed data allows predictions closer to actual species distributions. Several studies have provided evidence that the major contribution of remotely sensed variables to SDMs is improved explanatory power while having minimal to no-effect on the predictive power of models (Thuiller *et al.* 2004a). For instance, Truong *et al.* (2017) assessed the contribution of remote RS environmental variables in modelling 14 IAPs in Southeast Asia using Maxent. The comparative analysis of models with various predictor inputs showed that models built with the inclusion of remotely sensed data substantially reduced and refined modelled distributions compared to climatic or topographic models alone. Their study found that although composite models were generally found to perform best for non-native species, not all composite models achieved higher accuracies (Truong *et al.* 2017). Similar findings were realized in the works of Thuiller *et al.* (2004a) and Zimmermann *et al.* (2007) which illustrated that although improving the fit of species models, remotely sensed variables did not improve the cross validated accuracy of models. In contrast to climate SDMs that inevitably reflect species potential distributions, remotely sensed variables add a nuanced spatial detail resulting in predictions closer to actual species distributions (Cord *et al.* 2014). This is in agreement with Pearson *et al.* (2004) who noted that remotely sensed data has the ability to discriminate between suitable and unsuitable areas that cannot be delineated from climate predictors alone. However, studies such as Cord *et al.* (2010) and Malahlela *et al.* (2015) demonstrated that beyond refining modelled distributions, the inclusion of remote sensing predictors also improves the accuracy of models for predicting IAPs. In their study, Cord *et al.* (2010) analyzed the usefulness of Terra-MODIS vegetation index and land surface temperature in modelling the potential distribution of Tamarisk species. They found that remote sensing variables contributed significantly and improved model accuracies in all model scenarios with AUC scores ranging between 0.84 and 0.91. Hence, the inclusion of remotely sensed data is argued to have the potential of improving the reliability of SDMs for assessing biological invasions (Rocchini *et al.* 2015).

Although remotely sensed data has been proven to offer many valuable tools that deserve increased attention for SDMs, its full potential for modelling IAPs remains scarcely explored (Zimmermann *et al.* 2007). For example, ground surveys are laborious and associated with high economic costs as well as incomplete coverage of the landscape. Consequently, most SMDs are based on occurrence data that are poor proxies for invasive species occurrence or abundance (Rocchini *et al.* 2015). The use of RS has afforded the opportunity to improve the reliability of SDMs through generating large and statistically valid species distribution records (Stickler and Southworth 2008). Andrew and Ustin (2009) assessed the ability of advanced remote sensing in modelling habitat suitability of invasive pepperweed in California. In particular, they used hyperspectral classification to obtain presence-absence records of pepperweed and such approaches are still relatively underutilized. In addition, literature shows paucity of studies using remotely sensed climate data. The most common source of temperature and climate data is the freely available WorldClim global dataset. However, replacing these with remotely sensed data variables such as MODIS temperature and Climate Hazards group InfraRed Precipitation with Station data (CHIRPS) precipitation estimates resulted in more accurate predictions (He *et al.* 2015; Deblauwe *et al.* 2016). Although literature shows that inclusion of remotely sensed data in modelling studies has grown and improved the reliability of SDMs, its full potential in modelling IAPs is not yet fully explored.

2.5 Challenges in modelling IAPs

Many publications have highlighted unresolved issues in SDMs in general (e.g Austin, 2002; Guisan and Thuiller, 2005; Elith and Leathwick, 2009). Spatially explicit models are built based on underlying assumptions. They assume that (i) species are at quasi-equilibrium with the environment in which they occur, (ii) sampled data represents relevant environmental gradients, (iii) factors limiting species distributions have been identified and remain constant across space and time (Gallien *et al.* 2012; Uden *et al.* 2015; Mędrzycki *et al.* 2017). However, these assumptions have been noted to be in conflict with the ecological realities of IAPs. In the case of IAPs, SDMs are scarcely applicable because their continued spread through landscapes fails the equilibrium assumption, depending on the stage of their invasion (Mędrzycki *et al.* 2017). As a result, models become prone to predict substantial false presences and absences. For instance, the study by Václavík and Meentemeyer (2012) demonstrated that models calibrated in the early stages of IAP invasions were less accurate than those under scenarios closer to equilibrium. Hence, models calibrated for IAP closer to environmental equilibrium are more likely to be accurate and robust. However, Morecroft (2015) asserted that for the purpose of modelling, species are not required to occur everywhere they could possibly occur geographically; rather, requires that species encompass the entire environmental space of where it could potentially establish. Therefore, adequate sampling of all relevant environmental gradients would result in reliable and broadly-applicable IAP predictions (Uden *et al.* 2015).

The equilibrium assumption of SDMs has also been noted to be violated by Spatial Autocorrelation (SAC) in ecological data used in analyses of species distributions. This is because they are strongly influenced by dispersal and colonization processes (Václavík and Meentemeyer 2012). Whereas the effects of SAC in models has been investigated for native species, it has scarcely been examined in IAPs distribution models (Uden *et al.* 2015). The resultant clustering of range expansion during the early stages of invasion often leads to a mismatch between the potential and realized distributions of species which may distort model predictions. In that regard, considering the effects of SAC is important in modelling the distribution of IAPs. Uden *et al.* (2015) examined the effects of accounting for SAC in IAP models and found that it enhanced the predictive capability of models. Approaches to manage SAC include incorporating predictor variables that are able to quantify both spatial and temporal influences of species response (Smolik *et al.* 2010) or using models that account for SAC effects such as multivariate or machine learning models such as Maxent (Uden *et al.* 2015).

The knowledge of underlying patterns that drive the distribution of IAPs is important for estimating their niches (Smolik *et al.* 2010). For example, Rödder *et al.* (2009) demonstrated that when predicting invaded ranges, models built using variables directly related to the ecology of species resulted in higher discrimination capacity compared to those based on random variables. However, Jiménez-Valverde *et al.* (2011) and Uden *et al.* (2015) noted that this knowledge is still lacking for the vast majority of species and is exacerbated in IAPs as their behaviour varies in space and time with environmental conditions. Other limitations include the scale at which data are available. Young *et al.* (2012) examined how developing models with locally sampled abundance data for regional scale spatial extents and using regional data for predictions at local scale had associated errors. Hence, environmental variables used to develop models are often selected on the basis of availability and observed correlations between variables and species distributions (Austin 2007). In addition, to account for niche stability as IAP niches may differ between native and invaded ranges, Verbruggen *et al.* (2013) and Sosa *et al.* (2017) asserted that both data from the native and introduced ranges are important and should be used to account when modelling the potential distribution of IAPs. For example, when Beaumont *et al.* (2009) calibrated SDMs using the IAPs native and invade ranges, they found that the IAPs occupied different niches than those realized in their native range.

2.6 Future directions in invasive species modelling

Literature shows that there has been undeniable growth and progress in the application of SDMs for modelling IAPs. Whereas modelling frameworks for predicting native species distributions are relatively well developed, their application for IAPs has been rather challenged by equilibrium assumption. However, improvements to produce reliable IAPs distribution models have been achieved such as utilizing data from both native and invaded range as well as ensemble models (Stohlgren *et al.* 2010; Sosa *et al.* 2017). Although there has been progress in terms of widening the scope of environmental variables used for defining species ranges, most SDMs are still predominantly derived from climate data, exclusively. Hence, future research can explore how the selection of predictor variables affects the modelled species distribution and model performance to strengthen the need for explaining IAPs distribution beyond climatic variables.

The contributions of contemporary remote sensing are not only limited to widely available products at multiple spatial and temporal resolution, but have also been demonstrated to enhance model predictions. However, remotely sensed data has been incorporated usually as classified land cover maps and spectral indices. However, land cover maps have been argued to always contain an element of uncertainty that may result from classification errors. Therefore, another area that can be explored is the use of each channel (spectral values) of freely available and improved spatial and spectral multispectral imagery such as Sentinel 2 Multi Spectrometer Imager (MSI) in order to weigh and select optimal bands appropriate for the species of interest. Resultant bands can, therefore, further inform optimal spectral indices to use for improved reliable model predictions. Additionally, a quantitative review found that there is relatively a lack of studies modelling IAPs using SDMs in developing countries (Barbosa *et al.* 2012). Hence, this gap reveals the need for more research on modelling the spatial patterns of IAPs as they are a major concern for native ecosystems especially in regions such as Southern Africa that harbor high biodiversity. Future research and conservation managers could also adopt spatial modelling of IAPs and resultant risk maps as an iterative process which will inform a constant cycle of collecting new field information to perform new models (Uden *et al.* 2015). As a result, risk maps can form as a baseline constantly improved with new information to understand IAPs spatial patterns in order to track and monitor invasion.

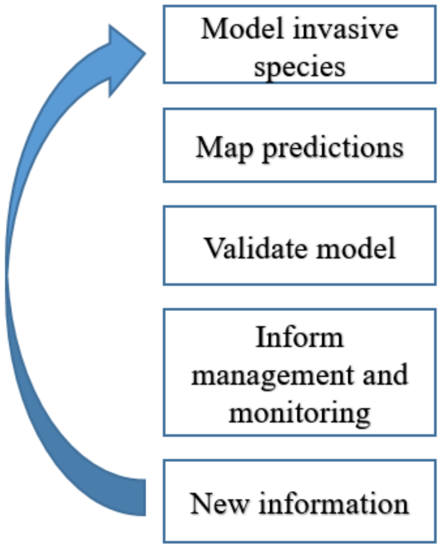

Figure 2.3: Generalized explanation of the iterative process of modelling invasive alien plants

2.7 Conclusion

This study has reviewed existing studies on the application of SDMs in modelling the current and potential spatial distribution of IAPs. Literature has revealed a growth in the application of SDMs in modelling IAPs which is attributable to IAPs being among the most formidable threats to ecosystems. Early modelling techniques such as logistic regression have proved to be useful in modelling IAPs. On the other hand, literature has shown that presence-only techniques such as Maxent have gained popularity in recent years. Comparative analyses revealed that models possessed different predictive capabilities with Maxent being one of the most robust modeling techniques. However, although not widespread, common shortfalls in modelling methods led to the ensemble approach which has proved promising for providing more robust predictions than individual models. Beyond climatic and topographic factors, remote sensing has provided opportunities to develop novel modelling frameworks presenting further exploration of factors underlying IAP distributions. Although SDMs are conceptually straightforward, their underlying assumptions have presented major challenges in the case of modelling IAPs. These include the equilibrium assumption, identification and sampling of all relevant environmental gradients which has proven to be in conflict with ecological realities of IAPs. However, despite these limitations, modelling approaches such as the inclusion of remotely sensed ancillary data have opened up possibilities for improving the capacity of SDMs in modelling IAPs. Hence, this study recommends that future research, especially developing countries of southern Africa, should focus on freely accessible modelling methods such as Maxent and freely available multispectral data to improve the potential distribution of IAP as well as their spread processes.

Chapter Three

Modelling the potential distribution of Bramble (*Rubus Cuneifolius*) in the KwaZulu Natal Drakensberg.

This chapter is based on:

Ndlovu, P., Mutanga, O., Sibanda, M., Odindi, J., & Rushworth, I (**accepted for publication**). *Modelling the potential distribution of Bramble (Rubus Cuneifolius) in the KwaZulu-Natal Drakensberg, South Africa*. Journal of Applied Geography, Elsevier, Manuscript number: JAPG_1941

Abstract

The American bramble (*Rubus cuneifolius),* a woody perennial invasive shrub, presents serious ecological and economic impacts, particularly in ecologically rich and protected landscapes. Since the ecological factors determining its geographic distribution are poorly understood, a comprehensive analysis and understanding of its potential distribution are essential to understand probable impacts and plan control interventions. Hence, this study sought to explore the use of Maximum Entropy (Maxent) modelling approach to determine the potential distribution of American bramble in the uKhahlamba Drakensberg Park (UDP), South Africa. Four sets of model scenarios based on topographic data, topographic and spectral reflectance data, topographic and bioclimatic data and a composite of all variables were generated using 73 occurrence points. Model performance was evaluated using Area Under the Curve (AUC), True Skill Statistic (TSS) and Kappa statistic. The model built using a composite of all variables yielded the highest accuracies, AUC score (0.957), indicating the best prediction of suitable and unsuitable areas for bramble. The inclusion of remotely sensed data improved model performance with bramble reflecting highly on the red edge bands. Elevation and rainfall of driest quarter were the most important variables associated with bramble distribution. The models predicted low elevation, warm and moist eastern parts as most suitable for bramble establishment and growth. Overall, all the models matched in terms of the geographic extent predicted as probable bramble distribution. Our results demonstrate that an integration of topographic, bioclimatic and remotely sensed variables are useful in determining landscape vulnerability to bramble invasion and provide a valuable tool for planning control strategies.

Keywords: Invasive alien plant species, American bramble, Maximum Entropy (Maxent)

3.1 Introduction

Invasive alien plant species have been recognized as a growing threat to the integrity of ecosystems in protected landscapes (Poona 2008; Jarošík *et al.* 2011; Hulme 2012; Fandohan *et al.* 2015). South Africa has seen a significant percentage of invasive alien plants (van Wilgen *et al.* 2008), notably, *Lantana camara*, parthenium hysterophorus and American bramble which have caused perceptible threats to the native biodiversity. *Rubus cuneifolius,* also known as the American bramble has been identified as one of the top ten most aggressive invasive plants prominent in grass-dominated landscapes of KwaZulu-Natal (Henderson 2007). Once established, it forms impenetrable dense and thorny stands that outcompete indigenous plant species and restrict access to grazing and water by wild and domestic animals (Hansen 2015; Hansen *et al.* 2018). Encroachment of bramble into grasslands is known to lower species richness and diversity resulting in grassland communities that are more characteristic of woodlands (Hansen 2015). Hence, its spread has stimulated interests in invasion ecology that include the predictive understanding of the spatial distribution. Invasion vulnerability maps are becoming increasingly important in invasion ecology as they offer great opportunities for mitigating spread (Jiménez-Valverde *et al.* 2011). For instance, such information can provide insights on species tolerances or their potential geographical range, useful for landscape management initiatives.

Specifically, Spatial Distribution Models (SDMs) offer possibilities to estimate probabilities of species occurrence in response to environmental conditions (Guisan and Thuiller 2005; Wisz *et al.* 2008). These empirical tools use data that assume species occur in suitable habitats (true presences) and are absent in unsuitable habitats (true absences) (Elith *et al.* 2006; Hirzel and Le Lay 2008). Their strength relies on the correlation between the observed species distribution and input parameters that represent suitable conditions such as climatic and landscape characteristics (Lemke and Brown 2012). To date, a wide variety of models have been adopted to predict invasive alien plant species distributions. These include Genetic Algorithm for Rule-set Production (GARP) (Peterson 2001; Zhu *et al.* 2007), logistic regression (Vanderhoof *et al.* 2009; Lemke *et al.* 2011), Generalized Linear Model (GLM) (Wang and Xu 2016) and Maximum Entropy (Maxent) (Lemke and Brown 2012; Fandohan *et al.* 2015). In comparative studies, Maxent, has been identified as a robust algorithm that can be used to optimally model the spatial distribution of species across a range of sample sizes (Elith *et al.* 2006; Hernandez *et al.* 2006; Wisz *et al.* 2008). Its success is attributed to its regularization procedure that avoids overfitting and compensates for small occurrence data (Phillips *et al.* 2004). Subsequently, Maxent has become invaluable in modelling the potential distribution of alien invasive species, and a potentially valuable source of information for grassland monitoring and management efforts (Ficetola *et al.* 2007; Evangelista *et al.* 2009; Stohlgren *et al.* 2010).

The adoption of GIS and Remotely sensed data has become increasingly appealing in SDMs as they offer a wide range of spatial information that includes surface conditions and interpolation of climate parameters (Hirzel and Le Lay 2008; Kozak *et al.* 2008). Whereas climate and topographic variables have been commonly used to evaluate species distributions, recent studies have proved that the integration of remotely sensed data as ancillary spatial variables improves the performance of SDMs (Cord and Rödder 2011; Rocchini *et al.* 2015; Truong *et al.* 2017). A range of remotely sensed datasets that include land cover (Pearson *et al.* 2004), spectral indices (Zimmermann *et al.* 2007), and surface spectral reflectance (Morán-Ordóñez *et al.* 2012) have been explored in SDMs. The use of spectral reflectance values, in particular, provides models with species information that is spectrally unique, aiding the model to discriminate between suitable and unsuitable areas that cannot be distinguished from topographic or bioclimatic factors alone (Zimmermann *et al.* 2007). More so, it gives the model freedom to select and weigh the bands appropriate for discriminating the species (Morán-Ordóñez *et al.* 2012). Therefore, we hypothesize that the inclusion of spectral reflectance characteristics as well as the determination of the optimal combination of bands that best reflect the ecological properties of bramble can be valuable in spatial distribution modelling.

The invasion of bramble in the KwaZulu-Natal grassland biome has become increasingly severe (Henderson 2011). In the study area, bramble has been observed to aggressively invade and re-establish on both pristine and disturbed grasslands (Hansen 2015), making it difficult and costly to control. Whereas the invasion of bramble has been observed to be problematic and detrimental to social and ecological systems, there is a paucity of literature on its habitat preferences and geographical distribution. In that regard, there is a need for predicting the spatial distribution of bramble based on its current occurrence in order to understand its ecological requirements and areas of potential distribution within a landscape. This study, therefore, sought to model the potential distribution of bramble (*Rubus Cuneifolius*) in the uKhahlamba Drakensberg Park (UDP) in relation to topographic, bioclimatic and remotely sensed data variables using the Maxent model. Specifically, this study sought to determine the most important variables (s) as well as testing how variable selection affected Maxent model performance in modelling bramble's spatial distribution.

3.2 Materials and methods

3.2.1 Study site

The study was conducted in the uKhahlamba Drakensberg Park, a UNESCO World Heritage site in KwaZulu-Natal province, South Africa (Nel 2009). The park covers an area of approximately 242 813 ha, with average elevation ranging between 1195 and 3451m above sea level (Krüger and Crowson 2004). The area is strongly seasonal and receives 75% of its annual rainfall between November and March, with mean annual rainfall highly related to altitude below 2100 meters above sea level (Nel and Sumner 2006; Nel 2009). Its climate is characterized by warm wet summers and cold dry winters (Hoerlé 2006). Mean temperatures vary between seasons and tend to decrease with increase in altitude. Temperature ranges from 13.1^{o} to 21.1^{o} in January and 9.2^{o} to 15.2^{o} in September (Bishop *et al.* 2014; Matongera *et al.* 2016). The associated vegetation species are highly responsive to climatic gradients, with grass and wood lands dominating lower altitudes with moist aspect (Matongera *et al.* 2016). The park is a national and international asset due to its biodiversity as it protects a high level of endemic and globally threatened species. More so, it is of major economic importance due to its catchment area that produces quality water from its diverse system of wetlands and rivers (Rushworth, 2011). However, invasive alien plant species have been identified as the most prominent threat to the biodiversity and water production of the park (Van Wilgen *et al.* 2001; Rushworth, 2011). These include American bramble which has been notably increasing exponentially with negative impacts on biodiversity, tourism and water production. The UDP comprises of 17 management units of which six (Monks cowl, Injesuthi, Highmoor, Kamberg, Mkhomazi and Cobham) are regarded as areas of high bramble infestations (Figure 3.1). These highly infested areas were the focus of the sampling strategy.

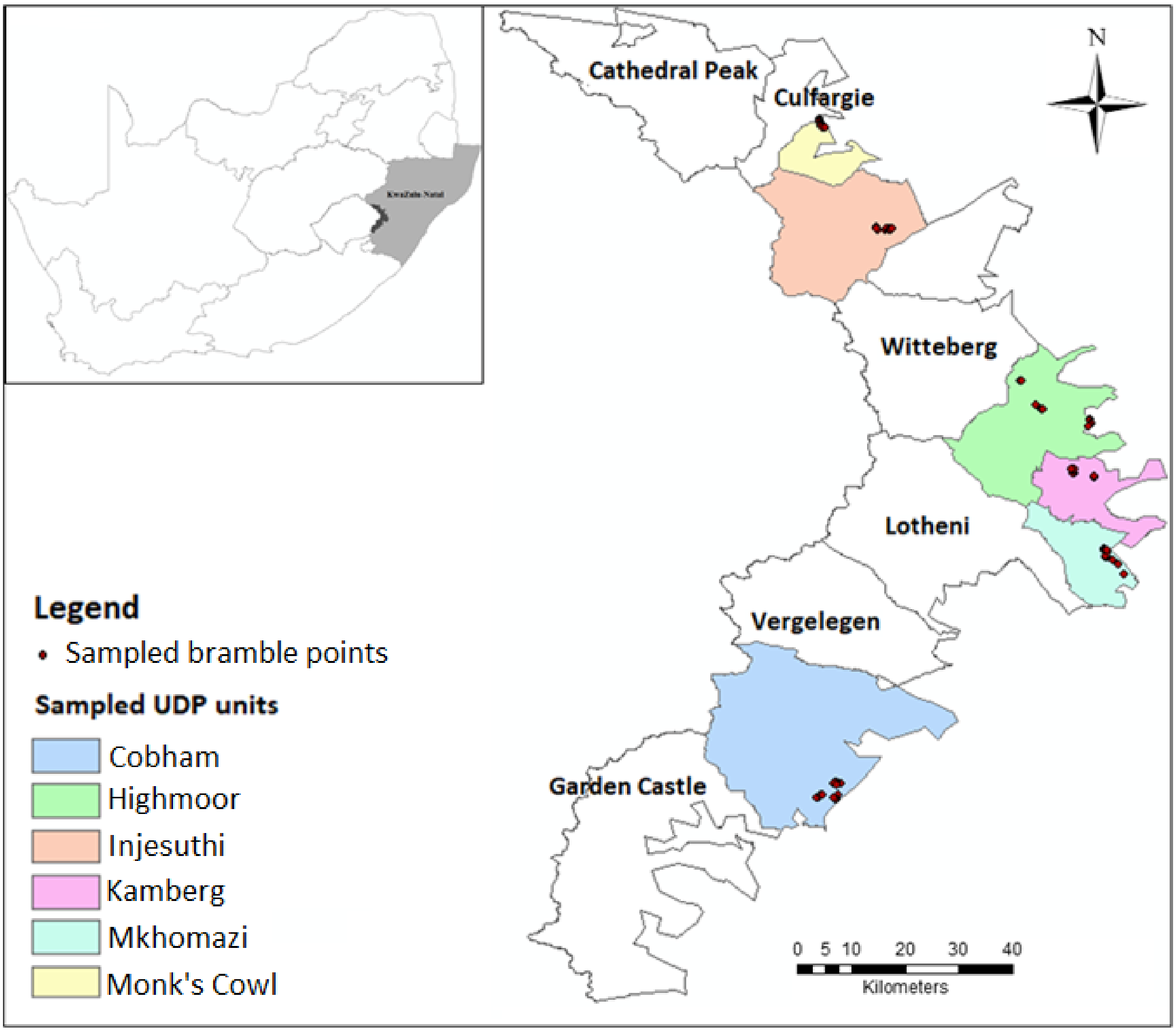

Figure 3.1: Location of study area in uKhahlamba Drakensberg Park (UDP) *and management units*

3.2.2 Field data

To determine bramble spatial distribution, two field surveys were conducted in October and December 2016. Field surveys were conducted during this period as bramble was at robust growth and distinct from other surrounding vegetation types (Figure 3.2). Using a hand held Trimble GeoHX 6000 Global Positioning System (GPS) with sub-meter accuracy; purposive sampling was used to collect data on bramble occurrence. The measured GPS points were converted to comma separated values (csv) compatible to Maxent. Ultimately, a total of 73 bramble locations were collected and used to model its spatial distribution.

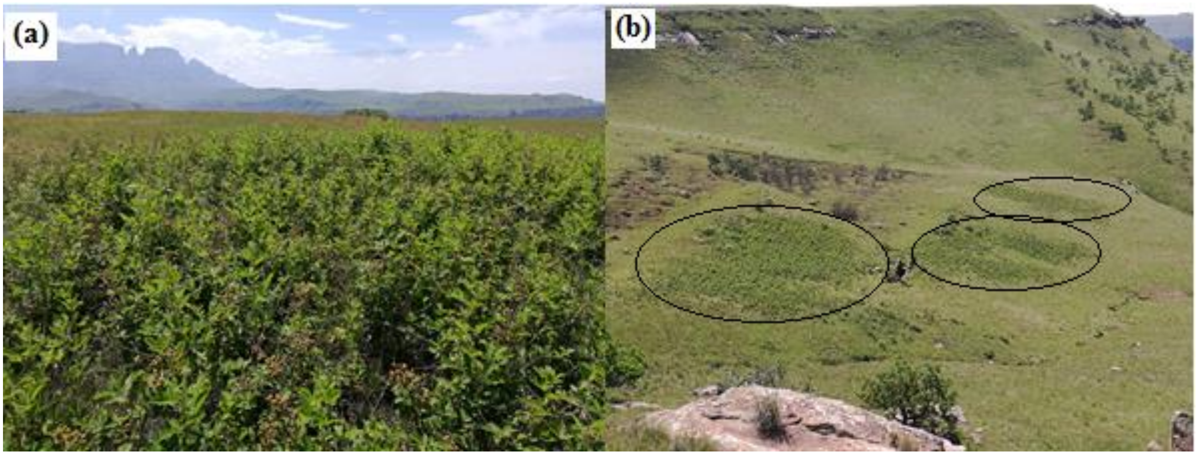

Figure 3.2: Bramble patches during (a) optimal growing season, and (b) at landscape scale

3.2.3 Remotely sensed data

Freely available Sentinel 2 - Multi-Spectral Instrument (MSI) imagery was used to test the utility of remotely sensed data in improving the prediction of bramble's spatial distribution. A cloud-free Sentinel 2 MSI image acquired on the 1st of January 2017 was downloaded from *Geocento* portal for analysis. The Sentinel 2 MSI sensor operates on 13 spectral bands ranging from the visible and near-infrared to shortwave infrared (Sibanda *et al.* 2015). Among the 13 spectral bands, four (2, 3, 4 and 8) acquire images with a 10m spatial resolution, six (5, 6, 7, 8a, 11 and 12) at 20m spatial resolution and three (1, 9 and 10) at 60m spatial resolution (Sibanda *et al.* 2015; Addabbo *et al.* 2016). The sensor's characteristic image acquisition at different wavelengths demonstrates its value in a range of investigations (Addabbo *et al.* 2016). More so, Sentinel 2 MSI has 3 bands in the red-edge region, which are absent in previous sensors like Landsat 8 (Sibanda *et al.* 2015). Atmospheric correction was performed using Sen2cor toolbox, a built-in algorithm within the Sentinel Application Platform (SNAP) tool version 4.0. Sen2cor performs atmospheric correction from Top-of-Atmosphere to Bottom-of-Atmosphere and converts Top-of-Atmosphere reflectance into canopy reflectance. Although the bands were correlating, the model was built using all bands so as to determine optimal band(s) for predicting bramble distribution (Table 3.1).

Table 3.1: Name and spatial resolution (m) of the corresponding Sentinel 2 MSI bands assessed in this study

Bands	**Band name**	**Resolution (m)**
Band 2	Blue	10
Band 3	Green	10
Band 4	Red	10
Band 5	Vegetation red edge	20
Band 6	Vegetation red edge	20
Band 7	Vegetation red edge	20
Band 8	Near Infrared (NIR)	20
Band 8a	Vegetation red edge	20
Band 11	Shortwave Infrared (SWIR)	20
Band 12	Shortwave Infrared (SWIR)	20

3.2.4 Topographic data

A 30 m Digital Elevation Model (DEM) generated from 1:50 000 contour lines was obtained from the Department of Agriculture and Environmental Affairs. Topographic variables were derived using the spatial analyst tools in ArcGIS. The topographic Wetness Index (TWI) quantifies the topographic influence on hydrological processes and was used as a measure for soil moisture (Raduła *et al.* 2018). A Topographic Position Index (TPI) was used to classify the landscape into slope position classes to identify the topographic positions preferred by bramble. Other topographic variables included elevation, slope and Aspect.

3.2.5 Bioclimatic data

Previous studies have identified bioclimatic variables as essential driving factors for species distributions (Kgosiesele 2010; Thuiller *et al.* 2004b). In this study, current climate data layers (1960-1990) generated through the interpolation of average monthly data using splining techniques were obtained from the WorldClim database (http://www.worlclim.org/) (Hijmans *et al.* 2005). The bioclimatic variables were derived from monthly temperature and rainfall values to produce variables that are biologically relevant (Hijmans *et al.* 2005). The variables were obtained in raster grid format with a 30'' arc seconds (1 km^2) spatial resolution. These variables were categorized into temperature and moisture variables (Table 3.2).

Table 3.2: Bioclimatic variables from WorldClim database (explained in Hijmans *et al.* 2005)

Abbreviation	Name	Units
	Temperature variables	
Bio01	mean annual temperature	^{o}C
Bio02	mean diurnal range in temperature	^{o}C
Bio03	*Isothermality (bio 02/bio 07)X100	^{o}C
Bio04	temperature seasonality	^{o}C
Bio05	maximum temperature warmest month	^{o}C
Bio06	minimum temperature coolest month	^{o}C
Bio07	annual temperature range	^{o}C
Bio10	mean temperature warmest quarter	^{o}C
Bio11	mean temperature coolest quarter	^{o}C
	Moisture variables	
Bio12	mean annual rainfall	mm
Bio13	rainfall wettest month	mm
Bio14	rainfall driest month	mm
Bio15	**rainfall seasonality (coefficient of variation)	mm
Bio16	rainfall wettest quarter	mm
Bio17	rainfall driest quarter	mm
Mi	annual moisture index	n/a

*Isothermality is the 'evenness' of temperature over the course of the year. It quantifies the length of day-to-night temperature oscillations compared to summer-to-winter oscillations (O'Donnell and Ignizio, 2012). It evaluates species that grow well in isothermal environments (constant temperature).

**Rainfall seasonality is an index of rainfall variability quantified through measuring the variation of monthly rainfall totals over the course of the year (O'Donnell and Ignizio, 2012). A large percentage of rainfall variability means there is greater rainfall variability.

Maxent requires that all explanatory variables have the same pixel size, extent and projection system. Therefore, all the other variables were resampled to 30m spatial resolution and projected to the Universal Transverse Mercator (UTM) projection to match topographic variables. The use of 30m resolution is supported by comparative studies of model performance at different spatial resolutions which showed improved model accuracies with decreasing spatial resolution (Ross *et al.* 2015; Manzoor *et al.* 2018). The input data was assessed for any autocorrelations based on Pearson correlation test as in Makori *et al.* (2017). Auto-correlated variables were not included in the model. The variables that had a correlation coefficient $-0.8 < r < 0.8$ were not selected for modelling bramble distribution. Resultant environmental predictor variables were classified into four categories (Table 3.3).

3.2.6 Modelling bramble distribution

The potential distribution of bramble was modelled using the freely available Maximum entropy version 3.4.0 (http://biodiversityinformatics.amnh.org/open_source/maxent/) (Phillips *et al.* 2017). Maxent models the probability of a distribution by estimating the most uniform distribution across the study area, while considering all constraints representing information about the distribution of the species (Phillips *et al.* 2006; Hernandez *et al.* 2006; Ficetola *et al.* 2007). The approach is based on the principle that the potential distribution must agree with the information that is inferred from the environmental conditions at the occurrence locations while avoiding unfounded constraints (Pearson *et al.* 2007). Therefore, the model evaluates each grid cell as a function of input environmental variables.

Maxent's major advantage is that it uses presence-only data with the ability to incorporate interactions between continuous and categorical data (Ficetola *et al.* 2007). Additionally, its algorithm is developed to converge with the optimal probability distribution and is less influenced by the number and spatial error of sample size as compared to related models e.g. GARP (Hernandez *et al.* 2006). For this study, the Maxent algorithm was performed under different modelling scenarios (Table 3.2). The data was split into 70% training and 30% test data randomly selected by the model within the study area. Model parameters were set to default replication of 1 with 500 iterations using cross-validation run type. Regularization multipliers were set to 4 to reduce overfitting (Cao *et al.* 2016) and the cloglog output format was used (Phillips *et al.* 2017). During training, Maxent also performs a jackknife test to assess the relative importance of predictor variables in explaining the distribution of the species and unique information provided by each variable (Phillips and Dudík 2008). Subsequently, the most important variable in the model's development is that which decreases the training gain when it is excluded and increases when its included (Phillips *et al.* 2006; Chikerema *et al.* 2013).

Table 3.3: Bramble model scenarios with different environmental inputs

Model scenario	Variables	No. of variables
Model 1	Elevation, slope, TWI, TPI and aspect	5
Model 2	Elevation, slope, TWI, TPI, aspect and Sentinel-2 MSI bands	15
Model 3	Elevation, slope, TWI, TPI, aspect, Bio 01, Bio 02, Bio 05, Bio 06, Bio 07, Bio 12,Bio 13, Bio 14, Bio 17 and moisture index	15
Model 4	Elevation, slope, TWI, TPI, aspect, Sentinel-2 MSI bands, Bio 01, Bio 02, Bio 05, Bio 06, Bio 07, Bio 12,Bio 13, Bio 14, Bio 17 and moisture index	25

3.2.7 Model evaluation

Model performance was evaluated using the threshold-independent (AUC) and threshold dependent (TSS and Cohen's Kappa) measures of accuracy. The AUC tests the agreement between the observed species presence and the estimated distribution, indicating whether the probability of presence (sensitivity) versus absence (specificity) was correctly ordered by the classifier (Phillips *et al.* 2006; Makori *et al.* 2017). Plots of sensitivity against 1-specificity are generated by the Maxent algorithm and produced as part of the outputs. The model with an AUC value less than or equal to 0.5 is considered to have a random prediction whereas AUC value greater than 0.5 is more than random, with values approaching 1 indicating a better prediction (Hernandez *et al.* 2006; Matawa *et al.* 2012). TSS and Kappa statistics were used as threshold-dependent measures of model accuracy. Kappa is the mostly used measure of model performance but has been criticized for dependence on prevalence (Allouche *et al.* 2006). TSS has become an alternative measure of accuracy that corrects this dependence while retaining the advantages of Kappa. The error matrix was used to derive specificity, sensitivity, Kappa and TSS values using background samples as absence data. The 10 percentile threshold value was used to evaluate classification accuracy. TSS is defined as specificity + sensitivity -1 (Allouche *et al.* 2006).

3.3 Results

3.3.1 Descriptive data analysis

Figure 3.3 shows the correlation between input environmental variables. A total of 25 variables with correlation -0.8< *r* <0.8 as recommended by Lemke and Brown (2012) and Young *et al.* (2012) were used for modelling bramble distribution. These variables included elevation (DEM), slope, aspect, Topographic Position Index (TPI), Topographic Wetness Index (TWI), spectral reflectance of Sentinel 2 MSI bands, Bio01, Bio 02, Bio 05, Bio 06, Bio07, Bio12, Bio 13, Bio14, Bio 17 and MI. Despite the observed correlation between the Sentinel 2 MSI bands, all bands were included in order to evaluate which bands best predicted bramble distribution.

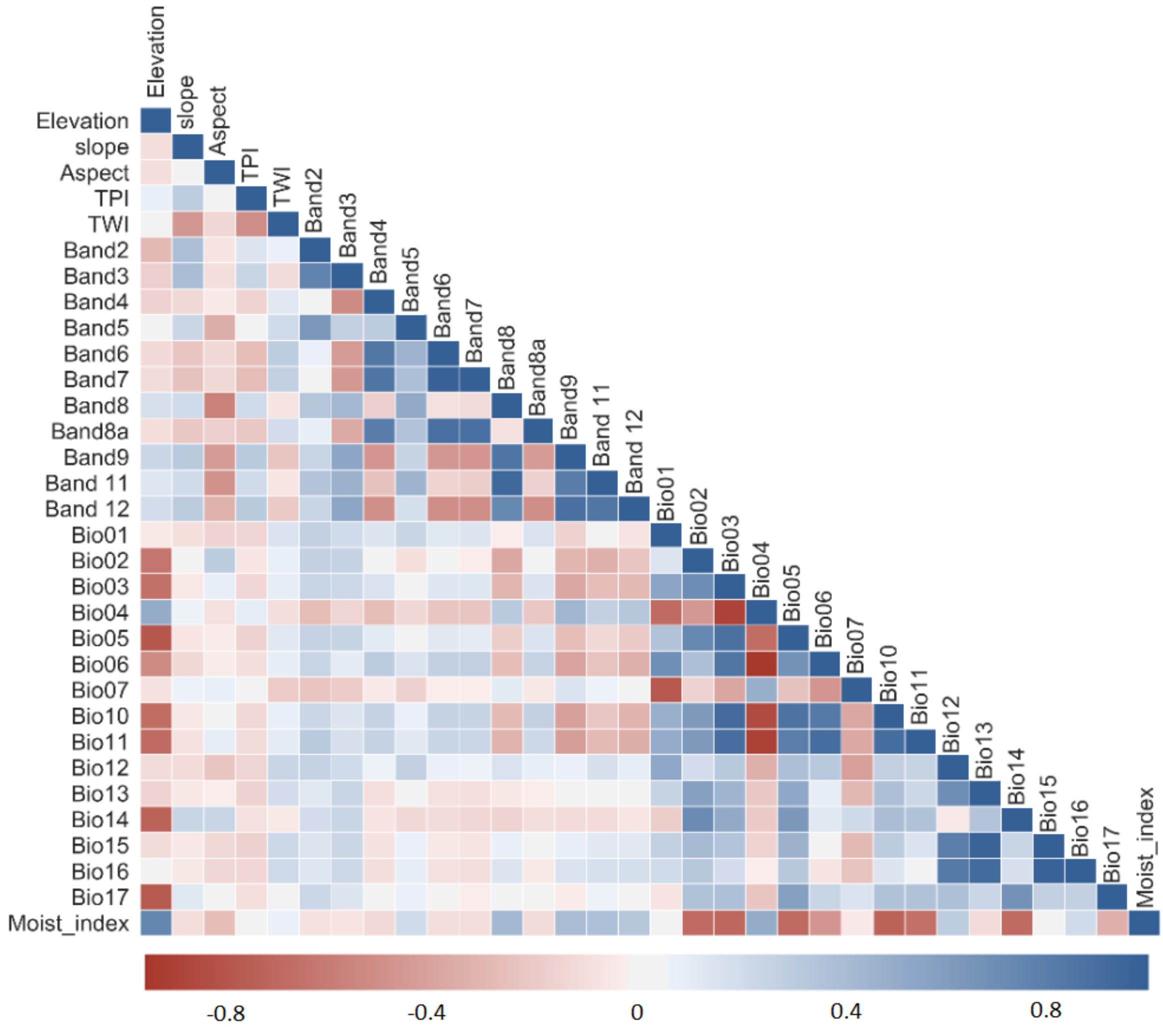

Figure 3.3: Pearson correlation test of input environmental variables

3.3.2 Model accuracy

Table 3.4 shows the threshold-independent (AUC) and threshold-dependent (TSS and Kappa) values of the randomly selected test data for bramble habitat prediction for the four models. The model built with a composite of all variables achieved the highest predictive accuracies (AUC 0.957, TSS 0.808 and Kappa 0.414). The addition of bramble spectral values as predictors improved model accuracy by 0.053 (AUC) indicating their additional information value. Model built on topographic data alone had the lowest performance (AUC 0.896, TSS 0.655 and Kappa 0.369). Although the increases of the cross-validated AUC scores were not large, they displayed a clear trend.

Table 3.4: Evaluation results for all model scenarios

Model scenarios	AUC	TSS	Kappa
Model 1	0.896	0.655	0.369
Model 2	0.949	0.781	0.409
Model 3	0.942	0.467	0.292
Model 4	0.957	0.808	0.414

Figure 3.4 shows the results of the jackknife test of variable importance. The analysis of variable importance ranked elevation as the most influential variable in predicting bramble distribution in models 1 and 2. Rainfall driest quarter (Bio 17) was the most important variable for models 3 and 4. These variables had the highest gain when used in isolation suggesting that they contained the most useful information, compared to other variables. Elevation also decreased model gain when omitted in models 1 and 2 indicating to have information not present in other variables. The inclusion of remotely sensed data in model 2 depicted band 8a as the most important variable indicating high bramble responses on the red edge band. The training gain for TPI and Band 3 maintained the least training gain in all respective models. Therefore, these variables have no influence on the outcome of the model.

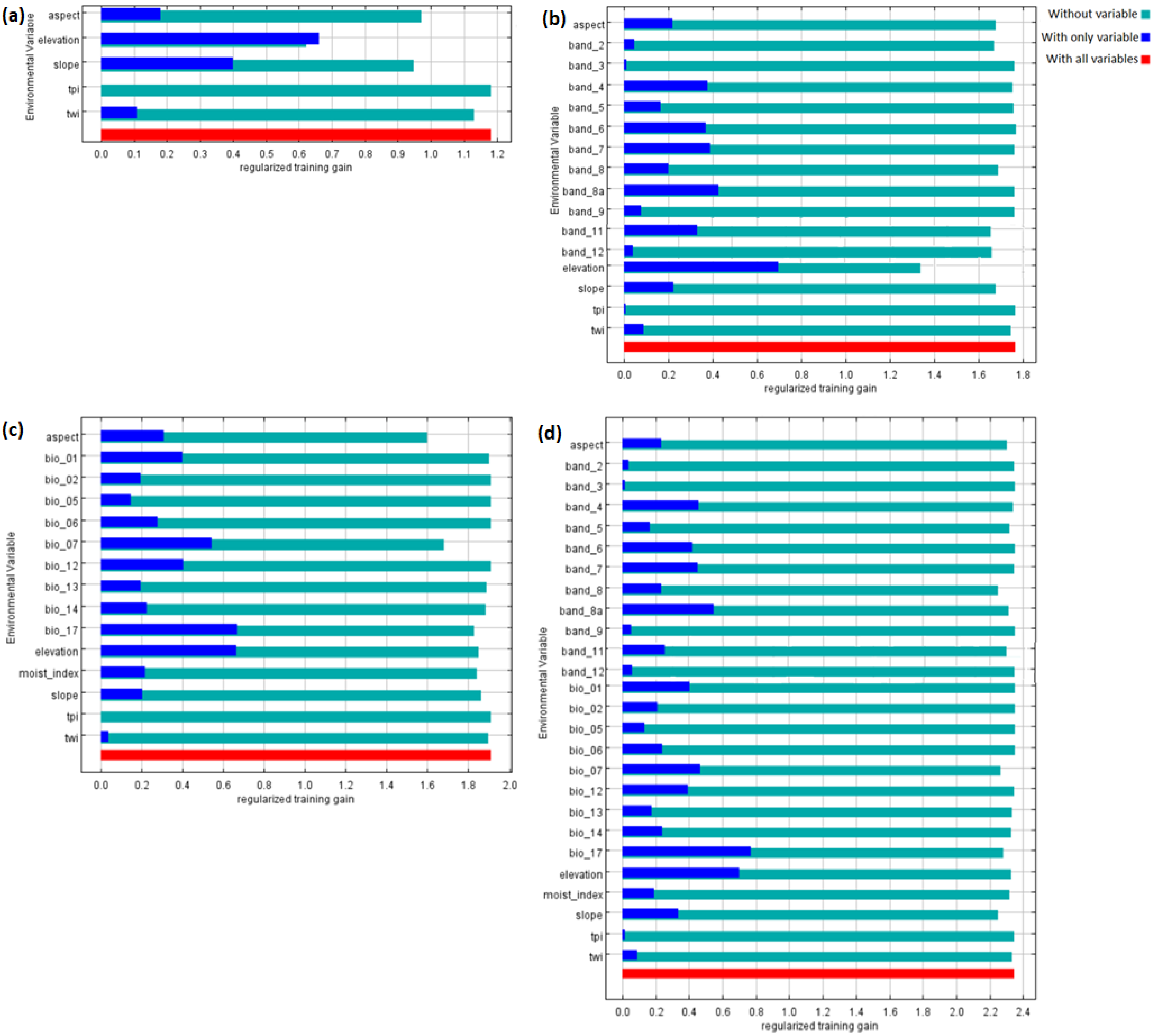

Figure 3.4: Jacknife test of variable importance. (a) Topographic variables (model 1), (b) topographic and Sentinel 2 imagery (model 2), (c) topographic and bioclimatic variables (model 3) and (d) composite of all variables (model 4) developed for bramble

3.3.3 Bramble spatial distribution

Figure 3.5 illustrates the potential geographic distribution of bramble across the study area based on the four model scenarios. Generally, all the model scenarios indicated similar geographic distribution making the comparison between the model outputs difficult. However, areas with a high suitability for bramble invasion were predicted in the eastern region of the study area. This spatial constriction corresponds to warm temperature and high rainfall regions. The models predicted absence in the western region coinciding with elevation greater than 2100 m, low temperatures, and limited rainfall. In all models, high suitability of bramble is observed in Injesuthi extending down to Hillside and Witteberg, north of the Kamberg, eastern part of Cobham and Garden Castle (Figure 3.5 a-d). Models built on topographic data (Figure 3.5a) and the inclusion of remotely sensed data (Figure 3.5b) showed a higher index of bramble occurrence at Cathedral Peak, which other models failed to capture. All the model predictions agreed well on the high suitability of bramble in areas such as Hillside, Witteberg and northern part of Kamberg (Figure 3.5).

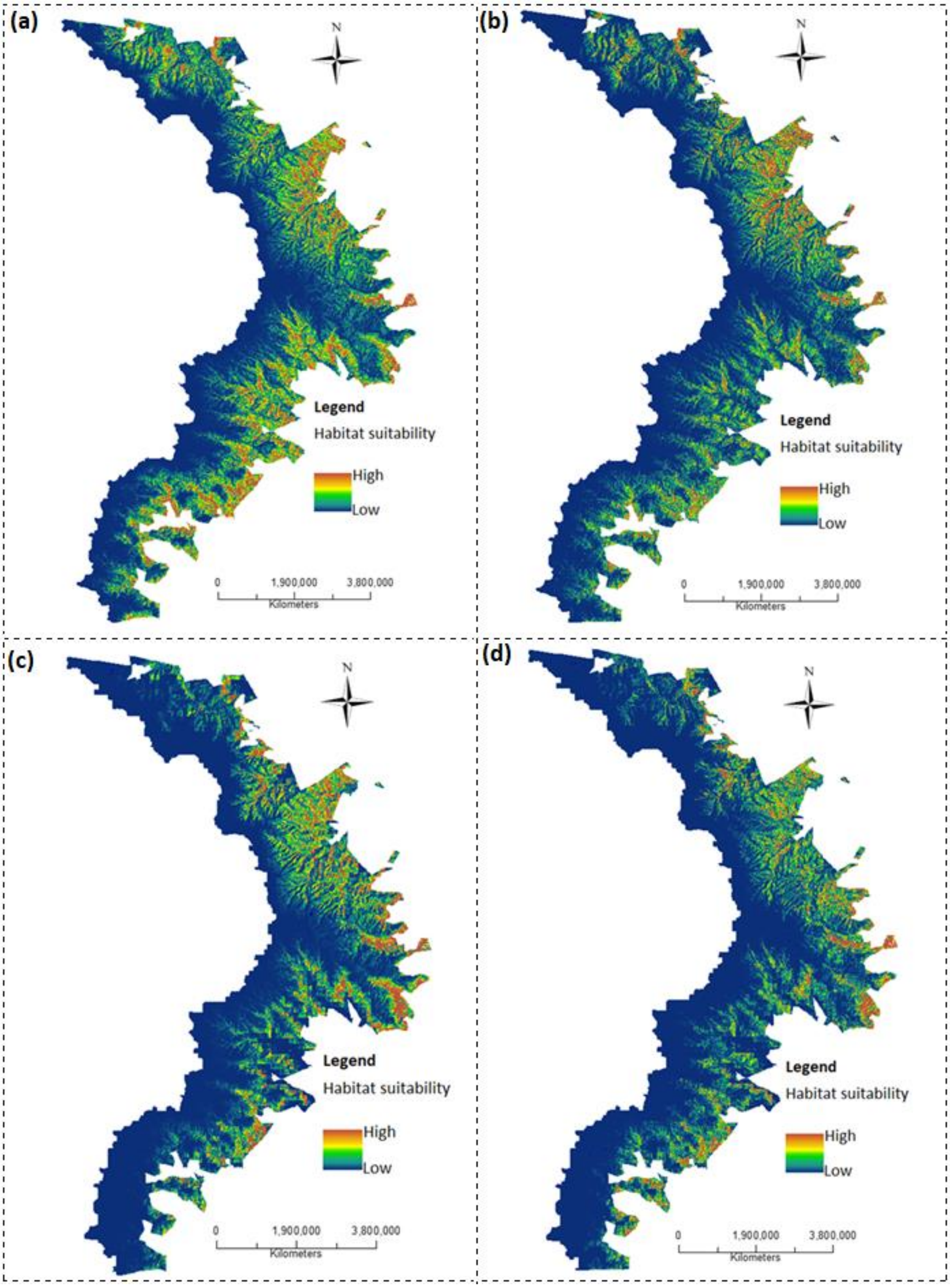

Figure 3.5: Maxent 'sub-models' results showing potential distribution of bramble using different input environmental variables. (a) Topographic variables (model 1), (b) topographic and Sentinel 2 imagery (model 2), (c) topographic and bioclimatic variable

3.4 Discussion

This study sought to model the spatial distribution of bramble invasion in the uKhahlamba Drakensberg Park (UDP), KwaZulu-Natal, South Africa. In the study, influential variables in estimating the potential distribution were determined and the influence of variable selection on model performance tested. Despite minor variances, results indicated better than random model prediction in all models (AUC values exceeding 0.75). These results are consistent with Phillips and Dudík (2008). In modelling probable bramble distribution, the model developed using a composite of all variables best estimated bramble distribution with a robust performance compared to other models. This supports previous studies that have also demonstrated that models built using a combination of variables were more reliable than those based solely on topographic or climatic predictors (Buermann *et al.* 2008; Saatchi *et al.* 2008; Parviainen *et al.* 2013). This finding demonstrates that a large number of factors enlarge the spectrum of environmental variables that define the spatial distribution of species. When variables are combined, they operate as complimentary predictors to disentangle distinct areas of absence, thus increasing model specificity (Parra *et al.* 2004; Cord and Rödder 2011). The addition of remotely sensed data improved model accuracy when compared to the model built solely on topographic variables. The results showed that differences derived from remotely sensed spectral information may be valuable in improving the calibration of SDMs (Morán-Ordóñez *et al.* 2012; Truong *et al.* 2017). The species-specific information provided by spectral reflectance aids the definition of the current distribution of its habitats (Pearson *et al.* 2004; Zimmermann *et al.* 2007). In line with the study by Ficetola *et al.* (2007), the integration of local bioclimatic data also increased model performance. The results of this study demonstrate that model complexity and the choice of variables used are critical since the accuracy of models varied according to the number of variables included.

Good model performance is an indication that key environmental variables associated with species suitability have been successfully identified for analysis (Evangelista *et al.* 2009; He *et al.* 2015). In this study, elevation and precipitation of the driest quarter were identified as the strongest predictors of bramble distribution for model scenarios 1 and 2, 3 and 4, respectively. Although elevation has no direct biological effect on plants, it is considered an important factor influencing patterns of vegetation distribution in mountainous areas due to its correlation with temperature and precipitation (Zhu *et al.* 2007; Mokarram and Sathyamoorthy 2015; Wang *et al.* 2017). The models showed high probabilities of bramble presence in areas with rainfall between 450-490 mm ($p > 0.5$) during the driest quarter. The results are in line with observations of moist warm areas being advantageous for bramble occurrence (Henderson 2011). Bramble distribution reflected highly on Bands 8a, 4 and 7, indicating the red edge, Near Infrared (NIR) and red band as the most optimal wavelengths for bramble discrimination. The sensitivity of bramble presence to the red edge and NIR bands was expected as the red edge is sensitive to chlorophyll status and vegetation reflects highly on the NIR band (Dlamini 2010; Addabbo *et al.* 2016). These results can be invaluable in informing the best optimal band combinations for mapping the ecological properties for bramble. TPI, TWI and Band 3 were poor predictors of bramble occurrence.

The models depicted a wider distribution than presently known, indicating potentially suitable habitats for bramble in new areas which have not yet been colonized. An overlay analysis of the resultant distribution and explanatory datasets revealed that the persistence of bramble in the eastern region of the study area is due to topographic gradients which further influenced the resultant climate (Pauchard and Alaback 2004; Haider *et al.* 2010). This study showed that occurrence of bramble was mainly in areas ranging from 1600 to 1850 m ($p > 0.5$) above sea level within the park. The preference of bramble in these areas reflects their range, with most of their known localities occurring at elevation between 1600 m and 1800 m, cool and moist eastern parts of South African grasslands (Henderson 2011). Bramble was predicted absent on the western boundary of the park, which corresponds with high and steep elevation gradients. This may represent a particular threshold for bramble to adapt in these areas. The preference of bramble in lower elevations may also represent early invasion stages wherein it could still slowly adapt to environments of steeper slopes and higher elevations. Areas of mean annual temperatures between 13^{o}C to 15^{o}C were a characteristic of areas suitable for the probability of bramble presence ($p > 0.5$). Interestingly, bramble occurrence was also observed in the area with low temperatures such as in Kamberg and Cobham. This confirms the Weed Management Guide (2003) observations that although bramble persisted in areas of temperate climates of warm summers and cool winters in Australia, they could also survive in lower rainfall areas when other environmental conditions are favourable.

Managing invasive species presents major economic and ecological challenges. Although fine-scale surveys provide insights on invasive plant species location and abundance, such approaches are not feasible for large regions (Lemke *et al.* 2011). The Maxent modelling approach has proven to be a simple and effective approach for predicting the potential distribution of bramble and its ecological requirements. In line with related studies such as Matawa *et al.* (2012) and Chikerema *et al.* (2013), such approaches have become valuable in landscape management that includes mapping species geographical distributions, understanding species tolerances or geographical ranges and identifying areas vulnerable to invasion. Bramble probable occurrence distribution maps can inform monitoring and management efforts such as burning, spraying and physical removal. Findings of this study also showed that combining complimentary predictors of topographic, climatic and remotely sensed data is a promising modelling strategy that improves prediction of species invasion range. SDMs like the one developed in this research can be used to inform landscape management and long-term monitoring programs.

3.5 Conclusion

This study demonstrated the value of the Maxent modelling approach in determining bramble distribution and identified factors influencing landscape vulnerability to bramble invasion. Based on the Maxent modelling technique our results conclude that:

- The accuracy of the models is better than random prediction with AUC scores >0.8. The inclusion of bramble spectral signature improved model accuracy. Models built using a composite of all variables performed better than those built solely on topographic or climatic data alone.
- Elevation and rainfall driest quarter were the most dominant variables with regards to estimating the distribution of bramble. The higher spectral resolution of Sentinel 2 MSI had an influence on model accuracy with the bramble reflecting high on the red edge and NIR bands.
- Mid-elevation, cool and moist eastern parts of the Drakensberg uKhahlamba region were more suitable for bramble infestation while cold and dry conditions at higher western elevation could be responsible for the spatial constriction.
- Such models can be integrated into conservation monitoring and management programs such as in the UDP. In this regard, distribution maps could aid in conservation planning as they can provide an indication of infested or areas vulnerable infestation, critical for bramble invasion management.

Chapter Four: Objectives reviewed and conclusions

4.1 Introduction

The focus of this study was to explore the application of SDMs in modelling the potential distribution of invasive American bramble at the Ukhahlamba Drakensberg Park, South Africa. This chapter evaluates the objectives presented in Chapter one against findings. Furthermore, the chapter highlights major conclusions, limitations and recommendations for future research.

4.2 Objectives reviewed

Reviewing the applications of SDMs in predicting the spatial distribution of invasive alien plants

The study reviewed the application of SDMs in modelling in predicting the distribution of IAPs across landscapes and at different spatial scales. SDMs have become a central tool yielding insights of current and future IAP distributions, invasion driving processes and spread. However, their suitability in IAPs is still challenged on the basis that SDMs are static in nature while the behaviour and effects of IAPs vary in space and time with environmental conditions. For example, IAPs by definition are not at equilibrium with their environment. Although the uncertainty that arises cannot be eliminated entirely, modelling frameworks have been developed and demonstrated to provide reliable results in modelling IAP distributions. Whereas traditional modelling approaches for predicting the distribution of IAPs were focused on identifying species' climatic space, a wide range of variables believed to influence IAP distributions are now encompassed. Presence-only methods such as Maxent have been found to have superior stability and prediction performance as they incorporate improved mathematical modelling methods compared to earlier SDMs. The ensemble modelling approach has proven to provide consensus predictions through combining the strengths of several distinct models. Recent studies have also incorporated remote sensing data providing predictor variables beyond climate conditions and enhancing model predictions. Therefore, adaptive practices have afforded the opportunities to provide timely information about areas of current and future invasion for informing effective management and monitoring strategies.

Modelling potential distribution of Bramble (*Rubus cuneifolius*) using topographic, remotely sensed and bioclimatic data in the KwaZulu-Natal Drakensberg, South Africa

The rapid proliferation of IAPs such as bramble has resulted in significant adverse impacts on native ecosystems. This study explored the Maxent technique to model the potential distribution of bramble in relation to environmental variables. The results showed that variable selection is an important step in SDMs as accuracy values varied according to the predictors used for building the model. As hypothesized, the inclusion of remotely sensed data improved model accuracy with the red edge and NIR bands being optimal band for discriminating bramble. Topographic variables elevation and slope had influence on bramble distribution with its probability of occurrence ($p > 0.5$) characterized by areas of elevation ranging from 1600m to 1800m asml and flatter angles of slope. The influence of these variables is believed to be as a result of their correlation with temperature and precipitation in the Ukhahlamba Drakensberg Park. This correlation explained the observed distribution of bramble being constricted to the eastern boundary of the park characterized by sufficient rainfall and higher temperatures compared to the western boundary. Our study demonstrated the ability of Maxent to integrate different geospatial data types for modelling and enhancing prediction efforts. Furthermore, this study provides a basis for identifying areas where management efforts should be focused.

4.3 Limitations and recommendations

- The strength of SDMs is determined, in part, by sample size and their spatial distribution. Although Maxent has been very effective in modelling the distribution of IAPs including with small sample sizes, further attempts should be made to collect more and well-distributed observation points.
- The choice of predictor variables used to model species distributions is influenced by data availability rather than species ecology or scale of study (Manzoor *et al*. 2018). For example, a lot of studies rely on bioclimatic variables that are available at a minimum of 1 km spatial resolution. Such large grids can result in spatial autocorrelation and bias the area identified as suitable when considering the scale of location points. As a result, this study used predictor variables characterized by a coarse resolution of 30 m. However, the 30 m was also coarse for the scale of the study thus resulted in spatial autocorrelation between sampled points. Hence, Future studies

should consider finer scale resolution of predictor variables and address spatial uncertainty such as avoiding multiple occurrence points within a single grid.

- Although river proximity is a crucial variable as bramble has also been observed to invade stream banks, its application in this study was challenged by the topographic diversity in the UDP. For example, areas 20 m away from a river may be completely different in terms of moisture, whereas other areas (flatter) can be similar 200 m away. Therefore, variables such as the Topographic Wetness Index (TWI) could be used as a proxy for moisture.
- As it was anticipated that moisture areas would be an important factor for bramble distribution, the models did not capture this for both river proximity and the TWI. This may be due to the TWI not being hydrologically corrected. Therefore, future studies should use a hydrologically corrected DEM for calculating TWI in mountainous.
- Bioclimatic data are interpolated from station data and therefore may not capture the local understanding of how rainfall decreases near the top of the mountains as most climate models assume a positive correlation with altitude/elevation. This could be misleading in this analysis. Therefore, future studies should explore remotely sensed climate data such as MODIS land surface temperature as they have demonstrated to be superior to ground station interpolations.
- The results of this study should be part of conservation planning for the UDP. Furthermore, it should form as a basis for an iterative process of constantly collecting new field information and performing new model runs to track and monitor spread patterns of bramble. This would enable conservation managers to prioritize control and eradication strategies.

4.4 Conclusions

Spatial Distribution Models (SDM) have become important and powerful tools in predicting IAPs distributions and ecological niches. In this study, Maxent was used to generate model scenarios of the potential distribution of American bramble in the Ukhahlamba Drakensberg Park (UDP) using topographic, bioclimatic and remotely sensed predictor variables. Although all models had better than random prediction, the strength of model predictions varied with the use of different variables. The inclusion of remotely sensed data as ancillary variables improved the predictive performance of Maxent and the composite model yielded the highest

AUC score. The predicted distribution of American bramble is far larger than its current geographic extent and models suggested bramble as a species characteristic of low elevation, warm and moist areas. The study demonstrated the usefulness of Maxent for predicting the distribution of American bramble and its ecological requirements. Potential distribution maps such as developed for this study can assist in informing control interventions for bramble invasion in the UDP.

References

Addabbo, P., Focareta, M., Marcuccio, S., Votto, C. & Ullo, S. L. 2016. Contribution of Sentinel-2 data for applications in vegetation monitoring. *Acta IMEKO,* 5(2).

Adhikari, D., Tiwary, R. & Barik, S. K. 2015. Modelling hotspots for invasive alien plants in India. *PloS one,* 10(7)**,** e0134665.

Allouche, O., Tsoar, A. & Kadmon, R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology,* 43(6)**,** 1223-1232.

Andrew, M. E. & Ustin, S. L. 2009. Habitat suitability modelling of an invasive plant with advanced remote sensing data. *Diversity and Distributions,* 15(4)**,** 627-640.

Araújo, M. B. & Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography,* 16(6)**,** 743-753.

Austin, M. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological modelling,* 157(2)**,** 101-118.

Austin, M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological modelling,* 200(1)**,** 1-19.

Austin, M. P. & Van Niel, K. P. 2011. Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography,* 38(1)**,** 1-8.

Barbosa, F. G., Schneck, F. & Melo, A. S. 2012. Use of ecological niche models to predict the distribution of invasive species: a scientometric analysis. *Brazilian Journal of Biology,* 72(4)**,** 821-829.

Beaumont, L. J., Gallagher, R. V., Thuiller, W., Downey, P. O., Leishman, M. R. & Hughes, L. 2009. Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions,* 15(3)**,** 409-420.

Benedetti, Y. & Morelli, F. 2017. Spatial mismatch analysis among hotspots of alien plant species, road and railway networks in Germany and Austria. *PloS one,* 12(8)**,** e0183691.

Bishop, T. R., Robertson, M. P., Rensburg, B. J. & Parr, C. L. 2014. Elevation–diversity patterns through space and time: ant communities of the Maloti-Drakensberg Mountains of southern Africa. *Journal of Biogeography,* 41(12)**,** 2256-2268.

Bjarnason, A., Katsanevakis, S., Galanidis, A., Vogiatzakis, I. N. & Moustakas, A. 2017. Evaluating Hypotheses of Plant Species Invasions on Mediterranean Islands: Inverse Patterns between Alien and Endemic Species. *Frontiers in Ecology and Evolution,* 5**,** 91.

Booth, T. H., Nix, H. A., Busby, J. R. & Hutchinson, M. F. 2014. BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Diversity and Distributions,* 20(1)**,** 1-9.

Bradley, B. A. 2014. The Biogeography of Invasive Plants–Projecting Range Shifts with Climate Change. *Invasive Species and Global Climate Change,* 4**,** 240.

Buermann, W., Saatchi, S., Smith, T. B., Zutta, B. R., Chaves, J. A., Milá, B. & Graham, C. H. 2008. Predicting species distributions across the Amazonian and Andean regions using remote sensing data. *Journal of Biogeography,* 35(7)**,** 1160-1176.

Campos, J. A., García-Baquero, G., Caño, L., Biurrun, I., García-Mijangos, I., Loidi, J. & Herrera, M. 2016. Climate and Human Pressure Constraints Co-Explain Regional Plant Invasion at Different Spatial Scales. *PloS one,* 11(10)**,** e0164629.

Cao, B., Bai, C., Zhang, L., Li, G. & Mao, M. 2016. Modeling habitat distribution of Cornus officinalis with Maxent modeling and fuzzy logics in China. *Journal Of Plant Ecology,* 9(6)**,** 742-751.

Cheng, D. & Xu, L. 2015. Predicting the potential distributions of Senecio vulgaris L. in China. *PeerJ PrePrints,* 3**,** e1612v1.

Chikerema, S., Murwira, A., Matope, G. & Pfukenyi, D. 2013. Spatial modelling of Bacillus anthracis ecological niche in Zimbabwe. *Preventive veterinary medicine,* 111(1)**,** 25-30.

Chiou, C.-R., Wang, H.-H., Chen, Y.-J., Grant, W. E. & Lu, M.-L. 2013. Modeling potential range expansion of the invasive shrub Leucaena leucocephala in the Hengchun peninsula, Taiwan. *Invasive Plant Science and Management,* 6(4)**,** 492-501.

Collingham, Y. C., Wadsworth, R. A., Huntley, B. & Hulme, P. E. 2000. Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. *Journal of Applied Ecology,* 37(s1)**,** 13-27.

Cord, A. & Rödder, D. 2011. Inclusion of habitat availability in species distribution models through multi-temporal remote-sensing data? *Ecological Applications,* 21(8)**,** 3285-3298.

Cord, A. F., Klein, D. & Dech, S. 2010. Remote sensing time series for modeling invasive species distribution: a case study of Tamarix spp. in the US and Mexico.

Cord, A. F., Klein, D., Gernandt, D. S., la Rosa, J. A. P. & Dech, S. 2014. Remote sensing data can improve predictions of species richness by stacked species distribution models: a case study for Mexican pines. *Journal of biogeography,* 41(4)**,** 736-748.

Cunze, S., Leiblein, M. C. & Tackenberg, O. 2013. Range expansion of Ambrosia artemisiifolia in Europe is promoted by climate change. *ISRN Ecology,* 2013.

Curtis, C. A. & Bradley, B. A. 2015. Climate change may alter both establishment and high abundance of red brome (Bromus rubens) and African mustard (Brassica tournefortii) in the semiarid southwest United States. *Invasive Plant Science and Management,* 8(3)**,** 341-352.

Dark, S. J. 2004. The biogeography of invasive alien plants in California: an application of GIS and spatial regression analysis. *Diversity and Distributions,* 10(1)**,** 1-9.

Deblauwe, V., Droissart, V., Bose, R., Sonké, B., Blach-Overgaard, A., Svenning, J. C., Wieringa, J., Ramesh, B., Stévart, T. & Couvreur, T. 2016. Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. *Global Ecology and Biogeography,* 25(4)**,** 443-454.

Dimitrakopoulos, P. G., Koukoulas, S., Galanidis, A., Delipetrou, P., Gounaridis, D., Touloumi, K. & Arianoutsou, M. 2017. Factors shaping alien plant species richness spatial patterns across Natura 2000 Special Areas of Conservation of Greece. *Science of The Total Environment,* 601**,** 461-468.

Dlamini, W. 2010. Multispectral detection of invasive alien plants from very high resolution 8-band satellite imagery using probabilistic graphical models. *Digital Globe,* 8**,** 1-17.

Duan, R.-Y., Kong, X.-Q., Huang, M.-Y., Fan, W.-Y. & Wang, Z.-G. 2014. The predictive performance and stability of six species distribution models. *PloS one,* 9(11)**,** e112764.

Dukes, J. S. & Mooney, H. A. 1999. Does global change increase the success of biological invaders? *Trends in Ecology & Evolution,* 14(4)**,** 135-139.

Duque-Lazo, J., Van Gils, H., Groen, T. & Navarro-Cerrillo, R. 2016. Transferability of species distribution models: The case of Phytophthora cinnamomi in Southwest Spain and Southwest Australia. *Ecological Modelling,* 320**,** 62-70.

Elith, J. 2015. Predicting distributions of invasive species.

Elith, J., Graham, C., Anderson, R., Dudık, M. & Ferrier, S. 2006. Novel methods improve prediction of species' distribution models. *Ecography,* 32**,** 66-77.

Elith, J. & Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics,* 40(1)**,** 677.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E. & Yates, C. J. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and distributions,* 17(1)**,** 43-57.

Erasmus, D. 1984. A. 3 BRA M BLE.

Evangelista, P. H., Stohlgren, T. J., Morisette, J. T. & Kumar, S. 2009. Mapping invasive tamarisk (Tamarix): a comparison of single-scene and time-series analyses of remotely sensed data. *Remote Sensing,* 1(3)**,** 519-533.

Fandohan, A. B., Oduor, A. M., Sodé, A. I., Wu, L., Cuni-Sanchez, A., Assédé, E. & Gouwakinnou, G. N. 2015. Modeling vulnerability of protected areas to invasion by Chromolaena odorata under current and future climates. *Ecosystem Health and Sustainability,* 1(6)**,** 1-12.

Fernandes, R. F., Vicente, J. R., Georges, D., Alves, P., Thuiller, W. & Honrado, J. P. 2014. A novel downscaling approach to predict plant invasions and improve local conservation actions. *Biological invasions,* 16(12)**,** 2577-2590.

Fernández, M. & Hamilton, H. 2015. Ecological niche transferability using invasive species as a case study. *PloS one,* 10(3)**,** e0119891.

Ficetola, G. F., Thuiller, W. & Miaud, C. 2007. Prediction and validation of the potential global distribution of a problematic alien invasive species—the American bullfrog. *Diversity and Distributions,* 13(4)**,** 476-485.

Franklin, J. 2010. *Mapping species distributions: spatial inference and prediction,* Cambridge University Press.

Fraterrigo, J. M., Wagner, S. & Warren, R. J. 2014. Local-scale biotic interactions embedded in macroscale climate drivers suggest Eltonian noise hypothesis distribution patterns for an invasive grass. *Ecology letters,* 17(11)**,** 1447-1454.

Friedman, J. H. 1991. Multivariate adaptive regression splines. *The annals of statistics***,** 1-67.

Gallien, L., Douzet, R., Pratte, S., Zimmermann, N. E. & Thuiller, W. 2012. Invasive species distribution models–how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography,* 21(11)**,** 1126-1136.

Gallien, L., Mazel, F., Lavergne, S., Renaud, J., Douzet, R. & Thuiller, W. 2015. Contrasting the effects of environment, dispersal and biotic interactions to explain the distribution of invasive plants in alpine communities. *Biological invasions,* 17(5)**,** 1407-1423.

Gallien, L., Münkemüller, T., Albert, C. H., Boulangeat, I. & Thuiller, W. 2010. Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions,* 16(3)**,** 331-342.

Gassó, N., Sol, D., Pino, J., Dana, E. D., Lloret, F., Sanz-Elorza, M., Sobrino, E. & Vilà, M. 2009. Exploring species attributes and site characteristics to assess plant invasions in Spain. *Diversity and Distributions,* 15(1)**,** 50-58.

Gastón, A. & García-Viñas, J. I. 2011. Modelling species distributions with penalised logistic regressions: A comparison with maximum entropy models. *Ecological modelling,* 222(13)**,** 2037-2041.

Gooden, B., French, K., Turner, P. J. & Downey, P. O. 2009. Impact threshold for an alien plant invader, Lantana camara L., on native plant communities. *Biological conservation,* 142(11)**,** 2631-2641.

Goslee, S. C., Beck, K. G. & Peters, D. P. 2003. Distribution of Russian knapweed in Colorado: Climate and environmental factors. *Journal of range management***,** 206-212.

Guisan, A., Edwards, T. C. & Hastie, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological modelling,* 157(2)**,** 89-100.

Guisan, A. & Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology letters,* 8(9)**,** 993-1009.

Haider, S., Alexander, J., Dietz, H., Trepl, L., Edwards, P. J. & Kueffer, C. 2010. The role of bioclimatic origin, residence time and habitat context in shaping non-native plant distributions along an altitudinal gradient. *Biological Invasions,* 12(12)**,** 4003-4018.

Hansen, S. 2015. *How flower visitation of remnant grassland patches is affected by commercial timber plantations and an invasive alien species (Rubus cuneifolius).* Stellenbosch University.

Hansen, S., Roets, F., Seymour, C. L., Thébault, E., Veen, F. & Pryke, J. S. 2018. Alien plants have greater impact than habitat fragmentation on native insect flower visitation networks. *Diversity and Distributions,* 24(1)**,** 58-68.

He, K. S., Bradley, B. A., Cord, A. F., Rocchini, D., Tuanmu, M. N., Schmidtlein, S., Turner, W., Wegmann, M. & Pettorelli, N. 2015. Will remote sensing shape the next generation of species distribution models? *Remote Sensing in Ecology and Conservation,* 1(1)**,** 4-18.

Hejda, M., Pyšek, P. & Jarošík, V. 2009. Impact of invasive plants on the species richness, diversity and composition of invaded communities. *Journal of Ecology,* 97(3)**,** 393-403.

Henderson, L. 2007. Invasive, naturalized and casual alien plants in southern Africa: a summary based on the Southern African Plant Invaders Atlas (SAPIA). *Bothalia,* 37(2).

Henderson, L. 2011. Mapping of invasive alien plants: the contribution of the Southern African Plant Invaders Atlas (SAPIA) to biological weed control. *African Entomology,* 19(2)**,** 498-503.

Henderson, L. & Wilson, J. R. 2017. Changes in the composition and distribution of alien plants in South Africa: An update from the Southern African Plant Invaders Atlas. *Bothalia-African Biodiversity & Conservation,* 47(2)**,** 1-26.

Hernandez, P. A., Graham, C. H., Master, L. L. & Albert, D. L. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography,* 29(5)**,** 773-785.

Higgins, S. & Richardson, D. 1996. A review of models of alien plant spread. *Ecological modelling,* 87(1-3)**,** 249-265.

Higgins, S. I., Richardson, D. M., Cowling, R. M. & Trinder-Smith, T. H. 1999. Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. *Conservation biology,* 13(2)**,** 303-313.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology,* 25(15)**,** 1965-1978.

Hirzel, A. H. & Le Lay, G. 2008. Habitat suitability modelling and niche theory. *Journal of Applied Ecology,* 45(5)**,** 1372-1381.

Hœrlé, S. 2006. Rock temperatures as an indicator of weathering processes affecting rock art. *Earth Surface Processes and Landforms,* 31(3)**,** 383-389.

Hoffman, J. D., Narumalani, S., Mishra, D. R., Merani, P. & Wilson, R. G. 2008. Predicting potential occurrence and spread of invasive plant species along the North Platte River, Nebraska. *Invasive Plant Science and Management,* 1(4)**,** 359-367.

Honig, M., Cowling, R. & Richardson, D. 1992. The invasive potential of Australian banksias in South African fynbos: a comparison of the reproductive potential of Banksia ericifolia and Leucadendron laureolum. *Austral Ecology,* 17(3)**,** 305-314.

Hulme, P. E. 2012. Weed risk assessment: a way forward or a waste of time? *Journal of Applied Ecology,* 49(1)**,** 10-19.

Jarošík, V., Pyšek, P., Foxcroft, L. C., Richardson, D. M., Rouget, M. & MacFadyen, S. 2011. Predicting incursion of plant invaders into Kruger National Park, South Africa: the interplay of general drivers and species-specific factors. *PloS one,* 6(12)**,** e28711.

Jiménez-Valverde, A., Peterson, A., Soberón, J., Overton, J., Aragón, P. & Lobo, J. M. 2011. Use of niche models in invasive species risk assessments. *Biological Invasions,* 13(12)**,** 2785-2797.

Kgosiesele, E. 2010. *Predictive Distribution Modelling of Timon Iepida in Spain,* University of Twente Faculty of Geo-Information and Earth Observation (ITC).

Kozak, K. H., Graham, C. H. & Wiens, J. J. 2008. Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution,* 23(3)**,** 141-148.

Kriticos, D., Sutherst, R., Brown, J., Adkins, S. & Maywald, G. 2003. Climate change and the potential distribution of an invasive alien plant: Acacia nilotica ssp. indica in Australia. *Journal of Applied Ecology,* 40(1)**,** 111-124.

Krüger, S. & Crowson, J. 2004. South Africa's uKhahlamba Drakensberg Park World Heritage Site Celebrates 30 Years of Wilderness. *International Journal of Wilderness,* 10(2)**,** 43-46.

Kühn, I., Brandenburg, M. & Klotz, S. 2004. Why do alien plant species that reproduce in natural habitats occur more frequently? *Diversity and Distributions,* 10(5-6)**,** 417-425.

Le Maitre, D. C., Thuiller, W. & Schonegevel, L. 2008. Developing an approach to defining the potential distributions of invasive plant species: a case study of Hakea species in South Africa. *Global Ecology and Biogeography,* 17(5)**,** 569-584.

Lemke, D. & Brown, J. A. 2012. Habitat modeling of alien plant species at varying levels of occupancy. *Forests,* 3(3)**,** 799-817.

Lemke, D., Hulme, P., Brown, J. & Tadesse, W. 2011. Distribution modelling of Japanese honeysuckle (Lonicera japonica) invasion in the Cumberland Plateau and Mountain Region, USA. *Forest Ecology and Management,* 262(2)**,** 139-149.

Liu, J., Liang, S. C., Liu, F. H., Wang, R. Q. & Dong, M. 2005. Invasive alien plant species in China: regional distribution patterns. *Diversity and distributions,* 11(4)**,** 341-347.

Lloret, F., Médail, F., Brundu, G., Camarda, I., Moragues, E., Rita, J., Lambdon, P. & Hulme, P. E. 2005. Species attributes and invasion success by alien plants on Mediterranean islands. *Journal of Ecology,* 93(3)**,** 512-520.

Lobo, J. M., Jiménez-Valverde, A. & Real, R. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography,* 17(2)**,** 145-151.

Lonsdale, W. M. 1999. Global patterns of plant invasions and the concept of invasibility. *Ecology,* 80(5)**,** 1522-1536.

Makori, D. M., Fombong, A. T., Abdel-Rahman, E. M., Nkoba, K., Ongus, J., Irungu, J., Mosomtai, G., Makau, S., Mutanga, O. & Odindi, J. 2017. Predicting Spatial Distribution of Key Honeybee Pests in Kenya Using Remotely Sensed and Bioclimatic Variables: Key Honeybee Pests Distribution Models. *ISPRS International Journal of Geo-Information,* 6(3)**,** 66.

Malahlela, O. E., Cho, M. A. & Mutanga, O. 2015. Mapping the occurrence of Chromolaena odorata (L.) in subtropical forest gaps using environmental and remote sensing data. *Biological Invasions,* 17(7)**,** 2027-2042.

Manzoor, S. A., Griffiths, G. & Lukac, M. 2018. Species distribution model transferability and model grain size–finer may not always be better. *Scientific reports,* 8(1)**,** 7168.

Marini, L., Gaston, K. J., Prosser, F. & Hulme, P. E. 2009. Contrasting response of native and alien plant species richness to environmental energy and human impact along alpine elevation gradients. *Global Ecology and Biogeography,* 18(6)**,** 652-661.

Martins, J., Richardson, D. M., Henriques, R., Marchante, E., Marchante, H., Alves, P., Gaertner, M., Honrado, J. P. & Vicente, J. R. 2016. A multi-scale modelling framework to guide management of plant invasions in a transboundary context. *Forest Ecosystems,* 3(1)**,** 17.

Matawa, F., Murwira, A. & Schmidt, K. S. 2012. Explaining elephant (Loxodonta africana) and buffalo (Syncerus caffer) spatial distribution in the Zambezi Valley using maximum entropy modelling. *Ecological Modelling,* 242**,** 189-197.

Matongera, T. N., Mutanga, O., Dube, T. & Lottering, R. T. 2016. Detection and mapping of bracken fern weeds using multispectral remotely sensed data: a review of progress and challenges. *Geocarto International***,** 1-16.

Matongera, T. N., Mutanga, O., Dube, T. & Sibanda, M. 2017. Detection and mapping the spatial distribution of bracken fern weeds using the Landsat 8 OLI new generation sensor. *International Journal of Applied Earth Observation and Geoinformation,* 57**,** 93-103.

Mędrzycki, P., Jarzyna, I., Obidziński, A., Tokarska-Guzik, B., Sotek, Z., Pabjanek, P., Pytlarczyk, A. & Sachajdakiewicz, I. 2017. Simple yet effective: Historical proximity variables improve the species distribution models for invasive giant hogweed (Heracleum mantegazzianum sl) in Poland. *PloS one,* 12(9)**,** e0184677.

Miller, J. 2010. Species distribution modeling. *Geography Compass,* 4(6)**,** 490-509.

Mokarram, M. & Sathyamoorthy, D. 2015. Modeling the relationship between elevation, aspect and spatial distribution of vegetation in the Darab Mountain, Iran using remote sensing data. *Modeling Earth Systems and Environment,* 1(4)**,** 30.

Morán-Ordóñez, A., Suárez-Seoane, S., Elith, J., Calvo, L. & de Luis, E. 2012. Satellite surface reflectance improves habitat distribution mapping: a case study on heath and shrub formations in the Cantabrian Mountains (NW Spain). *Diversity and Distributions,* 18(6)**,** 588-602.

Morecroft, M. D. 2015. Invasive species and global climate change. *Frontiers of Biogeography,* 7(4).

Mujuni, N. 2014. Use of species distribution modeling (SDM) in predicting distribution of invasive alien plants: A case study of Impatiens glandulifera in Europe.

Nel, W. 2009. Rainfall trends in the KwaZulu-Natal Drakensberg region of South Africa during the twentieth century. *International Journal of Climatology,* 29(11)**,** 1634-1641.

Nel, W. & Sumner, P. 2006. Trends in rainfall total and variability (1970–2000) along the KwaZulu-Natal Drakensberg foothills. *South African Geographical Journal,* 88(2)**,** 130-137.

O'Donnell, M.S., & Ignizio, D.A. 2012. Bioclimatic predictors for supporting ecological applications in the conterminous United States: U.S. Geological Survey Data Series 691, 10.

Ohlemüller, R., Walker, S. & Bastow Wilson, J. 2006. Local vs regional factors as determinants of the invasibility of indigenous forest fragments by alien plant species. *Oikos,* 112(3)**,** 493-501.

Padalia, H., Srivastava, V. & Kushwaha, S. 2014. Modeling potential invasion range of alien invasive species, Hyptis suaveolens (L.) Poit. in India: Comparison of MaxEnt and GARP. *Ecological informatics,* 22**,** 36-43.

Parra, J. L., Graham, C. C. & Freile, J. F. 2004. Evaluating alternative data sets for ecological niche models of birds in the Andes. *Ecography,* 27(3)**,** 350-360.

Parviainen, M., Zimmermann, N. E., Heikkinen, R. K. & Luoto, M. 2013. Using unclassified continuous remote sensing data to improve distribution models of red-listed plant species. *Biodiversity and conservation,* 22(8)**,** 1731-1754.

Pauchard, A. & Alaback, P. B. 2004. Influence of elevation, land use, and landscape context on patterns of alien plant invasions along roadsides in protected areas of South-Central Chile. *Conservation Biology,* 18(1)**,** 238-248.

Pearson, R. G., Dawson, T. P. & Liu, C. 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography,* 27(3)**,** 285-298.

Pearson, R. G., Raxworthy, C. J., Nakamura, M. & Townsend Peterson, A. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of biogeography,* 34(1)**,** 102-117.

Peterson, A. T. 2001. Predicting SPECIES'Geographic Distributions Based on Ecological Niche Modeling. *The Condor,* 103(3)**,** 599-605.

Peterson, A. T., Papes, M. & Kluza, D. A. 2003. Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science,* 51(6)**,** 863-868.

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E. & Blair, M. E. 2017. Opening the black box: an open-source release of Maxent. *Ecography,* 40(7)**,** 887-893.

Phillips, S. J., Anderson, R. P. & Schapire, R. E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological modelling,* 190(3)**,** 231-259.

Phillips, S. J. & Dudík, M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography,* 31(2)**,** 161-175.

Phillips, S. J., Dudík, M. & Schapire, R. E. Year: Published. A maximum entropy approach to species distribution modeling. *Proceedings of the twenty-first international conference on Machine learning*. ACM, 83.

Poona, N. 2008. Invasive alien plant species in South Africa: Impacts and management options. *Alternation,* 15(1)**,** 160-179.

Pyšek, P., Jarošík, V., Hulme, P. E., Pergl, J., Hejda, M., Schaffner, U. & Vilà, M. 2012. A global assessment of invasive plant impacts on resident species, communities and ecosystems: the interaction of impact measures, invading species' traits and environment. *Global Change Biology,* 18(5)**,** 1725-1737.

Raduła, M. W., Szymura, T. H. & Szymura, M. 2018. Topographic wetness index explains soil moisture better than bioindication with Ellenberg's indicator values. *Ecological Indicators,* 85**,** 172-179.

Ramírez-Albores, J. E., Bustamante, R. O. & Badano, E. I. 2016. Improved predictions of the geographic distribution of invasive plants using climatic niche models. *PloS one,* 11(5)**,** e0156029.

Richardson, D. M. & Van Wilgen, B. W. 2004. Invasive alien plants in South Africa: how well do we understand the ecological impacts?: working for water. *South African Journal of Science,* 100(1-2)**,** 45-52.

Robertson, M. P., Villet, M. H. & Palmer, A. R. 2004. A fuzzy classification technique for predicting species' distributions: applications using invasive alien plants and indigenous insects. *Diversity and Distributions,* 10(5-6)**,** 461-474.

Rocchini, D., Andreo, V., Förster, M., Garzon-Lopez, C. X., Gutierrez, A. P., Gillespie, T. W., Hauffe, H. C., He, K. S., Kleinschmit, B. & Mairota, P. 2015. Potential of remote sensing to predict species invasions: A modelling perspective. *Progress in Physical Geography,* 39(3)**,** 283-309.

Rödder, D., Schmidtlein, S., Veith, M. & Lötters, S. 2009. Alien invasive slider turtle in unpredicted habitat: a matter of niche shift or of predictors studied? *PLoS one,* 4(11)**,** e7843.

Rodríguez, J. P., Brotons, L., Bustamante, J. & Seoane, J. 2007. The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions,* 13(3)**,** 243-251.

Ross, L. K., Ross, R. E., Stewart, H. A. & Howell, K. L. 2015. The influence of data resolution on predicted distribution and estimates of extent of current protection of three 'listed' deep-sea habitats. *PloS one,* 10(10)**,** e0140061.

Rouget, M., Richardson, D. M., Nel, J. L., Le Maitre, D. C., Egoh, B. & Mgidi, T. 2004. Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions,* 10(5-6)**,** 475-484.

Rushworth, I. 2016. Alien and invasive species control plan for the South African portion of the Maloti Drakensberg Park World Heritage site,. Ezemvelo KwaZulu-Natal Wildlife (EKZNW) (unpublished document), Pietermaritzburg.

Rushworth, I. 2012. uKhahlamba Drakensberg Park World Heritage Site: Integrated Management Plan. Version 1.0, Ezemvelo KZN Wildlife, Pietermaritzburg.

Saatchi, S., Buermann, W., Ter Steege, H., Mori, S. & Smith, T. B. 2008. Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sensing of Environment,* 112(5)**,** 2000-2017.

Shabani, F. & Kumar, L. 2014. Sensitivity analysis of CLIMEX parameters in modeling potential distribution of Phoenix dactylifera L. *PloS one,* 9(4)**,** e94867.

Shezi, M. I. Z. & Poona, M. N. K. UKZN/DAEA.

Sibanda, M., Mutanga, O. & Rouget, M. 2015. Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. *ISPRS Journal of Photogrammetry and Remote Sensing,* 110**,** 55-65.

Simpson, M. & Prots, B. 2013. Predicting the distribution of invasive plants in the Ukrainian Carpathians under climatic change and intensification of anthropogenic disturbances: implications for biodiversity conservation. *Environmental conservation,* 40(2)**,** 167-181.

Smolik, M., Dullinger, S., Essl, F., Kleinbauer, I., Leitner, M., Peterseil, J., Stadler, L. M. & Vogl, G. 2010. Integrating species distribution models and interacting particle systems to predict the spread of an invasive alien plant. *Journal of Biogeography,* 37(3)**,** 411-422.

Sosa, A., Cardo, M. & Julien, M. 2017. Predicting weed distribution at the regional scale in the native range: environmental determinants and biocontrol implications of Phyla nodiflora (Verbenaceae). *Weed Research,* 57(3)**,** 193-203.

Steyn, C., Greve, M., Robertson, M. P., Kalwij, J. M. & Le Roux, P. C. 2017. Alien plant species that invade high elevations are generalists: support for the directional ecological filtering hypothesis. *Journal of Vegetation Science,* 28(2)**,** 337-346.

Stickler, C. M. & Southworth, J. 2008. Application of multi-scale spatial and spectral analysis for predicting primate occurrence and habitat associations in Kibale National Park, Uganda. *Remote sensing of environment,* 112(5)**,** 2170-2186.

Stohlgren, T. J., Ma, P., Kumar, S., Rocca, M., Morisette, J. T., Jarnevich, C. S. & Benson, N. 2010. Ensemble habitat mapping of invasive plant species. *Risk Analysis,* 30(2)**,** 224-235.

Sundaram, B. & Hiremath, A. J. 2012. Lantana camara invasion in a heterogeneous landscape: patterns of spread and correlation with changes in native vegetation. *Biological Invasions,* 14(6)**,** 1127-1141.

Taylor, S. & Kumar, L. 2013. Potential distribution of an invasive species under climate change scenarios using CLIMEX and soil drainage: A case study of Lantana camara L. in Queensland, Australia. *Journal of environmental management,* 114**,** 414-422.

Taylor, S., Kumar, L., Reid, N. & Kriticos, D. J. 2012. Climate change and the potential distribution of an invasive shrub, Lantana camara L. *PLoS One,* 7(4)**,** e35565.

Thuiller, W., Araujo, M. B. & Lavorel, S. 2004a. Do we need land-cover data to model species distributions in Europe? *Journal of Biogeography,* 31(3)**,** 353-361.

Thuiller, W., Brotons, L., Araújo, M. B. & Lavorel, S. 2004b. Effects of restricting environmental range of data to project current and future species distributions. *Ecography,* 27(2)**,** 165-172.

Thuiller, W., Richardson, D. M., PYŠEK, P., Midgley, G. F., Hughes, G. O. & Rouget, M. 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology,* 11(12)**,** 2234-2250.

Trethowan, P., Robertson, M. & McConnachie, A. 2011. Ecological niche modelling of an invasive alien plant and its potential biological control agents. *South African Journal of Botany,* 77(1)**,** 137-146.

Truong, T. T., Hardy, G. E. S. J. & Andrew, M. E. 2017. Contemporary remotely sensed data products refine invasive plants risk mapping in data poor regions. *Frontiers in plant science,* 8.

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D. & Kadmon, R. 2007. A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and distributions,* 13(4)**,** 397-405.

Uden, D. R., Allen, C. R., Angeler, D. G., Corral, L. & Fricke, K. A. 2015. Adaptive invasive species distribution models: a framework for modeling incipient invasions. *Biological invasions,* 17(10)**,** 2831-2850.

Václavík, T. & Meentemeyer, R. K. 2012. Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions,* 18(1)**,** 73-83.

Van Wilgen, B., Richardson, D., Le Maitre, D., Marais, C. & Magadlela, D. 2001. The economic consequences of alien plant invasions: examples of impacts and approaches to sustainable management in South Africa. *Environment, Development and Sustainability,* 3(2)**,** 145-168.

van Wilgen, B. W., Reyers, B., Le Maitre, D., Richardson, D. & Schonegevel, L. 2008. A biome-scale assessment of the impact of invasive alien plants on ecosystem services in South Africa. *Journal of Environmental Management,* 89(4)**,** 336-349.

Vanderhoof, M., Holzman, B. A. & Rogers, C. 2009. Predicting the distribution of perennial pepperweed (Lepidium latifolium), San Francisco Bay area, California.

Verbruggen, H., Tyberghein, L., Belton, G. S., Mineur, F., Jueterbock, A., Hoarau, G., Gurgel, C. F. D. & De Clerck, O. 2013. Improving transferability of introduced species' distribution models: new tools to forecast the spread of a highly invasive seaweed. *PLoS One,* 8(6)**,** e68337.

Vilà, M., Espinar, J. L., Hejda, M., Hulme, P. E., Jarošík, V., Maron, J. L., Pergl, J., Schaffner, U., Sun, Y. & Pyšek, P. 2011. Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology letters,* 14(7)**,** 702-708.

Wakie, T., Evangelista, P., Jarnevich, C. & Laituri, M. 2014. Mapping Current and Potential Distribution of Non-Native Prosopis juliflora in the Afar Region of.

Wan, J. Z., Wang, C. J., Tan, J. F. & Yu, F. H. 2017. Climatic niche divergence and habitat suitability of eight alien invasive weeds in China under climate change. *Ecology and evolution,* 7(5)**,** 1541-1552.

Wang, A., Goslee, S. C., Miller, D. A., Sanderson, M. A. & Gonet, J. M. 2017. Topographic variables improve climatic models of forage species abundance in the northeastern United States. *Applied Vegetation Science,* 20(1)**,** 84-93.

Wang, Y. & Xu, Z. 2016. Where are the Alien Species? Predictions of Global Plant Species Invasions under Current Environmental Conditions and the Human Footprint. *Polish Journal of Environmental Studies,* 25(4).

Weed Management Guide., 2003. Blackberry- Rubus fruticoses aggregates [online]. Weeds of national significance. Available from: http://www.environment.gov.au/biodiversity/invasive/weeds/publications/guidelines/won s/r-fruticosus.html. [Accessed 28 June 2017].

Wenger, S. J. & Olden, J. D. 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution,* 3(2)**,** 260-267.

Wisz, M. S., Hijmans, R., Li, J., Peterson, A. T., Graham, C. & Guisan, A. 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions,* 14(5)**,** 763-773.

Young, N. E., Stohlgren, T. J., Evangelista, P. H., Kumar, S., Graham, J. & Newman, G. 2012. Regional data refine local predictions: modeling the distribution of plant species abundance on a portion of the central plains. *Environmental monitoring and assessment,* 184(9)**,** 5439-5451.

Zhu, L., Sun, O. J., Sang, W., Li, Z. & Ma, K. 2007. Predicting the spatial distribution of an invasive plant species (Eupatorium adenophorum) in China. *Landscape Ecology,* 22(8)**,** 1143-1154.

Zimmermann, N., Edwards, T., Moisen, G., Frescino, T. & Blackard, J. 2007. Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of applied ecology,* 44(5)**,** 1057-1067.

Zimmermann, N. E., Edwards, T. C., Graham, C. H., Pearman, P. B. & Svenning, J. C. 2010. New trends in species distribution modelling. *Ecography,* 33(6)**,** 985-989.