Detecting and mapping the habitat suitability of the Cossid Moth, (*Coryphodema tristis)* on *Eucalyptus nitens* in Mpumalanga, South Africa

Samuel Takudzwa Kumbula

213568262

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.

Supervisor: Dr. Romano Lottering

Co-supervisor: Prof Paramu Mafongoya and Dr. Kabir Peerbhay

Pietermaritzburg, South Africa

October 2018

Abstract

Cossid moth *(Coryphodema tristis)* is an indigenous wood-boring insect that presents serious environmental, ecological and economic problems globally. An extensive analysis of the current spatial distribution of *Coryphodema tristis* is therefore essential for providing applicable management approaches at both local and regional scales. This aim of the study was to assess GIS and remote sensing applications combined with species distribution models (Maxent) to monitor habitat suitability of the *Coryphodema tristis* in Mpumalanga, South Africa. The first objective of the study focused on comparing the robustness of species distribution models using Maxent (presence-data only) and Logistic regression (presence-absence data) in characterizing the habitat suitability of the *Coryphodema tristis*. The second objective of the study evaluated the effectiveness of the freely available Sentinel 2 multispectral imagery in detecting and mapping the habitat suitability of the *C. tristis*. The models sought to identify the factors that can be used to predict habitat suitability for the *C. tristis* using environmental and climatic variables. Presence and absence records were collected through systematic surveys of forest plantations. The models were applied on *Eucalyptus nitens* plantations of the study area for habitat preferences. The overall accuracies indicated that Maxent (AUC = 0.84 and 0.810) was more robust than the Logistic regression model (AUC= 0.745 and 0.677) using training and testing datasets, respectively. In Maxent, the jackknife indicated that mean temperature for October, aspect, age, mean temperature for February, June, December and elevation as the most influential predictor variables. Meanwhile, age was the only significant variable in the Logistic regression model. Therefore, results concluded that temperature, aspect, age and elevation were optimal in modelling habitat suitability for the *Coryphodema tristis*.

For the second objective, model performance was evaluated using the Receiver Operating Characteristics (ROC) curve showing the Area Under the Curve (AUC) and True Skill Statistic (TSS), while the performance of predictors was displayed in the jackknife. Using only the occurrence data and Sentinel-2 bands and derived vegetation indices, the Maxent model provided successful results, exhibiting an area under curve (AUC) of 0.89. The Photosynthetic vigor ratio,

Red edge (705 nm), Red (665 nm), Green NDVI hyper, Green (560 nm) and Shortwave infrared (SWIR) (2190 nm) were identified as the most influential predictor variables. Results of this study suggests that remotely sensed derived vegetation indices from cost effective platforms could play a crucial role in supporting forest pest management strategies and infestation control. Overall, these results improve the assessment of temporal changes in habitat suitability of *Coryphodema tristis*, which is crucial in the management and control of these pests.

Keywords: Cossid moth, *Coryphodema tristis; Eucalyptus nitens* infestation, Sentinel 2, Environmental and climatic variables, Maxent model, Habitat suitability.

Preface

This study was conducted in the School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa, from February 2017 to October 2018 under the supervision of Dr. Romano Lottering, Prof Paramu Mafongoya and Kabir Peerbhay.

I declare that the work presented in this thesis has never been submitted in any form to any other institution. This work represents my original work except where due acknowledgements are made.

Samuel Takudzwa Kumbula: Signed …………………………. 　　　　　Date…………………

As the candidate's supervisor, I certify the aforementioned statement and have approved this thesis for submission.

Dr. Romano Lottering　　　　　Signed……………….. 　　　　　Date………………………..

Prof Paramu Mafongoya　　　　Signed……………….. 　　　　　Date………………………..

Dr. Kabir Peerbhay　　　　　　Signed……………….. 　　　　　Date………………………..

Declaration

I Samuel Takudzwa Kumbula, 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:……………………..   Date………………………..

Dedication

I dedicate this dissertation to my beloved family, for believing so greatly in me and in the potential that I have to achieve greatness. From day one you have had faith in me to travel this journey and now we have made it, I want to continue making you proud.

Acknowledgements

Special thanks goes to my family for their unwavering support, Mum and Dad you are the best and I'm forever grateful for all the strings you pulled ……"**Munondipasa manyemwe**". As for Edgar Mutape Chivunze aiwa babamudiki your graduation is coming the ball is still rolling, I'm still in the race. Babamudiki Finnet Kumbula "Oldman" thank you for your support you have been by my side all these journeys and I thank you. Above all thanks to my little brothers Finnet and Hugh Kumbula for your support gents we made it **"Mama we made it"**.

I would like to extend my gratitude to the University of KwaZulu-Natal, the School of Agricultural, Earth and Environmental Sciences and the Geography department. It has been a privilege to work and interact with all of you within the school at large, **Thank you**. My deepest gratitude goes to my supervisors, Dr. Romano Lottering, Dr. Kabir Peerbhay and Prof Paramu Mafongoya who spent day and night to ensure that this project smoothly sailed on till today. I have learnt a lot from all of you, it goes without saying "Your intellectual knowledge goes beyond what one can imagine". You have assembled me into a hardworking and disciplined researcher that focuses on the set targets and ensures that the job is done and I thank you for that. Special thanks to Prof Paramu Mafongoya as the NRF SARCHI for Rural Development and Agronomy for providing logistical support in everything much appreciated.

I would also like to extend my appreciation and special thanks to Dr. Mbulisi Sibanda for his commitment and rigorous efforts towards mentoring me during my studies. The long hours in the office were sometimes fun and at the same time unpleasant but you made sure that every day you came with new ways of thinking and motivation to keep me going and I have no words that can describe what I feel this moment, be blessed and carry on the good work.

I would also give special thanks to my friends and colleagues, Trylee "Chairman" Matongera, Rodney Muringai, Sivuyile Mkhulisi, Ngoni Chipendo, Lungile Pamela Madela, Shenell Sewell, Dr. Terrence Mushore, Mnqobi Mtshali, Dr. Sithabile Hlahla, Dr. Cletah Shoko, Charles Otunga, Mr. Donavon Devos for their encouragement during my study as well as for their unconditional psychological support.

Table of Contents

List of Tables

List of Figures

Chapter One

General Introduction

1.1 Introduction

Eucalyptus tree species are among the most planted trees in the world because of their economic value and rapid growth rate (Wingfield et al., 2008). In commercial forest plantations, *Eucalyptus* tree species have been widely grown and cultivated mainly for fuelwood, timber, pulp and paper (Swain and Gardner, 2003; Wingfield et al., 1996). Emerging insect pests and diseases have caused extensive damage to *Eucalyptus nitens* commercial forests threatening future sustainability of the forestry sector. *Coryphodema tristis* (commonly known as the Cossid moth) in particular, is one of the emerging pests that has adversely affected the growth and yield of *E. nitens* plantations (Adam et al., 2013; Boreham, 2006). Literature shows that the *C. tristis* was recorded on *E. nitens* plantations in South Africa in 2004. The *C. tristis* is native to South Africa and has been associated with grapevine, apple, quince and sugar pear trees (Bouwer et al., 2015). However, a sudden shift to infest *E. nitens* in South Africa has been observed and is associated with environmental conditions as well as the absence of natural enemies. *C. tristis* is an indigenous wood-boring insect that feeds on the bark of the *E. nitens* trees. It poses a major threat to the forestry sector as it affects the quality and quantity of the yield. During the initial stages, extensive tunneling in the sapwood and heartwood of the trees is observed. As infestation progresses, resin and sawdust from larval feeding will also be observed (Adam et al., 2013; Gebeyehu et al., 2005). The biology and impact of the *C. tristis* on *E. nitens* plantation forests will be further described in chapter two of this study. Currently, there is no biological agent to control the damage of the moth. Hence, understanding the spatial distribution as well as habitat suitability conditions of the *C. tristis* would be beneficial to the forestry stakeholders.

Species Distribution Models (SDM) have become increasingly important and widely used to determine the spatial distribution of forest pests (Michael and Warren 2009; Wisz et al. 2013). These methods use presence and absence or presence-only datasets to relate known locations of pests with the environmental conditions of the target area so as to estimate the response function, contribution of variables and use them to predict the potential spatial distribution of species (Matawa et al., 2016; Phillips et al., 2017; Yi et al., 2016). Several studies have challenged the

reliability of absence data in modelling forest pests, indicating that failure to observe does not necessarily signify absence (Baldwin, 2009; Phillips and Dudík, 2008). According to Elith et al., (2006) and Phillips et al., (2006) absence data are rarely available and costly to collect in traditional field surveys, especially in cases of rare or emerging species. In addition, traditional data collection methods are mostly time-consuming, labor-intensive and spatially restrictive resulting in subjective absence information (Ndlovu et al., 2018; Pause et al., 2016; Pietrzykowski et al., 2007). When comparing presence-only datasets and presence and absence datasets, presence-only datasets represent a convenient dataset that reduces the processing and handling costs (Sahragard and Ajorlo, 2018). Hence, presence-only datasets are appropriate for species distribution modelling, due to being readily available and cost-effective as compared to absence datasets. For example, several studies compared the prediction accuracy of Logistic regression, Maximum entropy, artificial neural network and other SDMs in the potential habitats of species. Maxent that uses presence-only data was found to be a robust algorithm among other SDM's (Elith et al., 2011; Phillips et al., 2017; Tarkesh and Jetschke, 2012).

Recently, integration of SDM's and GIS and remote sensing for assessing the sensitivity of data in detecting and mapping the habitat suitability of forest pests has become increasingly appealing (Kozak et al., 2008; Ndlovu et al., 2018). In South Africa, determining the vulnerability of *E. nitens* forests to the *C. tristis* has been currently conducted with traditional field surveys, climatic and topographic data only. Remote sensing plays a key role in the assessment and monitoring of forest health as well as condition of habitat suitability of plantation forest pests in real time (Ismail et al., 2007; Lottering et al., 2018; Oumar and Mutanga, 2013). Advances in multispectral remote sensing have improved the spatial and spectral capabilities of sensors even in the detection and mapping of forest plantation pests and diseases. Different studies utilizing spectral information such as wavebands, red edge bands and vegetation indices of multispectral sensors show that they have offered opportunities to enhance the capability of SDM's both spatially and temporally (Adelabu et al. 2013; Lottering and Mutanga 2016; Oumar and Mutanga 2011; Rullan-Silva et al. 2013). In addition, Light Detection and Ranging (LIDAR) has provided an extensive contribution to the monitoring of forest health (Lausch et al., 2017; Pause et al., 2016). For example, Müller and Brandl (2009) stated that derived predictor variables from LIDAR improved the modelling of the habitat suitability of forest beetles in Germany. On the other hand, Oumar and Mutanga (2013)

acknowledged that the addition of remotely sensed environmental predictors such as wavebands and vegetation indices improved the robustness of SDM's.

Due to the infestation outbreaks of the *C. tristis*, understanding the current and potential distribution of the *C. tristis* is essential for effective forestry management. Hence, the application of remote sensing would be beneficial to the forestry sector, because of its ability to cover large areas at a cheaper cost (Adelabu et al., 2014; Senf et al., 2017). The new generation of freely available multispectral sensors such as Sentinel-2 characterized by 13 spectral bands that cover the red edge region (Band 5, 6 and 7) acquired at 290 km orbital swath width, offers the potential to determine the habitat suitability of the *C. tristis* over a large landscape scale (Addabbo et al., 2016; Hawryło et al., 2018). The sensor is associated with a high revisit time of 5 days which provides an effective temporal resolution that can monitor forest plantation health. In addition, vegetation indices calculated from the Sentinel 2 wavebands are sensitive to vegetation health and have been widely used as predictor variables in mapping and monitoring of forest pests. Therefore, the current study aimed at assessing the application of remotely sensed data combined with SDMs in mapping the habitat suitability of the *C. tristis* in Mpumalanga, South Africa.

1.2 Aims and objectives

The overall purpose of the study was to model the potential habitat suitability of the Cossid moth (*Coryphodema tristis*)in Mpumalanga, South Africa. The following objectives were set:

➢ To evaluate the robustness of the Maxent approach in modelling the potential habitat suitability of the *C. tristis* on *E. nitens* using climatic, environmental and remotely sensed data in relation to the performance of Logistic regression.

➢ To understand the climatic and environmental variables that influence the suitability of the *C. tristis* on *E. nitens* plantation.

➢ To evaluate the effectiveness of the freely available Sentinel 2 multispectral imagery in detecting and mapping habitat suitability of the *C. tristis*.

1.3 Key research questions

➢ To what extent does the Maxent model successfully predict the potential habitats of the *C. tristis*?

➢ How can Maxent as a SDM identify the climatic and environmental variables that influence the suitability preference of the C. tristis on *E. nitens* plantation?

➢ How effectively does the freely available Sentinel 2 sensor detect and map the *C. tristis* habitat suitability?

1.4 Main hypothesis

➢ The integration of species distribution models and remotely sensed data has the potential to detect and map the spatial distribution of the *C. tristis* habitat suitability with acceptable accuracies.

1.5 General structure of the thesis

This thesis consists of four chapters. The first chapter is the general introduction that provides general background information on the subject at hand as well as the aim and objectives of the study. The two objectives of this thesis are presented in chapter two and three as standalone research papers that when combined answer the overarching aim of this study. The last chapter is the conclusion, which provides a synthesis of the overall research.

Chapter two assessed the habitat suitability of the *C. tristis* by comparing the robustness of two species distribution models (Maxent and Logistic regression) and testing the performance of remotely sensed data in modelling the suitability preference of the moth. In addition, it also investigates the climatic and environmental variables that contribute to the habitat preference of the *C. tristis*. Finally, the chapter highlights the advantages of presence-only datasets over presence-absence datasets in modelling and mapping the habitat suitability of the *C. tristis*.

Chapter three assessed the utility of the freely available Sentinel 2 multispectral instrument in detecting and mapping the habitat suitability of the *C. tristis*. The study tested the application of wavebands, red edge bands and vegetation indices in detecting and mapping the habitat suitability of the moth.

Modelling potential habitat suitability of *Coryphodema tristis* (Cossid moth) on *Eucalyptus nitens* plantations using Species Distribution Models

Abstract

The study sought to assess the robustness of species distribution models using Maxent (presence-data only) and Logistic regression (presence-absence data) algorithms to model the habitat suitability of *Coryphodema tristis*. The models were also used to identify climatic and environmental variables that can predict habitat suitability for the *C. tristis* in Mpumalanga, South Africa. Presence and absence records were collected through systematic surveys of forest plantations. Climatic and environmental variables included climate, topography and compartment-specific attributes. The overall accuracies indicated that Maxent (AUC = 0.840 and 0.810) was more robust than the Logistic regression model (AUC= 0.745 and 0.677) using training and testing data, respectively. In Maxent, the jackknife indicated that mean temperature for October, aspect, age, mean temperature for February, June, December and elevation were identified as the most influential predictor variables. Meanwhile, age was the only significant variable in the Logistic regression model. Therefore, results concluded that temperature, aspect, age and elevation were optimal in modelling habitat suitability for the *C. tristis*. Thus, these results improve the assessment of temporal changes in habitat suitability of *C. tristis*, which is crucial in the management and control of these pests.

Keywords: Cossid moth, climatic and environmental variables, Maxent model, Habitat preference.

2.1. Introduction

Coryphodema tristis (Lepidoptera: Cossidae), commonly known as the Cossid moth, is an indigenous wood-boring insect that has caused significant damage to commercial *Eucalyptus* plantations (Degufu *et al*. 2013). The native moth has suddenly been recorded in cold-tolerant areas of Mpumalanga that are prone to frost and snow, these conditions are conducive for *Eucalyptus nitens* plantations (Boreham 2006; Degefu et al. 2013). The *C. tristis* has recently shifted its hosts to *E. nitens* and this has been attributed to the absent or low numbers of natural enemies (Battisti and Larsson, 2015; Gebeyehu et al., 2005). *Eucalyptus* tree species have been widely grown and cultivated mainly for fuelwood, timber, pulp and paper (Swain and Gardner, 2003; Wingfield et al., 1996). Hence, the continuous infestation of *E. nitens* plantations reduces the quality and quantity produced by the commercial forest plantations, which may affect the gross

domestic product of the host country. Currently, the forest industry produces input raw materials for other sectors such as construction and textile. Therefore, the negative impact of pests and diseases reduces the income generated by the host country from commercial *E. nitens* plantations.

The biology of the *C. tristis* indicates that it takes between two or three years to complete its life cycle and its estimated that up to eighteen months is spent in its larval stage, which is the greater part of its life cycle (Adam et al. 2013; Bouwer et al. 2015). As a native species in South Africa, adult moths emerge from October to mid-December in the Western Cape province on fruit tree species of grapevine, apple, quince and sugar pear trees (Bouwer et al. 2015; Gebeyehu et al. 2005). Previous studies conducted in the Mpumalanga area indicated that occurrence times of the *C. tristis* are almost similar to those found in the Western cape (Adam et al. 2013; Boreham 2006). In July, signs and symptoms of larvae damage occurrence was recognized, this resulted in the development of the adult *C. tristis* that were seen between August to October (Gebeyehu et al. 2005). In addition, traditional field surveys reported the establishment of all stages in the month of October, which was seen by pupal cases protruding out on tree holes resembling existence of adult moths (Adam et al. 2013). The adult moths are rarely seen due to their dull colour and their short-lived duration, creating a challenge of identification of the moth (Ramanagouda et al. 2010). However, few studies regarding the habitat suitability of the *C. tristis* have been undertaken. Investigating and developing a habitat suitability model to estimate the spatial distribution, as well as habitat preferences of the *C. tristis* is crucial for the conservation of *E. nitens* plantations.

Over the years, several studies have been conducted to identify and understand how climatic and environmental variables influence the spatial distribution of pests. Climatic variables such as temperature and precipitation influence the development, reproduction, survival, geographic range and population size of insect pests (Jaworski and Hilszczański 2013; Petzoldt and Seaman 2006). A number of studies indicated that temperature changes have influenced warming in tropical areas and has resulted in tropical insects becoming sensitive to little changes (Biber-Freudenberger et al. 2016; Dillon et al. 2010). Change in temperature either negatively/positively impacts the surrounding conditions inducing pest's populations to either disperse, adapt or shift hosts (Deka et al. 2011). Different studies have highlighted that increased summer temperatures and shortened winter periods have resulted in rapid insect reproduction and faster growth (Kocmánková et al. 2009; Oumar and Mutanga 2013). Hence, changes in temperature reduce winter mortality and

increase the population size of pests which results in tree species becoming more vulnerable to infestation (Deka et al. 2011). Moisture (precipitation) availability and variability also contribute to the habitat preferences of pests as it affects insect pest predators, parasites, and diseases (Jaworski and Hilszczański 2013; Kutywayo et al. 2013). In addition, habitat preference is related to elevation gradients because without favorable matting, host foraging and ovipositional conditions the pest cannot reproduce and survive (Péré et al. 2013). Forest stakeholders such as agro foresters, ecologists and conservation practitioners need to understand the fundamental factors that shape species spatial distributions in order to develop effective management strategies (Meier et al. 2010). For that reason, we developed Species Distribution Models (SDM) as a function of location, climatic and environmental conditions.

SDMs model the geographic distributions of species using either presence and absence data or presence-only data (Michael and Warren 2009; Wisz et al. 2013). Corresponding mathematical environmental conditions and distribution data is utilized to estimate the suitable species habitat and projected onto the geographic area to determine the probability of habitat preferences (Elith and Leathwick 2009; Yi et al. 2016). To estimate suitable preferences, SDM's use true presences and true absences obtained either from traditional field surveys or georeferenced species records (Biber-Freudenberger et al. 2016; Wang et al. 2018). Several studies have indicated that it is very difficult to obtain absence data (Babar et al. 2012; Farzin et al. 2016; Michael and Warren 2009). According to Baldwin (2009), absence data is very difficult to verify because failure to observe the target species does not mean absence and this results in substantially biased species-habitat relationships. In addition, traditional data collection methods are mostly time-consuming, costly, labor-intensive and spatially restrictive hindering the collection of actual absence data (Pause et al. 2016; Pietrzykowski et al. 2007). To date, a number of models that use presence-absence and presence-only datasets such as Generalized Linear Model (GLM), Logistic regression, Genetic Algorithm for Rule-set Production (GARP), DOMAIN and Maxent have been used to predict species distributions. In recent comparative studies of these models, presence-only datasets have been identified as robust algorithms that can be used to optimally model the spatial distribution of species.

A previous study utilized the random forest species distribution model to map the presence or absence of *C. tristis* infestations on *E. nitens* forests in Mpumalanga (Adam et al. 2013). In their

endeavor, they only utilized climatic and topographic variables to determine the susceptibility of *E. nitens* forests to *C. tristis* infestations. Their study successfully identified four variables that included elevation, the maximum temperature for September and April as well as the median rainfall for April as influential to infestation of *E. nitens*. Previous application of climatic and environmental variables to evaluate species distributions has been commonly used. However, the recent integration of remotely sensed data into SDM has become increasingly appealing and considered to improve the performance of SDMs (Kozak et al. 2008; Ndlovu et al. 2018). Presence-only datasets have also proved to be cost-effective and statistically better for modelling species distribution as fewer costs are associated in collecting and processing the data in the field. As a result, this study selected Maxent because of its various advantages: (1) The input species data can be presence-only data; (2) both continuous and categorical data can be used as input variables; (3) its prediction accuracy is always stable and reliable, even with incomplete data, small sample sizes and gaps; (4) a spatially explicit habitat suitability map can be directly produced; and (5) the importance of individual environmental variables can be evaluated using a built-in jackknife. However, it is essential to compare the predictive efficiency of the Maxent model using a presence and absence SDM. Hence, the Logistic regression model was selected based on the criteria that both models use different input data type and modelling procedure. Considering the different capabilities of both models, there is a need for a more reasonable and cost-effective modeling approach in relation to the limitation of resources and budget constraints in the data collection process for large-scale operations.

In this study, using climatic and environmental variables we built SDM's for modelling the habitat suitability of the *C. tristis* using the following approach: 1) to validate Maxent's robustness in modelling the suitability preference of the *C. tristis*, we compared it with the Logistic regression model that uses presence and absence data; 2) to test the performance of remotely sensed data in modelling the suitability preference of the *C. tristis*; 3) determine the factors that influence the suitability of the *C. tristis* in Mpumalanga, South Africa.

2.2. Methods and Materials

2.2.1 Study area

The study was conducted in commercial *Eucalyptus* plantations of the Mpumalanga province of South Africa (Fig. 2.1). *Eucalyptus* plantations occupy an area of 23 928 Ha at an elevation that ranges between 1200m to 2100m. The mean annual precipitation for the area is 630–1600 mm and the mean annual temperature is $13 – 21^0$C. *E. nitens* is planted in this region due to the cold tolerance of the tree species. Compartments are managed for pulp and timber production and between 1 ha to 100 ha.

2.2.2 Species occurrence data

Commercial *Eucalyptus* compartments are annually assessed by Sappi for *C. tristis* induced infestation. The assessments are done following a two-tier approach. This is done during winter (June – July) and summer (August - October) seasons. Our field surveys were conducted during this period because the larvae stages occur between June and July and the adult moth is identified between August and October. The age of the *E. nitens* trees ranged between 4.5 to 6.7 years. Using a zigzag sampling technique, the number of infested trees were measured within a pre-determined number of transects across each stand (Boreham 2006). Transects were distributed evenly across each stand to ensure full representation. Each transect was made up of 100 live trees with the number of transects per stand area being proportional to the planted area of the stands. In each hectare, one plot was selected randomly and those less than one hectare was excluded from the survey (Adam et al. 2013). To determine the presence and absence of the moth, the boring dust on the stem or on the floor around the base of the tree was used an indicator (Boreham 2004, 2006, Adam et al. 2013). The number of infested trees per plot were then counted and expressed as a percentage for each surveyed stand. The attained percentages were used to indicate the suitable and unsuitable habitats of the *C. tristis*. According to the surveyed stands ($n = 77$), only 37 stands had signs of *C. tristis* infestation indicating suitable habitats while 40 compartments were free from infestation. Using ArcGIS 10.4, a polygon dataset was created to represent the suitable and unsuitable habitats of the *C. tristis*. These records were used to create presence and absence and presence only data to train (70%) and test (30%) the models. These recorded suitable and unsuitable datasets were then used to extract information using climatic and environmental and variables.
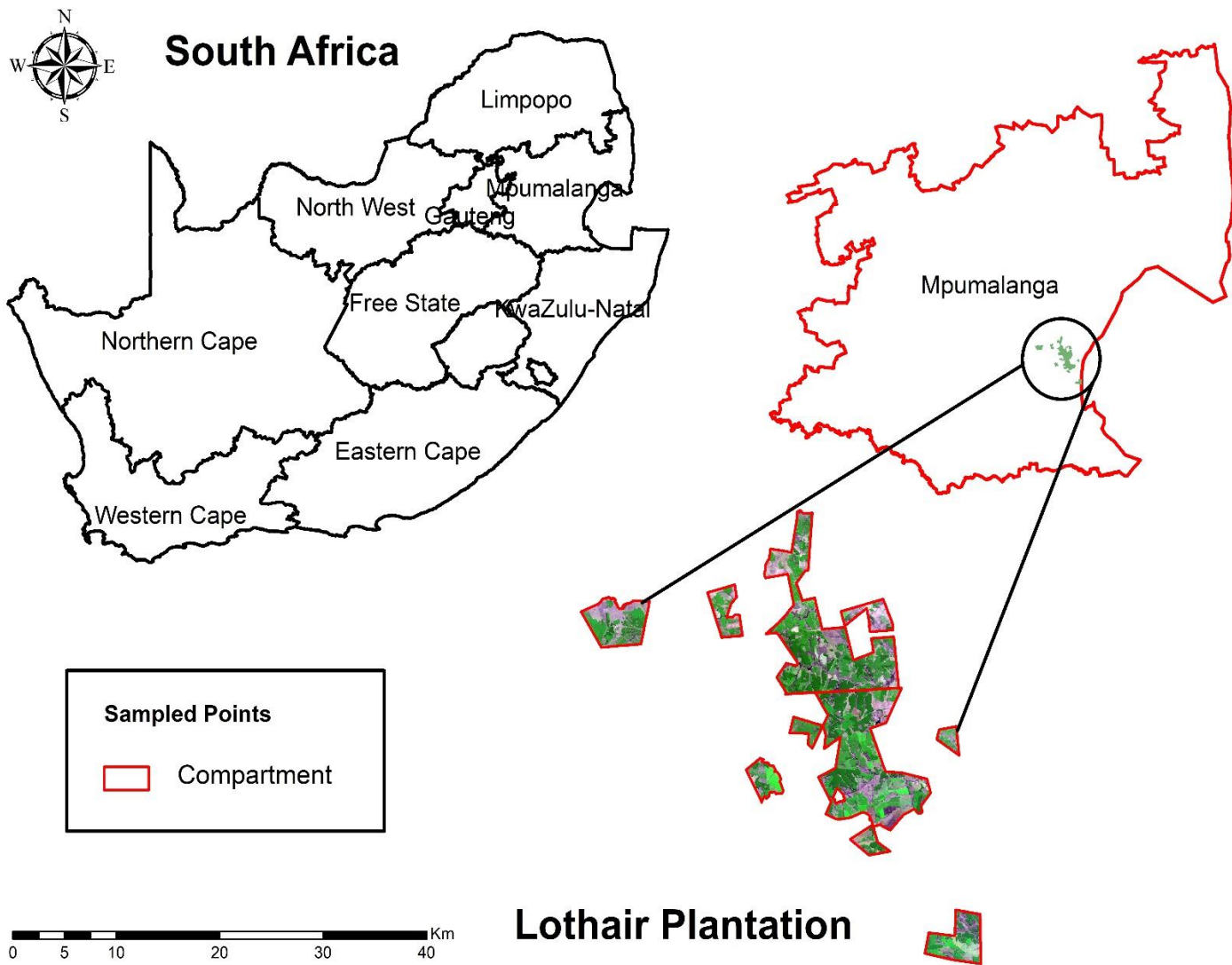
Figure 2. 1 Study area of Sappi plantations in Lothair, Mpumalanga, South Africa with a color composite of RGB (Red, NIR & Blue) using a Sentinel 2 image.

2.2.3. Climatic and Environmental predictors

A total of 32 environmental variables were considered when developing the *C. tristis* model. Multicollinearity between independent variables was checked through the calculation of the variance inflation factor (VIF) in the Logistic regression method. Variables that had a VIF lower than 10 were selected because they indicated that there was no multicollinearity between independent variables (Table 2.1). Precipitation and temperature variables were obtained from the WorldClim dataset (Fick and Hijmans 2017) at 30 arc-second (1 km x1 km grid cells) resolution. The dataset consisted of mean precipitation and temperature for the 12 months derived from historical records from weather stations across the globe, and it is available at http://www.worldclim.org (accessed 24 October 2016). Table 2.1 shows the selected variables to model the habitat suitability of the *C. tristis*.

Table 2. 1.Twelve variables selected for modelling of the suitability of the *C. tristis*.

Variable Type	Variables
1. **Environmental**	Age, Aspect, Elevation, Slope
2. **Climatic**	Mean Temperature- February, June, October, December
	Mean Precipitation- January, July, October, November

Both variables were calculated on the basis of monthly averages of rainfall and temperature and were significant because they contained the average information of the trends experienced during that particular month. Using a 1m DEM derived from LIDAR, topographic data comprised of slope, aspect and digital elevation were extracted using the spatial analyst tool in ArcGIS 10.4. However, Maxent is compatible with ASCII raster datasets and all the variables should have the same pixel size, extent and projection system in order to run the model (Ndlovu et al. 2018). Therefore, all the other variables were resampled to 1m spatial resolution and projected to the Universal Transverse Mercator (UTM) projection to match topographic variables. Hence, the conversion of all variables from raster to ASCII was carried out in ArcGIS to ensure all variables match.

2.2.4 Statistical Data analysis

    2.2.4.1 Maxent and Logistic Regression

Maximum entropy (Maxent) distribution model was used in this study. The model uses presence-data-only and the related environmental and climatic variables to model habitat suitability of the *C. tristis*. Maxent applies the maximum-entropy principle to fit the model and compares the interactions between the presence locations and variables to estimate the probability of species distribution (Berthon et al. 2018; Elith et al. 2011; Phillips et al. 2017). A complementary log-log (clog log) output was utilized as it strongly predicts areas of moderately high output (Phillips et al. 2017). The regularization multiplier was set at 4 to avoid overfitting of the test data. Model parameters were set to default replication of 1 with 500 iterations using cross-validation run type. Final outputs of the Maxent model predictions were exported to ArcGIS 10.4 for further analysis. The results from the model serve as an approximation of the suitable ecological niche for the moth under the studied environmental conditions.

In comparison, a Logistic regression model that depends on presence-absence datasets was utilized in this study based on principles to predict the causal relationship between predictors (independent variables) and predicted variables (dependent variables) (Gumpertz et al. 2000; Neupane et al. 2002). Using the Statistical Package for the Social Sciences (SPSS), *C. tristis* presence-absence datasets were used as the dependent variables, while the environmental and climatic variables were the predictors. To generate the best combination of predictors and approximate beta (β), we used a backward stepwise (conditional) entry of variables criteria and maximum likelihood method. In addition, for the model to accept species presence from the model prediction, a random threshold probability was required. Logistic regression is most sensitive to threshold effects because a given threshold can interact with species' prevalence (i.e. the frequency of suitability) to influence positive and negative prediction error (Gribko et al. 1995; Otunga et al. 2018). Hence, a probability threshold value greater than or equal to the selected threshold illustrates suitable habitats, while a threshold lesser than the selected value shows unsuitable habitats. As a result, only predictors with confidence levels above 95% or a p-value less than 0.05 were considered significant and used in fitting the Logistic regression function. Lastly, to validate the robustness of Maxent and Logistic

regression for mapping habitat suitability of the *C. tristis*, the dataset was randomly split into 70% training data and 30% test data and used for accuracy assessment.

2.2.4.2 Accuracy assessment

Receiver operating characteristic (ROC) area under the curve (AUC) method has been widely used for comparing species distribution model performances of Maxent and Logistic regression models (Bagheri et al. 2018; Cianci et al. 2015; Remya et al. 2015). The ROC plots the sensitivity values and the false-positive fraction for all available probability thresholds (Germishuizen et al. 2017). Sensitivity is the ability of a model to correctly identify known positive sites and specificity is the ability of a model to correctly identify known negative sites (Cianci et al. 2015; Phillips and Dudík 2008). AUC provides a single measure of model performance independent of any particular choice of threshold, making it an excellent index to evaluate model performance (Baldwin 2009). The AUC measures model performance that ranges from 0 to 1. Values close to 0.5 points to a random prediction, while a value of 1.0 indicates a perfect fit (Dicko et al. 2014; Fourcade et al. 2014).

Response curves are the most important aspects of species distribution modelling, because they can provide information on the relationship between the species and the environment (Baldwin 2009). Using Maxent and Logistic regression models, response curves showed how each of the environmental and climatic variables predicted habitat suitability of the *C. tristis*. The Maxent predictions (clog log output value) greater than 0.5 indicate conditions that are suitable and less than 0.5 showed unsuitable conditions for the distribution of the *C. tristis*. On the other hand, Logistic regression response curves depended on the significance of the relationship between the independent and predictor variables with alpha at 0.05 in determining the suitable habitat for the *C. tristis*. The Jackknife test was used to examine the importance of individual variables for Maxent predictions (Makori et al. 2017; Ndlovu et al. 2018).

2.3. Results

2.3.1 Evaluating the performance of Maxent and Logistic regression for detecting *C. tristis* presence.

Figure 2.2 displays the accuracies derived from estimating the suitable habitats for the *C. tristis*. Our results indicated that the Logistic regression and the Maxent have roughly similar efficiencies in predicting habitat suitability of the *C. tristis*. Using training and testing data respectively, Maxent produced a higher AUC of 0.840 and 0.810 when compared to Logistic regression (0.745 and 0.677). According to the sensitivity and specificity values, Maxent outperformed the Logistic regression.

2.2a

2.2b
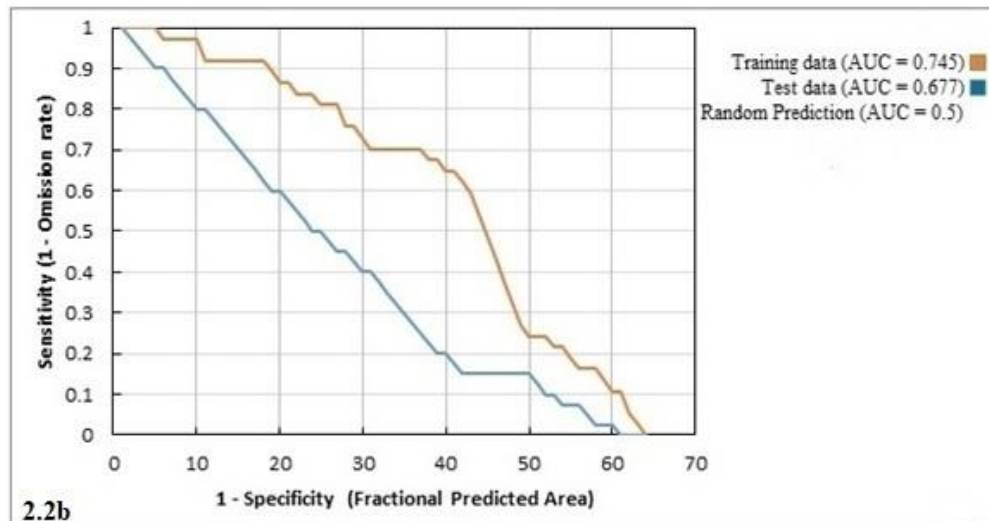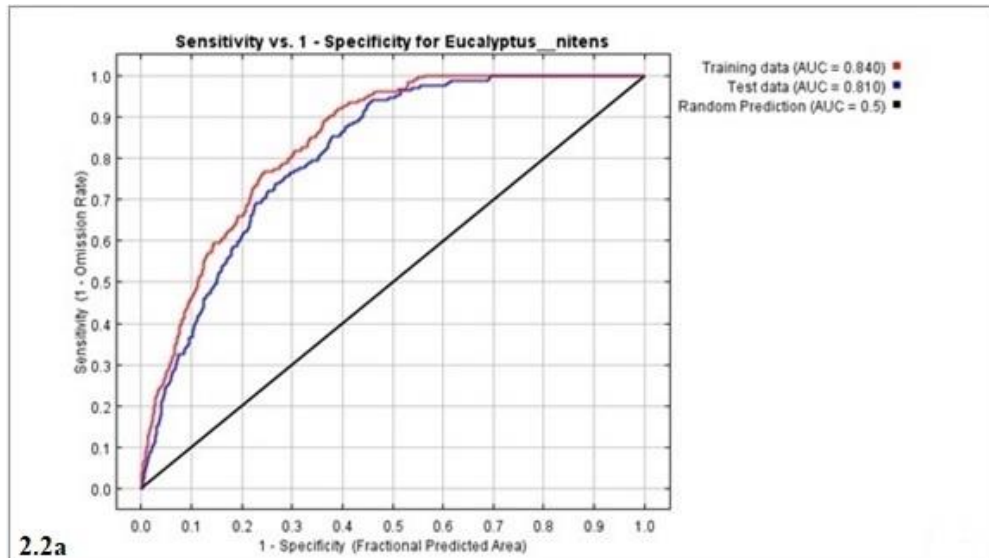
Figure 2. 2 AUC evaluating the performance of a) Maxent and b) Logistic regression in predicting the habitat suitability of the *C. tristis* using the selected variables.

2.3.2 Evaluating the significance of environmental and climatic predictors for *C. tristis* presence.

Figure 2.3 shows the contribution of the predictor variables in modelling the *C. tristis.* The Maxent model (Figure 2.3a) produced a test jackknife that indicated the relative importance of each variable in the modelling process. In Figure 2.3a, the most influential variables in the model were mean temperature for October, February, June, December and elevation respectively. As illustrated in Figure 2.3b, age was the only significant variable in the Logistic regression model.
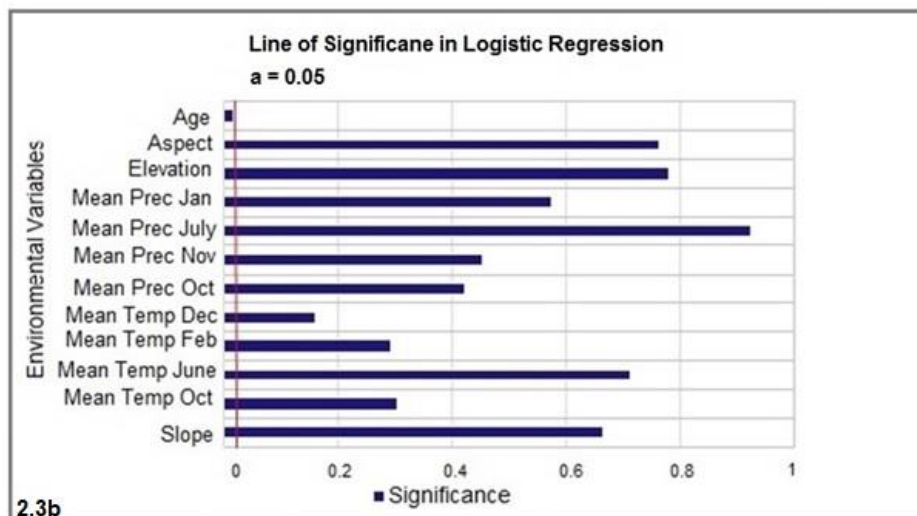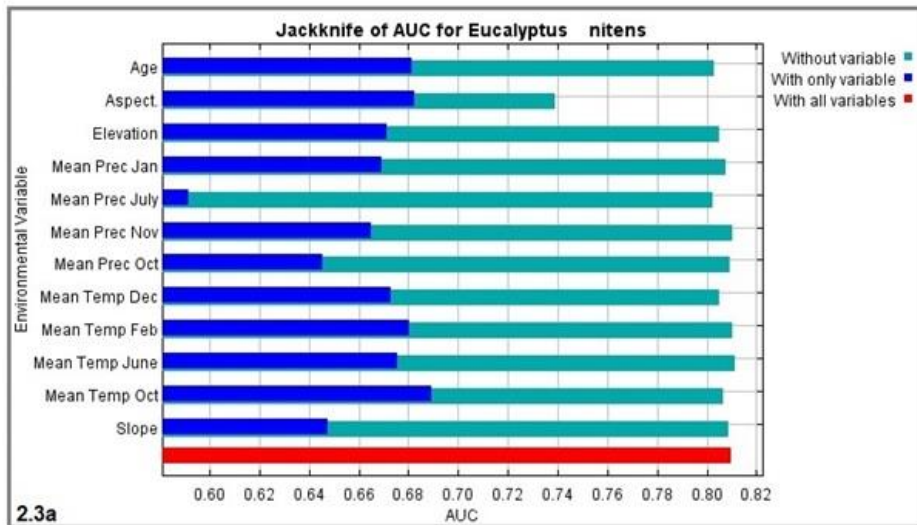
Figure 2. 3 a) Jackknife illustrating the variables that influence the prediction of the *C. tristis* using Maxent and b) indicating the significance and non-significant variables used to model the occurrence of the *C. tristis* using the Logistic regression.

Table 2.2 shows the results from Logistic regression that was used to determine the significant and non-significant variables in the model. Age was found to be the only significant variable among the 12 variables used in the model.

Table 2. 2 Results of the Logistic regression model.

	Estimate β	S.E.	Sig. p-value	exp(b)
Intercept	-59.3436	56.47802	0.293379	1.6881E-26
Age*	1.61089	0.674597	**0.016944***	5.00726634
Aspect	0.06948	0.23123	0.76381	1.07195083
Elevation	-0.06539	0.233941	0.779861	0.93670512
Mean Prec Jan	-0.12098	0.215493	0.574519	0.88605183
Mean Prec July	0.091586	0.963713	0.924287	1.09591131
Mean Prec Nov	0.217647	0.289818	0.452666	1.24314777
Mean Prec Oct	0.24557	0.306542	0.423076	1.27834945
Mean Temp Dec	16.36976	11.70297	0.161882	12861654.2
Mean Temp Feb	-8.75496	8.309308	0.292051	0.00015768
Mean Temp June	-1.5471	4.202071	0.712742	0.21286366
Mean Temp Oct	-7.06748	6.870387	0.303626	0.00085237
Slope	0.079346	0.183817	0.665991	1.08257837

Figure 2.4 shows the response curves of the seven most optimal environmental variables. The results in Fig 2.4 indicates that suitability of the *C. tristis* is associated with the mean temperature for October that is greater than 14.5 °C.
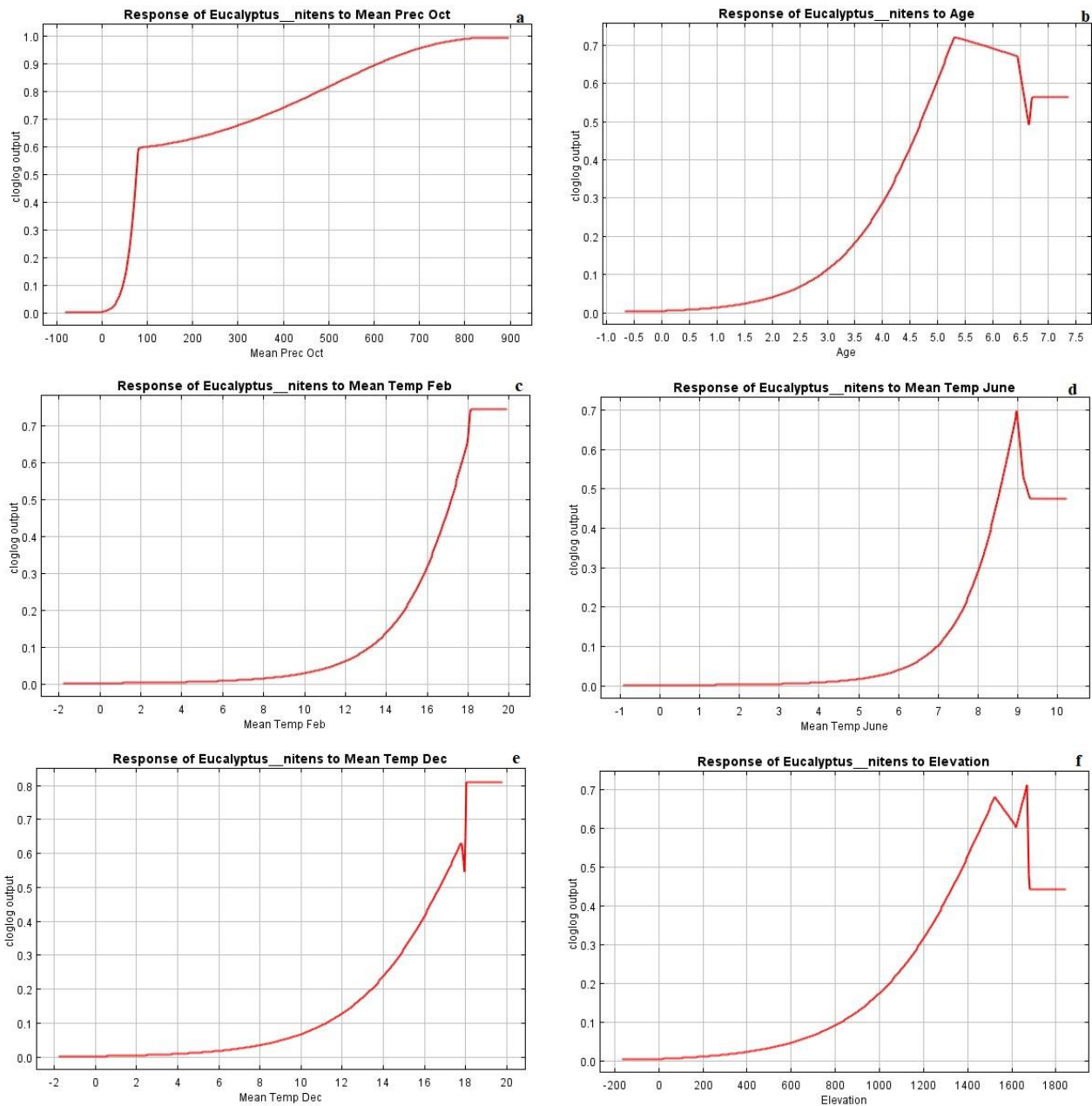
Figure 2. 4 Response curves of mean temperature for October (a), age (b), mean temperature for February (c), mean temperature for June (d), mean temperature for December (e) and elevation (f) that show how these selected variables affected the prediction of the *C. tristis* using Maxent.

Additionally, the results illustrate that aspect played a crucial role in identifying suitable conditions for the *C. tristis*. Figure 2.4 c shows that *E. nitens* plantations above the age of 4.7 provide a suitable habitat for the *C. tristis*. This means that plantations below 4.7 are likely unsuitable for the moth to occupy them. More specifically, the mean temperature of February (16.7 ^{o}C), June (8.5 ^{o}C) and December (16.4 ^{o}C) strongly influenced the favorable habitat for the *C. tristis*. Finally, results suggest that elevations between 1400m – 1650m have suitable conditions that favor the distribution of the *C. tristis* as compared to other areas in the study area (Fig 2.4 f). As a result, the contribution of remotely sensed data, the topographic data (aspect, elevation and slope) extracted from LIDAR helped improve the overall performance of both SDMs.

2.3.3 Spatial distribution of areas susceptible to *C. tristis* habitation

Figure 2.5 shows the suitability map of the *C. tristis* as predicted by Maxent and Logistic regression. From the maps, both models yielded good results using climatic and environmental variables. The visual assessment indicates that Maxent produced a highly suitable probability map when compared with the Logistic regression model. In this study, Maxent predictions of the moth corresponded to the rescaled suitability index (cloglog output), whilst the Logistic regression predictions corresponded to the probabilities of presence. Generally, the moth is projected to likely spread from the northern parts to the southern parts of the plantations. The spatial distribution corresponds to temperature conditions, age, elevation and precipitation.

Figure 2. 5 Maps showing the prediction of occurrence of the *C. tristis* using Logistic regression and Maxent

2.4 Discussion

The aim of this study was to investigate the habitat suitability of the *C. tristis* and to compare the performance of two species distribution models (Maxent and Logistic regression) as well to test the performance of remotely sensed data in modelling habitat suitability preference. The study utilised presence and absence and presence-only datasets to understand the contribution of multi-source data in mapping the suitable habitats 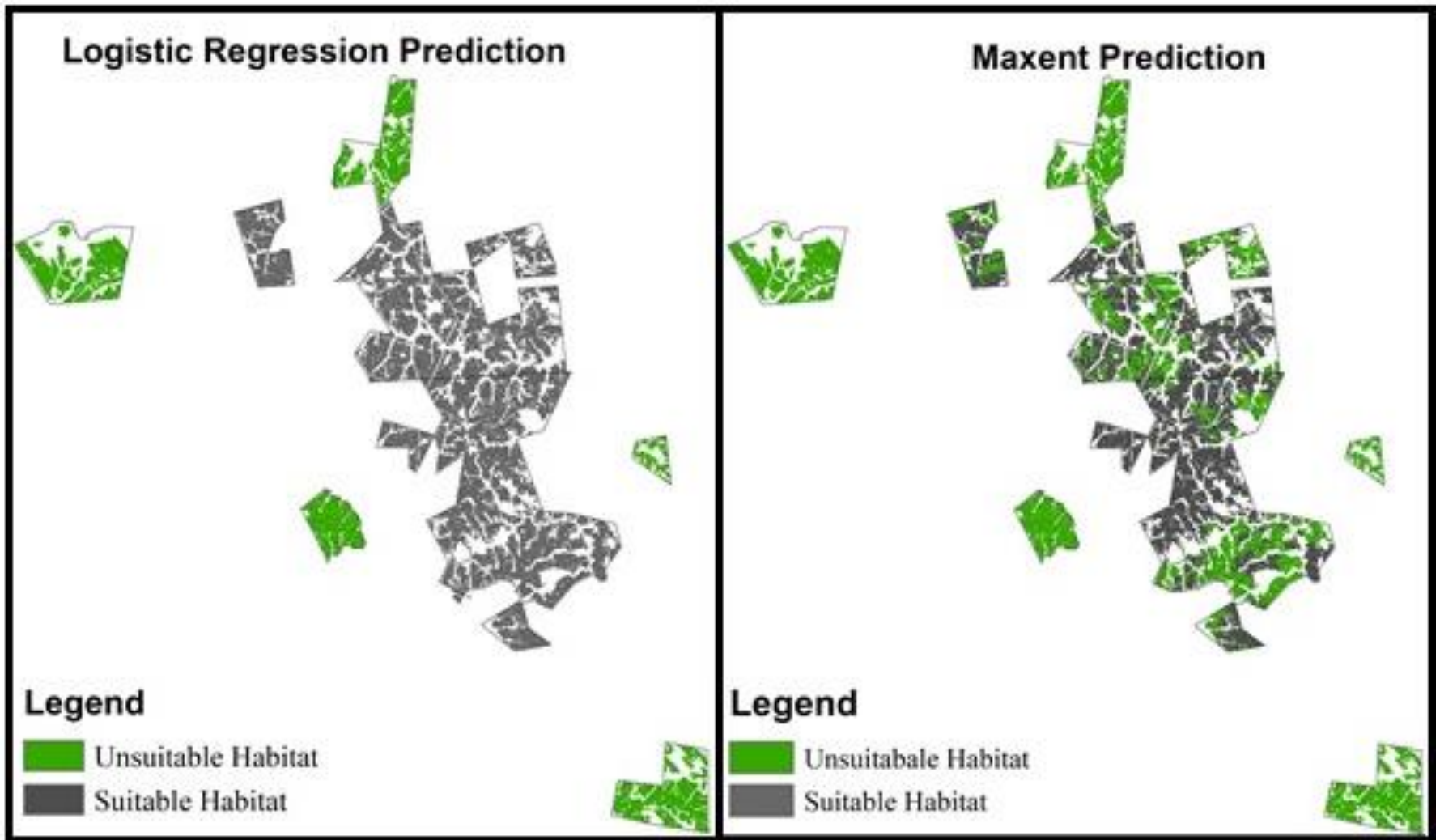of the *C. tristis*. AUC statistics of both models showed high values indicating a good model performance in relation to predicting suitable habitat distribution. Beyond describing species distributions, Maxent and Logistic regression have been considered as important and widely used decision making tools that can assist forest managers (Gribko et al., 1995; Gumpertz et al., 2000; Matawa et al., 2016; Ndlovu et al., 2018; Sahragard and Ajorlo, 2018).

Several studies have agreed that temperature influences the occurrence of the Lepidopteran defoliators (Adam et al. 2013; Boreham 2006; Michael and Warren 2009; Péré et al. 2013; Q. et al. 2017). It is not surprising that summer temperatures had the highest discriminatory power in predicting the highly suitable areas for the *C. tristis*. The current study established that mean temperature of February (16.7 ^{o}C) and December (16.4 ^{o}C) are strongly associated with the suitability preference of the moth on *E. nitens*. In addition, in October a mean temperature greater than 14.5 ^{o}C creates a conducive environment for the *C. tristis* to expand its habitat suitability. These results corresponded with previous studies that indicated that in October in the Lothair plantations all stages of the moth could still be found. According to Bentz el al. (2010 and Centre and Carroll (2006), ongoing expansion of the mountain pine beetle (*Dendroctonus ponderosae*) has been observed due to increased summer temperatures that have resulted in the beetle surviving in previously unsuitable ecological areas. Moreover, in Alaska and Yukon, the high summer temperatures have been associated with an outbreak of the spruce bark beetle (*Dendroctonus rufi collis*) on Engelmann spruce *(Picea engelmannii)* forests increasing its habitat suitability (Berg et al., 2006). Hence, increased summer temperatures and shortened winters influence the rapid insect reproduction, faster growth rates and mobility of insect pests, which influence the overall habitat preference (Battisti et al., 2006; Kocmánková et al., 2009).

The results showed that the habitat preference of the *C. tristis* increases on *E. nitens* tree species above 4.5 and decreases on trees species below 4.5. Boreham (2006) established similar results in characterizing the infestation of *E. nitens* tree species younger than 8 years of age using the Residual Maximum Likelihood (REML) method. *E. nitens* tree species are known as fast-growing trees species that produce high-quality timber in a short period of time. Hence, older trees have stronger barks that provide favorable larvae feeding conditions, which result in internal damage through infestation of the sapwood and hardwood (Adam et al. 2013). Currently, the *C. tristis* is regarded as a primary pest on *E. nitens* tree species that has the potential to become an epidemic pest due to the extensive population outbreaks. Hence, failure to manage and control the moth will result in tree mortality, which affects the production of high-quality timber. Once productivity is affected, this generates a problem for forest managers as the quality of timber decreases, thus reducing the net profits earned. Therefore, knowledge of the age of tree species that are vulnerable to infestation is crucial as it improves the management and monitoring programs.

Furthermore, the mean temperature for June (8.5 °C) was also a key element in determining the suitable habitat of the *C. tristis*. Previous studies stated that signs and symptoms of larvae damage were observed in July in the Lothair/Carolina area. The Highveld is associated with cold temperatures that are similar to parts of the Western Cape Province where the *C. tristis* has been recorded. Hence, the current study suggests that cold temperatures create favourable conditions that result in the larval stage of *C. tristis* occurrence. This effect has also been recorded on pine tree species as increased winter temperatures lead to better performance of the winter-feeding of larvae by the pine processionary moth (*Thaumetopoea pityocampa)* (Buffo et al., 2007). Furthermore, changes in the larval performance of the moth have strongly contributed to the progressive colonization of new areas increasing its habitats. In addition, the winter moth (*Operophtera brumata*) has also been reported to have expanded its habitat into the coldest continental landscapes and it is associated with the increasing winter temperature that lead to higher survival of overwintering eggs (Jepsen et al., 2008). As a result, cold temperatures play a vital role in the suitability of the habitat of the *C. tristis*. Certain temperature ranges as indicated by the results of this study trigger the *C. tristis* life cycle process influencing its habitat preference. According to the results of this study, precipitation variables did not perform as expected in modelling the habitat suitability of the *C. tristis*. Previous studies show that precipitation has not greatly influenced the habitat preference of the *C. tristis* and this area requires further studies.

Elevation is considered an important predictor that enhances the understanding of the distribution patterns of insect pests (Péré et al., 2013; Thomas et al., 2006). Results showed that the habitat preference of the *C. tristis* ranges between 1400 m and 1650 m. The relationship between suitable habitats of the *C. tristis* can be associated with matting, host foraging and ovipositional behaviors absent at an elevation less 1500m and greater than 1650m (Péré et al. 2013). Hence, without these three conditions, the development and survival of the moth are highly unsuitable. Furthermore, *E. nitens* plantation richness in different elevations also contributes to plantations being suitable for the moth's presence, because the greater the availability of the tree species the greater the risk of infestation. Hence, elevation plays a pivotal role in the occurrence of the *C. tristis*. Established results agree with the previous studies that showed the presence of the *C. tristis* ranges between 1500m and lower elevations below 1600m. According to Battisti and Larsson (2015), elevation and longitudinal expansion are regarded as the most common factor that influences the habitat suitability of insects pests. Also, it is worth noting the contribution of remotely sensed data in modelling habitat suitability of the *C. tristis*. The results in this study agree with Kozak et al. (2008) and Ndlovu et al. (2018) who indicated that integration of remotely sensed data into SDM's improves the overall performance of SDM in modelling species distributions. Moreover, LIDAR as a remote sensing tool is a promising tool for identifying suitability preferences of species that inhabit divergent climatic regimes.

Assessing the performance of both models was the main focus of this study, the results demonstrated that Maxent was more robust than Logistic regression. This outcome is not surprising as several studies have identified Maxent as one of the best alternatives in determining species distributions (Berthon et al. 2018; Elith et al. 2011; Phillips et al. 2017). Observation of the two distribution maps in Fig 2.5 indicated that the presence-only Maxent model produced a better habitat suitability map as compared to the presence-absence Logistic regression model. Dicko et al. (2014) had similar results demonstrating that only the Maxent model predicted an expert-based classification of landscapes correctly in their study as compared to Logistic regression. Noticeably, Logistic regression generated suitable habitats based on probabilities of presence data provided by data collectors. However, Fithian and Hastie (2013) challenged the availability of reliable absences records indicating that unreliability can be associated with identification errors and mostly inadequate knowledge of the target species. At the moment, the *C. tristis* is regarded as an emerging pest in forestry and less information is known about the moth on

E. nitens plantations. Consequently, the collection of the data of the moth can be affected by different aspects, such as; the lack of observer experience, identification errors and high costs associated with the process. As a result, comprehensive information (presence-absence data) of the *C. tristis* is essential to reduce the uncertainty in modelling the spatial distribution of the moth using the Logistic regression model. Hence, the results in this study confirm that suitability of the *C. tristis* on *E. nitens* plantations can be modelled using climatic and environmental variables and provides valuable information required by forest managers for effective inoculation and control of damaging pests, such as the wood boring *C. tristis*.

2.5 Conclusion

This study assessed the robustness of the Maxent model, compared with the Logistics regression model, in mapping the habitat suitability of the *C. tristis* in Mpumalanga, South Africa. Grounded in the results of this study, we conclude that:

- Temperature, aspect, age and elevation are optimal variables for modelling the suitability of the *C. tristis*
- Maxent model is a robust algorithm in relation to other methods such as Logistics regression model in mapping the habitat suitability of the *C. tristis*
- Integration of remotely sensed data from LIDAR improved the overall performance of SDMs

The results offer a useful tool to forest managers in understanding the climatic and environmental characteristics that influence the habitat suitability of the *C. tristis* on *E. nitens* compartments. At the moment, chemical control is not a feasible option as the use of systemic insecticides to kill the larvae would be impractical and expensive. Hence, the application of SDMs would benefit forest managers to formulate new suitable integrated pest management strategies to reduce infestation of un-colonized *E. nitens* plantations. However, the results in this study determined the habitat suitability of the moth based on the surrounding conditions and not the actual damage of plantations. Therefore, we recommend that future studies look at the utility of remote sensing and GIS to map and model the suitability distribution of the *C. tristis*.

Chapter Three

Using Multispectral remote sensing to map habitat suitability of the Cossid Moth in Mpumalanga, South Africa.

This chapter is based on:

Abstract

The study sought to model habitat suitability of the *Coryphodema tristis* on *Eucalyptus nitens* plantations in Mpumalanga, South Africa, using a Sentinel-2 multispectral instrument (MSI). Traditional field surveys were carried out through mass trapping in all compartments and positively identified 67 infested compartments. Model performance was evaluated using the receiver operating characteristics (ROC) curve showing the area under the curve (AUC) and True Skill Statistic (TSS) while the performance of predictors was displayed in the jackknife. Using only the occurrence data and Sentinel-2 bands and derived vegetation indices, the Maxent model provided successful results, exhibiting an area under the curve (AUC) of 0.89. The Photosynthetic vigour ratio, Red edge (705 nm), Red (665 nm), Green NDVI hyper, Green (560 nm) and Shortwave infrared (SWIR) (2190 nm) were identified as the most influential predictor variables for detecting the habitat suitability of the *C. tristis*. Results of this study suggest that remotely sensed derived vegetation indices from cost-effective platforms could play a crucial role in supporting forest pest management strategies and infestation control.

Keywords: Multispectral remote sensing, Eucalyptus *nitens*, *Coryphodema tristis* (Cossid moth), Sentinel 2, Maxent model.

3.1. Introduction

In South Africa, emerging forest pests have caused extensive damage to *Eucalyptus* plantations (Wingfield et al. 2001). Approximately 1.3 million hectares of South Africa's land is composed of both hard and softwoods, with the majority located on the eastern parts of the country; primarily in Mpumalanga (40.8%), KwaZulu-Natal (39.5%) and the Eastern Cape (11.1 %) (DAFF, 2015). These plantations contribute annually to South Africa's GDP with *Eucalyptus* plantations contributing over 9% to the total exported manufactured goods (DAFF 2017). These species are the most productive planted exotics that mostly offer timber, pulp and paper in South Africa (Albaugh et al. 2013; Swain and Gardner 2003; Wingfield et al. 2008). Therefore, a robust mechanism needs to be established to prevent excessive damage, as numerous investments have

been injected into the forestry sector, particularly the Mpumalanga province (SETA 2014). Since 2004, *Coryphodema tristis,* commonly known as Cossid moth, has been the major damaging agent destroying *Eucalyptus nitens* plantations across Mpumalanga, with forest managers requiring up to date information to support their forest protection interventions at the landscape level.

C. tristis is an indigenous wood-boring insect that commonly infests tree species, such as *Ulmaceae* (Elm Family*)*, *Vitaceae* (Wild Grape family)*, Rosaceae (*Rose family), *Scrophulariaceae* (figwort family), *Malvaceae* (Mallow family) and *Combretaceae* (Indian almond family) (Bouwer et al. 2015; FAO 2007). However, a sudden shift by the *C. tristis* to infest *E. nitens* in South Africa has been observed. According to Gebeyehu et al. (2005), the shift of the *C. tristis* to infest *E. nitens* trees may be as a result of few to the non-existence of natural enemies in the area. As a result, the absence of natural enemies influences the increase of pests in the ecological niche, due to less interspecific competition (Xing et al. 2017). This results in the moth breeding and multiplying at faster rates and increasing the intensities of *E. nitens* infestation. Adult female moths lay eggs on the bark of the *E. nitens* trees and the larvae feed on the bark damaging the cambium (Gebeyehu et al. 2005). The damage reduces the movement of water within the tree and also extend to the trunk and branches which turn black (Adam et al. 2013). Furthermore, as the larvae grow, it drills extensive tunnels into the sapwood and hardwood trees producing resin on trunks and branches and sawdust on the base of the forest floor (Bouwer et al. 2015). However, extensive tunneling by the moth results in severe damage to trees, thus increasing the probability of tree mortality. Additionally, pupal casings are found protruding on the holes tunneled or either at the base of the floor indicating the presence of the *C. tristis*.

In recent years, researchers have attempted to use environmental variables to predict the spatial distribution of the *C. tristis* (Adam et al. 2013; Boreham 2006). For example, Boreham (2006) conducted a study that investigated the outbreak and impact of the *C. tristis* on *E. nitens* in the Highveld of Mpumalanga, using environmental variables and the Residual Maximum Likelihood (REML) algorithm. Their results showed that older *E. nitens* trees (above 8 years) and lower elevation sites less than 1600m were the most susceptible to *C. tristis* infestations. Similarly, Adam et al. (2013) used climatic and topographical variables to map the presence and extent of *C. tristis* infestations on E. *nitens* plantations of Mpumalanga. Using a random forest classifier, their results indicated that September and April maximum temperature, April median rainfall and elevation

played a crucial role in identifying conditions that are suitable for *C. tristis* occurrence. Furthermore, their results predicted that areas with a maximum temperature greater than 23°C in September and 22°C in April were the most susceptible to infestation. While these studies have successfully utilized climatic and environmental variables to predict the presence of the moth. Different studies have identified a number of limitations regarding traditional data collection methods to determine the presence or absence of pests.

Different studies have stated that traditional methods are often time-consuming, costly, labor-intensive, spatially restrictive and likely unreliable as data collection is based on the knowledge of the surveyor (Pause et al. 2016; Pietrzykowski et al. 2007). Hence, a direct detection approach that provides real-time information and can be repeated regularly for up to date decisions is required. Furthermore, utilizing environmental or climatic variables only for mapping the spatial distribution of pests can be challenging since these variables focus precisely on the surrounding factors and not the actual damage of plantations. For example, Germishuizen et al. (2017) utilized environmental factors to determine the susceptibility of pine stands to bark stripping by Chacma baboons (*Papio ursinus*). Results indicated that indirect variables such as elevation and altitude provide a challenge in explaining the complex relationship of baboon-damage risk. Moreover, Donatelli et al. (2017) indicated that observed environmental datasets alone were no longer sufficient to predict the behavior of pests, due to climate change that has influenced the variability of temperature averages, rainfall means and distributions. Thus, requiring more traditional field surveys to confirm whether a particular area has been truly infested. Bouwer et al. (2015) indicated that actual confirmation of infestation was through tree felling, which is impossible for large-scale assessments. Hence, the inclusion of remotely sensed data with ancillary data such as environmental and climatic variables would provide an up to date, repeatable source of information for forest assessment and inventory.

Remote sensing has achieved unprecedented perspectives of forest-damaging pests using narrow and broad wavebands in the visible, near, shortwave-infrared and red edge regions (Lottering et al. 2016; Oumar and Mutanga 2013; Pietrzykowski et al. 2007). For example, Adelabu et al. (2014) sought to discriminate the levels of change in forest canopy cover instigated by insect defoliation using hyperspectral data in mopane woodlands. Results indicated that the overall accuracy of classification was 82.42% using random forest and was 81.21% using ANOVA. In another study,

Oumar and Mutanga (2013) successfully assessed the potential of WorldView-2 wavebands, environmental variables, as well as vegetation indices which resulted in the prediction of *Thaumastocoris peregrinus* infestations on *Eucalyptus* trees. Results indicated that WorldView-2 sensor bands and indices predicted *T. peregrinus* damage with an R^2 value of 0.65 and a root mean square error of 3.62% on an independent test data set. Similarly, Lottering et al. (2016) also found that vegetation indices derived from the red edge region correlated with *G. scutellatus*-induced vegetation defoliation using WorldView-2 satellite data. Furthermore, Pietrzykowski et al. (2007) assessed the presence and severity of defoliation and necrosis caused by the *Mycosphaerella* insect on *Eucalyptus globulus* plantation, using a multispectral imagery in north-western Tasmania, Australia. Their results indicated that the spectral bands performed well, producing an accuracy of 71% for defoliation and 67% for necrosis. Therefore, despite the optimal modelling accuracies attained using multispectral remotely sensed data in these studies, these data sets are expensive and limited to a local scale. In that regard, there is an urgent need for testing and assessing the utility of other cheaper data sets that could capture the disease and pest incidences at landscape scales.

This study, therefore, sought to model habitat suitability of the *C. tristis* on *E. Nitens* plantations in Mpumalanga, South Africa using the cost-effective Sentinel-2 multispectral instrument and derived vegetation indices. Sentinel 2 images across the valuable red edge portion of the electromagnetic spectrum are suitable for forest health applications related to pest and disease damage detection (Hojas-Gascon et al. 2015; Immitzer et al. 2016; Ng et al. 2017). The large swath width and a 16-day temporal resolution make this sensor suitable for repeatable monitoring over forest plantations and detect pest-related damage continuously for effective management and control. Therefore, we used Maxent a robust machine-learning algorithm to predict habitat suitability of the *C. tristis* using remotely sensed data.

3.2. Methods and Materials

3.2.1 Study area
The research was conducted in the Mpumalanga province of South Africa in the Lothair village also known as Silindile and is located in the Msukaligwa Local Municipality (Fig 3.1). The study site is located between 26° 26' 25.08" S and 30° 3' 59.4" E in the Highveld of Mpumalanga. It has an altitude that ranges between 1200 m and 2100 m.
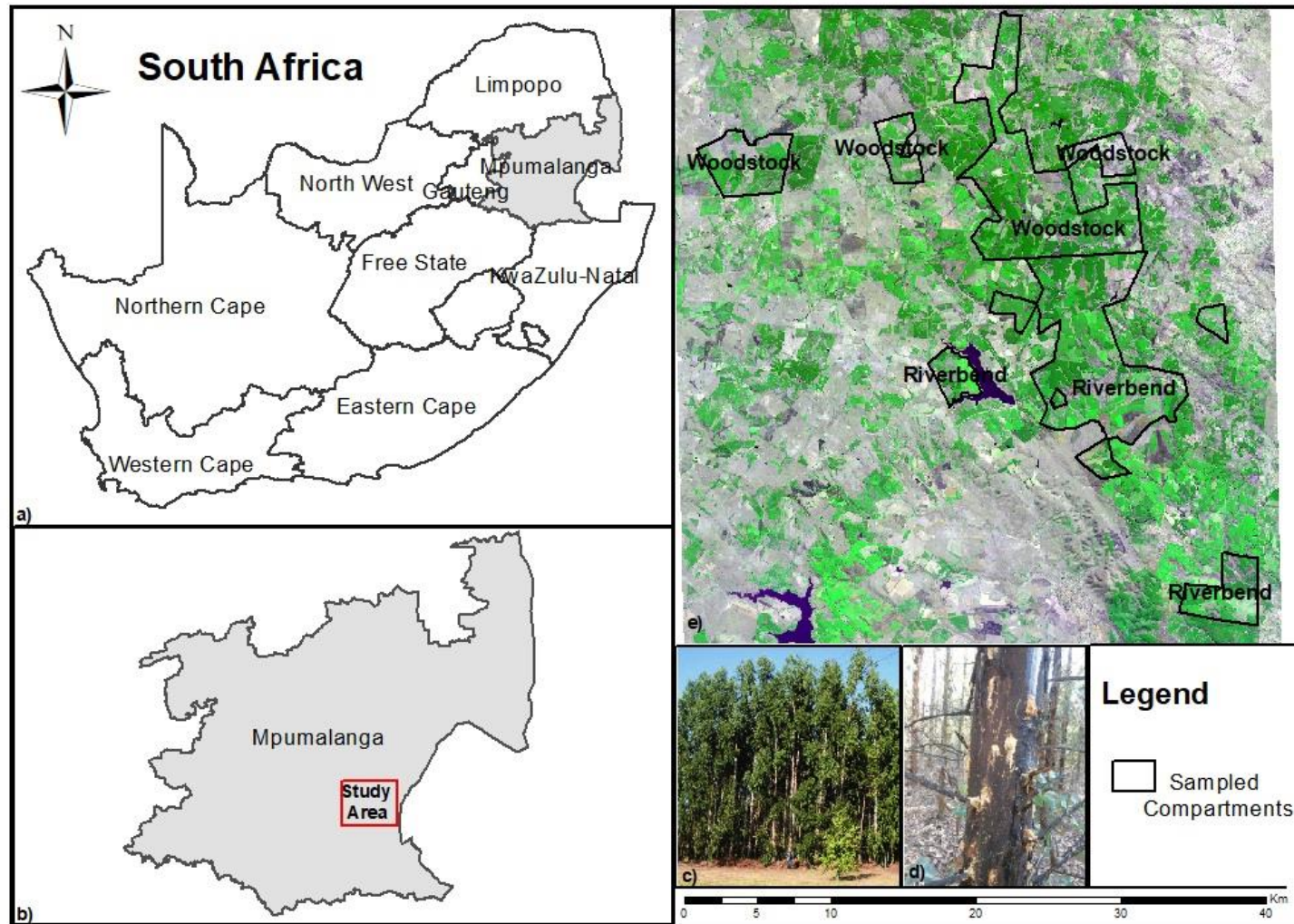
Figure 3. 1 a) Map of South Africa and b) the location of the Mpumalanga Province; (c) and (d) show healthy and infested *Eucalyptus nitens* and e) shows the sampled parts of the forest using the Sentinel 2 image with a colour composite of RGB (Red, NIR & Blue).

The area is associated with summer rainfall which ranges between 783–1200 mm per annum from November to March. The Highveld has a summer (October to February) to winter (April to August) temperature average of approximately 19º C, with average temperatures ranging between 8º C and 26º C in the contrasting seasons. The Highveld is among South Africa's highly productive commercial plantation forests that consist of Pine and *Eucalyptus* plantations. The greater parts of the Highveld are comprised of sandstone and granite derived soils, which the majority of commercial tree species are grown.

3.2.2 Image acquisition

A Sentinel 2 MSI image was acquired on the 19th of August 2016 under cloudless conditions, the sensor has a revisit time of 5 days making the detection of pest damage to vegetation instantaneous (Gascon et al. 2017; Hojas-Gascon et al. 2015; Immitzer et al. 2016). The satellite covers a large area with a swath width of 290 km for multispectral observations increasing the spatial coverage of the area of interests (Ng et al. 2017; Radoux et al. 2016). Sentinel 2 has thirteen spectral wavebands ranging from 443 nm to 2190 nm including four 10 m visible and near-infrared bands, six 20 m red edge, near infrared and shortwave infrared bands, and three 60 m bands visible, near-infrared and shortwave infrared bands. The narrow red edge wavebands cover spectral regions of 0.705 um, 0.740 um, 0.783um and 0.865um that can be utilized for monitoring vegetation status (Immitzer et al. 2016; Ng et al. 2017; Radoux et al. 2016).

3.2.3. Image processing and analysis

Atmospheric correction of the image was done using the Sentinel Application Platform (SNAP) software, which incorporates the plugin, Sen2Cor. In total, eleven bands were derived for modelling the suitable habitat of the *C. tristis*. In this study, 10 of the Sentinel 2 wavebands were used in the study excluding band 1, 9, and 10. These three wavebands were not incorporated in this study because they are not used for vegetation mapping. Using the Index Database, we selected vegetation indices with the best capacity to detect and map the *C. tristis* (see Table 3.1). Additionally, a number of published vegetation indices that have been effective in characterizing vegetation defoliation, many of which are sensitive to reflectance in the visible and NIR regions, were derived. However, vegetation indices with wavelengths from the red edge region were given more emphasis based on their ability to identify stressed vegetation (Lottering et al. 2016).

Table 3. 1: Sentinel 2 vegetation indices tested in this study

Vegetation indices	Abbreviation	Equation	Reference
Simple Ratio 800/500 Pigment specific simple ratio C1	PSSRc1	$\dfrac{NIR}{Blue}$	Blackburn (1998)
Simple Ratio 520/670	SR520/670	$\dfrac{Blue}{Red}$	Carter (1994)
Simple Ratio 774/677	SR774/677	$\dfrac{Vegetation\ Red\ edge}{Red}$	Zarco-Tejada et al. (2001)
Simple Ratio NIR/700-715	SRNir/700-715	$\dfrac{NIR}{(Red - Vegetation\ Red\ edge)}$	(Gitelson et al. 1996b)
Normalized Difference Vegetation Index	NDVI	$\dfrac{NIR - RED}{NIR + RED}$	Gitelson and Merzlyak (1997)
Normalized Difference 780/550 Green NDVI hyper	GNDVIhyper	$\dfrac{Vegetation\ Red\ edge - Green}{Vegetation\ Red\ edge + Green}$	Gitelson et al. (1996a)
Normalized Difference Salinity Index	NDSI	$\dfrac{SWIR\ (1.610) - SWIR\ 2.190}{SWIR\ (1.610) + SWIR\ 2.190}$	Richardson et al. (2002)
Normalized Difference 800/470 Pigment specific normalized difference C2	PSNDc2	$\dfrac{NIR - Blue}{NIR + Blue}$	Blackburn (1998)
Chlorophyll Green	Chlgreen	$\dfrac{(Vegetation\ Red\ edge)^{-1}}{Green}$	Gitelson et al. (2006)
Normalized Difference 550/650 Photosynthetic vigour ratio	PVR	$\dfrac{Green - Red}{Green + Red}$	Metternicht (2003)

3.2.4 Field data collection.

A field visit was conducted in two SAPPI plantations on the 19th of August 2016 to establish the presence of the pest in the area. Woodstock is located in the northern region of the plantation and consisted of 55 *E. nitens* plantations, whilst Riverbend located in the south comprised of 1145 plantations. Mass trapping of *C. tristis* was carried out in the field. Using a minimum of 19 and maximum of 348 traps randomly setup across all *E. nitens* stands. The number of traps used varied with the size of the compartments. Pheromones that match the chemical scent of a female adult moth was used to lure male moths into the traps that were located in the compartments (Bouwer et al. 2015). The sex pheromones altered the insect's behavior, disrupting their mating process. To determine the presence or level of infestation, the sawdust and resin on the stem or the base of the tree were used as indicators. Locations of these indicators were then measured using a handheld Global Positioning System (GPS). The dataset was then used to extract spectra from the Sentinel-2 image and develop training and testing datasets for statistical analysis.

3.2.5 Maxent modelling approach

The freely available Maxent approach (version 3.4.0) was used in this study and obtained from http://biodiversityinformatics.amnh.org/open_source/maxent/ (Phillips et al. 2017). Maxent is a machine learning technique that uses presence-only data to determine the potential spatial suitability preference of species (Ndlovu et al. 2018; Phillips et al. 2006). The model evaluates the probability of occurrence from a number of spatial environmental variables (Biber-Freudenberger et al. 2016; Matawa et al. 2016; Rebelo and Jones 2010). For Maxent to determine the suitability of a habitat and reduce uncertainty, it requires more presence information on the target species (Yi et al. 2016). The background dataset definition contributes to the model's output significantly and requires the species full environmental distribution of those areas that have been searched (Farzin et al. 2016). As a result, Maxent establishes a model with a maximum entropy in relation to the known knowledge of a species (Phillips et al. 2006; Phillips and Dudík 2008).

For this study, the data was split into 70% training and 30% test data randomly selected by the model within the study area. Sub-samples were used as the replicate run, and iterations were fixed to 500. The regularization multiplier was maintained at 4 to avoid overfitting of the test dataset (Phillips et al. 2006). The remaining model values were set to default values. A complementary log-log (clog log) output was utilized because it strongly predicts areas of moderately high output as compared to the logistic output (Phillips et al. 2017). During training, Maxent performs a

jackknife test that is used to assess the relative importance of predictor variables that explain the spatial distribution of the species and provide the performance of each variable (Phillips and Dudík 2008).

3.2.6 Model accuracy assessment

Presence data of the *C. tristis* infested locations ($n = 371$) within the compartments were randomly partitioned into two sets, 70% training data ($n = 260$) and 30% test data (n = 111). However, model performance was assessed using the area under the curve (AUC) of the receiver operating characteristics (ROC) (Hageer et al. 2017; Molloy et al. 2016; Rebelo and Jones 2010). ROC was measured by specificity as a function of sensitivity. As a result, the sensitivity which is regarded as the fraction of true-positives that are presently correctly predicted, while specificity is regarded as the fraction of false-positive absences that are correctly predicted as absences were assessed (Biber-Freudenberger et al. 2016; Germishuizen et al. 2017). Hence, the model was characterized as more accurate when the curve followed the left-hand border as compared to the right side because it attained a higher sensitivity value than a specificity value.

In that regard, the AUC ranged from 0 to 1 and the accuracy was classified as poor between 0.5 - 0.70, while 0.7 and 0.80 are good and above 0.90 are termed high (Tabet et al. 2018; Wakie et al. 2014). Additionally, the jackknife test was used to assess the contribution of each of the variable's to the model and highlighted the dominant ones (Rebelo and Jones 2010; Wang et al. 2018). Furthermore, True Skill Statistic (TSS), also known as the Hanssen–Kuipers discriminant was utilized to assess the accuracy of the model. TSS accommodates both sensitivity and specificity errors and success as a result of random guessing (Allouche et al. 2006). It ranges from − 1 to +1, whereby +1 indicates perfect agreement, whilst values of zero or less indicate random performance. The advantage of TSS as compared to Kappa is that TSS is not affected by prevalence making it a better accuracy assessment method (Liu et al. 2013; Thuiller et al. 2009). Table 3.2 shows the variables used in the three models that sought to model the habitat suitability of the *C. tristis*.

Table 3. 2: Variables used in the three analysis stages in Maxent model.

Simulation stage	Applied variables	List of variables used
	Spectral wavebands	Sentinel 2: Blue, Green, Red, Vegetation Red edge bands (band5, 6, 7 & 8A), NIR, 2 SWIR (band 11 and 12).
	Vegetation indices	NDVI, PVR, Green NDVI hyper, PSND, SR 520/670, SR 800/500, SR 774/667, SR NIR, Chlorophyll Green & ND: Salinity index.
	Spectral wavebands and vegetation indices	Combined variables.

3.3. Results

3.3.1 Prediction of the *C. tristis* using spectral bands and vegetation indices as independent datasets.

Figure 3.2 shows the prediction of habitat suitability of the *C. tristis* using Sentinel 2 spectral bands and vegetation indices as independent datasets. The red line represents the training data and the blue line represents the test dataset. Using spectral wavebands, an overall accuracy of test data = 0.898 and training data = 0.891 with a TSS value of 0.28 was achieved. While vegetation indices produced an overall accuracy of test data = 0.872 and training data = 0.875 with a TSS value of 0.32. When comparing the two models, the overall accuracy decreased by 0.026 test data and 0.04 training data. As a result, Sentinel 2 derived vegetation indices were outperformed by spectral indices in detecting and mapping the spatial distribution of the *C. tristis*.
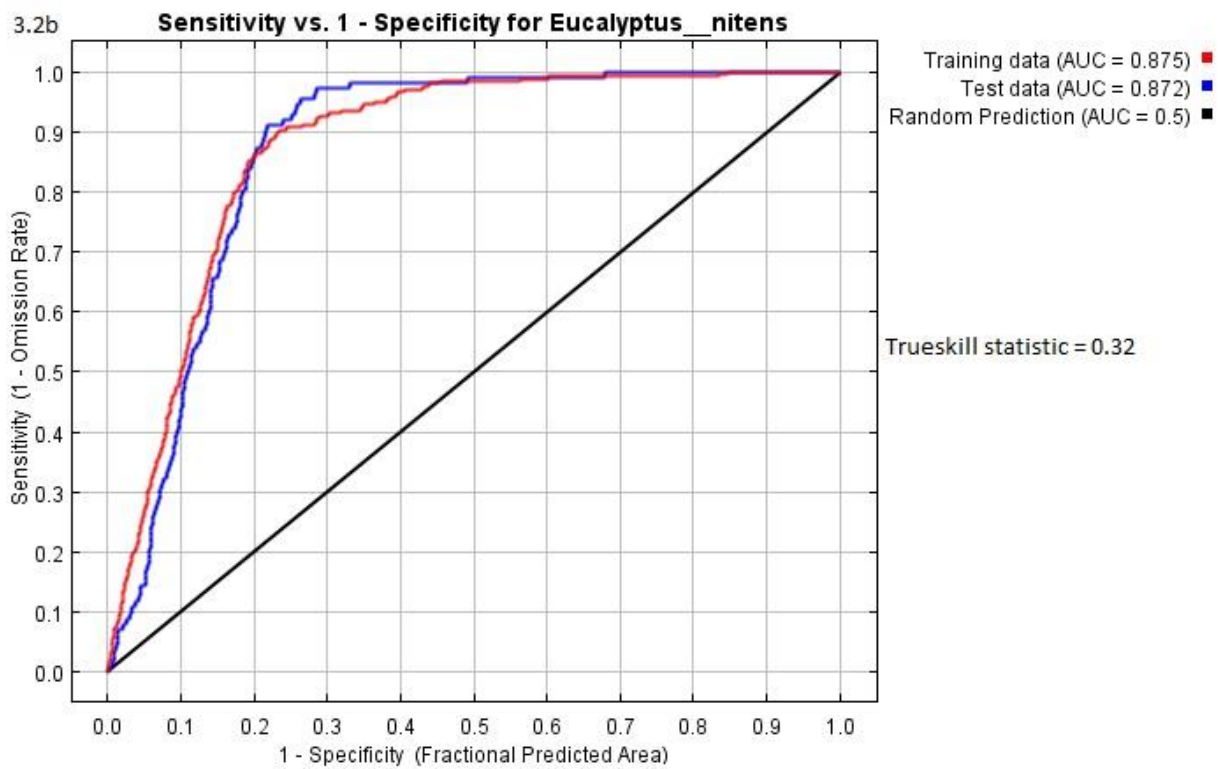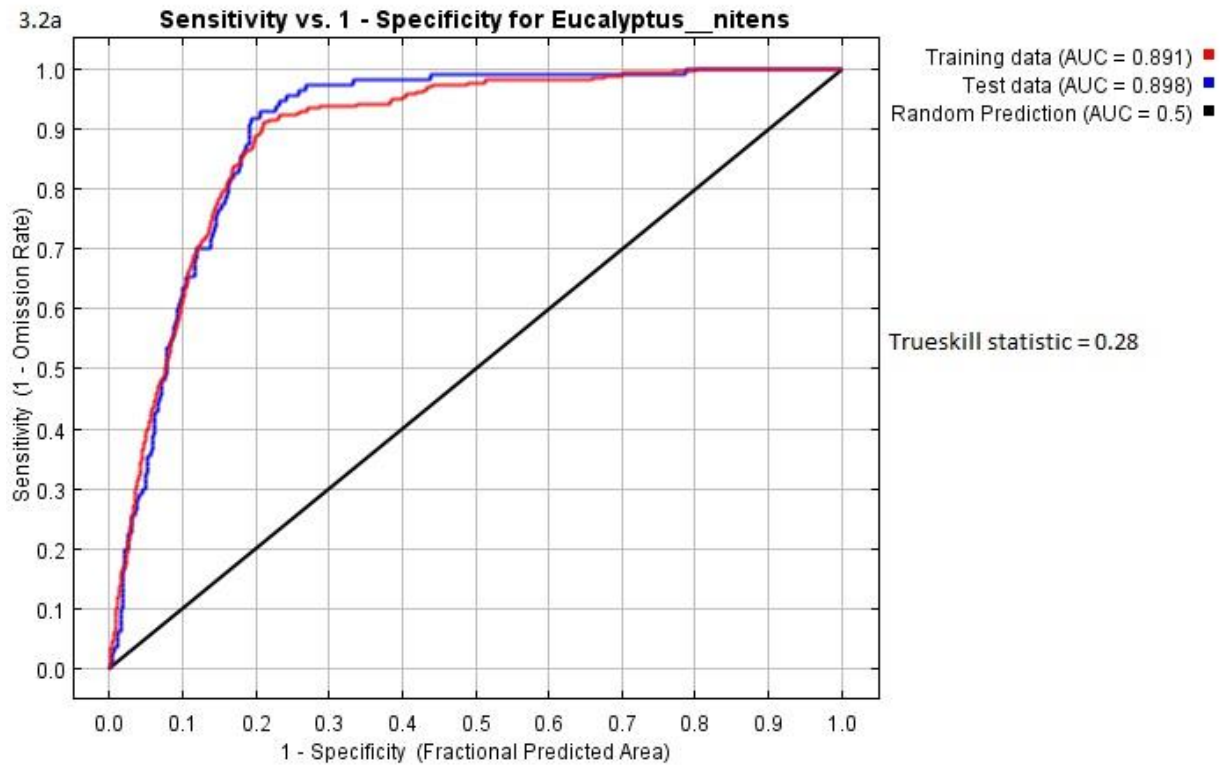
Figure 3. 2: The receiver operator characteristic curve that was used to measure the model accuracy of a) spectral wavebands and b) vegetation indices.

Respectively, the Maxent model produced a test jackknife that indicated the relative importance of each variable in the modelling process shown in Figure 3.3. In Figure 3.3 a, the most influential spectral bands in the model were Vegetation red edge (Band 5 at 705 nm), Red (Band 4 at 665 nm), Green (Band 3 at 560 nm), SWIR (Band 12 at 2190 nm), and Blue (Band 2 at 490 nm), respectively. As illustrated in Figure 3.3 b, Photosynthetic vigor ratio, Green NDVI hyper, Pigment specific normalized difference, Simple Ratio 774/667 and Salinity index were the most significant variables in the vegetation indices model, respectively.
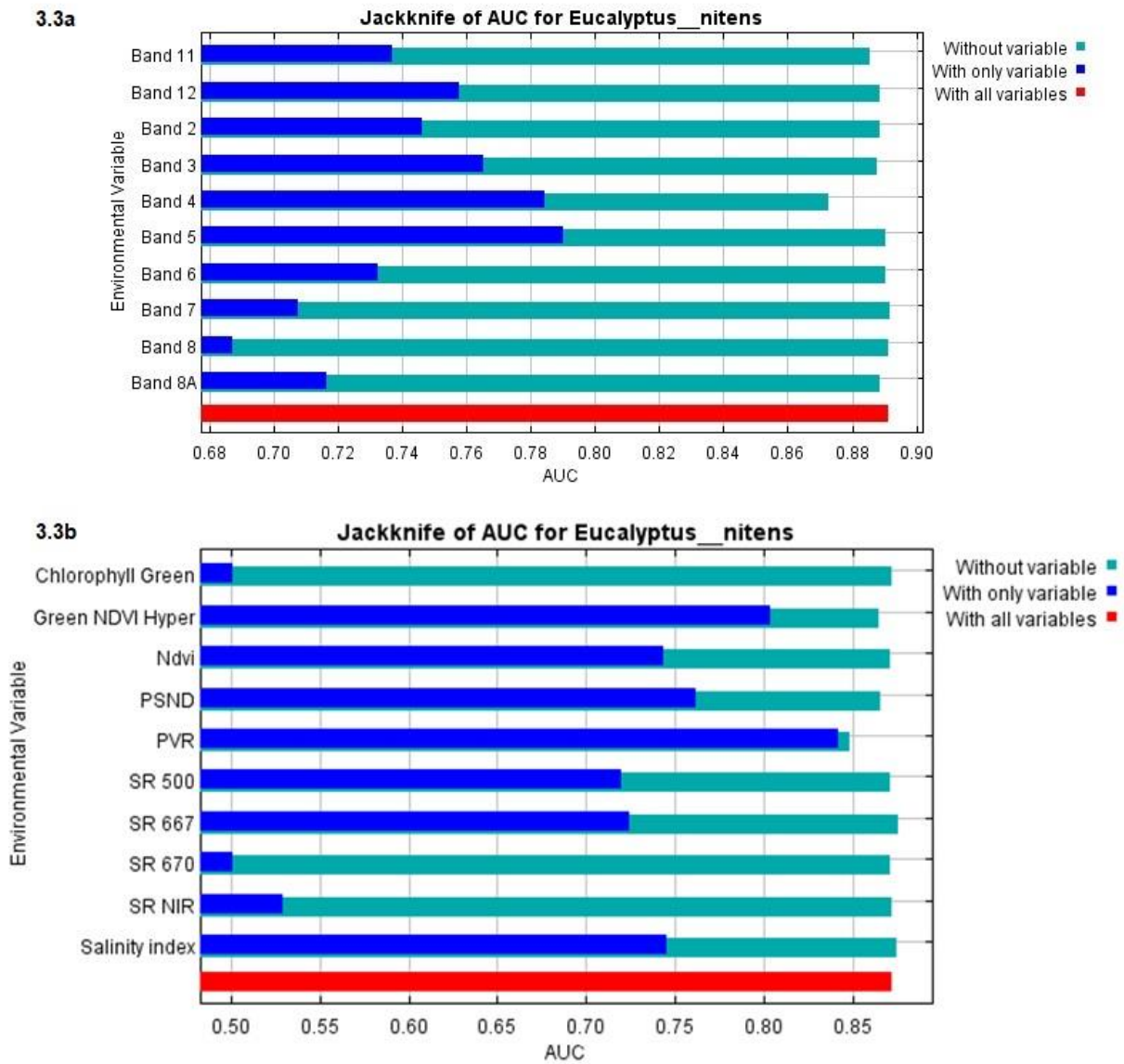
Figure 3. 3 Jackknife test variable importance graph of a) spectral wavebands and b) vegetation indices derived in modelling the spatial distribution of the *Coryphodema tristis*.

The Red edge waveband (705 nm) contributed significantly in the prediction of habitat suitability of the *C. tristis* with a variable importance of 0.814 (Fig 3.3 a). This shows the significance of the Red edge waveband in discriminating healthy and unhealthy *E. nitens* trees. In addition, the NIR (842 nm), Vegetation red edge (740 nm), Vegetation red edge (783 nm) and Vegetation red edge (865 nm) displayed a significant contribution above 0.65 each to the overall model. Moreover, the Sentinel 2 spectral wavebands in the Red (665 nm) were the second highest variable with a contribution of 0.802. The Red waveband (665 nm) recorded a decrease in the reflectance indicating the possibility of infested vegetation in the study area. Additionally, Fig 3.3 a illustrates that wavebands in the VIS had the highest contribution as Green (560 nm) was the third highest variable with a contribution of 0.793. Moreover, both SWIR bands performed well in the modelling of the *C. tristis*, SWIR (2190 nm) with a contribution of 0.784 was the fourth highest variable. The Blue (Band 2 at 490 nm) spectral waveband also yielded a contribution of 0.757 and was the fifth highest variable in the model. Sentinel 2 derived spectral bands demonstrated the high potential of predicting the likely spatial distribution of the *C. tristis*.

As shown in Fig 3.3b, PVR was the most prominent variable in the model with a contribution of 0.818. The index has the potential to detect any changes in chlorophyll content and identify weakly active vegetation affected by stress. The results showed that Green NDVI hyper was the second highest important variable with a contribution of 0.797. The test jackknife highlighted that the PSND was the third highest variable that performed well in the model with a contribution of 0.776. Both the ND: Salinity index and NDVI performed fairly equal with a contribution of 0.72. The remaining vegetation indices had a contribution above 0.50 on the independent dataset. The results obtained using Sentinel 2 derived vegetation indices alone produced slightly lower prediction accuracies when compared to those derived using the spectral bands as independent datasets.

3.3.2 Prediction of the *C. tristis* using combined variables.

The results in Fig 3.4 show prediction accuracies of both Sentinel 2 derived spectral wavebands and vegetation indices. Overall, the integration of spectral waveband information and vegetation indices produced higher prediction accuracy in this study. Using the combined data set, the model yielded high overall accuracies of 0.89 test dataset and 0.90 training dataset. Spectral wavebands

performed slightly weaker than vegetation indices. The ROC curves shows that the sensitivity value was higher than the specificity value. Therefore, the model performed above the random prediction of 0.5, indicating good results.
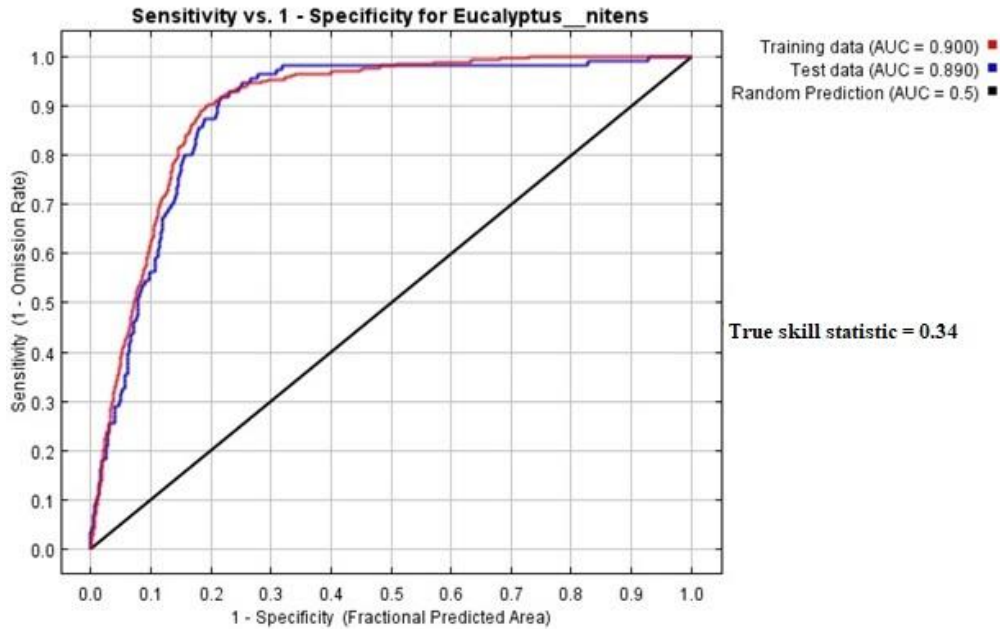

Figure 3. 4 The receiver operator characteristic curve of combined variables that was used to measure model accuracy.

Comparing the results attained in the analysis I and analysis II for each variable, it is evident that contribution accuracies did not significantly increase, indicating similar strength in the prediction of the occurrence of the *C. tristis*. Moreover, of all the three analysis conducted, PVR increased its contribution factor to 0.853 while Vegetation red edge (Band 5 at 705 nm) also increased to 0.821, resulting in vegetation indices outperforming the spectral bands. Furthermore, it was expected that vegetation indices (TSS = 0.32) would outperform the spectral wavebands (TSS = 0.28). However, the combined variables modeled produced a TSS value of 0.34, which is closer to +1 indicating a higher accuracy. Therefore, the results from the final analysis of both spectral and vegetation indices established a significant improvement on the overall contribution accuracies integrated into this study. Clearly, the results from the three models that surpassed the random prediction of 0.5, highlighted the great potential of the model to predict habitat suitability of the *C. tristis*.
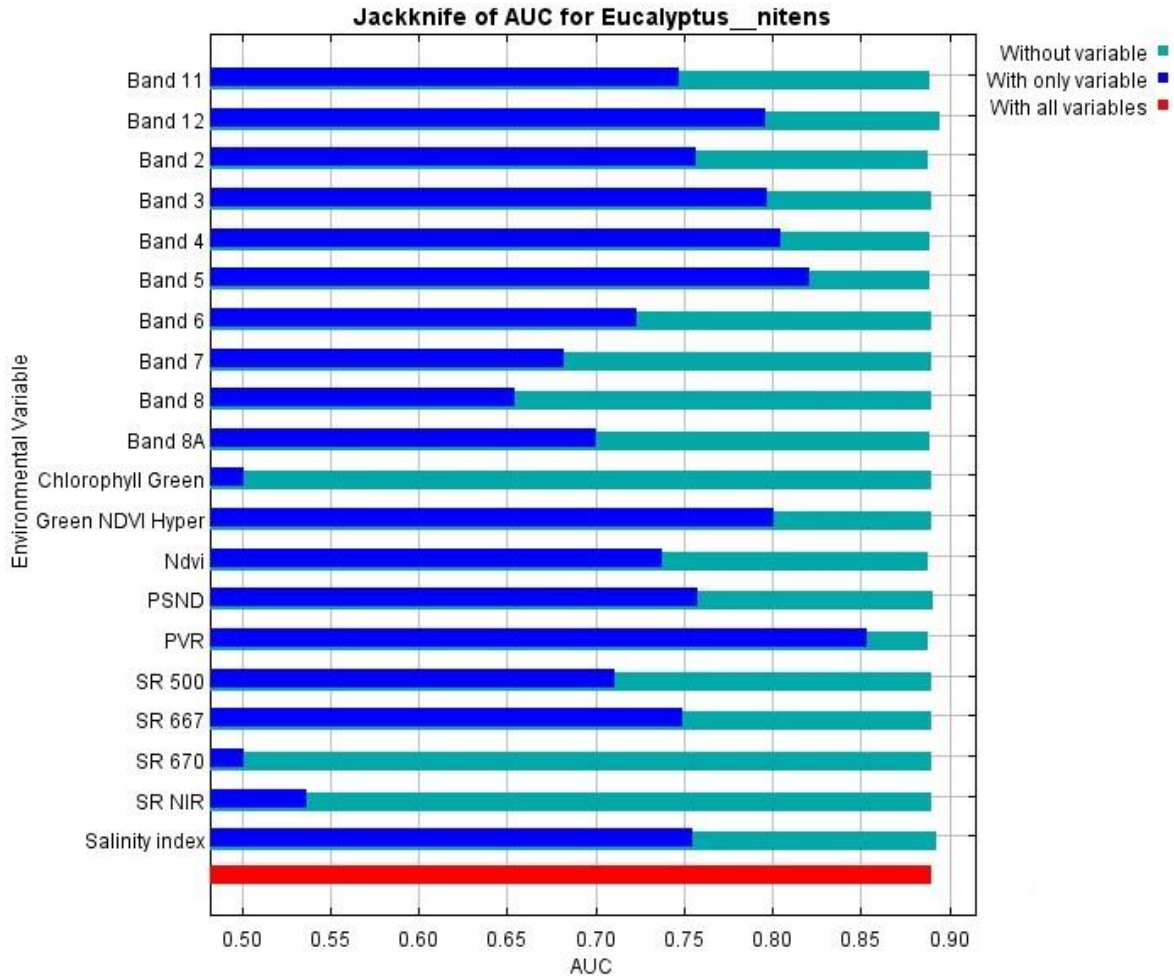
Figure 3. 5 Jackknife test variable importance graph of combined variables derived in modelling the spatial distribution of the *Coryphodema tristis*.

3.3.3 *C. tristis* spatial distribution

Fig. 3.6 illustrates the potential spatial distribution of areas highly suitable for the *C. tristis* across the study area using combined variables. The suitability preference is detected in the upper northern parts of the boundary in the Woodstock area descending towards the southern areas. In the middle of the Riverbend plantation, there are more suitable habitats as compared to unsuitable habitats of the moth. In the lowest parts of the study area it is seen that there are more unsuitable habitats. Generally, the *C. tristis* is more likely to occupy the upper parts of the study area as compared lower southern parts.
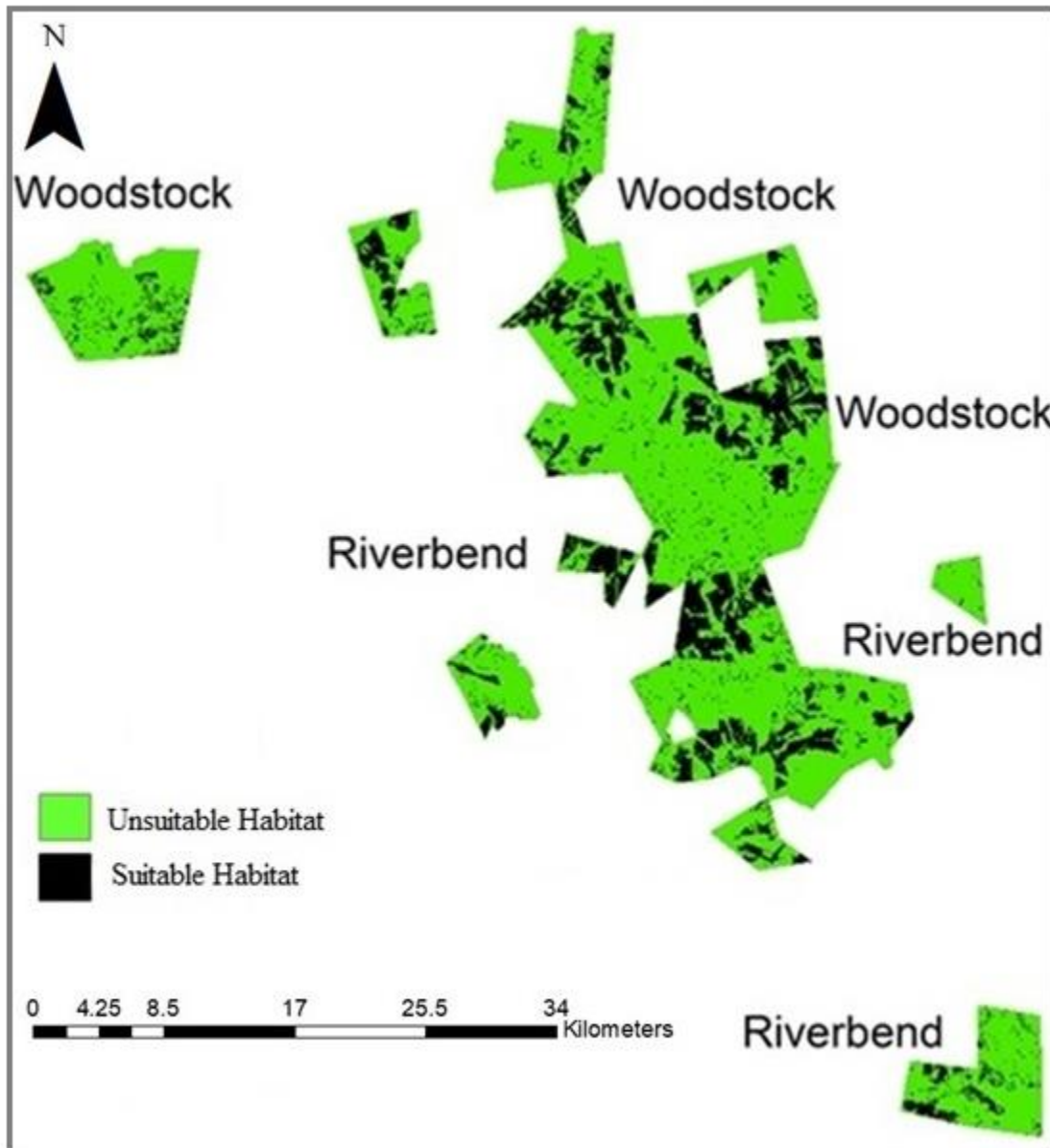
Figure 3. 6 Map showing the habitat suitability of *C. tristis* on the Lothair plantation.

3.4. Discussion

In this study, using remotely sensed data we modelled the habitat suitability of the *C. tristis* on *E. nitens* through the application of Maxent. Derived Sentinel 2 vegetation indices and spectral wavebands performed well in modelling habitat suitability of the *C. tristis*. The significance of vegetation indices as compared to spectral wavebands could be explained by their ability to detect

the health status of vegetation. The *C. tristis* damages the tree trunk and branches of *E. nitens* resulting in foliage turning black through chlorosis and it ultimately dies. As a result, there is a reduction in the absorption rates of the visible light as there are less green pigments available, which cause changes in the spectral reflection.

Results obtained in this study regarding the significance of vegetation indices concurs with previous studies of Minařík and Langhammer (2016), Metternicht (2003) and Hart and Veblen (2015). Gitelson and Merzlyak (1997) identified that healthy and unhealthy (stressed) vegetation is mostly observed in the green peak (0.665nm) and red edge (between 0.705nm and 0.783nm), hence vegetation indices such PVR and GNDVI yielded an outstanding performance in modelling the spatial distribution of the *C. tristis*. In addition, Metternicht (2003) highlighted that PVR detects any changes in the reflective properties originating from changes in chlorophyll content and produce low values for photosynthetically weakly active vegetation. Moreover, Gitelson et al. (1996a) stated that new vegetation indices such as GNDVIhyper have an extensive dynamic range as compared to NDVI, hence, they are more sensitive to chlorophyll changes. Therefore, this accounts for the high results yielded by GNDVIhyper in predicting the habitat suitability of the *C. tristis* in this study. Sanchez-Azofeifa et al. (2012) pointed out that SR and NDVI indices are used to estimate the chlorophyll concentration of vegetation as well as observing fundamental variations on leaf age, henceforth, these attributes boost its performance. Findings from this study showed that SR800/500, SR 774/667 and NDVI performed exceptionally well and can be credited to the above-mentioned. Also, a combination of two robust wavebands (NIR and Red) strengthens the probability of modelling and picking up vegetation characteristics that indicate the suitability preference of pests. Therefore, different studies have stated that the integration of NIR and Red wavebands (NDVI) and vegetation indices derived from the red edge wavebands have enhanced the prediction of pests (Lottering et al. 2016; Marx and Kleinschmit 2017; Matawa et al. 2013; Oumar and Mutanga 2013). For example, Hart and Veblen (2015) illustrated that the vegetation indices were the most important predictors to detect tree mortality caused by spruce beetle (*Dendroctonus rufipennis*) at grey-stage. Therefore, future studies should seek to improve the detection of the *C. tristis* and its associated impacts on *E. nitens* trees using powerful vegetation indices.

The results of this study also showed that the red edge wavebands were the most significant bands in determining habitat suitability of the *C. tristis* (Minařík and Langhammer 2016). There is a high correlation between red edge bands and chlorophyll content of leaves, so that the spectral signature of *E. nitens* after chlorosis due to being attacked by the *C. tristis* is easily detected on the red edge spectrum. Several studies that sought to detect and map the spatial distribution of insect pests affecting forest species confirmed that the red edge region played a significant role in the prediction of such pests (Adelabu et al. 2014; Atkinson et al. 2014; Eitel et al. 2011; Oumar and Mutanga 2013; Wulder et al. 2006). In support of these results, Oumar and Mutanga (2013), Murfitt et al. (2016) and Pietrzykowski et al. (2007) concluded that red edge bands perform slightly better than other wavebands in the detection of insect pests in forest damage. For example, Oumar and Mutanga (2013) illustrated that the red-edge and NIR wavebands of WorldView-2 were sensitive to stress-induced changes in leaf chlorophyll content, therefore, improved the potential to detect *T. peregrinus* infestations. In this regard, the Sentinel 2's red edge wavebands demonstrated its great potential in the monitoring the habitat suitability of the *C. tristis*, using its higher temporal and spatial resolution.

In determining habitat suitability of the *C. tristis*, results of this study also showed a significant potential of the SWIR region. This region has the ability to map vegetation statues, due to its sensitivity to changes in the water content of vegetation (Apan et al. 2005; Näsi et al. 2015). Generally, the larva of the *C. tristis* feeds on the cambium which is responsible for providing layers of phloem and xylem in *E. nitens* plantations. Therefore, damage to the cambium affects both phloem and xylem which ultimately alters the movement cycle of water from the roots through the trunk to the leaves of *E. nitens* trees (Näsi et al. 2015). This results in foliage and canopy water changes. It induces stress which leads to the reduction of the water content present in the main trunk and branches contributing to the change of color to black. Subsequently, the variations are then detected effectively by the SWIR portion of the electromagnetic spectrum. This then explains the optimal influence of the SWIR in detecting *E. nitens* stands that offer suitable habitat to *C. tristis*. Similarly to this study, Senf et al. (2017) accurately detected the infestations of bark beetle at the red-attack stage and grey-attack stage using the SWIR wavebands, which distinguished changes in the water content. In a similar study, Ismail et al. (2007) indicated that infestation caused by the *S. noctilio* on pine trees altered the water balance of the tree and wavebands within

the SWIR captured these changes and improved the overall prediction of the pests distribution. Furthermore, Hart and Veblen (2015) indicated that in the spruce beetle and mountain pine beetle-infested trees, reflection in the SWIR increased and decreased in the NIR due to the decrease in the foliar moisture content.

As a species distribution model (SDM), the Maxent model developed a spatial distribution map that shows the suitability preference of the *C. tristis* across the study area. High levels of suitable habitats of the moth spread across from the upper (Riverbend plantation) to the lower (Woodstock plantation) portions of the study area while medium presence along the center of the study area was recorded. The increase in suitable habitats of the moth from the upper portions to the lower portions might be characterized by the absence of natural enemies, hence this could explain the higher level of habitat suitability. The results were similar to Adam et al. (2013), which illustrated that in the upper portion of the study area there was a high presence of the *C. tristis* as compared to the lower portions indicating that the *C. tristis* is rapidly spreading. Hence, distribution maps of the *C. tristis* can help to formulate and improve on-going monitoring and management efforts to reduce the current infestation on *E. nitens* forests.

3.5. Conclusion

This study tested the utility of the new generation Sentinel 2 multispectral instrument in detecting and mapping habitat suitability of the *C. tristis* infestations on *E. nitens* plantations. Based on the findings of this study, we conclude that wavebands in the VIS, NIR and SWIR are significant in the modelling of the *C. tristis*. These three regions measure the spectral reflectance of vegetation that results in determining the amount of healthy and unhealthy vegetation. Additionally, the Red edge bands played a crucial role in the prediction of habitat suitability of the *C. tristis*. Consequently, vegetation indices derived from the VIS/NIR have demonstrated their influence in detecting changes in chlorophyll concentrations and improving the overall modelling concept in this study. Overall, these results underscore the significance of the Sentinel 2 sensor in detecting the *C. tristis* habitat suitability. The results are a platform towards the detection and mapping of the highest suitability preference of the *C. tristis*, using different multispectral sensors and their spatial resolution. The utility of remotely sensed data will improve the monitoring and management strategies used in forecasting the prevalence of pests as well as their spread.

Moreover, key stakeholders such as forest managers will be in a possession to control the damage of pests and devise proactive measures that are seemingly appropriate. This information is critical for preventing extensive damages in the forestry sector.

Chapter Four

Objectives reviewed and conclusions

4.1 Introduction

The widespread infestation caused by insect pest has become a cause for concern globally and locally. This requires an effective and efficient method to manage the damage encountered in the forest sector. The aim of the study was to assess the robustness of species distribution models in modelling the habitat suitability of *C. tristis*. Currently, SDM's have been utilized to provide current and potential distribution of insect's pests on forestry plantations and have produced vast knowledge in relation to insect pests. Several studies have strongly depended on traditional field surveys methods to identify and highlight the spatial distribution of pests using only presence and absence datasets. However, this has been rendered expensive and unreliable. The main focus of this study was to assess the application of remote sensing and species distribution models in modelling the potential habitat suitability of the *C. tristis* in Mpumalanga, South Africa. The aim of this study was to model the potential habitat suitability of the *C. tristis* (Cossid moth) in Mpumalanga, South Africa. The objectives of the study as indicated in chapter 1 were:

4.2 To evaluate the robustness of the Maxent approach in modelling the potential habitat suitability of the *C. tristis* on *E. nitens* using climatic, environmental and remotely sensed data in relation to the performance of Logistic regression.

Based on the findings in this study, Maxent outperformed the Logistic regression model in the prediction of the suitable habitats of the *C. tristis*. As a result, this indicated that presence-only datasets are effective in modelling habitat suitability of the *C. tristis*. In relation to the results, the margin of difference in both accuracies was small (10%), indicating that both models can predict the spatial distribution of the moth. However, Maxent receives more priority because it uses presence-only data, which is more accessible when compared to presence and absence data making it a cost-effective method.

The second objective intended to understand the climatic and environmental variables that influence the suitability of the *C. tristis* on *E. nitens* plantation. In relation to the results, Maxent highlighted the relative importance of variables using the Jackknife test. This is considered as an advantage of the model, because the Logistic regression doesn't indicate variable importance, but only significance in predicting the suitability of the moth. Temperature, aspect, age and elevation were identified as optimal variables that influence the suitability preference of the *C. tristis*. In addition, remotely sensed variables which include aspect and elevation derived from LIDAR increased the overall performance of the Maxent model. Hence, the inclusion of remotely sensed data into the SDMs boosts the performance of species distribution models.

4.3 To evaluate the effectiveness of the freely available Sentinel 2 multispectral imagery in detecting and mapping the habitat suitability of the *C. tristis*.

Adverse impacts on *E. nitens* commercial plantations is costly for the forestry sector as quality and quantity of yield are heavily affected. Forest stakeholders are under pressure to minimize the infestation endured from different pests and they seek to identify a fast and appropriate method to reduce the damage on commercial plantations. This study explored the utility of the Sentinel-2 multispectral instrument in modelling habitat suitability of the *C. tristis* on *E. nitens* through the application of Maxent. Based on the results, the utility of the Sentinel-2 sensor provides a cost effective opportunity for detecting and mapping the spatial distribution of the *C. tristis*. The sensor collects information using its high temporal resolution of 5 days, which allows the coverage of large areas at a short period of time. Additionally, the Sentinel 2 has a high spatial resolution with 13 wavebands which allow the construction of an image with more pixels to produce a greater detail of information. Therefore, utilizing remotely sensed data would be regarded as a cost-effective data collection method as compared to field surveys. Similarly, the jackknife from Maxent indicated the relative importance of variables showing that vegetation indices, red edge bands and wavebands determined the distribution of the moth, respectively. Our study demonstrated that the integration of remotely sensed data and Maxent improved the overall prediction of the habitat suitability of the *C. tristis*. Furthermore, this study provides a basis for identifying areas where management efforts should be focused on.

4.4 Conclusions

The major aim of this study was to assess the application of remote sensing as well as evaluating the robustness of Maxent a species distribution model in modelling the potential habitat suitability of the *C. tristis* in Mpumalanga, South Africa. Grounded in the findings, this study concludes that Maxent is an important and powerful tool in predicting the spatial distribution of the *C. tristis*. Maxent revealed the most important variables that influence the suitability preference of the moth. In addition, the application of the Sentinel 2 and LIDAR variables in modelling improved the performance of the overall models indicating their capability to offer long-term monitoring assistance on commercial forest plantations. These conclusions are coherent with the observations achieved throughout this thesis and they respond to the key research questions mentioned in the introduction chapter:

- **To what extent does the Maxent model successfully predict the potential habitats of the *C. tristis*?**

Based on the results (achieved in chapter 2 and chapter 3), Maxent successfully predicted the potential habitat of the *C. tristis* with good accuracies. All the Maxent models in this study had more than the random predictions and the difference in accucaries varied due to the different variables used in each model. Using presence-only datasets, Maxent generated predictive maps that showed the prospective habitat suitability of the moth. This indicated the significance of presence-only datasets and classified the model as a superior SDM with a good prediction performance in modelling the spatial distribution of the moth.

- **How can Maxent as a SDM identify the relevant variables that influence the suitability preference of the C. tristis on *E. nitens* plantation?**

Due to the distinctive design of the Maxent model, optimum variables that influence the suitability preference of the *C. tristis* on *E. nitens* plantation were successful identified. Maxent showed that temperature, aspect, age and elevation were the optimal variables that influenced the habitat preference of the *C. tristis* within the Mpumalanaga area. These variables corresponded with the previous studies conducted seeking to understand the moth's occurrence within the Mpumalanga area. Temperature was the most influential factor in identifying the habitat suitability of the moth and can be associated with climate change. Different studies have shown that climate change has played a critical role in the movement, shift of hosts and adaptation of insect pests across the globe.

As changes in temperature occur, they create favourable conditions which influence the suitability of the moth. Furthermore, Maxent indicated that the habitat suitability of *C. tristis* depended on the age of tree species. *E. nitens* tree species between the age of 4.5 and above were mostly vulnerable to infestation, creating a suitable habitat of the moth. Lastly, elevation that ranges between 1400 m and 1650 m was indicated as a conducive habitat for the moth. Clearly, the results show that Maxent effectively determined the variables influencing the suitability preference of the *C. tristis* on *E. nitens* plantation.

- **How effectively does the freely available Sentinel 2 sensor detect and map the *C. tristis* habitat suitability?**

Vegetation undergoing induced stress from either pests or diseases changes their spectral reflectance on the electromagnetic spectrum. As the chlorophyll content reduces, these changes are detected by the sensor. Based on the findings, the utility of Sentinel 2 derived vegetation indices, red edge bands and wavebands effectively modelled the habitat suitability of the *C. tristis* with acceptable accuracies. The combination of vegetation indices, red edge bands and wavebands provided a powerful tool in modelling the habitat suitability of the *C. tristis*. These variables managed to pick up stressed vegetation based on their spectral responses on the electromagnetic spectrum. In addition, the existence of the three red edge wavebands in the Sentinel 2 improved the capability of the sensor to detect any signs and symptoms of infestation on the on *E. nitens* plantations. Above that, the Sentinel 2 sensor has a revisit time of 5 days that allows the continous monitoring of vegetation status over a short period. This method is more cost-effective as compared to traditional field surveys and resulted in more information being collected. Hence, the freely available Sentinel 2 sensor detected the suitable habitats of the *C. tristis* and created platform towards the effective monitoring and Management of the moth.

Reference list

Adam, E., Mutanga, O., & Ismail, R. (2013). Determining the susceptibility of Eucalyptus nitens forests to Coryphodema tristis (cossid moth) occurrence in Mpumalanga, South Africa. I*nternational Journal of Geographical Information Science*, 27, 1924-1938

Adelabu, S., Mutanga, O., Adam, E., & Cho, M.A. (2013). Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image. *Journal of Applied Remote Sensing*, 7, 073480-073480

Adelabu, S., Mutanga, O., Adam, E., & Sebego, R. (2014). Spectral discrimination of insect defoliation levels in mopane woodland using hyperspectral data. IEEE *Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7, 177-186

Albaugh, J.M., Dye, P.J., & King, J.S. (2013). Eucalyptus and water use in South Africa. *International Journal of Forestry Research*, 2013

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, 1223-1232

Apan, A., Datt, B., & Kelly, R. (2005). Detection of pests and diseases in vegetable crops using hyperspectral sensing: a comparison of reflectance data for different sets of symptoms. In, Proceedings of the 2005 Spatial Sciences Institute Biennial Conference 2005: Spatial Intelligence, Innovation and Praxis (SSC2005) (pp. 10-18): Spatial Sciences Institute

Atkinson, J.T., Ismail, R., & Robertson, M. (2014). Mapping bugweed (solanum mauritianum) infestations in pinus patula plantations using hyperspectral imagery and support vector machines. IEEE J*ournal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7, 17-28

Babar, S., Amarnath, G., Reddy, C.S., Jentsch, A., & Sudhakar, S. (2012). Species distribution models: ecological explanation and prediction of an endemic and endangered plant species (Pterocarpus santalinus L.f.). Current Science, 102, 1157-1165

Bagheri, A., Fathipour, Y., Askari Seyahooei, M., & Zeinalabedini, M. (2018). Ecological Niche Modeling of Ommatissus Lybicus (Hemiptera: Tropiduchidae) De Bergevin. Annals of the *Entomological Society of America*, 111, 114-121

Baldwin, R.A. (2009). Use of maximum entropy modeling in wildlife research. *Entropy*, 11, 854-866

Berthon, K., Esperon-Rodriguez, M., Beaumont, L., Carnegie, A., & Leishman, M. (2018). Biological Conservation, 218, 154-162

Biber-Freudenberger, L., Ziemacki, J., Tonnang, H.E., & Borgemeister, C. (2016). Future risks of pest species under changing climatic conditions. *PloS one*, 11, e0153237

Blackburn, G.A. (1998). Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. In, *International Journal of Remote Sensing* (pp. 657-675)

Boreham, G. (2006). A survey of cossid moth attack in Eucalyptus nitens on the Mpumalanga

Highveld of South Africa. *Southern African Forestry Journal*, 206, 23-26

Bouwer, M.C., Slippers, B., Degefu, D., Wingfield, M.J., Lawson, S., & Rohwer, E.R. (2015). Identification of the sex pheromone of the tree infesting Cossid moth Coryphodema tristis (Lepidoptera: Cossidae). *PloS one*, 10, e0118575

Carter, G.A. (1994). Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *International Journal of Remote Sensing*, 15, 697-703

Cianci, D., Hartemink, N., & Ibáñez-Justicia, A. (2015). Modelling the potential spatial distribution of mosquito species using three different techniques. *International journal of Health Geographics*, 14, 10

DAFF (2015). A Profile Of The South African Forestry Market Value Chain

DAFF (2017). Forestry Regulation & Oversight: Facts and Figures On the Gross Domestic Product.

Degefu, D.T., Hurley, B.P., Garnas, J., Wingfield, M.J., Ahumada, R., & Slippers, B. (2013). Parallel host range expansion in two unrelated cossid moths infesting Eucalyptus nitens on two continents. *Ecological Entomology*, 38, 112-116

Deka, S., Barthakur, S., Pandey, R., Singh, M., Khetarpal, S., & Kumar, P. (2011). Potential effects of climate change on insects pest dynamics. Climate Change: Impacts and Adaptations in Crop Plants. New Delhi, Today and Tomorrow's Printers and Publishers, 301-312

Dicko, A.H., Lancelot, R., Seck, M.T., Guerrini, L., Sall, B., Lo, M., Vreysen, M.J., Lefrançois, T., Fonta, W.M., & Peck, S.L. (2014). Using species distribution models to optimize vector control in the framework of the tsetse eradication campaign in Senegal. Proceedings of the National Academy of Sciences, 111, 10149-10154

Dillon, M.E., Wang, G., & Huey, R.B. (2010). Global metabolic impacts of recent climate warming. *Nature*, 467, 704

Donatelli, M., Magarey, R.D., Bregaglio, S., Willocquet, L., Whish, J.P.M., & Savary, S. (2017). Modelling the impacts of pests and diseases on agricultural systems. *Agricultural Systems, 155, 213-224*

Eitel, J.U., Vierling, L.A., Litvak, M.E., Long, D.S., Schulthess, U., Ager, A.A., Krofcheck, D.J., & Stoscheck, L. (2011). Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. *Remote Sensing of Environment*, 115, 3640-3646

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, 677-697

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, 43-57

FAO (2007). Forest Health & Biosecurity Working Papers:Overview of Forest Pests South Africa

Farzin, S., Lalit, K., & Mohsen, A. (2016). A comparison of absolute performance of different

correlative and mechanistic species distribution models in an independent area. *Ecology and Evolution*, 6, 5973-5986

Fick, S.E., & Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302-4315

Fithian, W., & Hastie, T. (2013). Finite-sample equivalence in statistical models for presence-only data. *The Annals of Applied Statistics*, 7, 1917

Fourcade, Y., Engler, J.O., Rödder, D., & Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PloS one*, 9, e97122

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., & Gaudel-Vacaresse, A. (2017). Copernicus Sentinel-2A calibration and products validation status. *Remote Sensing*, 9, 584

Gebeyehu, S., Hurley, B.P., & Wingfield, M.J. (2005). A new lepidopteran insect pest discovered on commercially grown Eucalyptus nitens in South Africa: research in action. South African *Journal of Science*, 101, 26-28

Germishuizen, I., Peerbhay, K., & Ismail, R. (2017). Modelling the susceptibility of pine stands to bark stripping by Chacma baboons (Papio ursinus) in the Mpumalanga Province of South Africa. *Wildlife Research*, 44, 298-308

Gitelson, A.A., Kaufman, Y.J., & Merzlyak, M.N. (1996a). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58, 289-298

Gitelson, A.A., Keydan, G.P., & Merzlyak, M.N. (2006). Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. G*eophys. Res. Lett.*, 33, L11402

Gitelson, A.A., & Merzlyak, M.N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing*, 18, 2691-2697

Gitelson, A.A., Merzlyak, M.N., & Lichtenthaler, H.K. (1996b). Detection of Red Edge Position and Chlorophyll Content by Reflectance Measurements Near 700 nm. *Journal of Plant Physiology*, 148, 501-508

Gribko, L.S., Liebhold, A.M., & Hohn, M.E. (1995). Model to Predict Gypsy Moth (Lepidoptera: Lymantriidae) Defoliation Using Kriging and Logistic Regression. *Environmental Entomology*, 24, 529-537

Gumpertz, M.L., Wu, C.-t., & Pye, J.M. (2000). Logistic Regression for Southern Pine Beetle Outbreaks with Spatial and Temporal Autocorrelation. *Forest Science*, 46, 95-107

Hageer, Y., Esperón-Rodríguez, M., Baumgartner, J.B., & Beaumont, L.J. (2017). Climate, soil or both? Which variables are better predictors of the distributions of Australian shrub species? PeerJ, 5, e3446

Hart, S.J., & Veblen, T.T. (2015). Detection of spruce beetle-induced tree mortality using high-

and medium-resolution remotely sensed imagery. *Remote Sensing of Environment*, 168, 134-145

Hojas-Gascon, L., Belward, A., Eva, H., Ceccherini, G., Hagolle, O., Garcia, J., & Cerutti, P. (2015). Potential improvement for forest cover and forest degradation mapping with the forthcoming Sentinel-2 program. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40, 417

Immitzer, M., Vuolo, F., & Atzberger, C. (2016). First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sensing*, 8, 166

Ismail, R., Mutanga, O., & Bob, U. (2007). Forest health and vitality: the detection and monitoring of Pinus patula trees infected by Sirex noctilio using digital ultispectral imagery (DMSI). *Southern Hemisphere Forestry Journal*, 69, 39

Jaworski, T., & Hilszczański, J. (2013). The effect of temperature and humidity changes on insects development their impact on forest ecosystems in the expected climate change. *Forest Research Papers*, 74, 345-355

Kocmánková, E., Trnka, M., Juroch, J., Dubrovský, M., Semerádová, D., Možný, M., & Žalud, Z. (2009). Impact of climate change on the occurrence and activity of harmful organisms. *Plant Protection Science*, 45

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

Kutywayo, D., Chemura, A., Kusena, W., Chidoko, P., & Mahoya, C. (2013). The impact of climate change on the potential distribution of agricultural pests: the case of the coffee white stem borer (Monochamus leuconotus P.) in Zimbabwe. *PloS one,* 8, e73432

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, 40, 778-789

Lottering, R., & Mutanga, O. (2016). Optimising the spatial resolution of WorldView-2 pan-sharpened imagery for predicting levels of Gonipterus scutellatus defoliation in KwaZulu-Natal, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing,* 112, 13-22

Lottering, R., Mutanga, O., & Peerbhay, K. (2016). Detecting and mapping levels of Gonipterus scutellatus-induced vegetation defoliation and leaf area index using spatially optimized vegetation indices. *Geocarto International*, 1-16

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, 66

Marx, A., & Kleinschmit, B. (2017). Sensitivity analysis of RapidEye spectral bands and derived vegetation indices for insect defoliation detection in pure scots pine stands. *iForest-Biogeosciences and Forestry*, 10, 659

Matawa, F., Murwira, A., Zengeya, F.M., & Atkinson, P.M. (2016). Modelling the spatial-temporal distribution of tsetse (Glossina pallidipes) as a function of topography and

vegetation greenness in the Zambezi Valley of Zimbabwe. *Applied Geography*, 76, 198-206

Matawa, F., Murwira, K.S., & Shereni, W. (2013). Modelling the distribution of suitable Glossina Spp. habitat in the North Western parts of Zimbabwe using remote sensing and climate data. *Geoinform Geostast Overv*, 1-9

Meier, E.S., Kienast, F., Pearman, P.B., Svenning, J.C., Thuiller, W., Araújo, M.B., Guisan, A., & Zimmermann, N.E. (2010). Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, 33, 1038-1048

Metternicht, G. (2003). Vegetation indices derived from high-resolution airborne videography for precision crop management. *International Journal of Remote Sensing*, 24, 2855-2877

Michael, K., & Warren, P. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, 12, 334-350

Minařík, R., & Langhammer, J. (2016). Use of A Multispectral Uav Photogrammetry for Detection and Tracking of Forest Disturbance Dynamics. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 41

Molloy, S.W., Davis, R.A., & van Etten, E.J. (2016). Incorporating field studies into species distribution and climate change modelling: a case study of the koomal Trichosurus vulpecula hypoleucus (Phalangeridae). *PloS one*, 11, e0154161

Murfitt, J., He, Y., Yang, J., Mui, A., & De Mille, K. (2016). Ash decline assessment in emerald ash borer infested natural forests using high spatial resolution images. *Remote Sensing*, 8, 256

Näsi, R., Honkavaara, E., Lyytikäinen-Saarenmaa, P., Blomqvist, M., Litkey, P., Hakala, T., Viljanen, N., Kantola, T., Tanhuanpää, T., & Holopainen, M. (2015). Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree-level. *Remote Sensing*, 7, 15467-15493

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

Neupane, R.P., Sharma, K.R., & Thapa, G.B. (2002). Adoption of agroforestry in the hills of Nepal: a logistic regression analysis. *Agricultural Systems*, 72, 177-196

Ng, W.-T., Rima, P., Einzmann, K., Immitzer, M., Atzberger, C., & Eckert, S. (2017). Assessing the Potential of Sentinel-2 and Pléiades Data for the Detection of Prosopis and Vachellia spp. in Kenya. *Remote Sensing*, 9, 74

Otunga, C., Odindi, J., Mutanga, O., Adjorlolo, C., & Botha, J. (2018). Predicting the distribution of C3 (Festuca spp.) grass species using topographic variables and binary logistic regression model. *Geocarto International*, 33, 489-504

Oumar, Z., & Mutanga, O. (2011). The potential of remote sensing technology for the detection and mapping of Thaumastocoris peregrinus in plantation forests. *Southern Forests*, 73, 23-31

Oumar, Z., & Mutanga, O. (2013). Using WorldView-2 bands and indices to predict bronze bug (Thaumastocoris peregrinus) damage in plantation forests. *International Journal of Remote*

Sensing, 34, 2236-2249

Pause, M., Schweitzer, C., Rosenthal, M., Keuck, V., Bumberger, J., Dietrich, P., Heurich, M., Jung, A., & Lausch, A. (2016). In situ/remote sensing integration to assess forest health—A review. *Remote Sensing*, 8, 471

Péré, C., Jactel, H., & Kenis, M. (2013). Response of insect parasitism to elevation depends on host and parasitoid life-history strategies. *Biology letters*, 9, 20130028

Petzoldt, C., & Seaman, A. (2006). Climate change effects on insects and pathogens. Climate change and agriculture: Promoting practical and profitable responses, 3, 6-16

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, 887-893

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

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

Pietrzykowski, E., Sims, N., Stone, C., Pinkard, L., & Mohammed, C. (2007). Predicting Mycosphaerella leaf disease severity in a Eucalyptus globulus plantation using digital multi-spectral imagery. *Southern Hemisphere Forestry Journal*, 69, 175-182

Q., W.Y., F., M.J., Q., L.X., F., W.Y., S., C., T., X.A., F., Y.S., X., D.B., X., Z.W., X., Q.Y., F., X., Y., Z.Z., M., Z.X., Y., J.J., & P., D.Z. (2017). The distribution of Athetis lepigone and prediction of its potential distribution based on GARP and MaxEnt. *Journal of Applied Entomology*, 141, 431-440

Radoux, J., Chomé, G., Jacques, D.C., Waldner, F., Bellemans, N., Matton, N., Lamarche, C., d'Andrimont, R., & Defourny, P. (2016). Sentinel-2's potential for sub-pixel landscape feature detection. *Remote Sensing*, 8, 488

Ramanagouda, S., Kavitha Kumari, N., Vastrad, A., Basavanagoud, K., & Kulkarni, H. (2010). Potential alien insects threatening eucalyptus plantations in India. Karnataka Journal of *Agricultural Sciences*, 23

Rebelo, H., & Jones, G. (2010). Ground validation of presence-only modelling with rare species: a case study on barbastelles Barbastella barbastellus (Chiroptera: Vespertilionidae). *Journal of Applied Ecology*, 47, 410-420

Remya, K., Ramachandran, A., & Jayakumar, S. (2015). Predicting the current and future suitable habitat distribution of Myristica dactyloides Gaertn. using MaxEnt model in the Eastern Ghats, India. *Ecological Engineering*, 82, 184-188

Richardson, A.D., Duigan, S.P., & Berlyn, G.P. (2002). An evaluation of noninvasive methods to estimate foliar chlorophyll content. *New Phytologist*, 153, 185-194

Rullan-Silva, C., Olthoff, A., de la Mata, J.D., & Pajares-Alonso, J. (2013). Remote monitoring of forest insect defoliation-A Review. *Forest Systems*, 22, 377-391

Sanchez-Azofeifa, A., Oki, Y., Fernandes, G.W., Ball, R.A., & Gamon, J. (2012). Relationships

between endophyte diversity and leaf optical properties. Trees, 26, 291-299

Senf, C., Seidl, R., & Hostert, P. (2017). Remote sensing of forest insect disturbances: Current state and future directions. *International Journal of Applied Earth Observation and Geoinformation*, 60, 49-60

SETA (2014). Forestry and Wood products Sector: A profile of the forestry and wood products sub-sector.

Swain, T.-L., & Gardner, R.A. (2003). A summary of current knowledge of cold tolerant eucalypt species (CTE's) grown in South Africa. University of Natal, Institute for Commercial Forestry Research

Tabet, S., Belhemra, M., Francois, L., & Arar, A. (2018). Evaluation by prediction of the natural range shrinkage of Quercus ilex L. in eastern Algeria. *Forestist* (eski adıyla İstanbul Üniversitesi Orman Fakültesi Dergisi), 68, 1-1

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M.B. (2009). BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-373

Wakie, T.T., Evangelista, P.H., Jarnevich, C.S., & Laituri, M. (2014). Mapping current and potential distribution of non-native Prosopis juliflora in the Afar region of Ethiopia. *PloS one, 9, e112854*

Wang, R., Li, Q., He, S., Liu, Y., Wang, M., & Jiang, G. (2018). Modeling and mapping the current and future distribution of Pseudomonas syringae pv. actinidiae under climate change in China. *PloS one*, 13, e0192153

Wingfield, M., Roux, J., Coutinho, T., Govender, P., & Wingfield, B. (2001). Plantation disease and pest management in the next century. *The Southern African Forestry Journal*, 190, 67-71

Wingfield, M., Slippers, B., Hurley, B., Coutinho, T., Wingfield, B., & Roux, J. (2008). Eucalypt pests and diseases: growing threats to plantation productivity. Southern Forests: *A Journal of Forest Science*, 70, 139-144

Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dormann, C.F., Forchhammer, M.C., Grytnes, J.A., & Guisan, A. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological reviews*, 88, 15-30

Wulder, M.A., Dymond, C.C., White, J.C., Leckie, D.G., & Carroll, A.L. (2006). Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. *Forest Ecology and Management*, 221, 27-41

Xing, Z., Zhang, L., Wu, S., Yi, H., Gao, Y., & Lei, Z. (2017). Niche comparison among two invasive leafminer species and their parasitoid Opius biroi: implications for competitive displacement. *Scientific Reports*, 7, 4246

Yi, Y.-j., Cheng, X., Yang, Z.-F., & Zhang, S.-H. (2016). Maxent modeling for predicting the potential distribution of endangered medicinal plant (H. riparia Lour) in Yunnan, China. *Ecological Engineering*, 92, 260-269

Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H., & Sampson, P.H. (2001). Scaling-up and model inversion methods with narrow-band optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 1491−1507.