DEVELOPMENT OF A HABITAT SUITABILITY MODEL TO DETERMINE THE POTENTIAL DISTRIBUTION OF KLIPSPRINGER (*OREOTRAGUS OREOTRAGUS SUBSP. OREOTRAGUS*) IN TABLE MOUNTAIN NATIONAL PARK

BY

RICHARDT JOHN SMITH

Thesis submitted in fulfilment of the requirements for the degree

Master of Technology: Nature Conservation

In the Faculty of Applied Sciences

at the

CAPE PENINSULA UNIVERSITY OF TECHNOLOGY

Supervisor: Prof. R.G. Snyman
Co-supervisor: Dr. H.M. de Klerk

Cape Town

June 2015

CPUT copyright information
The thesis may not be published either in part (in scholarly, scientific or technical journals), or as a whole (as a monograph), unless permission has been obtained from the University.

DECLARATION

I, Richardt John Smith, declare that the contents of this thesis represent my own unaided work, and that the thesis has not previously been submitted for academic examination towards any qualification. Furthermore, it represents my own opinions and not necessarily those of the Cape Peninsula University of Technology.

Signed ……………………………………. **Date …………………………………**

ABSTRACT

The klipspringer (*Oreotragus oreotragus subsp. oreotragus)* population became extinct on the Cape Peninsula in 1930. Being re-introduced into Table Mountain National Park (TMNP) in 1999 it became one of the species of special conservation concern to monitor in the Park. Most klipspringer territories are known by Park management but the distribution of all potentially suitable habitats for this species in the Park is not known. The main aim of this study is to produce a distribution range map that is representative of all potentially suitable habitats for the klipspringer within TMNP, through the use of a species distribution modelling tool.

Since only presence data were available for this study, a popular presence-only modelling tool namely maximum entropy (MaxEnt) was used. The use of MaxEnt in species distribution modelling has become popular as it has proven to provide robust predictions of a species' geographic distribution. Klipspringer occurrence data and five environmental variables namely altitude, slope, aspect, vegetation, and distance to urban edge were used as model input. Occurrence data were sourced through existing databases and employing a stratified random sampling technique of dividing the Park into different habitat subtypes to survey the Park for more klipspringer occurrences. These habitat subtypes consisted of a variety of vegetation communities or vegetation types and altitudinal and slope ranges available in the Park. Grid size for all the raster layers used was 10x10 m. Spatial filtering of one point per 100 m^2 grid was used to eliminate clumping of points. Six models were run at different regularisation multiplier (RM) values namely 0.25, 0.5, default (1), 2, 4 and 7. To assist in better understanding of the spatial extent of the occurrence data and the areas inhabited by the klipspringer, home range analyses were carried out. This was done through kernel density estimation in the Geospatial Modelling Environment (GME). All six bandwidth parameters in GME namely smoothed cross validation (SCV), biased cross-validation (BCV), a second BCV algorithm, plug-in estimator, least squares cross validation and the likelihood cross validation (CVh) were tested.

The smoothed cross validation and likelihood cross validation bandwidth algorithms provided the best visual output of klipspringer home ranges and territories. Home range sizes from the SCV output ranged from about 3 – 11 ha across the study area, and home range size for the CVh output ranged from 0.6 – 2.5 ha. The output from the CVh algorithm was interpreted as territories rather than home ranges, as it is based on a univariate kernel unlike, the SCV algorithm that produces rotated bivariate kernels.

The default regularisation multiplier of 1 provided the best probability distribution output, whilst values lower than the default tended to underestimate the prediction and those values higher than the default were tending towards overestimations. Response curves for the default RM also gave the most ecologically meaningful responses of the klipspringer to each environmental variable. Model evaluation in the form of area under the receiver operating characteristic curve (ROC AUC) showed that all models performed well. Therefore, the choice of the “best” model was based on the ability to provide ecological interpretation, on the shape of the response curve and the probability distribution maps. Consequently, the default RM model was considered the best, with an AUC score of 0.903. Altitude and vegetation contributed the most to suitable habitat and therefore indicates that klipspringer in the Park do prefer high altitudinal areas with the right vegetation to feed on. Suitable altitudinal ranges are from 400 m.a.s.l. and higher and ericaceous fynbos is the most preferred vegetation community. Slope, aspect and distance to urban edge played a less important role in suitable klipspringer habitat.

The probability map and an additional binary map produced at the 10 percentile training logistic presence threshold showed that suitable habitat for the klipspringer occurs in all three sections of the Park in different proportions. These maps can be used by Park management to prioritise conservation efforts and future re-introductions.

ACKNOWLEDGEMENTS

I wish to thank:

- Professor Reinette Snyman of Cape Peninsula University of Technology (CPUT) and Dr Helen de Klerk of Stellenbosch University, for their guidance and assistance throughout this research.
- Chad Cheney (Table Mountain National Park: Planning Department) for the occurrence data and shapefiles provided as well as useful discussions on klipspringers in TMNP.
- Zishan Ebrahim (Cape Research Centre) for shapefiles provided as well as guidance in some GIS issues in the early stages of this research.
- City of Cape Town: GIS Department for the provision of the digital elevation model.
- All members of the Stellenbosch University Geoinformatics research group of 2013 – 2015 who provided a platform for sharing ideas, discussion and solving GIS related problems.
- Mr Andrew Rand (CPUT) for his guidance on GIS related issues.
- Mr Guy Palmer from CapeNature for a useful discussion on klipspringer ecology and klipspringers in CapeNature reserves.
- Ms Corlie Hugo from CapeNature for a useful discussion on klipspringers in CapeNature reserves
- The Senior Section Rangers, Section Rangers and other Rangers of the different sections of TMNP that assisted during field visits either through providing transport or assistance in occurrence data collection.
- Dr Tony Rebelo for providing klipspringer occurrence points recorded in the Ispot database.
- Ms Heather Edwards for providing klipspringer occurrence points collected in the southern section of the Park.
- My family away from home at the His People City Bowl church for their tremendous support and encouragement throughout the duration of this research.
- My sister Anneline Felicity Smith and dad, Johannes Smith for their words of encouragement to not give up and keep on pushing.

The financial assistance of the National Research Foundation towards this research is acknowledged. Opinions expressed in this thesis and the conclusions arrived at, are those of the author, and are not necessarily to be attributed to the National Research Foundation.

DEDICATION

This thesis is dedicated to my late mom, Anna Wilhelmina Smith and my baby sister Emmerencia Venecia Smith.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

1. Terms/Definitions

Area of distribution: The geographic range of a species i.e. areas in a landscape where a species is likely to occur (Brown *et al.,* 1996).

Ecological niche model: Fundamentally the same as a species distribution model (Elith & Leathwick, 2009; Sillero, 2011).

Environmental space: The combination of environmental predictor variables that define an area (Elith & Leathwick, 2009), as represented in a conceptual ecosystem.

Geographical space: Represented by latitude and longitude to identify location, allowing one to link a location on a landscape to two-dimensional coordinates or in the form of a digital elevation model to distinguish geographical features (Elith & Leathwick, 2009).

Habitat suitability map: The output of niche models and distribution models (those areas with the right environmental conditions for a species to survive).

Niche: An area in an ecosystem where a species has found a combination of optimal environmental conditions to fulfil its needs to live in and reproduce, together with its effects on those environmental conditions (Chase & Leibold, 2003).

Species distribution model: A statistical tool that combines GPS locality data with data on environmental variables, to produce predictions of a species' geographic distribution (Guisan & Thuiller, 2005; Elith & Leathwick, 2009).

Home range: That area an animal population would occupy and use for their daily activities like feeding and reproduction (Getz *et al.*, 2007).

Utilisation Distribution: A percentage of area an animal utilises e.g. a 50% area of core activity or 95% area of general use, represented by isopleths/contours that demarcate those areas utilised by the individuals of a population (Getz *et al.*, 2007).

2. Acronyms

AUC – Area under the receiving operating characteristic (ROC) curve.

BAM diagram – Representation of biotic, abiotic, & movement characteristics in environmental space.

BIOCLIM – Bioclimatic Envelope Algorithm

BRT – Boosted Regression Trees

CSV – Comma Separated Value

DEM – Digital Elevation Model

ENM – Ecological Niche Model/ling

GAM – General Additive Model

GARP – Genetic Algorithm for Rule-set Prediction

GBIF – Global Biodiversity Information Facility

GIS – Geographic Information System.

GLM – Generalised Linear Model

GPS – Global Positioning System

GME – Geospatial Modelling Environment

GUI – Geographic User Interface

KDE – Kernel Density Estimation

MAR – Mean Annual Rainfall

MaxEnt – Maximum Entropy

RM – Regularisation Multiplier

SANPARK – South African National Parks

SDM – Species Distribution Model/ling

TMNP – Table Mountain National Park

UD – Utilisation Distribution

3 – Terms used interchangeably

Occurrence points, locality points, presence points – all referring to geo-referenced GPS points that indicate where the klipspringer was recorded in the study area.

Probability of presence – probability of occurrence – probability of suitable habitat

Presence – suitable

Absence – unsuitable

Isopleth – contour

Training – calibration

Vegetation community subtype – habitat subtype

CHAPTER ONE: GENERAL INTRODUCTION AND LITERATURE REVIEW

1.1 Background and Introduction

The klipspringer (*Oreotragus oreotragus*) population of the Cape Peninsula became extinct in about 1930 (Skead, 2011). This extinction was rumoured to be the result of excessive hunting and the extensive habitat use by the Himalayan tahr (*Hemitragus jemlahicus)* (Skead, 2011). An effort to re-establish a viable population of klipspringer on the Cape Peninsula mountains, of which the boundaries of Table Mountain National Park (TMNP) encompass this mountain range, was initiated in 1999 (Cheney pers. comm. 2012[1]). A total number of 56 klipspringers were re-introduced in the Park in the period 1999 – 2005. Of these, five pairs were released in June 1999 (SANParks Scientific Services, 2010), nine individuals in July 1999 (SANParks Scientific Services, 2010), 25 individuals in October 2004 (Marell, n.d.) and six pairs in October 2005 (Marell, n.d.). All these individuals were supplied by CapeNature (SANParks Scientific Services, 2010; Marell, n.d.).

One of the agreements between South African National Parks (SANParks) and CapeNature was to report on the success of the re-introduction programme (Cheney pers. comm. 2012; Palmer pers. comm. 2013[2]). Consequently, a postgraduate study was initiated in 2005 to determine the habitat availability and carrying capacity for the klipspringer in TMNP (Cheney pers. comm. 2012). This study was however abandoned in 2007, making the continuation of this research vital. This initial study was adapted in the current study and the focus shifted to habitat availability, which is in essence habitat suitability, and the distribution of this potential suitable habitat for the klipspringer in the Park. The completed research would provide Park management with crucial information on how the released animals dispersed, and what habitat is probably suitable for klipspringer. This knowledge in return can guide Park management in their management decision-making.

Five of the introduced individuals were originally collared with global positioning system (GPS) collars to track their movement, however, some of them have died over the years and for others the operating machines of the collars have since stopped functioning (Cheney pers. comm. 2012). As a result, Park management have lost track of the movement of the collared individuals. Consequently, a follow-up study on where klipspringer individuals have established themselves was needed together with what their preferred habitat conditions are in TMNP, as well as an approximation of where those suitable habitat conditions exist in the Park. Such knowledge would be useful for the long-term conservation of klipspringers in the Park. By constructing a Geographic Information System (GIS) model which incorporates presence data and key environmental factors, it can be

---

[1] Mr Chad Cheney, Environmental Information Systems Officer – Table Mountain National Park.

[2] Mr Guy Palmer, Scientific Manager: Biodiversity – Cape Nature Scientific Services.

determined where these suitable habitats occur across the boundaries of the Park (Gross *et al.,* 2002; Rotenberry *et al.,* 2006).

Through the course of time since the re-introduction, most occupied territories of the klipspringers have become known to Park management, however not all suitable habitats across the Park have been identified through modelling and mapping. By employing species-specific habitat suitability models the potential distribution of suitable habitats for the klipspringers can be estimated. A popular tool for such modelling is the maximum entropy (MaxEnt) tool which will be used in this study (Phillips *et al.,* 2006; Phillips & Dudik, 2008). Ultimately this modelling process will assist in understanding how klipspringers utilise their habitat and have adapted in a fynbos habitat, as no published research is available for klipspringers in fynbos. In addition, home range analysis can also be performed to provide further understanding of how the klipspringers in TMNP utilise their habitat (Worton, 1989). Therefore, this study used home range analyses to derive an estimate of the minimum and maximum home range size for the klipspringers of TMNP. Home range studies for klipspringers in the Fynbos Biome are also lacking, as no literature on such analyses could be found on scientific databases. Home range analysis, however, will not be the primary focus of this study but rather habitat suitability modelling.

1.2 Klipspringer ecology

1.2.1 General description and geographic distribution

Klipspringer is an Afrikaans word that can be directly translated as “rock jumper”, although the English name is not in use. This common name is attributed to the general habitat klipspringers live in, namely rocky areas. Their hooves are specially adapted for great agility on rocky slopes (Skinner & Chimimba, 2005).

Amongst African antelope, the klipspringer (*Oreotragus oreotragus*), is considered to have the widest geographical distribution which stretches from the southern Cape mountains of South Africa northwards to the highlands of Ethiopia (Norton, 1980). The specific subspecies under study is *Oreotragus oreotragus subsp. oreotragus* (Appendix A) which occurs in three provinces of South Africa namely Northern Cape, Western Cape and Eastern Cape (Skinner & Chimimba, 2005).

O. oreotragus subsp. oreotragus has a yellow coat colour with brown speckles. Differences in coat colour amongst subspecies exist due to their area of origin (Skinner & Chimimba, 2005). A historical account of the early records published in Skead (2011), describe this animal’s elusive behaviour. The presence of this species has been observed on the Peninsula Mountains long before the establishment of European settlements on the lowlands of Cape Town. These historical records in

Skead (2011) make reference of the species being observed on Table Mountain, throughout Constantia and between Hout Bay and Cape Point.

Klipspringers are territorial and each territory is commonly occupied by a breeding pair or family group of three (Skinner & Chimimba, 2005). Klipspringers generally form a long-lasting pair bond (Dunbar & Dunbar, 1980; Estes, 1991), and pairs are usually in close proximity to each other during their normal daily activities (Dunbar & Dunbar, 1980; Tilson, 1980). During the day a dominant male can often be seen standing on a vantage point within his territory (Skinner & Smithers, 1990). Furthermore, the boundaries of territories are often marked with dung heaps and the territory holding pair will do scent-marking within the territory as a means of informing intruders that the area is occupied (Skinner & Chimimba, 2005). The latter authors together with Druce *et al.* (2009), state that this species does not depend on water, and thus is adapted to live in arid conditions. It will however utilise water when it is readily available (Skinner & Smithers, 1990).

Klipspringer family groups in Ethiopia were observed to have a territory size ranging from 6 – 8 ha (Dunbar & Dunbar, 1974). Locally, family groups in semi-arid Namaqualand and the medium rainfall southern Cape geographical areas, had a mean territory size of 49 ha and less than 15 ha respectively (Norton, 1980). This highlights the variation in territory size that can be found between different geographical areas of varying rainfall, and consequently varying availability of food. The present study considers the southern Cape as the closest approximation of TMNP and therefore it is assumed that territory sizes will also be less than 15 ha.

1.2.2 Habitat preferences

Klipspringers primarily inhabit open rocky terrain often in high mountainous areas (Norton, 1980; Estes, 1991; Skinner & Chimimba, 2005). Normally associated with such a terrain are low-growing shrubs, which serve as a food source and provide the klipspringer with clear views for spotting potential danger (Norton, 1980). Steep gorges and cliffs are mostly used to escape from predators (Norton, 1980), whilst the gentler rocky hill slopes are used for normal daytime activities like feeding (Dunbar & Dunbar, 1974). This species often moves to lower lying areas when food becomes scarce on the upper hills (Skinner & Chimimba, 2005), or in the case of individuals living in a desert canyon, they move to lower riverine areas to feed (Tilson, 1980). In summary, the distribution of the klipspringer in many areas that they occupy is patchy, because of their sparse and discontinuous habitat and specific habitat preferences for rocky areas (Skinner & Chimimba, 2005).

1.2.3 Feeding behaviour

Klipspringers are selective feeders (Jarman, 1974), an ecological term given mostly to browsers (dicotyledonous plant material feeders or non-grass feeders) for being particular in the plant part they would feed on. Browsers would typically select the plant parts with the highest protein content (Zhaowen & Takatsuki, 1999) and least amount of fibre, thus the softer parts of a plant e.g. new foliage, flowers, and fruits (McNaughton & Goergiadis, 1986). Most small body sized antelope are categorised into a feeding style of being very selective of the plant parts they eat. This feeding behaviour results in them feeding on a wide range of plant species, and remaining in one vegetation type and in one small home range (Jarman, 1974). This author further describes the food items of animals with this feeding style as being scattered, small, distinct items separated by varying distances from the next similar specimen or acceptable item.

The preference of klipspringers, and any other small herbivores, for high quality food, correlates with their physiology. They have a small gastro-intestinal system, thus have a shorter ingesta retention time, resulting in difficulty processing coarse vegetation (Van Soest, 1994; Hopcraft *et al.*, 2009). Therefore, the limiting factor in a small herbivore’s diet is its gastro-intestinal size. Furthermore, small herbivores use and lose energy quicker than larger herbivores, and therefore should select the highest energy forage to fulfil their energy needs (Hopcraft *et al.*, 2009).

Norton (1984) found that klipspringers in the southern Cape and Namaqualand are almost exclusively browsers, as they feed on shoots, flowers and fruits of shrubs and herbs. Only one grass species was recorded in their diet, thus illustrating the insignificance of graze in this antelope’s diet at the two study areas. An earlier study where the stomach contents of a female klipspringer in Kenya were examined during the dry season, found that this individual also fed mostly on dicotyledonous plant species occurring within its area of occupancy, with only 1.5% of grasses in its diet (Qvortrup & Blankenship, 1974). However, Dunbar (1979) found that klipspringer in Ethiopia included grasses in their diet almost throughout the year with a higher percentage being consumed during the wet season. This author highlighted the behavioural changes of different individuals in different ecosystems. This occasional alteration in diet means that the klipspringer can be described as an adaptable mixed feeder as defined by Kay *et al.* (1980).

The diet of the klipspringer includes a variety of shrubs from various fynbos genera, for example *Osteospermum, Hermannia, Aspalathus, Muraltia, Cliffortia, Agathosma, Maytenus* and *Metalasia* (Norton, 1984). Appendix B provides a detailed list of genera. These genera listed by the aforementioned author were highly favoured when in their growth phase of new young growing shoots and when in flower and bearing fruits (Skinner & Smithers, 1990; Skinner & Chimimba, 2005). Mature leaves browsed on were those of species where the older leaves do not become hard and leathery, whilst stems browsed on were only those that have a photosynthetic function as a result of reduced leaves e.g. *Euphorbia* species and *Thesium* species (Norton, 1984).

Klipspringers rarely feed in one place or from one section of a plant. They prefer to move between preferred browsed species in proximity, taking bites all over a plant (Norton, 1984; Skinner & Smithers, 1990; Skinner & Chimimba, 2005). Klipspringers would spend longer periods of time feeding on flowers and fruits when these are available (Norton, 1984). Furthermore, according to the latter author, the amount of browse being consumed, i.e. number of species in a given period of time; varies depending on the vegetation type and territory size. Although a bigger territory might not necessarily have more palatable and preferred species, it however appeared in the Norton (1984) study that the bigger territory had a greater number of species being fed on. More studies over longer periods in the different seasons are needed however, before conclusive deductions can be drawn about the amount of species fed on at particular sites with different territory sizes.

1.2.4 Activity patterns

Klipspringers are diurnal species with feeding peaks early in the morning and late afternoon, with only brief feeding activity during the course of the day (Norton, 1981). Occasional klipspringer nocturnal activity is suspected (Skinner & Chimimba, 2005) as signs have been observed in the Cape Fold Belt Mountains (Palmer pers. comm. 2013). The klipspringer rests mostly during the day to escape day-time temperature highs (Skinner & Chimimba, 2005). Therefore, resting can be associated with non-feeding periods.

In order to fulfil their high daily metabolic requirements as a species of small body size (Jarman, 1974; Hopcraft *et al*., 2009), klipspringers would continue to feed even in extreme temperatures, whether hot or cold (Norton, 1980; 1981). A characteristic called “standing resting” often occurs throughout the course of the day but was predominantly recorded as occurring in the early morning and late afternoon (Norton, 1981), thus it relates positively to feeding.

According to Norton (1981) females are often feeding for longer periods than males. This is associated with the increased energy requirements of females during times of gestation and lactation. In general klipspringer feed for fewer hours per day than most other antelope species. Furthermore, fluctuations in feeding times and activity patterns are positively correlated with the availability and distribution of preferred browse, and relates to each individual’s nutritional requirement (Norton, 1981).

1.3 Conservation status

Klipspringers are ranked number nine on the “Species of Special Concern” list for the Cape Peninsula and Table Mountain National Park (Rebelo *et al.*, 2010). Its global IUCN Red List status is however categorised as “Least Concern” (IUCN, 2014). This species’ extinction on the Peninsula mountain range and the consequent re-introduction onto the Cape Peninsula makes it an important species to monitor and study.

1.4 Home range analysis

Ecologists and protected area managers are often interested in knowing the home range of a species. Home range analysis assists in better understanding the spatial extent of the collected presence data and consequently the areas inhabited by a species (Keating & Cherry, 2009). A home range is defined as an area that an animal population would occupy and use for their daily activities like feeding and reproduction (Getz *et al.*, 2007). It is often difficult to define the exact extent of the home range of an antelope species because of its mobility, but an estimate can be derived from known data. Particularly, presence data used in habitat suitability studies can be useful in deriving a density estimate of the distribution of a species, and consequently its home range and utilisation distribution (UD) (Getz & Wilmers, 2004). The utilisation distribution is the percentage of area an animal utilises, e.g. 50% area of core activity or 95% area of active use, represented by isopleths/contours that demarcate those areas (Getz *et al.,* 2007).

1.5 Niche concepts

Three major niche concepts have been used by ecologists. With the developments in niche theory over the years, these niche concepts can be classified as traditional and modern niche concepts. This section will discuss these concepts from their earliest stage to their current status. Niche concepts will furthermore be discussed in terms of the biotic, abiotic movement (BAM) diagram and how these concepts pertain to the study (Figure 1).

1.5.1 Traditional niche concepts

The earliest niche concepts used were the Grinellian niche, Eltonian niche and the Hutchsonian niche. The Grinellian niche (Grinnell, 1917) is interpreted as that part of the habitat with the right environmental conditions that would enable a species to survive and produce offspring. Its emphasis therefore is on the habitat of a species. The Eltonian niche refers to the specific role a species fulfils in an ecosystem and consequently in its community (Elton, 1927), which therefore relates this concept to an animal’s size and the food web it is part of. This niche is therefore essentially based on the food web. The Hutchinsonian niche concept (Hutchinson, 1957), defines a niche as those multi-dimensions of environmental gradients, including resources, that fulfil a species’ daily needs to live successfully. Therefore, like the Grinnellian niche, the Hutchinsonian niche also refers primarily to the habitat of a species. Furthermore, the Hutchinsonian niche concept is more encompassing, and is therefore the niche concept largely used in species distribution modelling (SDM) (Kearney & Porter, 2009). Species distribution modelling has been developed as a means to provide ecological insight into the space that a species uses, as well as to predict the distribution of this space across a landscape (Elith & Leathwick, 2009). This occupied and used space is equivalent to a species’ niche. Attempts for niche clarification, however, did not stop with the Hutchinsonian niche of 1957. Further definitions of a niche have been developed although these definitions do relate to the three major niche concepts mentioned above.

1.5.2 Modern niche concepts

Soberon & Nakamura (2009) provided a recent description of three types of niches, which builds on the work of Connell (1961). These are the fundamental niche, potential niche, and realised niche. Fundamental niche refers to that combination of abiotic conditions that exist at a particular place and time to allow a species to occupy that space, and permits positive population growth. Potential niche refers to a subset of the fundamental niche that the species would actually occupy. Lastly, the realised niche, being similar to the Eltonian niche, is a smaller section of the potential niche, as a consequence of biotic interactions.

Connell (1961) illustrated how interspecific competition between two barnacle species can reduce one species’ fundamental niche and thereby give rise to its realised niche. Since a species needs a combination of abiotic and biotic factors (resources) to survive in a particular area, the aim would be to predict and map such an area where this overlap exists, i.e. the realised niche. The biotic, abiotic, movement diagram illustrates this overlap, which results in the realised niche (Figure 1).

In order to determine the extent and location of available niches, niche modellers develop models. The aim of the modelling process can be described as per the BAM diagram shown in Figure 1 (Soberon & Peterson, 2005; Soberon & Nakamura, 2009). Circle A refers to the area where abiotic conditions (e.g. temperature and rainfall) are perfect for the species to survive (Soberon & Peterson, 2005). In geographic space it corresponds with the fundamental niche of a species (Soberon & Peterson, 2005; Peterson & Soberon, 2012). Circle B represents the area that has the correct biotic or interspecific conditions (e.g. resources, predators, disease) for the species' existence. It must be noted that the negative factors in B, such as competitors, will be in moderation to still allow the species to have viable populations (Soberon & Peterson, 2005). The area where A and B overlaps, i.e. G (the sum of G*o* and G*i*) is considered the realised niche in a geographic space (Soberon & Peterson, 2005).

Circle M refers to the area in which movement of the species is not restrained and is accessible over periods of time (Peterson & Soberon, 2012), without barriers to movement and colonisation (Soberon & Peterson, 2005). The actual area of distribution, also known as the occupied geographic area or G*o* (Soberon & Nakamura, 2009; Peterson & Soberon, 2012), is the area of overlap between the three areas and therefore the area where individuals of a population can be viable. Therefore, G*o* corresponds to the geographical distribution of the species (Soberon & Peterson, 2005), which is generally the area desired to model. Establishing which areas are G*o* is not always the ultimate goal of a modelling exercise. Any of the overlaps or niches from the BAM diagram can be used by the species distribution modeller to address the specific question/s raised in the study being conducted. The goal of this study was to determine the areas where the combination of abiotic and positive biotic factors (resources) exists in the Park without determining how accessible these potentially suitable areas are. Lastly the overlap denoted as G*i*, is an area with the right biotic and abiotic conditions, that can be potentially occupied, but which is inaccessible to the species (Soberon & Nakamura, 2009; Peterson & Soberon, 2012).

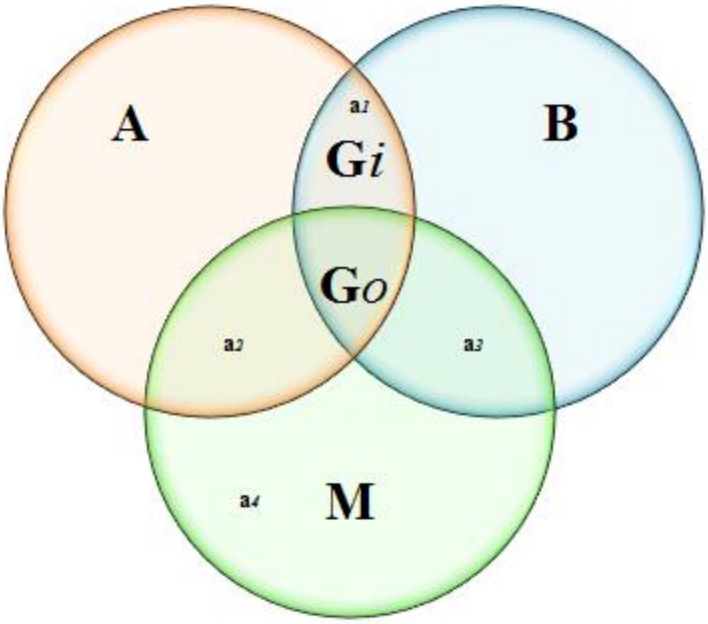


Figure 1: The BAM diagram explaining the three major components in ecological niche modelling. Circle A = the right abiotic conditions, B = the right biotic conditions, M = the areas accessible to the species. G*o* = the occupied geographic area where occurrence of the species is expected; G*i* = the area that can be potentially occupied provided it becomes accessible to the species in future (adapted from Soberon & Peterson, 2005)**.**

As a result of G*o* being the only area that can support a viable population, more absences than presences can be recorded in a geographical area. Figure 1 also shows some noteworthy absences: a*1* – where the species is absent because that area is inaccessible; a*2* – lack of biotic factors; a*3* – the abiotic conditions are not favourable; and a*4* – the correct abiotic and biotic conditions do not exist there (Sillero, 2011; Peterson & Soberon, 2012). Furthermore, absences can also exist within the predicted occupied area (G*o*), i.e. a species can occur partially although the conditions allow it to occupy any space within G*o* (Soberon & Nakamura, 2009; Peterson & Soberon, 2012).

Often absence data are lacking in research as a result of the difficulty in obtaining absence data accurately (Hirzel *et al.*, 2002). Three common reasons for absences are failure to detect individuals despite the fact that the animal was present; historical events that caused a species' absence even though the habitat is suitable; and true habitat unsuitability (Hirzel *et al.*, 2002). Soberon & Nakamura (2009) highlighted that all absences are informative and important to incorporate into the modelling process when the aim of the model is to determine an occupied area of distribution (G*o*).

As circles A, B and M are essential to be included in a SDM, or where the goal is to estimate G*o*, the question to consider is how to obtain data for these areas. Data for abiotic variables e.g. climate, temperature, topography etc. are the easiest to obtain (Soberon, 2010), as these are available on local government databases as well as from independent organizations, e.g. weather station data from the South African Weather Bureau Service database. Data for biotic variables are not easily available, as estimations of the specific parameters in B require large-scale observations over large spatial extents (Soberon, 2010) which can be time consuming and accompanied with high costs. M can be estimated by taking bio-geographical features into consideration, or alternatively by analysing models on the dispersal of individuals of a species (Soberon, 2010). Ultimately, an understanding of these niche concepts will assist in the improved interpretation of results obtained from species distribution modelling.

1.6 Species distribution modelling

There is no general consensus on the approach to niche modelling and distribution modelling concepts, as modelling is open to interpretation by the scientist and hence differ between studies (Jimenez-Valverde *et al.*, 2008; Soberon & Nakamura, 2009; Peterson & Soberon, 2012). Sillero (2011) and Elith & Leathwick (2009) consider both SDM and ecological niche modelling (ENM) as one concept because definitions of SDM and ENM mostly overlap. This conclusion is further supported by Elith & Graham (2009) who define a species distribution model as a model that explains a species' distribution through the use of environmental and geographic information. In addition, Soberon & Nakamura (2009), who define the usage of an ecological niche model as a tool to estimate the distribution of suitable areas for a species based on where the species was observed to occur (presences) or not occur (absences), also support this conclusion. Therefore, within this current study the specific modelling will be referred to as species distribution modelling instead of ecological niche modelling. Primary assumptions of species distribution modelling are that species are at equilibrium with their environment and that the relevant environmental gradients were sampled adequately (Elith & Graham, 2009; Sillero, 2011).

A habitat suitability map (Figure 2) for the klipspringer in TMNP will be the outcome of this research. Habitat suitability maps are used to illustrate areas where the species is likely to occur. The occurrence of the species in an area is based on areas that are ecologically similar to those where the species was observed (Soberon & Peterson, 2005). The habitat suitability map prediction is the result of the occurrence records plus environmental variables, and the specific algorithm used (Figure 2).

There are two types of models in SDM, namely correlative models and mechanistic models. Correlative models link environmental variables and species occurrence data to SDM's through GIS and statistical tools (Kearney & Porter, 2009), to predict an environment suitable for the species.

Mechanistic models on the other hand use functional trait data of a species, together with environmental variables, to predict habitat suitability (Kearney & Porter, 2009). A correlative modelling approach will be followed in the current study. One such a correlative model is the maximum entropy (MaxEnt) species distribution model which will be used in this study. The use of MaxEnt in SDM has become popular over other machine learning methods, as it has proven to provide robust predictions of a species' geographic distribution (Phillips *et al.*, 2006; Wisz *et al.*, 2008; Elith & Graham, 2009; Merow *et al.*, 2013). These other SDM methods include Generalised Linear Models (GLM) and General Additive Models (GAM) (Guisan *et al.,* 2002); Bioclimatic Envelope Algorithm (BIOCLIM) (Beaumont *et al.,* 2005); DOMAIN (Carpenter *et al*., 1993); BIOMOD (Thuiller *et al*., 2009); Genetic Algorithm for Rule-set Prediction (GARP) (Anderson *et al*., 2003) and Boosted Regression Trees (BRT) (Eskildsen *et al*., 2013) to mention a few.

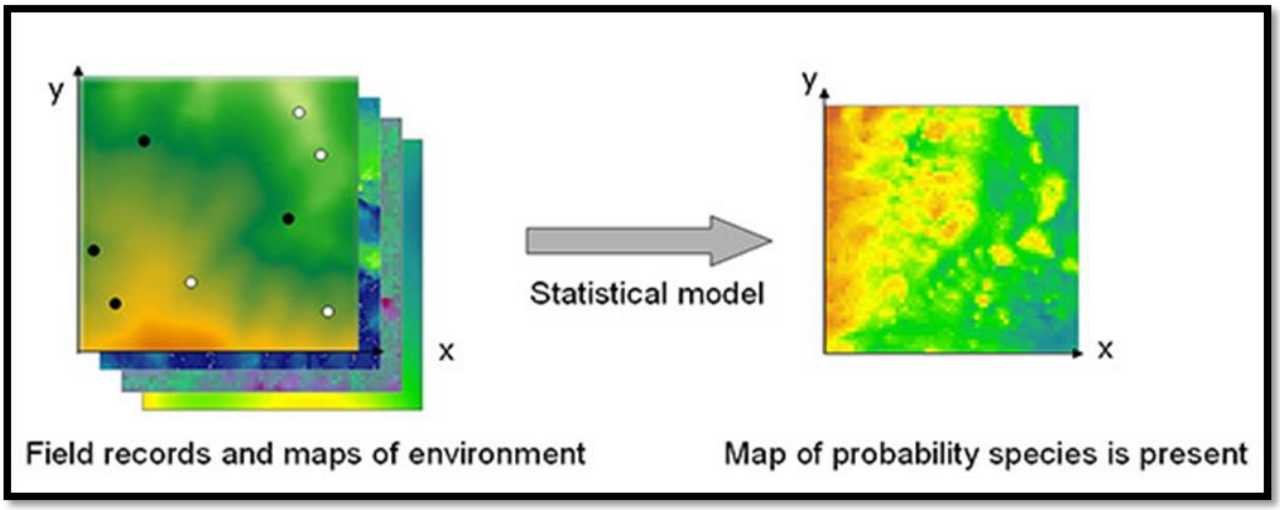


Figure 2: Species distribution modelling summary (Long, 2011).

1.6.1 Maximum Entropy (MaxEnt) modelling of a species' geographic distribution

MaxEnt is a multi-purpose machine learning-tool (Phillips *et al.*, 2006), that models a species' geographic distribution based on geo-referenced points where the species has been observed to occur. These occurrence points together with environmental variables are used as input for MaxEnt. The choice of environmental variables is a guideline for habitat suitability as they provide insight into aspects that would likely influence suitability of a site for a species (Araujo & Guisan, 2006; Phillips *et al.*, 2006; Pearson, 2007). The MaxEnt modelling tool is based on estimating the probability distribution with the maximum entropy which is the distribution that is most spread out, or closest to uniform (Phillips *et al.*, 2006). MaxEnt modelling is known as a presence-only modelling technique as it uses occurrence points recorded via a GPS. In the current study the occurrence points were obtained from individual klipspringer sightings or signs of its presence in an area. The probability distribution estimation is based on constraints and features, the latter being the environmental variables themselves or functions thereof. These environmental variables represent information that is known, and are therefore useful in the probability distribution modelling of the target distribution of suitable habitat. The constraints are the error bounds for the predicted probability distribution and are mathematically derived from the environmental variables (Phillips & Dudik, 2008).

As mentioned above, the environmental variables used in MaxEnt, named “features”, can be the variable itself or a function of it. From these features, MaxEnt generates response curves. Six feature types exist in MaxEnt. These feature types, as described by Phillips & Dudik (2008), Elith *et al.* (2011), and Merow *et al.* (2013) are:

1. Linear feature, referring to the variable itself and the constraint is the mean of a variable across the study area.
2. Quadratic features are the square of the variable thus this variable is constrained by the variance.
3. Product features use the product of two variables that is constrained by the covariance.
4. With threshold and hinge features a threshold is applied with the constraint being that values above the threshold are close to the observed value. Threshold features (stepwise function) models an arbitrary stepwise response once the threshold is reached while hinge features (linear function) fits a forward or backward arbitrary hinge response once the threshold is reached.
5. Discrete features refer to categorical environmental data, such as vegetation type used in the current study. With categorical features the variable is categorised into a specific number of categories that is present in the variable. A value of 1 is assigned where the specific category in that feature is present and a 0 where a specific category is absent. For example, the vegetation layer used in this study has 14 categories each representing a vegetation community or a vegetation type. When the model is run a value of 1 would be assigned for all the cells where e.g. ericaceous fynbos is present and 0 where ericaceous fynbos is absent and a probability distribution for this category will be calculated accordingly.

For a positive prediction, the specific constraints applied should be satisfied, i.e. the mean, variance, covariance etc. of that environmental variable should be close to values where an occurrence point was recorded (Phillips & Dudik, 2008). It is important to note that the cell values of an environmental variable where the species is recorded are matched within error bounds of the constraints, i.e. it is matched approximately and not matched exactly (Merow *et al.*, 2013). The number of occurrence points guides the choice of feature types used. If more than 80 occurrence points exist, the auto features option in MaxEnt (which automatically chooses feature types based on the sample size) will incorporate all the features (Phillips & Dudik, 2008).

Although MaxEnt only requires presence data, it can also process absence data making it a relatively flexible method (Phillips *et al.*, 2006). MaxEnt also has the ability to use both continuous variables (derived from measured values) and categorical variables (derived from a limited number of distinct values). The importance of each variable, i.e. the extent to which a variable influences the outcome, can also be determined through MaxEnt's jackknife tool (Phillips *et al.*, 2006). In particular the jackknife tool's outcome shows which variable has the highest gain (gain = a measure of goodness of fit) when used in isolation, therefore which variable contains the most useful information. In reverse it can also show which environmental variable reduces gain the most when being omitted, which in turn also reflects the contribution of each variable (Phillips, 2006). Furthermore, the regularisation parameter setting (a log-function) in MaxEnt minimises the chances of overfitting, i.e. when the predicted suitable habitats match the occurrence data too closely (Phillips *et al.,* 2006; Phillips and Dudik, 2008; Elith *et al.*, 2011). This is important since an overfitted model does not provide a good representation of other areas where the species can potentially occur, i.e. such a model does not provide much additional information. The regularisation parameter can be varied, and the most ecologically meaningful output can be visually inspected, as Radosavljevic & Anderson (2014) have done.

The probability of presence at a site cannot be determined from presence data only; the algorithm requires the use of absence data as well (Elith *et al.,* 2011; Merow *et al.,* 2013). In studies such as the present study where no absence data are available, MaxEnt uses background data or pseudo absence data (Phillips *et al.*, 2006). Background data are generated randomly by the MaxEnt model and the number of points is user specified. In essence this background data are raster cells or pixels that contain values representing the environmental conditions present in that pixel (since the environmental conditions present in the study area are represented as continuous raster layers of specified pixel or cell size).

Maximum entropy is rooted in the second law of thermodynamics (Phillips *et al.,* 2006) which specifies that in the absence of outside influences, processes moves towards maximum entropy. Its application to machine learning and specifically probability distribution estimation stems from the work of Jaynes (1957) and was further adapted by Phillips *et al*. (2004) to be applied to species distribution modelling. Maximum entropy can be interpreted as a means of getting the best probability distribution after all constraints were taken into consideration.

As mentioned above, MaxEnt uses environmental data and occurrence data to determine the distribution of suitable habitat for a species within a specified area. All environmental data are in the form of raster layers which consist of pixels of a specified cell size. Each cell or pixel has a value assigned to it which represent the information for a specific variable at a specific location e.g. for a variable rainfall, each cell will contain the amount of rainfall (in mm) that specific area (based on its position in the study area) receives. When the occurrence points are overlaid on the raster layers, one or more points will fall within a cell. MaxEnt then uses the cell value of each cell that has an occurrence point in it to calculate the mean (or one of the other five possible feature types) for that specific variable. This average value is what will be used to identify an area as suitable or not. Basically MaxEnt will use that average value to look for all other cells in the study area (those with occurrence points in it as well as those with no occurrence points in them) that has a value close to the average cell value within some specified constraints (Phillips & Dudik, 2008). This process is repeated for all environmental variables and a probability of occurrence map is created with a probability score between 0 and 1 given to each cell. This score indicates the probability that the study species can be present there based on the combination of suitable environmental conditions that exist within that specific cell (Figure 3). The cell values represented in the final map will be those of maximum entropy within the bounds of the constraints for each variable. Using MaxEnt's logistic output, a presence site would have a probability score of 0.5 and higher (Phillips & Dudik, 2008). A score closer to 1 indicates high occurrence probability and those lower than 0.5 indicate lower occurrence probability (Figure 3).

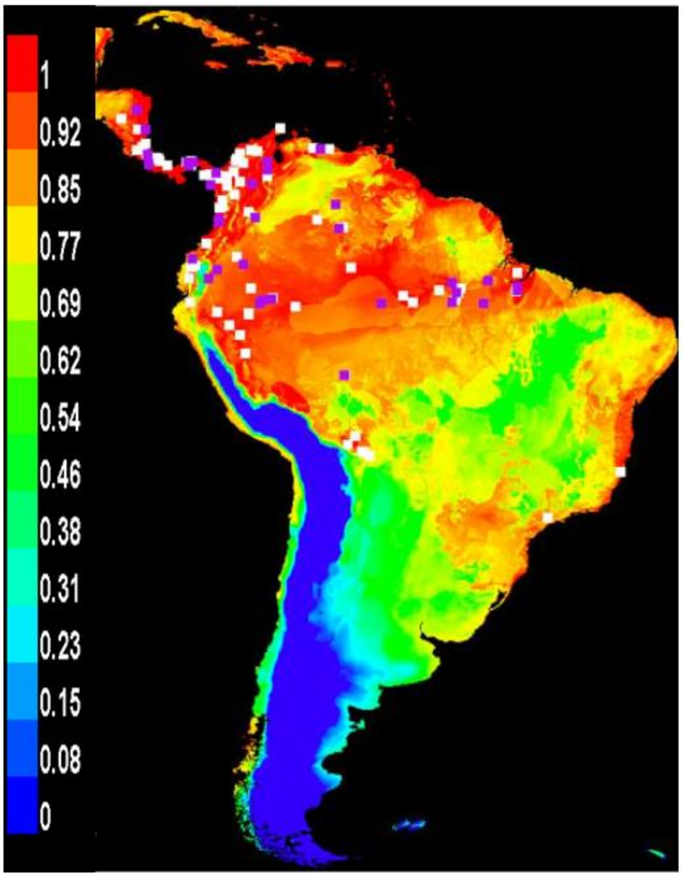


Figure 3: Probability of occurrence map as produced in MaxEnt. Red indicates high occurrence probability and blue low occurrence probability. The training and test occurrence points used are shown in white and purple respectively (Phillips, 2006).

1.6.2 Environmental variables

In species distribution modelling, multiple environmental variables can be included (Merow *et al.*, 2013). The choice of variables considered relates to the klipspringer's ecology and are those closely associated with the observed presence of the species. These include topographic variables, climatic variables and vegetation type, amongst others (Table 1). Topographic variables include altitude, slope and aspect (Appendix C). Slope is important as daily activities may vary along a gradient. The higher and steeper slopes (including cliffs and gorges) are useful for escaping predators whilst feeding can take place along the entire gradient (Dunbar & Dunbar, 1974). The klipspringer's general preference for open rocky and high mountainous areas makes elevation an important variable to include in the model training. Whether klipspringers prefer warmer north facing or cooler south facing slopes is currently unknown and therefore including aspect as a variable would be useful to identify if there is any correlation. Rainfall is linked to a klipspringer's territory size (Norton, 1980). Coupled with rainfall is the availability of their dietary plants as a result of the vegetation type that can establish and grow under the study area's climatic and edaphic conditions. The distance from the urban edge becomes important when attempting to understand the influence of urbanisation on a species (Blank & Blaustein, 2012).

The model will essentially identify areas in geographical space which have the same, or similar, combinations of environmental features as the existing klipspringer occurrence points. The environmental features used consist of a combination of continuous (e.g. altitude) and categorical variables (e.g. vegetation communities). GIS layers will be created for each environmental variable. Ultimately a species' niche will be reconstructed with these environmental layers and species occurrence data (Soberon & Peterson, 2005).

Table 1: Environmental variables that potentially affect klipspringer distribution within TMNP. Some of the layers for these variables are in raster format (i.e. as an image and therefore pixel size is relevant) whilst others are shape files in vector format produced at a specific ratio/map scale.

Variable	**Spatial resolution of layers**		**Source**	**Motivation**
	Image/raster pixel size (m)	**Ratio scale (vector)**		
Topographic variable: Elevation/Altitude	10x10 m		City of Cape Town (CoCT) 10 m Digital Elevation Model (DEM)	Species generally prefer high areas and therefore essential to know its preferred heights (Skinner & Chimimba, 2005).
Topographic variable: Slope	10x10 m		CoCT – derived from the DEM	Steep vs. gentle slope activity. These are used differently for different purposes e.g. feeding and escaping predators.
Topographic variable: Aspect	10x10 m		CoCT – derived from the DEM	Distinction between preference of warmer N-facing slopes or cooler S-facing slopes.
Climatic variable: Mean Annual Rainfall (MAR)	1x1 km		WordClim (Hijmans *et al.*, 2005)	Affects territory size. Availability of young shoots. Influence of droughts.
Vegetation type (communities)		1: 10 000	TMNP vegetation community map based on the work of Cowling *et al.* (1996).	Preferred food. Different vegetation communities and which microclimatic conditions of each vegetation community are preferred.
Distance from urban edge		1: 10 000	TMNP – Use Zones Classification (TMNP Planning Department)	Important to know what is the influence of urbanisation on the movement of individuals and the safest distance from the urban edge.

1.6.3 Model evaluation

Model evaluation is necessary to determine how well the chosen model performed (Elith & Leathwick, 2009; Merow *et al.,* 2013). In MaxEnt a subset of the occurrence data are withheld from the training data and are used for model evaluation, known as the test data, which were not used in training the model (Phillips *et al*., 2004). This set of occurrence data comes from the same set of data that were used to calibrate the model, however, using a setting in MaxEnt, *random test percentage*, the total set of data are split into training data and test data depending on the percentage specified. The data split is user defined e.g. 70% training and 30% test (Phillips *et al*., 2006) or 60% training and 40% test (Phillips & Dudik, 2008). Peterson *et al.* (2008) emphasised that probability distribution predictions involve two errors that can occur. The first is omission errors (presences/suitable areas that are not predicted correctly, also known as false negatives) and the second is commission errors (absences/unsuitable areas that are incorrectly predicted also known as false positives).

The area under the receiver operating characteristic curve (ROC AUC) is commonly used to measure performance of a machine learning modelling method (Bradley, 1997; Elith *et al*., 2006; Phillips & Dudik, 2008). This model evaluation tool is threshold independent and ranking based. It specifically measures the quality of a ranking of sites (Phillips & Dudik, 2008), which gives an account of a model's ability to classify an area as suitable or unsuitable (Bradley, 1997). This calculation is normally done with absence data; however, in cases where absence data are lacking, background data or pseudo-absences can be used to calculate the AUC (Phillips & Dudik, 2008). Since MaxEnt is a presence-only modelling tool it uses background data, which represents environmental conditions present in the study area and chosen at random (Phillips & Dudik, 2008). During a MaxEnt model simulation, model evaluations will be carried out to test for omission and commission error. An omission error curve as well as a ROC AUC curve is produced as part of the output. The AUC score ranges from 0 to 1, with a score of 0.5 indicating the prediction equates to a random assumption and a value of 1 indicating a perfect prediction (Elith *et al*., 2006). According to these authors, any score below 0.5 indicates poor predictive performance, which is even worse than a random guess.

The test data are chosen at random, similar to the background points. A good model will have low omission error on the test data (Anderson & Gonzalez, 2011) and is therefore able to correctly predict presence sites as present/suitable. In essence the AUC will allocate a higher ranking to randomly chosen presence sites compared to randomly chosen absence (background) sites (Phillips *et al.*, 2006). Thus, the AUC is based on how well the model performed in distinguishing between presence points and background points.

1.7 Research question, aim and objectives of the current study

Research Question

Which parts of Table Mountain National Park are potentially suitable habitats for the klipspringer?

Aim

The main aim of this study is to produce a distribution range map that shows the potentially suitable habitats for the klipspringer, (*Oreotragus oreotragus subsp. oreotragus*) within TMNP, through the use of a species distribution modelling tool.

Objectives

- To collect and source occurrence data for the klipspringers in the Park.

- To determine the home range extent of the klipspringer populations.

- To identify environmental variables that can explain klipspringer distribution.

- To produce environmental layers that can be used in a species distribution model.

- To run a species distribution model using the maximum entropy modelling tool.

- To determine the distribution of all suitable habitats for the klipspringer in TMNP.

CHAPTER TWO: MATERIALS AND METHODS

1. Study area

1.1 Location

Table Mountain National Park is situated in the Western Cape Province of South Africa (Figure 4), within the Fynbos Biome. Its grid coordinates are 18° 18' 12" E, -33° 54' 40" S for the top left corner and 18° 30' 3" E, -34° 21' 33" S for the bottom right corner. The Park's boundaries include the entire Cape Peninsula mountain range which stretches from Signal Hill in the north to Cape Point (south-western most tip of Africa) in the south. It is divided into three sections, namely northern section, central section and southern section. For some instances when data needed to be in a projection, Transverse Mercator (Lo) 19 was used as some layers were already in this projection (central meridian: 19.00000000, scale factor: 1.00000000, Cape datum). In other instances the WGS 1984 UTM Zone 34 S projection was used.

1.2 Geology

The underlying geology of the Fynbos Biome varies in rock composition, consisting of sandstone, quartzite, granite, shale and limestone (Mucina & Rutherford, 2006). These rock types are all present on the Cape Peninsula with the most prominent being quartzitic sandstone and granite outcrops (personal observation). Geologically classified as Cape Supergroup rocks, the Table Mountain subdivision of the Cape Supergroup is present in the study area, and forms part of the Cape Fold Belt Mountains (Compton, 2004). Dominating within the Table Mountain Group is the Peninsula Formation which consists entirely of quartzite (Mucina & Rutherford, 2006).

The Peninsula Formation supports the most dominant vegetation type in TMNP, namely Peninsula Sandstone Fynbos. This vegetation type grows on nutrient poor, acidic lithosol soils which originated from Ordovician sandstones of the Table Mountain Group (Mucina & Rutherford, 2006). In contrast, the northern-most section of TMNP, namely Lions Head and Signal Hill, is dominated by nutrient rich, clayey soils (Cowling *et al.,* 1996) derived mostly from shale and granite. At Lions Head and Chapman's Peak, the underlying granite within the study area is exposed.

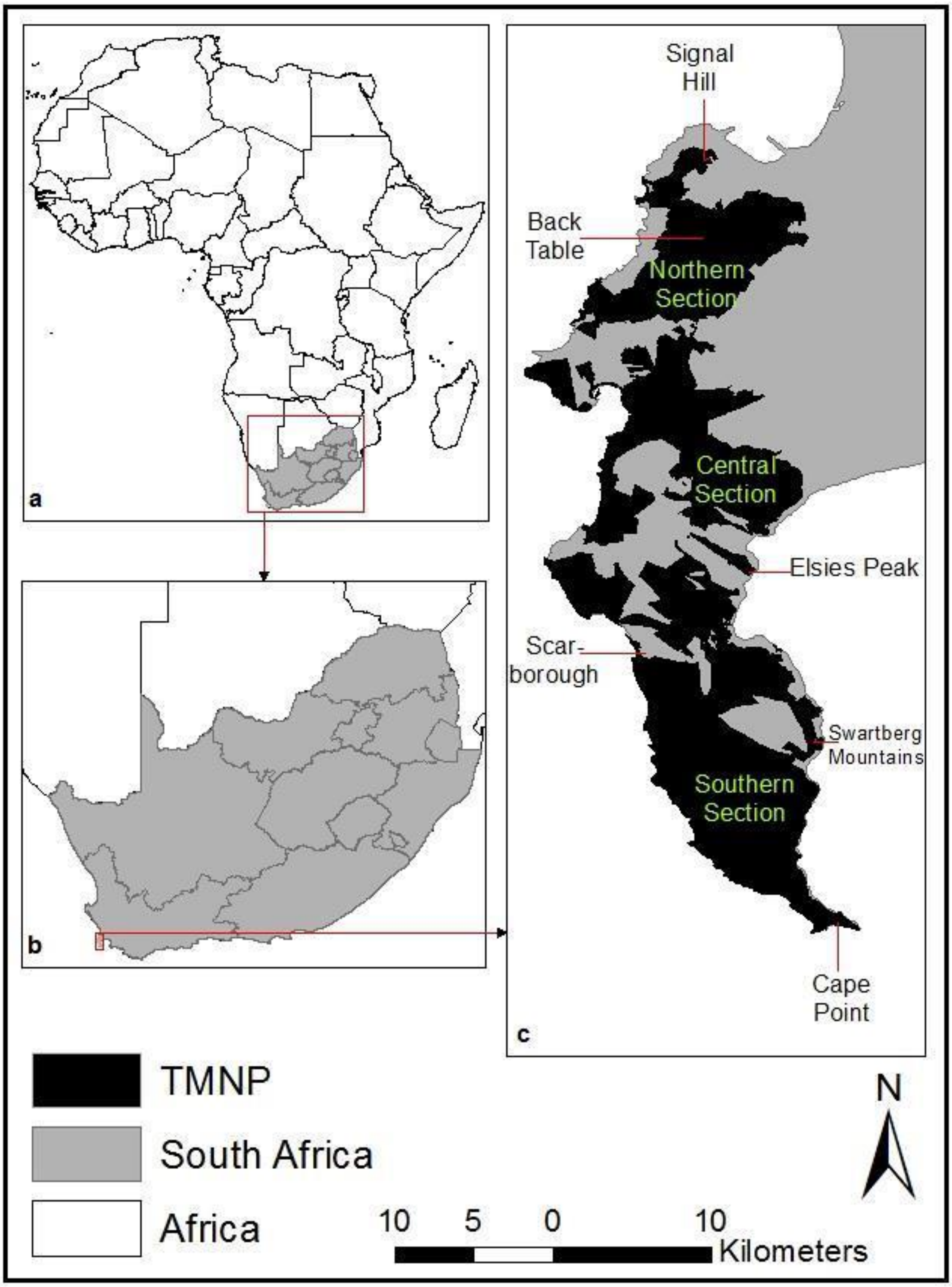


Figure 4: Table Mountain National Park (TMNP) situated in the Fynbos Biome of South Africa.

1.3 Climate

The study area is located in a winter rainfall region, and can be best described as Mediterranean with hot dry summers and cool wet winters (Goldblatt & Manning, 2002). Mean annual rainfall varies from 520 – 1690 mm per year, with the peak rainy season occurring between May and August. Temperatures vary between a mean daily maximum of 25°C in summer and mean daily minimum of 7.2°C in winter (Mucina & Rutherford, 2006). According to these authors, the climate is largely influenced by the south easterly wind in the summer months which results in the well-known “table cloth”, and consequently results in rainfall at higher altitudes and southern and eastern slopes.

1.4 Vegetation

The vegetation of the study area is described as fynbos, as it falls within the Fynbos Biome. This Biome occupies most of the Cape Fold Mountain Belt (Mucina & Rutherford, 2006). Furthermore, it forms part of the Cape Floral Region. This floral region’s profusion of plant diversity led to its proclamation as a world heritage site in 2004 by the United Nations Educational, Scientific and Cultural Organisation (UNESCO, 2004). The fynbos vegetation type dominating the study area is Mountain Fynbos, specifically Peninsula Sandstone Fynbos. Renosterveld is present to a lesser extent as well as remnant pockets of Forest and Coastal Thicket (Cowling *et al.*, 1996).

1.4.1 Cape Peninsula vegetation types and communities

Three major vegetation types occur on the Cape Peninsula, namely Cape Fynbos Shrubland being synonymous with Peninsula Sandstone Fynbos (92%), Renosterveld Shrublands and associated grasslands (>5%), and Forest and Thicket (±3%) (Cowling *et al.,* 1996). Table 2 provides an account of the major vegetation communities of the Cape Peninsula as well as their structure and typical environment. Eleven fynbos vegetation communities exist within the study area namely dune asteraceous fynbos, coastal restioid fynbos, wet restioid fynbos, upland restioid fynbos, sandplain proteoid fynbos, mesic oligotrophic proteoid fynbos, mesic mesotrophic proteoid fynbos, wet oligotrophic proteoid fynbos, wet mesotrophic proteoid fynbos, ericaceous fynbos, and undifferentiated cliff communities (Table 2 & Figure 5). All the aforementioned vegetation communities form part of the Peninsula Sandstone Fynbos vegetation type. Additionally it would include wetlands and their associated wetland plant species.

Renosterveld and its associated grasslands species occurs on nutrient rich clay-soils of shale and granite origin. Typical species include shrubs from the Asteraceae family, geophytes, as well as an understory of grasses (Mucina & Rutherford, 2006). Forest and thicket species are largely broad-leaved evergreen trees and shrubs. Furthermore these two vegetation types are considered to be species poor at both local and regional scales (Cowling *et al.*, 1996). Plant species growing in forest and thicket have a preference for moist habitats and therefore occur on fire protected kloofs (ravines), scree slopes, stream banks and on the coastal Margin (Cowling *et al.*, 1996; Mucina & Rutherford, 2006).

1.4.2 Peninsula Sandstone Fynbos

This vegetation type is largely confined to the Cape Peninsula of the Western Cape Province of South Africa. It predominantly stretches from Lions Head and Table Mountain in the north to Cape Point in the south and is found on altitudes of 20-1086 m.a.s.l. (Mucina & Rutherford, 2006). Its original extent is estimated to be 23 000 ha (Government Gazette, 2011).

Peninsula Sandstone Fynbos is considered least threatened as 90% of it is statutorily conserved in TMNP (Mucina & Rutherford, 2006; Government Gazette, 2011). It is estimated that 91% of this veld type remains in its natural form and it supports about 140 endemic plant species and 66 Red Data List plant species (Government Gazette, 2011). Alien infestations pose a threat to the continual existence of this vegetation in its pristine state (Roura-Pascual *et al.,* 2009; Foxcroft & McGeoch, 2011). Factors cited as leading to this vegetation type's transformation include urban sprawl and alien vegetation particularly pine plantations (Mucina & Rutherford, 2006).

Structurally it is described as a medium dense, tall proteoid shrubland over a dense moderate tall, ericoid-leaved shrubland. Thus it consists of proteoid, ericaceous and restioid fynbos, together with asteraceous fynbos. It grows on gentle to steep slopes. The vegetation is well adapted to its nutrient poor, acidic substrate (Mucina & Rutherford, 2006).

Table 2: Major vegetation communities of the Cape Peninsula (adapted from Cowling *et al.*, 1996).

Vegetation community	**Structural characteristics**	**Common species**	**Environment**
Dune asteraceous fynbos	Low ericoid shrubland	*Metalsia muricata, Ischyrolepis eleocharis, Searsia glauca, Phylica ericoides, Euclea racemose, Passerina paleacea, Ficinia lateralis, Helmuthia membranacea, Otholobium fruticans, Ehrharta villosa*	Recent calcareous coastal dune sands; MAR = 678 mm; mean slope = 17.0°; mean aspect = S
Coastal scree asteraceous fynbos	Low-medium ericoid and broad-leaved shrubland	*Coleonema album, Euclea racemosa, Tarchonanthus camphoratus, maytenus oleoides, Eriocephalus africanus, Pentachistis curviflora, Felicia fruticosa, Colpoon compressum, Cassine piragua, Olea europaea* subsp. *Africana*	Sandstone scree coastal slopes, subject to strong salt-laden onshore winds; MAR = 656 mm; mean slope = 30°; mean aspect = S
Wet restioid fynbos	Low restioid herbland	*Ischyrolepis cincinnata, Tetraria cuspidate, Elegia filacea, Thamnochortus lucens, Cliffortia subsetacea, Erica imbricata, Leucadendron laureolum, Pentachistis curviflora, Restio quinquefarius, Restio bifurcus*	Shallow seasonally waterlogged sands on sandstone at low altitudes; MAR = 634 mm; mean slope = 13°; mean aspect = WSW
Upland restioid fynbos	Low restioid herbland	*Thamnochortus nutans, Chondropetalum ebracteatum, Ursinia nudicaulis, Restion bifidus, Ehrharta setacea, Watsonia borbonica subsp. Borbonica, Penaea mucronata, Cliffortia ruscifolia, Erica hispidula, Chondropetalum mucronatum*	Shallow seasonally waterlogged sands on sandstone; MAR = 1404 mm; mean slope = 32°; mean aspect = SW
Sandplain proteoid fynbos	Medium-height proteoid shrubland with a low ericoid and restioid understorey	*Thamnochortus erectus, Metalasia muricata, Euclea racemosa, Carpobrotus edulis, Searsia laevigata, Leucadendron coniferum, Searsia glauca, Eriocephalus africanus, Agathosma imbricata, Diosma hirsute*	Old marine (deep and well-drained) sands, occasionally on calcrete or limestone; MAR = 804 mm; mean slope = 16°; mean aspect = S
Mesic oligotrophic proteoid fynbos	Medium-height proteoid shrubland with a low ericoid and restioid understorey	*Thamnochortus lucens, Elegia stipularis, Metalasia muricata, Phylica imberbis, Elytropappus scaber, Salixis axillaris, Struthiola ciliata, Ischyrolepis cincinnata, Hypodiscus aristatus, Tetraria eximia*	Shallow, leached sands on sandstone; MAR = 765 mm; mean slope = 21°; mean aspect = SW
Mesic mesotrophic proteoid fynbos	Medium-height proteoid shrubland with an ericoid understorey	*Erica plukenetti, Penaea mucronata, Protea lepidocarpodendron, Cliffortia falcata, Elytropappus scaber, Searsia lucida, Phylica imberbis, Passerina vulgaris, Leucadendron salignum, Cliffortia stricta*	Mainly deep, sandy loams associated with colluvium or granites on lower mountain slopes; MAR = 947 mm; mean slope = 30°; mean aspect = SSW
Wet oligotrophic proteoid fynbos	Medium-height proteoid shrubland with an ericaceous and restioid understorey	*Anthospermum galioides, Leucadendron xanthoconus, Penaea mucronata, Elegia racemosa, Bobartia gladiata, Erica plukenetii, Leucadendron salignum, Searsia lucida, Otholobium fruticans, Myrsine Africana*	Shallow, leached sands on sandstone; MAR = 1168 mm; mean slope = 30°; mean aspect = SSW
Wet mesotrophic proteoid fynbos	Medium-height proteoid shrubland with an ericoid understorey	*Searsia tomentosa, Myrsine africana, Penaea mucronata, Widdringtonia nodiflora, Protea nitida, Searsia lucida, Leucadendron xanthoconus, Stoebe cinerea, Maytenus oleoides, Erica plukenetii*	Mainly deep, sandy loams associated colluvium or granites on lower mountain slopes; MAR = 1136 mm; mean slope = 30°; mean aspect = SSE
Ericaceous fynbos	Low ericaceous and restioid shrubland	*Penaea mucronata, Hypodiscus aristatus, Erica hispidula, Leucadendron xanthoconus, Pseudopentameris macrantha, Corymbium glabrum, Erica plukenetii, Clutia polygonoides, Thamnochortus nutans, Tetraria cuspidate*	Shallow, leached, organic-rich sands at high altitude subject to intermittent condensation from orographic cloud in summer; MAR = 1197 mm; mean slope = 30°; mean aspect = SSW

Undifferentiated cliff communities	Variable, mainly low and sparse ericoid shrubland	*Cliffortia ruscifolia, Helichrysum cymosum, Cullumia ciliaris, Lampranthus falciformis, Cliffortia odorata, Ehrharta ramosa* subsp. *Aphylla, Peucedanum galbanum, Anthospermum galioides, Anthospermum aethiopicum, Pelargonium cucullatum*	Steep to vertical sandstone cliffs; MAR = 1168 mm; mean slope = 47°; mean aspect = SSW
Forest and Thicket	Low-medium (thicket), or medium to tall (forests) closed-canopy, broad-leaved formation with a sparse understorey	*Rapanea melanophloeos, Diospyros whyteana, Cassine peragua, Knowltonia capensis, Myrsiphyllum scandens, Kiggelaria Africana, Olea capensis, Olinia ventosa, Secomone alpine, Chionanthus foveolatus*	Colluvium or granite derived soils on wet slopes or fire protected kloofs (ravines) and coastal margins; MAR = 992 mm; mean slope = 30.0°; mean aspect = SSW
Renosterveld and associated grasslands	Low grassland/low-medium ericoid and broad-leaved shrubland	*Searsia lucida, Chrysocoma coma-aurea, Helichrysum patulum, Anthospermum spathulatum, Helichrysum cymosum, Salvia africana-caerulea, Hyparrhenia hirta, Mohria caffrorum, Searsia glauca, Merxmuellera stricta*	Sandy loams on clay subsoil, shale or granite derived; MAR = 826 mm; mean slope = 24°; mean aspect = SW
Wetlands	Medium-height ericoid shrubland with an ericoid and restioid understorey	*Penaea mucronata, Berzellia abrotanoides, Platycaulos compressus, Leucadendron laureolum, Berzellia lanuginosa, Pentaschistis curviflora, Osmitopsis astericoides, Watsonia tabularis, Psoralea pinnata, Restio quinquefarius*	Seepage sites with shallow medium depth sandy soils with high organic matter over sandstone bedrock; MAR = 779 mm; mean slope = 12°; mean aspect = SSW

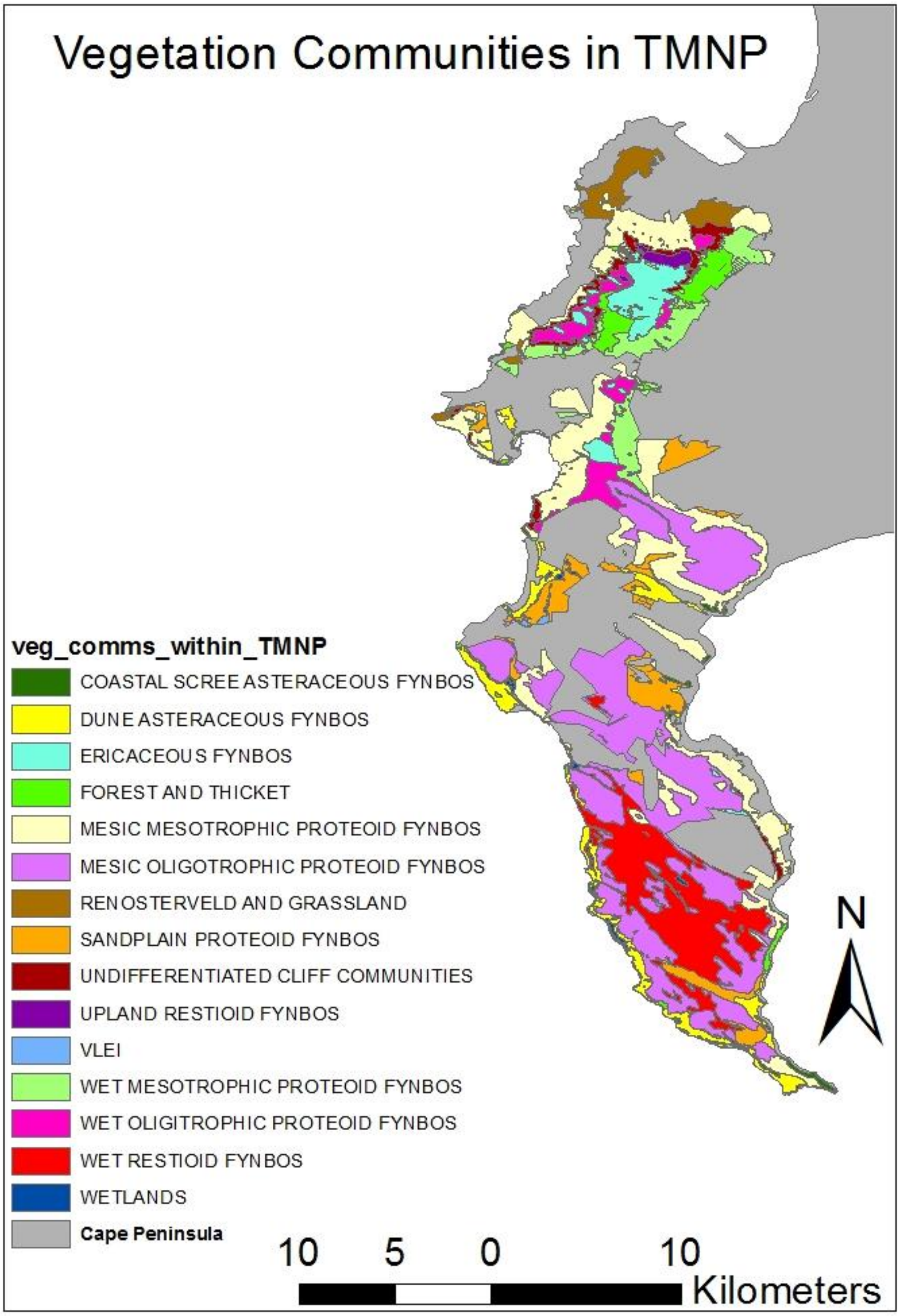


Figure 5: Vegetation communities (adapted from Cowling *et al.*, 1996) clipped to only the boundaries of Table Mountain National Park (TMNP,) as represented before some of these communities were transformed either by urbanisation, agriculture, and Pine plantations.

2. Data collection

2.1 Occurrence data

The first batch of klipspringer locality data was collected by Mr Emile Marell (former post graduate student at the Nelson Mandela Metropolitan University who first initiated this research) and his assistants during the years 2004 – 2007. The occurrence data were collected by intermittent patrols between November 2004 and September 2007. Collected data included that of dung middens, recognisable spoor or tracks, and visual observations. The first targeted areas that were previously surveyed for the species' presence were the known territories. These are the territories established in the Cape of Good Hope section of TMNP (also referred to as Cape Point), those at the Swartberg Mountains above Simons Town, as well as those on the Back Table section of TMNP (Figure 4). Furthermore, information gained from hiker sightings were also used as a guide to the areas to be surveyed for signs of the species' presence (Cheney pers. comm., 2013).

In addition, other locality data were obtained from the collared klipspringers that had been reintroduced into TMNP in October 2005. The klipspringers, collared with Hawk 105 GPS satellite collars, were of both sexes as the aim was to release the individuals in breeding pairs (Marell, n.d.). GPS collars were used as they are useful devices for tracking and monitoring mountain antelopes (Haller *et al.,* 2001). The GPS collar has a radio receiver which receives information from satellites from which an animal's location (i.e. latitude and longitude) can be determined and downloaded onto another device or sent directly to the researcher (Mech & Barber, 2002).

Occupancy was a key concept for collecting the second batch of presence data in this study. Occupancy describes a site in a landscape where individuals of a species are present (Mackenzie & Royle, 2005). According to these authors, one of the approaches to determine occupancy of a species is to apply stratified random sampling. Consequently, before sampling commenced, the study area was categorised according to vegetation communities and further stratified according to altitude and slope. The aforementioned stratification ensures that the sampling effort would extend beyond only sampling within the known territories. Within ESRI's ArcMap (ArcGIS 10.1) the polygon layers for slope and altitude were classified into categories (Figure 6) to assist in the planning of the field survey. This was done to ensure that a variation of slope and altitude was surveyed for klipspringer occupancy as well as all the vegetation communities. These three layers were merged together as one layer from which vegetation community subtypes (or habitat subtypes) were created to guide the random stratified sampling.

The established klipspringer population occurs permanently within the boundaries of TMNP. Therefore, the entire Park can be positively sampled for occupancy instead of just use. Use is defined by Mackenzie (2005) as the animals only being present in the area of interest during a certain period of time in a season whilst occupancy refers to the species being permanently present in some portions of the study area over an entire season/-s. A general assumption of occupancy studies is that the system is closed to changes in species occupancy while repeated surveys are being conducted (Mackenzie & Royle, 2005). Therefore in cases where the species is not detected, it will either be a direct result of the sampling method, the sampling effort (Mackenzie, 2005) or the individuals having remained in their territories and therefore were not detected outside their territories.

Other potential sources for klipspringer locality data were also explored. These included the Global Biodiversity Information Facility (GBIF) which is a global database for species distribution data. No distribution data for the klipspringer in the study area was available in this database. A second exploration was the citizen scientist platform for uploading species sightings namely I-spot. Five sightings were found and the GPS coordinates of these sightings were provided by the creator of this platform, Dr Tony Rebelo from the South African National Biodiversity Institute. These sightings were added to existing presence records. Lastly seven locality records from a National Diploma: Nature Conservation student from the Cape Peninsula University of Technology, Heather Edwards, were included in the occurrence layer of klipspringer observations in the southern section of the Park. The latter locality points were obtained in 2010 as part of the former student's unpublished Work Integrated Learning mini-research project that she carried out in the southern section of the Park.

2.2 Survey routes

Survey routes were designed within ArcMap (ArcGIS 10.1). Shapefile layers of the footpaths and roads in TMNP as well as the park boundary were used to overlay aerial images of the study area. The aerial images were used to study and inspect the terrain to ascertain information such as the ease of access to areas. It was also used to select survey routes that are a representative portion of the study area. In order to ensure surveys were carried out in all different habitats represented in the study area, a vegetation community subtypes layer had to be created.

The vegetation shapefile representing the vegetation communities or vegetation types within Table Mountain National Park was used as the base for preparing the habitat subtypes layer. These subtypes are essentially all the vegetation communities at a variation of slope and altitudinal ranges. A 10x10 m Digital Elevation Model (DEM) raster file was used to determine altitude and slope. Slope was derived from the DEM using the Spatial Analyst tool of ArcGIS 10.1. This ensured a separate layer for the before mentioned attribute to be used in the preparation of the vegetation subtypes.

Altitude was calculated from the DEM layer. The raster layers for slope and altitude were converted to vector layers before they were joined with the vegetation community shapefile. These three layers (vegetation, altitude and slope) were then merged within ArcMap (Figure 6). Before the merge the slope layer was categorised into five categories (0-10; 11-20; 21-30; 31-40, and those greater than 40 degrees). Altitude was categorised into five categories namely, below 0 (as a result of a water body recorded in the layer), 1-250; 251-500; 501-750; and 751-1084 m.a.s.l.

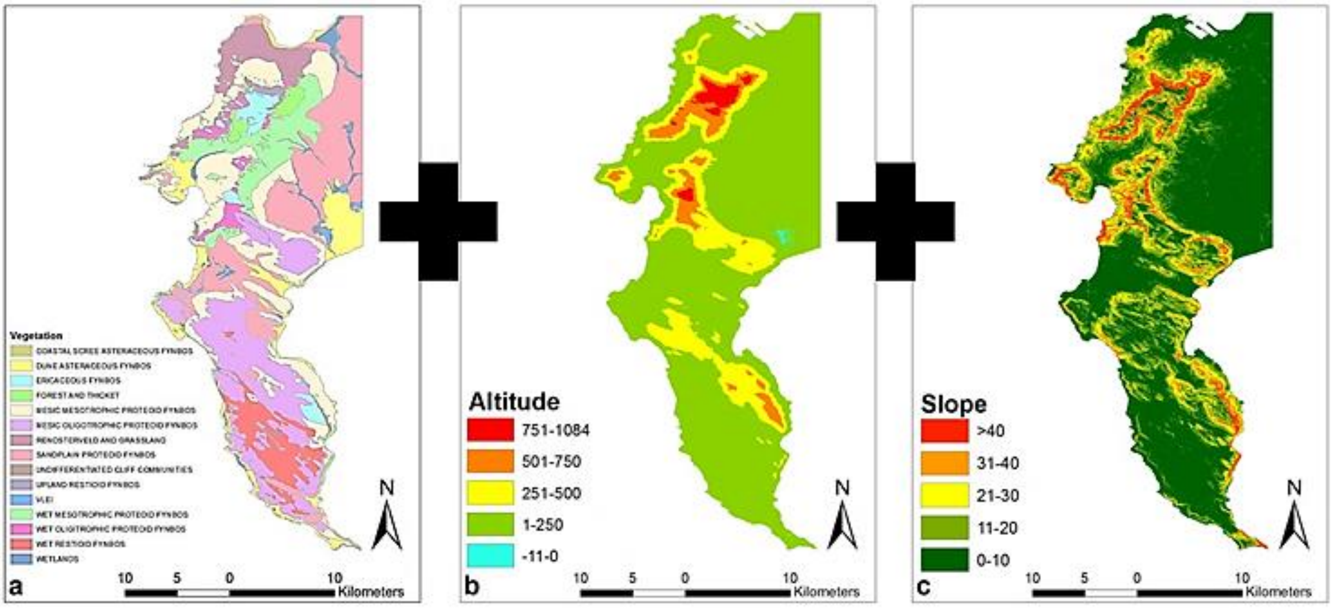


Figure 6: The process of creating habitat subtypes, whereby three shapefiles namely vegetation communities (a), altitude (b) and slope (c) were merged to produce the habitat subtypes at a variation of slope and altitude.

The merged layer was then clipped to the exact boundaries of TMNP. At this point a total of 65623 habitat subtypes (i.e. combinations of altitude, slope and vegetation communities) existed. As it is impossible to sample such a vast amount of subtypes, steps had to be taken to reduce this number. In an effort to reduce the number of subtypes those smaller than one hectare were eliminated. This was done by creating a new layer of only the subtypes greater than one hectare which reduced the number significantly to only 2412. Following this was the elimination of the extreme steep slopes as they were considered inaccessible to climb by humans. Higgins *et al*. (1999) described steep slopes as those that had a slope greater than 40 degrees. Thus subtypes with slopes of 41 – 85 degrees were eliminated.

The next step was to use a summarise function to determine how many polygons of each unique subtype existed within the entire study area. This was done using the dissolve tool of ArcGIS, which combines polygons with the same attributes. Consequently, the number of subtypes was reduced to

119 (Figure 7). The frequency tool was used afterwards to provide an improved summary of the vegetation community subtypes (Figure 8).

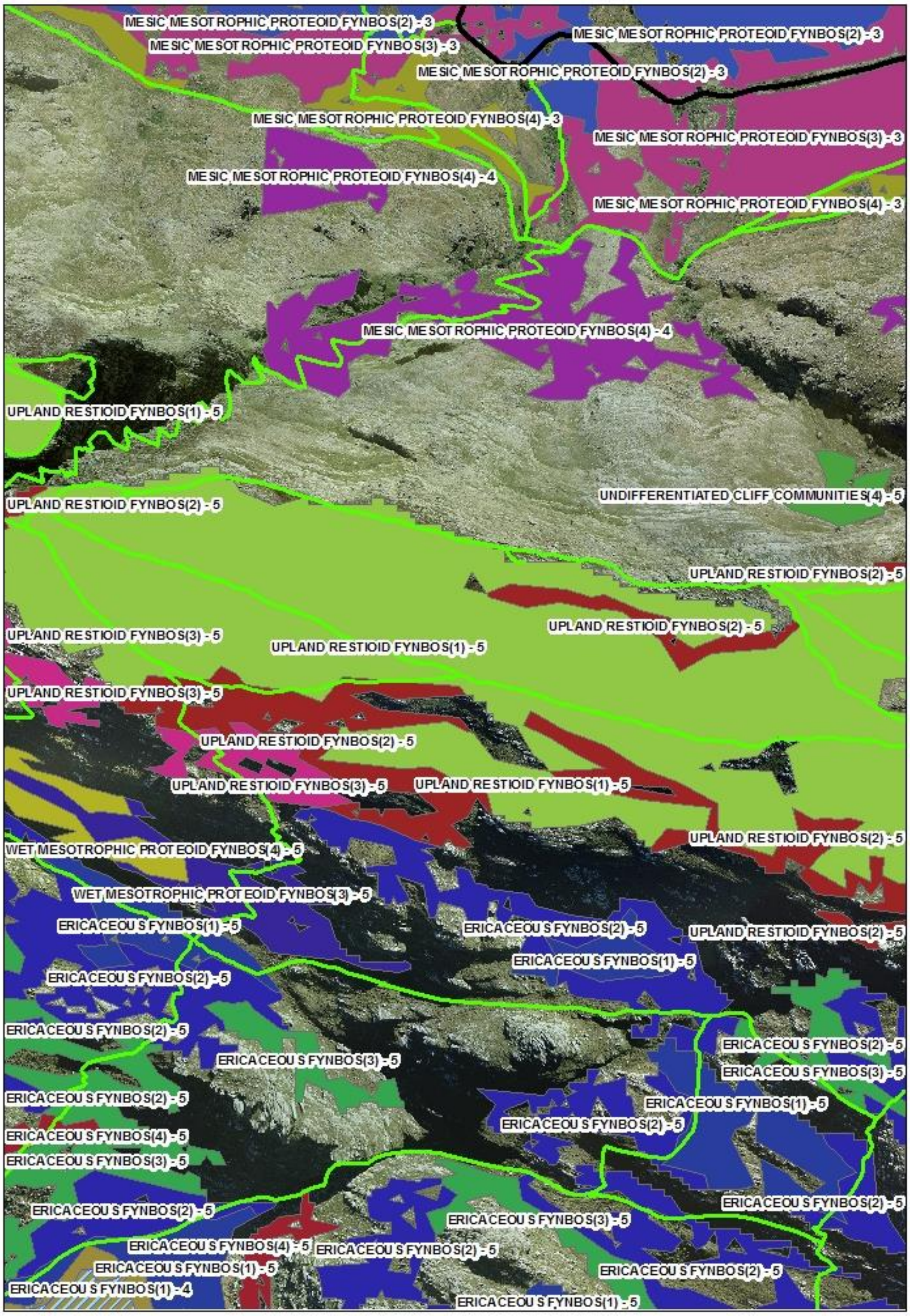


Figure 7: A number of vegetation community subtypes in the northern section of Table Mountain National Park. The first number shows the category or class for slope and the second number the class for altitude.

FID	Shape *	landclasse	Freq1_copy	Areaha
49	Polygon	MESIC OLIGOTROPHIC PROTEOID FYNBOS22	121	699.396
39	Polygon	MESIC MESOTROPHIC PROTEOID FYNBOS32	117	816.63
50	Polygon	MESIC OLIGOTROPHIC PROTEOID FYNBOS23	106	693.42
36	Polygon	MESIC MESOTROPHIC PROTEOID FYNBOS22	94	860.249
40	Polygon	MESIC MESOTROPHIC PROTEOID FYNBOS33	84	558.262
42	Polygon	MESIC MESOTROPHIC PROTEOID FYNBOS42	76	337.59
43	Polygon	MESIC MESOTROPHIC PROTEOID FYNBOS43	76	415.596
53	Polygon	MESIC OLIGOTROPHIC PROTEOID FYNBOS33	72	242.604
46	Polygon	MESIC OLIGOTROPHIC PROTEOID FYNBOS12	70	2772.069
5	Polygon	DUNE ASTERACEOUS FYNBOS22	54	203.823
37	Polygon	MESIC MESOTROPHIC PROTEOID FYNBOS23	53	352.805
47	Polygon	MESIC OLIGOTROPHIC PROTEOID FYNBOS13	51	1354.896
52	Polygon	MESIC OLIGOTROPHIC PROTEOID FYNBOS32	50	146.723
33	Polygon	MESIC MESOTROPHIC PROTEOID FYNBOS12	42	413.931
105	Polygon	WET OLIGITROPHIC PROTEOID FYNBOS24	40	162.115
94	Polygon	WET MESOTROPHIC PROTEOID FYNBOS33	37	176.019
98	Polygon	WET MESOTROPHIC PROTEOID FYNBOS43	36	137.568
56	Polygon	MESIC OLIGOTROPHIC PROTEOID FYNBOS43	35	99.193
69	Polygon	SANDPLAIN PROTEOID FYNBOS22	31	147.003
116	Polygon	WETLANDS12	31	177.254
14	Polygon	ERICACEOUS FYNBOS25	30	98.171
21	Polygon	FOREST AND THICKET12	29	121.574
108	Polygon	WET OLIGITROPHIC PROTEOID FYNBOS34	28	87.597
4	Polygon	DUNE ASTERACEOUS FYNBOS12	27	966.348
62	Polygon	RENOSTERVELD AND GRASSLAND32	26	194.549
13	Polygon	ERICACEOUS FYNBOS24	24	59.458
6	Polygon	DUNE ASTERACEOUS FYNBOS32	22	61.239
102	Polygon	WET OLIGITROPHIC PROTEOID FYNBOS14	22	207.22
27	Polygon	FOREST AND THICKET32	21	45.244
91	Polygon	WET MESOTROPHIC PROTEOID FYNBOS23	21	112.289
93	Polygon	WET MESOTROPHIC PROTEOID FYNBOS32	21	53.456
17	Polygon	ERICACEOUS FYNBOS35	19	48.28
90	Polygon	WET MESOTROPHIC PROTEOID FYNBOS22	19	301.771
24	Polygon	FOREST AND THICKET22	18	191.133
34	Polygon	MESIC MESOTROPHIC PROTEOID FYNBOS13	18	86.996
115	Polygon	WET RESTIOID FYNBOS22	17	37.734
10	Polygon	ERICACEOUS FYNBOS14	16	134.69
11	Polygon	ERICACEOUS FYNBOS15	16	59.215
28	Polygon	FOREST AND THICKET33	16	83.7
60	Polygon	RENOSTERVELD AND GRASSLAND22	16	173.578
3	Polygon	COASTAL SCREE ASTERACEOUS FYNBOS42	15	34.706
66	Polygon	SANDPLAIN PROTEOID FYNBOS12	15	1394.103
87	Polygon	WET MESOTROPHIC PROTEOID FYNBOS12	15	196.228

Figure 8: The frequency tool of ArcGIS, used to produce a frequency table of the number of unique habitat subtypes (119) after the dissolve tool of ArcGIS combined the polygons with the same attributes.

With a more manageable number of subtypes left that can be potentially included for survey routes to intersect them; the next step was to decide how to go about sampling i.e. traversing the study area for klipspringer occurrences. To reduce trampling of vegetation it was decided to sample along existing footpaths (hiking trails). Thus, the footpaths layer played a vital part in the planning process.

]

A total of 13 survey routes were designed across all habitat subtypes, and based on what was deemed representative of the different habitats within the study area (Figure 9). Five routes were in the southern section, three in the central section, and another five in the northern section of TMNP (Figure 9 and Appendix D).

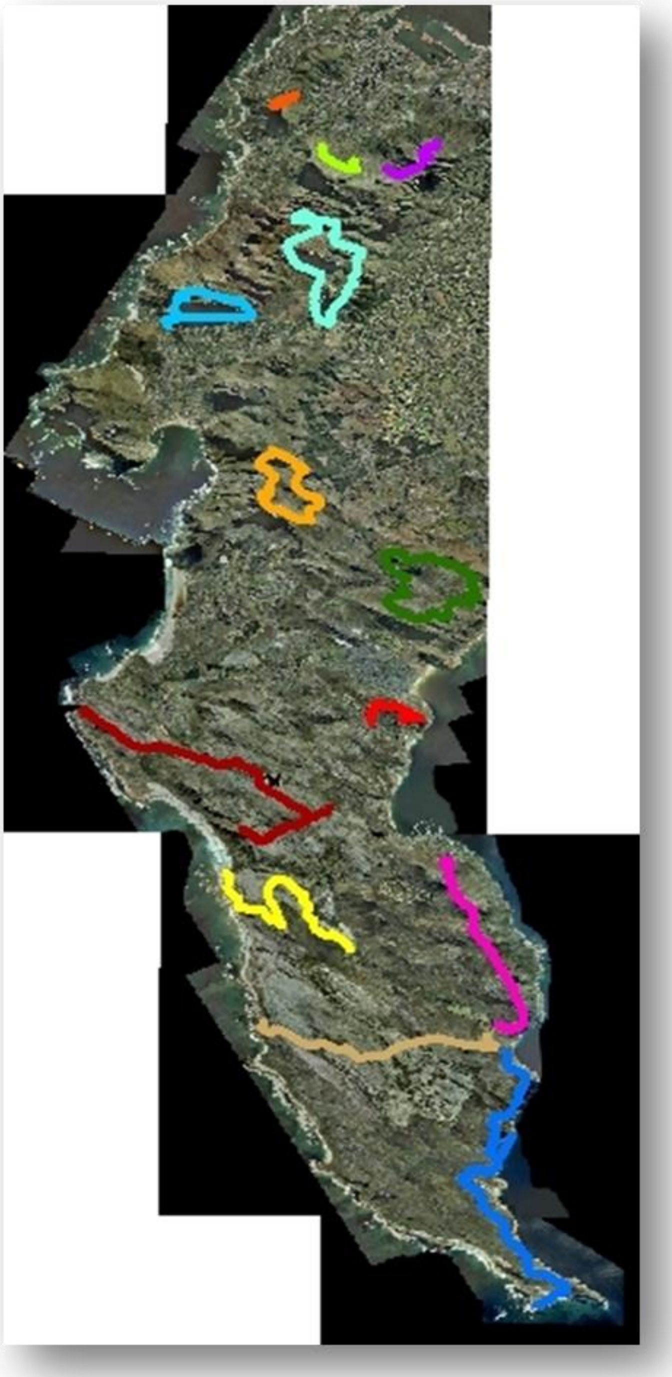


Figure 9: The thirteen survey routes in Table Mountain National Park.

Using the editing tool within ArcMap, each survey route was split into different segments as it was observed that a single route was crossing many different subtypes (Figure 10). This was done to ascertain the exact extent of each habitat subtype being surveyed for klipspringer occurrences. After each survey route was split into its numerous segments, the habitat subtype and distance walked in each of them was recorded in an Excel spread sheet (Appendix E). Some segments crossed some of the eliminated polygons (those <1 ha, >40 degrees slope, or outside the park boundary) since the Park's footpaths crossed many different types of areas (Figure 10). Thus, sampling the deleted polygons was inevitable; especially taking into consideration the numerous steps taken to get to the final 119 subtypes (described above).

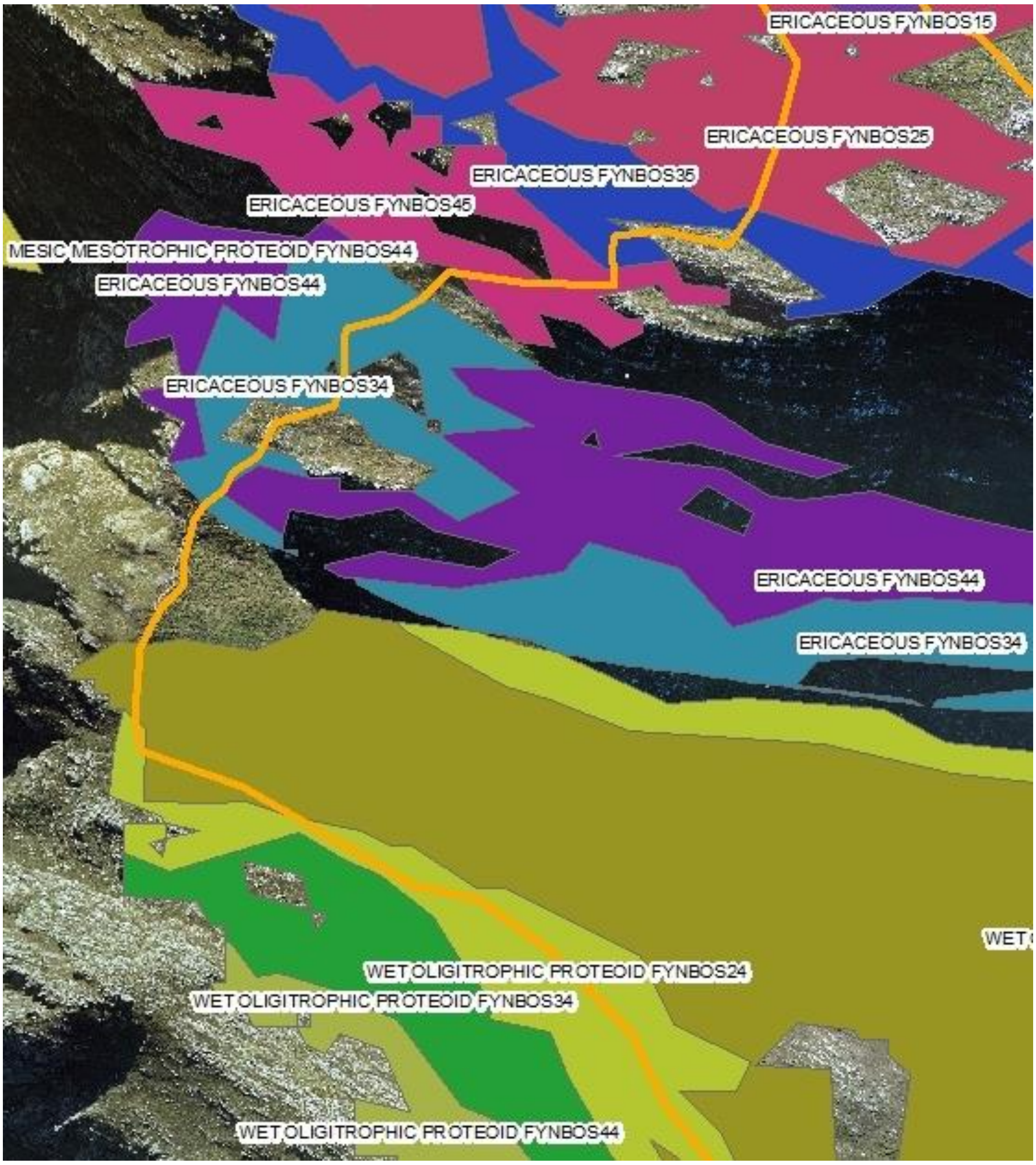


Figure 10: An example of a survey route in Table Mountain National Park crossing many different vegetation community subtypes as well as blank areas that were occupied by the deleted subtypes

.

The segments were summarised using the pivot table tool in Microsoft Excel (Table 3). This summary provided the total distance walked in each vegetation community. Furthermore, the original vegetation community layer was used to determine the proportion of each of the 14 vegetation communities or vegetation types within TMNP (Table 4). In order to establish whether sufficient sampling effort had been put into each vegetation community or type in relation to its size in the Park, a ratio was calculated by dividing the proportion of each vegetation community or type by the distance walked in each of them (Table 3). A ratio close to 1 would indicate sufficient sampling, therefore most vegetation communities received sufficient sampling except for ericaceous fynbos and wet oligotrophic proteoid fynbos. Ericaceous fynbos already had a high number of occurrence points recorded within it (i.e. dense cluster of points in the southern section) therefore no additional survey effort was added in this vegetation community. For wet oligotrophic proteoid fynbos on the other hand no additional survey effort was possible due to logistical constraints.

Fourteen field surveys were carried out between 23 July 2014 and 24 November 2014. Thus, surveys largely occurred in the winter and spring seasons. Essential materials used included a Garmin eTrex 30 GPS device and a printed A3 map of each section of the study area, each containing a survey route (Appendix F). These maps were gridded with latitude-longitude coordinates to assist with navigation in the field and thereby ensuring the exact survey route was followed. Furthermore, these gridded maps also served to assist in pinpointing and marking a location where an individual was sighted a distance away from the observer and thereby an occurrence point could still be recorded. Locality points were recorded for individual sightings, recognisable spoor and dung middens or dung pellets (Appendix G). Where a number of spoor records were close to one another, the points recorded were spaced out according to the discretion of the researcher to what was deemed reasonable. All relevant data were recorded onto a field data sheet. This field work added a total of 57 additional occurrence records in eight vegetation communities (eleven habitat subtypes with some points within the deleted polygons, recorded in Table 5). Figure 11 shows the distribution of the aforementioned points together with the collar data and other data sourced either from Ispot, CPUT alumnus, or the manually collected points from the previous researcher.

Table 3: The total distance walked in each major vegetation community or vegetation type of Table Mountain National Park as well as the ratio that this distance walked represents for each vegetation community.

Vegetation community or type	**Total length (km)**	**Ratio**
Mesic oligotrophic proteoid fynbos	34.77	0.80
Mesic mesotrophic proteoid fynbos	12.27	1.74
Wet restioid fynbos	6.61	1.86
Sandplain proteoid fynbos	3.33	2.11
Wet mesotrophic proteoid fynbos	6.89	0.84
Dune asteraceous fynbos	3.52	1.60
Ericaceous fynbos	8.77	0.46
Forest and Thicket	0.7	5.63
Wet oligotrophic proteoid fynbos	8.94	0.42
Renosterveld and associated grasslands	1.88	1.77
Undifferentiated cliff communities	1.93	1.41
Wetlands	0.10	8.23
Coastal scree asteraceous fynbos	0.15	5.38
Upland restioid fynbos	0.41	1.45

Table 4: The proportion of each vegetation community or vegetation type within Table Mountain National Park from largest to smallest.

Vegetation community or type	**Area (ha)**	**Area (km²)**	**Proportion in study area (%)**
Mesic oligotrophic proteoid fynbos	7064.128	70.641	27.8
Mesic mesotrophic proteoid fynbos	5409.735	54.097	21.3
Wet restioid fynbos	3121.441	31.214	12.3
Sandplain proteoid fynbos	1785.895	17.859	7.0
Wet mesotrophic proteoid fynbos	1470.84	14.708	5.8
Dune asteraceous fynbos	1429.898	14.299	5.6
Ericaceous fynbos	1028.094	10.281	4.0
Forest and Thicket	1001.164	10.012	3.9
Wet oligotrophic proteoid fynbos	948.674	9.487	3.7
Renosterveld and associated grasslands	845.571	8.456	3.3
Undifferentiated cliff communities	693.335	6.933	2.7
Wetlands	213.283	2.133	0.8
Coastal scree asteraceous fynbos	202.303	2.023	0.8
Upland restioid fynbos	149.02	1.49	0.6
Vlei (water bodies included as part of shapefile)	32.515	0.325	0.1
Total	**25395.9**	**253.958**	**100**

Table 5: Number of occurrence points recorded in the different vegetation communities and habitat subtypes in the 2014 survey of klipspringer occurrence in Table Mountain National Park.

Vegetation community or type	**Habitat subtype**	**No. of points**
Coastal asteraceous fynbos	Coastal asteraceous fynbos 1 2	1 (Marginal with mesic oligotrophic proteoid fynbos)
Ericaceous fynbos	Ericaceous fynbos 4 4	2
Ericaceous fynbos	Ericaceous fynbos (deleted polygons)	12
Forest and Thicket	Forest and Thicket (deleted polygon)	1 (Marginal with ericaceous fynbos)
Mesic oligotrophic proteoid fynbos	Mesic oligotrophic proteoid fynbos 2 3	2
Mesic oligotrophic proteoid fynbos	Mesic oligotrophic proteoid fynbos 2 2	12
Mesic oligotrophic proteoid fynbos	Mesic oligotrophic proteoid fynbos 1 2	3
Mesic oligotrophic proteoid fynbos	Mesic oligotrophic proteoid fynbos 3 2	1
Mesic oligotrophic proteoid fynbos	Mesic oligotrophic proteoid fynbos 4 4	1
Mesic oligotrophic proteoid fynbos	Mesic oligotrophic proteoid fynbos (deleted polygons)	5
Mesic mesotrophic proteoid fynbos	Mesic mesotrophic proteoid fynbos 2 2	2
Mesic mesotrophic proteoid fynbos	Mesic mesotrophic proteoid fynbos 3 3	1
Mesic mesotrophic proteoid fynbos	Mesic mesotrophic proteoid fynbos (deleted polygons)	9
Undifferentiated cliff communities	Undifferentiated cliff communities (deleted polygon)	1 (Marginal with ericaceous fynbos)
Wet oligotrophic fynbos	Wet oligotrophic fynbos 1 4	3
Wet restioid fynbos	Wet restioid fynbos 2 2	1 (Marginal with mesic oligotrophic proteoid fynbos)
Total	**16**	**57**

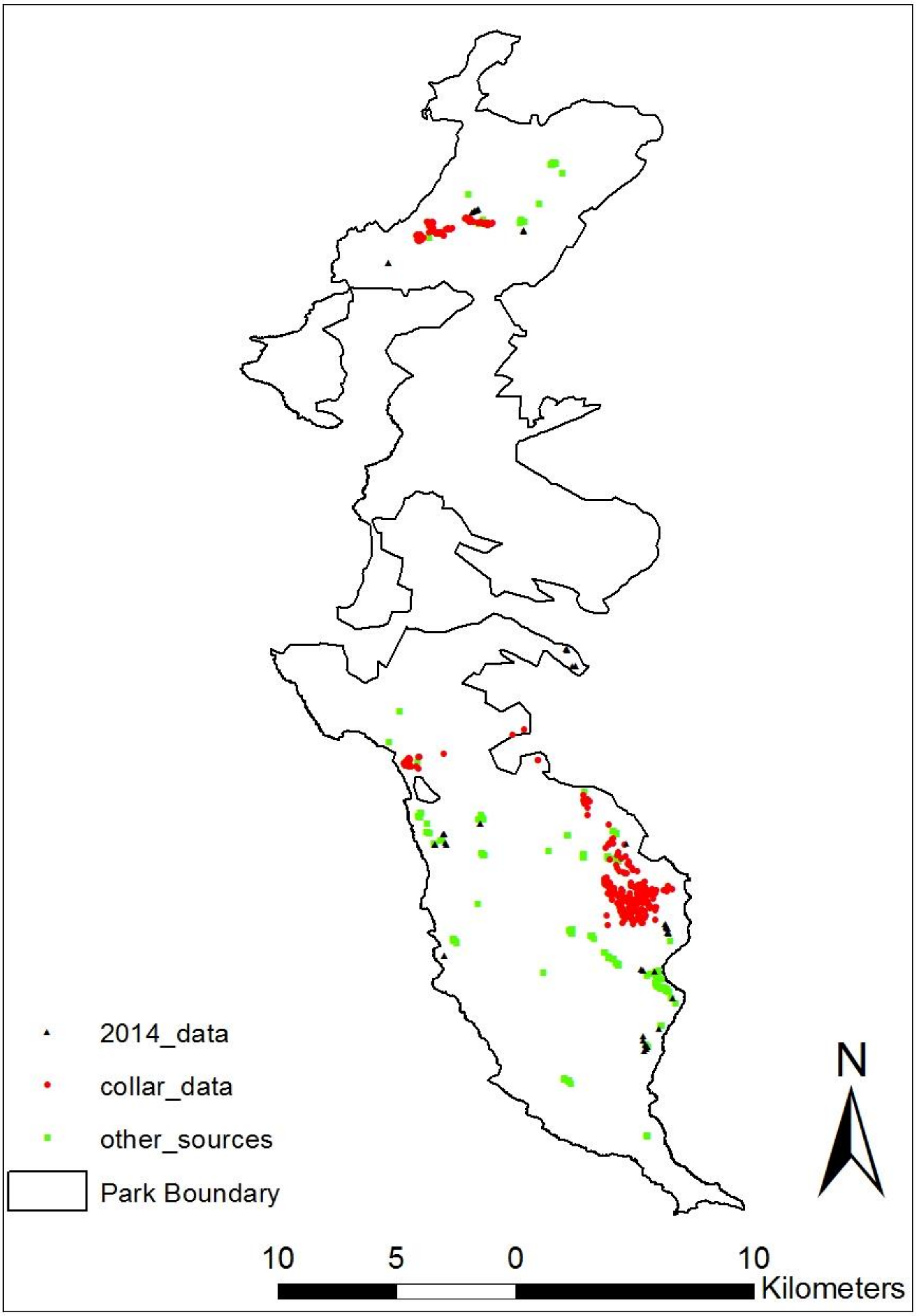


Figure 11: All occurrence point data used in this study recorded in Table Mountain National Park. The Park boundary refers to the general outline of the Park, but does not show the exact portions within this general boundary that is actually protected land under Table Mountain National Park.

3. Data analysis

3.1 Home range analysis

The kernel density function of the Geospatial Modelling Environment (GME) software was used to determine the possible extent of the klipspringer's home range. Kernel density estimation (KDE) was developed from the histogram, as a means of smoothing the frequency distribution (Figure 12) (Duong, 2001) KDE output presents a probability density surface as contours. This probability density estimate is largely dependent on the bandwidth (width of a kernel), also referred to as the smoothing parameter (Worton, 1989; Hemson *et al.,* 2005), as well as on the two dimensional (latitude-longitude) occurrence points. Several ways of calculating the bandwidth size are available in GME, which are based on algorithms that calculate the density as well as incorporating error (Table 6). These include the smoothed cross validation (SCV), biased cross-validation (BCV), a second BCV algorithm (BCV2), plug-in estimator (PLUGIN), least squares cross validation (LSCV) and the likelihood cross validation (CVh). All bandwidth algorithms in the GME are calculated as rotated bivariate kernels, except CVh which are computed as a univariate kernel (Rudeen, 2012). Horne & Garton (2006), recommended CVh as a good bandwidth estimator to use, taking into consideration its statistical properties which to some extent deal with autocorrelation, and location acquisition bias which ultimately results in better estimates of a species' utilisation distribution. In GME, a fixed kernel algorithm is applied for all bandwidth algorithms, thus all occurrence points are weighted equally with the same smoothing parameter (Beyer, 2014). This approach contrasts with an adaptive kernel where bandwidths are varied based on the density of points (Worton, 1989).

All six of the above mentioned bandwidth algorithms were tested on the klipspringer data in order to decide on the best algorithm to use (Beyer, 2014) (Figure 13). The SCV and CVh bandwidth provided the best output (i.e. not over generalised, Figure 13a and 13b) of the klipspringer's home range, followed closely by PLUGIN (Figure 13c). The BCV, BCV2 and LSCV overestimated the home range (Figures 13d, e and f). The output raster cell size was tested for values of 0.5; 5 and 10 m. The cell size was found to have no effect on the output based on visual inspection of how the output contours fit the location points (Figure 14). The output raster cell size was thus set at 10bm as this is the cell size of the DEM, and it matches the accuracy of the 1:10 000 layers too.

GME has three kernel type options namely the Gaussian, quadratic, and uniform kernel. The default Gaussian kernel was used. Several isopleth runs were also conducted in GME to visualise the utilisation distribution (percentage area utilised) and decide on the optimal UD. These runs were conducted at the 50%, 60%, 70%, 75%, 80%, 85% and 95% contour level for the two bandwidth algorithms that gave best output. The optimal utilisation distribution was set at 50% for the bivariate kernel (SCV) and 80% for the univariate kernel (CVh). This decision was based on the contours that visually present the home ranges best.

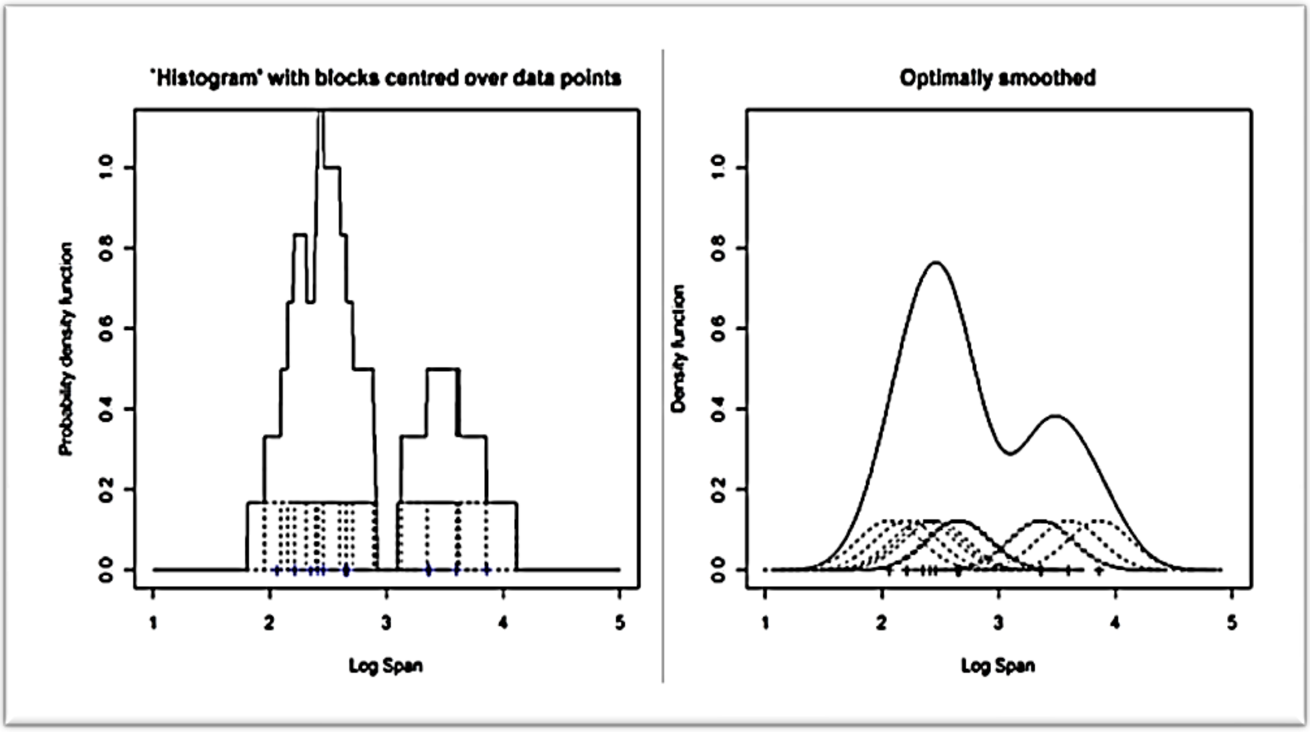


Figure 12: An example of the conversion of the histogram to an optimally smoothed kernel density estimate. (Duong, 2001).

Table 6: Bandwidth algorithms in GME and the type of kernel each of these were used as (adapted from Rudeen, 2012).

Bandwidth algorithm	**Kernel type 1**	**Kernel type 2 (determines shape of kernel)**	**Description of kernel type 2**
Smoothed cross validation (SCV)	Gaussian	Rotated bivariate	Determined by the bandwidth of the X dimension, Y dimension and the covariance of XY. The rotated kernel is a result of the covariance of X and Y.
Biased cross-validation (BCV)	Gaussian	Rotated bivariate	Determined by the bandwidth of the X dimension, Y dimension and the covariance of XY. The rotated kernel is a result of the covariance of X and Y.
Second BCV algorithm (BCV2)	Gaussian	Rotated bivariate	Determined by the bandwidth of the X dimension, Y dimension and the covariance of XY. The rotated kernel is a result of the covariance of X and Y.
A plug-in estimator (PLUGIN)	Gaussian	Rotated bivariate	Determined by the bandwidth of the X dimension, Y dimension and the covariance of XY. The rotated kernel is a result of the covariance of X and Y.
Least squares cross validation (LSCV)	Gaussian	Rotated bivariate	Determined by the bandwidth of the X dimension, Y dimension and the covariance of XY. The rotated kernel is a result of the covariance of X and Y.
Likelihood cross validation (CVh)	Gaussian	Univariate	Uses X dimension only. A circular kernel with the same dimension is created in all directions around each location.

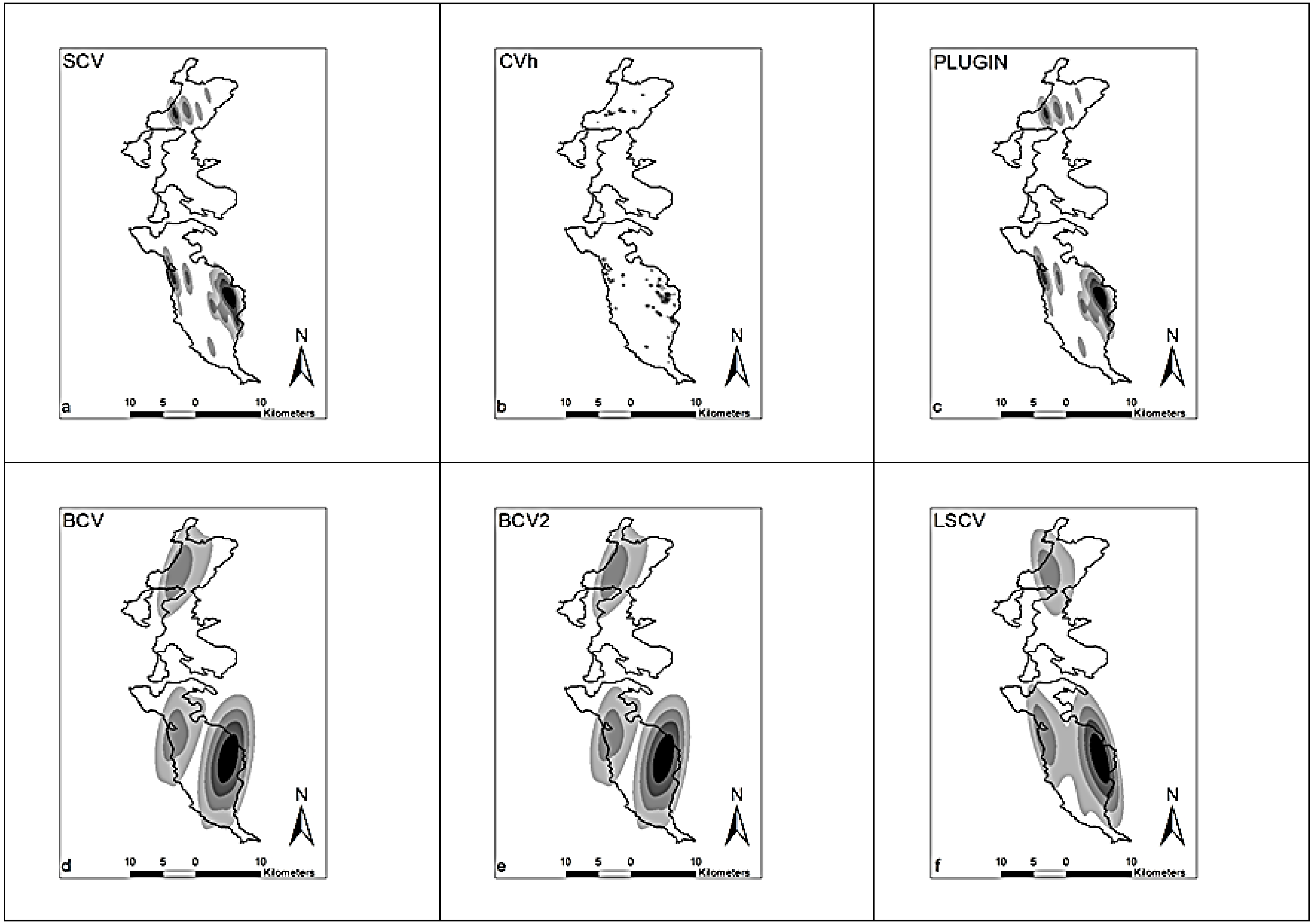


Figure 13: Kernel Density Estimation (KDE) output for all six bandwidth algorithms of the Geospatial Modelling Environment (GME) tested on the same locality point data: SCV (a), CVh (b), PLUGIN (c), BCV (d), BCV2 (e) and LSCV (f). An increase in density is shown as the contours get darker (grey to black).

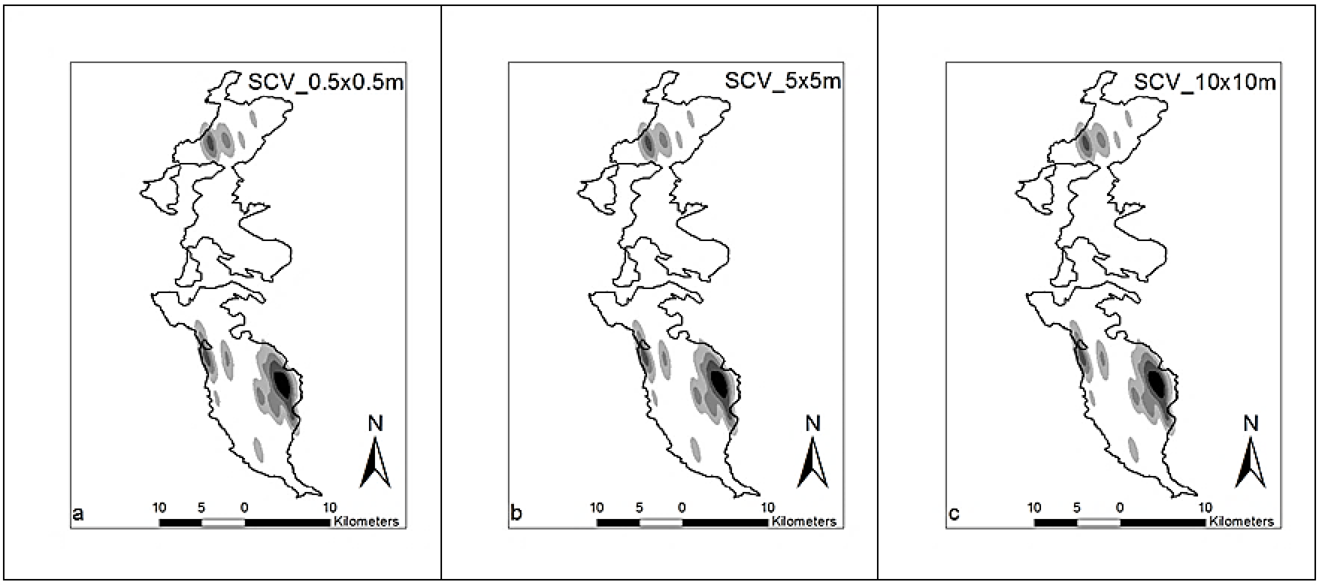


Figure 14: Kernel Density Estimation (KDE) output on three different cell sizes: 0.5x0.5 m (a), 5x5 m (b) and 10x10 m (c) using the SCV bandwidth algorithm. Note that an increase in cell size does not result in any visible difference in the density estimation.

GME is coded and runs simulations in the R programme language (R Core Team, 2013). Input is in the form of a point locality shapefile that has to be in the projected coordinate system for conducting a kernel density estimation. A kernel function (weighting function in the form of a normal curve) is placed over each locality point (Silverman, 1986; Worton, 1989) (Figure 12). The sum of all kernels at each locality point recorded gives an estimate of the density within the study area (Silverman, 1986). The kernel density is produced as contour lines that give an indication of the total area in which the species possibly moves. Therefore, it gives an approximation of which areas the species might be occupying, although not observed at the time of data collection. It can be concluded that the KDE function provides a higher probability of occurrence.

3.2 Data preparation for MaxEnt

3.2.1 Occurrence data

All occurrence points were recorded in an excel spreadsheet. A point layer file was created in ArcMap to visualise the distribution of locality points within the Park (Figure 15). This layer was also used for the home range analysis carried out.

MaxEnt is sensitive to sampling bias (Phillips *et al.*, 2009; Kramer-Schadt *et al.,* 2013; Fourcade *et al.*, 2014). In an attempt to reduce sample selection bias e.g. uneven sampling effort, a spatial filtering approach was undertaken. The most noticeable bias in the occurrence points were in the collar data where a high number of points were recorded at the release sites before the species had spread out properly in their new area (Figure 15). The filtering approach applied entails reducing the number of points to one point per 10x10 m or 100 m^2 grid cell. This is similar to the one point per grid approach followed by Elith *et al.* (2010) to reduce dense clusters of points although they worked at a much larger scale. Thus all cells that had more than one occurrence point were highlighted and these extra points were deleted manually using the editing tool in ArcMap. This reduced the number of points from 845 to 608. Thereafter the point layer was converted to an Excel spread sheet. Visual scrutiny of this Excel spread sheet revealed that some duplicate points existed, as a consequence of these points that would have been exactly on top of each other in ArcMap and therefore it would appear as one point per grid when in fact there are two. Existing duplicate points were deleted manually in Excel which further reduced the number of locality points to 591. To prepare the data for import into MaxEnt (version 3.3.3k), column headings in excel were set in the following order, in order for the MaxEnt Geographic User Interface (GUI) to read it: Species, Longitude, and Latitude. The spreadsheet was saved as a comma-delimited (comma separated value, csv) file.

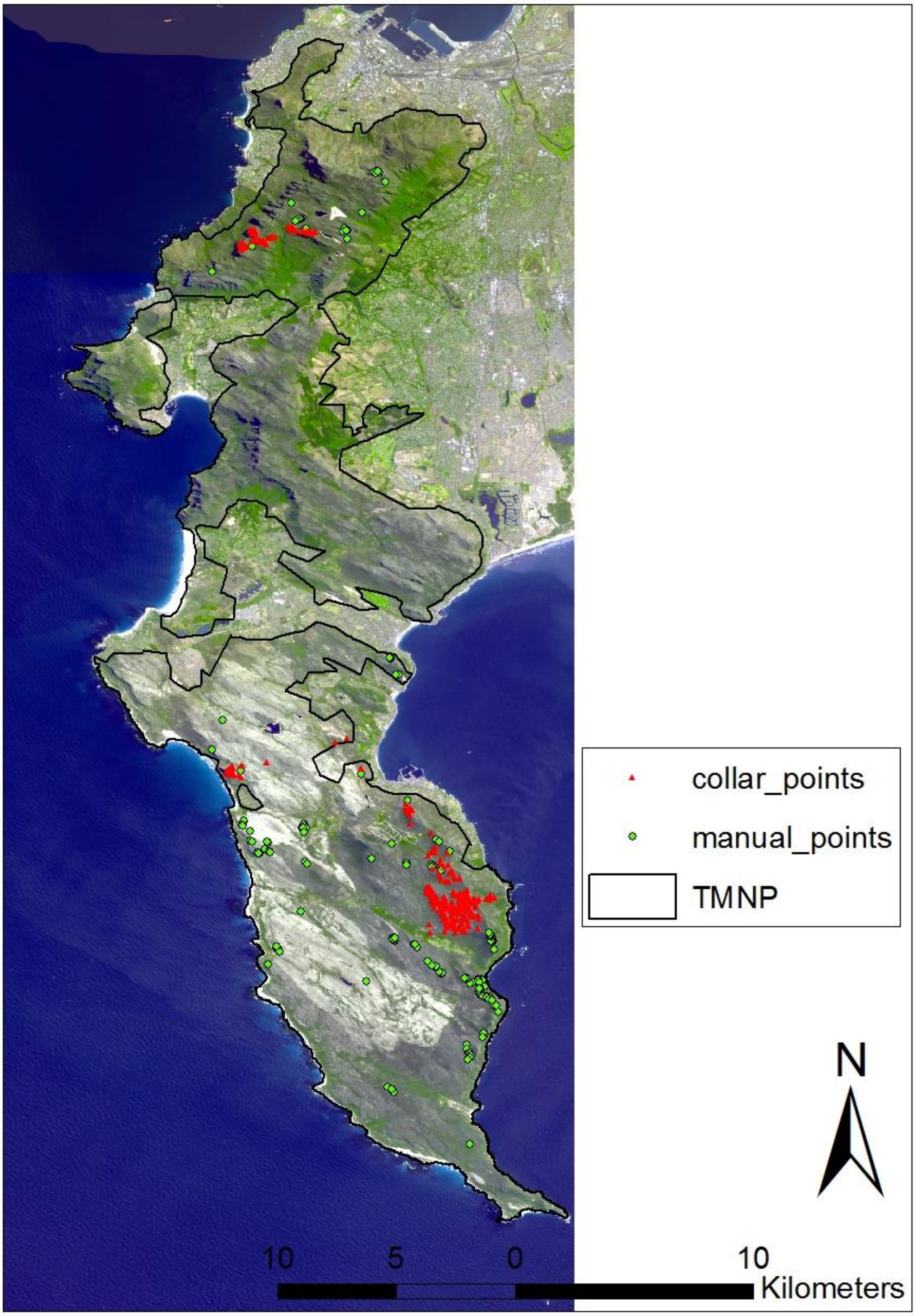


Figure 15: All occurrence points (GPS collar and manually collected) within Table Mountain National Park.

3.2.2 Environmental data

All environmental variable raster layers have to be in the geographic coordinate system and be converted to ASCII files (.asc) in order to be used in MaxEnt. Furthermore, all layers need to have the same geographic extent as well as the same cell size. ArcMap (ArcGIS 10.1) was used to prepare all the necessary layers. As mentioned in Section 2.2 of this chapter, a 10x10 m DEM was used as the base for altitude and to derive slope as well as aspect from using the spatial analyst tool of ArcGIS (Appendix C). The original DEM layer was first clipped to the study area using the extent of an existing vector layer that represented the vegetation of the Cape Peninsula as created by Cowling *et al.* (1996) (Figure 16). Note that the vector layer that the DEM was clipped to, extend beyond the exact boundary of TMNP as this boundary is not one complete connected area, but rather consist of disjointed sections as shown in Figure 4.

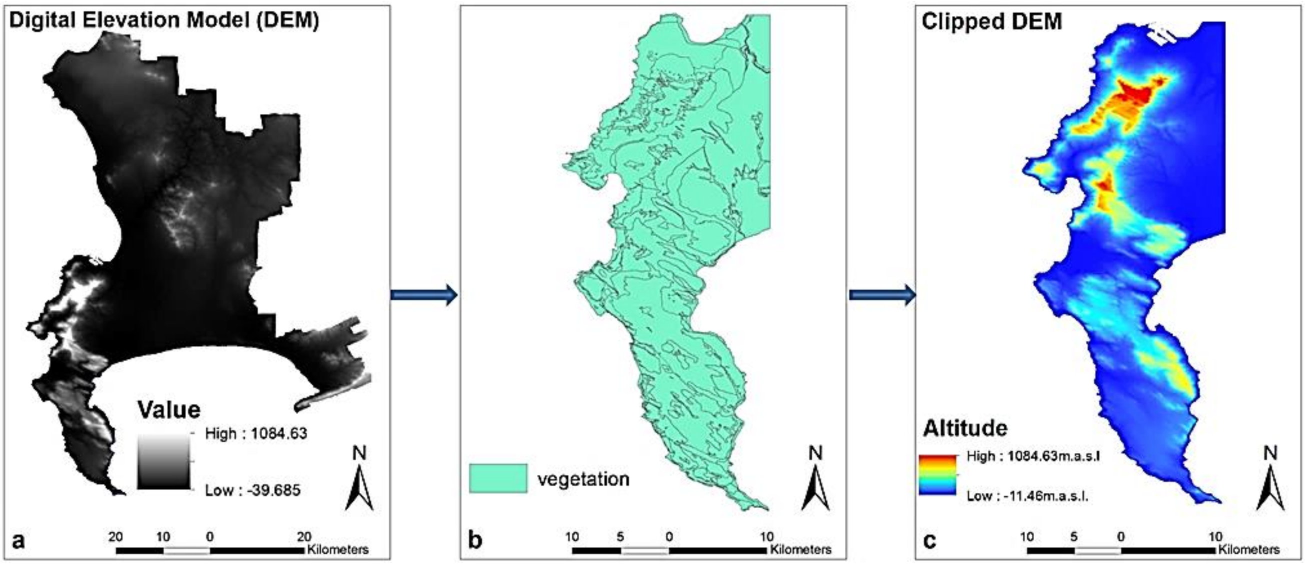


Figure 16: Clipping process of the Digital Elevation Model (DEM) to the general extent of the study area. A DEM (a) covering a greater area was clipped to the extent of a vegetation polygon layer (b) to produce a DEM for the study area (c).

As the vegetation community layer was in polygon (vector) format, it had to be converted to a raster first before it was converted to an ASCII file. The majority rule was used when converting the vegetation community polygon layer into a 10x10 m raster. This is the only categorical variable and subsequently this has to be specified within the MaxEnt GUI. Average monthly rainfall data were sourced from the global climate database, WordClim (Hijmans *et al.,* 2005), and prepared through masking and clipping it to the study area. The distance to urban edge layer was created using the euclidean distance tool in ArcGIS. A polygon layer representing the urban areas surrounding the

Park was used for this calculation (Appendix H). This layer had to be in a projection (used Transverse Mercator Lo (19)) for the euclidean distance to be calculated where after it was ‘unprojected’ into ‘geographic projection’ for use in MaxEnt.

Covariance among variables should be avoided in any statistical analysis (Field, 2005). MaxEnt is no exception (Kumar & Stohlgren, 2009; Boria *et al*., 2014), and when two variables are highly correlated with one another MaxEnt will use the one that is listed first and not incorporate the other variable to its fullest capacity in the model simulation. Thus variables were tested for spatial autocorrelation, specifically the relationship between the climatic variables and the topographical variables, as these were the variables most likely to be correlated. Spatial autocorrelation occurs in geographic space when environmental variables and geo-referenced locality points lack independence from one another (Legendre, 1993). A two tailed Pearson’s correlation (Field, 2005) was used in the statistics software, SPSS (IBM SPSS Statistics 22) to test for autocorrelation between all individual independent variables.

Rainfall for each month showed over 50% correlation with altitude as well as with one another (Appendix I). With $p<0.01$ for the aforementioned variables (Appendix I), it indicates that there is a highly significant relationship between these variables. Altitude has a higher contribution to habitat suitability for klipspringers (Skinner & Chimimba, 2005) compared to rainfall and therefore rainfall was more likely to be excluded from the model. Before the final exclusion, however, a model was run with rainfall included and a second model without rainfall to test how model performance varies with rainfall included and excluded. Both these models were run at 1 km^2 as this is the resolution that the rainfall data was at and could therefore not be resampled into a smaller scale. AUC values, which indicates model performance, changed from 0.835 (when MaxEnt models were run with the monthly rainfall variables) to 0.751 (when MaxEnt models were run without the monthly rainfall variables). This further indicates that the model still performed well in predicting suitable habitat without the rainfall being included as a variable.

3.2.3 Cell size/scale

The grid cell size at which a species distribution model is run plays an important role, because when the cell size that environmental variable layers are prepared at, are too coarse, it might result in the loss of vital information, e.g. microclimatic conditions or specific topographical features and vegetation compositions. On the other hand, a grid cell size that is too fine might results in non-real patchiness (Francis & Klopatek, 2000). Model runs were carried out with four environmental variables at three different spatial scales to determine what the most appropriate cell size for this specific study should be. To do this the four environmental variables were prepared in ArcGIS at three different cell sizes namely 10 m (initial cell size), 100 m and 1000 m using the raster resampling tool, which is based on the majority rule. The regularisation multiplier was kept at default; 1 and 10 replicate runs were carried out. From the MaxEnt output of AUC and omission rates, a cell size of 10 and 100 m gave better results (Figures 17 and 18). This was concluded from the lower AUC value at 1000 m (0.691) compared to the AUC value of the two smaller cell sizes which were both 0.893 (Figure 17). This is almost a 20% difference in model performance between the 10 m and 1000 m cell size as well as the 100 m and 1000 m cell size. The mean and standard deviation (red and blue line in Figure 17) for the 1000 m cell size moving closer to the random prediction (black line), indicates that this model performance is poorer than the other two models of smaller cell size.

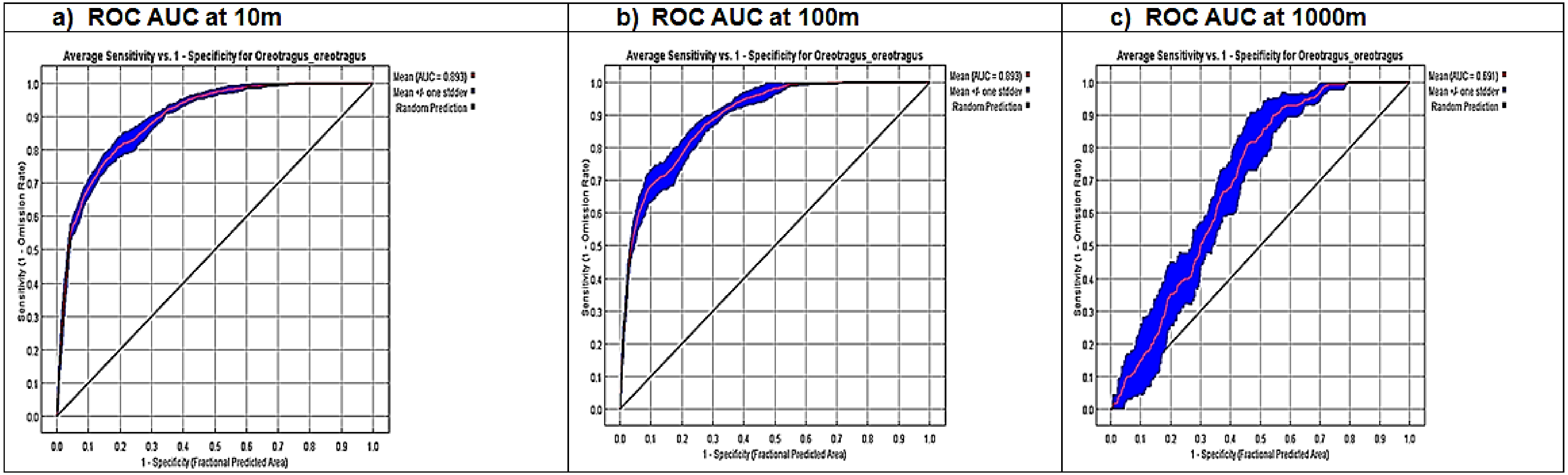
Omission error was also higher at 1000 m and departs from the predicted omission line (Figure 18). The 1000 m probability map also did not give the best visual output of the potential distribution of suitable areas, with almost no cells having a probability of greater than 70% and a larger questionable part of the Park appeared to be suitable or tending towards suitability from this output (Figure 19). The 10 m and 100 m cell size showed the suitable habitats in better detail; thus improving the ability to distinguish between areas with a high predicted probability of occurrence and those with a low predicted probability. Response curves for the two smaller cell sizes also showed in better detail which values of a particular variable, e.g. altitude, corresponds positively to a predicted presence and therefore suitable habitats (Figure 20). The 1000 m model's altitude curve was over generalised thus showing underfitting and clear conclusions cannot be drawn on how the species responds to this variable.

Additional considerations with regards to cell size, that will give sufficient information for the species under study, were done through incorporating the territory size of klipspringers. Norton (1980) studied the habitat and feeding ecology of klipspringer and recorded that territory size can be affected by the annual rainfall an area receives. For instance, in Norton's Springbok study area, which is a semi-arid winter rainfall region receiving a mean annual rainfall (MAR) of about 160 mm, it was recorded that territory sizes can be up to 50 ha, whilst in Gamka (medium rainfall area with MAR of between 350 – 450 mm) the territory size was <15 ha. Therefore, for the Cape Peninsula, which receives a MAR of between 520 – 1690 mm, the territory size was reasoned to be smaller than 15 ha, about 10 ha and below. A simple calculation, using the 10 ha territory size, to determine the appropriate cell size that will capture all the environmental conditions in a Cape Peninsula klipspringer territory, was performed:

10 000 m^2 x10 = 100 000 m^2

square root of 100 000 m^2 = 316 m; rounded off to 300 m.

A sufficient cell size would then be 300x300 m and below. Both the 10x10 m and 100x100 m cell sizes were tested and were shown to perform well. Consequently the cell size of 10x10 m was selected to match all environmental layers (rasters), as there was no clear impetus to generalise these data to 100x100 m for this modelling exercise.



a) ROC AUC at 10m

b) ROC AUC at 100m

c) ROC AUC at 1000m

Figure 17: The area under the receiving operating characteristic curve (ROC AUC) results for three models at cell size 10 m (a), 100 m (b) and 1000 m (c). These curves give a measure of a model's ability to classify an area as suitable or not. Both AUC values for 10 m and 100 m were given as 0.893 (indicating good model performance) and 1000 m was given as 0.691, which is still better than random but not as good as the two smaller cell sizes.

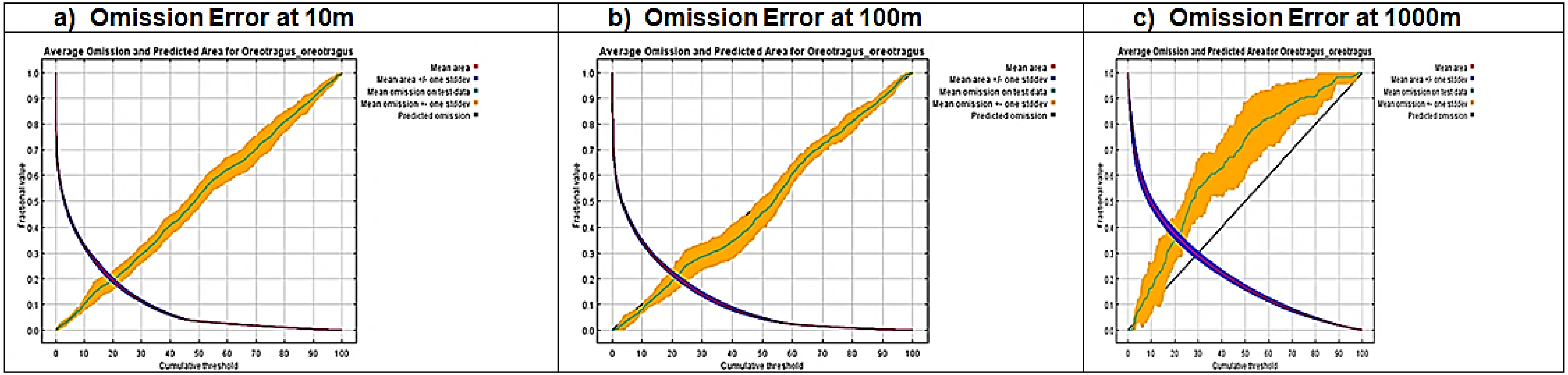


Figure 18: Omission error for the three models at cell sizes 10 m (a), 100 m (b) and 1000 m (c). An omission error close to the predicted omission (black line) reflects a good model. Clearly the 1000 m cell size had a higher omission error than what was predicted for that model and consequently was not such a good model.

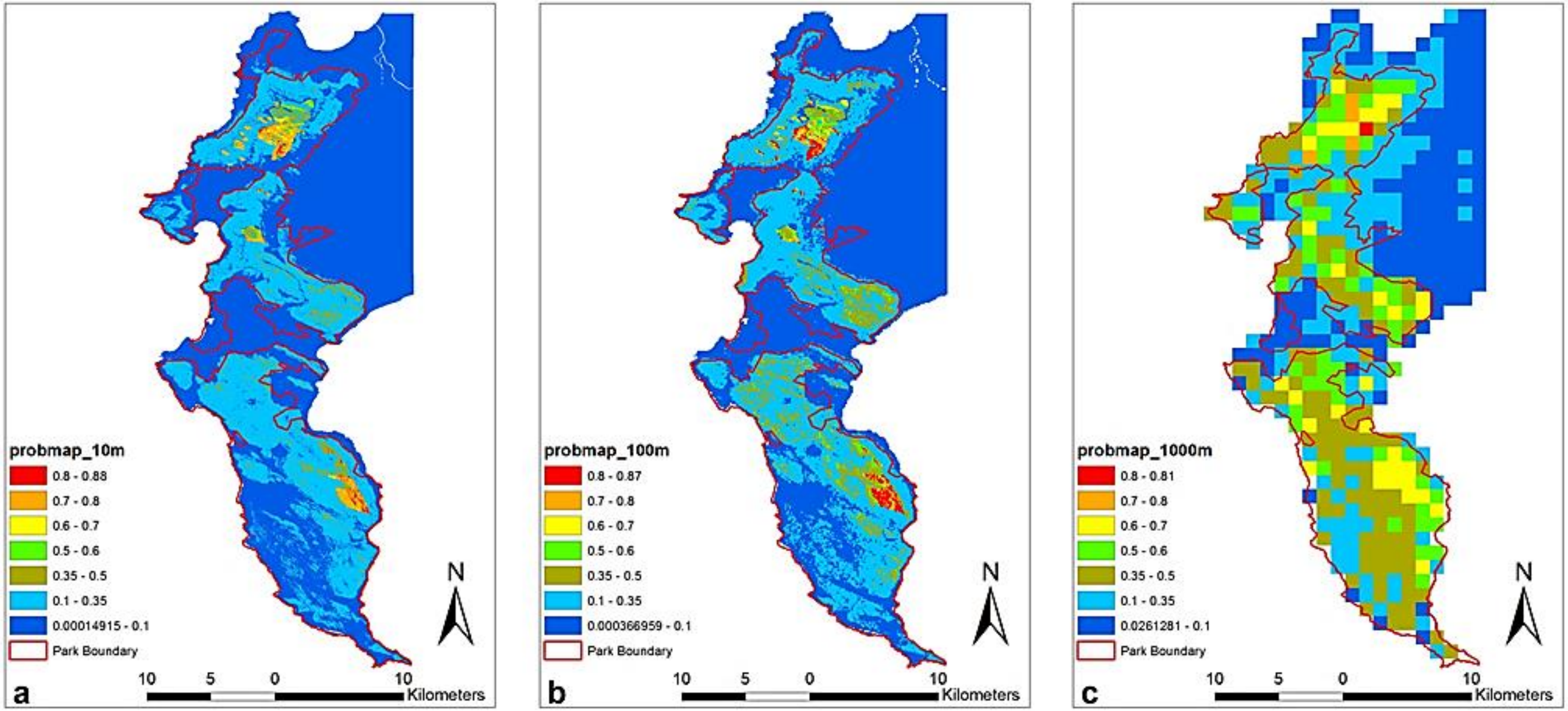
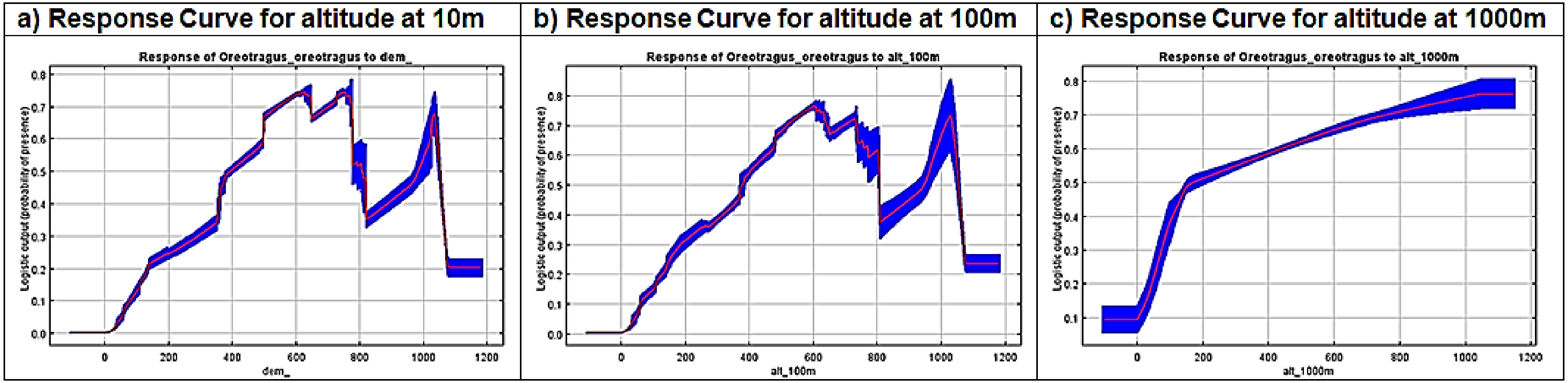


Figure 19: Probability of occurrence maps for the 10 m (a) and 100 m (b) cell size were similar. In contrast the 1000 m (c) cell size over predicts the occurrence probability and therefore gives an incorrect display of suitable habitat.



a) Response Curve for altitude at 10m

b) Response Curve for altitude at 100m

c) Response Curve for altitude at 1000m

Figure 20: The response curve for altitude at a cell size of 10 m (a) and 100 m (b) displayed the altitudinal range where a presence can occur more clear than what the 1000 m (c) cell size does which over-smoothed the response.

3.3 MaxEnt settings

It is important to do more than one simulation on the MaxEnt model as the tool selects the points used for training and testing the model at random. Since each simulation generates different output results, the combined average of all the output results is used. The number of result replicates in MaxEnt has been limited to 10 as in Phillips (2006). The occurrence data was split into 70% training data and 30% testing data to evaluate the model. Background data was left at the default of 10000 as in Phillips *et al.* (2006). Phillips & Dudik (2008) also found that this number of background points was sufficient for their occurrence points that ranged from two to 5822. Thus it is presumed that for the 591 occurrence points in this study, 10 000 background points would suffice as well. For the feature types, the auto features option was selected since more than 80 occurrence points are available and therefore the model will incorporate all features. In the study by Phillips & Dudik (2008) the default regularisation multiplier of 1 delivered good results on a number of presence data sets. A default regularisation value exists for each feature type, however, when choosing a value for the regularisation multiplier, that value is used instead of the default (Phillips *et al.,* 2006; Warren & Seifert, 2011). To test the effect of different regularisation multiplier values this study ran six models at the following values: 0.25, 0.5, 1, 2, 4, and 7 at 10 replicate runs each. Results were inspected to decide which value presented the probability distribution best and are discussed in Chapter 4.

Prevalence, the proportion of sites in which a positive observation of occurrence can be made (Phillips and Elith, 2013), was left at the default value of 0.5. A prevalence of 0.5 would mean that the species is likely to be found in 50% of the study area compared to not setting a prevalence value, which might indicate that the species can occur anywhere in the study area (Phillips and Elith, 2013). Prevalence cannot be determined with occurrence data only (Phillips *et al*., 2006) and as a consequence only the probability distribution can be estimated.

The MaxEnt parameter ‘threshold’ determines a value which allows a binary classification of suitable and not suitable habitat. Any value above the calculated threshold equates to suitable sites, and values below the threshold will indicate unsuitable habitat. MaxEnt model output provides threshold values in the “MaxEntResults.csv file” to select from. The two most common thresholds, namely minimum presence training logistic threshold and the 10 percentile training presence logistic threshold, was tested (Pearson *et al*., 2007; Escalante *et al.,* 2013; Wakie *et al*., 2014). The former threshold is based on the lowest predicted value for any of the pixels with an occurrence record in it (Pearson, 2007). The latter threshold ensures that the predicted suitable habitat includes at least 90% of the calibration data, i.e. 90% of the occurrence points has been correctly predicted as present and omission error does not exceed 10% (Young *et al.,* 2011).

CHAPTER THREE: RESULTS AND DISCUSSION: HOME RANGE ANALYSIS

The smoothed cross validation (SCV) and likelihood cross validation (CVh) bandwidth algorithm provided the best visual output for the specific set of occurrence data used and will be the only algorithms reported on in this chapter (Figures 21 and 22). Home range sizes from the SCV output ranged from about 3 km to 11 km across the study area. This equates to a 3 - 11 ha home range size. The univariate CVh bandwidth provided many smaller and more localised home ranges (Figure 22). Home range size from the latter bandwidth algorithm ranged from 0.6 km to 2.5 km, which equates to a 0.6 - 2.5 ha home range size. These home ranges of less than 15 ha relates to the findings of Norton (1980) for a klipspringer home range in a similar habitat. Noticeable in the study by Norton (1980) was that the amount of rainfall an area received was directly related to a klipspringer’s territory size. For instance, in Norton’s Springbok study area, which is a semi-arid winter rainfall region receiving a mean annual rainfall of about 160 mm, it was recorded that territory sizes can be up to 50 ha, whilst in Gamka (medium rainfall with MAR of between 350 – 450 mm) the territory size was <15 ha. Therefore, for the Cape Peninsula, which receives a MAR of between 520 – 1690 mm, the assumption is that the territory size will be much smaller than 15 ha. This makes the estimated home range and territory sizes in this study acceptable in the absence of other studies besides Norton (1980) to compare it with. Both SCV and CVh presented results which required visual inspection in order to establish the extents which were ecologically and biologically meaningful.

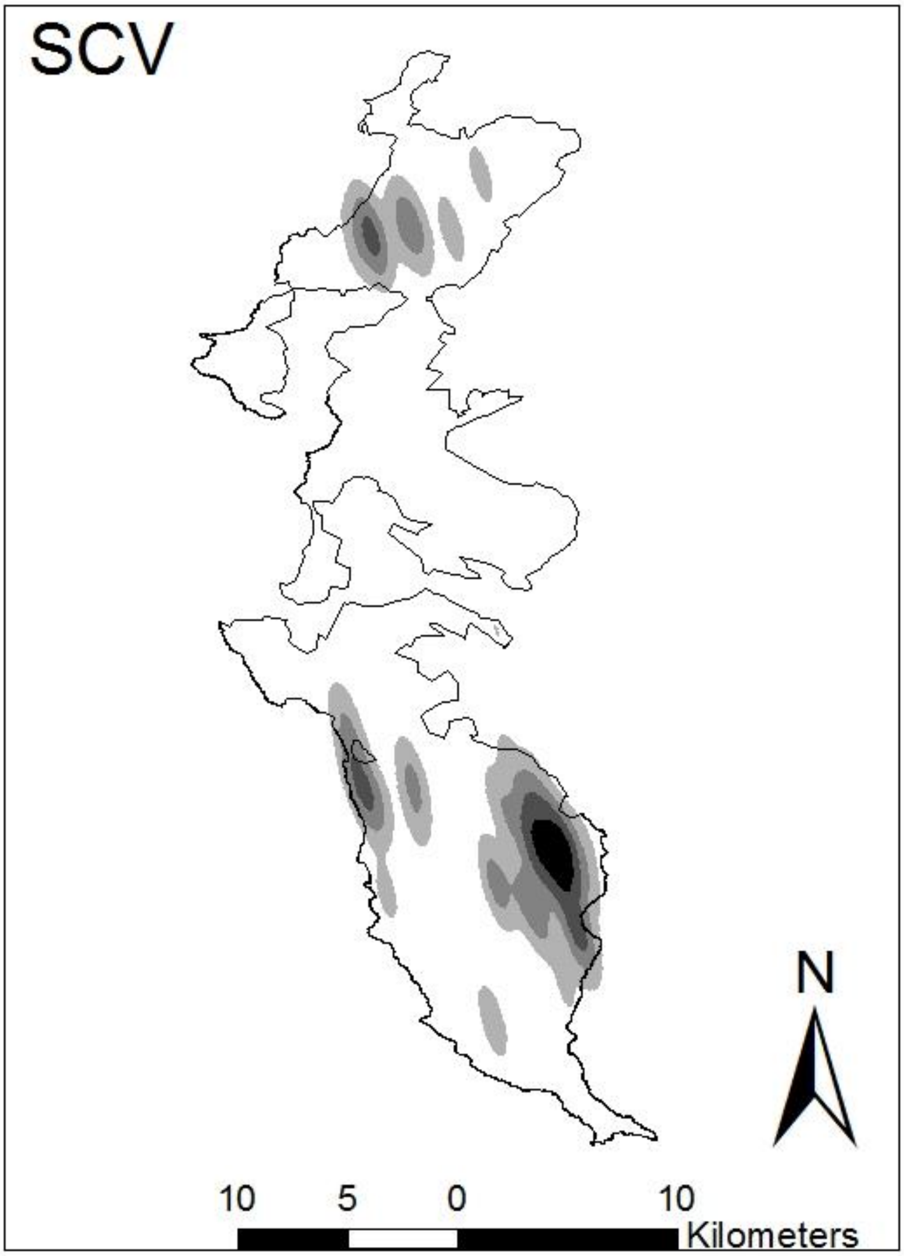


Figure 21: Estimated klipspringer home ranges from a kernel density estimation using the smoothed cross validation (SCV) bandwidth algorithm in the Geospatial Modelling Environment (GME). The contours reflect an increase in density distribution from light grey to black and have the typical shape as obtained from a rotated bivariate fixed kernel.

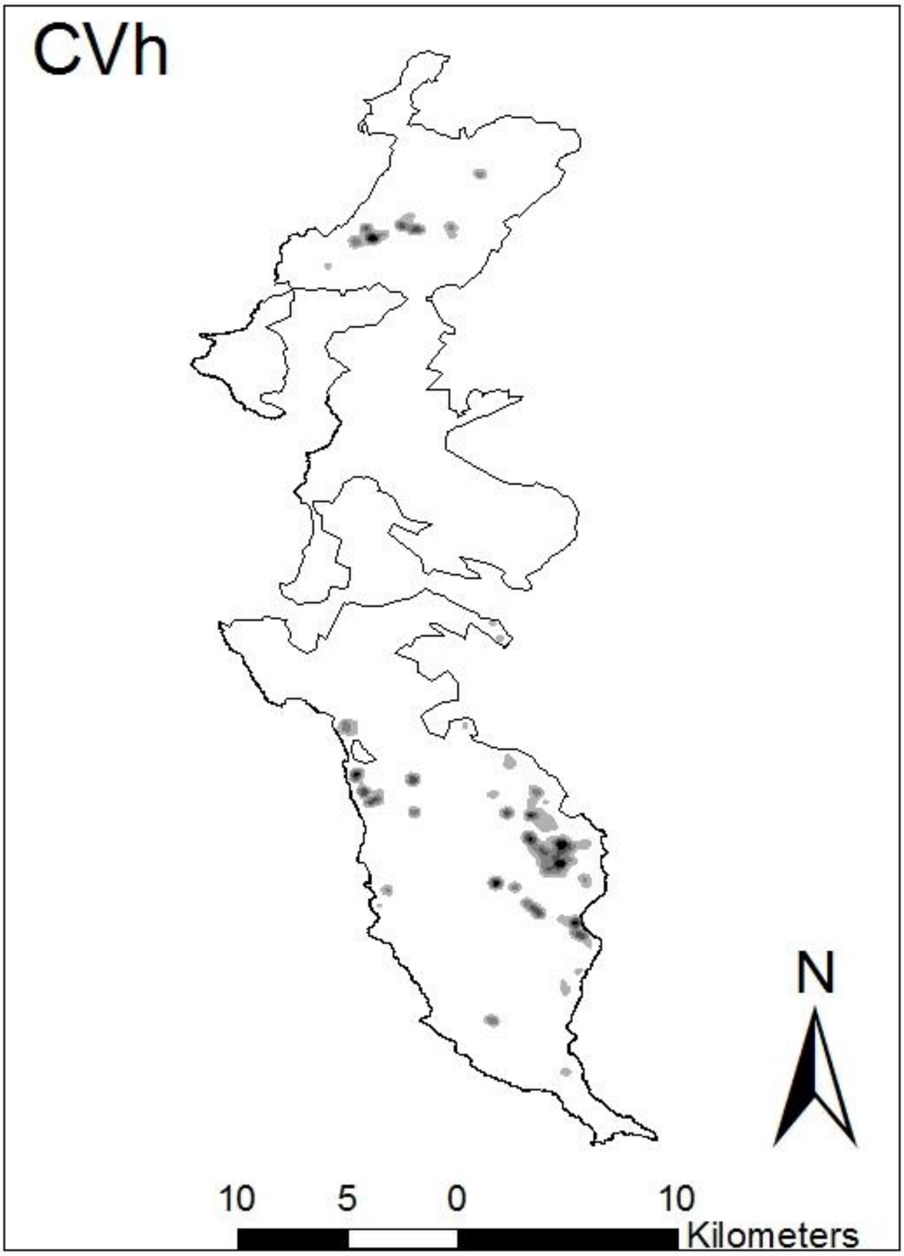


Figure 22: Estimated klipspringer home ranges from a kernel density estimation using the likelihood cross validation (CVh) bandwidth algorithm in the Geospatial Modelling Environment (GME). The contours reflect the shape obtained from a univariate fixed kernel.

After several isopleth runs on the SCV output, the 50% isopleth provided the most ecologically meaningful output results, although this still has to be interpreted with caution (Figure 23a). Isopleths greater than 50% included areas that visually would not be areas that klipspringers would roam into, namely the ocean and residential areas (Figure 23b). The 80% isopleth gave the best representation of the utilisation distribution when the CVh output was used (Figure 24b). Isopleths smaller than the 80% location density excluded some territories where klipspringer presences were confirmed in the 2014 survey, for example the territories on Elsies Peak above Fish Hoek (Figure 24a). It is important to interpret these areas of core activity (50%) and active use (80%) as a guide to klipspringer utilisation rather than an absolute.

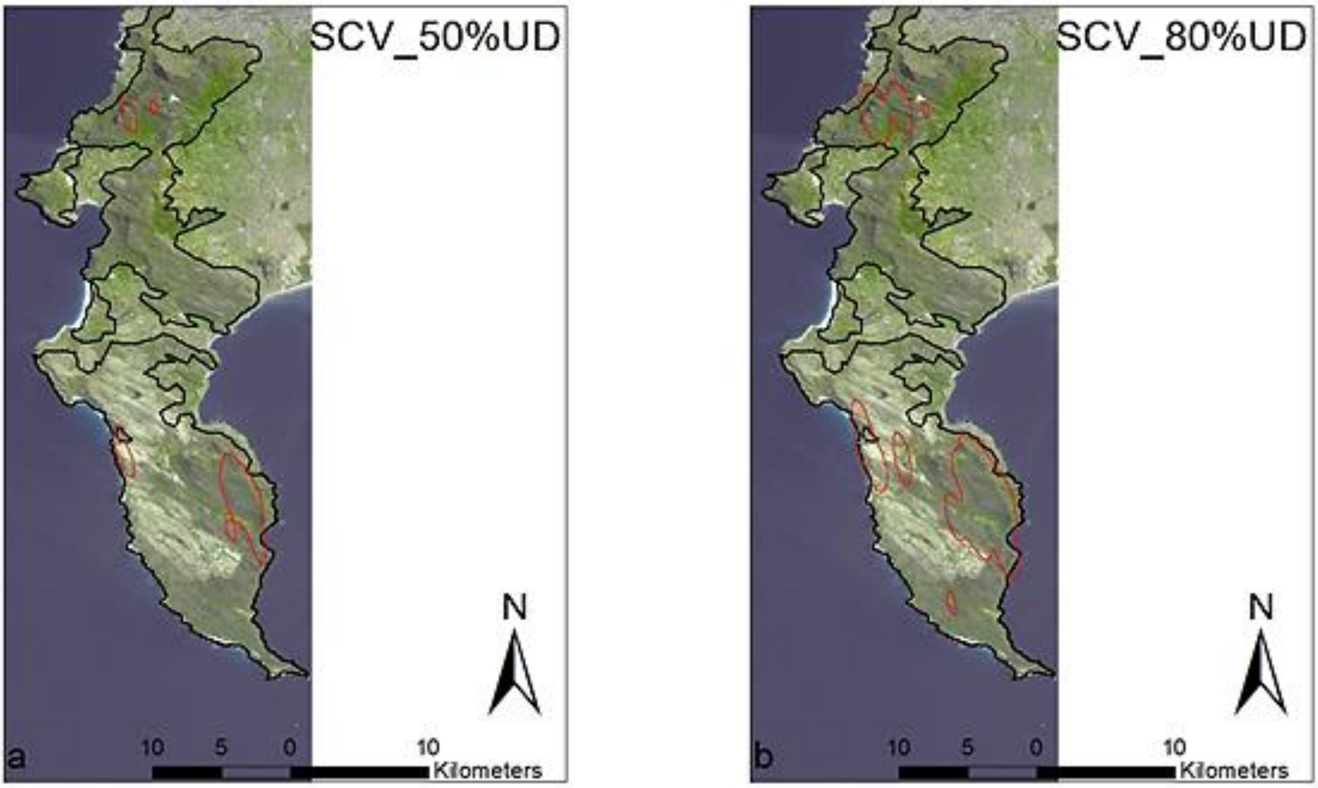


Figure 23: Smoothed cross validation (SCV) bandwidth algorithm showing the 50% (a) and 80% (b) utilisation distributions. Note how the 80% contour extends, at places, beyond the Park boundary into residential areas and the ocean.

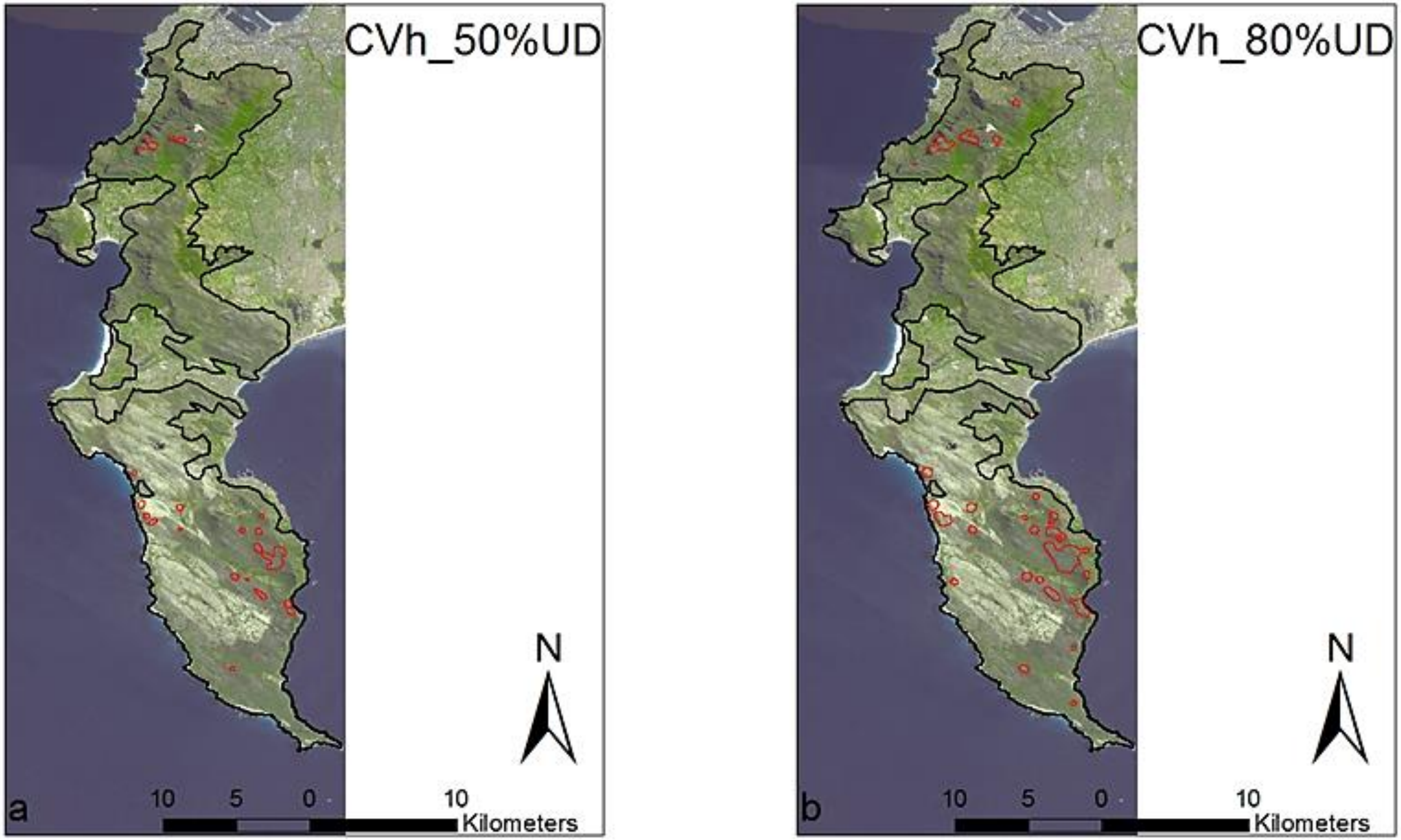


Figure 24: The KDE output using the likelihood cross validation (CVh) bandwidth presented results that can be interpreted as territories rather than home range extents, as shown by the 50% (a) and 80% (b) utilisation distributions. The 80% contours included more territories and thus gave a better representation of where klipspringers occur, unlike the 50% contour that excluded some known territories.

The kernel density tool estimates the probability of an animal occurring within an area based on the collected presence points (Worton, 1989; Seaman *et al.,* 1999). Consequently, the number of points and the period of data collection play an important role in the output of any home range analysis. Thus, home range extents, using statistical software tools, are a direct reflection of the nature of the occurrence data.

The smaller home ranges from the CVh output were interpreted as territories rather than home ranges and therefore useful to look at when interested in territories. For a more comprehensive picture of klipspringer home ranges the 50% utilisation distribution derived from the SCV can be used. Rudeen (2012) and Meck (2013) found that the CVh estimator produced the best ecologically meaningful home range estimates for their respective study species. Horne & Garton (2006) also found some promising output using CVh compared to the more common LSCV.

A combination of abiotic and biotic conditions (food source in the form of dicotyledonous plants) exists at these 50% and 80% density distributions obtained from the home range analysis. Therefore, these areas can be described as the realised niche of the klipspringers (Soberon & Peterson, 2005; Soberon & Nakamura, 2009). It can also be assumed that predators (e.g. caracal (*Caracal caracal*)) would either be absent at some of these habitats or are found in small enough numbers for the klipspringer populations to still persist there.

A drawback of GME is that it only has the option for fixed kernel methods. Worton (1989) highlights that the adaptive kernel with bandwidth algorithm LSCV is a better option to use when accuracy of the UD is important. However, Seaman & Powell (1996) found in their study that the adaptive kernel overestimates home ranges and consequently fixed kernels performed better. Similarly, Pebsworth *et al.* (2012) also found that a fixed kernel method performed well for their specific study whilst taking into consideration the ecology of their study species. Thus, the best KDE output is largely influenced by the specific set of data, the choice of bandwidth algorithm (smoothing parameter) as well as the biology and ecology of the species being studied.

Deciding subjectively on the best smoothing parameter/-s after a number of analyses, is not uncommon in home range studies using KDE (Horne & Garton, 2006; Pebsworth *et al.,* 2012). Statistical methods for calculating an optimum bandwidth is also given in Silverman (1986); Worton (1989); and Sheather (2004), but were not explored in this study because the home range analysis tool decided on, namely GME, comes with its own fixed options of bandwidths. Much emphasis is placed on the importance of choosing the most suitable bandwidth algorithm as this directly influences the probability density estimation (Seaman & Powell, 1996; Seaman *et al.,* 1998; Hemson *et al.,* 2005). Therefore within this current study all bandwidth algorithms in GME were explored. Cell size was found to have little effect on the output, except that a smaller cell size increases the running time.

Telemetry data are commonly used in home range analysis (Hemson *et al.,* 2005; Getz *et al.,* 2007; Pebsworth *et al.,* 2012). This study used a combination of telemetry data and manually collected presence points. Of concern is that the collar data (three readings per day) from the five collared individuals were not collected for the same period of time: three weeks, one month, one month and two weeks, two months, and five and half months respectively. This was either as a result of death of an individual or battery failure. Consequently, some areas, e.g. the Swartberg Mountain chain above Simonstown to Smitswinkel Bay, had a higher concentration of points resulting from points being collected there for a longer period of time, which the KDE output reflected as an area of high utilisation even after the spatial filtering of one point per grid. Spatial filtering at a larger scale could probably have provided a better reflection of home ranges as it could have resulted in less clumping of points on the Swartberg Mountains. A second drawback is the lack of repeat sampling for the manually collected data. A "better" set of data for future use would be telemetry data collected in different seasons and for the same period of time in each season in order to give a better reflection of the home ranges of the Park's klipspringer populations.

Irrespective of the accuracy of the output, home range analysis is still a good way to represent occurrence data in geographical space. Home range analysis in return aids in understanding and interpreting the collected points and visualising the distribution of the focal species within their study areas. The use of home range analysis on the specific set of data also highlighted what should be done differently in a future study in order to obtain the best representation of the klipspringer home ranges in TMNP.

CHAPTER FOUR: RESULTS AND DISCUSSION: SPECIES DISTRIBUTION MODELLING

The six models at different regularisation values (two below the default value of 1, and three above the default value) produced maps that broadly seem to resemble one another (Figures 25 and 26). However, closer visual inspection showed that as the regularisation value is increased (i.e. relaxing the constraints); the proportion of potentially suitable habitat also increased. This illustrates that lower RM values than the default can potentially underestimate the prediction and higher values can lead to an over prediction, as found by Radosavljevic & Anderson (2014), suggesting that the default RM is probably appropriate for this study. The two probability maps at lower RM than the default also show a strange model fit in the lower part of the southern section (Figure 25a and b), possibly due to the overfitting inclination that accompanies RM values lower than the default (Phillips *et al.*, 2006). Visual inspection of aerial images in that portion of the southern section could not reveal any particular topographical feature/-s that could be directly linked to the specific shape of that clear cut bands. From the default RM onwards (Figure 25c and Figure 26a – c) the possible tight overfit that created the strange bands in the lower part of the southern section started disappearing. This further supports that the lower RM values are not fitting for this study and the default RM onwards are better options to be considered for the best model after taking into consideration other factors. Figure 27 shows the response of klipspringers to altitude as an example with different RM values. When the RM values are set below the default, the response curves are complex and show detailed peaks and troughs that have no ecological meaning. As the RM values are set to values much larger than the default, the response curves over-smoothed the responses. The default RM value of 1 would therefore suffice in this specific study. All response curves for all other variables showed similar response curves. The AUC values for all six models indicated good model performance (Table 7; Appendix J: Figures J1, J2 and J3), thus the choice of the “best” model could not be based on the AUC values, and is primarily based on the ability to provide ecological interpretation, on the shape of the response curve and the probability distribution maps. Using MaxEnt’s logistic output, the probability of presence of a species or habitat suitability is from 0.5 and higher (Phillips & Dudik, 2008). These values are represented by the green, yellow, orange, and red colours (Figures 25 and 26). The olive green to dark blue colour represents lower probabilities to complete absence (<0.5 – 0) (Figures 25 and 26). A common trend was visible amongst these six maps on which areas had a probability of presence score of 0.6 and above, i.e. those indicated from yellow to red.

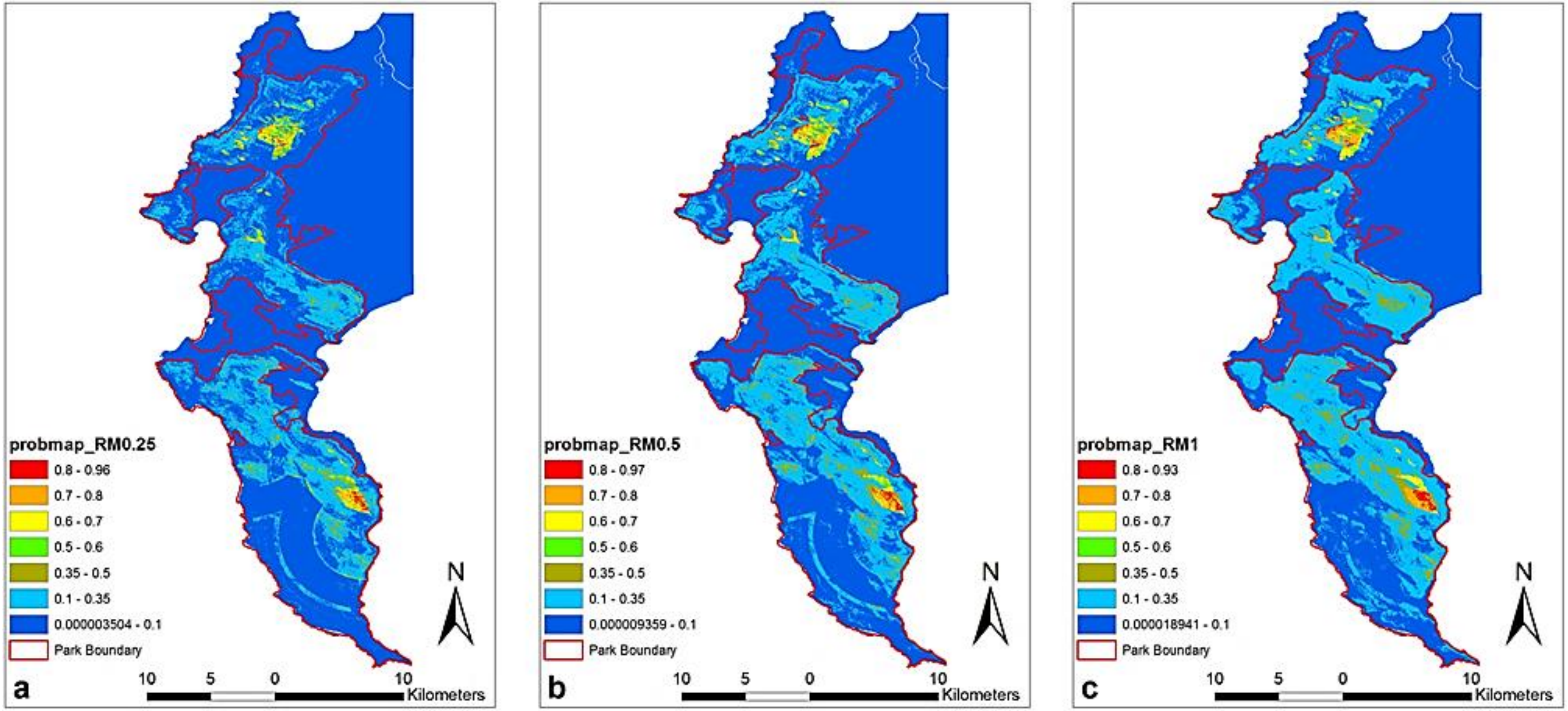


Figure 25: The probability distribution of suitable habitat across the Park for the models at regularisation multiplier (RM) values of 0.25 (a), 0.5 (b), and default (1, c).

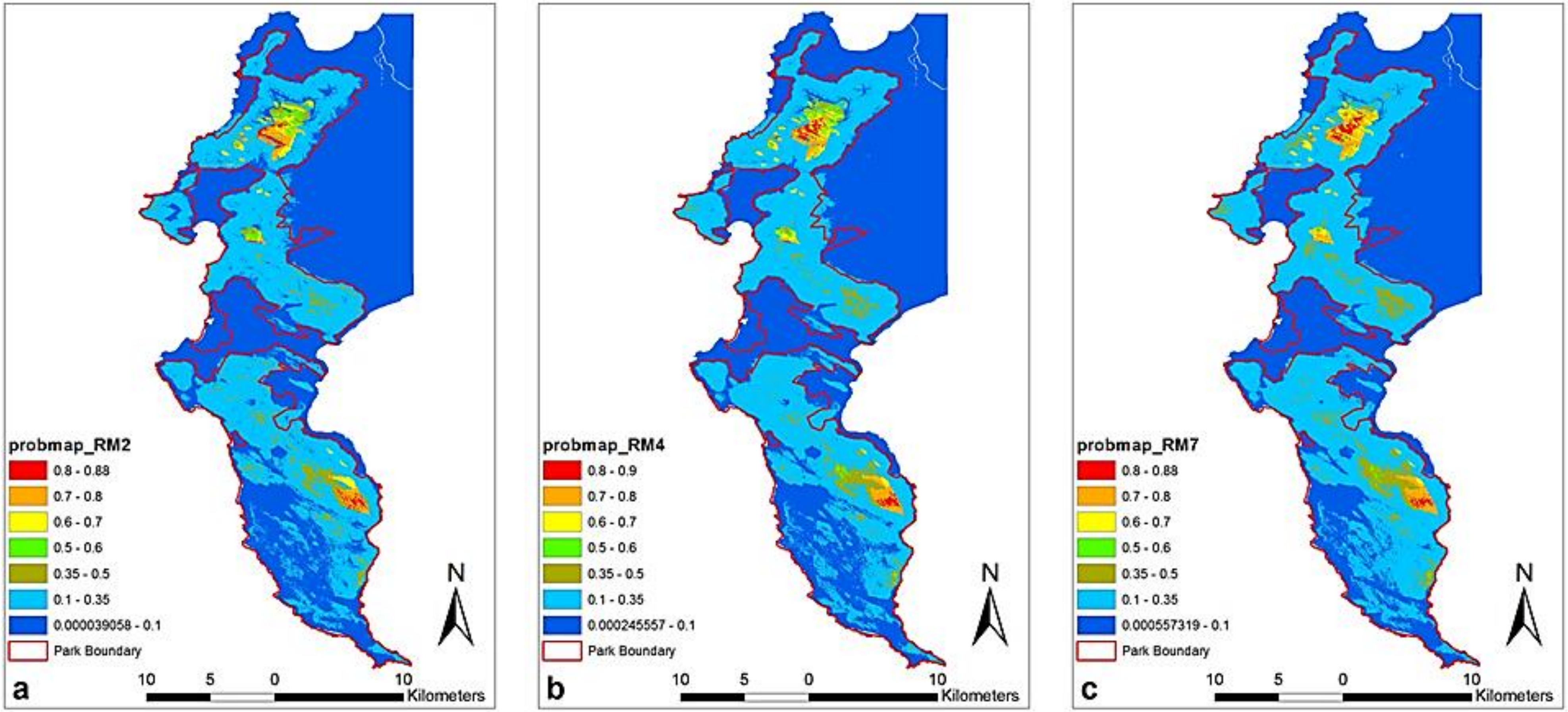
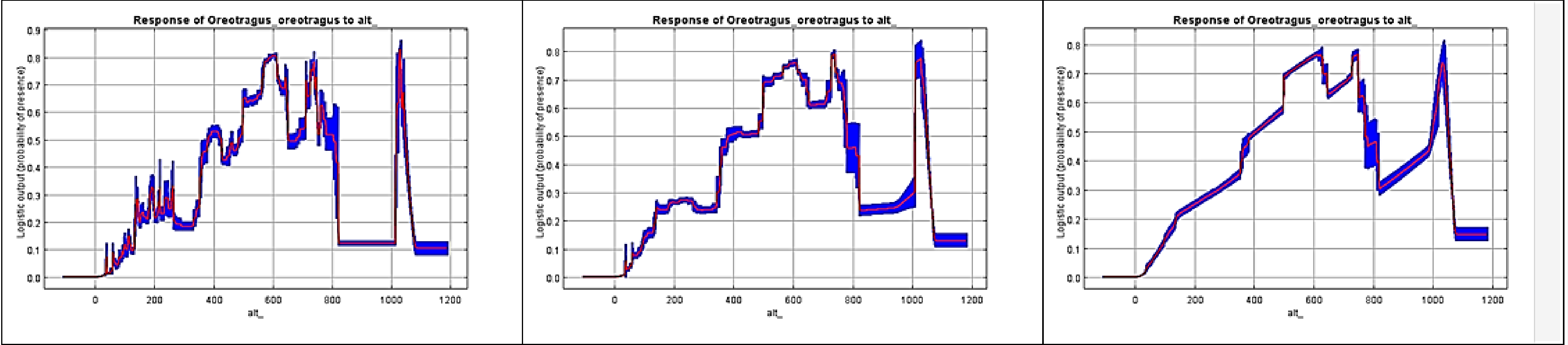


Figure 26: The probability distribution of suitable habitat across the Park for the models at regularisation multiplier (RM) values of 2 (a), 4 (b) and 7 (c).

a) Klipspringer Response to Altitude RM 0.25 **b) Klipspringer Response to Altitude RM 0.5** **c) Klipspringer Response to Altitude RM 1**



d) Klipspringer Response to Altitude RM 2 **e) Klipspringer Response to Altitude RM 4** **f) Klipspringer Response to Altitude RM 7**

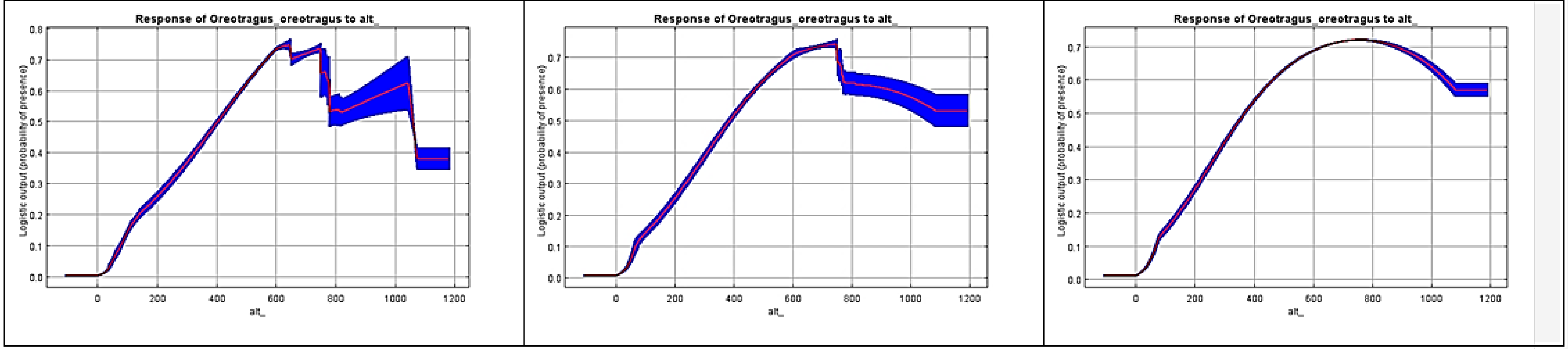


Figure 27: Response curves for the six models showing the response of klipspringers to altitude at different regularisation multiplier (RM) values: RM 0.25 (a), RM 0.5 (b), RM 1 (c), RM 2 (d), RM 4 e) and RM 7 (f). Red indicates the mean response over 10 replicate runs and blue is the mean +/- one standard deviation.

Table 7: AUC values for the six models at different regularisation multiplier values.

Model	**Regularisation multiplier**	**Mean AUC on test data over 10 replicate runs**	**Mean +/- one standard deviation on test data over 10 replicate runs**
1: All five variables	0.25	0.906	0.007
2: All five variables	0.5	0.911	0.006
3: All five variables	1	0.903	0.008
4: All five variables	2	0.904	0.009
5: All five variables	4	0.893	0.011
6: All five variables	7	0.883	0.008

The omission error (suitable areas not predicted by the model) was close to the predicted omission error (Figure 28). This indicates that the model performed well as there will be very little omission error (Phillips, 2006). Deviations higher and lower than the predicted omission would indicate the model could not effectively account for omission error. Most of the other models (using other RM values) also showed low omission error, although the RM values of 0.25 and 7 showed considerable deviation from the predicted omission line, and the omission errors for RM 0.5, 2 and 4 did not fit the omission line as tightly as the default RM value (Appendix J: Figures J4, J5 and J6). The ROC AUC which plots Sensitivity (presences correctly predicted) against 1 – Specificity or commission error (absences incorrectly predicted) showed exceptional model performance with its score of 0.903 (Figure 29). This confirms that the model performed well at predicting presences thus suitable habitat.

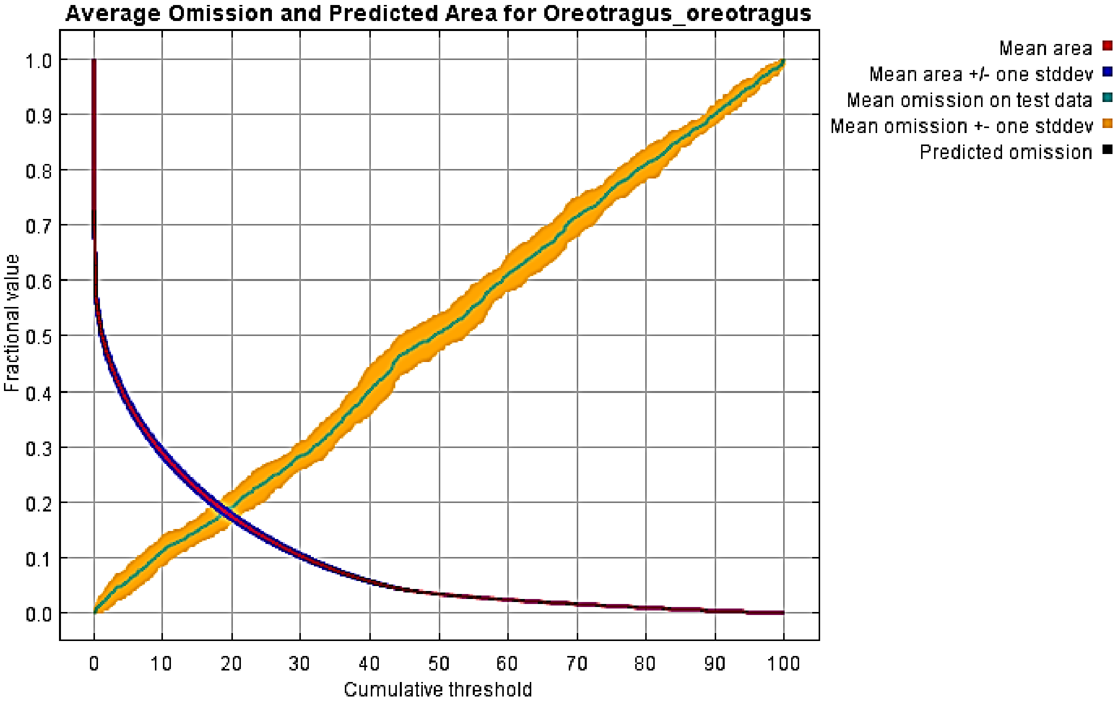


Figure 28: The omission error (suitable areas not predicted correctly), thus an indication how well the model performed. The mean omission (orange and green line) is close to the predicted omission (straight black line which is not visible as the mean omission resembles it closely and therefore covers it completely) which indicates good model performance.

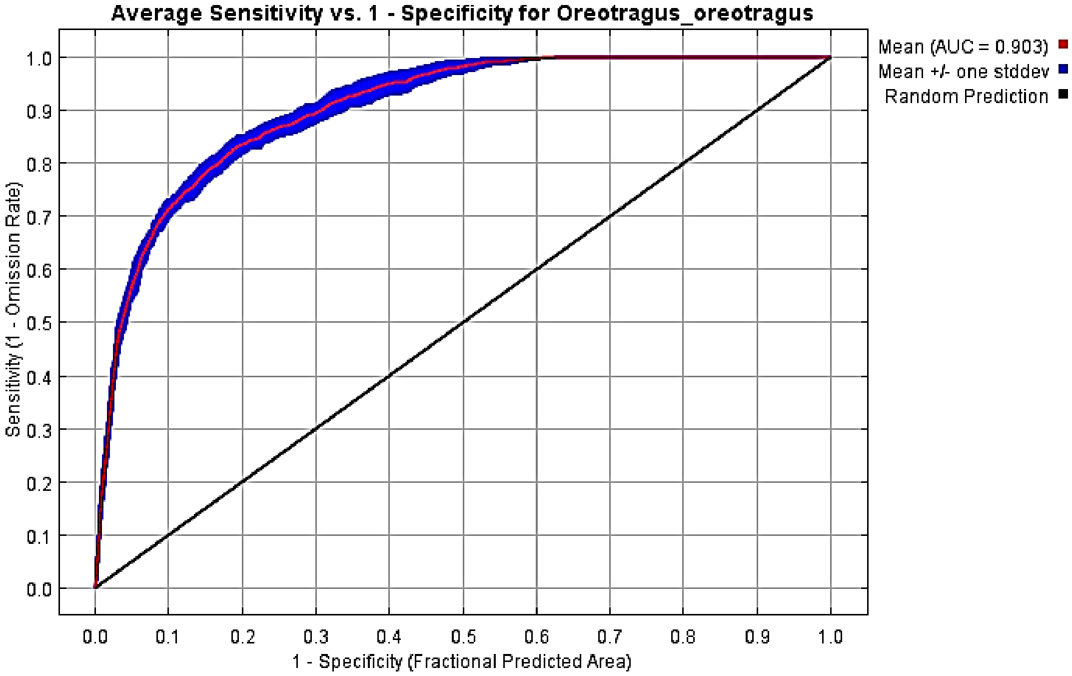


Figure 29: Area under the receiver operating characteristic (ROC) curve (AUC) testing whether the model was able to rank presence sites higher than random background sites. The AUC score of 0.903 indicates exceptionally good model performance.

MaxEnt creates response curves to show how each variable responds in the model. The mean response from the 10 replicate runs is indicated in red and one standard deviation around the mean is indicated in blue (Figures 30 – 34). The range of values displayed on the x– axis reflects the information that is present in each layer. The response curve in essence reflects how the species responds to the specific range of environmental conditions that are present in the study area, and this response relates to where the occurrence points were collected (Baldwin, 2009). The response of klipspringers to altitude shows that the probability of klipspringer occurrence is from an elevation of about 400 m.a.s.l. (Figure 30). The peak of this response curve is at a probability of presence score of about 0.76 and an elevation of about 620 m.a.s.l. Thereafter two consecutive drops occur at 620-650 m.a.s.l. and 740-810 m.a.s.l. of which this second drop goes down to a probably of presence of 0.3. Inspecting aerial images of the study area and the altitude and slope layer revealed that these drops are resultant of steep cliffs, especially in the northern section of the Park. After the second drop another suitable elevation range is reached just after 1000 m.a.s.l. of which the final drop is a result of the elevation in the study area not exceeding 1084 m.a.s.l. The curve shows that although other authors found that klipspringers do descend into river valleys (Tilson, 1980) and low slopes under certain conditions (Skinner & Chimimba, 2005) in search of browse, this behaviour is seldom recorded in the current study area. The response to slope shows that klipspringers will occupy slopes from 15 to about 48 degrees, thereafter the slope becomes too steep (Figure 31). The highest peak for this curve was reached at a probability score of only 0.59 at a slope of between 15-20 degrees. Note again the low probability of occurrence for klipspringer in this study area in flatter areas, which suggests they prefer steep rocky slopes and rarely use flat areas as Dunbar & Dunbar (1974) also found.

For aspect the probability of presence was slightly higher than that of slope at about 0.63 with a general preference for south west (202.5 degrees to 247.5 degrees) to west facing (247.5 degrees to 292.5 degrees) slopes being shown (Figure 32). This can partly be attributed to the vegetation communities that the model has output as most preferred having a mean aspect of south west and south-south west (ericaceous fynbos and upland restioid fynbos respectively) (Figure 34 and Table 2). A mean aspect of south, and south-west facing slope is also prevalent for most of the other vegetation communities (Table 2). Thus with more areas in the Park facing south, south-west and west, it is not surprising to find klipspringer occupying these slopes as well. North, north-east, east, and south-east facing slopes in the study area contain most steep cliffs, thus explaining the low probability of presence score between 0.35 and 0.45. Steep cliff faces are also present for south and west facing slopes although to a lesser extent. Many north-west facing slopes also contain steep cliffs and consequently the probability of presence drops below 0.5 (i.e. after 337.5 degrees).

Although klipspringers do not occur right on the urban area boundary, it appears that urban areas do not affect where klipspringers would establish territories, as a probability of presence is predicted as close as 2.5 km to 1.5 km to the nearest urban areas (Figure 33). It is only in the southern section (Cape of Good Hope) where there is Park land that is more than 3 km away from an urban area (Appendix H), thus only klipspringer individuals recorded in this section would contribute to a probability of presence above 0.5 scored after 3 km. As a result of fewer and fewer individuals being recorded after 3 km, the response curve continuously dropped, although a small peak occurred after 5.5 km to 6 km (although still not significant). This is possibly attributed to the 10 occurrence points within this distance band of 5.5 km to 6 km (Appendix H).

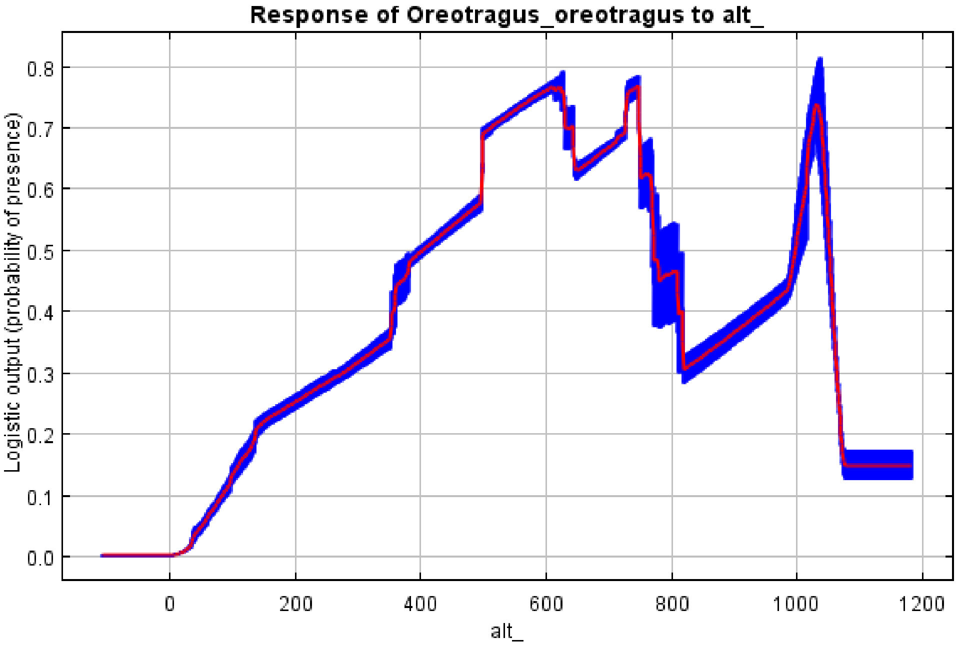


Figure 30: The response of klipspringers to altitude of which the probability of klipspringer occurrence is from an elevation of about 400 m.a.s.l. and above. This indicates the preference of klipspringers for high altitudinal areas.

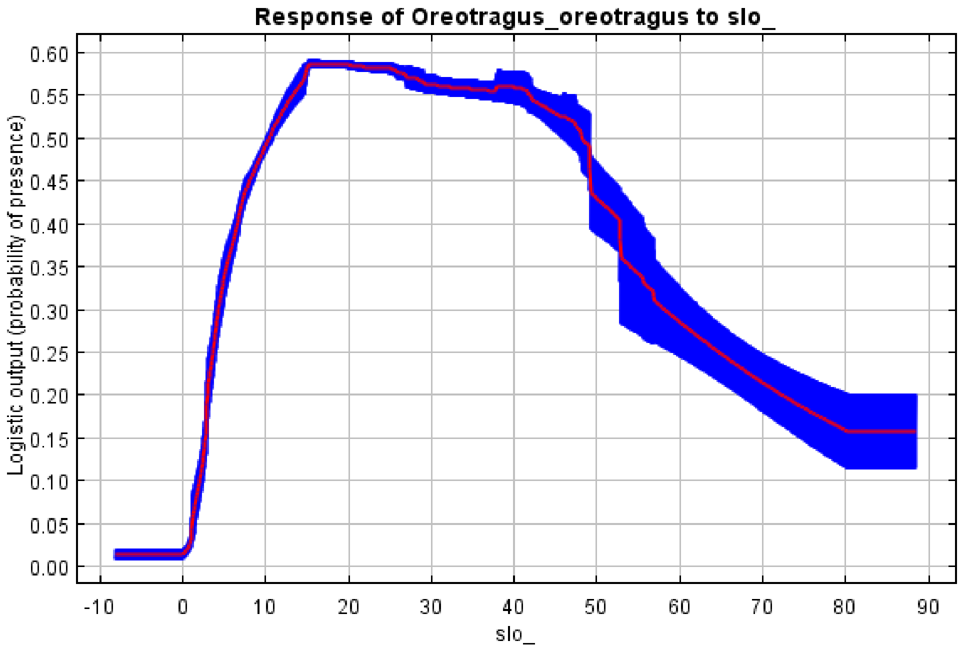


Figure 31: The response curve of klipspringers to slope shows that klipspringers occupy slopes from 15 to about 48 degrees, thereafter it becomes too steep.

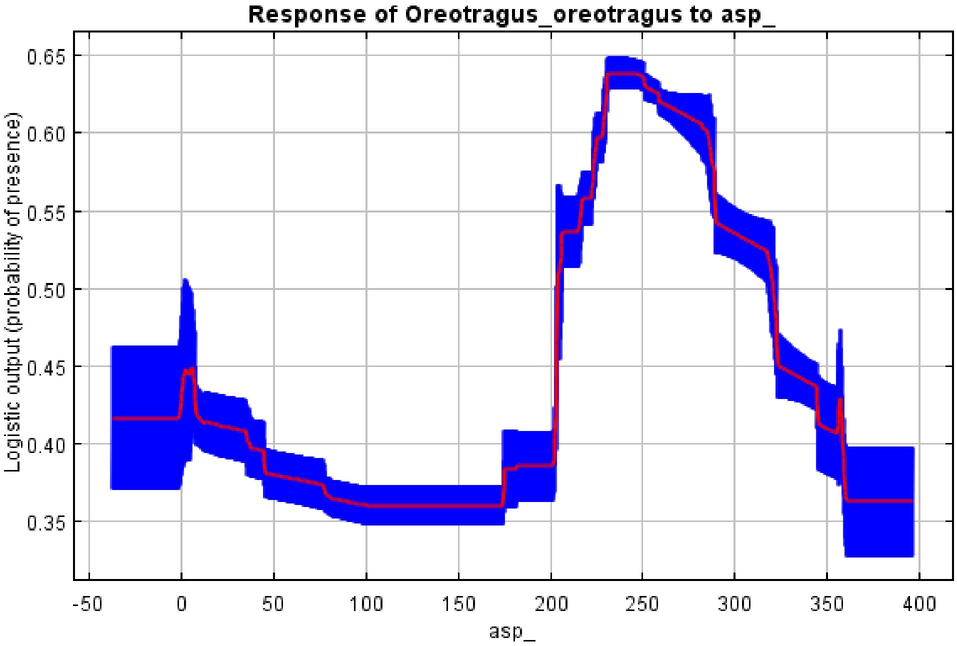


Figure 32: The response of klipspringers to aspect. South-west to west facing slopes shows a higher probability of occurrence than the other slopes.

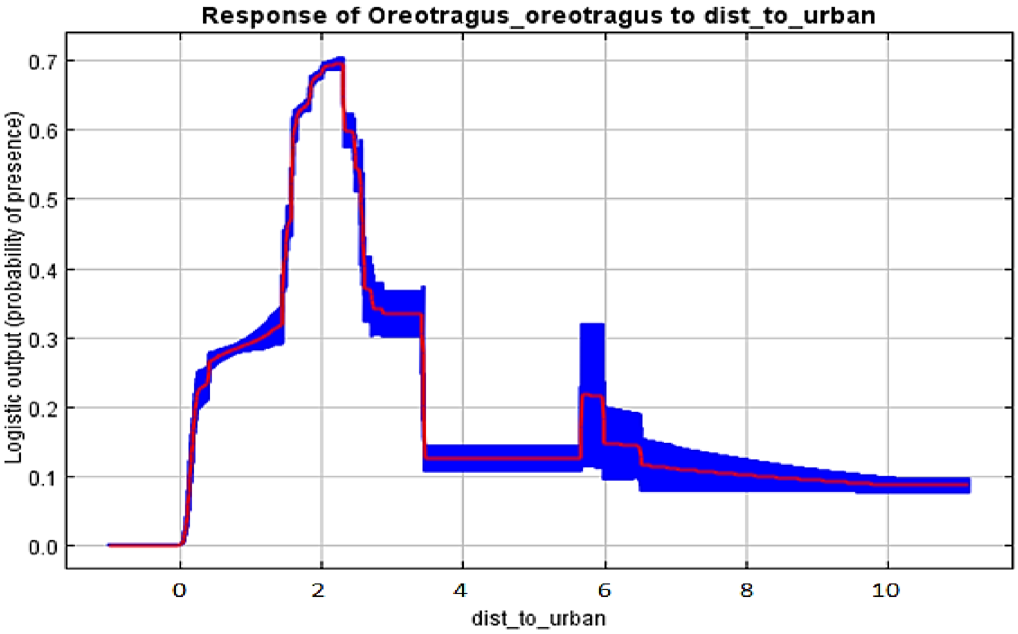


Figure 33: The response of klipspringers to distance to urban edge shows that the probability of presence is predicted as close as 2.5 km to 1.5 km from the nearest urban areas.

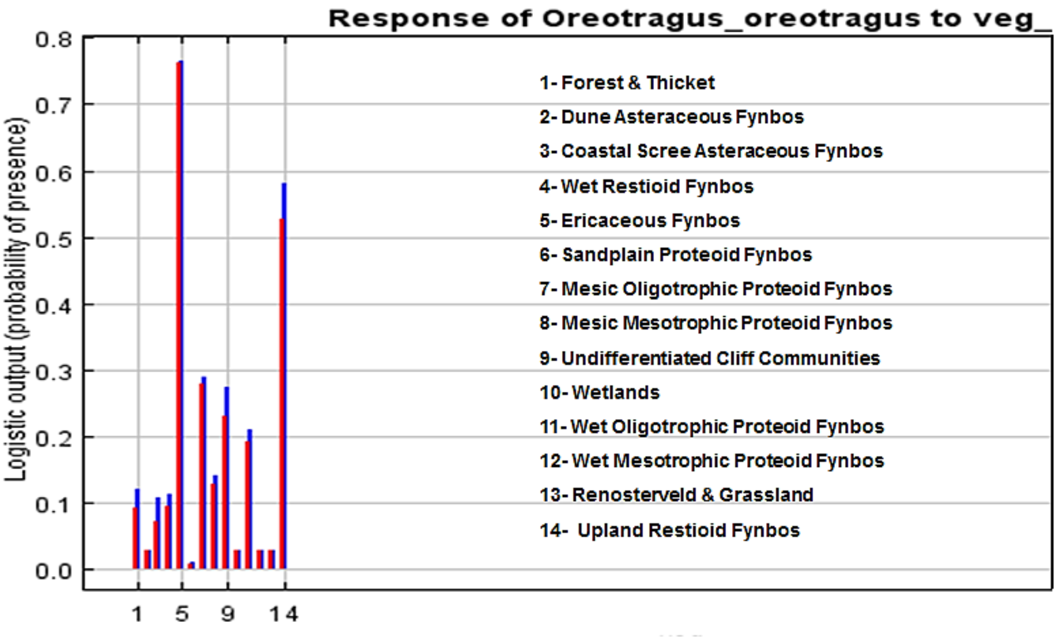


Figure 34: The response of klipspringers to vegetation communities shows that ericaceous fynbos is the most preferred vegetation community followed by upland restioid fynbos.

The response of klipspringers to vegetation only showed two vegetation communities to have a high contribution to the probability distribution of suitable habitat (Figure 34), these being ericaceous fynbos with a probability of presence estimated at 0.77 and upland restioid fynbos at 0.53. All the other communities had probability values lower than 0.3 which indicates a low contribution of these vegetation communities to suitable habitats.

MaxEnt's output also includes an analysis of variable importance in two formats; as a table and in the form of jackknife graphs. Table 8 shows the percent contribution and the permutation importance of each variable. The percent contribution indicates how each variable contributed to the overall model gain. Permutation importance on the other hand shows how each variable contributed to the final model when the values of each variable were randomly permuted. If the percentage value in the permutation importance column is lower than the value displayed in the percent contribution column, it indicates that it is indeed an important variable (Kalle *et al.,* 2013). This reduction in percentage value can be seen for vegetation and altitude which reflects their importance as a variable for klipspringers. Aspect decreased with 0.1 percent and thus this reduction is considered insignificant. Distance to urban edge and slope increased in percentage value and thus they do not play a large role in klipspringer habitat selection.

Table 8: Analysis of variable contribution. The percent contribution indicates how each variable contributed to the overall model gain whilst permutation importance shows how each variable contributed to the final model when the values of each variable were randomly permuted. A decrease in percentage value from the percent contribution indicates that it is indeed an important variable.

Variable	Percent contribution	Permutation importance (%)
Vegetation	58	37.1
Altitude	25.3	15.5
Distance to urban edge	12.6	37.7
Slope	2.9	8.6
Aspect	1.2	1.1

The jackknife results confirmed the order of variable importance as listed in the table above (Figure 35). Jackknife output shows which variable contributes the most to the prediction when used in isolation as well as how the model gain is affected when a variable is excluded. The dark blue line shows the amount of model gain when only one variable was used. The turquoise line gives an indication of model gain when a specific variable is excluded. The first jackknife figure (Figure 35a) is from the training data and the second one (Figure 35b) from the test data after 10 replicate runs. In both these jackknife figures vegetation had the highest gain when used in isolation. This indicates that the vegetation layer had more useful information compared to the other layers (Phillips, 2006). Thus vegetation is the most important variable followed closely by altitude. When vegetation was excluded it decreased the model gain the most, once again indicating the importance of the information present in this layer. Vegetation was also found to be the determining factor for habitat selection by klipspringers in the Kruger National Park in that it provides both shelter and food (Kruger, 2001). Both altitude and distance to urban areas had noteworthy contributions to the model prediction as well. The jackknife analysis on the test data showing similar results as the jackknife analysis on the training data ensured confidence in the order of variable importance. This same order was seen in the jackknife analysis on training and test data for the other models (Appendix J: Figure J7) as variables affecting klipspringer distribution, especially the true environmental variables. Slope and aspect did not contain as much important information as the other three variables and therefore contributed to less model gain when used in isolation and when excluded the model gain was not much reduced.

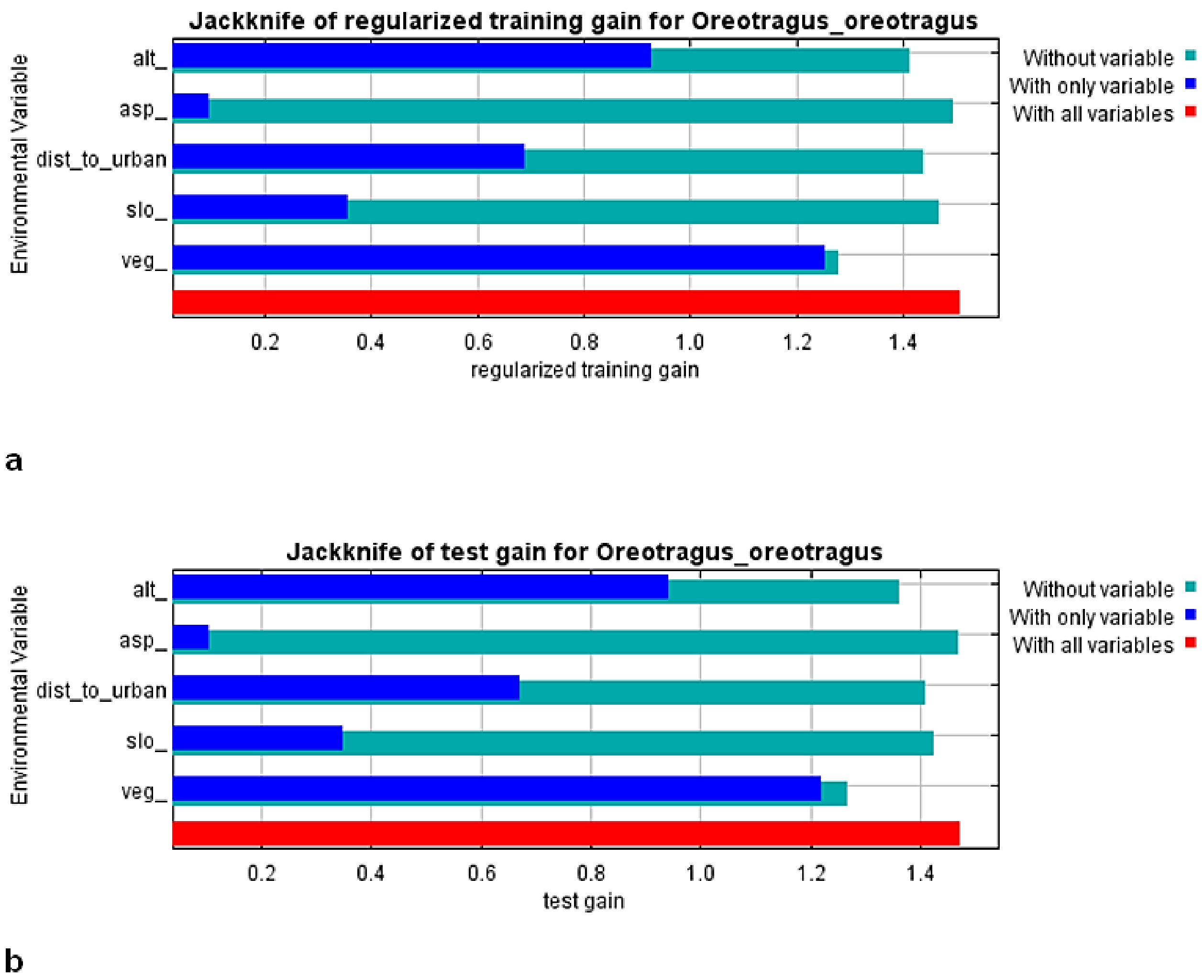


Figure 35: The jackknife output of variable contribution on regularised training gain (a, using the training/calibration data) and test gain (b, using the test data) shows that vegetation contributes the most to suitable habitat for klipspringers followed by altitude. Dark blue indicates model gain with only the variable, turquoise indicates model gain when the variable was excluded, and red indicates the total model gain.

Testing different regularisation multiplier values was useful to confirm that the default was applicable to be used in this study in the same way that it was in Phillips & Dudik (2008). Visual inspection of probability maps confirmed that a RM higher than the default can potentially lead to a model that under fits the occurrence points and therefore overestimate suitability and lower than the default can potentially overfit occurrence points. Thus it was decided that the default RM is most appropriate to use since the “best” model were in the vicinity of the default RM.

RM 7 had the lowest AUC score, 0.894. This, however, is still a good model evaluation score (Swets, 1988; Hosmer & Lemeshow, 2000) thus a definite conclusion cannot be reached that RM 7 was not performing well. Although visually the areas with a probability score of 0.1 – 0.35 (light blue) and 0.35 – 0.5 (olive green) became more prominent as the RM value was increased it is still below

0.5 and thus those areas are not suitable when interpreting it from the logistic output maps (Figures 25 and 26). This is based on a score for a typical presence site being 0.5 (Phillips & Dudik, 2008). In the study by Radosavljevic & Anderson (2014), a RM of 6 evidently over predicted some areas for the Caribbean spiny pocket mouse (*Heteromys anomalus)* from north-western South America and the RM of 0.25 and 1 were fitting the data too closely. A RM of 2 proved to show the probability distribution best in their study. Shcheglovitova & Anderson (2013) assessed optimal regularisation multiplier values in terms of omission rates and found that a higher RM value in both their data sets were associated with lower omission rates. Their highest and consequently optimal RM value was 2. This shows that it is useful to do such a test on the best RM for the specific study as the default would not always be the best choice. An understanding of the study species' ecology and biology as well as good knowledge of the study area is also helpful in identifying when a model has completely over or under predicted suitability. In addition, sample size is another important aspect to keep in mind when determining and deciding on the best regularisation multiplier value for a specific study (Anderson & Gonzalez, 2011) together with the features used i.e. model complexity (Anderson & Gonzalez, 2011; Shcheglovitova & Anderson, 2013; Warren & Seifert, 2013).

After a model run, MaxEnt provides possible threshold values that can be used if there is a need to create binary maps for the study. In an attempt to better distinguish between suitable and unsuitable habitats, such a binary map was produced to show exactly which areas are suitable and which areas are not suitable (Figure 36). The current study used the minimum training presence logistic threshold and the 10 percentile training presence logistic threshold to produce Figure 36a and b. Young *et al.* (2011) stated that if there is certainty in the quality of the presence data, then the minimum presence threshold can be used. Since the 2014 survey confirmed the klipspringers are still largely found in the same areas as from the data collected in 2004 to 2007, this provided more confidence in the quality of the presence data, suggesting the use of the minimum presence threshold. Using the minimum presence threshold of 0.0121 (Figure 36a), however, resulted in a map that clearly overestimated suitable habitat, compared to the 10 percentile threshold of 0.1614 (Figure 36b) that gave a better representation of the suitable versus unsuitable habitat. This minimum presence threshold is based on the lowest predicted value for any of the pixels with an occurrence record in it (Pearson, 2007; Young *et al.*, 2011), thus an overestimation of suitable habitat is a possibility.

Similarly, Pearson *et al.* (2007) found that a smaller threshold resulted in an increase in the proportion of area predicted as present in their study area. They, however, interpreted the higher proportion of predicted suitable habitat as identifying additional regions of potentially suitable habitat for their species. Thus interpretation can differ depending on whether species distribution modelling is performed at a regional scale or a local scale. Hence, for this study the 10 percentile threshold was the better option to present suitable and unsuitable habitat for klipspringers in TMNP as it captured many of the areas where occurrence points were recorded. The 10 percentile threshold binary map also captured those high altitudinal areas as potentially suitable habitat following from the jackknife output that indicated and confirmed altitude as an important variable. Escalante *et al.* (2010) tested four thresholds for their study and also found that the 10 percentile threshold gives the best binary output. This illustrates that with species distribution modelling there is no clear and precise way to represent suitable habitat of a species. The estimate of the probability of presence should therefore only be used as a guide to where a species of interest can potentially be found. Consequently, to best illustrate where in TMNP all the potentially suitable habitats for the klipspringer are, it is suggested that the default RM probability map and the 10 percentile training presence logistic threshold map should be used in conjunction with one another (Figures 36b, 37 and 38).

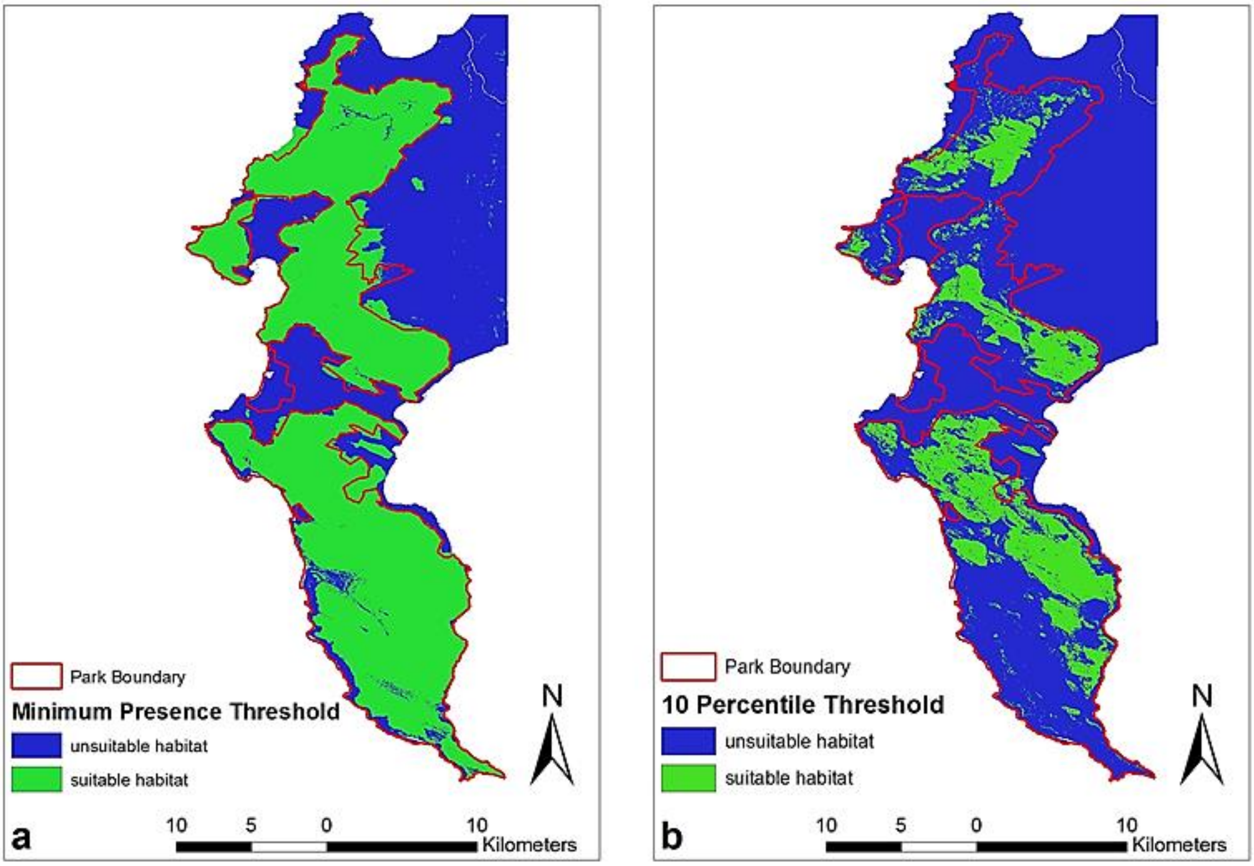


Figure 36: Binary maps showing suitable versus unsuitable habitat at the minimum training presence logistic threshold of 0.0121 (a) and the 10 percentile training presence logistic threshold of 0.1614 (b). The minimum presence threshold clearly overestimates suitable habitat and therefore not a good indication of suitable klipspringer habitat in TMNP.

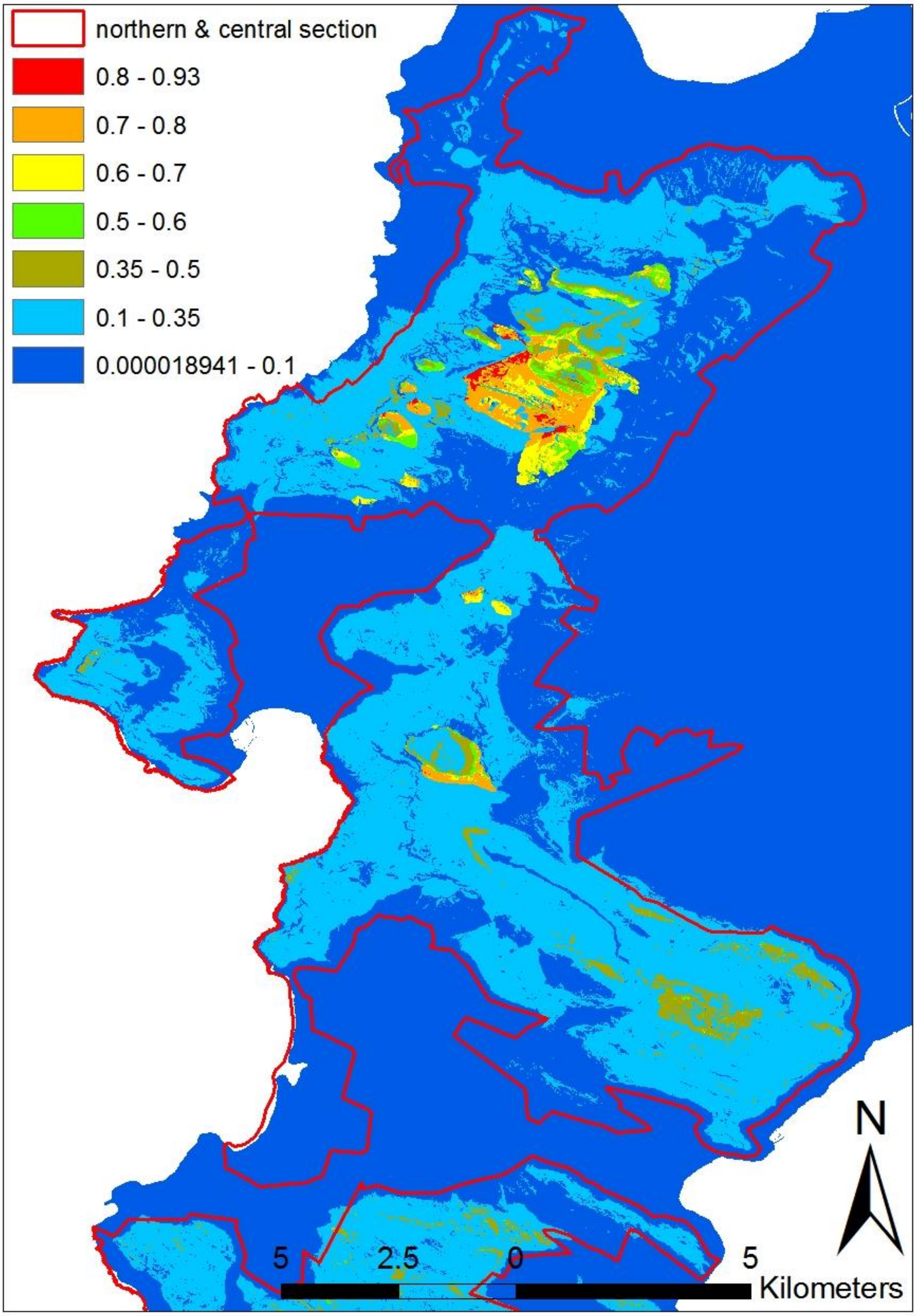


Figure 37: The probability distribution map at the default regularisation multiplier of 1 showing the probability of presence and thus suitable habitat for the klipspringer in the northern and central section of the Park indicated with a score of 0.5 and above.

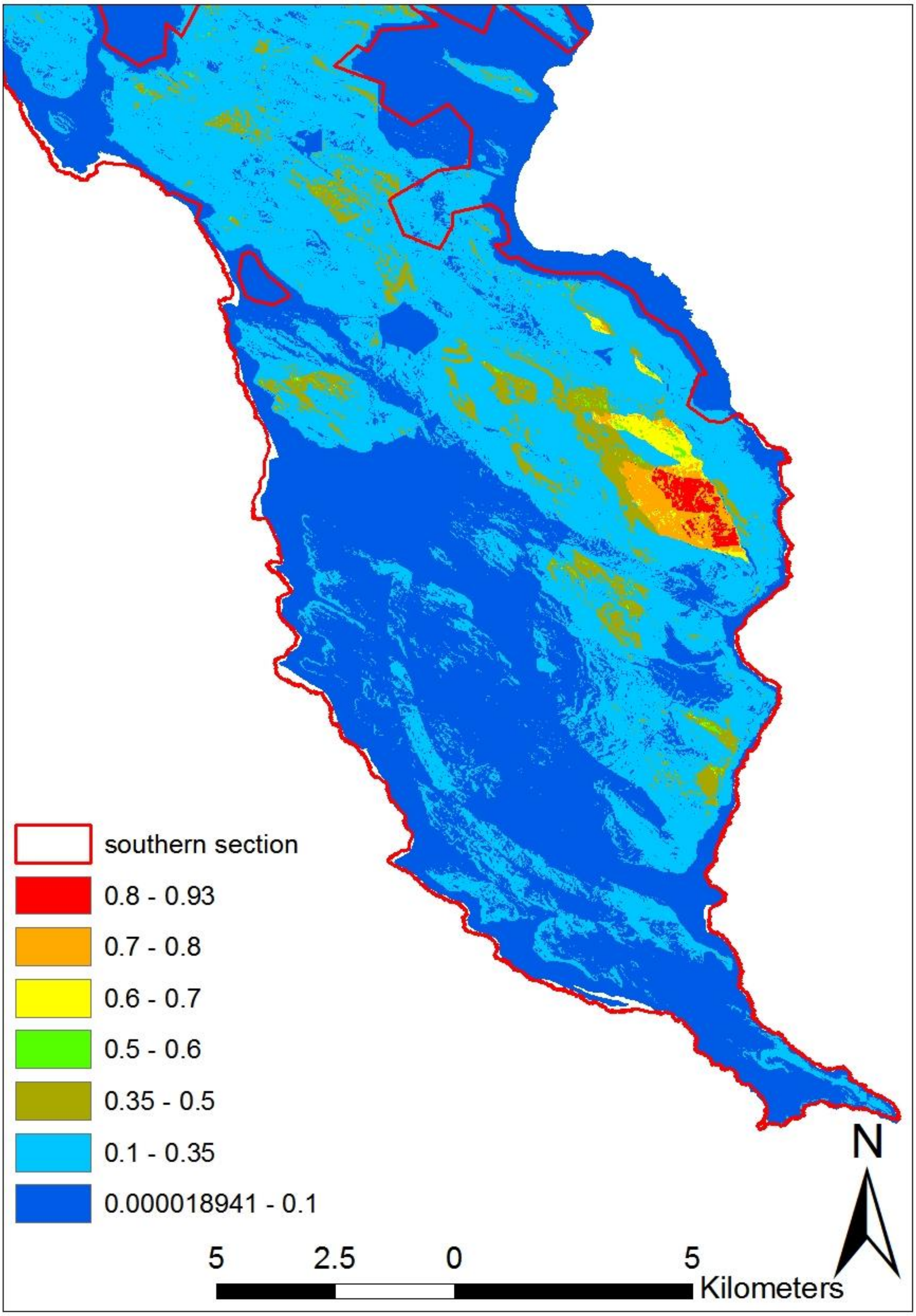


Figure 38: The probability distribution map at the default regularisation multiplier of 1 showing the probability of presence and thus suitable habitat for the klipspringer in the southern section of the Park indicated with a score of 0.5.and above.

The AUC value of 0.903 (RM1) indicates that the model performed well at predicting presences which translates to suitable habitat. This result indicates that the model could distinguish between suitable (presence) and unsuitable (absence) sites and consequently was able to assign a higher ranking to presence cells than absence ones (Phillips *et al.*, 2006). ROC AUC results from the other five models also indicated good model performance, as can be seen from their high AUC values. Lobo *et al.* (2008) highlighted that AUC is not always a good measure of model performance, as models that overestimate or underestimate can also give a high AUC score and consequently indicate good model performance when it is in actual fact not such a good model. Yackulic *et al.* (2013) also questioned the heavy reliance on the AUC for presence-only modelling as an indicator of model performance. Their concern results from ROC AUC initially having been designed to classify presences against absences, whereas MaxEnt uses presence and background data, of which these background sites are selected randomly and thus include absence sites as well as presence sites. Using the default number of 10 000 background points can however ensure that sufficient pseudo-absence sites were included in the model evaluation since the study area is small. Model evaluation is important in any modelling exercise because if the model was relatively good at predicting the true occurrence sites as present then it can be inferred that it would have predicted areas with similar conditions to the occurrence sites as present as well.

Reducing sampling bias is an important step in species distribution modelling. Therefore, applying a spatial filtering approach of one point per grid in this study as a means to dealing with the spatial clumping of points was essential. Kramer-Schadt *et al*. (2013) and Boria *et al*. (2014) found that their model trained with spatially filtered data performed better than those not trained with filtered data. In the current study, after the spatial filtering approach, the general clumped appearance of occurrence points on the Swartberg Mountain chain was however still evident (Figure 15). Clumped points almost always contribute to the environmental variables where these points are to explain much of the distribution, i.e. they contribute a lot to model gain (Segurado *et al.,* 2006). This is evident in the model output of the most preferred vegetation community, namely ericaceous fynbos, where a high concentration of points was found even after reducing it to one point per 10 m grid. This is a case of occurrence points being autocorrelated with one another and with the environmental or predictor variable. Running models at 100x100 m and consequently filtering occurrence points to one point per 10000 m² grid could potentially have reduced this effect of clumped points. However, without spatial filtering applied and tested at this larger scale, no conclusions can be drawn that it would have resulted in a different scenario. Segurado *et al.* (2006) used a systematic subsampling approach to reduce occurrence points and consequently spatial autocorrelation. They found that this did not completely eliminate spatial autocorrelation. However, even with miniscule effects a filtering approach and other means of correcting for sampling bias is useful in species distribution modelling and therefore should be incorporated in a study when needed (Syfert *et al.*, 2013; Fourcade *et al.*, 2014; Stolar & Nielsen, 2015).

As a consequence of more occurrence points in ericaceous fynbos, it contributed to the cells where this community is present to have high probability values. Almost all the areas indicated in yellow to red in the probability maps (Figures 25, 26, 37 and 38) are the exact outline of each pocket of this vegetation community. This vegetation community also occurs in the higher altitudinal classes i.e. classes 3 – 5 being from 250-1084 m.a.s.l. As a consequence of altitude being shown to play an important role in klipspringer habitat, it follows that klipspringers can be expected to roam in the vegetation communities associated with high elevation. Upland restioid fynbos having the second highest probability of presence score was unexpected, as not so many occurrences were recorded in this vegetation community compared to mesic oligotrophic proteoid fynbos and mesic mesotrophic proteoid fynbos. This vegetation does occur at a high altitude (class 5, 751-1084 m.a.s.l.), and therefore high altitudinal conditions coinciding with upland restioid fynbos can possibly explain this result. Both ericaceous and upland restioid fynbos consist of a low vegetation structure (Cowling *et al.*, 1996). Thus the animal's vision will not be obstructed in these vegetation communities, which is important, as the klipspringer relies on flight to respond to predators instead of hiding (Tilson & Norton, 1981). Since all vegetation communities in the Park were surveyed it can be confirmed that klipspringers currently only occur in Peninsula Sandstone Fynbos. As expected Forest and Thicket areas are avoided except for one occurrence point that was recorded on the margin of this vegetation type that bordered with ericaceous fynbos. Renosterveld and associated grasslands are also avoided. This latter vegetation type is mostly on Signal Hill where there are frequent human activities (Table Mountain National Park, 2011).

Accessibility to the predicted suitable areas was not incorporated in the modelling. Thus as a reference to the biotic, abiotic, movement diagram (Chapter 1, section 1.5) it would not be the occupied geographic area (G*o*) that has been predicted but rather areas of overlap between circle A (abiotic conditions) and circle B (biotic conditions, Soberon & Peterson, 2005). This is essentially the realised niche (Soberon & Nakamura, 2009). Also, not all biotic factors were included e.g. the effect of predators and diseases. Acquisition and incorporation of such biotic data could change the predicted distribution of suitable habitat for TMNP's klipspringers as reviewed in Wisz *et al.* (2013) using several plant and animal species across the globe.

Considering the 10 percentile binary map and the default RM probability map as shown in Figures 36b and 37, those areas in the central section of the Park where habitat suitability is predicted is possibly the result of high altitudinal ranges that exist there. This section largely has not had klipspringer individuals establishing there, except for the recent individual/-s that established a territory on Elsies Peak, above Fish Hoek, which is roughly where the central section ends. The central section is largely classified as a remote area in TMNP's conservation development framework report of visitor use zones (Table Mountain National Park, 2011). Thus, too frequent human activities are not the reason for klipspringers not establishing in most parts of this section. A lack of sufficient corridors to reach areas in the central section or undiscovered corridors by this antelope might be a possible reason for non-occupancy in the greater portion of the central section at this stage.

Model output will without a doubt be affected by the environmental variables used, as well as the modelling tool and scale (Elith & Leathwick, 2009). Spatial scale was tested at a resolution of 10x10 m, 100x100 m and 1000x1000 m, of which the result led to the smallest resolution being used in this study. Other factors that can affect the model results include the degree of correlation amongst the predictor variables, spatial autocorrelation resulting from clumped occurrence points and sampling bias (Segurado *et al*., 2006; Boria *et al*., 2014; Fourcade *et al.*, 2014). Consequently, spatial autocorrelation was tested between the climatic data and the topographical data. This led to the exclusion of the monthly rainfall variables from the model. The effects of clumped points were reduced through a spatial filtering approach of one point per 100 m^2 grids. Efforts were made in the sampling design of the 2014 field work to improve possible sampling bias and a stratified sampling approach was employed to this end. This ensured that areas other than only the known territories were sampled. Furthermore, it also ensured that sampling took place within all the vegetation communities or vegetation types present in the study area, at variable slope and altitudinal ranges. All these have to be taken into consideration when interpreting and reporting the model output for a specific study, although they were also all tested to some extent in this study.

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS

Home range analysis and species distribution modelling for klipspringers in South Africa, including Table Mountain National Park, is still a new field of research. Furthermore, producing home range estimates and suitable habitat predictions is not a clear cut exercise as was seen in this research. However, the results obtained for the afore-mentioned analyses can serve as a base line for future research to build on.

Home range analyses provided useful results to visualise the distribution of klipspringer home ranges in the Park. The interpretation of the CVh output as territories also assisted in interpreting the output of a univariate fixed kernel. These estimated home range sizes and territories as well as the areas of core activity had to be interpreted with caution. The home range analyses showed that the distribution and density of the occurrence points together with the home range tool used largely affects the prediction.

Species distribution modelling confirmed that altitude plays an important role in klipspringer habitat selection. It also showed that the vegetation communities accompanied by their moisture and structural conditions contribute to where klipspringers will establish a territory. Klipspringers are also not much affected by human settlements on the lower slopes of their territories. This is useful to know since TMNP is surrounded by an urban area and has a variety of human leisure activities occurring within the Park (Table Mountain National Park, 2011). The probability of occurrence map and to a certain extent the binary map is a useful map representation of suitable habitat for klipspringers. Thus Park management can incorporate these maps into their conservation efforts for the klipspringer species. These maps can also guide decision making for future release sites if more individuals are acquired from other protected areas.

Future home range and distribution studies can be improved through GPS collaring of all territorial males and collecting data in summer and winter. This can contribute to a better understanding of the movement patterns of klipspringers in a fynbos environment and national park where different degrees of human activities take place (Table Mountain National Park, 2011). Sufficient funding for GPS collars and all logistical requirements would need to be secured though for such an operation. An alternative home range approach to the fixed kernel approach used in this study is the adaptive kernel approach which can be explored for future studies since it places a higher weight on sparsely distributed points and less weight on dense points. As with home range analyses tools, there will be drawbacks in the tool decided on. This also applies to species distribution modelling. Therefore, it is important to tune these tools according to the specific study species and the goals of the study (Anderson & Gonzales, 2011).

A commitment of Park management to implement a continual monitoring programme for this species would prove important for future studies. Monitoring can include park rangers using GPS devices during their patrols and other monitoring initiatives. In this way the true home range and territories of the klipspringers can be determined more accurately. Improved unbiased data would also mean more home range analyses tools can be explored and ultimately a better representation and understanding of home ranges and territories for the klipspringers of TMNP.

Species distribution modelling or habitat suitability modelling fulfils a vital function in the field of conservation ecology (Guisan *et al*., 2013). It contributes to the long term survival and conservation of a species through identification of environmental requirements for a species over a large geographic space. It can lead to improved species management plans for rare species and species of special conservation concern. For the klipspringer species, which is a species of special conservation concern in TMNP (Rebelo *et al*., 2010), such a species management plan can be developed with the model, indicating that all three sections of the Park have suitable habitat available for this species. Some areas, especially in the central section, have potential suitable habitat that is unoccupied currently and can therefore be targeted as an area for possible future reintroductions.

.

REFERENCES

Anderson, R.P.; Lew, D.; Peterson, A.T. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, 162: 211-232.

Anderson, R.P.; Gonzalez Jr., I. 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with MaxEnt. *Ecological Modelling*, 222: 2796-2811.

Araujo, M.B.; Guisan, A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33: 1677-1688.

Baldwin, R.A. 2009. Use of maximum entropy modeling in wildlife research. *Entropy*, 11: 854-886.

Beaumont, L.J.; Hughes, L.; Poulsen, M. 2005. Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling*, 186: 250-269.

Beyer, H.L. 2014. *Geospatial Modelling Environment*. Available at http://www.spatialecology.com/gme/kde.htm

Blank, L.; Blaustein, L. 2012. Using ecological niche modelling to predict the distributions of two endangered amphibian species in aquatic breeding sites. *Hydrobiologia*, 693: 157-167.

Boria, R.A.; Olson, L.E.; Goodman, S.M.; Anderson, R.P. 2014. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, 275: 73-77.

Bradley, A.P. 1997. The use of Area Under the ROC Curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7): 1145-1159.

Brown, J.H.; Stevens, G.C.; Kaufman, D.M. 1996. The geographic range: size, shape, boundaries and internal structure. *Annual Review of Ecology and Systematics*, 27: 597-623.

Carpenter, G.; Gillison, A.N.; Winter, J. 1993. DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation*, 2: 667-680.

Chase, J.M.; Leibold, M. 2003. *Ecological Niches: Linking classical and contemporary approaches.* The University of Chicago Press: Chicago.

Compton, J.S. 2004. *The rocks & mountains of Cape Town*. Double Storey Books: Cape Town.

Connell, J.H. 1961. The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. *Ecology*, 42 (4): 710-723.

Cowling, R.M.; Macdonald, I.A.W.; Simmons, M.T. 1996. The Cape Peninsula, South Africa: physiographical, biological and historical background to an extraordinary hot-spot of biodiversity. *Biodiversity Conservation*, 5: 527-550.

Druce, D.J.; Brown, J.S.; Kerley, G.I.H.; Kotler, B.P.; Mackey, R.L.; Slotow, R. 2009. Spatial and temporal scaling in habitat utilisation by klipspringer (*Oreotragus oreotragus*) determined using giving-up densities. *Austral Ecology*, 34: 577-587.

Dunbar, R.I.M. 1979. Energetics, thermoregulation and the behavioural ecology of klipspringer. *African Journal of Ecology*, 17: 217-230.

Dunbar, R.I.M.; Dunbar, E.P. 1974. Social organisation and ecology of the klipspringer (*Oreotragus oreotragus*) in Ethiopia. *Zeitschrift fur Tierpsychologie*, 35: 481-493.

Dunbar, R.I.M.; Dunbar, E.P. 1980. The pair bond in klipspringer. *Animal Behaviour*, 28: 219-229.

Duong, T. 2001. Seminar Notes: Weatherburn lecture series for the Department of mathematics and Statistics, at the University of Western Australia. Available at http://www.mvstat.net/tduong/research/seminars/seminar-2001-05/index.html

Elith, J.; Graham, C.H. 2009. Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 32: 66-77.

Elith, J.; Leathwick, J.R. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, 40: 677-697.

Elith, J.; Kearney, M.; Phillips, S. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1: 330-342.

Elith, J.; Phillips, S.J.; Hastie, T.; Dudik, M.; Chee, Y.E.; Yates, C.J. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity Distrib*ution, 17: 43-57.

Elith, J.; Graham, C.H.; Anderson, R.P.; Dudik, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huetmann, F.; Leathwick, J.R.; Lehmann, A.; Li, J.; Lohmann, L.G.; Loiselle, B.A.; Manion, G.; Moritz, C.; Nakamura, M.; Nakazawa, Y.; Overton, J.McC.; Peterson, A.T.; Phillips, S.J.; Richardson, K.S.; Scachetti-Pereira, R.; Schapire, R.E.; Soberon, J.; Williams, S.; Wisz, M.S.; Zimmermann, N.E. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29: 129-151.

Elton, C. 1927. *Animal Ecology*. Ch. V: The Animal Community. Sedgwick and Jackson: London.

Escalante. T.; Rodriguez-Tapia, G.; Linaja, M.; Illoldi-Rangel, P.; Gonzalez-Lopez, R. 2013. Indentification of areas of endemism from species distribution models: Threshold selection and nearctic mammals. *TIP Revista Especializada en Ciencias Quimico-Biologicas*, 16(1): 5-17.

Eskilden, A.; le Roux, P.C.; Heikkinen, R.K.; Hoye, T.T.; Kissling, W.D.; Poyry, J.; Wisz, M.S.; Luoto, M. 2013. Testing species distribution models across space and time: high altitude butterflies and recent warming. *Global Ecology and Biogeography*, 22: 1293-1303.

Estes, R.D. 1991. *Behaviour guide to African mammals: including hoofed mammals, carnivores, primates*. Russell Friedman Books: South Africa.

Field, A. 2005. *Discovering statistics using SPSS*. 2nd ed. SAGE Publications: London.

Fourcade, Y.; Engler, J.O.; Rodder, 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(5): e97122. doi:10.1371/journal.pone.0097122.

Foxcroft, L.C.; McGeoch, M.A. 2011. *SANParks biodiversity monitoring programme: Invasive alien species*. Cape Research Centre: Tokai.

Francis, J.M.; Klopatek, J.M. 2000. Multiscale effects of grain size on landscape pattern analysis. *Geographic Information Sciences*, 6(1): 27-37.

Getz, W.M.; Wilmers, C.C. 2004. A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography*, 27: 489-505.

Getz, W.M.; Fortmann-Roe, S.; Cross, P.C.; Lyons, A.J.; Ryan, S.J.; Wilmers, C.C. 2007. LoCoH: Nonparameteric Kernel methods for constructing home ranges and utilisation distributions. *PLoS ONE*, 2(2): e207. doi:10.1371/journal.pone.0000207.

Goldblatt, P.; Manning, J.C. 2002. Plant diversity of the Cape Region of Southern Africa. *Annals of the Missouri Botanical Garden*, 89(2): 281-302.

Government Gazette. 2011. No. 34809. *National list of ecosystems that are threatened and in need of protection*. Department of Environmental Affairs: South Africa.

Grinnell, J. 1917. The niche-relationships of the California Thrasher. *The Auk*, 34(4): 427-433.

Gross, J.E.; Kneeland, M.C.; Reed, D.F.; Reich, R.M. 2002. GIS-based habitat models for mountain goats. *Journal of Mammalogy*, 83(1): 218-228.

Guisan, A.; Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8: 993-1009.

Guisan, A.; Edwards Jr., T.C.; Hastie, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, 157: 89-100.

Guisan, A.; Tingley, R.; Baumgartner, J.B.; Naujokaitis-Lewis, I.; Sutcliffe, P.R.; Tulloch, A.I.T.; Regan, T.J.; Brotons, L.; McDonald-Madden, E.; Mantyka-Pringle, C.; Martin, T.G.; Rhodes, J.R.; Maggini, R.; Setterfield, S.A.; Elith, J.; Schwartz, M.W.; Wintle, B.A.; Broennimann, O.; Austin, M.; Ferrier, S.; Kearney, M.P.; Possingham, H.P.; Buckley, Y.M. 2013. Predicting species distributions for conservation decisions. *Ecology letters*, 16: 1424-1435.

Haller, R.; Filli, F.; Imfeld, S. 2001. Evaluation of GPS-technlogy for tracking mountain ungulates: VHF-transmitters or GPS-collars? In: Sibbald, A.; Gordon, I.J. (Eds). *Tracking animals with GPS*. Macaulay Land Use Research Institute, Aberdeen, SD, USA. pp 61-66.

Hemson, G.; Johnson, P.; South, A.; Kenward, R.; Ripley, R.; McDonald, D. 2005. Are kernels the mustard? Data from Global Positioning System (GPS) collars suggests problems for kernel home-range analysis with least squares cross-validation. *Journal of Animal Ecology*, 74(3): 455-463.

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

Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatol*ogy, 25: 1965-1978. Available at http://www.worldclim.org/

Hirzel, A.H.; Hausser, J.; Chessel, D.; Perrin, N. 2002. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? *Ecology*, 83 (7): 2027-2036.

Hopcraft, J.G.C.; Olff, H.; Sinclair, A.R.E. 2009. Herbivores, resources and risks: alternating regulation along primary environmental gradients in savannas. *Trends in Ecology & Evolution*, 25 (2): 119-128.

Horne, J.S.; Garton, E.O. 2006. Likelihood cross-validation versus least squares cross-validation for choosing the smoothing parameter in kernel home-range analysis. *The Journal of Wildlife Management*, 70(3): 641-648.

Hosmer, D.W.; Lemeshow, S. 2000. *Applied logistic regression*. 2nd ed. Wiley: New York.

Hutchinson, G.E. 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology,* 22: 415-427.

IUCN. 2014. The IUCN red list of threatened species. Available at http://www.iucnredlist.org/details/15485/0

Jarman, P.J. 1974. The social organisation of antelope in relation to their ecology. *Behaviour*, 48 (3/4): 215-267.

Jaynes, E.T. 1957. Information theory and statistical mechanics. *The Physical Review*, 106(4): 620-630.

Jimenez-Valverde, A.; Lobo, J.M.; Hortal, J. 2008. Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity Distribution,* 14: 885-890.

Kalle, R.; Ramesh, T.; Qureshi, Q.; Sankar, K. 2013. Predicting the distribution pattern of small carnivores in response to environmental factors in the western Ghats. *PLoS ONE*, 8(11): e79295. doi:10.1371/journal.pone.0079295.

Kay, R.N.B.; van Engelhardt, W.; White, R.G. 1980. The digestive physiology of wild ruminants. In: Ruckebusch, Y.; Thivend, P. (Eds). *Digestive physiology and metabolism in ruminants. Proceedings of the 5th international symposium on ruminant physiology*. Clermont-Ferrand, France. 3-7 September 1979. pp 743-761. MTP: Lancaster.

Kearney, M.; Porter, W. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters,* 12: 334-350.

Keating, K.A.; Cherry, S. 2009. Modelling utilization distributions in space and time. *Ecology*, 90(7): 1971-1980.

Kramer-Schadt, S.; Niedballa, J.; Pilgrim, J.D.; Schroder, B.; Lindenborn, J.; Reinfelder, V.; Stillfried, M.; Heckmann, I.; Scharf, A.K.; Augeri, D.M.; Cheyne, S.M.; Hearn, A.J.; Ross, J.; Macdonald, D.W.; Mathai, J.; Eaton, J.; Marshall, A.J.; Semiadi, G.; Rustam, R.; Bernard, H.; Alfred, R.; Samejima, H.; Duckworth, J.W.; Breitenmoser-Wuersten, C.; Belant, J.L.; Hofer, H.; Wilting, A. 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity Distribution*, 19: 1366-1379.

Kruger, M. 2001. The status, distribution and ecology of the klipspringer in the Kruger National Park. MSc. Thesis. University of Pretoria.

Kumar, S.; Stohlgren, T.J. 2009. Maxent modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia. *Journal of Ecology and Natural Environment*, 1(4): 94-98.

Legendre, P. 1993. Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74(6): 1659-1673.

Lobo, J.M.; Jimenez-Valverde, A.; Real, R. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17: 145-151.

Long, P.J. 2011. Species distribution modelling – an increasing powerful tool. Available at http://www.biodiversityscience.com/2011/04/27/species-distribution-modelling/

Mackenzie, D.I. 2005. What are the issues with presence-absence data for wildlife managers? *Journal of Wildlife management,* 69(3): 849-860.

Mackenzie, D.I.; Royle, A. 2005. Designing occupancy studies: general advice and allocating survey effort. *Journal of Applied Ecology*, 42: 1105-1114.

Marell, E. n.d. *Re-establishment of klipspringer in the Table Mountain National Park*. Info and progress report.

McNaughton, S.J.; Georgiadis, N.J. 1986. Ecology of African grazing and browsing mammals. *Annual Review of Ecology and Systematics*, 17: 39-65.

Mech, L.D.; Barber, S.M. 2002. *A critique of wildlife radio-tracking and its use in national parks: A report to the U.S. National Parks service*. U.S. Geological Survey, Northern Prairie Wildlife Research Center: Jamestown, ND.

Meck, S.R. 2013. Range-use estimation and encounter probability for juvenile Steller Sea Lions (*Eumetopias jubatus*) in the Prince William Sound-Kenai Fjords Region of Alaska. MSc. Thesis. Oregon State University.

Merow, C.; Smith, M.J.; Silander, J.A. Jr. 2013. A practical guide to MaxEnt for modelling species' distributions: what it does, and why inputs and settings matter. *Ecography*, 36: 001-012.

Mucina, L.; Rutherford, M.C. (Eds). 2006. *The vegetation of South Africa, Lesotho, and Swaziland*. Strelitzia 19, South African National Biodiversity Institute: Pretoria.

Norton, P.M. 1980. The habitat and feeding ecology of the klipspringer (*Oreotragus oreotragus*) in two areas of the Cape Province. MSc. Thesis. University of Pretoria.

Norton, P.M. 1981. Activity patterns of klipspringers in two areas of the Cape Province. *South African Journal of Wildlife Research*, 11: 126-134.

Norton, P.M. 1984. Food selection by klipspringers in two areas of the Cape Province. *South African Journal of Wildlife Research*, 14: 33-41.

Pearson, R.G. 2007. Species' distribution modeling for conservation educators and practitioners. Synthesis. American Museum of Natural History. Available at http://ncep.amnh.org

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

Pebsworth, P.A.; Morgan, H.R.; Huffman, M.A. 2012. Evaluating home range techniques: use of Global Positioning System (GPS) collar data from chacma baboons. *Primates*, 53(4): 345-355.

Peterson, A.T.; Soberon, J. 2012. Species distribution modelling and ecological niche modeling: Getting the concepts right. *Natureza & Conservacao (Brazilian Journal for Nature Conservation)*, 10(2): 102-107.

Peterson, A.T.; Papes, M.; Soberon, J. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, 213: 63–72.

Phillips, S.J. 2006. *A brief tutorial on MaxEnt*. AT & T Research. Available at https://www.cs.princeton.edu/~schapire/MaxEnt/tutorial/tutorial.doc

Phillips, S.J.; Dudik, M. 2008. Modelling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. *Ecography*, 31: 161-175.

Phillips, S.J.; Elith, J. 2013. On estimating probability of presence from use-availability or presence-background data. *Ecology*, 94(6): 1409-1419.

Phiilips, S.J.; Dudik, M.; Schapire, R.E. 2004. A maximum entropy approach to species distribution modeling. *Proceedings of the 21st International Conference on machine Learning.* Banff, Canada.

Phillips, S.J.; Anderson, R.P.; Schapire, R.E. 2006. Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, 190: 231-259.

Phillips, S.J.; Dudik, M.; Elith, J.; Graham, C.H.; Lehmann, A.; Leathwick, J.; Ferrier, S. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19(1): 181-197.

Qvortrup, S.A.; Blankenship, L.H. 1974. Notes and records: Food habits of the klipspringer. *East African Wildlife Journal*, 12: 79-80.

R Core Team 2013. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3-900051-07-0, Available at http://www.R-projecy.org/.

Radosavljevic, A.; Anderson, R. 2014. Making better MaxEnt models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography,* 41: 629-643.

Rebelo, T.G.; Freitag, S.; Cheney, C.; McGeoch, M. 2011. Prioritising species of special concern for monitoring in Table Mountain National Park: The challenge of a species-rich, threatened ecosystem. *Koedoe*, 53(2): 1-14.

Rotenberry, J.T.; Preston, K.L.; Knick, S.T. 2006. GIS-based niche modelling for mapping species' habitat. *Ecology*, 87(6): 1458-1464.

Roura-Pascual, N.; Richardson, D.M.; Krug, R.M.; Brown, A.; Chapman, R.A.; Forsyth, G.G.; Le Maitre, D.C.; Robertson, M.P.; Stafford, L.; van Wilgen, B.W.; Wanneburgh, A.; Wessels, N. 2009. Ecology and management of alien plant invasions in South African fynbos: Accommodating key complexities in objective decision making. *Biological Conservation*, 142: 1595-1604.

Rudeen, C.E. 2012. Applying home range models to Greater Sage-Grouse biotelemetry data. MSc. Thesis. Idaho State University.

SANParks Scientific Services, 2010. *Update report: Klipspringer in the Cape of Good Hope section of Table Mountain National Park*.

Seaman, D.E.; Powell, R.A. 1996. An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology*, 77(7): 2075-2085.

Seaman, D.E.; Griffith, B.; Powell, R.A. 1998. KERNELHR: a program for estimating animal home ranges. *Wildlife Society Bulletin,* 26(1): 95-100.

Seaman, D.E.; Millspaugh, J.J.; Kernohan, B.J.; Brundige, G.C.; Raedeke, K.J.; Gitzen, R.A. 1999. Effects of sample size on kernel home range estimates. *The Journal of Wildlife Management*, 63(2): 739-747.

Segurado, P; Araujo, M.B.; Kunin, W.E. 2006. Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, 43: 433-444.

Shcheglovitova, M.; Anderson, R.P. 2013. Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. *Ecological Modelling*, 269: 9-17.

Sheather, S.J. 2004. Density estimation. *Statistical Science*, 19(4): 588-597.

Sillero, N. 2011. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling,* 222: 1343-1346.

Silverman, B.W. 1986. *Density estimation for statistics and data analysis. Monographs on statistics and applied probability.* Chapman & Hall: London.

Skead, C.J. 2011. *Historical incidence of the larger land mammals in the broader Western and Northern Cape.* 2nd ed. Centre for African Conservation Ecology, Nelson Mandela Metropolitan University: Port Elizabeth.

Skinner, J.D.; Smithers, R.H.N. 1990. *The mammals of the southern African sub-region.* University of Pretoria: Pretoria.

Skinner, J.D.; Chimimba, C.T. 2005. *The mammals of the southern African sub-region*. Cambridge University Press (African Branch): Cape Town.

Soberon, J. 2010. Niche and area of distribution modeling: a population ecology perspective. *Ecography*, 33: 159-167.

Soberon, J.; Peterson, A.T. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2: 1-10.

Soberon, J.; Nakamura, M. 2009. Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, 106: 19644-19650.

Stolar, J.; Nielsen, S.E. 2015. Accounting for spatially biased sampling effort in presence-only species distribution modelling. *Diversity and Distributions*, 21: 595-608.

Swets, J.A. 1988. Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293.

Syfert, M.M.; Smith, M.J.; Coomes, D.A. 2013. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PLoS ONE*, 8(2): e55158. doi:10.1371/journal.pone.0055158.

Table Mountain National Park, 2011. *Conservation development framework report: volume 1*. 2006 – 2011. Table Mountain National Park: Cape Town.

Thuiller, W.; Lafourcade, B.; Engler, R.; Araujo, M.B. 2009. BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, 32: 369-373.

Tilson, R.L. 1980. Klipspringer (*Oreotragus oreotragus*) social structure and predator avoidance in a desert canyon. *Madoqua*, 11(4): 303-314.

Tilson, R.L.; Norton, P.M. 1981. Alarm duetting and pursuit deterrence in an African antelope. *The American Naturalist*, 118(3): 455-462.

UNESCO, 2004. Committee decisions. Available at http://whc.unesco.org/en/decisions/95

Van Soest, P.J. 1994. *Nutritional Ecology of the Ruminant*. 2nd ed. Cornell University Press: USA.

Wakie, T.T.; Evangelsita, 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(11): e112854. doi:10.1371/journal.pone.0112854.

Warren, D.L.; Seifert, S.N. 2011. Ecological niche modelling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21(2): 335-342.

Wisz, M.S.; Hijmans, R.J.; Li, J.; Peterson, A.T.; Graham, C.H.; Guisan, A.; NCEAS predicting species distribution working group. 2008. Effects of sample size on the performance of species distribution models. *Diversity Distribution,* 14: 763-773.

Wisz, M.S.; Pottier, J.; Kissling, W.D.; Pellissier, L.; Lenoir, J.; Damgaard, C.F.; Dormann, C.F.; Forchhammer, M.C.; Grytnes, J.; Guisan, A.; Heikkinen, R.K.; Hoye, T.T.; Kuhn, I.; Luoto, M.; Maiorano, L.; Nilsson, M.; Normand, S.; Ockinger, E.; Schmidt, N.M.; Termansen, M.; Timmermann, A.; Wardle, D.A.; Aastrup, P.; Svenning, J. 2013. The role of biotic interactions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, 88: 15-30.

Worton, B.J. 1989. Kernel methods for estimating the utilization distribution in home range studies. *Ecology,* 70(1): 164-168.

Yackulic, C.B.; Chandler, R.; Zipkin, E.F.; Royle, J.A.; Nichols, J.D.; Grant, E.H.C.; Veran, S. 2013. Presence-only modelling using MaxEnt: when can we trust the inferences? *Methods in Ecology and Evolution*, 4: 236-243.

Young, N.; Carter, L.; Evangelista, P. 2011. A Maxent model v3.3.3.e Tutorial (ArcGISv10). Colorado. Available at http://ibis.colostate.edu/WebContent/WS/ColoradoView/TutorialsDownloads/A_Maxent_Model_v7.pdf

Zhaowen, J.; Takatsuki, S. 1999. Constraints on feeding type in ruminants: a case for morphology over phylogeny. *Mammal Study,* 24: 79-8.

APPENDICES

Appendix A: Male and female klipspringer. Sexual dimorphism exists in the form of males having horns.

Figure A1: Male klipspringer (Marell, n.d.).

Figure A2: Female klipspringer (Marell, n.d.).

Appendix B: Browse species recorded as part of the klipspringer diet in two study areas (Norton, 1984**).**

Table B1: Browse plant species recorded in the diet of klipspringers in Gamka and Springbok.

Plant species browsed on in Gamka (southern Cape)	**Plant species browsed on in Springbok (Namaqualand)**
Manochlamys albicans	*Agathosma sp.*
Didelta spinosa	*Maytenus oleoides*
Berkheya spp.	*Aspalathus hirta*
Asparagus capensis	*Phylica purpurea*
Solanum burchellii	*Cineraria sp.*
Mesembryanthemum karroense	*Metalasia gnaphaloides*
Hebenstreitia crassifolia	*Eriocephalus ajricanus*
Asparagus asparagoides	*Chrysanthemoides monilifera*
Albuca altissima	*Montinia caryophyllacea*
Othonna spp.	*Solanum tomentosum*
Pteronia spp.	*Colpoon compressum*
Montinia caryophyllacea	*Berkheya angustifolia*
Lycium oxycarpum	*Elytropappus gnaphaloides*
Galenia ajricana	*Muraltia alopecuroides*
Veltheimia capensis	*Pelargonium spp.*
Nenax dregei	*Polygala fruticosa*
Hermannia spp.	*Thesium nigromontanum*
Antizoma miersiana	*Asparagus spp.*
Thesium lineatum	*Themeda triandra*
Pharnaceum spp.	*Leucadendron salignum*
Polymita albiflora	*Relhania squarrosa*
Pelargonium grandicalcaratum	*Searsia lucida*
Lotonis longiflora	*Pentzia dentata*
Crassula brevijolia	*Euphorbia sp.*
Amoebophyllum angustum	*Searsia tomentosa*
Eriocephalus ericoides	*Elytropappus adpressus*
Moquiniella rubra	*Carpobrotus sp.*
Drosanthemum albens	*Manochlamys albicans*

Cineraria canescens	*Felicia filifolia*
Eriocephalus africanus	*Cliffortia ramosissima*
Dodonaea viscosa	*Cliffortia pulchella*
Diospyros ramulosa	*Muraltia ericaefolia*
Diospyros ramulosa	*Dioscorea elephantipus*
Arctotis revoluta	*Metalasia langebergensis*
Pelargonium dasyphyllum	*Dodonaea viscosa*
Pentzia incana	*Hermannia spp.*
Asparagus aspergillus	*Diospyros dichlorophylla*
Ruschia/Leipoldtia spp.	
Euphorbia decussata	
Osteospermum spp.	
Searsia undulata	
Tetragonia spp.	
Euphorbia mauritanica	
Indigofera spinescens	
Galenia fruticosa	
Zygophyllum spp.	

Appendix C: Topographical variables used in this modelling exercise, namely altitude (a), slope (b) and aspect (c). The Digital Elevation Model (DEM) was used as the base for the topographical variables.

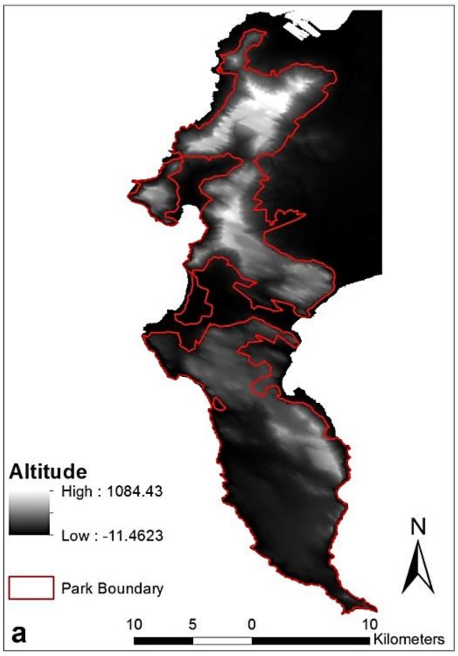


Figure C1: Altitude layer for Table Mountain National Park.

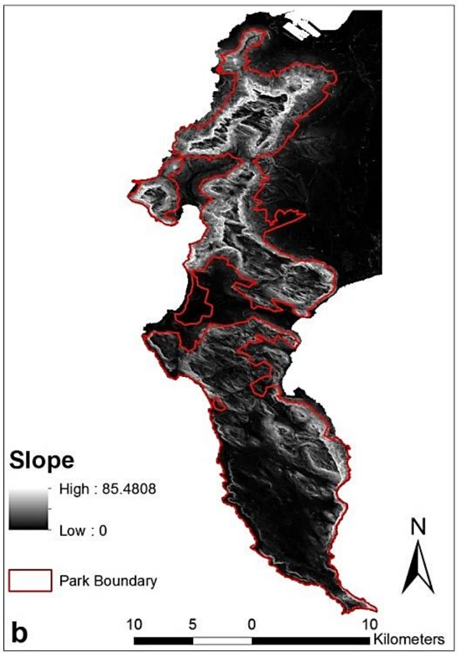


Figure C2: Slope layer for Table Mountain National Park.

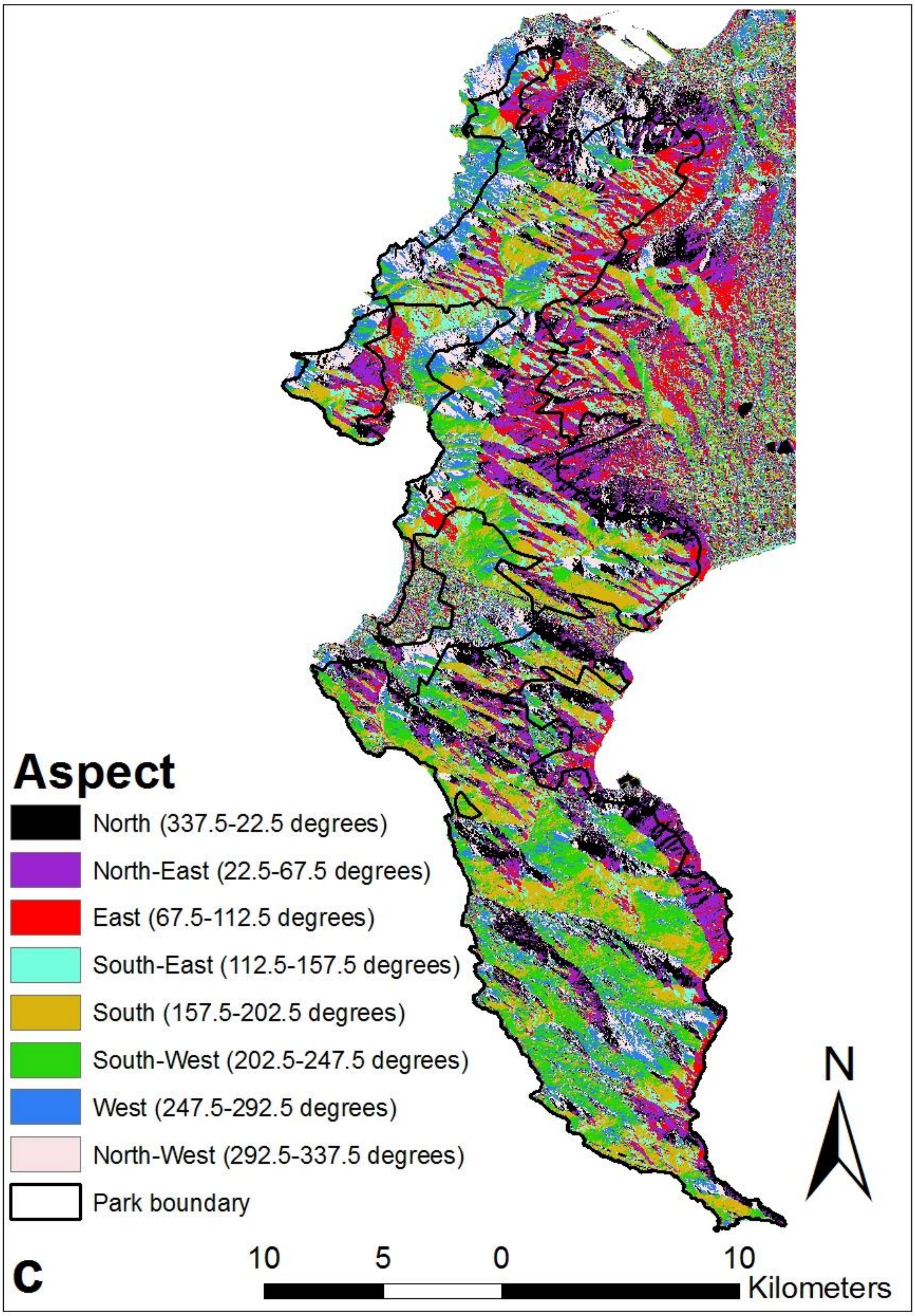


Figure C3: Aspect layer for Table Mountain National Park.

Appendix D: Survey routes enlarged for the three sections of the TMNP, namely northern, central, and southern section**.**

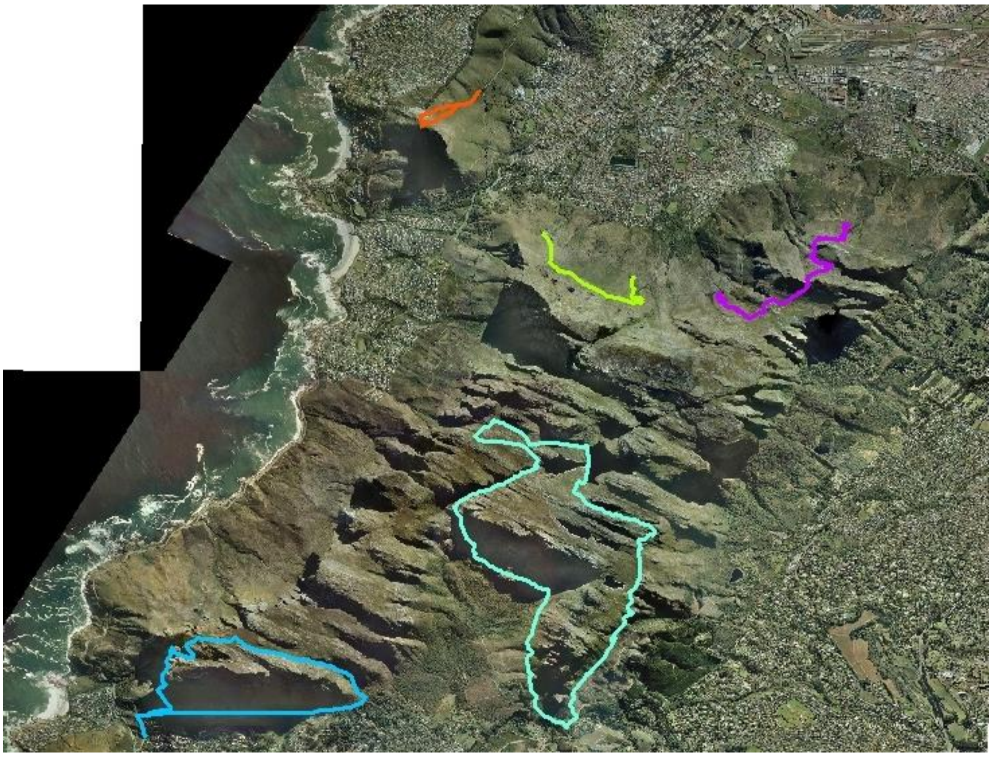


Figure D1: The five survey routes in the northern section of TMNP.

Figure D2: The three survey routes in the central section of TMNP.

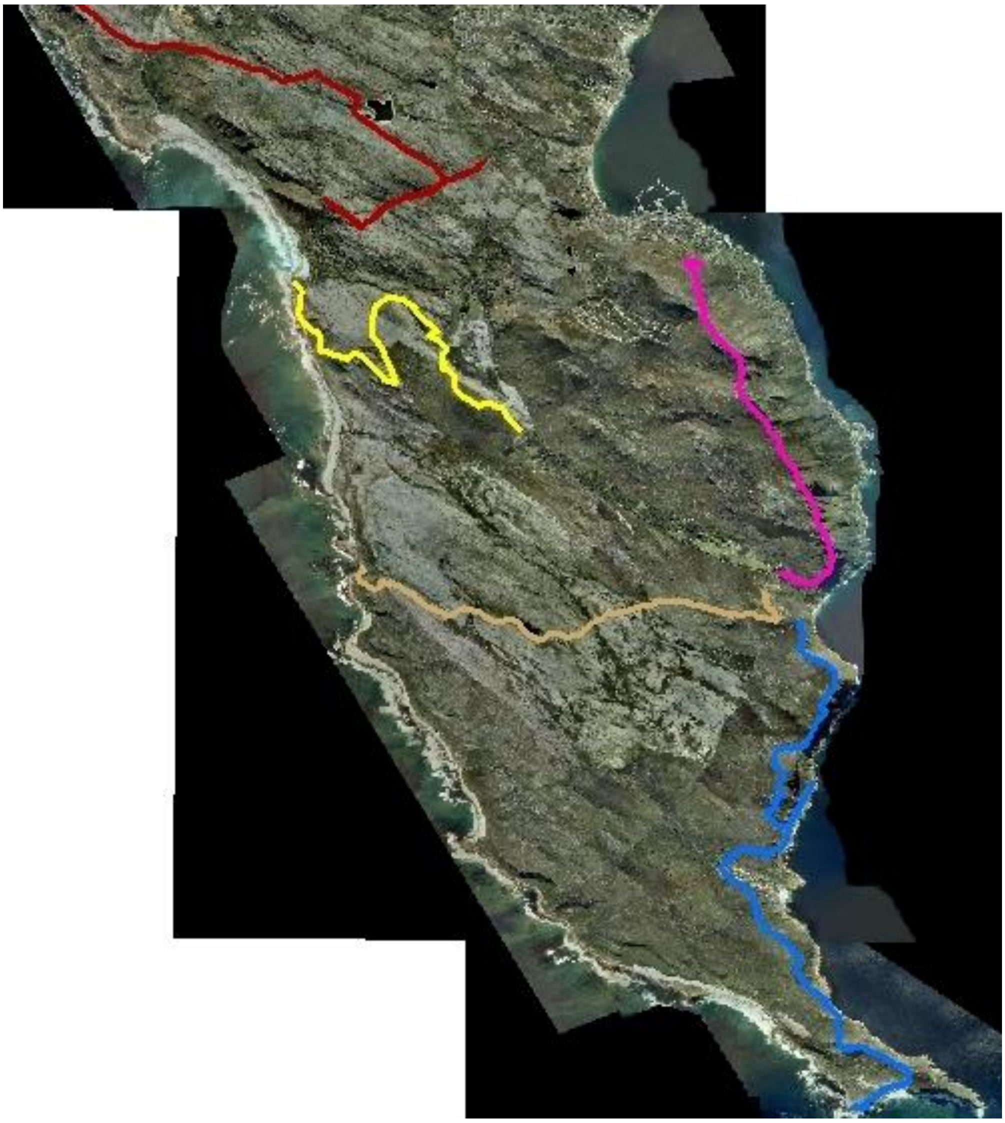


Figure D3: The five survey routes in the southern section of TMNP.

Appendix E: Extract from the Excel spread sheet showing an example of the process followed to summarise each segment of a survey route in order to determine the total length of a specific survey route passing through a specific vegetation community subtype.

Table E1: Extract of the process followed to summarise the distance walked in each vegetation community subtype.

Route 1: Southern section: Cape Maclear within CoGH to entrance gate of CoGH		
Segment number	**Vegetation community + Slope + Altitude**	**Distance**
1	Dune asteraceous fynbos 12	0.052
2	Dune asteraceous fynbos 22	0.022
3	Dune asteraceous fynbos 32	0.015
4	Deleted polygon	0.077
5	Dune asteraceous fynbos 32	0.019
6	Deleted polygon	0.173
7	Dune asteraceous fynbos 22	0.023
8	Deleted polygon	0.116
9	Dune asteraceous fynbos 22	0.086
10	Dune asteraceous fynbos 12	0.138
11	Deleted polygon	0.681
12	Mesic mesotrophic proteoid fynbos 12	0.053
13	Mesic mesotrophic proteoid fynbos 22	0.164
14	Deleted polygon	0.225
15	Mesic mesotrophic proteoid fynbos 22	0.042
16	Deleted polygon	0.231
17	Mesic mesotrophic proteoid fynbos 22	0.018
18	Deleted polygon	0.063

Appendix F: An example of one of the gridded maps used in the 2014 survey.

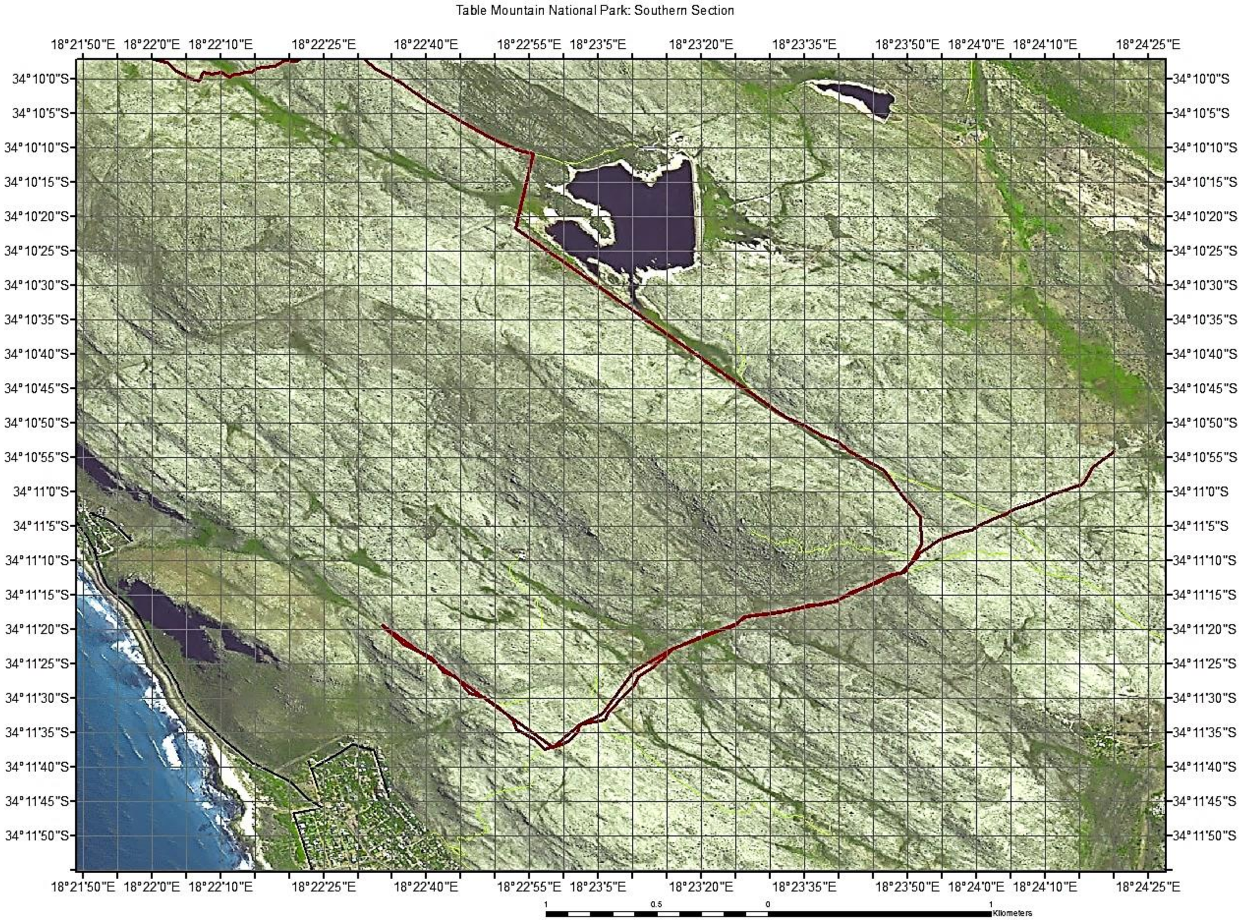


Figure F1: Example of a map gridded with latitude-longitude coordinates used to assist in the klipspringer surveys of 2014 in TMNP.

Appendix G: Klipspringer spoor and dung pellets recorded in TMNP.

Figure G1: A GPS point was recorded when recognisable klipspringer dung pellets or spoor (tracks) were observed during a survey.

Appendix H: Distance to urban edge layer classified into 0.5 km distance bands.

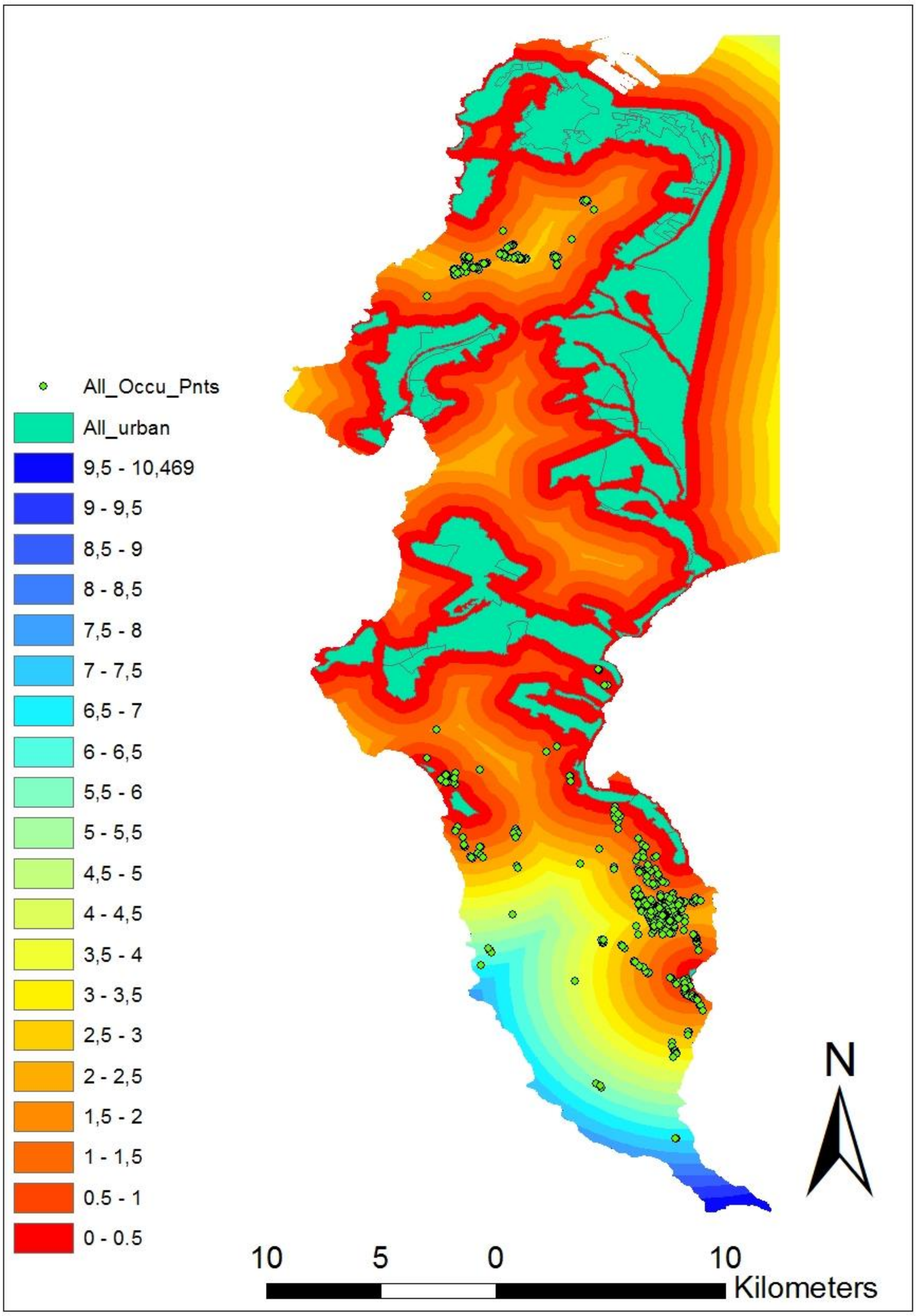


Figure H1: Distance to urban edge layer as produced from the polygon layer named "All_urban" using the euclidean distance tool and classifying it into distance bands of 0.5 km

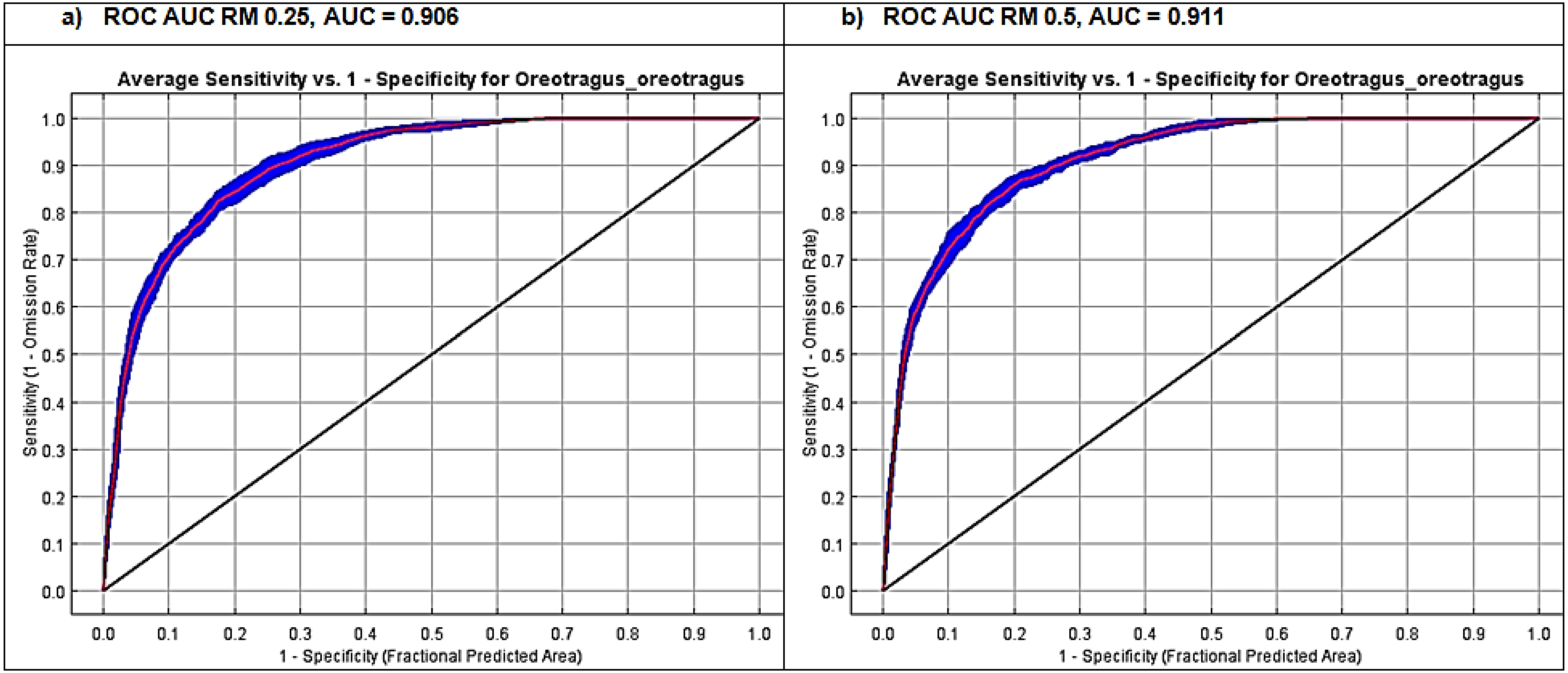
Appendix I: Correlation test output using the Pearson's two tailed correlations test.

Table I1: Two tailed Pearson's correlation test results. Highlighted blocks show the correlation of altitude with monthly rainfall, as well as the correlation of monthly rainfall with one another. [**. Correlation is significant at the 0.01 level (2-tailed), *. Correlation is significant at the 0.05 level (2-tailed)].

		Aspect	**Altitude**	**Slope**	**June rainfall**	**May rainfall**	**April rainfall**	**February rainfall**	**March rainfall**	**January rainfall**	**December rainfall**	**November rainfall**	**October rainfall**	**September rainfall**	**August rainfall**	**July rainfall**
Aspect	Pearson Correlation	1	.178**	-.083*	.211**	.222**	.223**	.266**	.194**	.202**	.231**	.259**	.240**	.221**	.202**	.208**
	Sig. (2-tailed)		.000	.016	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
Altitude	Pearson Correlation	.178**	1	.081*	.532**	.724**	.602**	.567**	.696**	.741**	.751**	.658**	.571**	.681**	.587**	.580**
	Sig. (2-tailed)	.000		.019	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
Slope	Pearson Correlation	-.083*	.081*	1	-.050	.000	-.022	-.040	-.005	-.001	.001	-.029	-.042	-.016	-.041	-.037
	Sig. (2-tailed)	.016	.019		.148	.989	.518	.240	.877	.966	.977	.405	.217	.633	.232	.281
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
June rainfall	Pearson Correlation	.211**	.532**	-.050	1	.957**	.992**	.987**	.966**	.943**	.944**	.978**	.995**	.976**	.995**	.998**
	Sig. (2-tailed)	.000	.000	.148		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
May rainfall	Pearson Correlation	.222**	.724**	.000	.957**	1	.981**	.963**	.995**	.997**	.998**	.989**	.970**	.997**	.975**	.974**
	Sig. (2-tailed)	.000	.000	.989	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
April rainfall	Pearson Correlation	.223**	.602**	-.022	.992**	.981**	1	.991**	.984**	.969**	.972**	.991**	.996**	.992**	.993**	.996**
	Sig. (2-tailed)	.000	.000	.518	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
February rainfall	Pearson Correlation	.266**	.567**	-.040	.987**	.963**	.991**	1	.959**	.944**	.954**	.990**	.997**	.976**	.981**	.986**
	Sig. (2-tailed)	.000	.000	.240	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
March rainfall	Pearson Correlation	.194**	.696**	-.005	.966**	.995**	.984**	.959**	1	.995**	.991**	.982**	.970**	.996**	.983**	.981**
	Sig. (2-tailed)	.000	.000	.877	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
January rainfall	Pearson Correlation	.202**	.741**	-.001	.943**	.997**	.969**	.944**	.995**	1	.996**	.978**	.954**	.991**	.967**	.963**
	Sig. (2-tailed)	.000	.000	.966	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
December rainfall	Pearson Correlation	.231**	.751**	.001	.944**	.998**	.972**	.954**	.991**	.996**	1	.985**	.959**	.993**	.964**	.963**
	Sig. (2-tailed)	.000	.000	.977	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
November rainfall	Pearson Correlation	.259**	.658**	-.029	.978**	.989**	.991**	.990**	.982**	.978**	.985**	1	.991**	.994**	.984**	.986**
	Sig. (2-tailed)	.000	.000	.405	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
October rainfall	Pearson Correlation	.240**	.571**	-.042	.995**	.970**	.996**	.997**	.970**	.954**	.959**	.991**	1	.984**	.992**	.995**
	Sig. (2-tailed)	.000	.000	.217	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
September rainfall	Pearson Correlation	.221**	.681**	-.016	.976**	.997**	.992**	.976**	.996**	.991**	.993**	.994**	.984**	1	.988**	.988**
	Sig. (2-tailed)	.000	.000	.633	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
August rainfall	Pearson Correlation	.202**	.587**	-.041	.995**	.975**	.993**	.981**	.983**	.967**	.964**	.984**	.992**	.988**	1	.998**
	Sig. (2-tailed)	.000	.000	.232	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845
July rainfall	Pearson Correlation	.208**	.580**	-.037	.998**	.974**	.996**	.986**	.981**	.963**	.963**	.986**	.995**	.988**	.998**	1
	Sig. (2-tailed)	.000	.000	.281	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	N	845	845	845	845	845	845	845	845	845	845	845	845	845	845	845

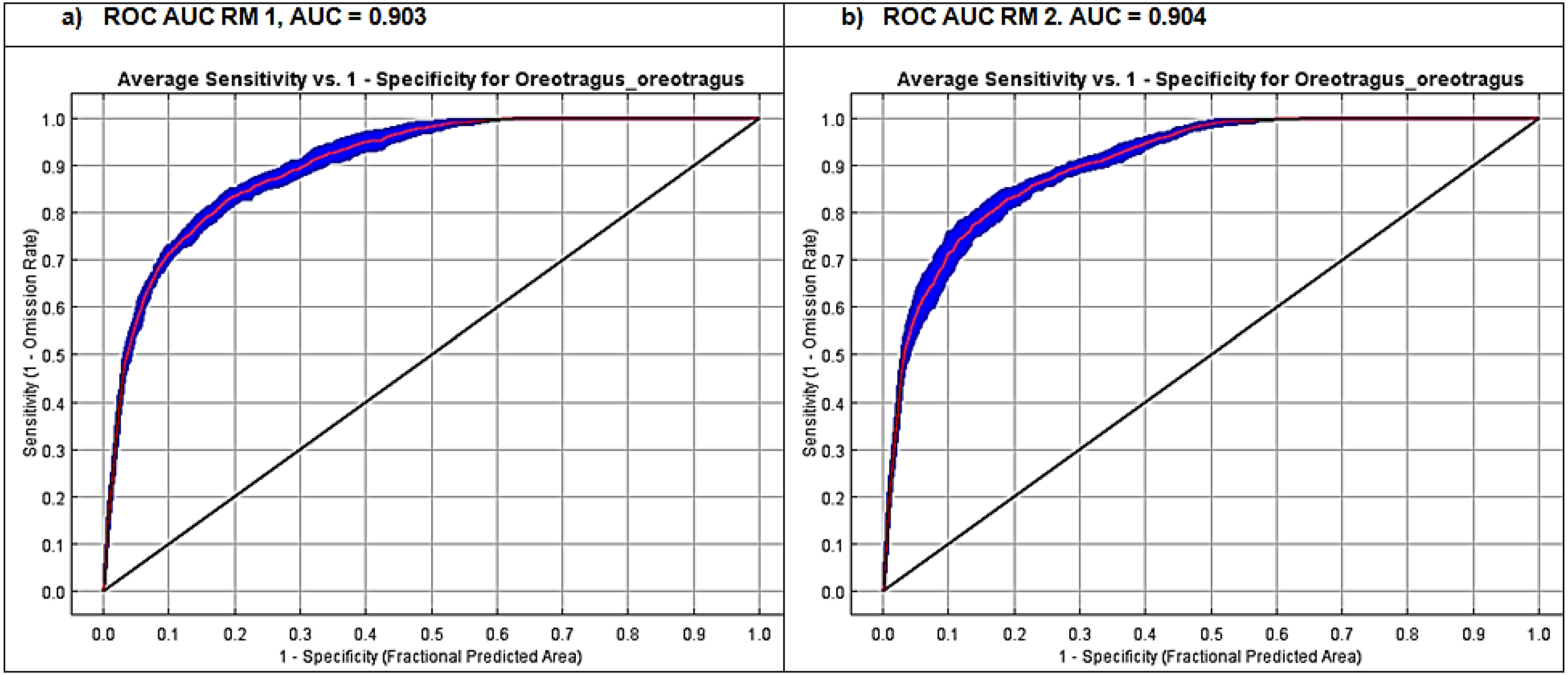
Appendix J: ROC AUC, omission error and jackknife output for the six models.



a) ROC AUC RM 0.25, AUC = 0.906

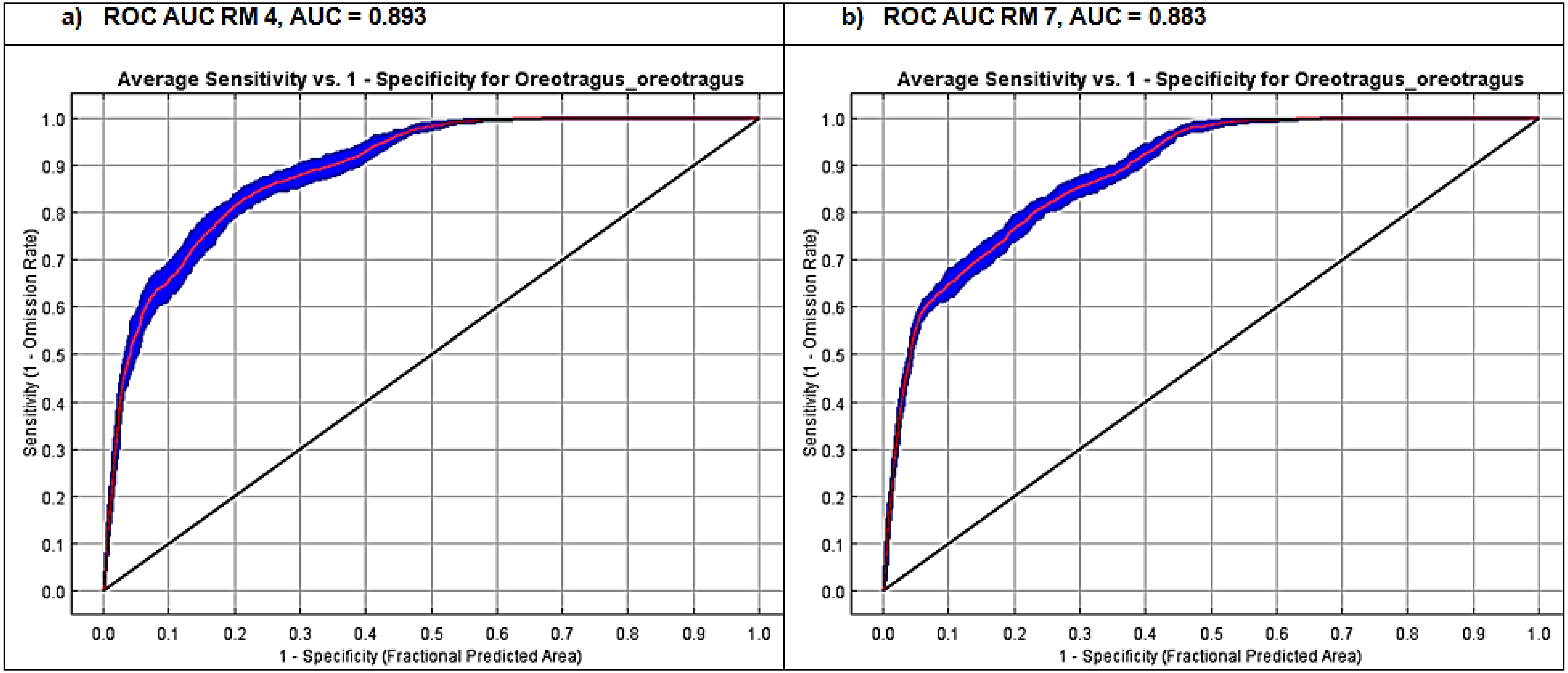
b) ROC AUC RM 0.5, AUC = 0.911

Figure J1: ROC AUC model evaluation results for the models produced at RM 0.25 (a) and RM 0.5 (b).with an AUC score of 0.906 and 0.911 respectively.


a) ROC AUC RM 1, AUC = 0.903

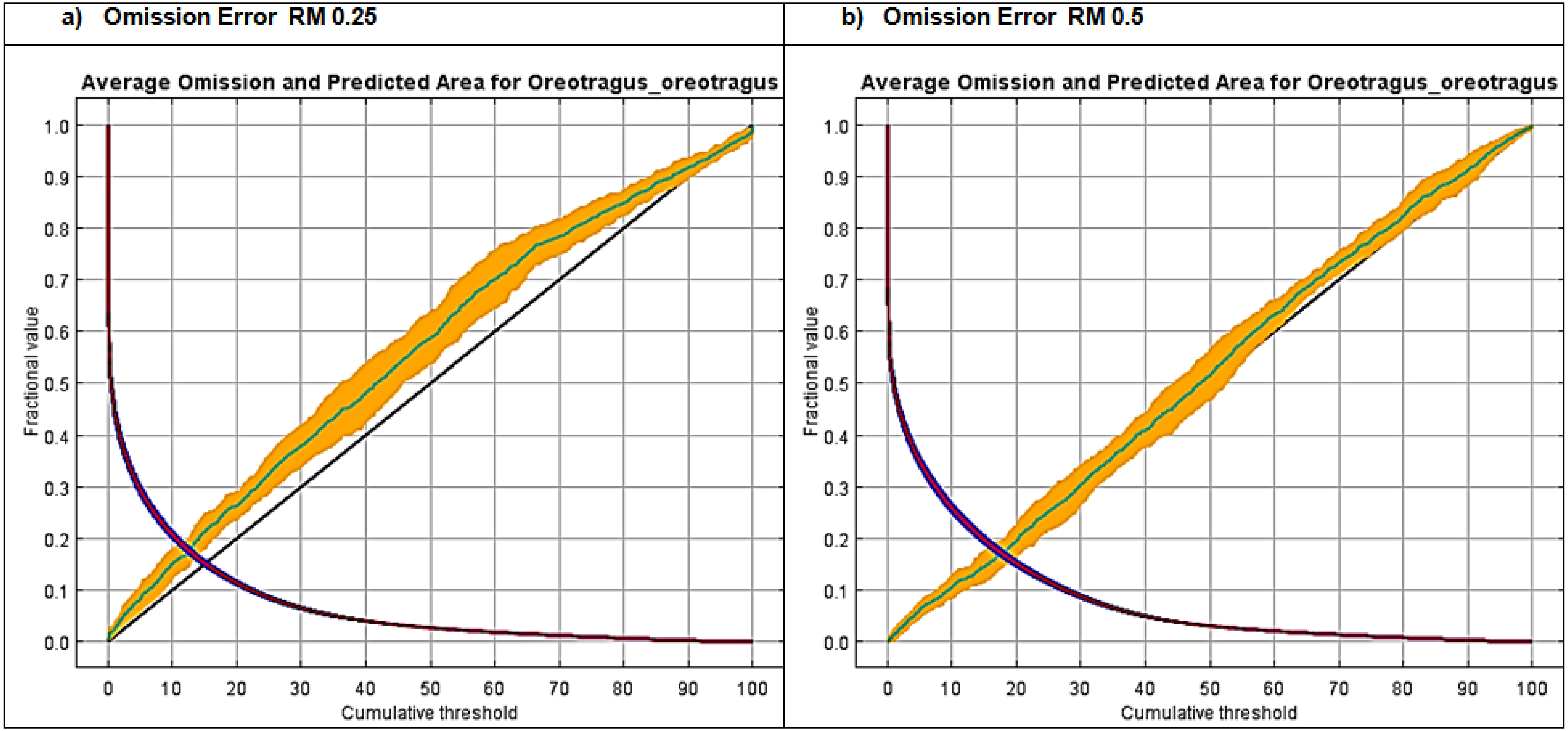
b) ROC AUC RM 2. AUC = 0.904

Figure J2: ROC AUC model evaluation results for the models produced at RM 1 (a) and RM 2 (b) with an AUC score of 0.903 and 0.904 respectively.


a) ROC AUC RM 4, AUC = 0.893

b) ROC AUC RM 7, AUC = 0.883

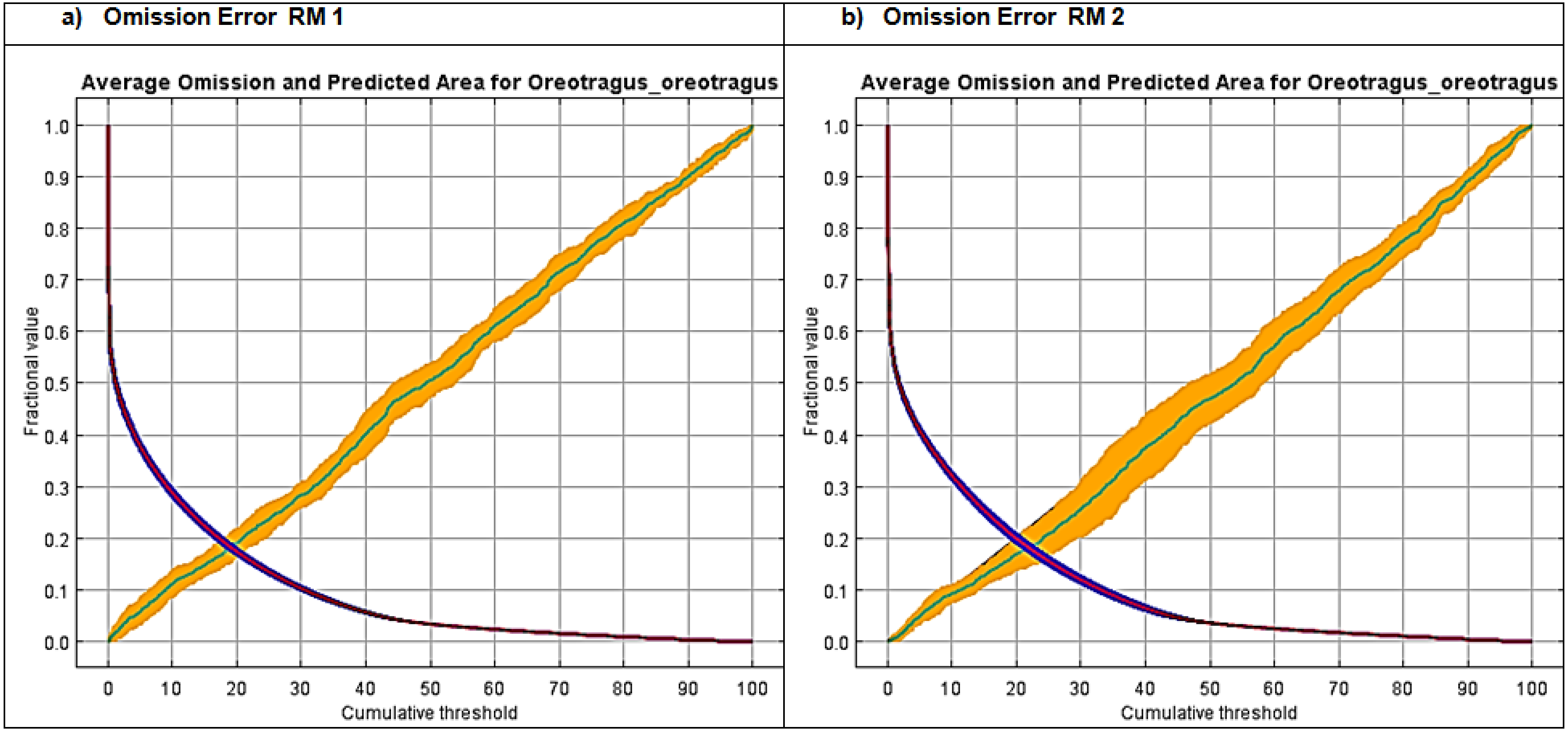
Figure J3: ROC AUC model evaluation results for the models produced at RM 4 (a) and RM 7 (b).with an AUC score of 0.893 and 0.883 respectively.



a) Omission Error RM 0.25

b) Omission Error RM 0.5

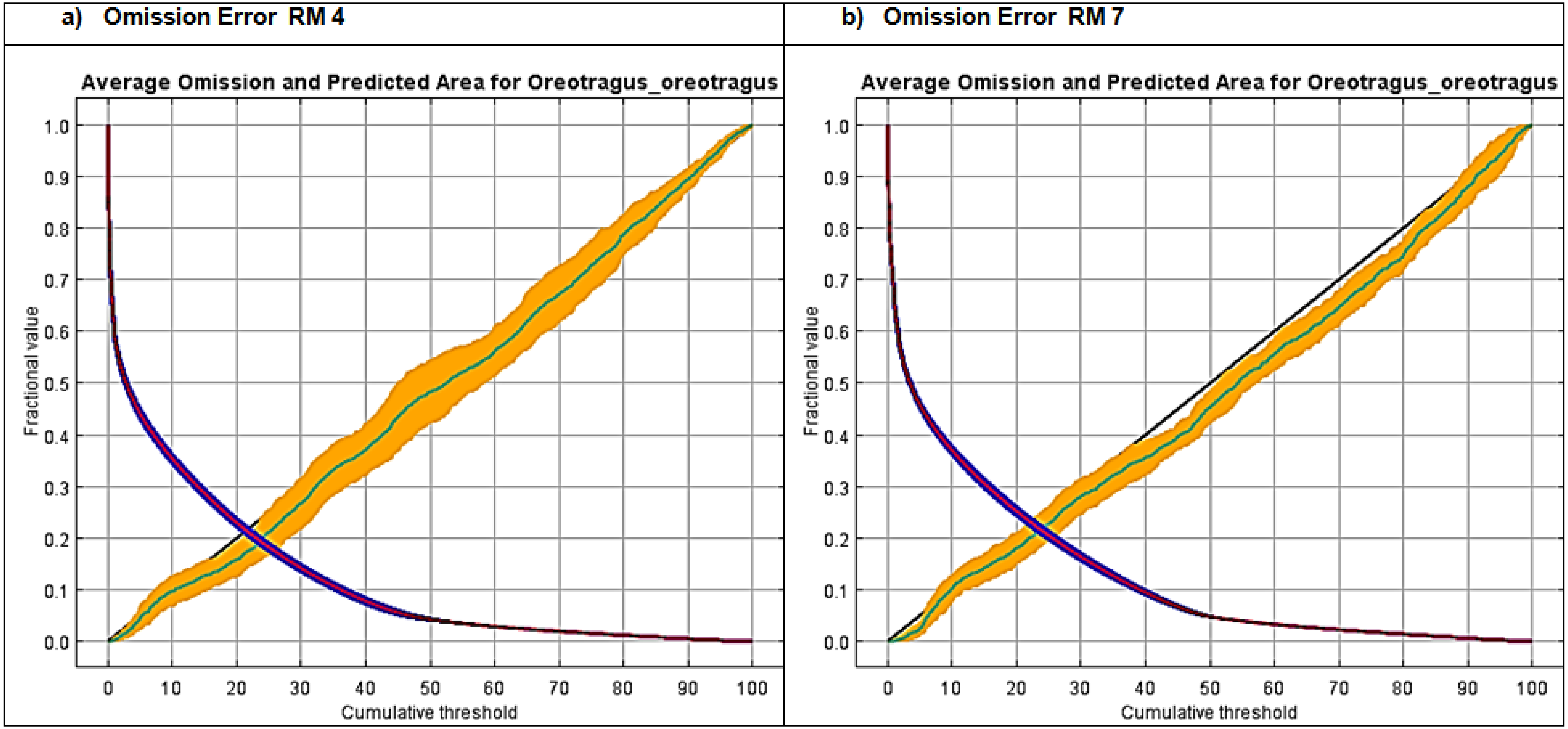
Figure J4: Omission error results for the models produced at RM 0.25 (a) and RM 0.5 (b).



a) Omission Error RM 1

b) Omission Error RM 2

Figure J5: Omission error results for the models produced at RM 1 (a) and RM 2 (b).

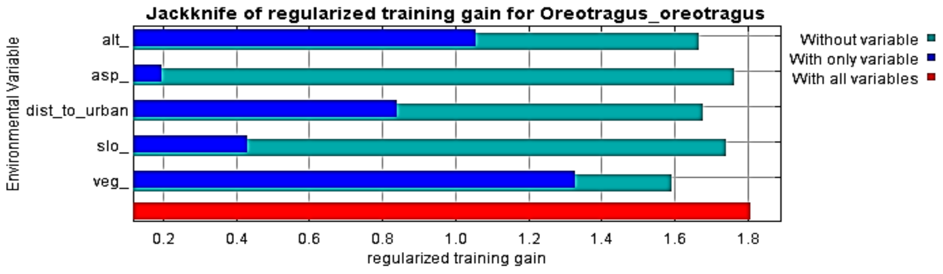


a) Omission Error RM 4

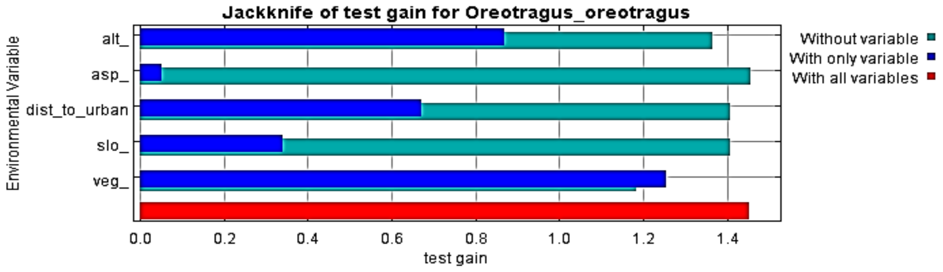
b) Omission Error RM 7

Figure J6: Omission error results for the models produced at RM 4 (a) and RM 7 (b).

Jackknife of regularized training (a) and test gain (b) at RM 0.25

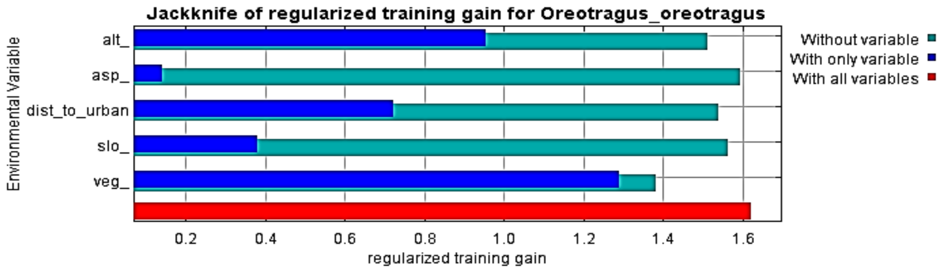


a

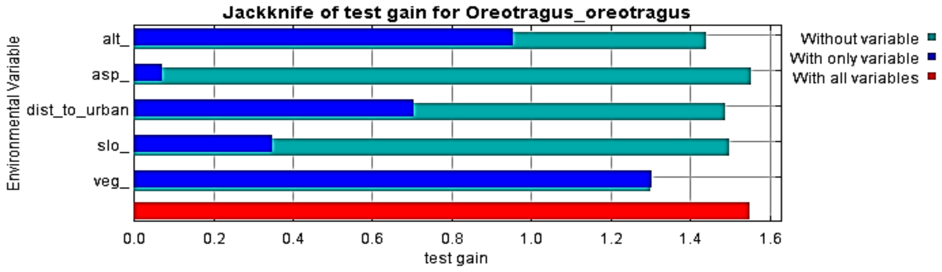


b

Jackknife of regularized training (c) and test gain (d) at RM 0.5



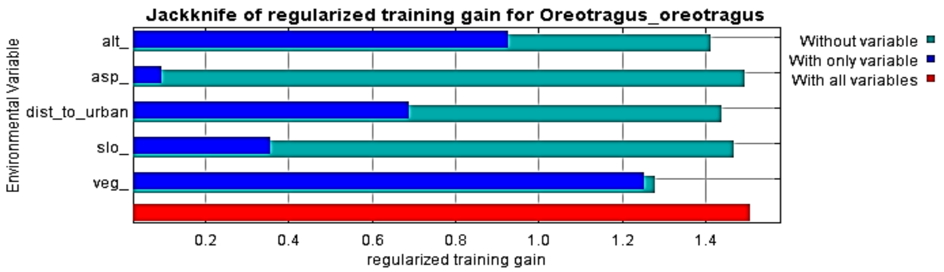
c



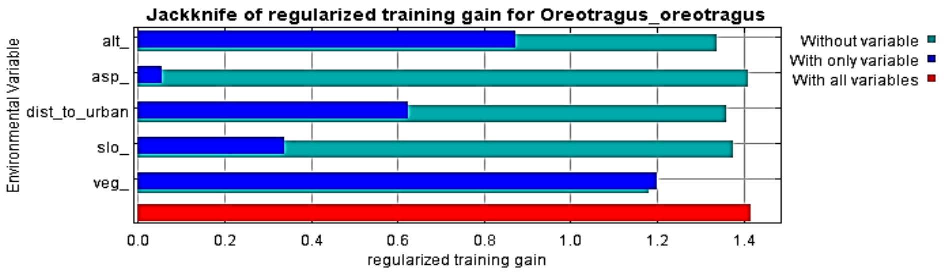
d

Jackknife of regularized training (e) and test gain (f) at RM 1

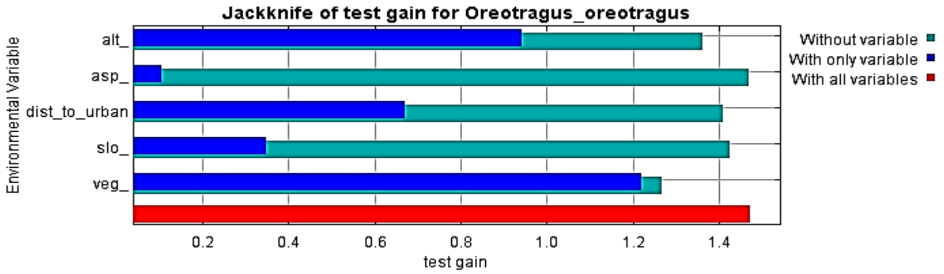
Jackknife of regularized training (g) and test gain (h) at RM 2



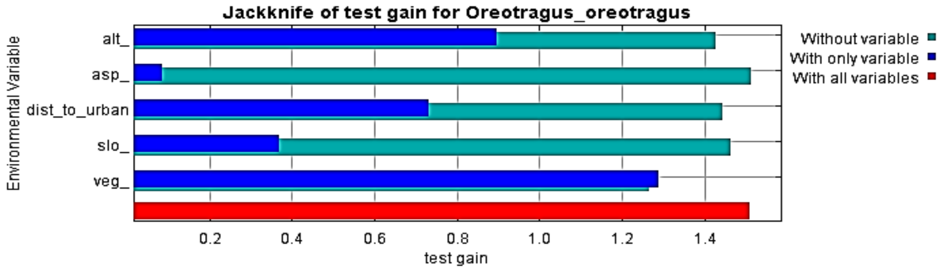
e



g



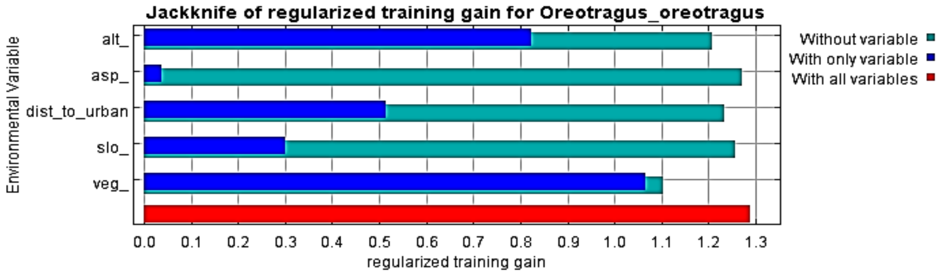
f



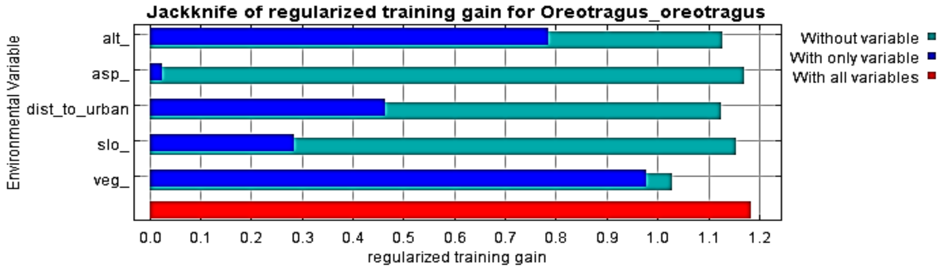
h

Jackknife of regularized training (i) and test gain (j) at RM 4

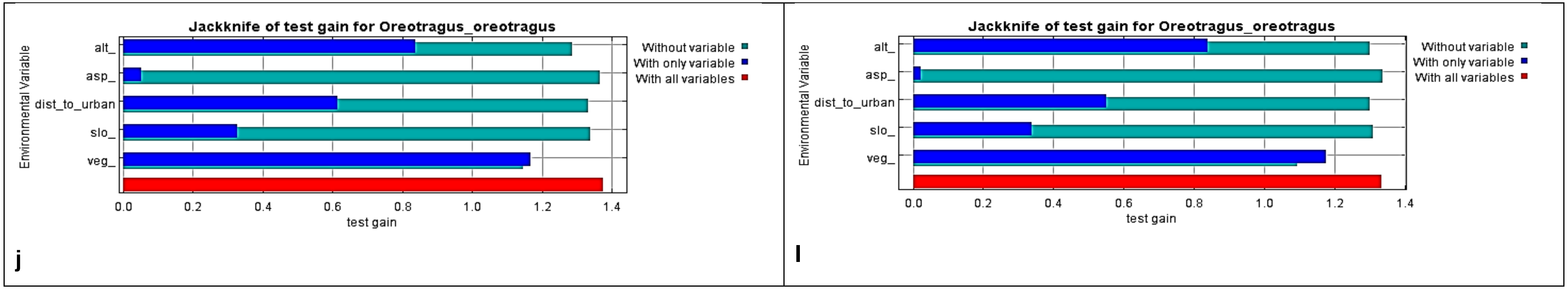
Jackknife of regularized training (k) and test gain (l) at RM 7



i



k



j

l

Figure J7: Jackknife of variable importance on training and test data for the six models at RM 0.25 (a & b), RM 0.5 (c & d), RM 1 (e & f), RM 2 (g & h), RM 4 (i & j) and RM 7 (k & l). In all these models vegetation and altitude were predicted as the most important variables and aspect as the least important.