ASSESSING ELAND HABITAT SUITABILITY

AN EXPLORATIVE STUDY ON VEGETATION AND SPECIES DISTRIBUTION IN THE SNEEUBERG NATURE RESERVE, SOUTH AFRICA

March 2016

Thesis presented for the requirements of

Master International Land and Water Management

Specialization A - Sustainable Land Management

Chairgroup: Soil Physical and Land Management

Course: SLM-80336

Wageningen University,

The Netherlands

in collaboration with Living Lands,

South Africa

Sofie Anna te Wierik

Student registration number: 930609951050

03-05-2016

Abstract

In the Sneeuberg Nature Reserve, South Africa, reintroduction of indigenous eland (*Taurotragus oryx*) has left landowners with an eland population that has grown remarkably. To reduce the risk of overgrazing from eland overabundance, ecologically sound ecosystem management that supports sustainable stocking rates is required. Understanding spatial and temporal dynamics of landscape utilization by species is pivotal to determine sustainable stocking rates. This research explores how eland utilizes his landscape by analyzing which and how environmental factors can explain spatial distribution of eland and produces a habitat suitability map for eland in the Sneeuberg Nature Reserve, representative for the dry winter season, using the commonly used species distribution model MaxEnt.

Variables that appear to explain eland distribution are altitude, net primary productivity and vegetation types. Eland density is highest in low lying riparian areas with low abundance of *Merxmuellera disticha* within an optimized plant productivity range (NDVI values of 0.2 – 0.6). The non-linear response of the predicted eland distribution to altitude suggests that soil type also plays a role in the species' distribution. Although eland is known to mainly browse, no relation between woody browse abundance and eland density was observed. However, as the quality of the woody browse species is probably more important, more research on eland dietary requirements is needed. Additionally, as predicted species-habitat associations might only be representative for the dry season, exploration of multi-temporal habitat suitability for eland is required to approximate sustainable stocking rates.

Table of Contents

Chapter 1. Introduction

South Africa is a country known for its high level of biodiversity and fascinating wildlife that are protected in numerous national parks. The statutory conservation efforts in these parks have made considerable achievements in protecting the country its natural landscape. However, the current state budgets fall short to support sustainable park management and the size of statutory protected area is too small to conserve biodiversity on the long run and possibilities for expansion are low. With approximately 80% of the land in private hands, there is increasing attention for the role of private wildlife ranching as a conservation tool (Cousins et al. 2008). The benefits of wildlife ranching for conservation are manifested in the protection of various vegetation types that are protected in eco-tourism based private wildlife reserves (Langholz and Kerley, 2006, as quoted in Cousins et al. 2008). For *wildlife conservation* in particular, economic incentives for conservation in the trophy hunting industry on private wildlife ranches have encouraged reintroduction and recovery of various African mammals, including the white rhinoceros (*Ceratotherium somum simum)* (Flack, 2003; Leader-Williams et al., 2005, as quoted in Lindsey, 2007).

In the semi-arid environment of the Karoo, South Africa, conservation efforts on privately owned land have shown to be successful: not only did they significantly complement state protected areas in terms of biomes, but also tripled the conservation targets when their achievements were taken into account (Gallo, 2008). Private conservation efforts in the Eastern Cape are stimulated through the *Eastern Cape Biodiversity Stewardship program*. This program aims to secure biodiversity through mutually beneficial public-private partnerships that enable governments to operate outside of state-protected areas and make private parties *stewards* of biodiversity through supporting management and conservation of nature on their own land. The *Compassberg Protected Environment* (CPE) is a pilot site of the Eastern Cape Biodiversity Stewardship Program. The CPE is located in the Karoo and comprises of eight private landholdings. Landowners in the CPE aim to restore important ecosystem functioning of the area to conserve the scenic beauty, support sustainable livelihoods and, most importantly, protect the area from mining activity including fracking.

One of the landholdings is the Sneeuberg Nature Reserve. The Sneeuberg has a long history of livestock herding that has resulted in an overall degraded landscape. European settlement and the consecutive increase in grazing pressure has resulted in the severe erosion and the formation of bad-lands and gully systems (Keay-Bright and Boardman, 2007). Although the Sneeuberg is characterized by a slow rate of soil formation and recovery rate from overgrazing (ECPTA, 2011),

reduced stocking rates from the 1930's onwards seem to have decreased the extent of badlands and stabilized ones that were already present (Keay-Bright and Boardman, 2009).

When the Compassberg Protected Environment was established, the indigenous eland (*Taurotragus oryx*) was reintroduced in the Sneeuberg (ECPTA, 2011). The eland is the second largest antelope species and occurs throughout one-third of Africa. Although eland has been eliminated from a large part of its former habitat by humans, IUCN labels it on the Red List of Threatened Species as 'Least Concern' (IUCN, 2015). It is an intermediate feeder that feeds on grasses, herbs, tree leaves, bushes and succulent fruits (Pappas, 2002). Currently, the eland population size has resulted in a spill-over of the eland to surrounding parts of the Sneeuberg. In some areas, the eland is invading the grazing area of the domestic stock where it is believed to compete for food resources (Steven de Bie, personal communication 03-06-2015). Landowners within the CPE are currently exploring the extent of the eland problem to come up with a long term eland management plan.

To examine the extent of this *human-wildlife conflict*, local NGO Living Lands has recently embarked on research in the Sneeuberg Nature Reserve. They aim to explore eland population sizes, species movements and diet to see how eland competes with cattle in the area and how conflict could be mitigated while simultaneously conserving sustainable wildlife populations. Population sizes and species distributions over the landscape must be taken into careful consideration for sustainable management as implied by historical overstocking and resulting degradation of the landscape. Understanding the mechanisms of overstocking and subsequent overgrazing requires spatial and temporal indications of landscape utilization by herbivores (Coughenour, 1991).

To understand how eland uses its environment, field observations that captured eland presence have been carried out. As animals species use their environment selectively, by using some areas intensively and avoiding others, the observations show areas of high and low eland densities. Selective use of habitat results from the species' distribution being constrained abiotic and biotic landscape factors that are spatially variable such as topography, vegetation cover, water availability and minerals (Bailey et al., 1996; Coughenour, 1991). How eland selects habitat in the Sneeuberg landscape is still unexplored. A study on eland in Mountain Zebra National Park, Eastern Cape (MZNP), a landscape in close proximity to Sneeuberg, has shown that the eland is mainly a browser as woody species contribute to the majority (84%) of its annual diet (Watson and Owen-Smith, 2000). It prefers open plains and may graze more during the wet season due to the higher availability of grasses and browse during dry periods (Pappas, 2002; Watson and Owen-Smith, 2000).

A first step to understand the eland in relation to his environment is indicating its use of space. Through mapping species *habitat suitability*, one can gain information on 'where they are, why they are there, and where else they could be' (Aarts, 2008). This research focuses on landscape utilization by eland though mapping habitat suitability using eland observations data from the dry winter season (May-June 2015) and various ecologically relevant environmental predictors. Predictions of habitat suitability will be made with MaxEnt, a powerful and commonly used species distribution model. Not only does MaxEnt return a spatial indication of preferred habitat, it also indicates *how* environmental variables influence the species distribution. Accordingly, indications of habitat suitability hold important information about the way eland utilizes its environment and can be used to estimate sustainable stocking rates of eland in the Sneeuberg Nature Reserve.

1.2 Study area

The Sneeuberg Nature Reserve is one of eight privately owned landholdings that together form the Compassberg Protected Environment and covers more than 40,000 ha. The Sneeuberg Nature reserve comprises of 14,500 ha and is located in the mountainous escarpment of the Karoo, at the border of the Grassland and Nama Karoo biomes (ECPTA, 2011). Landowners keep a mix of livestock and cattle, and various wildlife species are present that can roam freely throughout the area as fences between the landholdings have been taken down.

The area receives very unpredictable rainfall with an average annual of 421 mm, while the average evaporation rate amounts to 1,500-2,500 mm per year. Most precipitation falls in summer (72-80%).
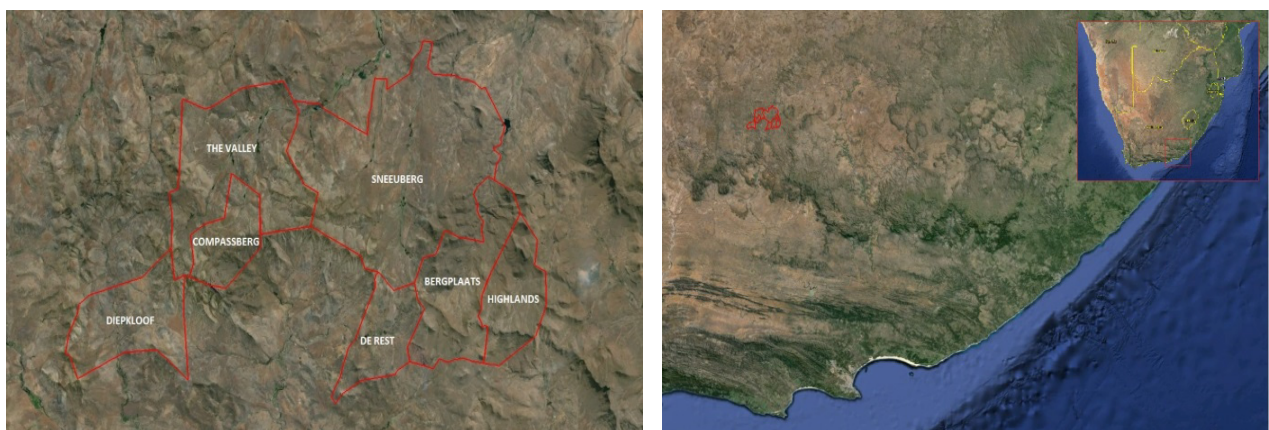

Figure 1 Overview of the research area

There is a strong altitudinal gradient in rainfall, with four times more rainfall in the top of the escarpment compared with the lower lying areas. Extreme temperatures occur, ranging from -15 °C to 44 °C (ECPTA, 2011). The Sneeuberg is an erosional landscape that is characterized by a mixture of blue and grey mudstones and sandstones with dolerite intrusive material. Layers of calcium carbonate are deposited on the near surface through continuous soil water evaporation containing dissolved calcium from sedimentary deposits. Valleys or *vleilands* consist of alluvial fans and water courses that actively, although not perennial, drain the valley (ECPTA, 2011). Soil formation is slow and soils are generally shallow and poorly developed. A distinctive A-horizon and vegetation litter on the topsoil is lacking as the topsoil consists mostly of exposed B- or C-horizons. Although the extent of degraded areas has decreased with 15% in the last 50 years, the rate of soil erosion on badlands still amounts 5-6 mm yr^{-1} with soil losses estimated to be 54.8 t ha yr^{-1} (Keay-Bright and Boardman, 2004b, as quoted in Keay-Bright and Boardman, 2007). In the valley bottoms that drain the area, the accumulated colluvial material is being eroded actively, resulting in gully systems that reach the bedrock and can be 1-4 m in depth (ECPTA, 2011). The soils are rich in minerals (especially lime) and mineral leaching is low due to the arid climate and poor drainage. Locally, high water tables can occur in alluvial sediments due to dolomite dykes cutting across a gradient, functioning as an underground dam. This has resulted in the formation of black peaty soils in some areas (ECPTA, 2011).

The general vegetation type on these semi-arid rangelands is karroid *Merxmuellera* mountain veld where the species *Merxmuellera disticha* is dominant (Acocks, 1953 - from Keay-Bright and Boardman, 2007). Grasses and shrubs are interspersed and trees are scarce. Plant productivity is constrained by the rainfall and temperature variability. However, large altitudinal ranges create high structural complexity in the area, resulting in a variety of landforms and heterogeneous vegetation patterns that are associated with the Karoo biome (ECPTA, 2011). This complexity leads to a high flowering plant species richness of 1200 species recorded in the Sneeuberg (Clark et al., 2009). The currently existing vegetation map of the Sneeuberg is derived from the national vegetation map by Mucina & Rutherford (2011). Four vegetation types are identified in the research area (description from Clarck (2010):

Figure 2 Gullies encountered in the research area

Eastern Upper Karoo: Covers the majority of the Sneeuberg landscape and occurs on the plains in the north of the Sneeuberg and is transitional between the Grassland Biome and the Nama-Karoo Biome (Mucina and Rutherford, 2011)

Upper Karoo Hardeveld: Vegetation type that includes most of the steep south-facing slopes of the Great Escarpment. Considered as one of the richer floras of the Nama-Karoo Biome, but can hardly be distinguished from thicket or other vegetation units.

Karoo escarpment grassland: Dominant grassland on the Karoo Escarpment. Occurring grass species include *Ehrharta calycina, Melica decumbens, Merxmuellera macowanii, M. disticha, Tetrachne dregei, Karroochloa purpurea, Helictotrichon spp.,* and specialist grasses such as *Festuca spp.* and *Brachypodium bolusii* at the base of scarps and *Pentaschistis airoides subsp. jugorum* on high peaks. When altitude reaches a height of 2100 m, grassland is generally replaced with 'arid fynbos'.

Southern Karoo Riviere: Riparian vegetation at lower altitudes (1000 m) that is typically dominated by *Acacia karroo, Tamarix usneoides* and *Salsola spp.* (Mucina and Rutherford, 2011).

1.3 Problem statement

Landowners in the Compassberg Protected Environment are currently looking for management strategies to conserve the scenic beauty and restore important ecosystem functioning of the area to ensure sustainable livelihoods in the area. One of the issues that they are dealing with is the re-introduction of eland that has grown to such proportions that it has caused a spillover of eland in

surrounding areas. As landowners within the CPE have reached consensus neither on the extent of the eland problem nor on the solutions that are proposed, the extent of this *human-wildlife conflict* requires further exploration. The population size and spatial extent of eland over the landscape must be taken into careful consideration as implied by historical overstocking and resulting degradation of the landscape. Understanding the mechanisms of overstocking and subsequent overgrazing, requires spatial and temporal indications of landscape utilization by herbivores.

1.4 Aim and objectives of the research

A first step to understand how eland interact with its environment, is to indicate how it utilizes the landscape. Hence, the aim of this research is to explore landscape utilization by eland in the dry winter season by 1) analyzing which and how environmental factors can explain spatial distribution of eland and 2) producing a habitat suitability map for eland in the Sneeuberg Nature Reserve.

This will contribute to a better understanding of interactions of eland with its landscape and accordingly, support sustainable wildlife management strategies for the Sneeuberg Nature Reserve or the Compassberg Protected Environment as a whole.

1.5. Research questions

The main research question is formulated as follows:

Main research question: *What environmental factors explain spatial distribution of eland and how do they define habitat suitability in the Sneeuberg Nature Reserve?*

To answer this question, environmental factors that are potentially explanatory for eland distribution will be selected. As limited data on vegetation types (as a potential explanatory factor) is available, the first research question is hence defined as:

Sub question 1: What vegetation types can be identified from aerial photography and field data collection?

Subsequently, the second and third sub-question relate to species-habitat associations and the spatial extent of predicted suitable habitat.

Sub question 2: *What environmental factors best explain eland distribution and how?*

Sub question 3: *What is the spatial extent of suitable habitat for eland?*

Figure 3 Eland with calves in Sneeuberg Nature Reserve

Chapter 2. Conceptual Framework

2.1 Habitat suitability and species-habitat associations

A *species habitat* is defined as '*any portion of the surface of the earth where the species is able to colonize and live*' (Fretwell and Lucas, 1970) and is made up of a specific combination of environmental factors. Accordingly, species are expected to be non-randomly distributed throughout the landscape (Hirzel et al., 2002). Environmental factors that structure habitats are, amongst others, vegetation, food availability, climate, soil, breeding and refuge sites, interspecific effects and individual/racial preferences (Grinnell, 1917 - quoted in Whittaker et al, 1973).

The *spatial distribution* of species is described as '*the density of animals (or, their usage) over geographical space*' (Aarts, 2008). The density of animals varies over space as they actively *select* their habitat. The *habitat matching rule* (Pulliam & Caraco, 1984, as quoted in Cassini, 2011) assumes that the species density is directly linked to the quality of the habitat. Habitat quality is species specific and is determined by a species' *ecological niche,* which is defined by Hutchinson (1957, as quoted in Hirzel and Le Lay, 2008) as: '*the volume in the environmental space that permits positive growth'*, but many other definitions exist (Hirzel and Le Lay, 2008). Estimations of habitat suitability strongly lean on the concept of ecological niche as observed species distributions in environmental space are assumed to reflect a species *realized niche*. In the ecological niche theory, a distinction between *realized* and *fundamental niche* is made. A species *realized niche* reflects the spatial distribution of species only in a part of their *fundamental niche* due to exclusion by a competitor or predator. The fundamental niche reflects that area of suitable environmental conditions that assure a population growth rate of ≥ 1 (Gausan et al., 2005). The importance of this population growth rate is stressed by Gausan et al. (2005), as habitat suitability estimations that are solely based on presence-absence observations of species and do not take into account a sexual reproduction parameter (i.e. population growth rate), may falsely reflect a species true niche.

As the eland population in Sneeuberg has shown positive growth rates, the observed species distribution is likely to reflect the species niche and hence is expected to correspond to the suitability of habitat. A precondition of this expectation however, is that each habitat is *equally accessible.* When animals are obstructed in their movement (e.g. by fencing), low observed species density may reflect inaccessibility rather unsuitable habitat (Pulliam, 2000, as quoted in Hirzel and Le Lay, 2008). As fences have been taken down throughout the Compassberg Protected

Environment to facilitate the free roaming of wildlife and livestock in the area, all areas are theoretically equally accessible to eland.

2.1.1 Selecting ecologically relevant environmental variables

When exploring species-habitat associations, selecting ecologically relevant environmental variables is important and must be *'considered from the perspective of the species'*. When considering the prediction of habitat suitability in particular, a number of studies shows that pre-selection of variables based on ecological relevance of that variable to the species improved model predictive power (Elith & Leathwick, 2009). On the other hand, the a priori choice of environmental variables can be misleading as it does not allow exploration of variables that are (unexpectedly) important (Porter and Church, 1987; Phillips et al., 2006). Nevertheless, possibilities for exploration are generally bounded by the data and knowledge available. Potential ecologically relevant variables are further elaborated on below.

Vegetation structure and composition and *topography* seem to be the main drives of African herbivore spatial distributions (Melton, 1987). Habitat selection and diet composition of eland in MZNP confirms the importance of these variables as eland showed to prefer vegetation comprising of woody species and selectively uses topographically different habitats (Watson and Owen-Smith, 2000). Additionally, a study on explanatory factors of spatial distribution of various African grazers in Kruger National Park showed that most distribution patterns related to *forage quality and quantity*, *surface water* and *habitat openness* (Smit, 2011). As *forage quality and quantity* are generally difficult to represent spatially throughout a landscape, a proxy of *net primary productivity*, the Normalized Difference Vegetation Index (NDVI), is often used in species distribution modelling to represent *forage quantity* (Smit, 2011; Pettorelli, 2011). A study on the relation between NDVI and species densities of 13 African ungulates generally shows a positive relationship between NDVI values and species distributions (Pettorelli et al., 2009). In contrast, a study by Verlinden & Masogo (1997) on eland distribution in the Kalahari showed *no significant response* of eland distributions to NDVI values. However, this is very likely due to local adaptations to the extreme environmental conditions in the Kalahari and the influence of hunting pressure that influenced eland spatial distribution, the latter acknowledged by the authors as an influential factor.

The *forage quality* as described by Smit (2011) was approximated based on geological formations and associated clay contents of the overlying soils. The importance of soil type and the availability of clay content and micronutrients is suggested by a research on herbivore distribution by Mills &

Fey (2005) as it influences the quality in terms of nutrients availability of the overlying vegetation. Eland distribution in MZNP seems to confirm this suggestion as the majority of avoided habitat occurred on '*shallow leached lithosols*' (Watson & Owen-Smith, 2000). However, as data on soil types at a relevant spatial scale was not available for the research area, this information could not be included.

The importance of *surface water* for grazer distributions is confirmed by a similar study by Redfern el al. (2003), as the availability of *surface water* showed to put a constraint of the spatial distribution of African ungulates in Kruger National park, especially in the dry season. This constrain is species specific and although for eland indifferent relations between rivers and distribution patterns were found, a positive correlation was shown to exist between artificial surface water holes and eland distribution (Smit, 2007).

Habitat openness is not considered in this research as the measures of openness generally relate to the type of vegetation ('closed woodland', 'open grassland' etc.) and the Sneeuberg is a homogenous entity of open grassland in that regard.

At last, the limited water availability in rangelands (particularly in the dry season) has shown that spatio-temporal patterns of *precipitation* greatly influences herbivore dynamics through differences in vegetation cover, composition and productivity (Vetter, 2005). Although precipitation itself does not explain distributions but rather through associated differences in vegetation cover, composition and productivity, exploration of the relevance of this variable seems worthwhile.

2.1.2 Temporal context of habitat suitability

Habitat suitability needs to be understood in a seasonal context, as both landscape as species requirements can change over seasons. Many African herbivores migrate during seasons, meaning that their suitable habitat types differs throughout the year. For eland specifically, research by Watson and Owen-Smith (2000) identified a different use of *topographical space* between seasons: in the early wet season, the plateau habitats were used, while valleys and slopes were more favorable in the late wet and dry season. Furthermore, seasonal variability in *vegetation structure and productivity* have shown to result in large spatial-temporal differences in predicted habitat suitability in a research on Mongolian antelope (Mueller et al., 2008). Eland dietary patterns observed in MZNP confirm this temporal variability as feeding behavior showed a high grass consumption in the beginning of the wet season (December), when grass is young and easy to digest. In January, the grass consumption rapidly declined, which is suggested to be due to eland its

browser-like stomach structure that cannot digest the mature grasses as well as grazers (Watson and Owen-Smith, 2000). As the eland observation data is documented in May and June 2015 (dry season), the predicted species-habitat associations and spatial extent of suitable habitat in this study might only be representative for the dry winter season.

2.2. Stocking rates

Ecosystem sustainability is determined by interactions between herbivore spatial distribution and density, vegetation growth, responses to vegetation to grazing and the topography of the landscape. The impact of these interactions is dependent on the herbivore abundance, expressed as the number of animals per unit land (per unit time) or per unit of plant production (Coughenour, 1991). Hence, the amount of herbivores that a rangeland can support is a fundamental question in wildlife management (McLeod, 1997). Overabundance of herbivores and subsequent overgrazing must be prevented as it generally results in a loss of productive capacity due to the impact on the vegetation community and the soil (Hunt et al., 2014). In the Karoo in particular, high historical domestic stocking rates have shown to affect vegetation structure and soil erosion dynamics (Du Toit, 2011; Keay-Bright and Boardman, 2007). When stocking rates regularly exceed the carrying capacity, which is the amount of grazing pressure that can be exerted *sustainably* on a certain vegetation state of a landscape (Kessler, 1994), the land conditions decline as perennial grasses disappear and barren soil surface and subsequent soil loss increases (Hunt et al., 2014). To prevent stocking rates exceeding carrying capacity, a common approach is the use of *ecologically safe utilisation rates* (Hunt, 2008 as quoted in Hunt et al., 2014). Generally, stocking rates that result in the utilization of 20-30% of the annual herbage growth is considered safe. This utilization rate approximation can evidently be adjusted to fit specific field conditions. At any rate, understanding of the way in which species utilize the landscape is important to identify sustainable stocking rates.

Chapter 3. Modelling framework

3.1 Introduction to species distribution modelling

An important aspect in conservation and restoration planning is identifying the spatial distribution of animals (Paudel et al, 2015). Species distribution models provide a powerful tool to use typically sparse data on species observations and environmental data (Elith & Leathwick, 2009). By linking species observations to environmental variables, one can indicate the *observed versus the expected observation of space use* in different habitat types. One assumes that if the *observed observation* is higher than the *expected observation* in a certain habitat type, the animal is actively selecting one habitat over another (Porter and Church, 1987). Expected observations are based on the hypothesis that *'the spatial distribution of a population of unconstrained random walkers within a region of space is approximately uniform'* (Aarts, 2008). If the observed observation shows a deviation of this expected uniform distribution, it is thus assumed that species is selectively choosing its habitat.

Most species distribution models are used to produce *habitat suitability* maps as an output (Cassini, 2011), aiming to predict the space use of species based on a priori associations of the species with environmental variables (Porter & Church,1987). As described earlier, one of the pivotal concepts in habitat suitability modelling is the concept of the *ecological niche.* Although the concept of ecological niche relates to the *fitness of a species* in relation to environmental variables, and the habitat suitability relates to the *occurrence of species* in relation to environmental variables, Hirzel and Le Lay (2008) try to integrate these concepts by stating that habitat suitability models can be seen as *'operational applications of the ecological niche'*.

3.2 Exploring different modelling approaches

In the field of species distribution modelling, many different approaches exist to model habitat suitability. Generally, they can be categorized into two classes: models that require presence/absence data or use presence-only data (Elith et al., 2011).

For *presence/absence studies*, a range of multivariate statistical approaches, are used for habitat suitability modeling that all build on the same principles of calibrating functions of environmental variables in relation to presence and absences localities (Hirzel et al., 2002). Because accurate absence data is difficult to obtain (e.g. species was not present at the time of observation), alternative approaches that work with *presence-only* data have been developed (Hirzel et al., 2001;

Hirzel et al., 2002; Paudel et al., 2015). It must be noted that limitations of presence-only data have been identified that relate to the lacking value of *prevalence* (which reflects the proportion of occupied sites) in the landscape. Another fundamental limitation is the *sample selection bias* that has a much stronger effect on presence-only data. This sampling bias can result in an environmental bias that may lead to inaccurate models (Philips et al., 2009). To correct for sampling bias, various options are proposed by Fourcade et al (2014). A systematic sampling method performed best and is the simplest way to prevent sampling bias. Nevertheless, a *sampling bias grid* performs fair as a post-sampling correction method. Some modelling approach (e.g. MaxEnt) have an inbuilt function that allows the input of a sampling bias grid. This grid reflects the *relative sampling effort* over geographical space and implicitly downgrades the weight of presence localities at locations of high survey effort. Otherwise, the model implicitly assumes that habitat types where the sampling effort was low or zero, are not important for the species distribution. For instance, if surveys at higher altitudinal zones are not carried out due to inaccessibility to the observers (hence no eland observations are documented at high altitudes), it does not specifically mean that these areas are unsuitable. Vice versa, areas that show high eland observations densities might have been easier accessible. By downgrading the relative contribution of the environmental variables at locations that probably have a higher sampling effort, sampling bias is aimed to be corrected for.

Various modelling approaches that support presence-only data have been developed (e.g. Ecological Niche Factor Analysis, Genetic Algorithm for Rule Set Prediction (GARP) and MaxEnt). MaxEnt is an extensively used species distribution model that shows a relatively high performance and is freely available. In contrast to other species distribution models, MaxEnt not only delineates species distributions but also determines the importance of habitat *characteristics* for the species of interest (Baldwin, 2009). In the light of this research, this aspect is considered important and hence MaxEnt is used in this research to model eland distribution. A more comprehensive explanation of MaxEnt is given below.

3.3 MaxEnt

MaxEnt is a stand-alone modelling package that is commonly used to model species distribution based on presence-only data. It has been used extensively in the field of biogeography, conservation biology and ecology (Elith et al, 2011) and can make accurate predictions from small data sets (<100 presence localities) (Philips et al., 2004).

What it basically does is estimating a target distribution of maximum entropy from the environmental variables (*covariates)* at presence localities, which means that the *estimated distribution* is closest to the *observed distribution*. The measure of maximum entropy is defined by different coefficients that derive from covariate transformations (*features*). By definition, different features put different constraints on the estimated probability (e.g. linear features ensure an equal mean between the observed and predicted distribution of each covariate, whereas quadratic features ensure equal variance). MaxEnt is based on a sequential-update algorithm that changes the weights given to each feature at each iteration, optimizing the estimated distribution accuracy. For a comprehensive explanation of the statistical procedure of MaxEnt, see Phillips & Dudík (2004), Elith (2011) and Phillips et al. (2006).

Chapter 4. Methodology

4.1 Data collection

4.1.1. Eland surveys

Eland herd observations in the Sneeuberg Nature Reserve were carried out by Living Lands from 7th of May until the 30th of June 2015. Data was collected through a (semi-)systematic sampling method of driving fixed routes at fixed hours and documenting eland herd locations. In total, 141 eland herd observations were documented, of which some were observations of the same herds that were documented at a certain time interval. Despite this semi-systematic sampling method, sufficient data is collected that can be used to estimate the species' distribution in MaxEnt.

Figure 4: Method of eland herd location estimation (documented variables are A = GPS location of observation; r = distance (m); and h = radial distance from the north. B = estimated GPS location of the herd)

Herd observations were documented by measuring distance (r) from the point of observation (A) to the herd location (B) with a rangefinder, which is an accurate measurement method according to Ransom & Pinchak (2003). The geographical position of the point of observation A was documented with a GPS. Additionally, the bearing (h), which is the radial distance from the north of point B relative to point A is measured (Figure 4). Subsequently, the GPS locations of the observed herds could be calculated. For each observation, the date and time of observation were recorded, as well as the herd size and herd composition. The herd composition was described in terms of age and sex by counting the amount of young males and females, and old males and females.
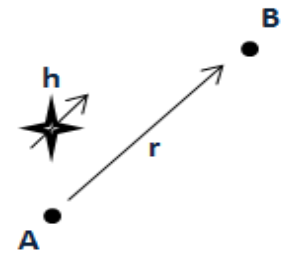
4.1.2 Environmental variables (covariates): collection and pre-processing

The covariates used in this research are environmental factors that appeared (from literature) ecologically relevant to describe habitat suitability. The environmental variables relate to *topography* and *vegetation*, as this has shown to explain eland movements (Melton, 1987). In many studies, topography is represented by a Digital Elevation Model (DEM), representing *altitude* and from which the *slope* and *aspect* can also be derived (Elith & Leathwick, 2009). Vegetation is important in terms of *composition* and *productivity*. Vegetation *composition* is described in a categorical *vegetation map* whereas vegetation *productivity* is be approximated with a *Normalized Difference Vegetation Index* (NDVI). Surface water data was used to spatially indicate Euclidean distance to surface water. At last, *precipitation* was included as it appears to greatly influences herbivore dynamics through differences in vegetation cover, composition and productivity. An overview of the data background can be found in Appendix I: Overview and sources of environmental data. Below, the pre-processing of these variables is described:

A *DEM* was created from contour lines (10 m) of the research area using the 3D Analyst Toolbox in ArcGIS 10.2. The pixel size of the DEM was chosen based on the method proposed by Hengl (2006), which can be approximated from:

$$p = \frac{A}{2 \cdot \sum l} \qquad\qquad\qquad (Eq.\ 2)$$

Where P is the optimal pixel size, A the total size of the study area and $\sum l$ is the total cumulative length of all contours. The resulting pixel size was 21.321194 m. The DEM was projected onto a Projected Coordinate System WGS UTM Zone 35S. Subsequently, all other variables described below have been fit to this projection in ArcGIS 10.2 to ensure equal cell sizes and raster outlines. An *aspect* and *slope* raster file were created from the DEM with ArcGIS 10.2 using the Aspect and Slope function, respectively, in the Spatial Analyst Toolbox. Satellite imagery from the Landsat 8 Satellite was used to calculate the *Normalized Difference Vegetation Index (NDVI)*. A trade-off between image quality on the one hand and temporal consistency of the observation data and environmental variables on the other, resulted in selecting imagery from 31st of March 2015 (which deviates slightly from the period that the eland observations were carried out, but still falls in the dry season). This is important as Phillips et al. (2006) highlights that discrepancy between the time of species observations and environmental data collection violates the ecological assumption of *temporal correspondence* (i.e. data used in modelling must be temporarily conformable) that is implicit for species distribution modelling. To enhance image resolution, panchromatic sharpening of the 30 m resolution image was performed. Subsequently, NDVI was calculated in ArcGIS 10.2 with the Image Analysis NDVI function.

A national vegetation map from Mucina & Rutherford (2011) was used to support the development of a more comprehensive *vegetation map* of the research area. In the next subchapter (

4.1.2. Vegetation data collection) , the pre-analysis, field work and vegetation analysis required to create the map are described in more detail.

Precipitation data from WorldClim represents average monthly precipitation measured from 1950 to 2000. Because the resolution is only 1000 m, the data was interpolated using bilinear interpolation in ArcGIS 10.2. Three different variables were created that represented the average precipitation of the months in which the observations were carried out (May and June) and the yearly average precipitation. A *surface water* map was made from river outlines that were extracted from a national data set of major river systems. Additionally, locations of surface water that were observed and documented when the eland survey was carried out were used to produce a Euclidean distance to surface water raster file.

4.1.2. Vegetation data collection

Although a vegetation map of the Sneeuberg is already available, this research partly aimed to produce a more comprehensive vegetation map of the research area. The current map derived from a national vegetation map of South Africa, Lesotho and Swaziland from Mucina & Rutherford (2011) and only identifies four vegetation types in the Sneeuberg. This subchapter describes how preliminary classification of aerial photographs and Google Earth Imagery in combination with a field-based vegetation sampling survey served as reference data to produce a more comprehensive vegetation map.

Pre-analysis in Google Earth and ArcGIS

The most recent existing vegetation map from Mucina & Rutherford (2011) was used as baseline information. The four vegetation types that are currently identified in this map (Southern Karoo Riviere, Karoo Escarpment Grass, Upper Karoo Hardeveld and Eastern Upper Karoo) were divided into vegetation sub-types by means of visual interpretation of aerial photographs and Google Earth imagery. Sub-types were delineated in ArcGIS 10.2.

Line intercept method for field data collection

To assess the vegetation composition in the distinguished vegetation sub-types, field work was carried out to collect vegetation data. The spatial heterogeneity and patchiness of vegetation in grass- and shrublands such as the Sneeuberg, requires a sampling unit and size that takes this into account (Pilliod and Arkle, 2013). Methods commonly used in grass- and shrub lands are *point-intercepts methods* as they sample along gradients thus lower the risk of capturing only a one-sided fragment of the vegetation type. The *Line Point Intercept* is such a point-intercept method and is an accurate and time-efficient method to sample vegetation along a line segment (Pilliod and Arkle, 2013; Herrick et al. 2005). The line often extends between 30-70 meters and sample points are taken at a regular interval. The data at the sample points is collected by a pin that is dropped to the ground. The type of cover (vegetation, soil or rock) is recorded. The plant species are collected and identified later.

A transect of 50 meters was used, with sample points taken at an interval of 1 meter, resulting in 50 recorded data points. At each data point, the species are documented as horizontal layers if more than one plant hits the pin. The data collection forms used can be found in Appendix II and Appendix III. Three transects were set out in each vegetation subtype. The locations were chosen at random in ArcGIS 10.2 with a 100 m buffer from the edge of the delineation to reduce edge effect. From the transect location points, all transects were directed northwards (see Figure 5 for the transect locations).
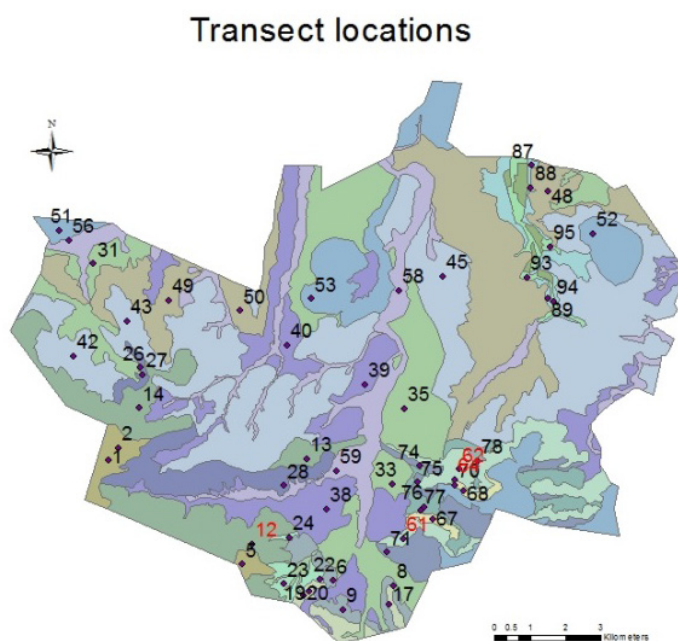
Figure 5 Transect locations for vegetation sampling

Data analysis

The Line Point Intercept method provided data on *ground cover, canopy cover to undergrowth ratio, species composition, species richness* and *species frequency* for each vegetation type. Mean and variances of these variables were calculated as three transects were set out for each subtype. Collected plant species were identified using books (Van Oudtshoorn, 2012), herbaria with plants from a similar region, the plant database of JSTOR Global Plant (JSTOR, 2015), and the help of a botanist at Grootfontein Agricultural Development Institute. Within the initial vegetation types

identified by Mucina & Rutherford (2011), the vegetation sub-types were compared, and if needed, merged.

4.2 Data preparation

4.2.1 Input formatting for species distribution modelling

Firstly, a selection of all eland observation data was made. The interval data was removed from the data set and only the first observation was kept to reduce the effect of responsive movement of herds to the observers (Thomas, et al. 2010). A final presence localities data set of 76 observations was used to run the model.

As pointed out by Paudel et al. (2015), cross-correlated variables create biased outcomes and must be prevented as much as possible. Cross-correlation was tested for in ArcGIS 10.2 based on Pearson's correlation matrix. Variables where correlated when r > 0.7. However, as this coefficient can only describe the relation between *continuous* variables (and aspect and vegetation are categorical data), a measure of association is calibrated using the popular Cramer's V, which is based on Pearsons's chi-squared statistic and returns a measure of association between 0 and 1. This requires the transformation of all continuous variables to categorical data. This dichotomization was done in ArcGIS 10.2 based on the natural break points that are identified based on a maximum variance per class. Cramer's V was calculated in Matlab using the Measure of Effect Size Toolbox.

4.2.2. Sampling bias

To correct for sampling bias, an estimation of the relative sampling effort over space was composed. An accurate estimation can be made with a Gaussian Kernel Density map of the presence localities (Elith et al., 2010) that performed relatively well according to Fourcade et al. (2014). Accordingly, a sampling bias grid was created based on the selected 76 presence localities using an external Species Distribution Model Toolbox (SDM Toolbox) for ArcGIS 10.2 (Brown, 2014).

4.3 MaxEnt model parametrization

The default settings of MaxEnt have been developed based on validation over an extensive dataset of various species distributions in different environmental conditions, which, according to the

developers, justifies the use of the default settings (Philips & Dudík, 2008). Others claim otherwise: unfamiliarity with the modelling method however results in limited adjustment of the model parameters. The default settings are often used where alternative settings might be more appropriate. Merow et al (2013) have identified six key issues relating to parameterization of the model and suggest to test model fit of model outcomes with different settings, acknowledged by Halvorsen (2015). The six key issues relate to 1) different background sampling sizes 2) the use of feature types 3) adjusting the value of the regularization parameter β 4) correcting for sampling bias 5) the use of output types for model analysis and 6) model evaluation methods. For an explanation of these settings, see Appendix VI.

Reflecting on the six key issues mentioned above:

Background sampling sizes were kept in default mode as adjusting the *background sampling settings* is particularly relevant when species distribution is estimated for large areas (e.g. continental scale) or is extrapolated to other areas. Model parameterization has hence been carried out for *feature types, regularization parameter values* and *sampling bias correction*. Subsequently, the *raw output type* was used to compare model performance to minimize arbitrary assumptions made on output data, as recommended by Merow et al. (2013). It is recommended by various research not to use the logistic output (Merow et al., (2013); Halvorson et al., (2015). The cumulative output will only be used in the final model to visualize the geographic extent of suitable habitat.

The *model evaluation* was based on training AUC from K-fold cross validation data where k = 10, as proposed in the MaxEnt tutorial developed by Phillips et al. (2004). Additionally, the response curves of the model will be examined to interpret the importance of explanatory variables (Merow et al., 2013). For all runs the following parameter settings were adjusted to:

- Maximum iterations: set to 1000 rather than default 500 since preliminary test runs showed that the maximum amount of iterations needed to get to the default log loss lies between 500 and 1000.
- K-fold cross validation: with k = 10.

The following subchapters describe the various modelling scenario's that were executed to optimize model performance. A workflow overview with model outcomes is given in the Results (Chapter 5).

4.3.1. Feature types

When more than 80 presence localities are available, MaxEnt uses all feature types. Otherwise, an automatic subset is used. The presence localities used in this research consist of 76 observations, which means that the Product and Threshold Features are omitted by default. Elith et al. (2011) adverts the use of *only hinge features* because they produce smooth fitted functions. Sometimes model simplification by using linear functions is desirable. Threshold features may be relevant when known ecological thresholds exists. Hence the following scenarios are tested. The other test parameters are kept in default mode.

    A. Auto Features (default)

    B. No Threshold Feature Types

    C. Only Linear Feature Types

    D. Only Hinge Features

    E. All Feature Types

4.3.2. Regularization parameter β

Testing for a range of regularization coefficients is recommended to improve model performance (Merow et al., 2013). Generally, model performance is improved with regularization parameter values that decrease when the number of occurrence records in a dataset increases (Philips & Dudík, 2008). The default value of the regularization parameter $\beta = 1$. Accordingly, the model was tested on a higher and lower bound of the default value. The performance of the different parameter settings were compared with the default based on the AUC. The following scenarios are tested:

    A. $\beta = 1$ (default)

    B. $\beta = 0.5$

    C. $\beta = 2$

4.3.3. Sampling bias

MaxEnt support the input of a grid that represents the relative sampling effort. A Gaussion Kernel Density Map with a sampling distance of 500 m was tested under the following scenario's:

A. No sampling bias correction (default)

B. Gaussion Kernel Density sampling bias correction

The model fit of different parameter settings were compared in order to make the most accurate prediction of the species distribution.

4.4 Final model run and evaluation

Based on the outcomes of the test scenarios above, a final model with optimized parameter settings was run. Besides raw output data, cumulative output was returned in order to produce a habitat suitability map that reflects habitat suitability and unsuitability based on a modelled threshold value. Choosing a threshold value requires a (subjective) decision that determines at what predicted probability of presence the location is classified as suitable (Freeman & Moisen, 2008). Depending on the aims of the modelling outcome, an appropriate threshold value must be selected. Outcomes of an extensive comparative study on the performance of various threshold criteria by Freeman & Moisen (2008) recommend to use continuous outcomes rather than binary. Therefore, the raw data that returns the probability of presence per pixel will be used to show the continuous distribution probability as a proxy for habitat suitability. Additionally, since this research aims to show the spatial extent of suitable habitat for eland, a binary map is constructed using the cumulative output. MaxEnt provides a threshold value where the model sensitivity equals the specificity C_{ss} (Merow et al., 2013). This is a commonly used metric that entails the cumulative value '*where positive and negative observations have equal chance of being correctly predicted*' (Freeman & Moisen, 2008). Logistic output is not used as it leans heavily on arbitrary values, as mentioned earlier.

To evaluate the importance of the environmental variables, response curves and jackknife test were returned. Response curves show how the individual covariates influence the model when 1) that particular variable is altered and all other variables are kept constant and 2) when only that particular variable is used to create the model. Jackknife tests are an alternative way of showing variable importance by excluding each variable in turn and creating a model with each variable in isolation.
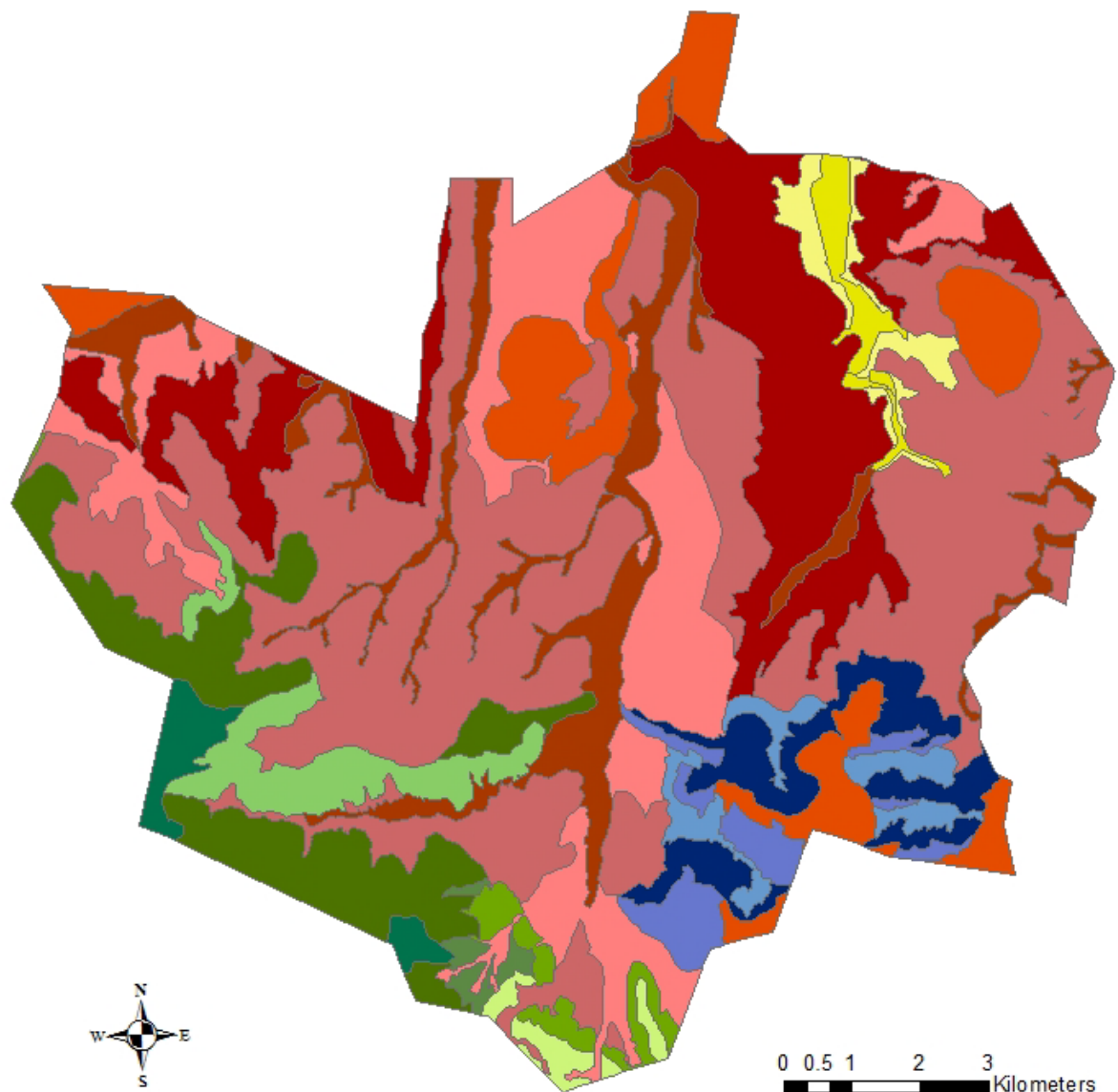
Chapter 5. Results

5.1 Vegetation map

A total of 18 sub-types were initially identified based on visual interpretation of vegetation types from aerial photographs and Google Earth Imagery. A description of the characteristics that were used to distinguish vegetation sub-types can be found in Appendix IV: Vegetation type names and identification characteristics**.** Analysis of the collected vegetation data pointed out that two identified sub-types were very similar in terms of cover and vegetation composition and were therefore merged, resulting in a final map that distinguishes 16 vegetation subtypes. An extensive description and comparison of the subtypes can be found in Appendix V: Vegetation type description.

Figure 6 shows the final vegetation map. The four color families (green, red, yellow and blue) show the outlines of the initial vegetation map from Mucina & Rutherford (2011).

5.2 Data preparation

The outcome of the multicollinearity test to identify cross-correlated variables is shown in Table 1. The Pearson correlation coefficients between all environmental layers are given, although they are trivial for relations with categorical variables (aspect and vegetation). High correlation ($r > 0.9$) exists between all precipitation variables (May, June and Annual Average) and the DEM. This high correlation resulted in the omission of all precipitation variables from the model. The outcome of the measure of association using Cramer's V is given in Table 2. Precipitation variables are not taken into account since they have been omitted already.

Legend

KEG heterogenous riparian (M. disticha)	SKR heterogeneous riparian (Pteronia tricephala)
KEG south sloping hills (M.disticha-Renosterbos)	SKR homogenous grassland (M. disticha)
KEG sparse heterogeneous highlands	UKH high diversity altitudinal plateau
EUK mountain shrubland (Helichrysum rosum)	UKH mountain valleys (Renosterbos)
EUK riparian grasslands (Eragrostis curvula)	UKH northern escarpment (Themeda triandra)
EUK stony grassland (Aristida vestita)	UKH northern grassland (Themeda triandra)
EUK tufted grassland (M. disticha)	UKH southern escarpment (Chrysocoma ciliate)
EUK variable floodplain (Eragrostis lehmannia)	UKH southern grassland (Eragrostis lehmannia)

Figure 6 Vegetation map of the Sneeuberg Nature Reserve (blue colors (KEG) = Karoo Escarpment Grassland; red colors (EUK) = Eastern Upper Karoo; yellow colors (SKR) = Southern Karoo Riviere; green colors (UKH) = Upper Karoo Hardeveld)

Table 1: Results for Pearson's correlation matrix of multiple collinearity tests

Variable	Surface water	DEM	Aspect	NDVI	Precipitation May	Precipitation June	Vegetation	Slope	Precipitation (average)
Surface water	1	0.12258	-0.02451	-0.18218	0.05229	0.07552	0.00718	0.18836	0.0944
DEM		1	-0.0501	0.06784	**0.95763**	**0.94573**	-0.49486	0.52037	**0.95533**
Aspect			1	-0.07368	-0.04097	-0.0373	0.08308	-0.0602	-0.03657
NDVI				1	0.1097	0.10651	-0.11643	-0.03784	0.09544
Precipitation May					1	**0.98231**	-0.46669	0.46657	**0.98865**
Precipitation June						1	-0.43278	0.47254	**0.98774**
Vegetation							1	-0.28393	-0.45414
Slope								1	0.48077
Precipitation (average)									1

Table 2: Results for Cramer's V measure of association

	Aspect	DEM	Hydro	NDVI	Slope	Vegetation
Aspect	1	**0.3267**	**0.3659**	**0.2393**	**0.3744**	**0.2441**
DEM		1	**0.3731**	**0.2645**	**0.2642**	**0.3673**
Hydro			1	**0.3255**	**0.354**	**0.3437**
NDVI				1	**0.2762**	**0.2098**
Slope					1	**0.346**
Vegetation						1

5.3 Model parameterization

5.3.1. First step of model parameterization

To optimize the model its predictive power, parameter settings associated with maximized AUC values for each parameterization factor were selected. Figure 6 shows the first step of the parameterization optimization workflow and corresponding AUC scores. As becomes apparent, the model scenario's using *All Feature Types* and a *β regularization coefficient of 2* scored best (scenario 4.3.1 E and 4.3.2 B). Although the use of a sampling bias grid slightly reduced the AUC, the use of a sampling bias grid is stressed by Phillips (2009). He also acknowledges the possible reduction in AUC but assures it improving the true distribution estimation. A comprehensive explanation of the importance of correcting for sampling bias can be found in Phillips (2009). The

optimized model settings were aggregated in an *optimized model* (shown in the blue bar in Figure 7). The returned AUC score was 0.8301. Because this score was slightly lower than the individual optimized AUC scores for each model scenario, model parameterization was further explored.

5.3.2. Second step of model parameterization

The strong tendency of regularization parameter *reduction* stressed by Philips and Dudík (2008)
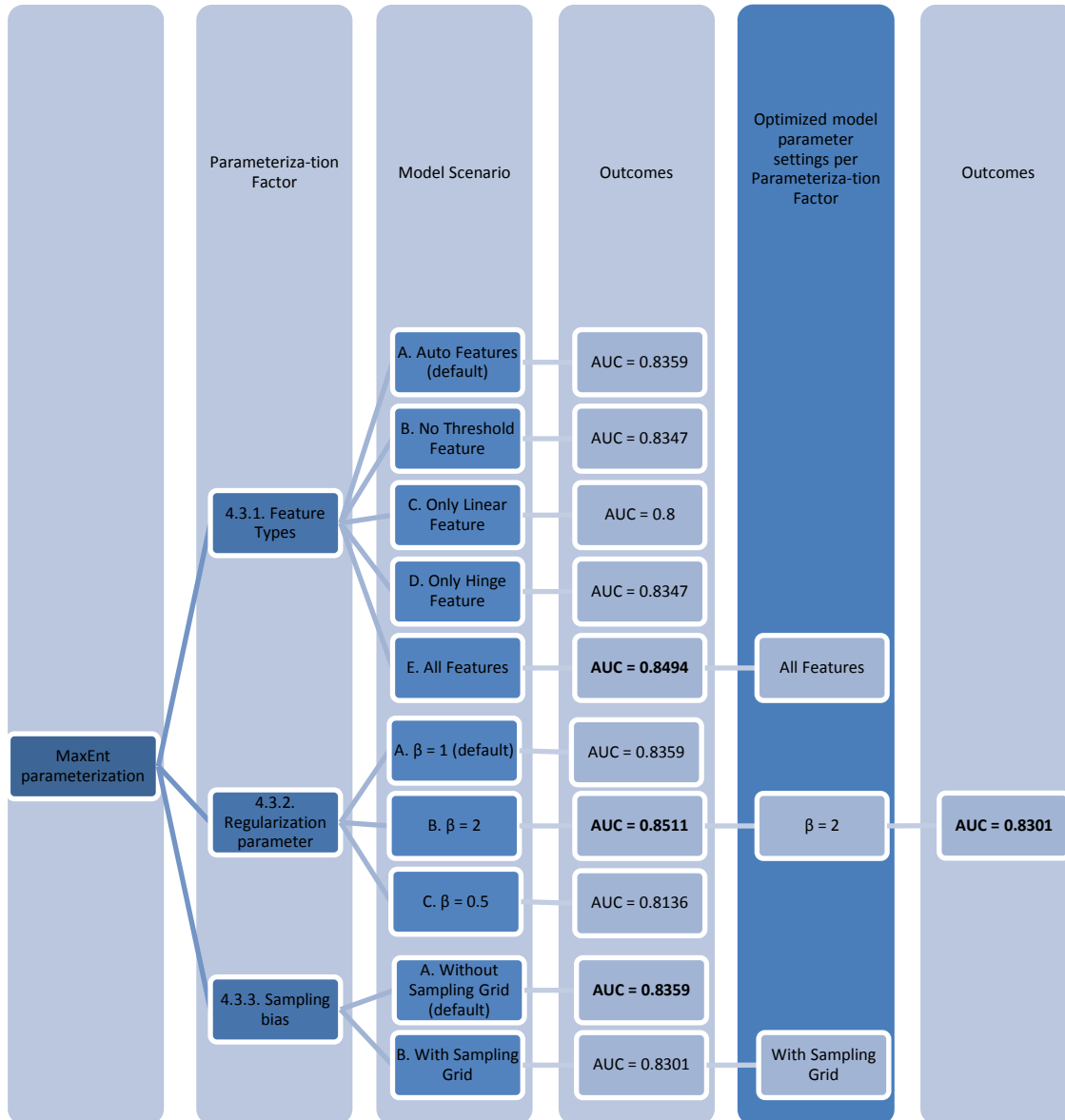

Figure 7 FIRST STEP OF MODEL PARAMETERIZATION. Workflow and resulting test gain AUC values for model parameterization.

brought about testing a lower regularization parameter. Figure 9 shows the AUC scores and the marginal response curves of the second step of parameterization. In line with the postulation of Philips & Dudík (2008), reducing the regularization coefficient improves model performance in terms of test AUC score. When examining the *response curves* of the model scenarios, it becomes apparent that the variable responses associated with lower regularization coefficients (model scenario B and C) are less smooth in comparison to scenario A. As is inherent to the restriction of the confidence interval of the estimated species distribution (i.e. reducing the regularization coefficient), the variable responses become highly non-linear and might pick up more noise rather than showing clear trends.

5.3.3. Third step of model parameterization

When examining the response curves of the first step of model parameterization (response curves not shown here), the '*Only Hinge Feature'* model scenario produces smooth response curves while preserving a relatively high AUC. This observation is supported by experimental tuning of MaxEnt parameters by Philips & Dudík (2008) that showed that the use of hinge features significantly improved model performance. Additionally, they empirically show that the best performing regularization coefficient with Hinge features has a value of $\beta = 0.5$. Similarly, Elith et al. (2011) produces a well performing model using Only Hinge Features. The smooth response curves and high AUC were also observed with the default 'All Feature Types' model scenario. Because this function makes a selection of the features types that explain the species distribution best when the observation data counts < 80 presence localities (with Product and Threshold features to be excluded), the performance of this default setting with a regularization coefficient of $\beta = 0.5$ was tested as well. Modelling outcomes are shown in Figure 8. In terms of the AUC, the Auto Features model scenario performs slightly better that the Only Hinge Feature scenario. From the response curves, no significant difference could be observed.

Based on the AUC scores and response curves, model scenario B from Figure 8 (using Auto Features and a regularization coefficient of $\beta = 0.5$) seems most adequate to model species distribution.
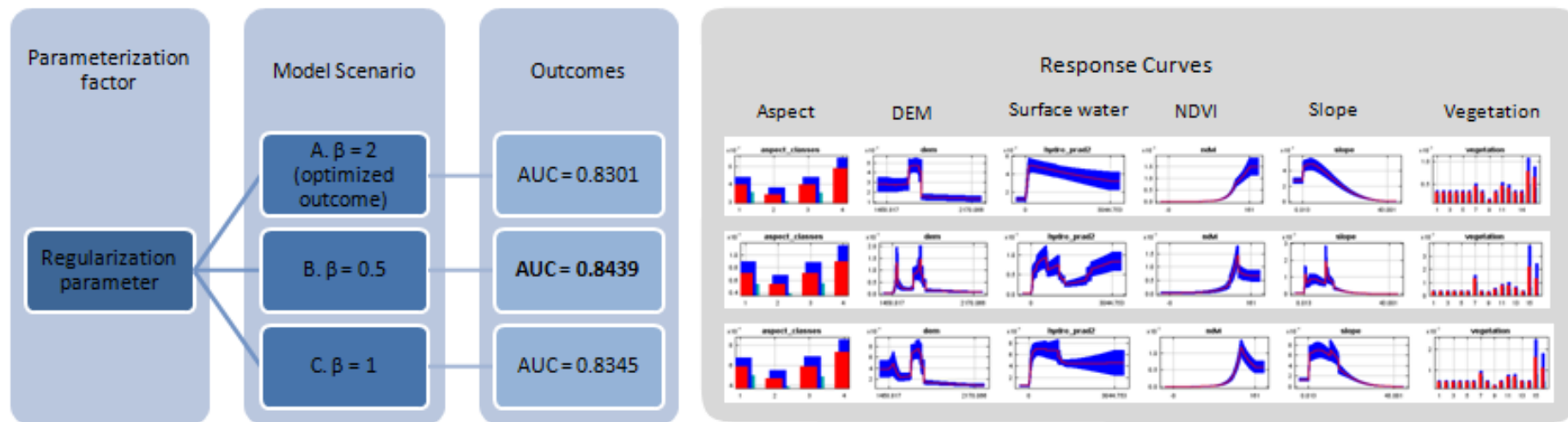
Figure 9: SECOND STEP OF MODEL PARAMETERIZATION. Response curves show how logistic prediction changes as each environmental variable is varied, keeping all other environmental variables at the average sample value (model settings: With sampling bias correction and All Feature Types).
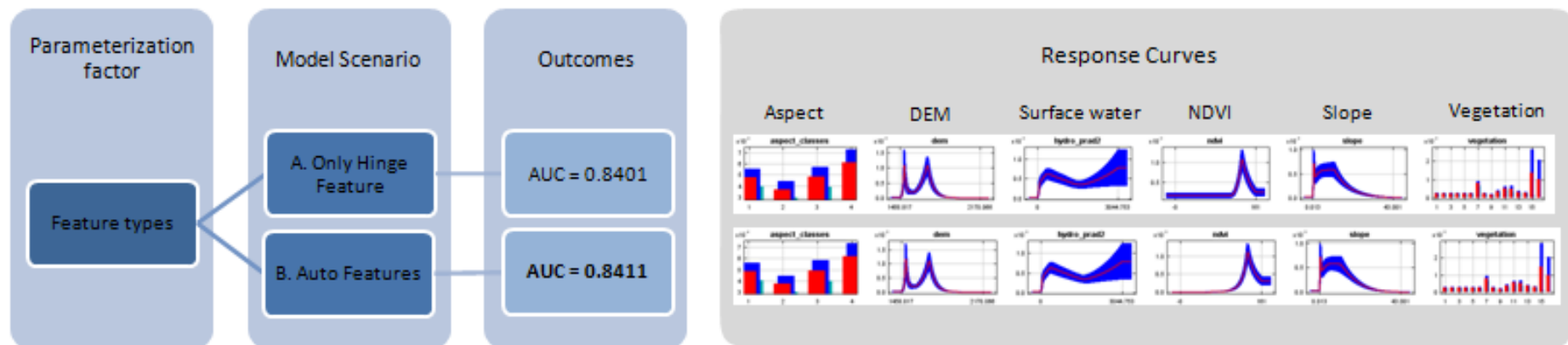
Figure 8 THIRD STEP OF MODEL PARAMETERIZATION. Final comparison of optimized model outcomes (model settings: With sampling bias correct and β =0.5). The Auto Features (default) setting returns the highest test AUC while keeping smooth and interpretable response curves.

5.4 Modelling outcomes

The optimized model as defined above (Auto Feature Types, β= 0.5, with sampling bias correction) was run with 10-fold cross validation. The general responses of the model (AUC and spatial predicted probability of presence) elaborate on the model performance and the spatial distribution of suitable habitat based on the predicted probability of presence, whereas response curves and measures of variable importance describe which and how environmental variables best explain eland habitat suitability.

5.4.1 General responses

The metric of comparison used in model parameterization, Area Under Curve (AUC), derives from the Receiver Operator Curve (ROC) that is averaged over multiple runs and has a mean training $AUC_{T\mu}$ = 0.9077, a mean test $AUC_{t\mu}$ = 0.8411 with a standard deviation of $AUC_{t\sigma}$ = 0.0623. The ROC is shown in Figure 10 where the red line represents the mean test $AUC_{t\mu}$ and the blue area the standard deviation of $AUC_{t\sigma}$. According to Thuiller et al. (2006), models with AUCtµ values between 0.8 and 0.9 AUC are fair.
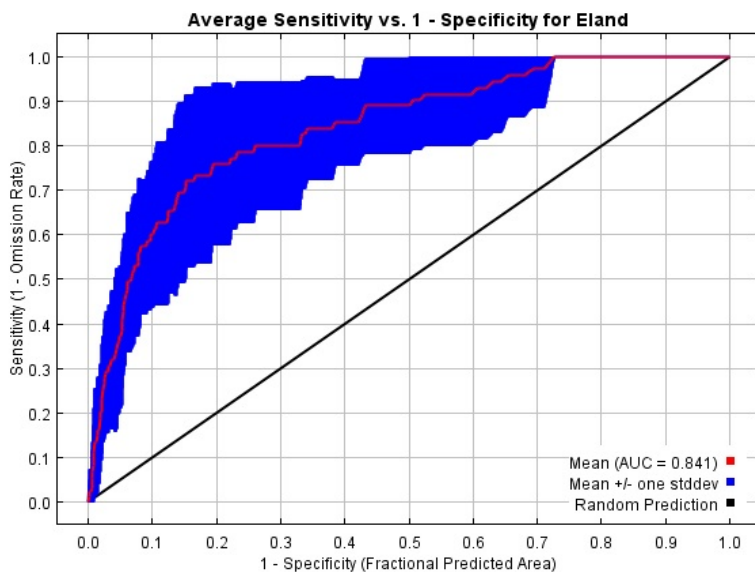
Figure 10: Receiver Operator Curve (ROC), plotting mean average sensitivity versus 1-specificity of 10-old cross validation model Mean test AUC of 0.841 (red line) and standard deviation of 0.0623 (in blue).
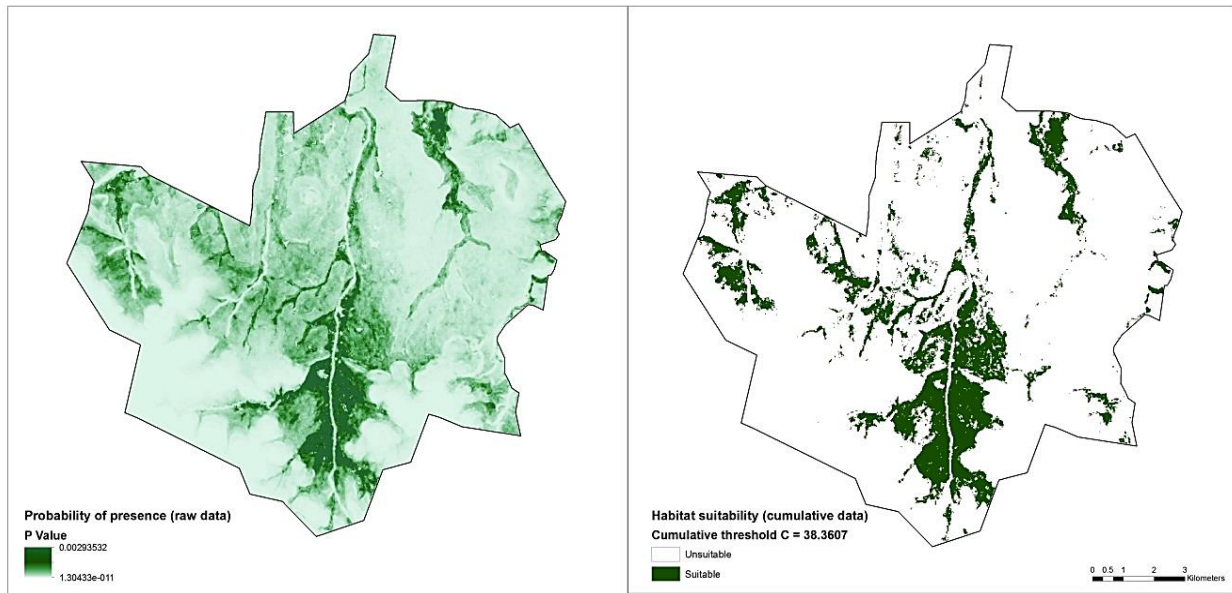
Figure 11: A Continuous probability of presence and B predicted habitat suitability from cumulative data based on threshold C_{ss}

The probability of presence for eland is calculated for each pixel on the raster and shown in Figure 11A. The figure shows the continuous estimated probability that eland is found in that specific location. The threshold value where the model sensitivity equals the specificity amounts C_{ss} = 38.3607. Thresholding suitability on this value for the cumulative output resulted in the map shown in Figure 11B.

5.4.2 Response curves

Response curves show how the individual covariates influence the model when that particular variable is altered and all other variables are set to their average value (*marginal response curves*) and when only that particular variable is used to create the model *(alternative response curves)*.

The average response curve of each covariate is shown in Figure 13, showing the marginal response curves whereas Figure 12 shows the alternative response curves.
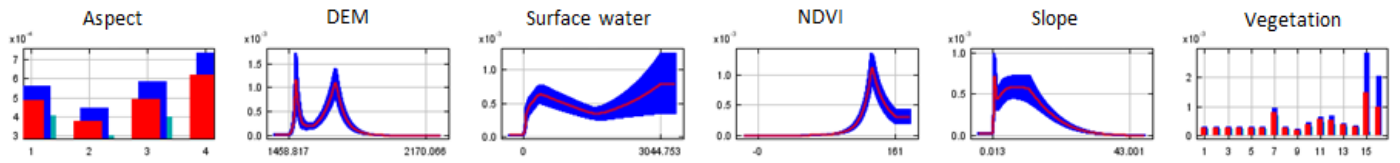
Figure 13: Marginal response curves
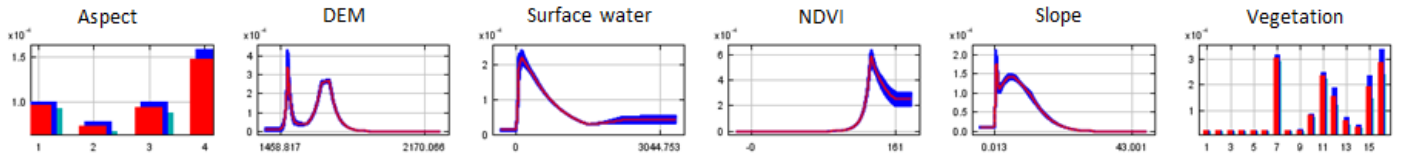
Figure 12: Alternative response curves.

The x-axis shows the covariate values and y-axis the predicted probability of suitable conditions at that value. The red line shows the average response curve of the 10 replicate runs and the standard deviation in blue.

When comparing the marginal and alternative response curves, most covariates show similar trends. Only *surface water* and *vegetation* show a deviation in the predicted presence. Surface water predicted suitability generally decreases with increased distance to surface water when this variable was used independently, while a less linear response was observed in the marginal response curve. The vegetation covariate shows that the predicted suitability of various vegetation types increases relative to other types that score higher in the marginal response curve.

5.4.3. Variable contribution

To identify what variables are most important, MaxEnt keeps track of the covariates that contribute mostly to the model fit and converts their contribution to percentages. Additionally, the permutation importance shows the resulting drop in performance if the variable is randomly permuted. Both model outcomes are given in Table 3.

Variable	Percent contribution	Permutation importance
Vegetation	31.9	12.1
DEM	26.3	49.6
NDVI	25.7	18.6
Slope	7.6	9.8
Surface Water	6.3	8.8
Aspect	2.3	1.1

Vegetation, DEM and *NDVI* show a high contribution to the model fit. Although *vegetation* scores highest in terms of variable contribution, it has a relatively low permutation importance whereas *DEM* shows a very high permutation importance. *Slope* and *surface water* both show a relatively low contribution to the model and permutation importance. *Aspect* scores extremely low on both contributing to the model and permutation importance.

Jackknife tests show the importance of each variable in terms of *regularized training gain, test gain* and *AUC.* The jackknife graphs (Figure 14, Figure 15, Figure 16) show a similar trend, where *Vegetation*, *DEM* and *NDVI* provide the highest scores when the model is ran with only these variables individually. All jackknife results show that the scenario's *without* the *vegetation* variable score very high - for *test gain* and *AUC* even higher than the scenario where all variables are included (red bar). For *test gain*, the scenario with only *vegetation* clearly scores lower than *DEM* and *NDVI*, whereas in the other jackknife tests, these scores are fairly equal.

The variable that decreases the gain most when it is omitted is *DEM*, which according to MaxEnt output indicates that *'it appears to have the most information that isn't present in the other variables'*.
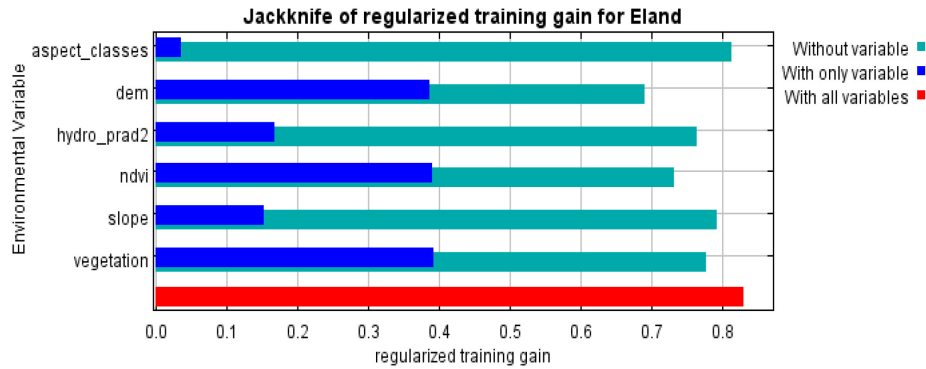
Figure 14 Jackknife results for Regularized training gain
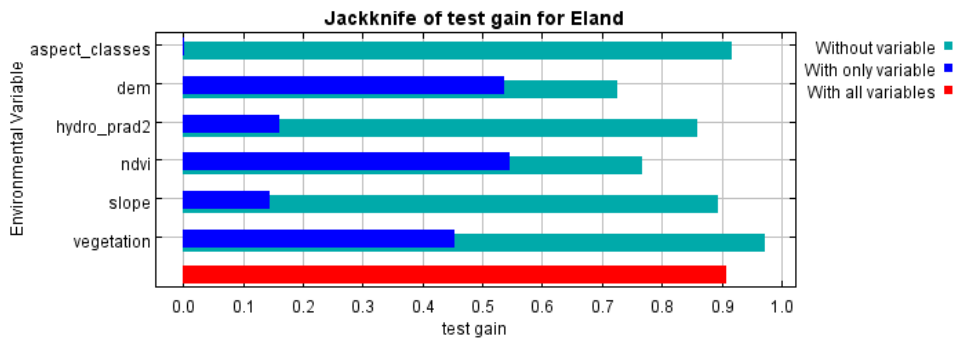
Figure 15 Jackknife results for Test gain
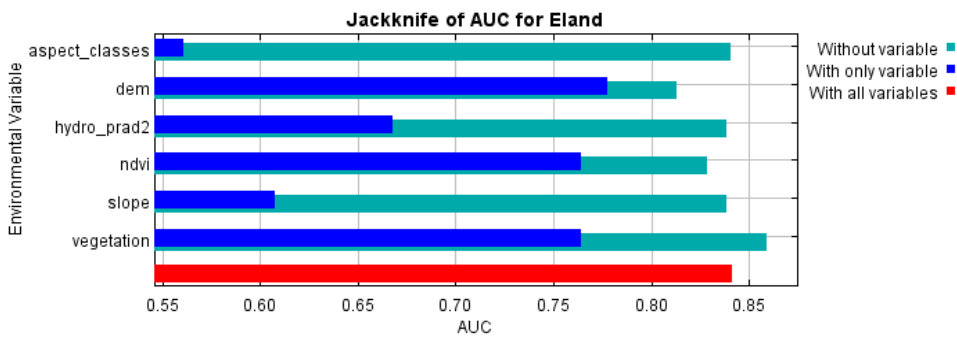
Figure 16 Jackknife results for AUC

Taking the importance of these three variables *(DEM, NDVI* and *Vegetation*) into consideration, the response curves of these variables indicate *how* they influence the estimated probability distribution. Marginal response curves can be misleading when variables are correlated and as Cramer's V does show some association between variables, the alternative response curves are examined to explain how the variable influence the predicted probability of suitable conditions.
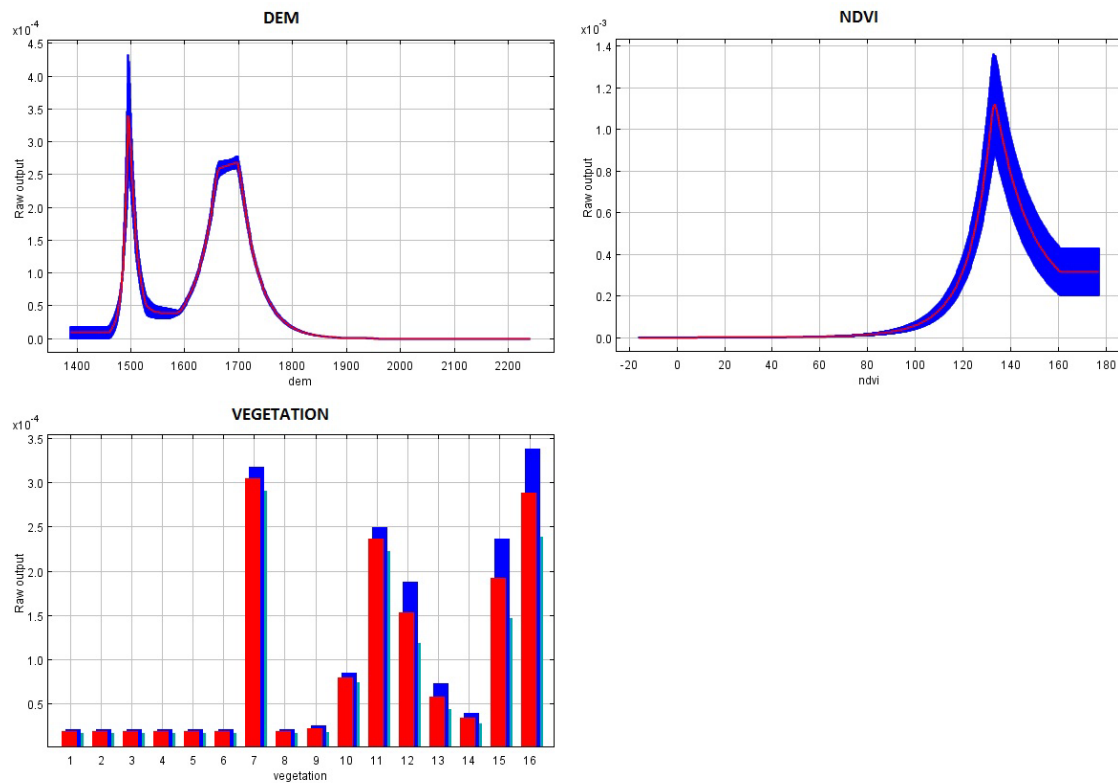
Figure 17 Response curves of eland distribution to elevation (DEM), net primary productivity (NDVI) and vegetation types (upper left corner, upper right corner and lower left corner respectively). Numbers in the Vegetation graph correspond to vegetation types as described in Appendix IV)

DEM shows a bimodal response curve with a clear peak at 1500 m and one between 1600 and 1700 m. Values higher than 1700 show a clear drop in the probability distribution.

NDVI probability distribution shows a high peak at a value of 120. Although NDVI values are scaled between -1 and 1, the NDVI analysis in ArcGIS rescales these values from 0 to 200 to easily render with a colour map. The peak between 120 and 160 is corresponding with NDVI values ranging from 0.2 to 0.6.

The Vegetation Types show high responses for the vegetation 7, 16, 11 and 15, which are Eastern Upper Karoo Riparian Grassland (*Eragrostic curvula),* Southern Karoo Riviere Heterogenous Riparian (*Pteronia tricephala*), Eastern Upper Karoo Variable Floodplain (*Erasgrostis lehmannia*) and Southern Karoo Riviere Homogenous grassland (*Merxmuellera disticha*), respectively. These vegetation types encompass the following characteristics:

Eastern Upper Karoo Riparian Grassland: Heterogenous grasslands surrounding the riverbed with patches of palatable grass species and a high vegetation cover in general. Most dominant grasses are *Eragrostis curvula*, *Tragus koeleroides*, and *Bromus cartharticus*. Woody species and forbs occur in patches and comprise mainly of species in the *Asteracea* family (*Euryops annae* and *Pentzia quinquefida*).

Southern Karoo Riviere Heterogenous Riparian: Riverine vegetation with alternately barren soil and vegetation patches. Mostly dominated by grass *Miscanthus capensis* and woody shrub *Pteronia tricephala*. High variance in species composition was observed with a relative high shrub abundance. Remarkable is the very low *Merxmuellera disticha* abundance in this vegetation type compared to the homogenous riverine grassland of SKR.

Eastern Upper Karoo Variable Floodplain: High variance in community composition. The vegetation cover is generally lower but the vegetation composition is comparable to the EUK Riparian grassland in terms of relative grass and shrub abundance. Grass species mostly dominated by *Eragrostis* species (as in EUK Riparian Grassland) and *Karoochloa* species.

Southern Karoo Riviere Homogenous grassland: Grass dominated plains with high grass canopy cover and homogenous grass cover of mostly *Merxmuellera disticha* and *Aristida vestida* (and some *Eragrostis* species). High diversity of small woody species (small shrubs of the *Asteracea* family).

Chapter 6. Discussion

6.1 Vegetation mapping

The overall low vegetation cover in the semi-arid environment of the Sneeuberg compromised the identification of vegetation types from aerial photographs and Google Earth imagery. Most apparent was the underlying geology rather than distinct vegetation types. This difficulty brought about classification procedure that was mainly focused on distinguishing ecologically distinct areas based on expected differences in vegetation (from geomorphological characteristics of the landscape) and slightly observable differences in patterns (e.g. higher shrub densities).

The 16 subtypes in the final vegetation map were the outcome of a comparison of means and variances from calculated variables (e.g. ground cover, species frequency) for each subtype (see Appendix V: Vegetation type description). In general, the variance of most variables was high, indicating that within subtypes, large difference in vegetation composition, ground cover or species frequencies were observed. Nonetheless, clear difference in variable means and variances between subtypes corresponded to expected patterns of vegetation (e.g. lower mean vegetation cover on escarpments compared to the lower lying valleys and sparse vegetation on south directed slopes).

Furthermore, it must be noted that no statistical analysis could be carried out with the data since transect data was only collected in triplicate. Additionally, Milton & Dean (2006) underline the fact that vegetation cover in the Karoo varies from year to year due to differences in precipitation, causing many plants to disappear in dry periods. Since the data was collected in October, the collected data may, at least partly, be only representable for a restricted period of the year. As annuals ('opslag') often flourish only in a short period of time (generally after the rain) (Milton & Dean, 1996), recorded species may not have been present in the period that eland observations were carried out. However, national rainfall data from the South African Weather Service show that most rainfall from the wet season only arrived in November-December 2015 hence no major difference in species composition and cover is expected relative to the time of eland observations (May-June 2015).

6.2 Importance of environmental variables for eland distribution

One of the sub research questions aimed to shed light on factors that explain the spatial distribution of eland and how. As the variable pre-selection is based on known ecological relevance and the

availability of data, the model is per definition exclusive: there might be more variables that explain distribution.

The variables that influence the model fit mostly were *DEM*, *NDVI* and *vegetation*. These variables and their ecological relevance are further discussed below. Figure 18 shows a spatial exploration of the variable values that corresponded to peaks of the response curves (peaks reflect high predicted eland presence probabilities). The lower bounds of the response curves (reflecting low predicted eland presence probabilities) are not relevant and given in white.

Spatial exploration of response curves
(DEM, NDVI and Vegetation Types)

DEM
altitude (m)

	1,458 - 1,480
	1,480 - 1,520
	1,520 - 1,620
	1,620 - 1,750
	1,750 - 2,171

NDVI
Rescaled values ranging from 0-200

	0 - 120
	120 - 160
	160 - 177

Vegetation Types

	EUK variable floodplain (E. lehmannia)
	EUK riparian grasslands (E. curvula)
	SKR homogenous grassland (M. disticha)
	SKR heterogeneous riparian (Pteronia tricephala)
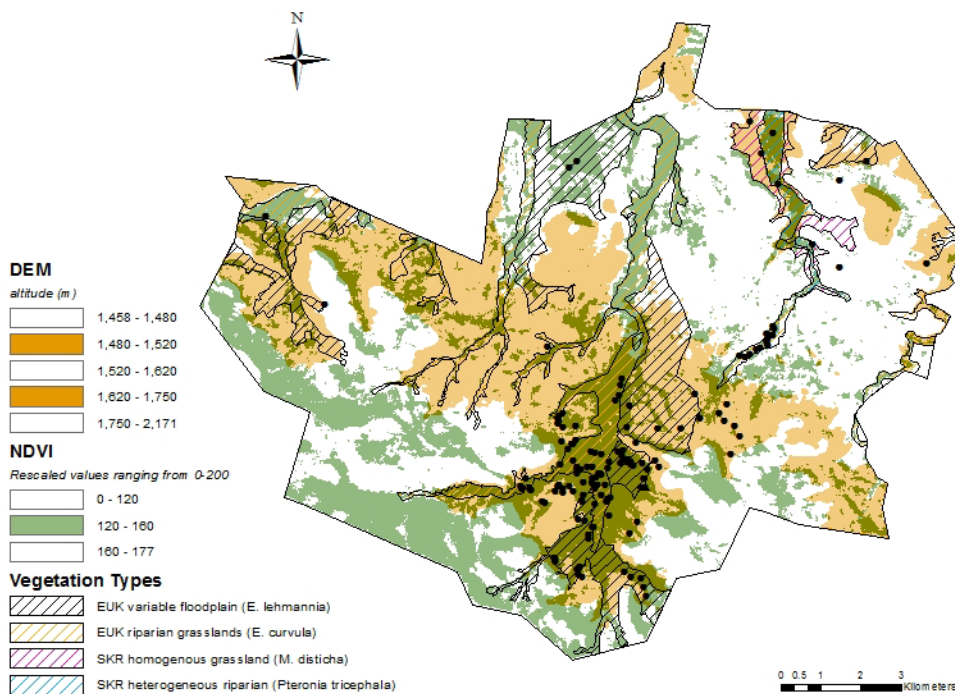
Figure 18 Spatial exploration of the variable values that contribution most to the predicted eland distribution

6.2.1 DEM

The high permutation importance of the DEM shows that differences in *elevation* hold most information regarding species distribution. The relatively high response in areas of lower elevation corresponds to the observation of Watson & Owen-Smith (2000) who observed higher use of low lying valleys in the dry season. Generally, the curve shows that species density generally decreased

with increasing altitude. This decreasing trend appeared to be non-linear however, as the bimodal response curve shows a large drop in variable importance between ± 1520 and 1620 m.

When exploring the extent of this altitudinal range (1520-1620 m) in relation to the eland presence localities (Figure 19), less herds seem to be observed in this region. Accordingly, one explanation of the low predicted presence may be that the model has picked up noise rather than a complex non-linear trend, due to *biased sampling*. Although sampling bias is explicitly corrected for in the model, this might still deviate from *true* sampling effort (as the sampling bias correction is based on merely an *estimation* of the relative sampling effort).

Spatial exploration of drop in DEM response

Legend
Eland
DEM
altitude (m)
1,458 - 1,520
1,520 - 1,620
1,620 - 2,171

Figure 19: Spatial exploration of drop in response between ±1520 and 1620 m (indicated in orange)

Another explanation might be the landscape variability as a result of *latent variables* (underlying variables that are not incorporated in the model but correspond to a variable that is used). Elith & Leathwick (2009) state that associated changes of other variables along elevation gradients are often more important to explain species distributions. An environmental variable that has shown to explain eland distribution in particular is *soil type*, as eland in MNZP avoided shallow leached lithosols (Watson & Owen-Smith, 2000) which is designated to a different clay content and micronutrient availability that in turn influence the *quality* of the overlying vegetation (Mills & Fey, 2005). The alternating geological formations of dolerite, sand- and mudstones and calcareous sediments that are present in the Compassberg Protected Environment lay the foundation for the variety of soil types that can be found (ECPTA, 2011). These differences in soil types might explain the drop observed species density in the area between 1520 and 1620 m. However, comprehensive data on soil types is lacking and more research on that account is required.

6.2.2. NDVI

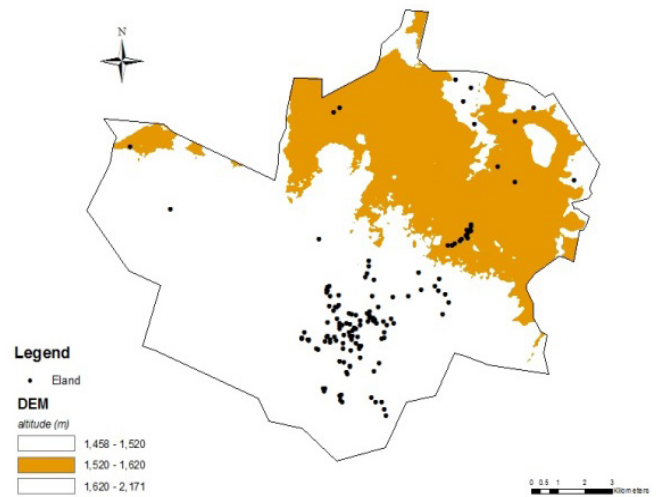The response curve of the NDVI shows that the variable importance increases with NDVI values from 0 to 0.35, which corresponds to the expected positive correlation between NDVI and herbivore densities (Petorelli et al., 2009; Evans et al., 2005).

The observed drop in the response curve (from NDVI values of 0.35 onwards) however, does not match earlier observations of eland responses to plant productivity values. An explanation of this drop might be found in the trade-off that herbivores make between food *quality* and *quantity*. For instance, as grass species mature, the fibre content increases which makes it more difficult to digest. Although these grasses are highly productive (i.e. show a high NDVI value), they have a relatively low value when it comes down to *feeding efficiency*. A species feeding efficiency is determined by the characteristics of the digestive system (Reydon & Hemerik, 2005). When a relatively high amount of energy is needed in order to digest the intake (e.g. mature grasses that are hard to digest), the feed efficiency is low (Pettorelli, 2011). This implies that higher NDVI values do not need to be desirable for herbivores as it may reduce a species feeding efficiency.

Feeding behavior for eland observed by Watson & Owen-Smith (2000) seems to confirm the above for Eland in MZNP: high grass consumption (46% of monthly diet) was observed in December (beginning of the wet season) when grass is young and easy to digest. The grass consumption rapidly declined in January, which the authors assign to eland its browser-like stomach structure that cannot digest the mature grasses as well as grazers. The bulk of eland diet was composed of dwarf shrubs and shrubs that seem to have a relatively low fibre content. The observed peak in NDVI response might thus reflect an optimized plant productivity range where feeding efficiency is highest for eland. In a study on habitat suitability for Mongolian gazelles (*Procapra gutturosa*) on the Eastern Steppes of Mongolia, a similar optimized NDVI ranges is identified (Mueller et al., 2008). The study illustrates that NDVI values within that range provided the best forage quantity as well as quality. Areas with low NDVI areas were avoided due to *limited ingestion rates* (food availability), whereas high NDVI regions were avoided for its *low digestibility*.

When examining the spatial extent of NDVI values corresponding to the drop in response (from 0.35 onwards), aerial photographs show that mostly areas with *trees along drainage lines* reflect these high NDVI values. These trees occur on small patches and are generally planted alien species (mostly poplars). Although foliage of certain tree species is known to be part of eland diet (e.g. *Acacia, Combretum, Commiphora, Diospyros, Grewia, Rhus,* and *Ziziphus*) (Pappas, 2002), the

contribution of poplars to eland diet is has not yet been recorded. Whether the poplars contribute to suitable eland habitat in another way (e.g. by providing shade), the drop in variable importance at least suggests that these areas are likely to be of little value for eland habitat suitability. The optimal plant productivity range for eland hence seems to lie between NDVI values of ± 0.2 and 0.6 (Figure 20).
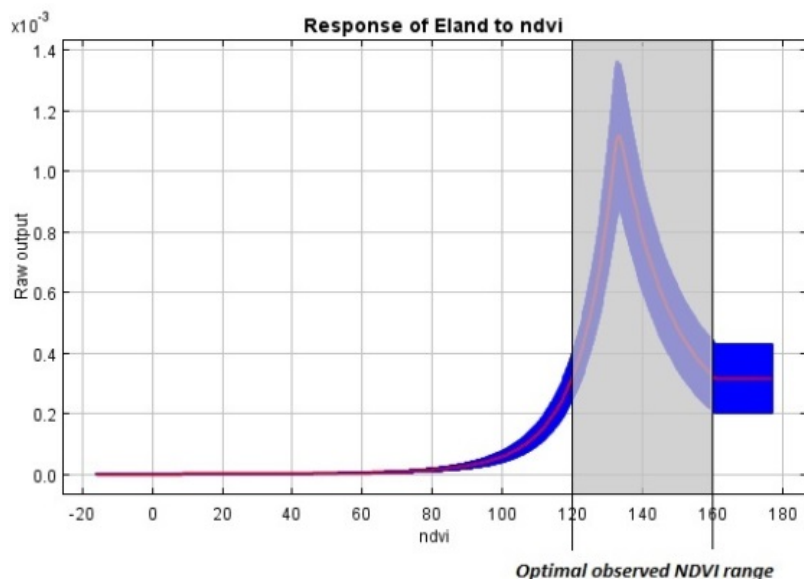
Figure 20: Response curve and estimated optimal plant productivity range (NDVI 0.2 – 0.6) for eland

6.2.3 Vegetation

The four vegetation types observed to be most important (Eastern Upper Karoo Riparian Grassland, Southern Karoo Riviere Heterogenous Riparian, Eastern Upper Karoo Variable Floodplain and Southern Karoo Riviere Homogenous Grassland) seem to be generally heterogeneous grass-dominated riparian zones surrounding drainage lines in the landscape. Although *Merxmuellera disticha* was the most commonly found grass species in the Sneeuberg, the relatively low abundance of *M. disticha* in all these vegetation types (except SKR Homogenous grassland) is remarkable as *M. disticha* is known to be an unpalatable grass species (van Oudtshoorn, 2012). Eland thus seems to select areas with a low abundance of *M. disticha*.

Grass species that are known to be dominant in eland diet include *Setaria* (not found in Sneeuberg), *Themeda* and *Cymbogon plurinodis* (Hillman, 1979 as quoted in Pappas, 2000; Watson & Owen-Smith, 2000). *Themeda thiandra* was generally highly abundant in the Sneeuberg, but occurred mostly on rocky slopes at higher altitudes. *C. plurinodis* was found in lower abundance to *T.*

Thiandra but seem to occur in the same vegetation types and associated altitudinal ranges. The vegetation types that showed most important to predict eland densities did not show a higher availability of *Themeda* or *C. plurinodis*. However, this can likely be explained by eland avoiding the higher altitudinal ranges where these grass species are occurring, as indicated by the drop in the DEM response curve with increasing altitude.

The species composition in the preferred vegetation types seems dissimilar to known eland diet *in terms of grasses*. However, as mentioned earlier, grasses make out a major part of the diet only during a peak in the wet season (46 %), as the overall *annual* diet is made up of 84.4% of *woody plants* for eland in MZNP (Watson & Owen-Smith, 2000). As eland showed dominant browsing behavior in the dry period (Pappas, 2002; Watson & Owen-Smith, 2000), distribution seems to be driven by *woody species* rather than *grass species*. The forb and shrub species known to be consumed mostly belong to the *Asteracea* familiy (Buys, 1990; Fabricius and Mentis, 1990; Kingdon, 1997; Skinner and Smithers, 1990, quoted in Pappas, 2002) and dwarf shrubs such as *Felicia Muricata*, *Helichrysum dregeanum* and *Walafrida geniculata* (Watson & Owen-Smith, 2000). The latter three dwarf shrubs showed to form between 42% and 63% of browse eaten by eland in MZNP.

Considering the collected vegetation data of the Sneeuberg in relation to the preferred dwarf shrubs, *Felicia muricata* was found only once in the whole area (in Southern Karoo Riviere Heterogenous Riparian) and for *Helichrysum* species, not clear pattern could be observed. Species of the *Walafrida* genus were not found in the research area at all. Nonetheless, many woody species of the *Asteraceae* family were observed in the preferred vegetation types (mainly a remarkably high species diversity of *Asteraceae* in the Southern Karoo Riviere Homogenous Grassland).

When comparing the *species composition* of the preferred vegetation types to other vegetation types, no clear pattern that showed higher shrub abundance could be observed from the collected data as the mean shrub abundance in the preferred vegetation types was similar to the mean abundance of all vegetation types. Watson & Owen-smith (2000) suggest that the habitat selection for browsing was mainly determined by the *quality* of woody browse rather than the quantity: as a result of the generally low leaf-to-stem ratio, woody plant species often have high fibre contents. To prevent excessive fibre intake, eland selects woody species that have a relatively low fibre content.

As the exact dietary composition for eland in Sneeuberg is yet unexplored, linkages between dietary requirements and vegetation types are difficult to construct. Nonetheless, the species documented in these preferred vegetation types are may hold valuable information about species-habitat

associations (e.g. for further exploration of eland dietary patterns) and made areas of vegetation importance spatially explicit.

6.3 Spatial distribution of suitable habitat

The habitat suitability map shows that predicted eland occurrence probability is concentrated in areas that surround the drainage lines of the landscape. The binary suitability map spatially indicates suitable and unsuitable habitat by thresholding at a cumulative probability value. Although this threshold is commonly used and is statistically legitimized, it does not imply that this threshold value is ecologically sound (i.e. that it is truly representative for the threshold between suitable and unsuitable habitat for eland in particular). Hence, the binary probability map must be interpreted as a map that spatially indicates core areas that appeared to be suitable rather than giving an absolute quantification of suitable habitat available in the research area.

Additionally, as mentioned earlier, predicted habitat suitability must be understood within a seasonal timeframe. Environmental factors may be influential in a different manner between seasons. Considering that dietary requirements play a major role in species distributions, changes in feeding strategies for eland that has been observed throughout the year (Watson & Owen-Smith, 2000) will most likely cause temporal differences in habitat suitability. Nevertheless, the predicted spatial distribution of habitat suitability most likely reflects eland requirements for the *dry winter season.* Expected changes in predicted habitat suitability over time can be explored with multi-temporal habitat suitability studies, requiring species observation and environmental data over different seasons. For further surveys on the species to be carried out, a systematic sampling method is recommended to reduce sampling bias when modeling species distributions over multiple seasons. The software Distance supports the design, collection and analysis of systematic wildlife sampling surveys. As it is freely available and widely applied, the use of this software is recommended when further sampling surveys are carried out.

6.4 Implications for management

Species distribution models have the potential to fulfill a critical role in conservation decision making, but their utility and applicability remains unclear as Guisan et al. (2013) point out that evidence of supportive use of SDMs in conservation management is scarce hence little guidance exist on how to use SDMs particularly for conservation decision-making. They stress the need to

specifically design models to the decision-making context and preliminary reflect on the decisions need to be taken to match modelling outcomes with identified knowledge gaps (Guisan et al., 2013).

Decision-making for sustainable management of the Sneeuberg Nature Reserve is focused towards eland population management. To explore the extent of the eland problem (i.e. the suggested conflict with cattle and potential risk of overgrazing by overabundance of eland), the relationships between the environment and eland ecology must be understood to see how eland currently utilizes its environment. The species-habitat associations (preferred altitudinal zones, indications of soil importance, NDVI ranges and vegetation types) and spatial indications of suitable habitat that are the outcome of this research indicate how eland uses its environment in the dry winter season. Accordingly, overlap of suitable habitat for eland with habitat used by cattle could be identified and approximations of sustainable stocking rates could be made based on availability of suitable habitat in the area and utilization rates that are considered 'safe and sustainable' (Hunt, 2008, as quoted in Hunt et al., 2014). Identifying these utilization rate however, requires a better understanding of the requirements of eland (i.e. in terms of dietary requirements).

Additionally, as habitat suitability must be understood within a temporal context, estimations of sustainable stocking rates might only be representative for the dry winter season. Hence, further exploration of landscape utilization eland in different seasons (i.e. multi-temporal research on habitat suitability) is recommended to account for possibly variable stocking rates that are considered sustainable throughout the year.

Chapter 7. Conclusion

The reintroduction of eland in the Compassberg Protected Environment appears to have resulted in a *human-wildlife conflict*, where the fast growing eland population has caused a spill-over to other areas and is believed to compete for resources with domestic stock. To explore the extent of this conflict, local NGO Living Lands is carrying out research on eland population sizes, species movements and diet to see how eland competes with cattle in the area and how conflict could be mitigated while simultaneously conserving sustainable wildlife populations. Population sizes and species distributions over the landscape must be taken into careful consideration for sustainable ecosystem management as implied by historical overstocking and resulting degradation of the landscape. Understanding the mechanisms of overstocking and subsequent overgrazing requires spatial and temporal indications of the interaction of herbivores with the landscape (Coughenour, 1991). A first step to understand how eland interacts with its landscape, is to indicate how it utilizes his landscape. By mapping species habitat suitability, one gains insight in the spatial distribution of habitat utilization by eland and environmental factors that explain this distribution. Species distribution models that are based on observations of species provide a powerful tool to estimate habitat suitability. To predict habitat suitability and identify species-habitat associations for eland in the Sneeuberg Nature reserve, the commonly used species distribution model MaxEnt was used in this research.

The model provided accurate estimations based on the mean test $AUC_{t\mu} = 0.8411$ with a 10-fold cross validated model. The environmental variables that showed to be most important to estimate eland spatial distribution altitude (DEM), net primary productivity (NDVI) and vegetation types.
The response curve of the DEM showed that eland observations generally decreased when altitude increased. Although the bimodal response shows a drop between approximately 1520 and 1620 m, it is suggested that either noise is picked up from biased sampling or differences in geological formations and hence soil type might play a role, as it influences micronutrient availability in plants hence the quality of the overlying vegetation (Mills & Fey, 2005). However, no spatial data on soil types or underlying geology was available to validate this hypothesis. NDVI values showed a positive relation with predicted eland densities, with a steep increase from 0 onwards (when net primary production is measured). After 0.35 the response curve drops again indicating that very high NDVI values are not desirable and an optimal plant productivity range seems to occur between NDVI values of approximately 0.2 and 0.6. A detailed vegetation map was made based on the already existing vegetation map of Mucina & Rutherford (2011). In total, 16 vegetation subtypes

were identified. The model showed that eland preferred vegetation types that were located closely to drainage lines in the landscape. Generally, these types seemed less dominated by the commonly found *Merxmuellera disticha*. In terms of grass species, the preferred vegetation types showed a dissimilarity with grasses that are known to be part of eland dietary composition. However, as grasses showed only to be a marginal part of eland diet in the nearby Mountain Zebra National Park, woody species composition is likely to be more important. Although no clear pattern was observed in relation to woody species abundance in these vegetation types that were known to be part of eland diet, many woody species of *Asteraceae* family were found in the preferred vegetation types. As suggested by Watson and Owen-Smith (2000), the quality of woody browse rather than the quantity is important for eland distributions. As dietary patterns and hence qualitative value of plant species for eland diet are yet unexplored for eland in Sneeuberg, no links between diet requirements and vegetation types could be made. Nonetheless, documented species in these preferred vegetation types may hold valuable information about species-habitat associations and made areas of vegetation importance spatially explicit. Based on these observed species-habitat associations, it is recommended that further research focuses on 1) the role of geology and soil types on species distribution and 2) quality of vegetation for eland (i.e. eland diet requirements).

Although predicted habitat suitability should be interpreted with caution, outcomes of this research allow for further clarification of the eland problem (i.e. the suggested conflict with cattle and potential risk of overgrazing by overabundance) as it has made eland habitat spatially explicit and indicated which and how environmental factors explain the species' distribution in the dry winter season.. Accordingly, this information can support decision making on sustainable stocking rates of eland. However, as these estimations of sustainable stocking rates might only be representative for the dry winter season, further exploration of landscape utilization by eland in different seasons (i.e. multi-temporal research on habitat suitability) is recommended to account for possibly variable stocking rates that are considered sustainable throughout the year.

Acknowledgements

First of all, I want to thank Saskia Keesstra for stepping out of her research-expertise-comfort-zone and guiding me through an adventurous project that was theoretically speaking new to both of us. Secondly, I am sincerely grateful to Ignas Heitkonig, who was willing to step in just under the wire and provide substantive support on the modelling section that has been truly priceless.

From the other side of the world, I want to thank Justin Gird for his support and complete devotion, in- and outside of office hours, and Colin Tucker for bearing two weeks of vegetation sampling in fit-mode.

At last but not least, I want to express my gratitude to my family and herd of friends who have inspired, supported and guided me, from moments of defeat to pure joy.

References

Aarts,G., MacKenzie, M., McConnell, B., Fedak, M., Matthiopoulos, J. (2008) Estimating Space-use and habitat preference from wildlife telemetry data. Ecography, 31, pages 140-160

Bailey, D.W., Gross, J.E, Laca, E.A., Rittenhouse, L.R., Coughenour, M.B., Swift, D.M, and Sims, P.L. (1996) Mechanisms that result in large herbivore grazing distribution patterns. *Journal of Rangeland Management, 49,* pages 386-400.

Baldwin, R.A. (2009) Use of Maximum Entropy in Wildlife Research. *Entropy, 11,* pages 854-866

Brown, J.L. (2014) SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and
species distribution model analyses. Methods in Ecology and Evolution

Buys, D. (1990) Food selection by eland in the western Transvaal. South African Journal of Wildlife Research, 20, 1, pages 16-20.

Cassini, M.H. (2011) Ecological principles of species distribution models: the habitat matching rule. *Journal of biogeography, 38,* pp 2057-2065.

Clark, V.R. (2010) The Phytogeography of the Sneeuberg, Nuweveldberge and Roggeveldberge (Great Escarpment): Assessing Migration Routes and Endemism. PhD Thesis for Doctor of Philosophy with Rhodes University, 2010.

Clark, V.R., Barker, N.P. and Mucina, L. (2009) The Sneeuberg: A new centre of floristic endemism on the Great Escarpment, South Africa. *South African Journal of Botany, 75,* pages 196-238

Cousins, J. A., J. P. Sadler, and J. Evans (2008) Exploring the role of private wildlife ranching as a conservation tool in South Africa: stakeholder perspectives. *Ecology and Society* 13(2): 43

Coughenour, M.B. (1991) Spatial components of plant-herbivore interactions in pastoral, ranching and native ungulate ecosystems. *Journal of Rangeland Management, Vol. 44,* 6, pages 530 - 541

Du Toit, G., Snyman, H.A. and Malan, P.J. (2011) Physical Impact of Sheep Grazing on arid Karoo subshrub/grass rangeland, South Africa. *South African Journal of Animal Science, Vol. 41,* 3, pages 280-288

ECPTA (2011) Management Plan for the Compassberg Protected Environment 2013-2018. ECBiodiversity Stewardship: Program of the Eastern Cape Parks & Tourism Agency.

Elith, J. & Leathwick, J. (2009) The Contribution of Species Distribution Modelling to Conservation Prioritization. Spatial Conservation Prioritization: Quantitative Methods & Computational Tools. Chapter 6. (ed. by A. Moilanen, K.A. Wilson & H.P. Possingham). Oxford University Press.

Elith, J., Kearney, M., Philips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution, 1*, pages 330-342

Elith, J., Phillips, S.J., Hastie, T., Dudík, Miroslav, En Chee, Y. and Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. Diversity and Distributions, 17, pp. 43-57

Evans, K.L., Warren, P.H., Gaston, K.J. (2005) Species-energy relationships at the macro-ecological scale: a review of the mechanisms. *Biol. Rev. Camb. Philos. Soc.*

Fourcade, Y., Engler, J.O., Rodder, D and Secondi, J. (2014) Mapping Species Distribution with MAXENT Using a Geographically Biased Sample of Presence Data: a Performance Assessment of Methods for correcting Sampling Bias. PLOS ONE, Vol. 19, 5: e97122. doi:10.1371/journal.pone.0097122

Fretwell, S.D. and Lucas, H.L. (1970) On Territorial Behavior and Other Factors Influencing Habitat Distribution in Birds. Theoretical development. *Acta Bioth. 19,* pages 16-36

Freeman, E. A. and Moisen, G.G. (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling, 217*, pages 48-58

Gallo, J.A., Pasquini, L., Reyers, B. And Cowling, R.M. The role of private conservation areas in biodiversity representation and target achievement within the Little Karoo region, South Africa. *Biological Conservation, Vol. 142,* pages 446-454

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

Guisan, A., Tingley, R., Baumgartner, J.B., Najokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Broton, L., McDonald-Madden, E., Mantyja-Pringle, C., Martin, T.G., Rhodes, J.R., Maggina, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S. Kearney, M.R. and Possingham, H.P and Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters, 16*, pages 1424-1435.

Halvorsen, R., Mazzoni, S., Bryn, A., Bakkestuen, V. (2015) Opportunities for improved distribution modelling via a strict maximum likelihood interpretation of Maxent. *Ecography, 38,* pages 172-183

Hengl, T. (2006) Finding the right pixel size. Computers and Geosciences, 32, 9 pages 1283-1298.

Herrick (2005) Monitoring Manual for Grasslands. Retrieved from: http://www.fwspubs.org/doi/suppl/10.3996/052014-JFWM-039/suppl_file/052014-jfwm-039r1-s04.pdf

Hirzel, A.H., Hausser, J. & Perrin, N. (2001) Biomapper 1.0, Lausanne Laboratory for Conservation Biology. URL: www2.unil.ch/biomapper

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,* pages 2027-2036

Hirzel, A.H. and Le Lay, G. (2008) Habitat suitability and niche theory. *Journal of Applied Ecology, 45,* pages 1372 – 1391

IUCN SSC Antelope Specialist Group (2008). *Tragelaphus oryx*. The IUCN Red List of Threatened Species. Version 2014.3. <*www.iucnredlist.org*>. Downloaded on 10th of April 2015.

JSTOR (2015) JSTOR Global Plants online plant specimen database. Retrieved from https://plants.jstor.org/

Keay-Bright, J. and Boardman, J. (2007) The influence of land management on soil erosion in the Sneeuberg Mountains, Central Karoo, South Africa. *Land Degradation & Development, Vol. 18*, 4, pages 423-439

Keay-Bright, J and Boardman, J. (2009). Evidence from field-based studies of rates of soil erosion on degraded land in the central Karoo, South Africa, *Geomorphology, Vol. 103,* 3, pages 455-465

Kessler, J.J. Usefulness of the human carrying capacity concept in assessing ecological sustainability of land-use in semi-arid regions. *Agriculture, Ecosystems and Environment, 48*, pages 273-284

Lindsey, P.A., Roulet, P.A. and Roman, S.S. Economic and conservation significance of the trophy hunting industry in sub-Saharan Africa. *Biological Conservation, 134,* pages 455-469.

McLeod, S.R. (1997) Is the Concept of Carrying Capacity Useful in Variable Environments? Oikos Vol. 79, No. 3, pp. 529-542

Melton, D.A. (1987). Habitat selection and resource scarcity. *South. African Journal of Sc*ience, *83,* pages 646-651.

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, pp. 1058-1069.

Mills, A, J. and Fey, M,V. (2005) Interactive response of herbivores, soils and vegetation to annual burning in a South African savanna. *Austral Ecology, Vol. 30,* 4, pages 435-444.

Milton, S.J. and Dean, R.J. (1996) *Karoo Veld: ecology and management*. Agricultural Research Centre. Eds. Ellis, R. Published by ARC-Range and Forage Institute, South Africa (1996).

Mucina, L. & Rutherford, M.C. (2011) The vegetation of South Africa, Lesotho and Swaziland. Strelitzia 19. South African National Biodiversity Institute, Pretoria. ISBN 13: 978 1 919976 21 1.

Mueller, T., Olson, K.A., Fuller, T.K., Schaller, G.B., Murray, M.G. and Leimgruber, P. (2008) In search of forage: predicting dynamic habitats of Mongolian gazelles using satellite-based estimates of vegetation productivity. *Journal of Applied Ecology, 45*, pages 649-658.

Pappas, L.A. (2002) Taurotragus oryx. *Mammalian Species, 689,* pages 1-5.

Paudel, P.K , Hais, M., and Kindlmann, P. (2015) Habitat suitability models of mountain ungulates: identfying potential areas for conservation

Pettorelli, N., Bro-Jørgensen, J., Durant, S.M., Blackburn, T., Carbone, C. (2009) Energy Availability and Density Estimates in African Ungulates. *The American Naturalist, 173*, pages 698-704

Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jędrzejewska, B., Lima, M. & Kausrud, K. (2011) The Normalized Difference Vegetation Index (NDVI) : unforeseen successes in animal ecology. *Climate Research, 46*, pages 15-27.

Phillips, S.J, Dudík, M. and Schapire, R.E. (2004) A Maximum Entropy Approach to Species Distribution Modelling. *Proceedings of the Twenty-First International Conference on Machine Learning*. Pages 655-662

Phillips S, Anderson R, and Schapire R. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190*, pages 231–259.

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

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

Pilliod, D.S. And Arkle, R.S. (2013) Performace of Quantitative Vegetation Sampling Methods Across Gradients of Cover in Great Basin Plant Communities. *Rangeland Ecology & Management, Vol. 66*, 6, pages 634-647.

Porter, W.F. and Church, K.E. (1987) Effects of Environmental Pattern on Habitat Preference Analysis. *The Journal of Wildlife Management, Vol. 51, pages 681-685.*

Ransom, D. Jr., and Pinchak., W.E. (2003) Assessing accuracy of a Laser Rangefinder in Estimating Grassland Bird Density. Wildlife Society Bulletin, Vol. 31, No.2, pages 460-463.

Redfern, J.V., Grant, R., Biggs, H. and Getz, W.M. (2003) Surface-water constraints on herbivore foraging in the Kruger National Park, South Africa. *Ecology, 84,* 8, pages 2092-2107.

Reydon, T.A.C. & Hemerik, L. (2005) Current themes in Theoretical Biology: a Dutch Perspective. Springer.

Smit, I.P.J., Grant C.C. & Devereux B.J. (2007) Do artificial waterholes influence the way herbivores use the landscape? Herbivore distribution patterns around rivers and artificial surface water sources in a large African savanna park. *Biological Conservation, 136*, pages 85-99.

Smit, I.P.J. (2011) Resources driving landscape-scale distribution patterns of grazers in an African savanna. *Ecography, 34*, pages 67-74.

Steven de Bie (2015) Personal communication 03-06-2015.

Thomas, L, Buckland, S.T., Rexstad, E.A., Laakes, J.L., Stringberg, S. Hedley, S.L., Bishop, J.R.B, Marques, T.A. and Burnhams, K.P (2010) Distance Sampling: design and analysis of distance sampling surveys for estimating population size.

Thuiller W, Broennimann O, Hughes G, Alkemade JRM, Midgley GF, Corsi F (2006) Vulnerability of African mammals to anthropogenic climate change under conservative land transformation assumptions. *Global Change Biology, 12*, pages 424–440.

Van Oudtshoorn, F. (2012) Guide to grasses of Southern Africa. Briza Publications; Third Edition edition (March 15, 2012).

Verlinden, A. and Masogo, R. Satellite remote sensing of habitat suitability for ungulates and ostrich in the Kalahari of Botswana, *Journal of Arid Environments, 35*, pages 563-574

Vetter, S. (2005) Rangelands at equilibrium and non-equilibrium: recent developments in the debate. *Journal of Arid Environments, 62,* pages 321-341.

Walker, B.H. (1970) An Evaluation of Eight Methods of Botanical Analysis on Grasslands in Rhodesia. *Journal of Applied Ecology, Vol. 7*, 3 , pp. 403-416.

Watson, L.H and Owen-Smith, N. (2000) Diet composition and habitat selection of eland in semi-arid shrubland. *African Journal of Ecology*, *38*, pages 130-137.

Whittaker, R.H, Levin, S.A., and Root, R.B. (1973) Niche, habitat and ecotope. *The American Naturalist,Vol. 107,* 955, pages 321-338.

Appendix I: Overview and sources of environmental data

Environmental variable	Format and resolution	Source	Notes
Digital Elevation Model (DEM)	GRID 21.32199m	Contour lines (Living Lands Database)	
Aspect	GRID 21.32199m	Contour lines (Living Lands Database)	*Created from DEM with ArcGIS 10.2* *Cell values of 1 represent north facing, 2 east facing, 3 south facing and 4 west facing slopes.*
Slope	GRID 21.32199m	Contour lines (Living Lands Database)	*Created from DEM with ArcGIS 10.2* Cell values reflect the continuous inclination in degrees.
NDVI	GRID 15m	Landsat 8 satellite imagery downloaded from http://glovis.usgs.gov Date = 31/03/2015 Cloud cover = 0% Quality = 9	*Initial resolution is 30m but panchromatic sharpening is applied to enhance imagery to 15m resolution*
Vegetation type	Shapefile	• Biodiversity GIS SANBI http://bgis.sanbi.org/vegmap/project.asp • Fieldwork (October 2015)	*Vegetation map from SANBI (Mucina & Rutherford, 2011) used as basis to create a more comprehensive vegetation map*
Precipitation	GRID 1000m	WorldClim Global Climate data http://www.worldclim.org/	
Surface water	Shapefile	• National Hydrology (Living Lands Database) • Eland Survey KML file (from fieldwork May and June 2015)	

Appendix II: Transect form for vegetation survey

Transect	

"date _vegetation type_vegetation subtype_transect number" (e.g. 0710_SKR_Agriculture_8)

Vegetation sampling pilot Sneeuberg 9/10/11 October 2015

TRANSECT FORM

P	Top canopy	LC 1	LC 2	LC3	Soil Surface		Top canopy	LC 1	LC 2	LC 3	Soil Surface
1						26					
2						27					
3						28					
4						29					
5						30					
6						31					
7						32					
8						33					
9						34					
10						35					
11						36					
12						37					
13						38					
14						39					
15						40					
16						41					
17						42					
18						43					
19						44					
20						45					
21						46					
22						47					
23						48					
24						49					
25						50					

CODE DESCRIPTION:

Top canopy (**NONE**, Grass/Shrub/Forb/Tree_NR)

Lower canopy (Grass/Shrub/Forb/Tree_NR, **L** (herbaceous litter), **W** (woody litter)

Soil surface (**R** (rock), **BR** (bedrock), **EL** (embedded litter), **M** (moss), **LC** (lichen crust), **S** (unprotected soil)

Appendix III: Field checklist for vegetation survey

Transect	

"date _vegetation type_vegetation subtype_transect number" (e.g. 0710_SKR_Agriculture_8)

FIELD CHECKLIST

- Start with the FIELD CHECKLIST at GPS location *('point 0')*

Vegetation type characteristics	Explanation		Notes
Mean vegetation height *(cm)*	Estimation of mean height		
Grass/shrub dominated *(grass/shrub)*	Estimation of grass/shrub dominated vegetation type		
Trees (yes/no)	Indicate presence of trees, the species and amount if present		
Interspersion of shrub (random/systematic/contagious)	The pattern in which various habitat features are arranged.		
Interspersion of grass (random/systematic/contagious)	"		

Photographs	

Appendix IV: Vegetation type names and identification characteristics

Describes the characteristics that could be distinguished from aerial photographs and Google Earth Imagery and on which the identification and delineation of vegetation subtypes is based. The names given to the initial vegetation subtypes reflect quick and dirty interpretations of the identified differences and hold no relevant information.

Vegetation type (Mucina & Rutherford, 2011)	Vegetation subtype (initial name)	Description of identification characteristics (from Google Earth and aerial photographs)	Vegetation subtype (new name)	Number used in MaxEnt
Eastern Upper Karoo	Floodplain 1	Flat floodplain surrounding riparian areas, less homogenous surface compared to floodplain 2 (mixture of grasses/shrubs)	EUK variable floodplain (*Eragrostis lehmannia*)	11
	Floodplain 2	Floodplain on slope area, with a more homogenous soil cover compared to floodplain 1 (more grass dominated?)	EUK tufted grassland (*Merxmuellera disticha*)	10
	Grasslands	Flat homogenous plains (looks like floodplain 2)	EUK tufted grassland (*Merxmuellera disticha*)	10
	Mountainous Floodplain	Surrounding riparian areas with relief, alternately grass and shrub areas	EUK stony grassland (*Aristida vestita*)	9
	Mountainous Shrubland	Higher areas, steep slope, heterogeneous thick shrub cover	EUK mountain shrubland (*Helichrysum rosum*)	8
	Riverine	Riparian areas	EUK riparian grassland eragrostis curvula	7
Karoo Escarpment Grassland	Riverine Floodplain	Flat areas surrounding riverine vegetation (riverine vegetation included since very sparse). Dominated by small shrubs	KEG heterogenous riparian (*Merxmuellera disticha*)	12
	S Hills (Shrubs)	South directed hills with a dense vegetation cover (shrubs?)	KEG south sloping hills (*M.disticha-Renosterbos*)	13
	Shrubby N Hills	North directed hills with heterogenous cover of shrubs. Underlying geology clearly visible	KEG sparse cover heterogeneous highlands	14
	Upper Shrubs	Higher altitude flatter areas with heterogenous vegetation cover	KEG sparse cover heterogeneous highlands	14
Southern Karoo Riviere	Floodplain	Flat homogenous vegetation cover	SKR homogenous grassland (*Merxmuellera disticha*)	15
	Riverine	Heterogenous vegetation dominated by shrubs along the stream	SKR heterogeneous riparian (*Pteronia tricephala*)	16
Upper Karoo	High Plateau	Flat area at high altitude with	UKH high diversity	1

Hardeveld		homogenous buy quite sparse vegetation cover (underlying geology of escarpment visible)	altitudinal plateau	
	N Hills	North directed slope at lower altitude with shrub dominated vegetation cover	UKH northern grassland grassland (*Themeda triandra*)	2
	N Mountains	North directed steep slope at high altitude underlying geology visible, sparse vegetation	UKH northern escarpment (*Themeda triandra*)	3
	S Hills	South directed slope at lower altitude with more homogenous vegetation cover compared to N Hills	UKH southern grassland (*Eragrostis lehmannia*)	4
	S Mountains	South directed steep slope at high altitude, underlying geology less visible and denser vegetation compared to N mountains	UKH southern escarpment (*Chrysocoma ciliate*)	5
	Shrubby Floodplain Hills	Dense vegetation cover in lower lying mountain valleys	UKH mountain valleys (*Renosterbos*)	6

Appendix V: Vegetation type description

Southern Karoo Riviere

Initial Name:		Floodplain
New Name:		**SKR Homogenous grassland _Merxmuellera disticha_**
	Ground cover	Lower percentage of barren soil compared to other class, some bedrock present on the surface in one transect. High variance in ground cover in general, organic cover ranges from 34 to 80%.
	Undergrowth	Clearly lower undergrowth ratio (with low variance): high canopy cover. The lower layers have a higher relative cover than the lower layers of the Riverine vegetation, but because it's overall cover and canopy cover is lower than the Floodplain vegetation, the undergrowth ratio is relatively higher
	Species composition	High relative abundance of grass, especially in top canopy layer
	Species richness	Similar to other subtype
	Dominant species	
	- Grass	Merxmuellera disticha & Aristida vestita
	- Forb	Pseudognaphalium undulatum
	- Shrub	Rosenia Humilis & Eriocephalus ericoides & Helichrysum Rosum
	Summary	Grass dominated plain with high grass canopy cover, homogenous grass cover. Merxmuellera disticha and Aristida vestida dominant.

Initial Name		Riverine
New Name		**SKR heterogeneous riparian _Pteronia tricephala_**
	Ground cover	Higher percentage of barren soil with a low variance. Lower basal cover
	Undergrowth	Higher undergrowth ratio relatively to the grasslands, but a lower absolute amount of cover in most layers. Vegetation seems to occur more in patches between barren soil.
	Species composition	The relative abundance of shrubs in the top canopy is much higher compared to the Floodplain vegetation. Where the Floodplain vegetation canopy is mainly composed out of grasses, the Riverine top canopy has a more equal distribution of grasses compared to shrubs (although the variance in shrub abundance is high). The overall relative and absolute grass cover is lower, but the variance in grass cover between transects is high (and higher than the Floodplain vegetation).
	Species richness	Similar to other subtype
	Dominant species	

- *Grass*	Miscanthus capensis	
- *Forb*	Arctotis sulcocarpa & Anchusa capensis	
- *Shrub*	Pteronia tricephala & Selago gracilis	
Summary	Riverine vegetation with alternately barren soil and vegetation patches. Dominated mostly by Miscanthus capensis (and an unknown grass species) and shrub Pteronia tricephala. Hardly any M. disticha present	

Karoo Escarpment Grassland

Initial Name		Riverine floodplain
New Name		**KEG heterogenous riparian *Merxmuellera disticha***
	Ground cover	Higher percentage of barren soil compared to other classes, a merely equal distribution of soil to organic material. Most of the organic material is made out of litter, the basal cover is relatively low (19%)
	Undergrowth	Canopy undergrowth ratio is clearly lower than all other classes, meaning there is relatively larger undergrowth. Seems like (similar to the riverine vegetation type of SKR) as if the vegetation occurs in patches amongst patches of barren soil. The absolute cover of all layers is higher compared to other vegetation types.
	Species composition	Grass most dominant in all layers of the vegetation. Relatively more grasses than shrubs in the undergrowth compared to the top canopy. The top canopy composition has a higher relative abundance of shrubs compared to the other vegetation types
	Species richness	Similar to other subtypes (14 species)
	Dominant species	
	- *Grass*	Merxmuellera disticha & Aristida vestita & Aristida diffusa
	- *Forb*	
	- *Shrub*	Elytropappus rhinocerotis & Euryops nodosus & Euryops annae
	Summary	Riverine grassland dominated by with Merxmuella disticha and Aristida vestida and Aristida diffusa. Canopy cover relatively more composed of shrubs. Dominant shrubs are Renosterbos, Euryops nodosus and Euryops annae. Differs in dominant species from SKR riverine vegetation type

Initial Name		S Hills
New Name		**KEG south sloping hills *Merxmuellera disticha – Renosterbos***
	Ground cover	Lower percentage of barren soil. Soil is mostly covered with small gravel. Some rock and a relatively high percentage (20%) of bedrock present. Still, in comparison with the North direction hills, the organic material

		cover is much higher (both litter and basal cover)
	Undergrowth	High ratio of canopy cover to undergrowth. Vegetation cover in all three layers is lower compared to other classes, but the canopy cover is quite similar.
	Species composition	Grass is dominant in all layers, but shrub abundance is higher in the top canopy compared the shrubby N hills. The overall relative abundance of shrubs is also higher, but the variance is very high.
	Species richness	Similar to other subtypes (16)
	Dominant species	
	- Grass	Merxmuellera disticha & Themeda triandra & Enneapogon scoparius
	- Forb	Plantago laceolata
	- Shrub	Elytropappus rhinocerotis & Euryops nodosus
	Summary	Grass dominated south slopes with dense vegetation cover and higher abundance of small shrubs (renosterbos) compared to northern directed hills. Grasses dominated by Merxmuellera disticha, Themeda triandra, and Enneapogon scoparius.

Initial Name		**Shrubby N hills and Upper shrubs (merged)**
New Name		**KEG sparse heterogeneous highlands**
	Ground cover	Higher percentage of barren soil and bedrock ground cover compared to the south directing hills and a lower organic cover (both basal and litter).
	Undergrowth	Lower ratio than S hills, indicating a higher undergrowth level
	Species composition	Some big shrubs randomly interspersed (although not sampled) that are clearly visible from aerial photographs. Grass is more dominant and lower shrub cover compared to S hills. Grass is more abundant in the top canopy layer compared to all vegetation types. High forb abundance in one of the transects.
	Species richness	similar to other subtypes
	Dominant species	
	- Grass	Merxmuellera disticha & Aristida sp. & Themeda trianda
	- Forb	Indigofera alterans & Lotonis azureoide
	- Shrub	Passerina corymbosa & Chrysocoma ciliata
	Summary	North directing slopes and high altitudinal plateau's and cliffs with a lower vegetation cover compared to southern slopes. Higher undergrowth level might indicate that vegetation occurs in patches amongst rock and bedrocks. Scattered large shrubs (up to 1 m) amongst bedrock cliffs and gravel covered soil. Although from aerial photographs large shrubs were clearly distinctive, the shrub abundance is lower compared to southern slopes. Larger shrubs visible from aerial

Eastern Upper Karoo

Initial Name		Riverine
New Name		**EUK riparian grasslands *Eragrostis curvula***
	Ground cover	Relatively low percentage of barren soil and high organic matter + basal cover. Some bedrock is present in two of the transects, but the variance is very high
	Undergrowth	High canopy to undergrowth ratio, indicating a relatively low undergrowth level (although variance is high as well)
	Species composition	Clearly grass dominated, shrubs occur randomly in patches. Shrub abundance is quite high in the top canopy layer compared to other classes, but the variance in relative shrub abundance is high. One of the transects (59) was located in a patch of short grazed grass.
	Species richness	Similar to other subtypes
	Dominant species	
	- Grass	Eragrostis curvula & Tragus koeleroides & Bromus cartharticus
	- Forb	Asteraceae sp.
	- Shrub	Euryops annae & Pentzia quinquefida
	Summary	Variable grasslands surrounding the riverbeds with patches of palatable grasses and a relatively high vegetation cover. Shrubs occur in patches in the grassland.

Initial Name		Mountainous shrubland
New Name		**EUK mountain shrubland *Helichrysum rosum***
	Ground cover	Relatively high percentage of soil and bedrock
	Undergrowth	Relatively more forbs in the undergrowth
	Species composition	Relatively high shrub and forb abundance although grass remains dominant. Forbs mainly present in the undergrowth
	Species richness	Lowest of all classes but does not seem significant
	Dominant species	
	- Grass	Aristida vestita & Merxmuellera disticha

	- Forb	Fabaceae sp.
	- Shrub	Helichrysum rosum & Eriocephalus ericoides
	Summary	High altitude shrubland with more forbs and less grasses compared to other classes.

Initial Name		**Mountainous floodplain**
New Name		**EUK stony grassland *Aristida vestita***
	Ground cover	Low percentage of barren soil and high amount of bedrock (ground cover was estimated only for one transect due to an error made in the other two)
	Undergrowth	High canopy to undergrowth ratio, indicating a high canopy cover (74%), that is mainly composed of grass and a small percentage of shrubs and little undergrowth
	Species composition	Clearly grass dominated: high grass abundance and little shrub species.
	Species richness	Similar to other subtypes
	Dominant species	
	- Grass	Aristida vestita & Heteropogon contortus
	- Forb	
	- Shrub	Elytropappus rhinocerotis & Chrysocoma ciliate
	Summary	Sparsely covered vegetation that is dominated by grasses, occurring among present bedrock. Aristida vestita is the dominant species in all three transects.

Initial Name		**Grassland + Floodplain 2 (merged)***
New Name		**EUK tufted grassland *Merxmuellera disticha***
	Ground cover	Higher percentage of barren soil compared to the floodplain and a much higher organic material cover, mainly composed of litter.
	Undergrowth	Lower undergrowth ratio than the mountainous floodplain, indicating that more species are occurring below the canopy
	Species composition	Some shrubs occurring in patches within the grasslands. Grass is dominant in all layers.
	Species richness	Similar to other subtypes
	Dominant species	
	- Grass	Merxmuellera disticha
	- Forb	

- Shrub	Chrysocoma ciliata	
Summary	Tufted grassland where small patches of various shrubs species occur. Tufts systematically cover the ground and between tufts, herbaceous litter and barren soil is found.	

Initial Name	Floodplain 1
New Name	**EUK variable floodplain *Eragrostis lehmannia***
Ground cover	Lower vegetation cover compared to grasslands and more barren soil (although only based on one transect due to error in others).
Undergrowth	Seems like less undergrowth is present due to a very high canopy to undergrowth ratio is one transect. High variance indicates alternating vegetation structure (confirmed by field experience: field was characterized by grass and shrub dominated patches)
Species composition	Lower vegetation cover in all layers compared to grassland. The relative abundance of grass, shrubs and forbs is comparable to grassland vegetation type
Species richness	Slightly higher than other classes (15)
Dominant species	
- Grass	Eragrostis lehmannia & Merxmuellera disticha & Karoochloa sp.
- Forb	
- Shrub	Felicia fascularis & Pentzia incana
Summary	Overall high variance in community composition

* The initial identification of vegetation types based on aerial photographs was already based on an idea of Floodplain 2 having a more homogenous and grass dominated cover compared to Floodplain 1. Comparison of the transect data confirms that the vegetation types are very similar, in both composition and species dominance

Upper Karoo Hardeveld

Initial Name	S mountains
New Name	**UKH southern escarpment *Chrysocoma ciliate***
Ground cover	Higher percentage of present bedrock. Vegetation cover (basal and litter) is quite similar to N mountains.
Undergrowth	Higher ratio of canopy to undergrowth compared to N mountains, indicating a lower undergrowth
Species composition	Less vegetation cover in all layers compared to N mountains. More shrubs in top canopy
Species richness	Similar to other subtypes

Dominant species		
	- Grass	Merxmuellera disticha
	- Forb	
	- Shrub	Chrysocoma ciliate & Elytropappus rhinocerotis
	Summary	High altitudinal, steep south sloping part of escarpment. Slightly higher shrub abundance in the canopy (Renosterbos and Chrysocoma ciliate) and less grass than on N mountains. Overall vegetation cover is lower than on N mountains, which is not in line with expectations regarding north facing slopes are hotter and drier, creating less favourable conditions for plant species. Different from N mountains in higher bedrock percentage and species dominance

Initial Name		**N Mountains**
New Name		**UKH northern escarpment *Themeda triandra***
	Ground cover	Lower bedrock cover but more loose rocks. Organic cover is comparable to the south sloping mountains, although there is a huge variance.
	Undergrowth	Lower undergrowth ratio, more grasses and shrubs in the lower layers compared to south slopes
	Species composition	Higher vegetation cover in all layers. High grass abundance in the top canopy, mainly consisting of T. thriandra and M. disticha
	Species richness	Similar to other subtypes
	Dominant species	
	- Grass	Themeda triandra & Merxmuellera disticha
	- Forb	
	- Shrub	Elytropappus rhinocerotis
	Summary	High altitudinal, steep, north sloping part of escarpment. Although unexpected, the vegetation cover is higher on these slopes and less bedrock is present. Still, ground is covered with rocks and gravel. High variance in species composition and ground cover, indicating more heterogeneous vegetation type (due to high geological variation and creation of microclimates on slopes?)

Initial Name		**N Hills**
New Name		**UKH northern grassland *Themeda triandra***
	Ground cover	Less bedrock and more barren soil compared to N mountains. Organic cover is quite similar although the basal cover is lower compared to N mountains.

	Undergrowth	Higher canopy to undergrowth ratio. The vegetation cover in all layers is lower compared to the mountains.
	Species composition	More grass dominated than N mountains and much less undergrowth
	Species richness	Similar to other subtypes
	Dominant species	
	- Grass	Themeda triandra & Merxmuellera disticha
	- Forb	
	- Shrub	Melolobium candicans
	Summary	Similar in composition and species dominance compared to steeper slopes of N mountains, although grass is more dominant.

Initial Name		**S Hills**
New Name		**UKH southern grassland *Eragrostis lehmannia***
	Ground cover	More barren soil and less rock and bedrock than N hills (and much less compared to S mountains)
	Undergrowth	Similar to N Hillls
	Species composition	Similar to N Hills
	Species richness	Similar to other subtypes
	Dominant species	
	- Grass	Eragrostis lehmannia & Aristida vestita
	- Forb	
	- Shrub	Elytropappus rhinocerotis
	Summary	Composition and cover quite similar to N hills, but the (grass) species present are very different: hardly any rooigras and more eragrostis and aristida species. Compared to the S mountains, the grass composition is higher and there are less shrubs (but more shrubs compared to N Nills)

Initial Name		**High plateau**
New Name		**UKH high diversity altitudinal plateau**
	Ground cover	High organic cover (mainly basal cover), little rock and bedrock
	Undergrowth	Relatively low, indicating that there is quite a high level of undergrowth

Species composition	High vegetation cover in all layers, vegetation occurred in dense cushions in some places. Forb and shrub abundance is much higher than in other classes.	
Species richness	Higher than all other classes	
Dominant species		
- Grass	Pentaschistis sp. & Merxmuellera disticha	
- Forb	Stachys sp & Indigastrum sp.	
- Shrub	Helichrysum sp. & Passerina corymbosa & Melolobium candicans	
Summary	High altitude small slope lands with a relatively high vegetation cover and high shrub and forb abundance. Overall variance is high, indicating the existence of microclimates. High in species richness.	

Initial Name	**Shrubby floodplain**	
New Name	**UKH mountain valleys *Renosterbos***	
Ground cover	Relatively high basal cover	
Undergrowth	Very high variance	
Species composition	Less grass abundance than the S and N hills and more shrubs, in both top canopy and lower layers. Merxmuellera disticha is significantly dominant in the top canopy for all transects, ensuring a high vegetation cover from the tufted grasses.	
Species richness	Similar to other subtypes	
Dominant species		
- Grass	Merxmuellera disticha & Eragrostis lehmannia	
- Forb		
- Shrub	Elytropappus rhinocerotis	
Summary	Mountain valleys where vegetation grows among drainage lines (mainly gullies) with a relative high vegetation cover (basal and canopy cover) and less grasses compared to surrounding hill slopes.	

Appendix VI: Explanation of MaxEnt setting

1. Background sampling

By default, MaxEnt assumes that all areas within the landscape hold the same probability of being selected as a background sample in accordance to its frequency of occurrence. Different background samples result in different response curves. The background samples should reflect the landscape of interest based on prior ecological knowledge and the observed species presence localities (Merow et al., 2013). The default value of background sampling points is 10.000.

2. Feature types

MaxEnt produces various response curves of the input variables in relation to the predicted species distribution. Different feature types put different constraints on descriptive statistical outcomes of each environmental variable at the predicted distribution in relation to the observed distribution. Five different feature types are used in MaxEnt:

- **Linear features** ensure that the mean observed value of the input variables where species occur equals the **mean** predicted values
- **Quadratic features** ensure equal **variance**
- **Product features** ensure equal **covariance** with other predictors
- **Threshold features** make continuous variables **binary** (with values of 0 below the threshold and 1 above)
- **Hinge features** are like **threshold** features but use a **linear function** rather than a step function

The user is free to choose which feature types (hence constraints) will be used in the model. By default, a subset of feature types is selected when the dataset counts < 80 presence localities (Merow et al, 2013). By means of a sequential-update algorithm, the feature types most suited to estimate the target probability distribution are selected. The regularization parameter (described below) determines the weight imposed by the constraints of each feature type.

3. Regularization parameter β

MaxEnt selects the feature types that improves the model fit using regularization. Regularization reduces over-fitting by loosening the constraints in such a way that the predicted distribution approximates rather than perfectly fits the observed distribution for each variable. Regularization parameter β_j determines the range of the error bound accepted for feature j. It is defined by:

$$\beta_j = \beta \sqrt{\frac{s^2[f_j]}{m}} \qquad \qquad \textit{(Eq. 3)}$$

Where β is the regularization parameter that can be adjusted and determines the power of regularization parameter β_j. If the default value of β = 1 is kept, βj is defined by an estimate of the standard deviation of the empirical average of feature f_j (Philips & Dudík, 2008). Increasing the regularization parameter simplifies the model by removing more features, but it also undermines the theorem of analogous uniform probability distribution of the observed and predicted species distribution. Reducing the parameter is recommended when a large observation dataset is available. With smaller data sets, the model tends to capture highly non-linear response curves of the environmental variables resulting in a complex, overfitting model (Merow et al., 2013). Without regularization (β = 0), the estimated probability distributions are required to be exactly equal to the observed distributions.

According to Philips & Dudík (2008), model predictive power is generally optimized when regularization coefficients decrease with increasing availability of presence localities. This is consistent with the notion that relatively more presence localities allow a more accurate estimation of the species distribution and thus tolerate smaller error bounds (error bounds represent the accepted deviation of the estimated to the observed species distribution for each variable).

4. Sampling bias

As mentioned in subchapter 4.2.2. Sampling bias, sampling bias can result in an environmental bias that may lead to inaccurate models (Philips et al., 2009). To correct for sampling bias, MaxEnt supports the incorporation of the spatial sampling effort or an estimation thereof (The FactorBiasOut Method as described by Phillips et al. (2009). An evaluative study of sampling bias correction methods by Fourcade et al. (2014) shows that the use of a bias grid that represents an estimated sampling effort based on a Gaussian Kernel Density distribution of the observation localities, as used by Elith et al. (2010), is a suitable method to account for sampling bias in geographic space.

5. Output types

MaxEnt returns three forms of output data:

- **Raw**: returns the *probability that a species is found at pixel X.* It is calculated based on the probability density of covariates at locations where the species is present divided by the probability density of covariates across the landscape of interest (Elith et al., 2011). This

ROR (Relative Occurence Rate) of the raw output sums to unity (1) if all locations on the landscape are included in the background (Merow et al, 2013).

- **Logistic:** returns the *probability of presence of a species*, using an intercept between the logistic transformation of the raw data and the raw data itself. Uses an 'arbitrary' parameter $T = 0.5$ (Elith et al., 2011).

- **Cumulative:** Rescales the raw values to lie between 0 and 100. Cumulative output can interpreted in terms of an omission rate because thresholding at a value of C to predict presence surface will omit C% of the presences (Merow et al 2013)

Both the Logistic and Cumulative outputs are a post-transformation of the raw output data.

6. Model evaluation

In order evaluate the modelling outcomes, MaxEnt produces a range of metrics that allow for model evaluation and comparison. The most common measure of model evaluation if the Area Under Curve (AUC), which is a metric that quantifies 'the proportion of test localities correctly predicted present relative to the proportion of all map pixels predicted to have suitable conditions for the species'. An AUC value of 0.5 implies that the model performs no better than random, a score < 0.8 is poor, between 0.8 and 0.9 is fair and higher than 0.95 is excellent (Thuiller et al., 2006). It must be underlined that AUC scores are often higher for species that have a narrow range relative to the study area because the localized occurrences are likely to be coupled with narrow environmental ranges (niches) that are easily predictable by MaxEnt.

A popular way to evaluate model outcomes uses *k-fold cross-validation* of the presence localities: the data set is divided in *k* independent subsets, where for each run, the model is trained with *k-1* subsets and tested with the kth subset (Merow et al., 2013).The model returns two values of the AUC: one for the training data and one for the test data. MaxEnt supports two other ways of model replication: Repeated subsampling (presence points are repeatedly split into random training and testing subsets) and bootstrapping (where training data is selected by sampling with replacement from presence points).