Kalahari Wildlife Landscape Connectivity Analysis

Phase 1 Report

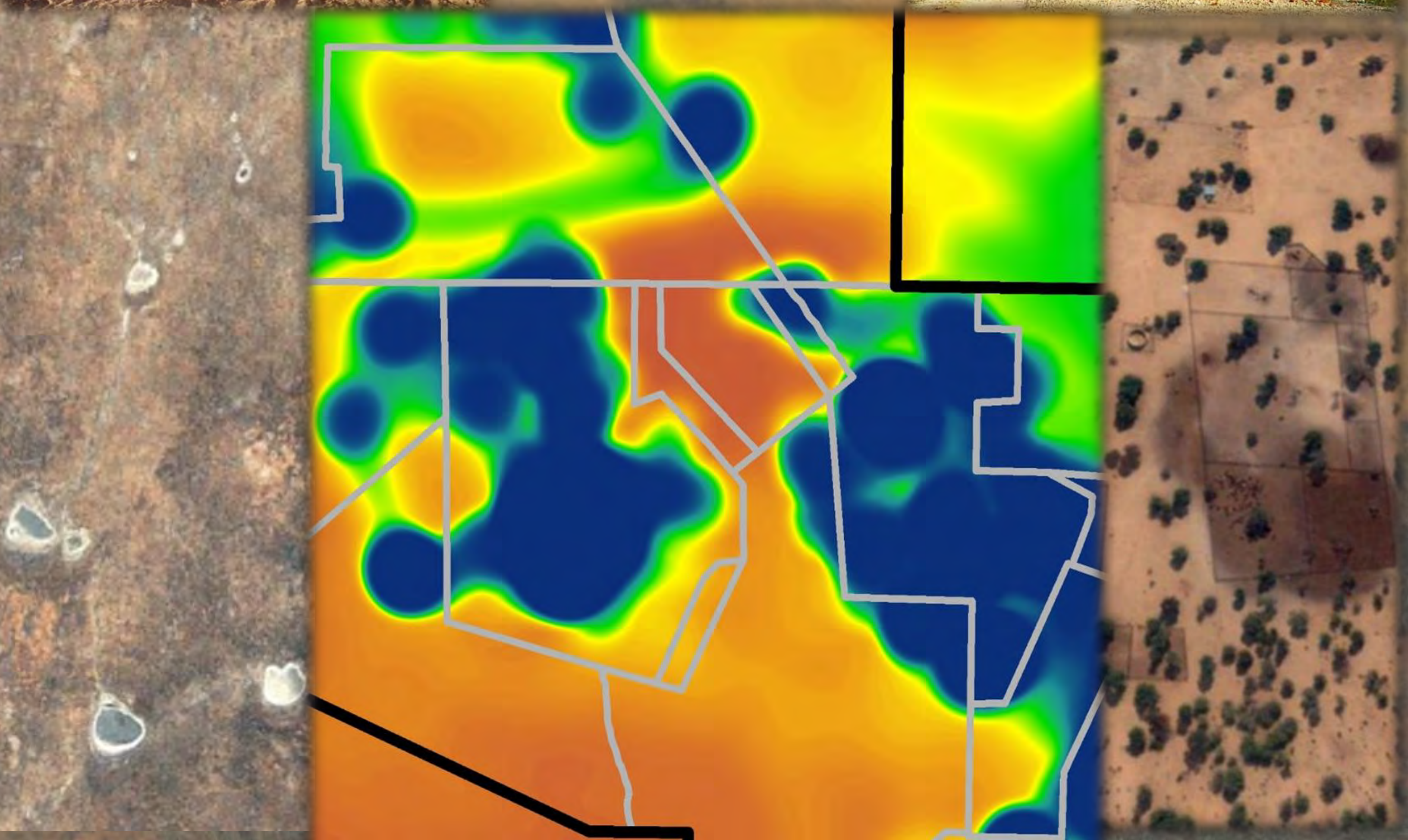
For

UNITED NATIONS DEVELOPMENT PROGRAMME

Kgalagadi-Ghanzi Drylands Ecosystem Project

Derek Keeping PhD
Zaneta Kaszta PhD
Samuel A. Cushman PhD

February 2022



Acronyms and Abbreviations

AOI	Area of Interest (aka KWLCA study area)
AUC	Area under ROC Curve
CGA	Communal Grazing Area
CI	Conservation International
CKGR	Central Kalahari Game Reserve
DWNP	Department of Wildlife and National Parks
GEF	Global Environment Fund
GPS	Global Positioning System
GH	Ghanzi District
HS	Habitat Suitability
ILUMP	Integrated Land Use Management Plan
KD	Kgalagadi District
KGDEP	Kgalagadi and Ghanzi Drylands Ecosystem Project
KW	Kweneng District
KWLCA	Kalahari Wildlife Landscape Connectivity Analysis
KTP	Kgalagadi Transfrontier Park
MIR	Model Improvement Ratio
NDVI	Normalized Difference Vegetation Index
PAR	Pastoral/Arable/Residential
PCC	Percent Correctly Classified
RF	Random Forest
RAD	Remote Area Dweller
ROC	Receiver Operating Characteristic
SDM	Species Distribution Model
SO	Southern District
TGLP	Tribal Grazing Lands Policy
ToR	Terms of Reference
TSS	True Skill Statistic
UNDP	United Nations Development Programme
UTM	Universal Transverse Mercator
VCF	Vegetation Continuous Field
WKCC	Western Kgalagadi Conservation Corridor
WMA	Wildlife Management Area

Executive Summary

The purpose of this report is to present and discuss Phase 1 results of the Kalahari Wildlife Landscape Connectivity Analysis (KWLCA). This is an intermediate stage of the KWLCA marked by the completion of comprehensive wildlife species habitat suitability (HS) models for 32 species, presented as occurrence probability surfaces across the project landscape. These HS models will be inverted into resistance surfaces for select species then combined with other landscape resistances (e.g. fences, roads) for Phase 2 wildlife movement (connectivity) modelling of present time and future scenarios.

In this Phase 1 we subjected a unique animal track-based dataset of comprehensive multi-species wildlife occurrence to a rigorous spatial analysis of habitat suitability using diverse multi-scale landscape variables that included novel fine-resolution mapping of human-livestock disturbance.

Results show that although most species responded negatively to human-livestock disturbance, many of the smaller herbivores, carnivores, insectivores are widespread and have at least some level of tolerance to intermediate kraals density, i.e. borehole allocations at enforced spacing (6-8 km) in PAR/CGA land use zones. Further examination of connectivity concerns for these species is unnecessary as a) land use is unlikely to intensify beyond this level over large areas of the landscape in the foreseeable future; and b) fences are largely permeable to their movements.

By contrast, most large herbivores (antelopes) and carnivores are highly sensitive to pastoral activity in the landscape. They are therefore dependent on land uses that exclude or restrict human-livestock disturbance, and are threatened by habitat loss and landscape fragmentation (loss of connectivity) due to pastoral encroachment. Gemsbok, eland, and lion proved the most disturbance sensitive. These two antelopes had the strongest performing predictive models among all species, due in large part to their exceedingly strong negative responses to kraals density at the largest spatial scale examined (32 km radius kernel). Their spatial responses to fenced livestock operations were similarly negative, indicating that fences do not form hard boundaries separating pastoral areas from wildlife areas but rather negative impacts extend far beyond the fence lines often assumed to contain them.

Habitat suitability and combined biodiversity maps indicate the highest value wildlife core comprised of KTP and adjacent WMAs (KD1,2,12,15). The occurrence probability surfaces of the most disturbance-sensitive species indicate continuity through 2 strong corridors linking the aforementioned southern core to the northern core (CKGR, KW2, GH10 WMAs), comprising the ‘central’ corridor (KD5,6,12 WMAs) and the ‘western’ corridor (GH11,13, KD1). The relative importance of these two corridors appears to differ depending on species. The formal connectivity analyses to be conducted in Phase 2 will be illuminating in assessing the functional connectivity of these corridors for the most sensitive species.

The present KGDEP is timely considering the already allocated, planned and proposed agricultural developments (RAD borehole allocations, WMA dezonings, fenced ranch expansions) which will impact precisely these precariously remaining corridors of wildlife connectivity. The impact of these specific detailed landscape change scenarios will be examined, quantified and reported in Phase 2 to more incisively and explicitly inform land use planning. Given exceptional gemsbok model prediction performance and strong response to modelled disturbance variables that stand to change markedly in the future landscape scenarios, gemsbok is nominated as an umbrella in Phase 2 connectivity modelling for

slightly less disturbance-sensitive species, to be supported by eland and lion in a subset of packaged scenario options.

The habitat suitability maps presented herein represent an already outdated or somewhat historical state of landscape in terms of encroachment (i.e. the situation for wildlife can appear in these maps better than it presently is). This is because models are based on borehole allocations that were already developed as operational cattleposts during the past decade coinciding with track data capture. There has been a somewhat rapid allocation of new borehole rights, particularly among RAD communities situated in proximity to key areas of connectivity in the landscape, and there is marked lag affect between time of landboard allocation and time these new locations become developed into cattleposts to influence wildlife distributions. In Phase 2 will we update models to reflect the present time situation, before venturing into future scenarios including those based on borehole allocations already approved but not yet developed on the ground.

Note we use 'habitat suitability' and 'probability of occurrence' synonymously and interchangeably throughout the report. We use 'human-livestock disturbance' loosely interchangeably with 'kraals' or 'kraals density'.

Table of Contents

1 INTRODUCTION

The Government of Botswana, with the support of the United Nations Development Programme (UNDP) is undertaking a Global Environment Fund (GEF)-financed project entitled "Managing the human-wildlife interface to sustain the flow of agro-ecosystem services and prevent illegal wildlife trafficking in the Kgalagadi and Ghanzi Drylands", more simply known as the Kgalagadi and Ghanzi Drylands Ecosystem Project (KGDEP).

Key to the KGDEP is development of an Integrated Land Use Management Plan (ILUMP) for Kgalagadi and Ghanzi Districts. These two Districts (along with lesser parts of Southern and Kweneng Districts) comprise what remains of the core free-ranging Kalahari wildlife landscape in southern Africa. The main objective of the ILUMP is to address the competition and conflict between land uses in this landscape and to provide a framework for maintaining wildlife migratory corridors between Kgalagadi Transfrontier Park (KTP) and Central Kalahari Game Reserve (CKGR).

Accordingly, a project area expert consultant was procured to conduct a quantitative Kalahari wildlife landscape connectivity analysis (KWLCA) to generate the appropriate spatially explicit guidance for the ILUMP that ensures the KGDEP achieves its component goals. The KWLCA requires computationally intensive state-of-the-art landscape modelling, so the consultant necessarily collaborated with leading world experts at EcoAnalytica LLC, a world-leading ecological modeling and informatics consultancy based in the United States.

The Inception Report for the KWLCA (May 2021) described the background, objectives, methodology and proposed timeline for the KWLCA, as well as initial ideas for integration with the early phase Situational Analysis of the ILUMP. This report describes the actual methodological approaches taken in detail, including the wildlife and landscape data, steps involved in their processing, and the analyses workflow to produce the results. Phase 1 results are presented predominantly in the form of habitat suitability maps, including predictive performance measures of the models, models structure with variables coefficients and strengths, along with plsmo spline plots showing responses to specific variables. We provide tentative inferences about the current state of wildlife core areas and connectivity from these results and justify species selection for Phase 2 connectivity modelling and scenario planning.

The Deliverables for Phase 1 reporting from the ToRs are reiterated here:

- Assessment of the current state of the comprehensive multi-species free-ranging Kalahari wildlife landscape (from data based on multi-scale modelling of distribution and abundance), including the identification of core areas and existing connectivity between KTP and CKGR.
- Identification of the key environmental driver(s) of wildlife species spatial use in the Kalahari landscape, including quantification of spatially explicit response gradients of wildlife species occurrence and abundance in relation to those key environmental driver(s).
- Identification of the subset of most disturbance-sensitive species for further analysis.

2 METHODOLOGY

2.1 Study Area

The KWLCA area of interest (AOI) was necessarily defined for all subsequent data layers preparation and analyses. The AOI encompasses the full extent of the remaining free-ranging wildlife landscape in between KTP and CKGR (Figure 1). It includes most of Kgalagadi District, much of Ghanzi District, and parts of Southern and Kweneng Districts. Present land use in between the two protected areas is largely comprised of unfenced Wildlife Management Areas (WMAs) and Communal Grazing Areas (CGAs), plus lesser portions enclosed with fencing as ranches, farms, and ploughing fields.

The AOI extent is bound by the international border fences with Namibia and South Africa in the west through southeast, by a reasonable buffer into Ghanzi District CGAs and TGLP ranch block in the northwest - and similarly in the east, and an acceptable distance into CKGR. The CKGR extent was determined as a tradeoff between our reluctance to over-extend modelling prediction in the northeast too far up the rainfall gradient with limited wildlife locational data on the one hand, and on the other hand the need to include a large enough core area of CKGR for source populations required in Phase 2 modelling.

The AOI is 170,128 square kilometres in extent.

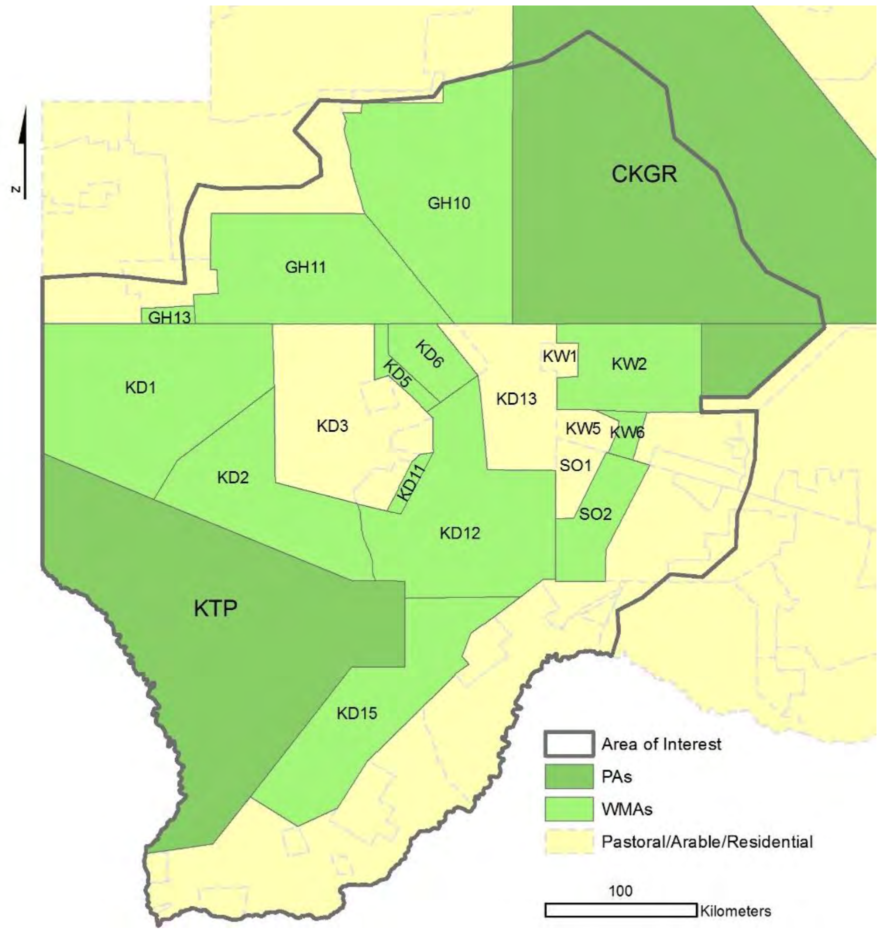


Figure 1 AOI for the KWLCA. Select land use blocks within the study extent are labelled with their alpha-numeric identifiers.

2.2 Wildlife geolocational data – track (spoor) transects

2.2.1 Transect coverage

Wildlife track data were collected along various transect coverages over the period August 2008 – April 2018. Track transects of varying lengths typically coincided with low-traffic sand roads, cutlines, and 4x4 trails, but also include lesser lengths along upgraded calcrete roads and the meridians of paved roads, and some coverage off-road. Multiple criteria guided transect selection. Some were selected to sample disturbance gradients i.e. from densely utilized CGAs through the last cattlepost at the frontier edge, continuing through WMAs to the protected area boundaries. Other transects were selected to randomly sample land use types. All suitable linear features that bisected potential corridor areas in the AOI were sampled throughout their continuous lengths. The majority of transects were replicated over multiple years, in different seasons. Overall, coverage includes extremes in disturbance from the outskirts of major villages to the core of the KTP. Total transect coverage analyzed was 9,274 km (Figure 2).

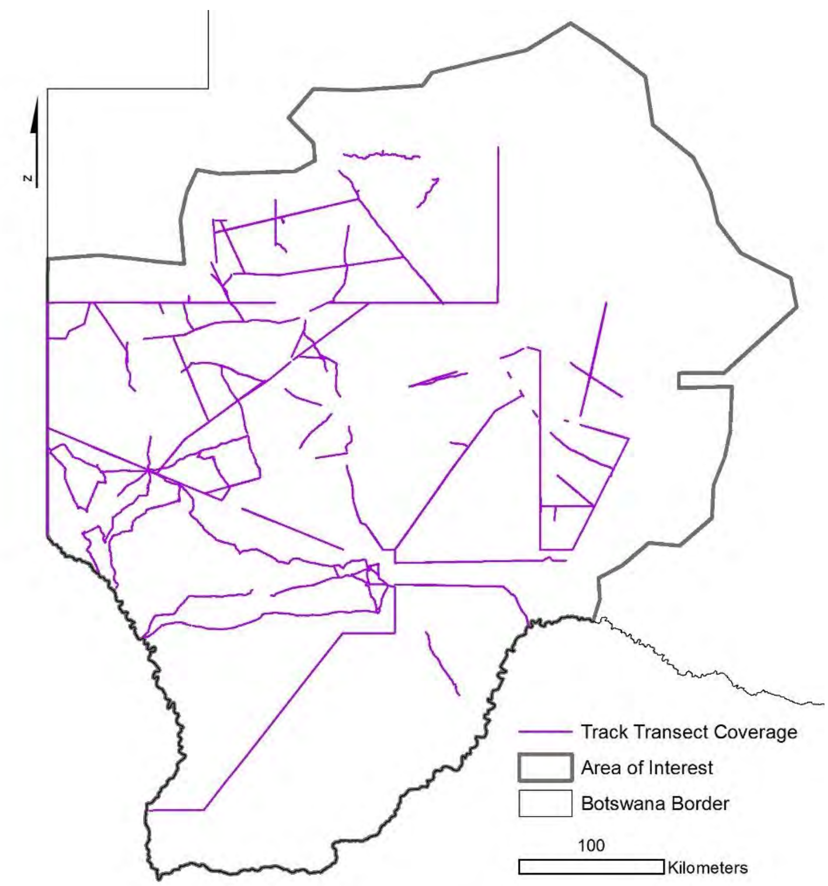


Figure 2 Wildlife track transect coverage within the AOI. The majority of transects within the spatial coverage were temporally replicated.

2.2.2 Transect sampling

A key distinction between transects is those for which comprehensive species > 0.2 kg were sampled vs those for which the subset of large species only were sampled. Large species are those generally > 25 kg body mass, i.e. the antelopes besides dwarf species (steenbok, duiker), ostrich, large cats, hyenas and wild dog. The distinction between these two groups was naturally born of convenience and practicality of tracker search image related to size of tracks.

2.2.2.1 Comprehensive species transects

A subset of transects were sampled in 2008-2014 for comprehensive species (i.e. all mammalian wildlife > 0.2 kg body mass, plus large terrestrial birds). These transects were precleared of old tracks by dragging a heavy steel beam behind a vehicle, so that fresh tracks accumulated over the following 24-hr period before sampling. Without pre-clearing the transect first, track accumulation of comprehensive species is too great, especially including the smaller abundant animals (e.g. hare, springhare, jackal, steenbok) to allow collection of data in an efficient and standardized way. Additionally, some of the smaller and lighter foot-loading species are difficult to see on compacted and/or vegetated ground. Preclearing therefore improved detectability, minimizing false absences. Surveys began early morning and were conducted by two observers (Panana Sebati – expert traditional tracker from Ngwatle village, and Derek Keeping – also a competent tracker) on specialized 'tracker seats' mounted to the front of the vehicle. Progressing at a meticulous rate between 6 and 8 kph, all track intersections with the transect were recorded as species and numbers with global positioning system (GPS) locations. The advantage of two simultaneous observers minimized missed tracks but also allowed difficult and obscure spoors to be examined in detail until consensus on identity reached, using Liebenberg (1990) as a reference when necessary.

2.2.2.2 Exclusively large-bodied species transects

The remainder of transects sampled in 2014-2018 differed from the comprehensive species transects in that they were not precleared or prepared in anyway prior to sampling. As only large-bodied species with larger spoor were considered, sampling proceeded at faster speed, typically 8 – 15 kph. Observations were made similarly from the front of the vehicle by a minimum 2 expert observers (traditional trackers from villages Zutshwa, Ngwatle, Ukwi, Maake, Bere, Kacgae). Large herds were enumerated using handheld mechanical tally counters. The other distinction from comprehensive species transects was that older tracks were recorded, and ages of observations estimated (i.e. < 24 hrs, > 24 hrs, 3 days +, 1 week +, 1 month +).

2.2.3 Wildlife species occurrence data

36 mammalian and ground bird wildlife species above threshold size were detected with regularity on the track transects (Table 1). Secretary bird were also detected with regularity but omitted from further consideration based on low observations and the fact that they move more by air than ground. Additional wild mammals of target size detected but with irregularity included baboon, banded mongoose, giraffe and elephant. We consider these species uncommon or having peripheral ranges in relation to the project area of interest, although elephant appeared to be increasing towards the latter end of the sampling decade.

Table 1. Transect observations and numbers of 100m segments along the transects with presence (1) and absence (0) back-transformed from density for each of 36 species.

Species	Field data		Processed locational data		
	Track sets (individual intersections)	Occurrences (GPS locations)	Absence	Presence	Presence (%)
pangolin	44	42	29035	1050	3.5
wild dog	334	115	49945	2763	5.2
warthog	93	38	13687	805	5.6
suricate	149	32	8567	850	9.0
springbok	8109	842	29679	4559	13.3

black-footed cat	35	34	5917	1035	14.9
slender mongoose	159	123	7928	1489	15.8
lion	537	299	60805	11990	16.5
cheetah	423	270	61424	12143	16.5
spotted hyena	339	260	48742	10292	17.4
ground squirrel	471	173	7738	1679	17.8
leopard	554	519	56986	17383	23.4
eland	12489	1176	14988	6148	29.1
wildebeest	3923	1096	22791	10297	31.1
honey badger	373	296	6081	3136	34.0
aardwolf	308	288	5885	3532	37.5
kudu	3774	1692	19723	13550	40.7
caracal	297	271	5373	4044	42.9
duiker	2149	1770	8059	6418	44.3
ostrich	3616	1943	16891	18149	51.8
brown hyena	2110	1944	24695	34339	58.2
aardvark	634	590	3858	5559	59.0
genet	568	546	3728	5689	60.4
yellow mongoose	1714	1355	3464	5953	63.2
kori bustard	1316	1092	3388	6029	64.0
gemsbok	22505	8082	11102	21986	66.4
African wild cat	748	690	3128	6289	66.8
hartebeest	13360	6023	9472	25568	73.0
porcupine	2380	1515	2169	7248	77.0
striped polecat	1981	1651	2120	7097	77.0
bat-eared fox	4669	2228	2148	7269	77.2
cape fox	934	878	1905	7512	79.8
springhare	8271	4804	552	8865	94.1
jackal	7880	6283	149	9268	98.4
steenbok	16427	12066	20	9397	99.8
hare	12723	7845	0	9417	100.0

There was a large disproportion between originally recorded presences and absences for each species along the transects. To avoid zero inflated models, and to reduce fine scale stochastic noise below the scale of our analysis, we transformed presences and absences into continuous density values based on density of occurrence records within 100m transect segments calculated over 250m search radius.

We first prepared the point observations for each species separately. Next, we divided occurrence records for each species by seasons and year to match the seasonal variables. We defined seasons for this analysis as wet (October to March) and dry (March to September). We then applied a 500m buffer over the linear transect as a mask for kernel density calculations to allow a degree of spatial mismatch between the linear transect and point species occurrence observations. Sampling segments of 100m were then prepared by converting the linear transects into 100m resolution rasters, which were then converted back to a point layer for density values extraction.

To calculate density of species occurrence we ran a kernel density using *sp.kde* function in the *SpatialEco* R package (R Core Team. 2021) and applying 250m kernel width. This kernel width was chosen to ensure that the density is calculated over a small enough window to preserve fine scale variation in occurrence distribution and reduce autocorrelation due to overlapping neighborhoods. However, the window should also be large enough to capture neighboring occurrences and reduce the fine scale noise of interspersed zeros from stochastic factors of where and when particular animals crossed. This allowed us to reduce zero inflation caused by too many fine scale absences among scattered presences. We ran kernel density on the species point occurrence data separately for each species, season and year. Next, for each season and year, we extracted the species occurrence density values using the previously prepared 100m spaced sampling points along the transect. We used the same sampling points to then extract the values of all considered predictor variables (this step is further explained in the following section). If the transect/segment was sampled more than one time we averaged the kernel density values as well as values of the seasonal variables (precipitation and NDVI) to receive one spatial location and density/variable value per sampled segment.

To model species habitat selection based on spoor transect data we tested performance of several modelling approaches. These included random forest, generalized linear regression and logistic regression. To do so we tested both continuous species occurrence density values as well as density values back-transformation to binary presence/absence data. The latter was achieved by assigning value 1 (presence) to every location with density higher than zero and value zero (absence) to location with density equal zero. The back-transformation from density to binary data was done to reduce fine scale stochastic noise (at scales smaller than the environmental variation sampled) and the problem of zero inflation in the species data matrix.

2.3 Landscape data and variables

2.3.1 Landscape data

We used a set of 14 different source layers to then develop and test a biologically relevant set of covariates likely influencing habitat use of the studied species (Table 2). These included four anthropogenic and ten environmental factors.

Table 2. Data used to derive variables hypothesized to affect distribution of the study species

Layer	Abbreviation	Description	Original format	Resolution	Final units	Source
Kraals	• KraalsNF (non-fenced) • KraalsF (fenced)	Points representing kraals. Divided into kraals in free-ranging landscape i.e. cattleposts (KraalsNF), and kraals within fenced enclosures like farms and ranches (KraalsF).	Points		Number of kraals /km^2	Ground truth during surveys and digitalized based on Google Earth imagery representing years 2008 - 2018
Human population	HP	Population number per settlements	Points		Estimated human population/km^2	Statistics Botswana Census 2011
Land use	• PA (Protected areas) • WMA (Wildlife Management Areas) • PAR (Pastoral/ Arable/ Residential)	Designated main land use categories	Polygon		Focal mean	Adapted from DWNP BASIS program layers (2009)
Roads	• Roads calcrete • Roads paved • Roads sand	Main types of roads	Polyline		Focal mean	Corrected by digitalization National Roads Database
Artificial Water Points	AWP	Wildlife provisioned water	Points		Number of water points/km^2	Adapted from KCS, and groundtruthing
Pans	Pans	Unvegetated pan surfaces	Polygon		Focal mean	Digitalized based on Google Earth imagery
Percent tree cover	VCF200817	MODIS Terra Vegetation Continuous	Raster	250m	%	NASA data accessed and processed in Google Earth

		Fields Yearly (averaged for years 2008-2018)				Engine
Normalized Difference Vegetation Index	NDVI	Landsat7 8-Day NDVI Composite averaged over each season	Raster	30m	-1 to 1	NASA data accessed and processed per season and year in Google Earth
Precipitation	Prec	Monthly averaged precipitation from 2008 to 2018	Raster	1000m	mm	OpenLandMap Precipitation Monthly accessed and processed per season in Google Earth Engine
Bulk density of the fine earth fraction	Soilbdod	0-5cm mean bulk soil density as measure of soil compaction	Raster	250m	cg/cm^3	SoilGrids ISRIC – World Soil Information accessed and processed in Google Erath Engine
Soil clay content	Soilclay	0-5cm mean proportion of clay particles (< 0.002 mm) in the fine earth fraction	Raster	250m	g/kg	SoilGrids ISRIC – World Soil Information accessed and processed in Google Erath Engine
Soil nitrogen content	SoilN	0-5cm mean of total nitrogen	Raster	250m	g/kg	SoilGrids ISRIC – World Soil Information accessed and processed in Google Erath Engine
Soil sand content	Soilsand	0-5cm mean proportion of sand particles (< 0.05 mm) in the fine earth fraction	Raster	250m	g/kg	SoilGrids ISRIC – World Soil Information accessed and processed in Google Erath Engine

Soil organic carbon content	Soildsoc	0-5cm mean of soil organic carbon content in the fine earth fraction	Raster	250m	g/kg	SoilGrids ISRIC – World Soil Information accessed and processed in Google Erath Engine

2.3.1.1 Kraals

The prevailing human land use influence in the Kalahari landscape is pastoral. We strived to model the most relevant variable representing human-livestock disturbance in the project AOI, and ultimately settled on kraals for the following reasons. Kraals are either steel fenced or thorn branch bomas into which livetock (cattle, goats and sheep) are periodically herded and contained in dense concentration. They develop in the vicinity of any human settlement in the Kalahari, and around any fixed points of water provision (i.e. borehole, or alternatively bowsered water) colloquially known as 'cattleposts'. At fine spatial scale, kraals develop in a clustered pattern around provisioned water points to form a 'cattlepost'. In addition to cattle and/or goats/sheep, a lesser number of horses, donkeys, dogs and chickens are often present at cattleposts. Kraals also form larger denser clusters around the outskirts of Kalahari villages and towns – essentially configured as scaled up oversize cattleposts. Thus, kraals not only signify the point locations where livestock are concentrated within the landscape, but also permanent human settlements and micro-settlements, i.e. the people who tend the livestock. They are therefore the appropriate point disturbance locations from which the synergistic effects of livestock and people attenuate outwards into the surrounding landscape. One familiar impact is the changes to vegetation caused by overgrazing with time resulting in the development of 'piospheres' (Andrew 1988; Perkins 2018). We hypothesize that illegal wildlife harvest and human-wildlife conflict also exhibit predictable attenuating impacts in relation to kraal proximity. Finally, and importantly, at the landscape scale the locations of kraals can be practically managed via land use designations and the allocation of borehole rights entrusted by landboards.

Kraals are readily visible in high resolution satellite imagery as dark-stained features (contrasting with predominantly light-colored soils in the region) caused by the concentrated accumulation of livestock dung. We manually digitalized kraal locations across the project landscape from the best resolution imagery available in Google Earth corresponding to the period over which track data were collected (2008-2018). Within the project AOI, we marked a total 26,843 kraals (Figure 3).

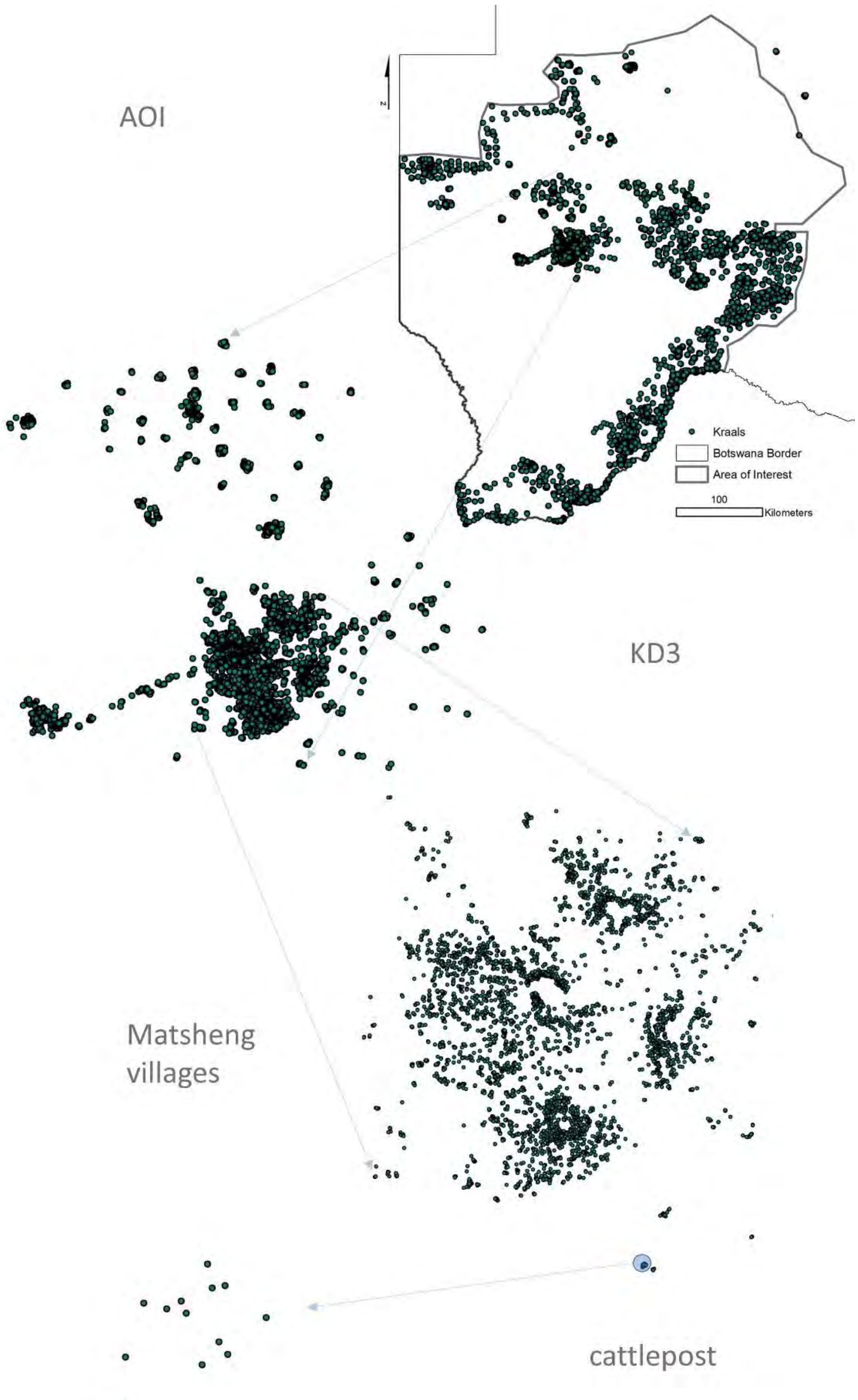


Figure 3 Kraal locations (26,843 points) at AOI scale zoomed to typical cluster of kraals comprising cattlepost scale.

During track data collection, the cattleposts in closest proximity to transects were either groundtruthed, or their vehicle accesses geo-referenced to be crosschecked and verified later in satellite imagery. We are therefore confident in the accuracy and completeness in temporal match of cattleposts (kraals) in nearest proximity to track transects. Notably, in no instances were cattleposts discovered on the ground during the establishment of transects or track data collection that were not visible in satellite imagery.

Several groundtruthing efforts were made during 2019-20 in key areas of the landscape known to be experiencing an expansion of cattleposts (e.g. RAD village development radii) to verify newly founded cattleposts since the time that track data were collected. This limited number of locations that were not relevant at the time of track data collection were omitted from the present habitat suitability modelling and will only be considered in Phase 2 modelling to update a present time picture of the landscape and model future scenarios.

Our analysis distinguished two categories of kraals, those enclosed within fenced farm and ranch areas (KraalsF) and those that were not (KraalsNF). To clarify differentiation between KraalsNF vs KraalsF variables: we do not claim any visible or fundamental distinction between the kraals that develop in free-ranging areas around villages and cattleposts (KraalsNF), and those that develop on farms and ranches (KraalsF). The kraals themselves are of course similar in structure and all fenced in order to hold livestock. The KraalsF vs KraalsNF distinction is rather whether those kraals (or typically kraal clusters) are located within a larger fenced enclosure (i.e. farm/ranch) which separates them from the greater free-ranging wildlife landscape, or not. In fact, many KraalsF begin as functional KraalsNF and only later is a farm or ranch boundary fence erected to enclose them.

Fences were not included as a variable in the habitat suitability modelling because we decided it is most important to measure the spatial impact of kraals inside fenced farms and ranches (KraalsF), independent of fenced boundaries. Fences, however, are obviously important features functioning as filters and barriers to wildlife movement and will therefore be introduced more appropriately into the Phase 2 connectivity models and scenario planning.

2.3.2 Variables

All layers, except of kraals, human population and land use, were rasterized and resampled to 100m resolution and re-projected to UTM 34S coordinate system.

From these data layers we generated in total 135 covariates tested in the habitat suitability modeling. These were derived from multiscale analysis of the original variables. Seventeen of the original variables were calculated at seven different spatial scales, and two variables (kraals fenced and non-fenced) were calculated at eight different spatial scales to represent their spatial effect on wildlife habitat use (McGarigal et al. 2016). Since the studied ecosystem is strongly influenced by seasonality, two of the variables, NDVI and precipitation, were calculated seasonally.

2.3.2.1 Seasonal variables (NDVI and precipitation)

For seasonal averages of NDVI and precipitation we considered wet season as October through March and dry season as April to September. Since precipitation data were represented by monthly averages across the 2008-2018 study period, we generated two precipitation layers: one representing wet season and one for dry season. In the case of NDVI, we calculated seasonal averages of 8-days NDVI for each surveyed year, resulting in 22 NDVI rasters, which were then temporally matched with species occurrence data. The calculations of the seasonal layers and data extraction were performed in Google Earth Engine.

2.3.2.2 Human population

To reflect the spatial distribution and density of the human population in the study region we used census data of the human population for each settlement that is officially recognized and censused by the Botswana government in the Statistics Botswana Census 2011. We converted these census population sizes into a spatial point layer reflecting the spatially distributed populations. We produced a spatial layer reflecting the extent of these populated areas using a calibration approach described by Elliot et al. (2014). Specifically, we calculated the size of the largest town in the study area (Tshabong) using as a base Google Earth Imagery. We then used the ratio of the Tshabong population size to its area to calculate the calibrated expected spatial extent (area) of all other settlements based on their population number provided by the census. Using the estimated settlement area, we calculated a radius of a circle, reflecting spatial extent of the settlement weighted by human population. Over each settlement we generated a buffer with the given radius and we then rasterized this buffer at 30m resolution and turned it back to a point layer reflecting spatially distributed population size per settlement (as in Elliot et al. 2014). These analyses were performed in ArcGIS 10.x (ESRI 2012).

2.3.2.3 Multiscale variables

We considered and tested seven spatial scales reflecting the extent of an impact each anthropogenic or environmental factor can have on wildlife. For point features (human population and artificial water points we calculated point-based kernel density (km^2) of search radius: 250m, 500m, 1km, 2km, 4km, 8km, 16km. Due to importance and suspected wider range of kraals impact we calculated an additional radius of 32km for fenced and non-fenced kraals. The selection of scales to be tested follows other recently published multi-scale modeling studies (e.g. Mateo-Sanchez et al. 2014, Cushman et al. 2017) The scale range allows for multiscale assessment of how variables affect movement and occurrence at different scales (e.g. McGarigal et al. 2016) including fine scale habitat and movement path selection up to home range size for smaller species. These layers were produced in ArcGIS 10.x using *Kernel Density* tool.

In the case of factors represented as continuous raster variables we calculated focal mean by applying a gaussian kernel function. We determined the size of sigma (standard deviation of the Gaussian kernel) as half of each scale used to calculated point-based variables (given that 95% of a Gaussian distribution is within two standard deviations), thus: 125m, 250m, 500m, 1km, 2km, 4km, 8km. The focal mean layers were calculated using *focalMat* function in R *terra* package (R Core Team. 2021).

2.4 Habitat suitability models

We used the species density data and binary presence/absence data to test various modelling approaches of modelling habitat suitability for each of the species. These included random forest, linear regression and logistic regression. Random forest showed very good ability to explain the occurrence patterns along the transects; however, it performed poorly when predicting species occurrence outside of the transects and produced oddly patterned artifacts in the predicted probability of occurrence surface. Linear regression performed poorly fitting the density data. The best approach to model and predict habitat suitability with spoor data was logistic regression. We tested each of the three modelling approaches by applying three analytical steps described in detail below: scale optimization of each variable, multivariate model selection and model fit (e.g. Wasserman et al. 2012; Macdonald et al. 2019, 2020). However, since only the logistic regression produced high performing habitat suitability predictions across the study area for most species, below we will only describe this approach.

2.4.1 Scale-optimization and variables selection

To choose the best spatial scale (scale at which the considered factor affects strongest species habitat use) for each of the considered variables, we applied scale optimization using random forest (RF) Model Improvement Ratio (MIR) (Murphy et al. 2010). To do this we built a general random forest model including all variables at all scales (135 covariates) and applied a model selection approach using *rf.modelSel* function in *rfUtilities* R package (Murphy et al. 2010). This allowed us to identify the most parsimonious set of variables that maximizes model predictive ability while not including spurious relationships and statistical noise. This was done by selecting a RF model with the lowest error component based on the value of MIR (high MIR indicates most important covariates). For each tested variable we selected and retained the scale with the highest MIR value. We then ran the RF model with MIR model selection again to eliminate covariates (at their best scale) which were not adding to the model's parsimony and performance.

We also applied a correlation filter of 0.7 using Pearson's correlation to remove highly correlated predictor variables. For each pair of correlated predictor variables, we eliminated the one with lower MIR (calculated in the previous step of variables selection), starting from the variable with the highest MIR to ensure that the variables with the highest predictive performance were retained in the final set of variables used further on in the logistic regression.

2.4.2 Generalized linear models

To reduce autocorrelation and generate the most robust habitat suitability models using the entire dataset for each species we applied a bootstrapping approach similar to Wan et al. (2019). This involved randomly sampling a subset of the data 50 times, each time fitting the logistic regression function. The final model was produced by averaging the 50 bootstrapped runs, each independently validated.

To ensure that we maintained the correct proportion of presences and absences for each species (which, as shown by Cushman et al. 2017, is critical to unbiased estimates in SDM), in each bootstrap run, we sampled 20% of presences and 20% of absences for model training. The same proportion of presences and absences from the remaining data (independent validation) were then used for the validation of each bootstrap run. For five species (hartebeest, porcupine, stripped polecat, bat eared fox and cape fox) we were unable to sample 20% of the data due to high imbalance between presences and absences (Table 1). Therefore, for hartebeest, porcupine and striped polecat we sampled 15% of presences and 15% of absences, and for the remaining two species, with even lower proportion of absences, we sampled 10% of absences and presences.

Due to extremely high imbalance between presences and absences we were not able to run logistic regression for several species. Irregularly occurring species such as baboon, giraffe and secretary bird had too few and clustered presences (<500). Conversely, at the extreme end of the abundance spectrum springhare, jackal, steenbok and hare could not be modelled effectively due to too few absences (<560) in the back-transformed binary data (Table 1). In the case of the latter four species, we tested linear regression on the density data but results were not satisfactory. These 4 species can be considered common and ubiquitous throughout the Kalahari. Although their abundance precluded modelling their probabilities of occurrence in relation covariates across the landscape, we displayed these 4 common species' occurrence probabilities in relation to kraals (cattlepost) density for relative comparison with the other species (see section 3.2). This left us with total of 32 species for which we could model and predict habitat suitability.

For each bootstrap iteration we then fitted a binomial generalized linear regression with logit as a link function using glm algorithm in R stats package (R Core Team. 2021). The final habitat suitability model represented an average of the 50 model coefficients for each model variable. We also calculated averaged *p*-value for each variable and standard error of the coefficients and p-value across the 50 bootstrapped runs.

2.4.3 Model fit and prediction performance

We measured the model fit for each species by calculating deviance explained for each of the glm runs and then averaging these values and calculating standard deviation across all 50 runs.

Each bootstrapped glm model was also independently validated. We calculated several measures to assess predictive performance of each model: Kappa, PCC (Percent Correctly Classified), sensitivity, specificity, TSS (True Skill Statistic) and AUC (Area Under the ROC Curve) using the independent hold-out data. The PCC represents model accuracy calculated based on classification table (confusion matrix) and it ranges from 0-1 with higher values representing higher accuracy of the model prediction. Sensitivity measures the number of recorded present sites correctly predicted and divided by the total number of recorded present sites. Specificity measures the same but for absences. Kappa statistic summarizes all the information in the confusion matrix and it can be interpreted as the percent assignment between presences and absences better than chance. TSS is considered a superior measure of model accuracy which corrects the dependence of Kappa on prevalence, while maintaining all the advantages of Kappa. TSS ranges from -1 to 1, with TSS higher than zero representing models with performance better than random. Finally, AUC index (0-1), also independent of prevalence, is considered a highly effective measure of the model performance and is calculated as area under the receiver operating characteristic (ROC) curve. The curve is constructed by plotting all possible thresholds to classify the scores into confusion matrices (Pearce & Ferrier, 2000; Allouche et al. 2006). AUC ranges from 0.5 for random assignment of presences and absences to 1 for perfect assignment of presences and absences; as a rule of thumb it is often thought that AUC > 0.7 shows fair discrimination, > 0.8 good, and greater than 0.9 exceptional model performance. The final performance values for each of these performance matrices were averaged across all 50 bootstrapped runs.

We then generated habitat suitability maps for each 32 species as an average prediction of each 50 bootstrapped models. This was done by transforming z value (the direct value of logistic regression prediction) into p (probability of occurrence) by applying the following equation:

$$p = \exp(z)/(1 + \exp(z))$$

2.4.4 Biodiversity assessment

Based on the habitat suitability maps we generated three biodiversity maps to highlight particular combinations of species by adding the habitat suitability layers of individual species (*sensu* Penjor et al. 2021). The first map represented biodiversity of large mammals and consisted of: eland, gemsbok, wildebeest, hartebeest, kudu, springbok, ostrich, lion, leopard, cheetah, wild dog, spotted hyena, brown hyena. The second map calculated biodiversity of medium to small species, thus all remaining species. The last map represented the total biodiversity.

3 RESULTS and DISCUSSION

3.1 Habitat Suitability Models – species accounts

The following pages contain 32 species accounts in alphabetical order of common name, 1 per page.

Each page features:

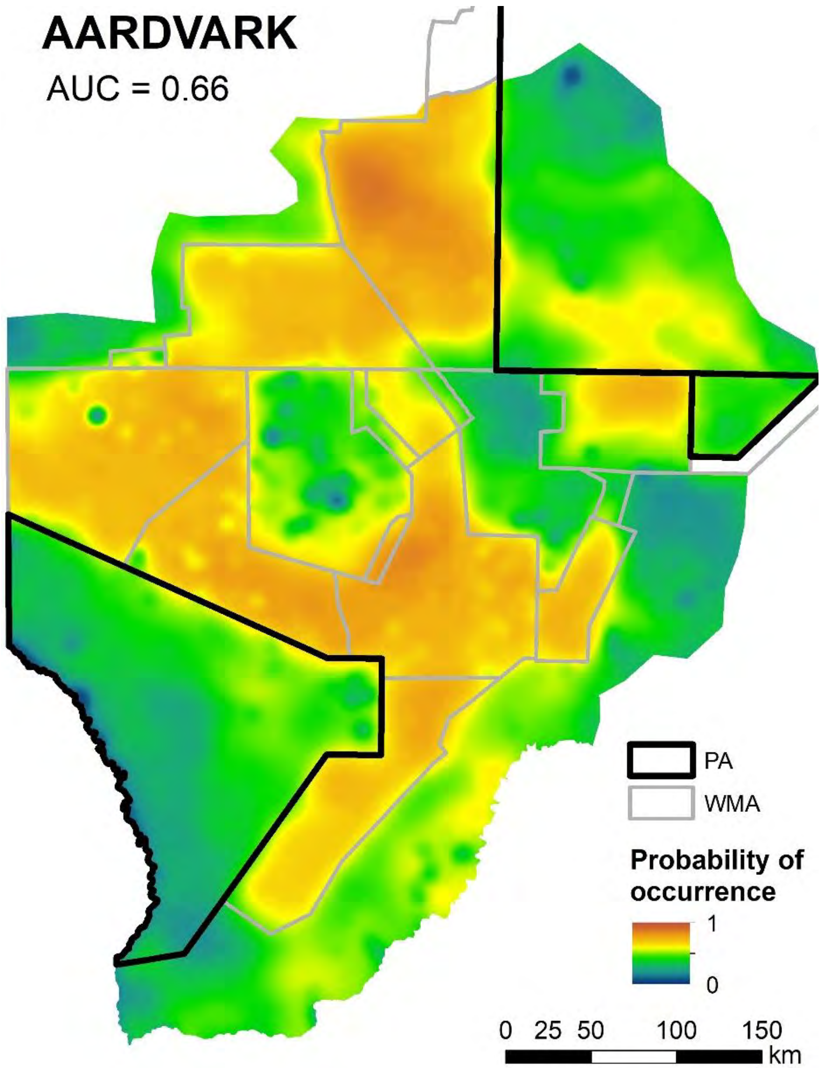
- Habitat suitability map displaying probability of occurrence surface across the AOI.

- Table displaying statistics describing prediction performance of the final species model, averaged from 50 bootstrapped model runs.

- Table displaying the best fit variables comprising the final model selected over 50 bootstrap runs. The strength of each variable is shown along with statistical significance indicated in bold and with asterisks at $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***).

- Where space permitted, plsmo splines (locally weighted regression splines, LOWESS) of responses to top variables ranked according to model improvement ratio (MIR). Note that some variables with the highest MIR where not selected for the final model by random forest selection process (i.e. for some species splines are presented for variables not in final model). Table 2 provides a key to variable abbreviations displayed on x-axes of plsmo plots.

Note that probability of occurrence is tightly correlated with abundance/population density, i.e. abundance typically increases proportionally as probability of occurrence increases.

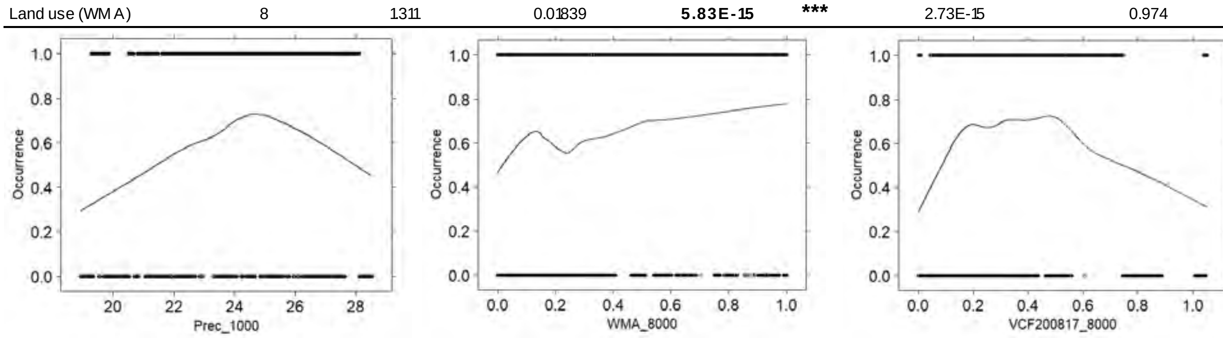


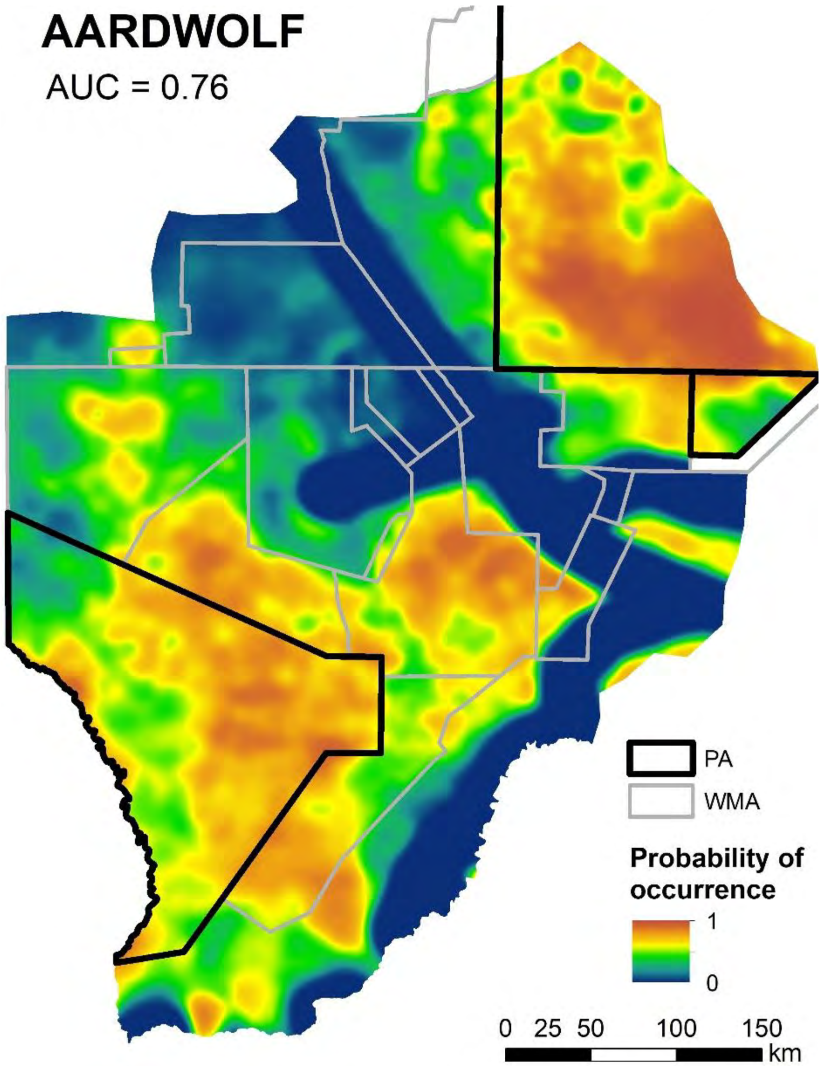
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.66	0.01
Kappa statistic	0.26	0.03
Sensitivity	0.48	0.09
Specificity	0.77	0.09
True skill statistic (TSS)	0.26	0.03
Percent correctly classified (PCC)	0.63	0.01
Deviance explained	0.93	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-11577	0.63696	0.147		0.022	
Pans	4	-20.946	0.77845	**0.021**	*	0.005	0.560
Roads (sand)	8	-17.919	1.13521	0.173		0.029	0.642
Soil (bulk density)	4	0.045	0.00395	0.358		0.041	0.583
Soil (clay content)	8	0.004	0.00046	0.413		0.038	0.875
Soil (Nitrogen)	8	-0.006	0.00016	**0.002**	**	0.001	0.665
Soil (sand content)	8	0.009	0.00035	**0.041**	*	0.010	0.694
Soil (organic carbon)	8	0.029	0.00177	0.134		0.024	0.789
Vegetation continuous field	8	0.294	0.05778	0.405		0.041	0.924
Land use (WMA)	8	1311	0.01839	**5.83E-15**	***	2.73E-15	0.974



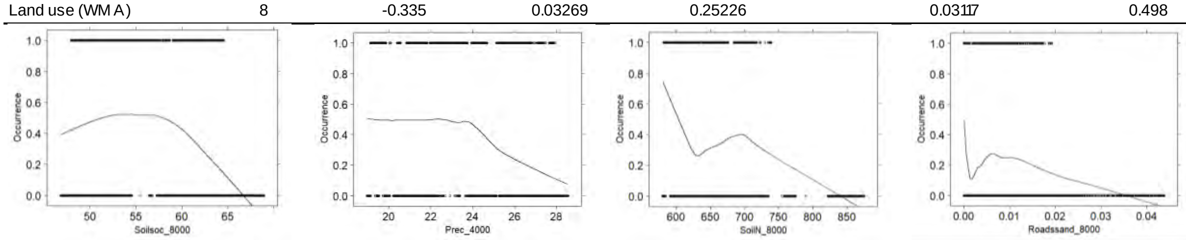


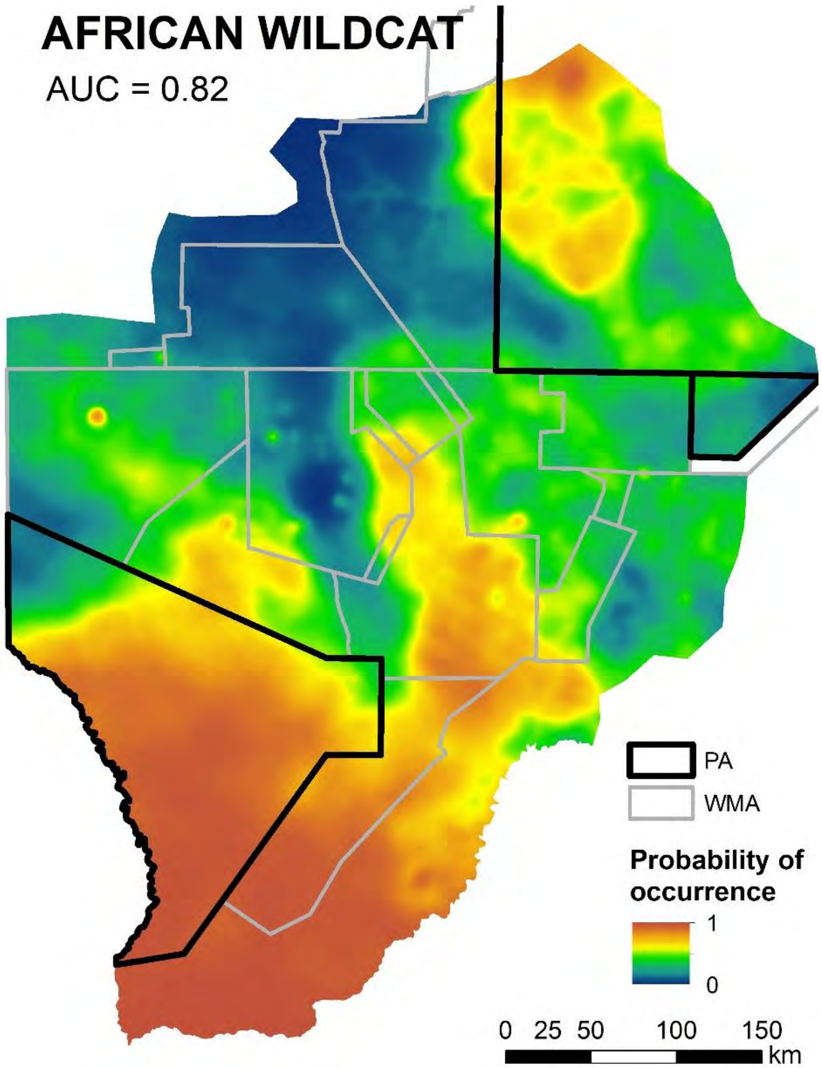
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.76	0.01
Kappa statistic	0.40	0.02
Sensitivity	0.78	0.18
Specificity	0.63	0.18
True skill statistic (TSS)	0.40	0.02
Percent correctly classified (PCC)	0.70	0.01
Deviance explained	0.80	0.02

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-9.099	122005	0.36968		0.04031	
Kraals	32	11749	0.77247	0.16831		0.03091	0.547
Pans	8	51444	2.49437	**0.04621**	*	0.01388	0.644
Land use (CGA)	8	-2.130	0.04283	**2.21E-05**	***	172E-05	0.326
Roads (calcrete)	8	-45.380	5.53038	0.39915		0.04105	0.130
Roads (paved)	8	-3625.470	100.11468	**0.00022**	***	4.31E-05	0.069
Roads (sand)	8	77.797	4.09846	0.05335		0.01500	0.701
Soil (bulk density)	4	-0.101	0.00753	0.18074		0.02969	0.532
Soil (clay content)	8	0.028	0.00082	**0.00061**	***	0.00038	0.653
Soil (Nitrogen)	8	-0.004	0.00032	0.20603		0.03630	0.708
Soil (sand content)	4	0.036	0.00074	**5.7E-10**	***	4.68E-10	0.614
Soil (organic carbon)	8	-0.050	0.00400	0.18252		0.03860	1000
Vegetation continuous field	8	-2.936	0.07998	**0.00073**	***	0.00046	0.601
Land use (WMA)	8	-0.335	0.03269	0.25226		0.03117	0.498



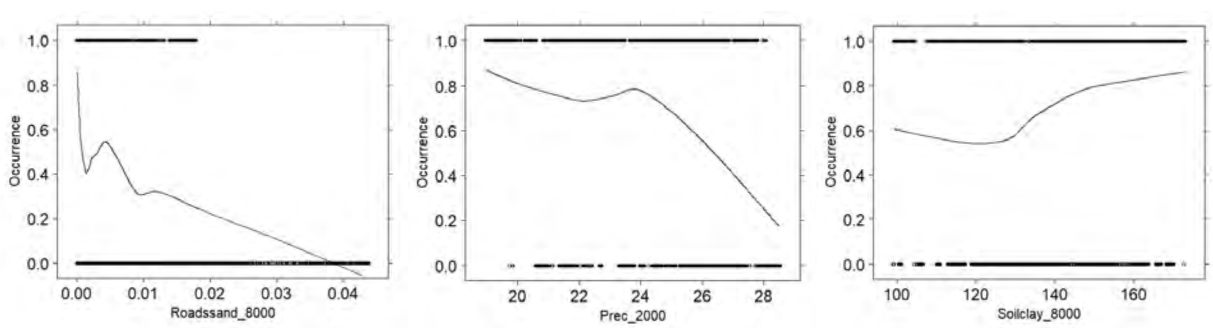


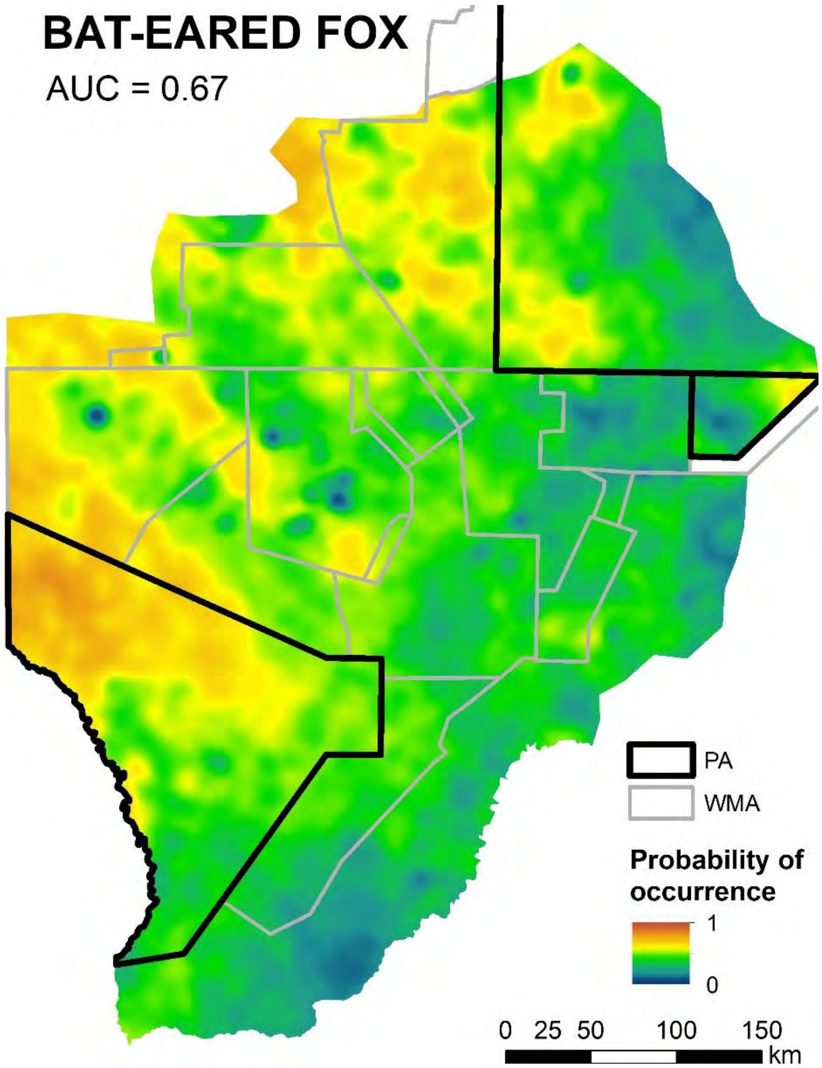
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.82	0.01
Kappa statistic	0.52	0.02
Sensitivity	0.72	0.03
Specificity	0.80	0.02
True skill statistic (TSS)	0.52	0.02
Percent correctly classified (PCC)	0.76	0.01
Deviance explained	0.77	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		5.16713	0.79982	0.40829		0.04261	
Precipitation	2	-0.57779	0.00563	**4.93E-33**	***	3.66E-33	0.810
Roads (sand)	8	-58.44541	2.08187	**0.00194**	**	0.00060	1000
Soil (bulk density)	4	-0.21363	0.00462	**6.81E-08**	***	5.84E-08	0.653
Soil (clay content)	8	0.06206	0.00055	**2.40E-35**	***	2.33E-35	0.654
Soil (sand content)	8	0.04333	0.00052	**1.82E-23**	***	132E-23	0.531



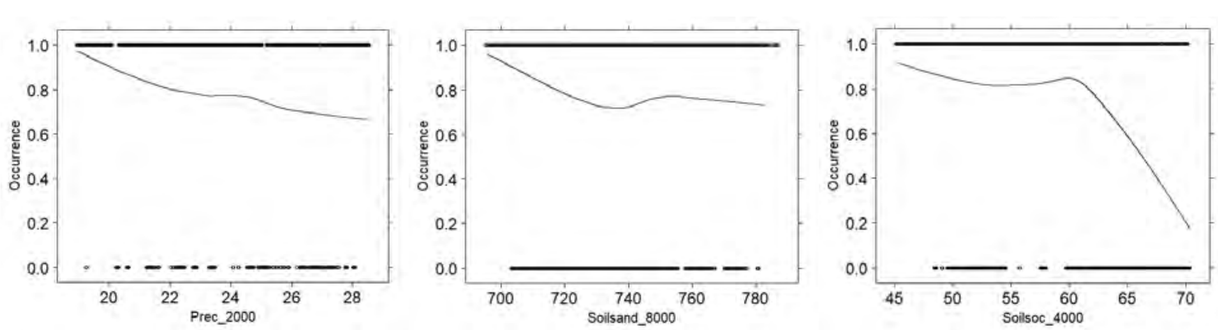


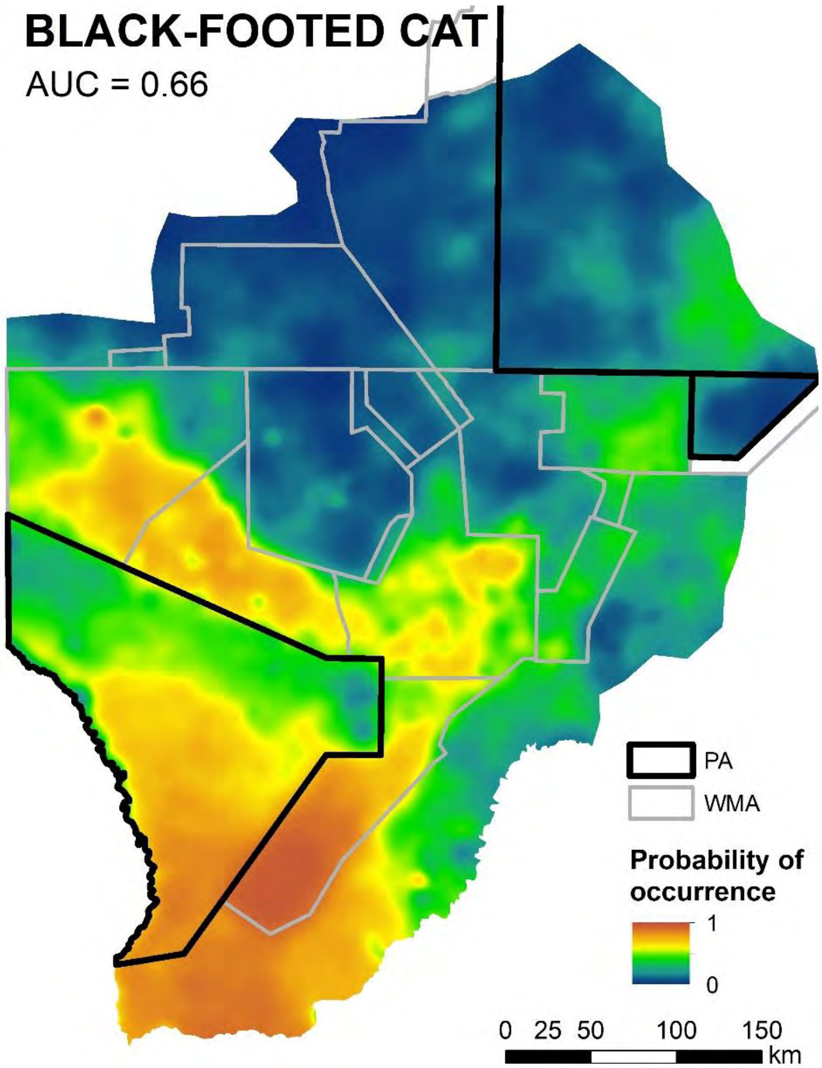
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.67	0.02
Kappa statistic	0.31	0.02
Sensitivity	0.90	0.05
Specificity	0.41	0.05
True skill statistic (TSS)	0.31	0.02
Percent correctly classified (PCC)	0.65	0.01
Deviance explained	0.94	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-23.367	0.51591	**0.00011**	***	5.78E-05	
Soil (bulk density)	4	0.219	0.00367	**4.57E-08**	***	2.85E-08	0.609
Soil (sand content)	8	-0.007	0.00031	0.06696		0.01570	0.738
Soil (organic carbon)	4	-0.065	0.00111	**1.16E-08**	***	7.45E-09	0.715



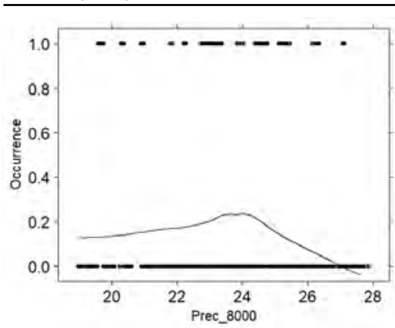


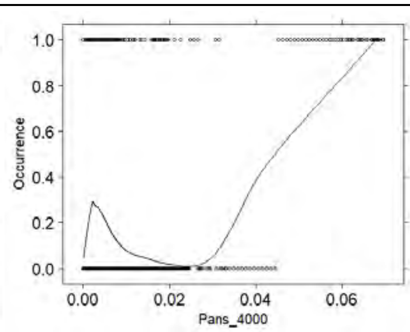
Model Prediction Performance

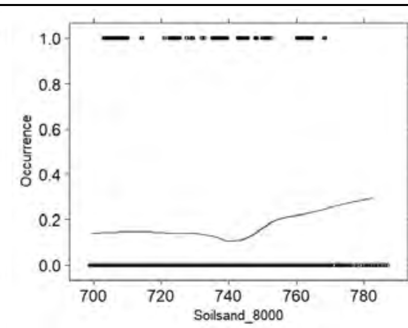
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.66	0.03
Kappa statistic	0.30	0.05
Sensitivity	0.77	0.15
Specificity	0.52	0.13
True skill statistic (TSS)	0.30	0.05
Percent correctly classified (PCC)	0.65	0.03
Deviance explained	0.90	0.02

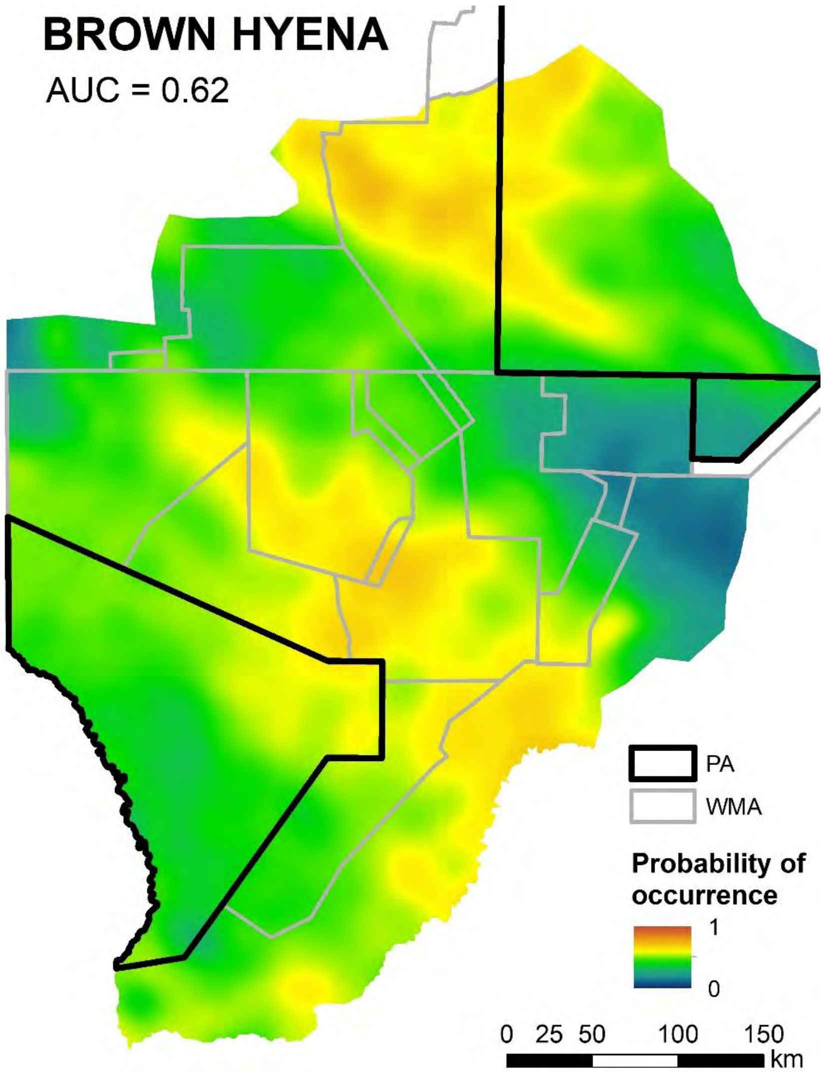
General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		34.64282	2.57557	0.18854		0.02971	
Pans	4	-27.71192	2.62857	0.26197		0.03488	0.852
Precipitation	8	-0.22070	0.01563	0.18232		0.03511	1.000
Soil (bulk density)	4	-0.26261	0.01465	0.09309		0.02714	0.517
Soil (clay content)	8	0.00710	0.00147	0.51438		0.04228	0.604
Soil (Nitrogen)	8	0.00111	0.00051	0.52065		0.03940	0.615
Soil (sand content)	8	0.01022	0.00141	0.38668		0.04007	0.683
Vegetation continuous field	8	-3.31418	0.18890	0.09995		0.02338	0.502
Land use (WMA)	8	1.80520	0.05463	**0.00112**	**	0.00064	0.560







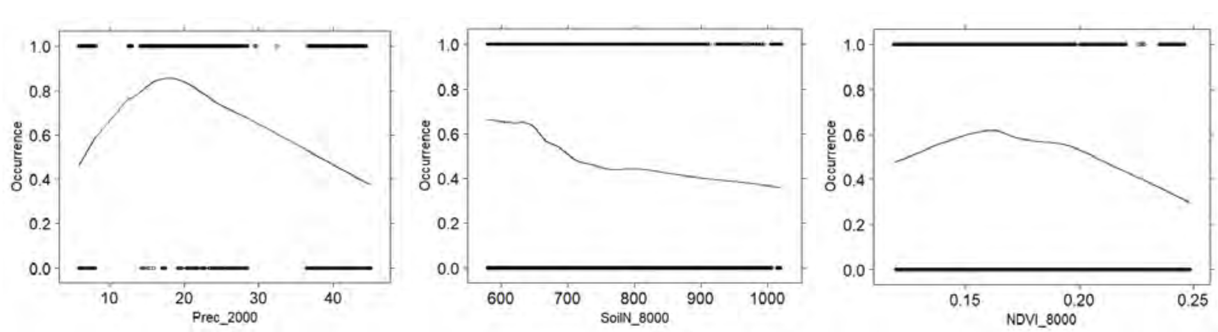


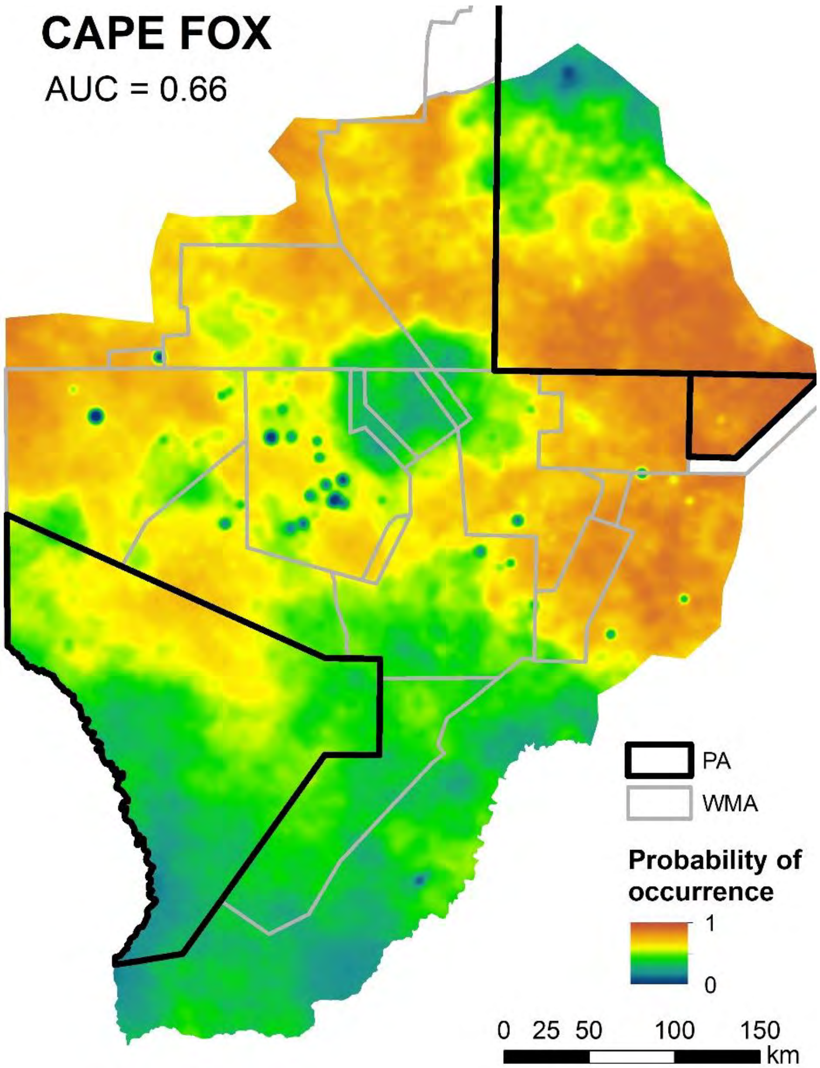
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.62	0.01
Kappa statistic	0.20	0.01
Sensitivity	0.58	0.06
Specificity	0.62	0.06
True skill statistic (TSS)	0.20	0.01
Percent correctly classified (PCC)	0.60	0.01
Deviance explained	0.97	0.00

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		2.46137	0.03754	**2.36E-13**	***	193E-13	
Vegetation greeness (NDVI)	8	6.71393	0.11096	**2.90E-07**	***	131E-07	0.606
Precipitation	2	-0.00319	0.00021	0.13548		0.02670	1000
Soil (clay content)	8	0.00417	0.00011	**0.00024**	***	0.00017	0.566
Soil (Nitrogen)	8	-0.00617	4.72E-05	**7.05E-63**	***	6.91E-63	0.773
Vegetation continuous field	8	0.38454	0.01218	**0.00316**	**	0.00104	0.589



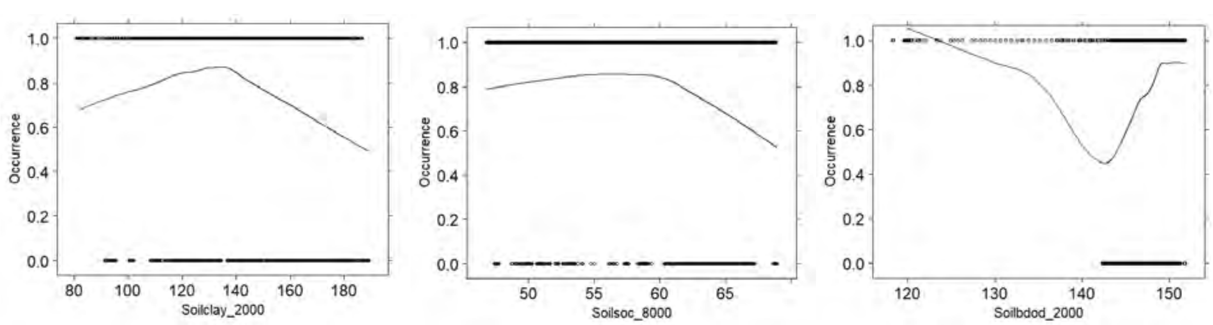


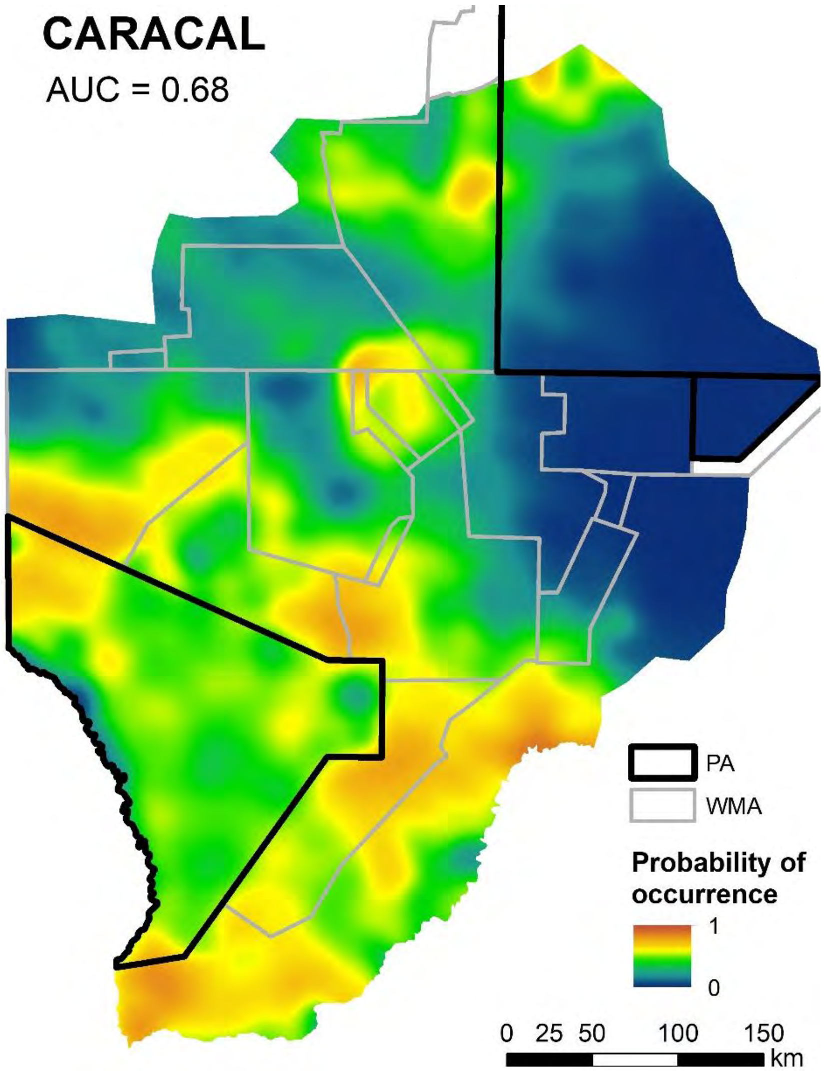
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.66	0.02
Kappa statistic	0.28	0.03
Sensitivity	0.61	0.05
Specificity	0.67	0.06
True skill statistic (TSS)	0.28	0.03
Percent correctly classified (PCC)	0.64	0.01
Deviance explained	0.95	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-3173244	114270	**1.30E-05**	***	8.28E-06	
Soil (bulk density)	2	0.16725	0.00713	**0.00025**	***	9.43E-05	0.736
Soil (clay content)	2	-0.00032	0.00045	0.53805		0.03898	1000
Soil (sand content)	8	0.01414	0.00051	**0.00411**	**	0.00202	0.699
Soil (organic carbon)	8	-0.05779	0.00172	**0.00056**	***	0.00028	0.969



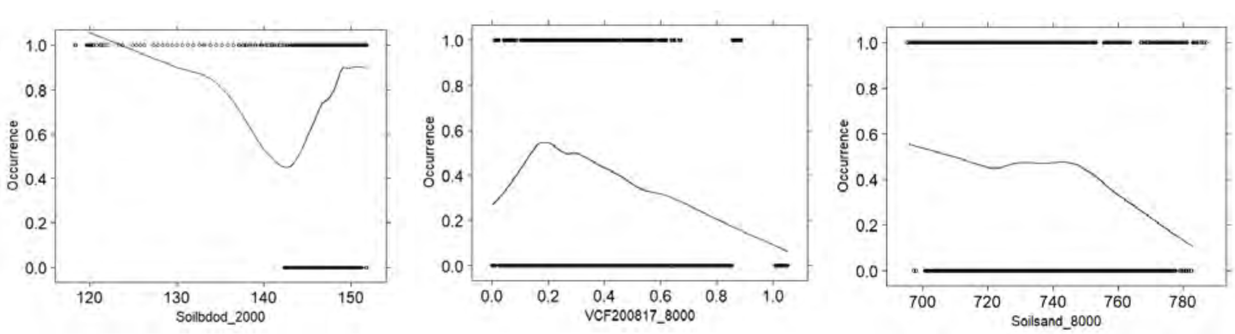


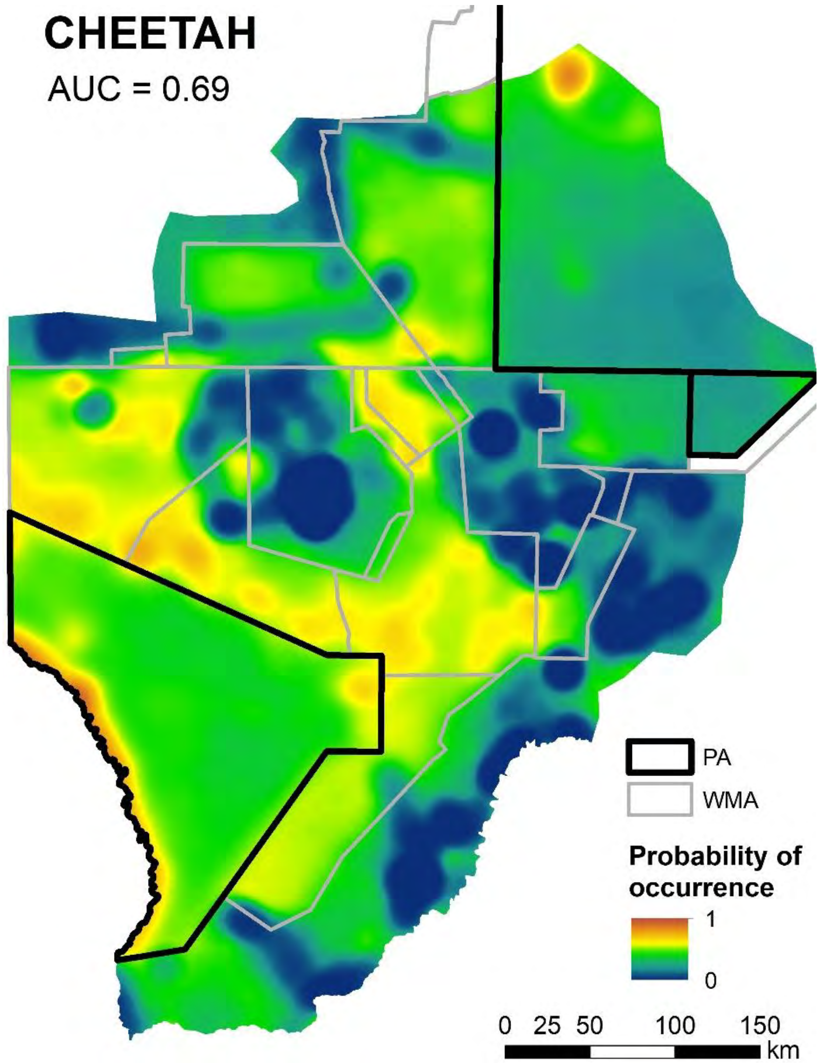
Model Prediction Performance

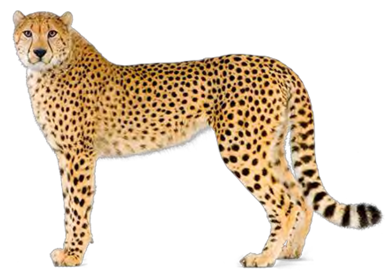
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.68	0.02
Kappa statistic	0.32	0.03
Sensitivity	0.52	0.09
Specificity	0.80	0.09
True skill statistic (TSS)	0.32	0.03
Percent correctly classified (PCC)	0.66	0.01
Deviance explained	0.93	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		24.02432	0.54004	**3.31E-06**	***	3.00E-06	
Land use (PA)	8	-0.86442	0.02471	**0.00328**	**	0.00226	0.744
Pans	8	-83.35296	180384	**2.17E-06**	***	110E-06	0.632
Soil (clay content)	8	0.00247	0.00040	0.49421		0.04145	1000
Soil (Nitrogen)	8	-0.00784	0.00022	**0.00047**	***	0.00027	0.746
Soil (sand content)	8	-0.02538	0.00061	**1.18E-05**	***	8.37E-06	0.856
Vegetation continuous field	8	0.50851	0.06379	0.33026		0.03355	0.857
Land use (WMA)	8	0.19702	0.02538	0.30556		0.04106	0.671





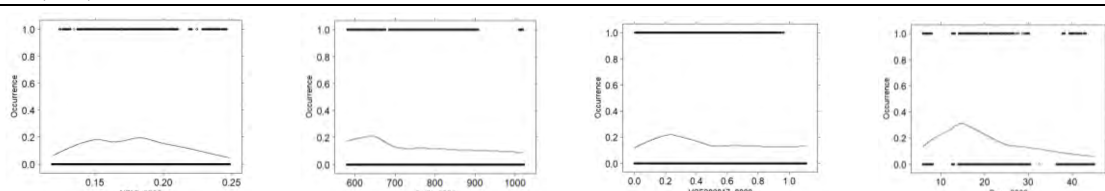


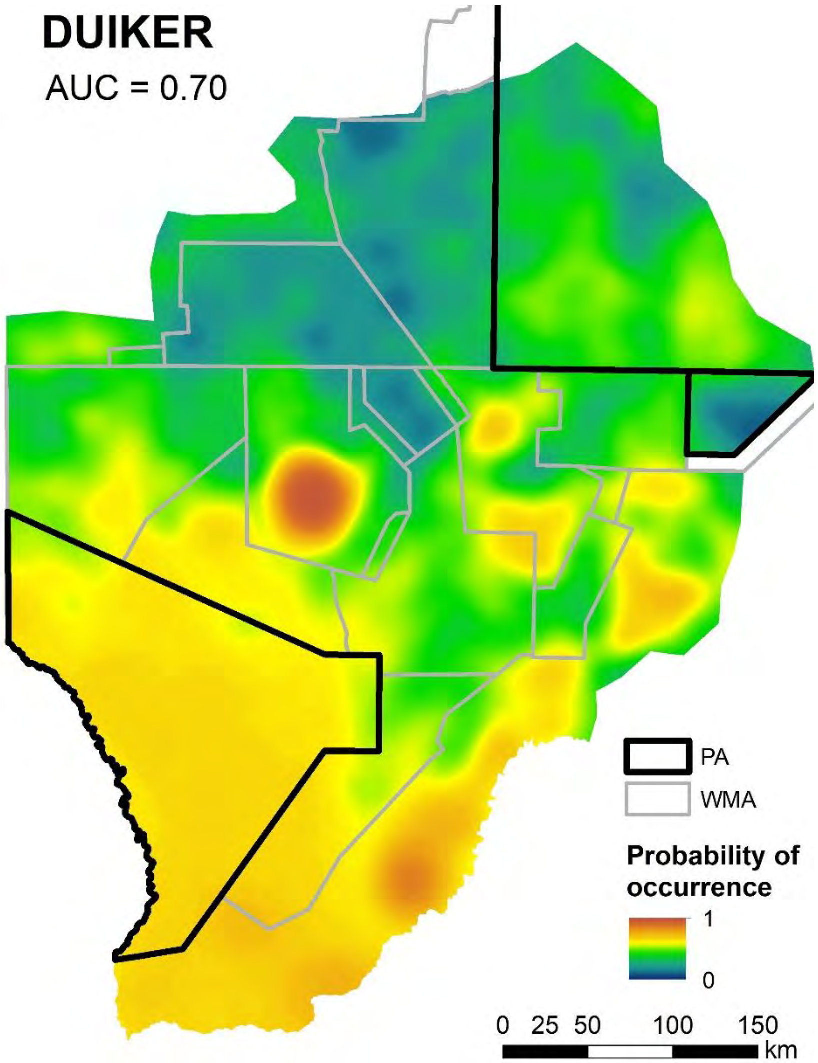
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.69	0.01
Kappa statistic	0.32	0.02
Sensitivity	0.75	0.06
Specificity	0.57	0.06
True skill statistic (TSS)	0.32	0.02
Percent correctly classified (PCC)	0.66	0.01
Deviance explained	0.90	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		10.39502	0.64701	0.12133		0.02774	
Kraals (inside fenced farms)	32	-4.58818	0.28427	0.16927		0.03080	0.289
Kraals	16	-4.78265	0.06204	**1.78E-15**	***	1.18E-15	0.501
Pans	8	56.74267	0.93455	**1.64E-10**	***	107E-10	0.662
Land use (CGA)	8	-0.06805	0.02891	0.49085		0.03324	0.359
Precipitation	2	-0.01572	0.00041	**0.00036**	***	0.00016	0.830
Roads (calcrete)	8	-246.03053	3.17345	**4.43E-14**	***	3.45E-14	0.381
Roads (paved)	8	62.13118	3.27976	0.06307		0.01550	0.078
Roads (sand)	8	35.25201	168573	**0.03987**	*	0.01595	0.428
Soil (bulk density)	4	-0.02894	0.00469	0.36551		0.04164	0.786
Soil (clay content)	8	-0.00256	0.00030	0.35517		0.04051	0.662
Soil (sand content)	8	-0.00774	0.00025	**0.00230**	**	0.00143	0.710
Soil (organic carbon)	8	-0.00590	0.00127	0.45883		0.03822	0.726
Vegetation continuous field	8	0.29856	0.02868	0.30612		0.04252	0.876
Land use (WMA)	8	0.78806	0.01961	**1.20E-05**	***	103E-05	0.663



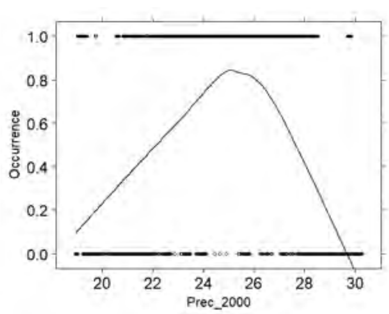


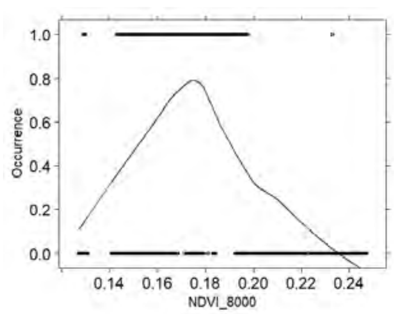
Model Prediction Performance

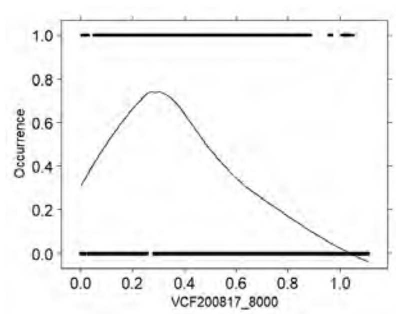
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.70	0.01
Kappa statistic	0.43	0.01
Sensitivity	0.93	0.01
Specificity	0.49	0.01
True skill statistic (TSS)	0.43	0.01
Percent correctly classified (PCC)	0.71	0.01
Deviance explained	0.88	0.01

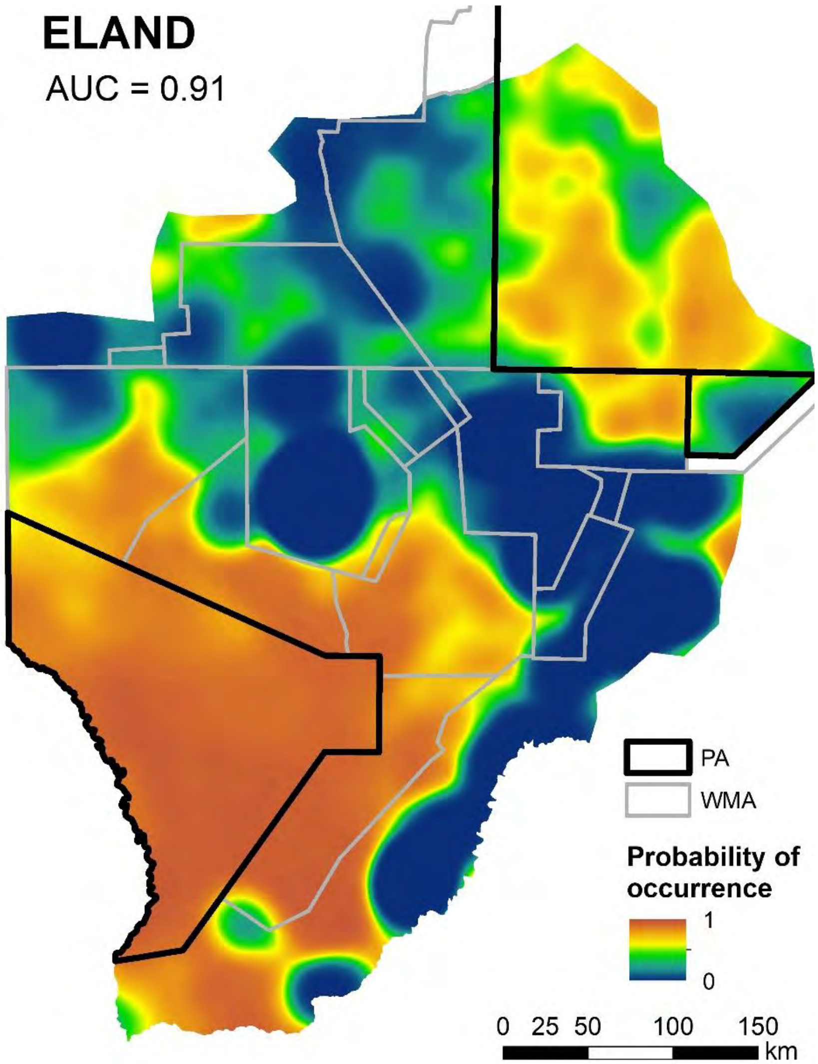
General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		0.73862	0.00835	**1.24E-19**	***	1.10E-19	
Kraals	32	0.89838	0.02313	**1.38E-10**	***	7.73E-11	0.411
Vegetation continuous field	8	-2.40388	0.01791	**6.37E-53**	***	5.93E-53	0.672







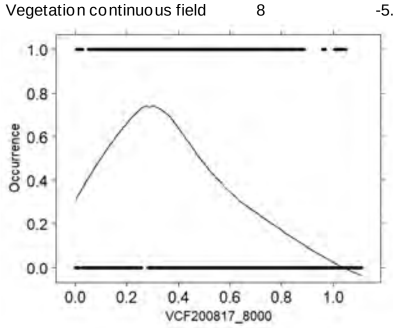


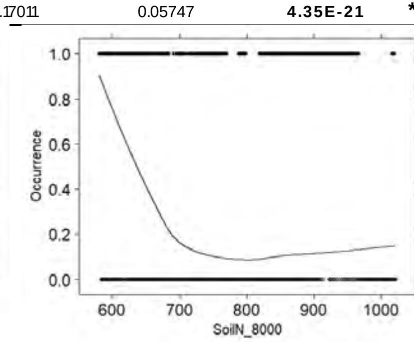
Model Prediction Performance

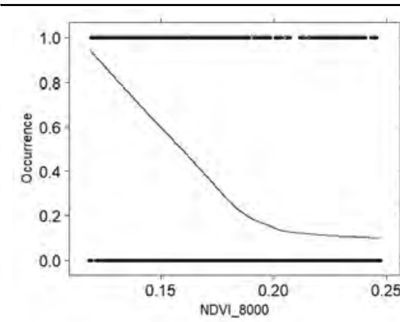
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.91	0.01
Kappa statistic	0.70	0.02
Sensitivity	0.83	0.01
Specificity	0.88	0.02
True skill statistic (TSS)	0.70	0.02
Percent correctly classified (PCC)	0.85	0.01
Deviance explained	0.55	0.01

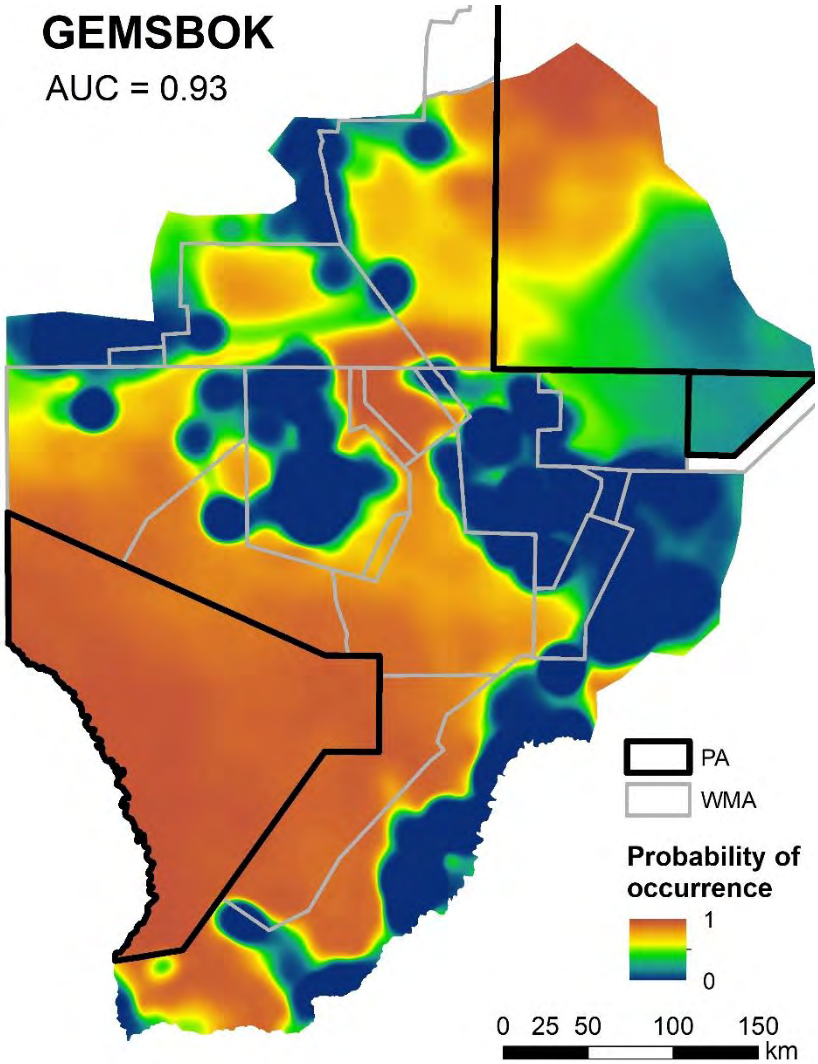
General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-171569	144325	0.52695		0.03708	
Kraals (inside fenced farms)	32	-9.62099	0.52534	0.11601		0.02628	0.533
Kraals	32	-15.10741	0.14357	**4.26E-33**	***	2.54E-33	0.692
Precipitation	2	0.03209	0.00067	**1.68E-05**	***	144E-05	0.765
Roads (calcrete)	8	-57.33275	5.71275	0.24393		0.03674	0.454
Soil (bulk density)	4	-0.00500	0.00864	0.56116		0.04000	0.508
Soil (clay content)	4	0.00414	0.00067	0.41326		0.03958	0.513
Soil (sand content)	8	-0.00087	0.00046	0.61397		0.04014	0.716
Soil (organic carbon)	8	0.08029	0.00232	**0.00024**	***	9.72E-05	0.517
Vegetation continuous field	8	-5.17011	0.05747	**4.35E-21**	***	2.31E-21	1000







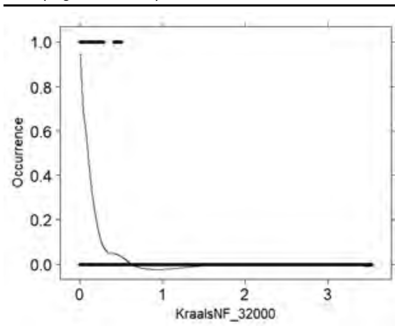


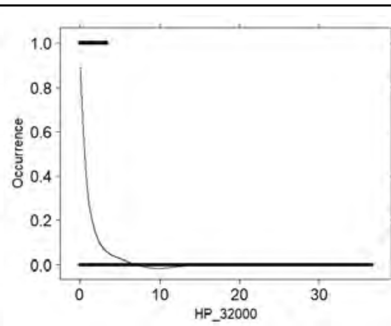
Model Prediction Performance

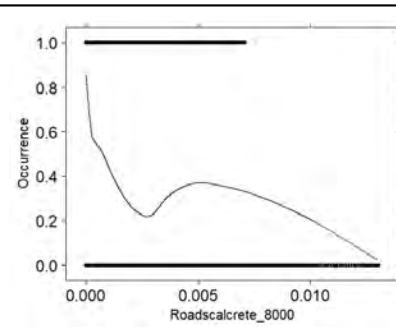
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.93	0.00
Kappa statistic	0.69	0.01
Sensitivity	0.86	0.04
Specificity	0.84	0.04
True skill statistic (TSS)	0.69	0.01
Percent correctly classified (PCC)	0.85	0.00
Deviance explained	0.52	0.01

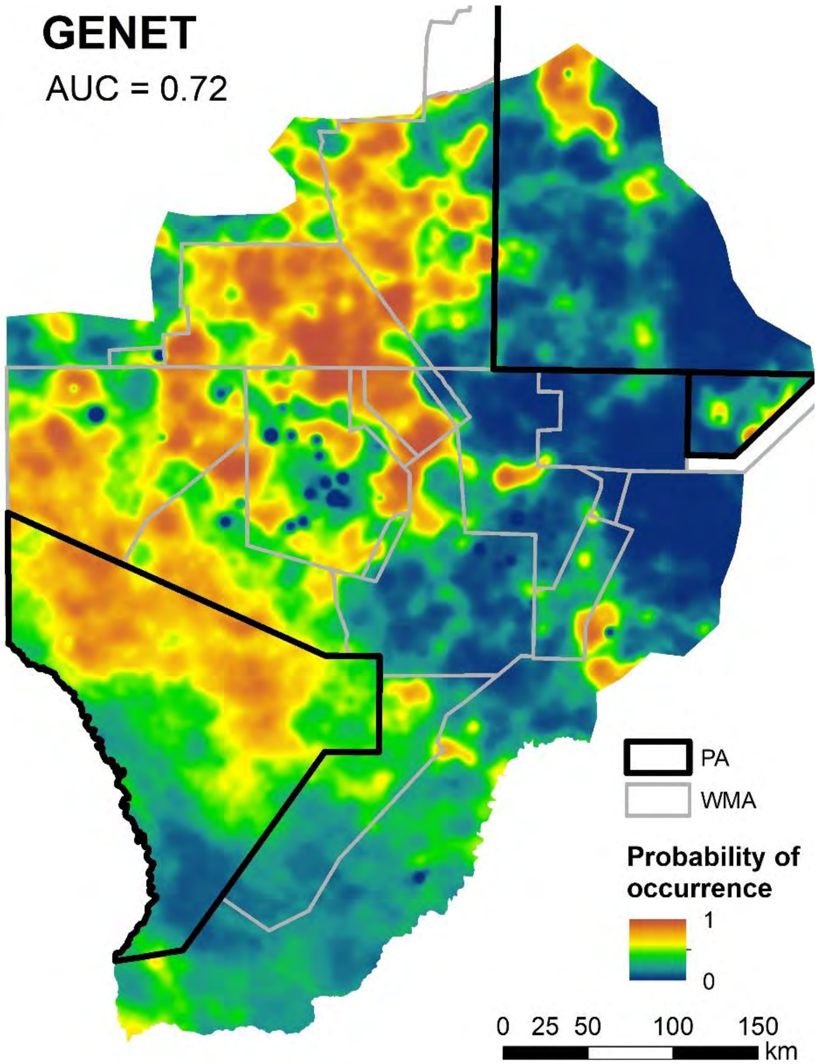
General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		13.50091	0.14472	**2.12E-14**	***	9.95E-15	
Kraals (inside fenced farms)	32	-34.09650	0.26945	**1.44E-43**	***	1.41E-43	0.470
Kraals	16	-18.86751	0.12272	**5.52E-170**	***	0	1000
Vegetation greeness (NDVI)	8	-8.85758	0.15367	**1.55E-07**	***	1.17E-07	0.453
Precipitation	8	-0.02191	0.00038	**1.91E-08**	***	1.32E-08	0.476
Roads (calcrete)	8	-266.85987	1.73294	**9.01E-52**	***	8.26E-52	0.663
Soil (clay content)	8	0.01266	0.00023	**7.38E-05**	***	5.02E-05	0.366
Soil (sand content)	8	-0.01741	0.00021	**3.34E-13**	***	1.67E-13	0.449
Soil (organic carbon)	8	0.03150	0.00114	**0.01088**	*	0.00571	0.388







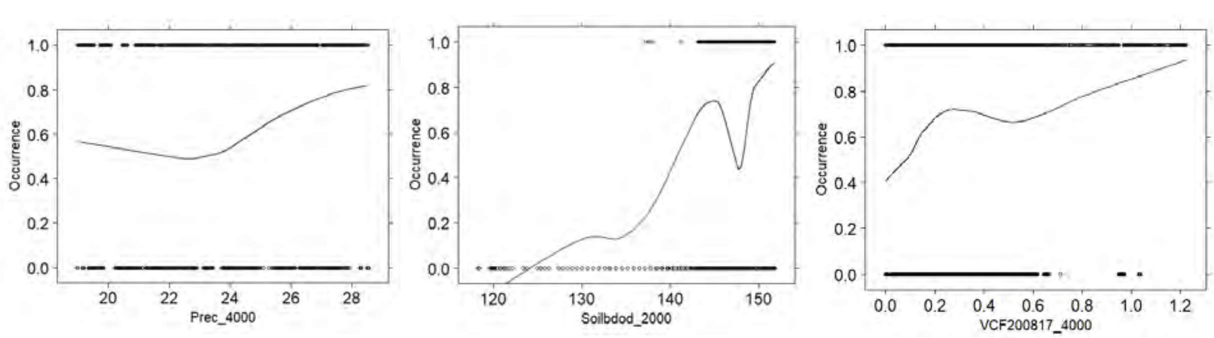


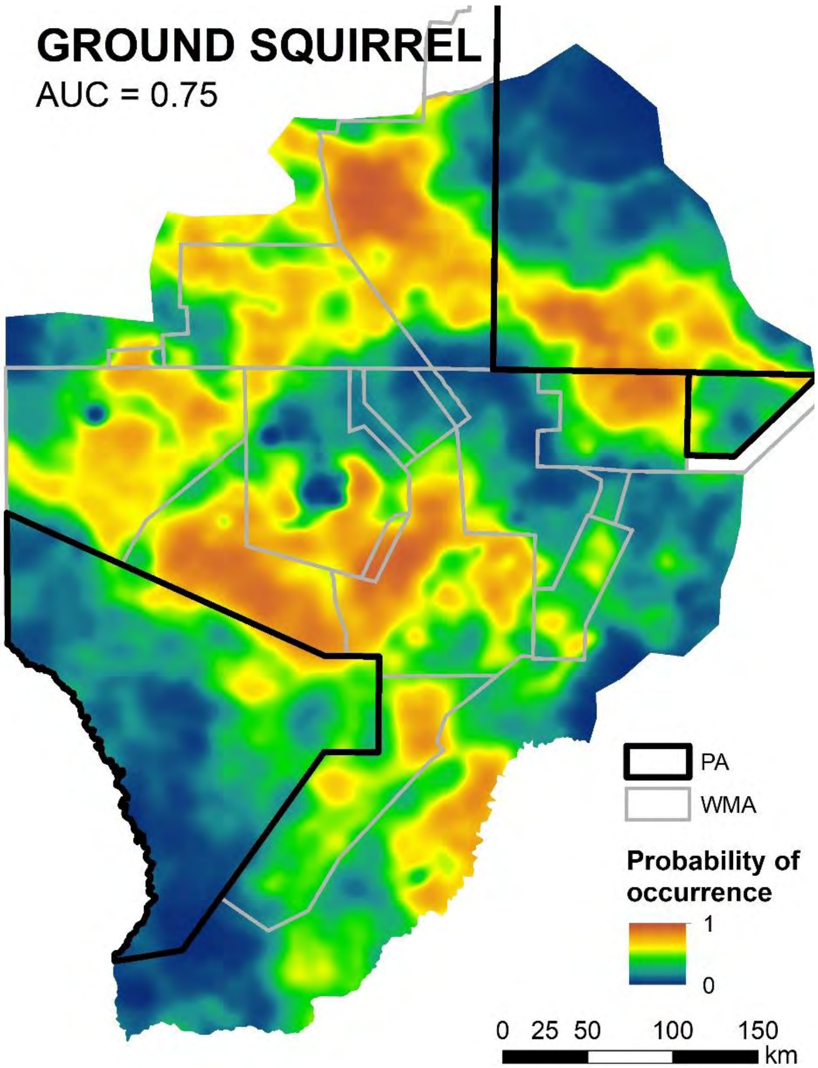
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.72	0.01
Kappa statistic	0.36	0.02
Sensitivity	0.67	0.06
Specificity	0.69	0.06
True skill statistic (TSS)	0.36	0.02
Percent correctly classified (PCC)	0.68	0.01
Deviance explained	0.90	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-46.25847	0.90247	**7.26E-11**	***	3.51E-11	
Vegetation greeness (NDVI)	8	-45.59596	0.66940	**6.52E-13**	***	5.04E-13	0.647
Soil (bulk density)	2	0.33492	0.00598	**6.67E-12**	***	4.74E-12	0.786
Soil (clay content)	8	0.02308	0.00053	**1.92E-08**	***	9.27E-09	0.610
Soil (Nitrogen)	8	-0.01038	0.00023	**6.98E-09**	***	5.73E-09	0.542
Soil (sand content)	4	0.00874	0.00033	**0.01179**	*	0.00544	0.617
Vegetation continuous field	4	5.96151	0.07979	**8.42E-24**	***	3.99E-24	0.671



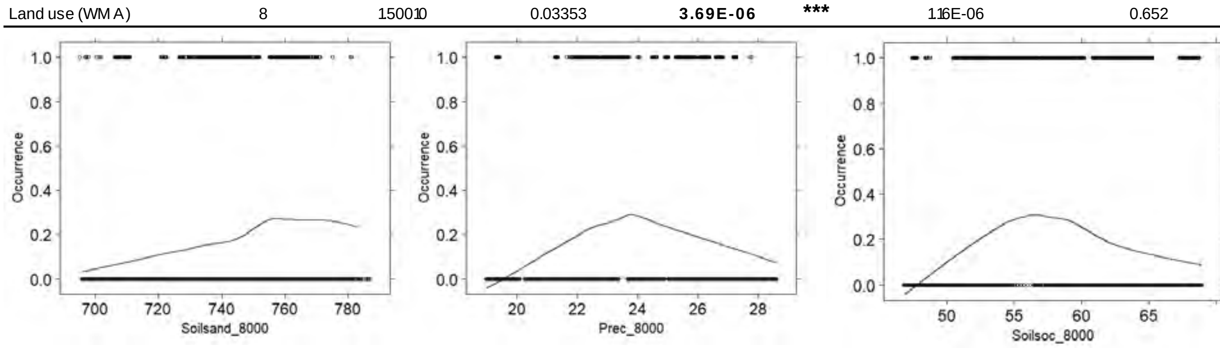


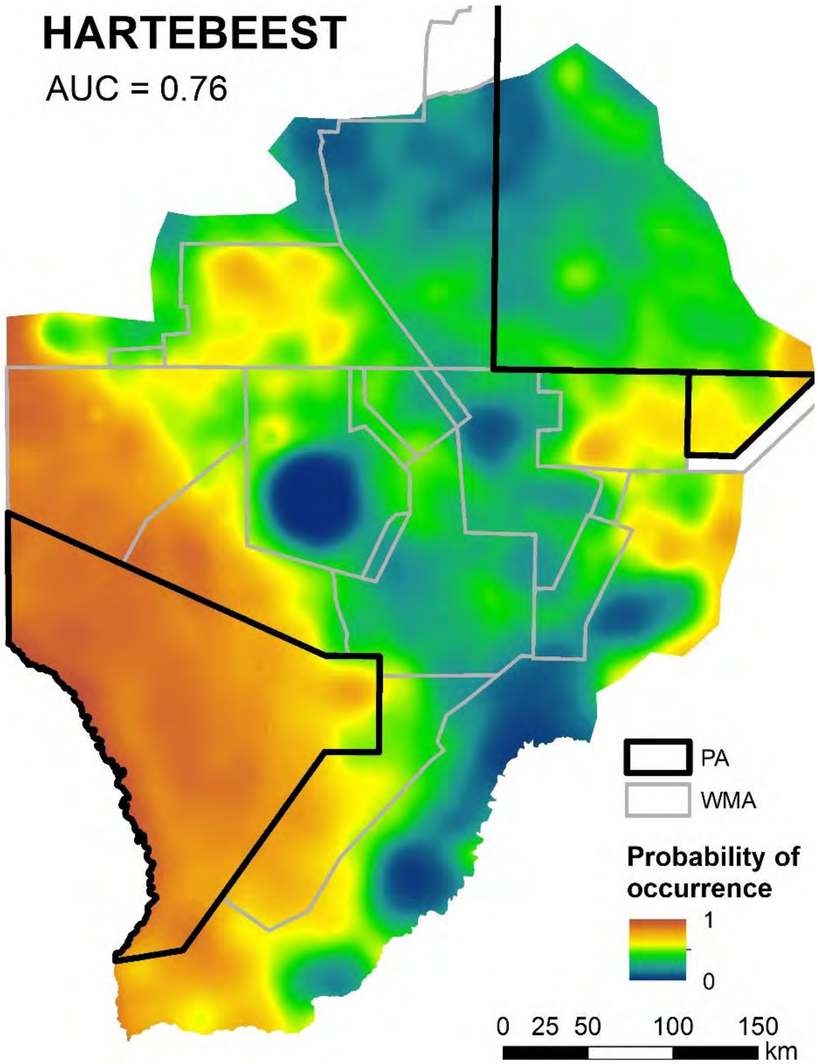
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.75	0.02
Kappa statistic	0.43	0.04
Sensitivity	0.78	0.07
Specificity	0.65	0.07
True skill statistic (TSS)	0.43	0.04
Percent correctly classified (PCC)	0.71	0.02
Deviance explained	0.82	0.02

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-78.36857	1.95183	**1.34E-05**	***	6.13E-06	
Pans	8	31.33744	3.53400	0.33038		0.03903	0.349
Roads (sand)	8	-115.03926	2.91257	**0.00120**	**	0.00040	0.359
Soil (bulk density)	4	0.41761	0.01136	**6.46E-05**	***	1.75E-05	0.523
Soil (clay content)	8	0.00707	0.00139	0.45165		0.04508	0.516
Soil (Nitrogen)	4	-0.01275	0.00039	**0.00171**	**	0.00065	0.793
Soil (sand content)	8	0.02264	0.00110	**0.04080**	*	0.01487	1.000
Soil (organic carbon)	8	0.11876	0.00437	**0.00500**	**	0.00137	0.865
Vegetation continuous field	8	-1.79786	0.11181	0.10119		0.02538	0.739
Land use (WMA)	8	1.50010	0.03353	**3.69E-06**	***	1.16E-06	0.652



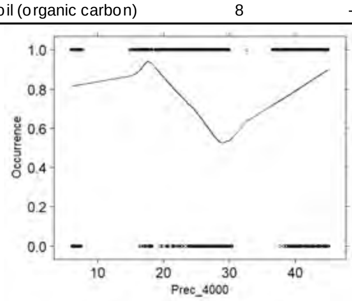


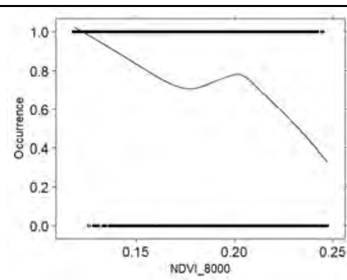
Model Prediction Performance

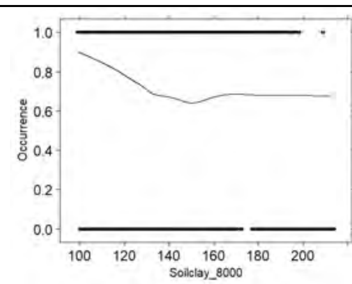
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.76	0.01
Kappa statistic	0.40	0.01
Sensitivity	0.90	0.01
Specificity	0.50	0.01
True skill statistic (TSS)	0.40	0.01
Percent correctly classified (PCC)	0.70	0.00
Deviance explained	0.83	0.00

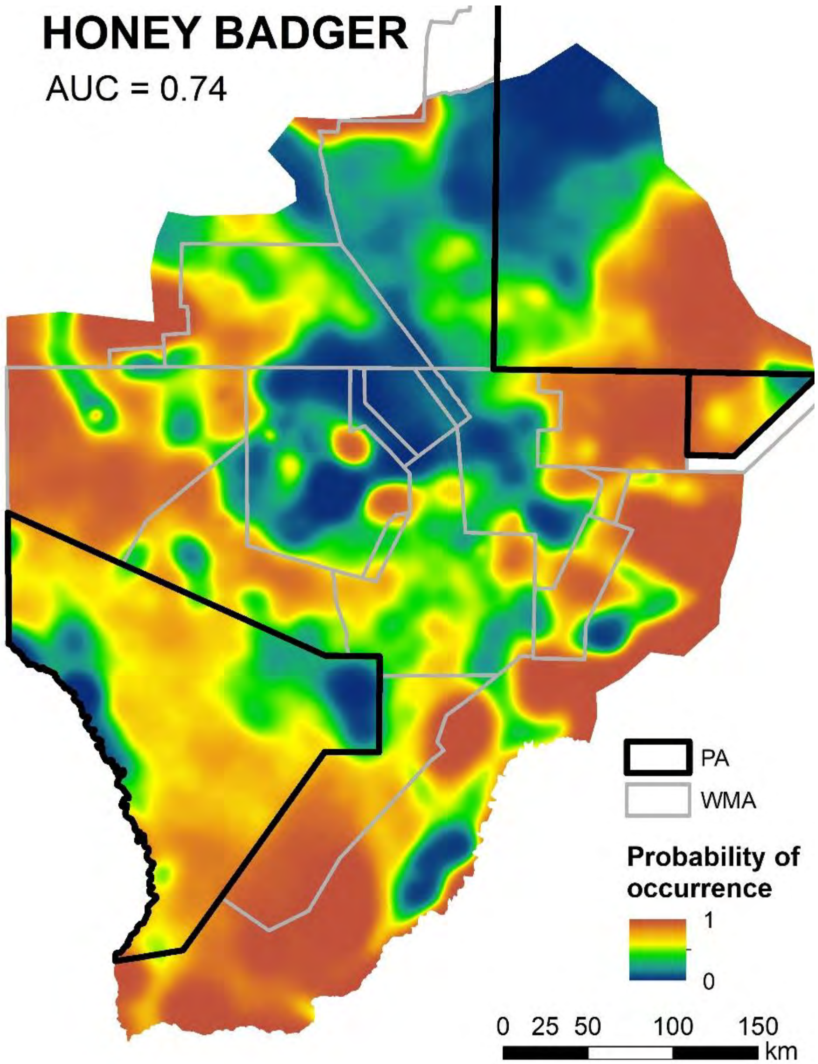
General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-5.03925	0.32694	0.18287		0.03454	
Kraals (inside fenced farms)	32	-8.94094	0.14697	**3.26E-06**	***	**1.46E-06**	0.397
Kraals	32	-2.38391	0.01294	**5.73E-61**	***	**5.51E-61**	0.637
Vegetation greeness (NDVI)	8	-26.60666	0.12663	**6.76E-94**	***	**6.60E-94**	0.985
Pans	8	59.65452	0.70521	**3.80E-13**	***	**2.71E-13**	0.566
Precipitation	4	-0.02778	0.00029	**2.11E-24**	***	**1.39E-24**	1.000
Soil (bulk density)	4	0.06318	0.00191	**0.00419**	**	**0.00160**	0.493
Soil (clay content)	8	-0.00633	0.00021	**0.00810**	**	**0.00498**	0.740
Soil (Nitrogen)	8	0.00580	5.31E-05	**6.45E-40**	***	**6.32E-40**	0.698
Soil (sand content)	8	0.00045	0.00018	0.53982		0.04068	0.572
Soil (organic carbon)	8	-0.03386	0.00081	**6.89E-05**	***	**3.12E-05**	0.680







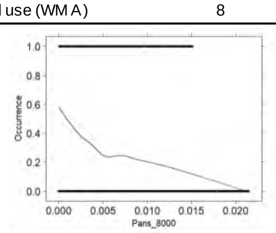


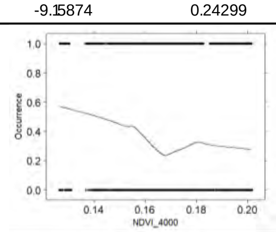
Model Prediction Performance

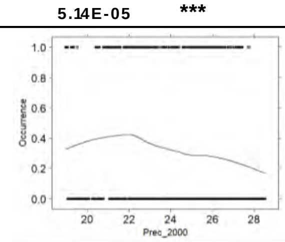
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.74	0.02
Kappa statistic	0.39	0.02
Sensitivity	0.86	0.06
Specificity	0.54	0.06
True skill statistic (TSS)	0.39	0.02
Percent correctly classified (PCC)	0.70	0.01
Deviance explained	0.84	0.01

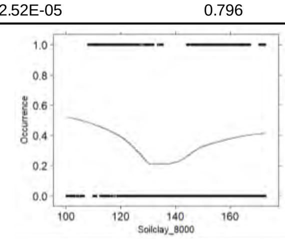
General Linear Model from 50 bootstrap runs

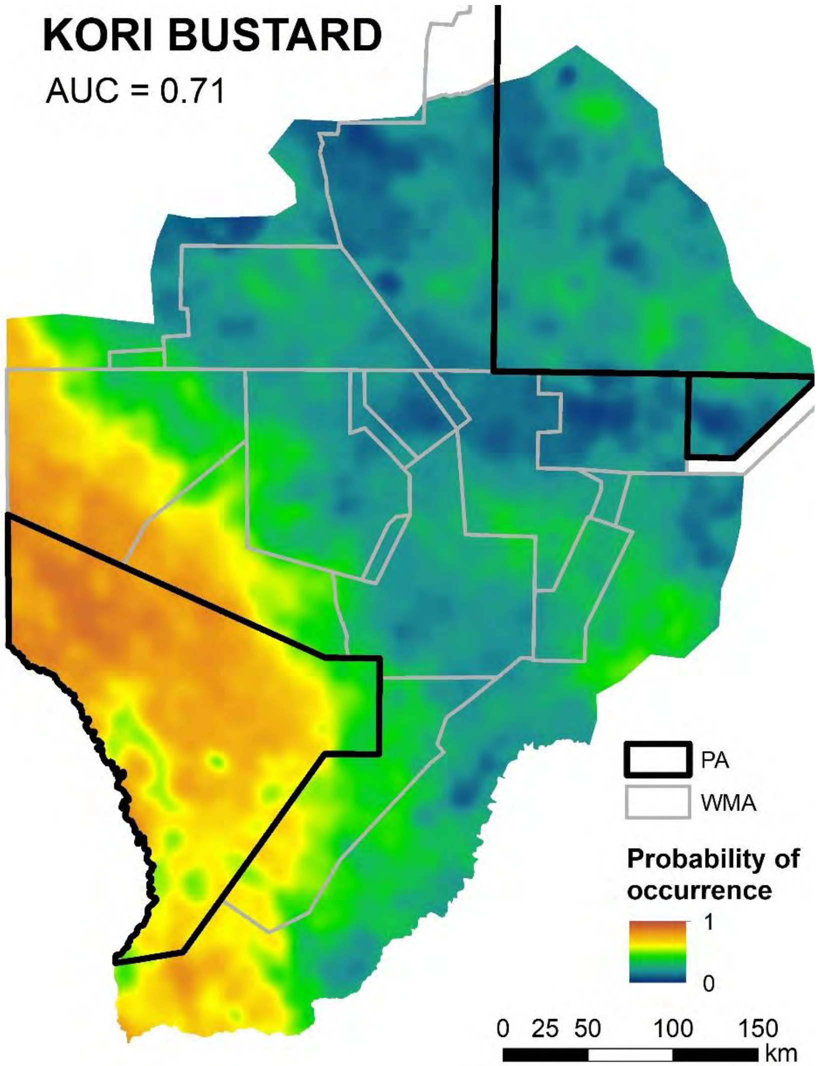
Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		54.12511	1.38449	**1.07E-05**	***	3.25E-06	
Kraals (inside fenced farms)	32	-7.82635	0.62061	0.23414		0.03229	0.218
Land use (PA)	8	-10.62171	0.24954	**3.01E-06**	***	1.57E-06	0.587
Pans	8	-258.59245	2.85245	**8.59E-22**	***	5.38E-22	1.000
Land use (CGA)	8	-9.99312	0.24093	**5.23E-06**	***	2.72E-06	0.322
Roads (calcrete)	8	113.66397	5.38519	0.05772		0.01122	0.289
Roads (paved)	8	-233.83405	10.19031	**0.03776**	*	0.01068	0.079
Roads (sand)	8	-48.29935	3.15064	0.15727		0.03145	0.837
Soil (bulk density)	4	-0.29126	0.00808	**2.17E-05**	***	7.52E-06	0.753
Soil (clay content)	8	-0.04540	0.00070	**8.65E-11**	***	6.96E-11	0.880
Soil (Nitrogen)	8	0.00419	0.00022	0.10004		0.02188	0.655
Soil (sand content)	8	0.00548	0.00063	0.38052		0.04412	0.685
Vegetation continuous field	8	-3.07373	0.08104	**0.00026**	***	0.00010	0.779
Land use (WMA)	8	-9.15874	0.24299	**5.14E-05**	***	2.52E-05	0.796









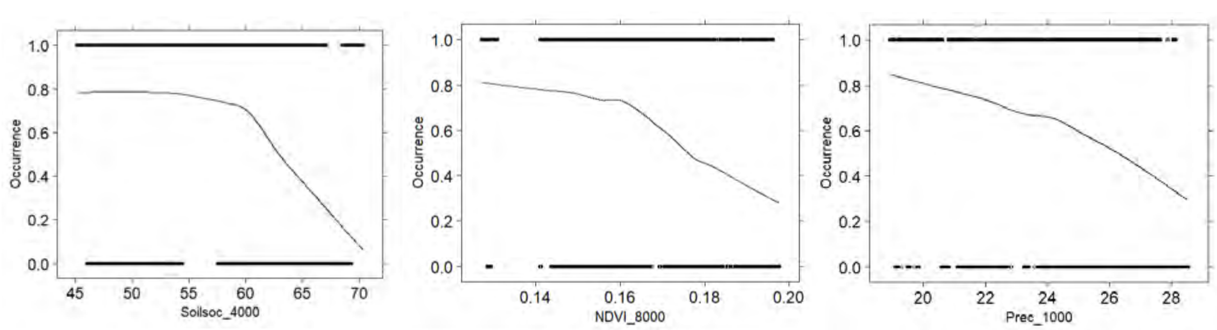


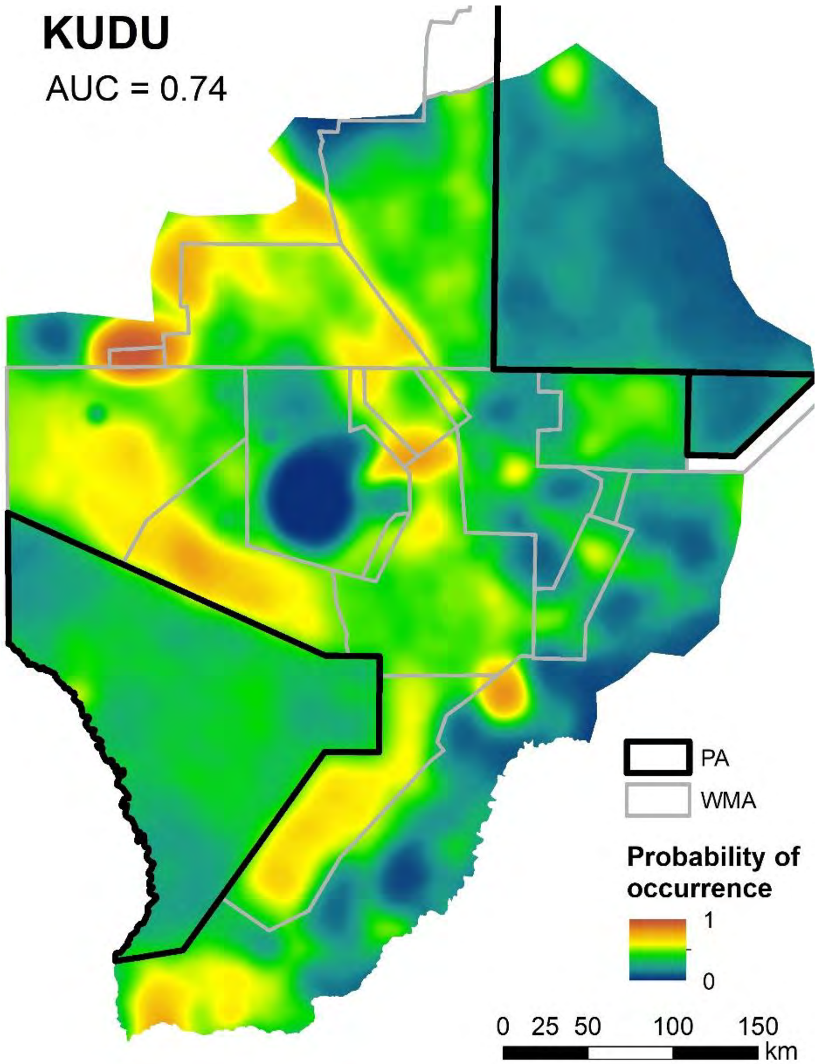
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.71	0.01
Kappa statistic	0.36	0.02
Sensitivity	0.75	0.07
Specificity	0.61	0.07
True skill statistic (TSS)	0.36	0.02
Percent correctly classified (PCC)	0.68	0.01
Deviance explained	0.91	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		8.85730	0.08027	7.02E-27	***	6.65E-27	
Vegetation greeness (NDVI)	32	-19.16991	0.43543	2.73E-05	***	178E-05	0.879
Soil (clay content)	8	-0.00304	0.00026	0.35222		0.03647	0.692
Soil (Nitrogen)	8	0.00009	0.00010	0.59546		0.03460	0.689
Soil (organic carbon)	8	-0.09111	0.00101	3.98E-15	***	3.23E-15	1000



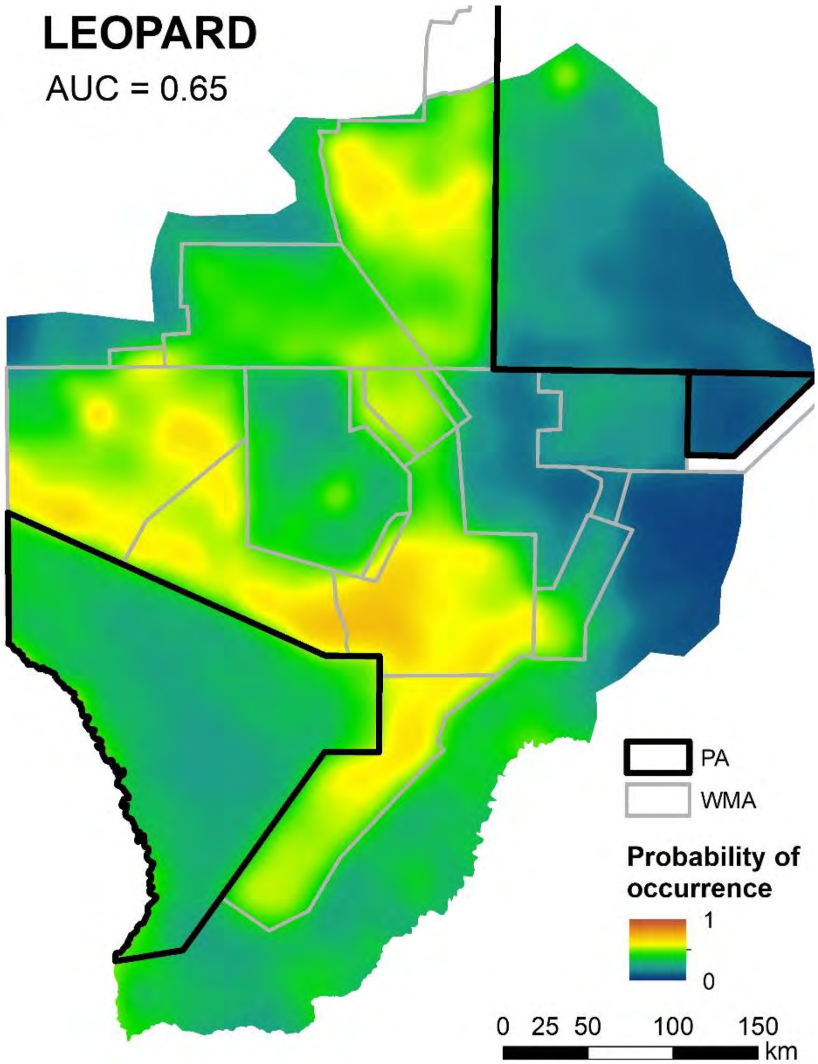


Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.74	0.01
Kappa statistic	0.37	0.01
Sensitivity	0.65	0.06
Specificity	0.72	0.06
True skill statistic (TSS)	0.37	0.01
Percent correctly classified (PCC)	0.68	0.01
Deviance explained	0.86	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-24.61441	0.60759	**3.57E-05**	***	196E-05	
Kraals (inside fenced farms)	32	30.82329	0.41329	**1.26E-15**	***	1.21E-15	0.238
Kraals	32	-3.31274	0.04525	**6.81E-17**	***	4.57E-17	0.614
Vegetation greeness (NDVI)	8	-13.48968	0.20623	**1.33E-11**	***	9.07E-12	0.687
Land use (PA)	8	140076	0.10121	0.24796		0.03491	0.323
Pans	8	35.29281	106237	**0.00600**	**	0.00185	0.559
Land use (CGA)	8	2.78232	0.10591	**0.02019**	*	0.00600	0.300
Precipitation	8	-0.05119	0.00046	**1.48E-36**	***	126E-36	1000
Roads (calcrete)	8	-28.27570	195834	0.15869		0.03014	0.276
Roads (paved)	8	135.49750	2.25287	**8.50E-12**	***	8.13E-12	0.226
Roads (sand)	8	-18.25081	136386	0.18915		0.02995	0.464
Soil (bulk density)	4	0.09968	0.00343	**0.00627**	**	0.00311	0.465
Soil (clay content)	8	0.01878	0.00040	**3.15E-07**	***	2.16E-07	0.539
Soil (sand)	8	0.01126	0.00024	**2.77E-06**	***	1.17E-06	0.499
Soil (organic carbon)	8	-0.00537	0.00115	0.47259		0.03939	0.494
Land use (WMA)	8	3.07828	0.09965	**0.00759**	**	0.00192	0.720

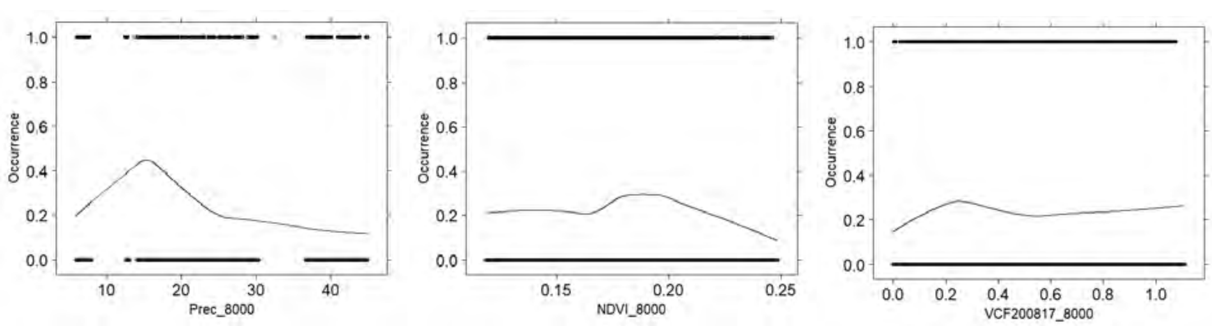


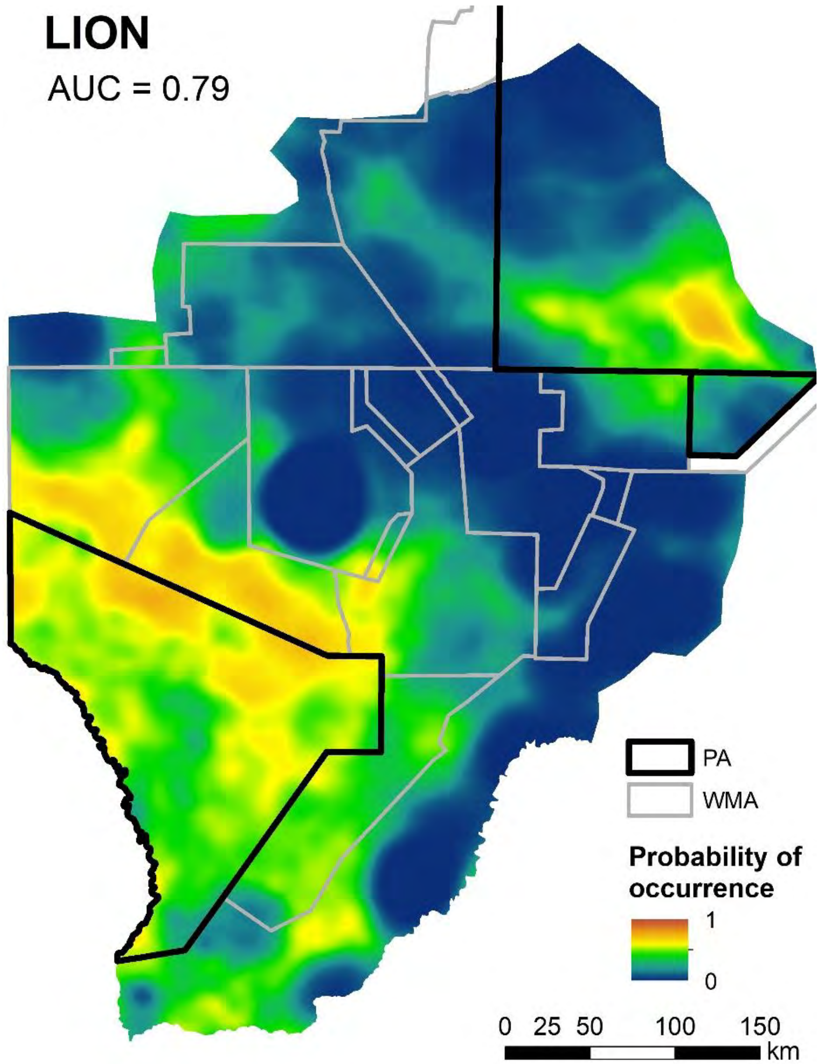
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.65	0.01
Kappa statistic	0.25	0.01
Sensitivity	0.68	0.05
Specificity	0.56	0.05
True skill statistic (TSS)	0.25	0.01
Percent correctly classified (PCC)	0.62	0.01
Deviance explained	0.95	0.00

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		7.84500	0.11311	**1.06E-10**	***	1.03E-10	
Vegetation greeness (NDVI)	8	6.82366	0.16808	**5.58E-05**	***	4.74E-05	0.711
Pans	8	17.18037	0.78001	**0.02814**	*	0.01120	0.450
Precipitation	8	-0.03135	0.00037	**5.02E-27**	***	2.85E-27	1000
Soil (clay content)	8	-0.00266	0.00024	0.22560		0.03974	0.509
Soil (Nitrogen)	8	-0.00377	6.10E-05	**3.10E-11**	***	2.86E-11	0.534
Soil (sand content)	8	-0.00969	0.00016	**1.44E-07**	***	1.40E-07	0.514
Soil (organic carbon)	8	0.01433	0.00102	0.15921		0.03224	0.507
Land use (WMA)	8	1.35881	0.00964	**1.16E-57**	***	8.55E-58	0.552



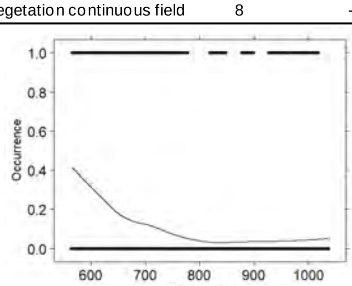


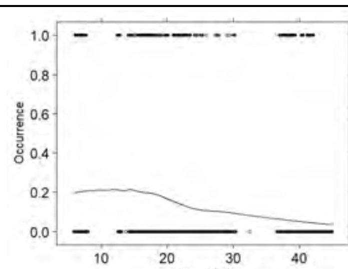
Model Prediction Performance

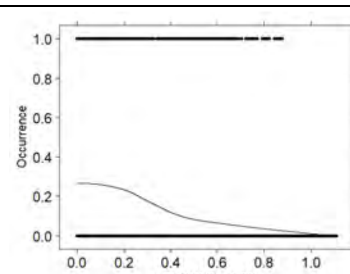
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.79	0.01
Kappa statistic	0.48	0.02
Sensitivity	0.87	0.04
Specificity	0.61	0.04
True skill statistic (TSS)	0.48	0.02
Percent correctly classified (PCC)	0.74	0.01
Deviance explained	0.78	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-8.12523	0.60966	0.22002		0.04175	
Kraals	32	-7.33859	0.05860	**1.11E-28**	***	6.76E-29	0.782
Pans	8	19.80785	0.82718	**0.02929**	*	0.00867	0.533
Precipitation	8	-0.03924	0.00047	**2.12E-19**	***	1.53E-19	0.944
Roads (sand)	8	-8.29998	1.54755	0.44813		0.04079	0.545
Soil (bulk density)	4	0.08497	0.00438	0.06811		0.01754	0.605
Soil (clay content)	8	-0.02082	0.00038	**6.36E-07**	***	6.22E-07	0.760
Soil (Nitrogen)	4	-0.00688	0.00010	**4.87E-13**	***	4.77E-13	1.000
Soil (sand content)	4	0.00037	0.00026	0.53056		0.04165	0.826
Soil (organic carbon)	8	0.06950	0.00148	**9.26E-05**	***	8.84E-05	0.735
Vegetation continuous field	8	-2.22358	0.03321	**3.91E-13**	***	3.83E-13	0.996

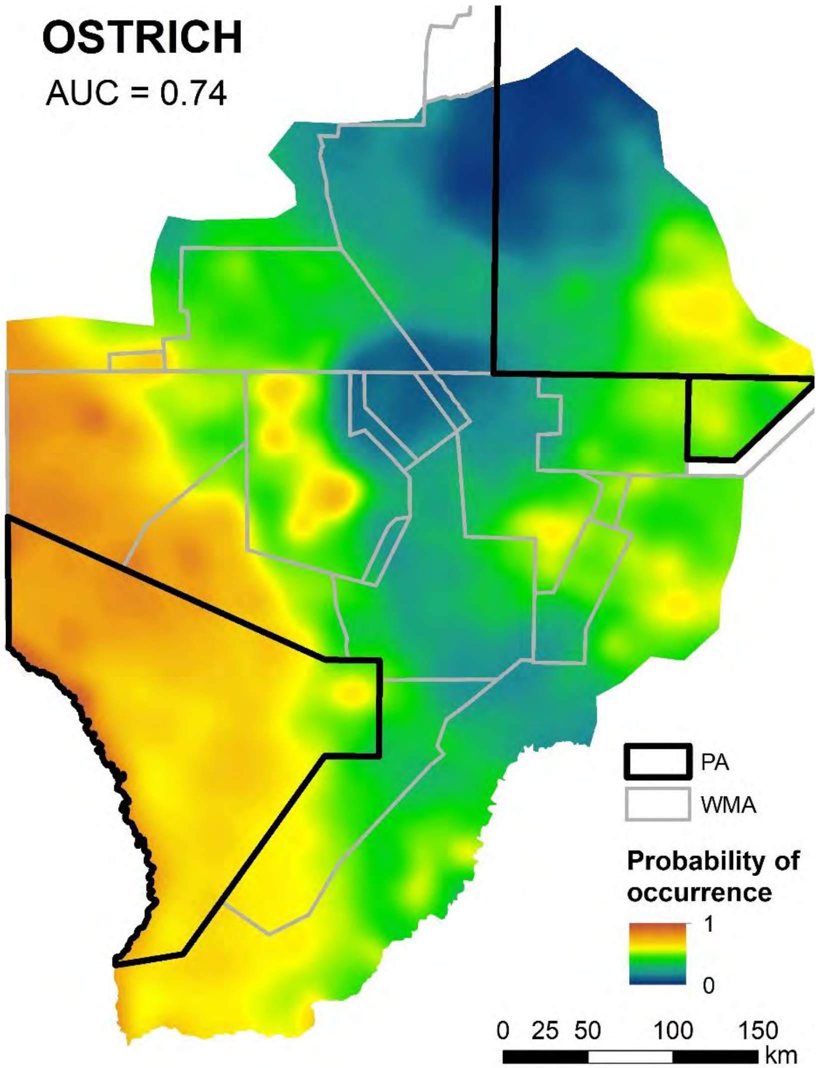






OSTRICH

AUC = 0.74

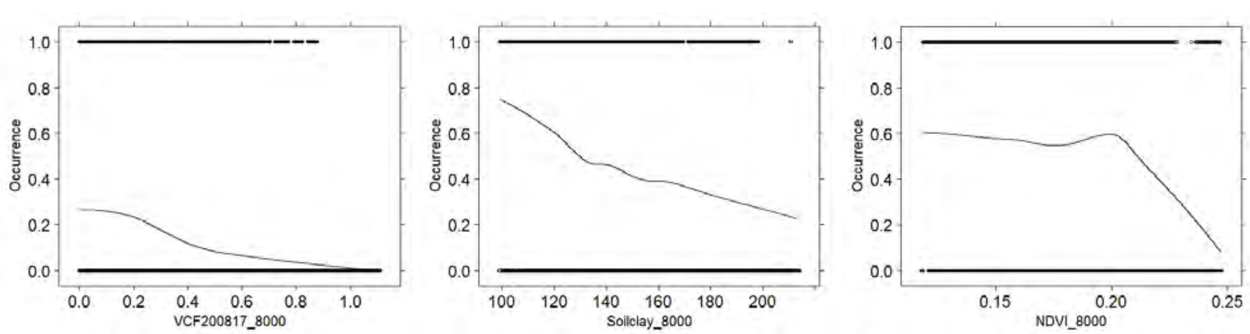


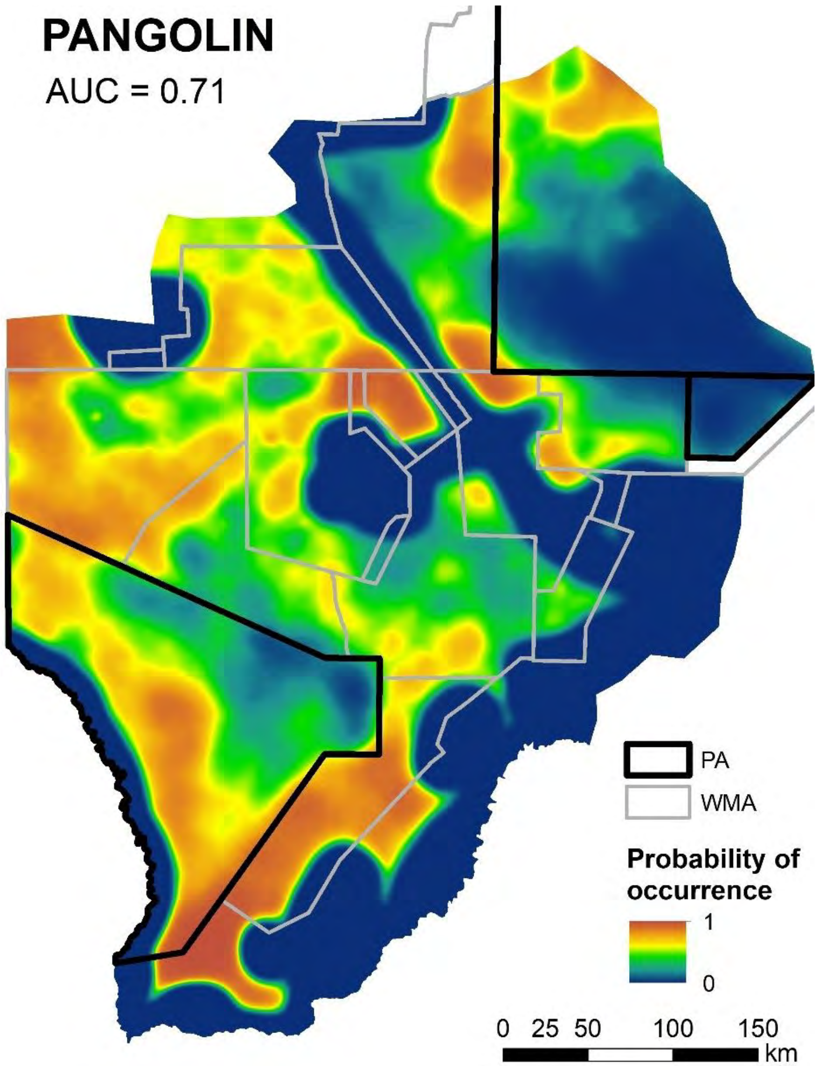
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.74	0.01
Kappa statistic	0.39	0.01
Sensitivity	0.68	0.04
Specificity	0.71	0.04
True skill statistic (TSS)	0.39	0.01
Percent correctly classified (PCC)	0.69	0.01
Deviance explained	0.87	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		7.10192	0.04736	**3.65E-61**	***	3.57E-61	
Vegetation greeness (NDVI)	8	-13.27717	0.12781	**6.02E-29**	***	5.33E-29	0.642
Pans	8	44.13687	0.88254	**7.87E-07**	***	5.89E-07	0.566
Precipitation	2	-0.07289	0.00034	**1.25E-132**	***	7.46E-133	1000
Soil (clay content)	8	-0.02242	0.00017	**8.91E-51**	***	8.61E-51	0.704
Soil (organic carbon)	8	0.00044	0.00075	0.55295		0.03818	0.465



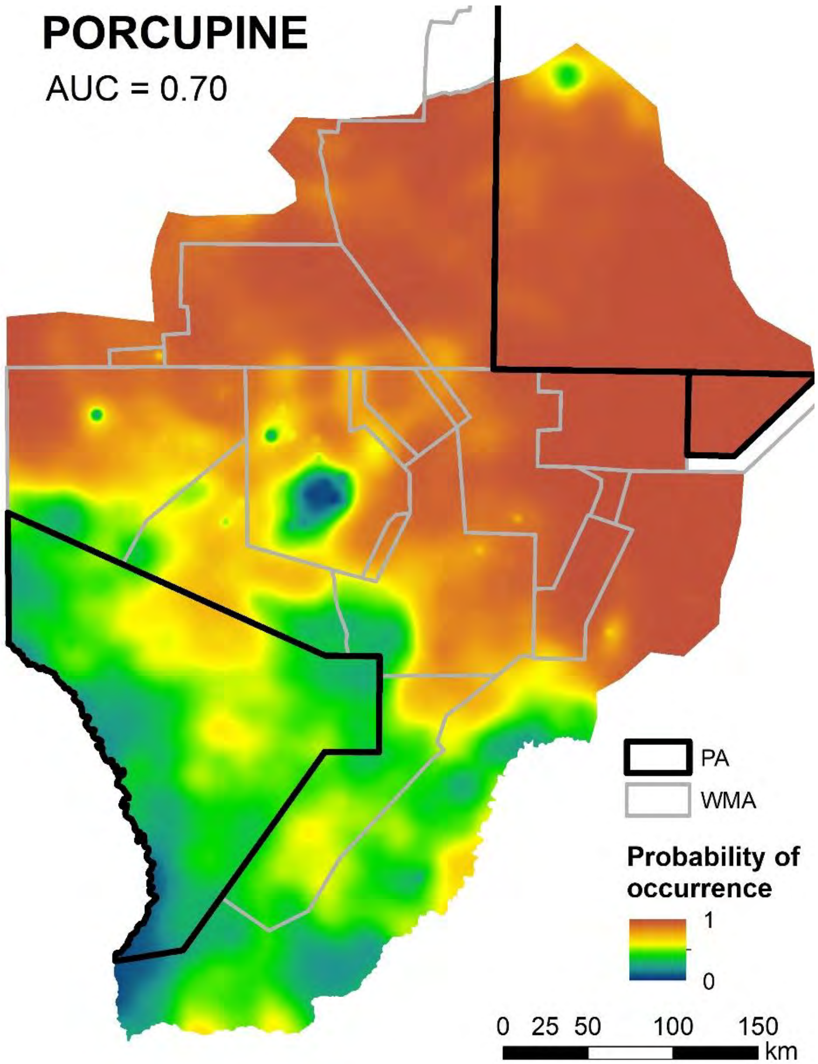


Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.71	0.01
Kappa statistic	0.44	0.01
Sensitivity	0.92	0.04
Specificity	0.51	0.04
True skill statistic (TSS)	0.44	0.01
Percent correctly classified (PCC)	0.72	0.01
Deviance explained	0.83	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-16.47896	2.02708	0.42358		0.04573	
Kraals (inside fenced farms)	32	0.66285	136673	0.67058		0.02792	0.106
Kraals	32	0.55866	0.06670	0.34584		0.04297	0.500
Vegetation greeness (NDVI)	2	-9.32024	103789	0.29875		0.03912	0.710
Land use (PA)	8	57.88922	128423	**0.00045**	***	0.00027	0.229
Pans	8	-142.23571	3.87522	**0.00048**	***	0.00032	0.518
Land use (CGA)	8	60.87967	132050	**0.00029**	***	0.00019	0.223
Roads (calcrete)	8	42.79255	5.75889	0.35892		0.04441	1000
Roads (paved)	8	-1933.17288	58.31364	**0.00816**	**	0.00144	0.027
Roads (sand)	8	87.11592	4.03334	**0.03588**	*	0.01071	0.657
Soil (bulk density)	4	-0.12290	0.01086	0.25098		0.03916	0.483
Soil (clay content)	8	0.00199	0.00158	0.44391		0.04148	0.645
Soil (Nitrogen)	8	0.01402	0.00058	**0.00686**	**	0.00273	0.603
Soil (sand content)	8	-0.03512	0.00078	**2.35E-05**	***	186E-05	0.600
Soil (organic carbon)	8	-0.10871	0.00795	0.14015		0.02582	0.583
Vegetation continuous field	4	-0.08778	0.10270	0.50365		0.03969	0.678
Land use (WMA)	8	60.18960	130134	**0.00030**	***	0.00019	0.737

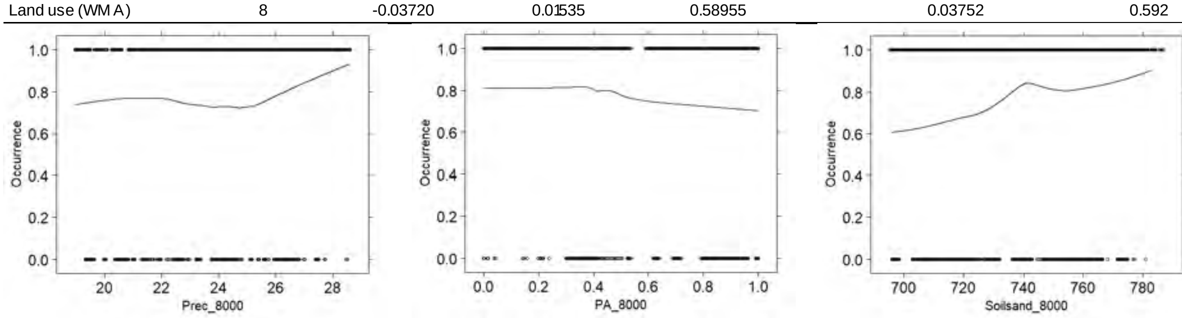


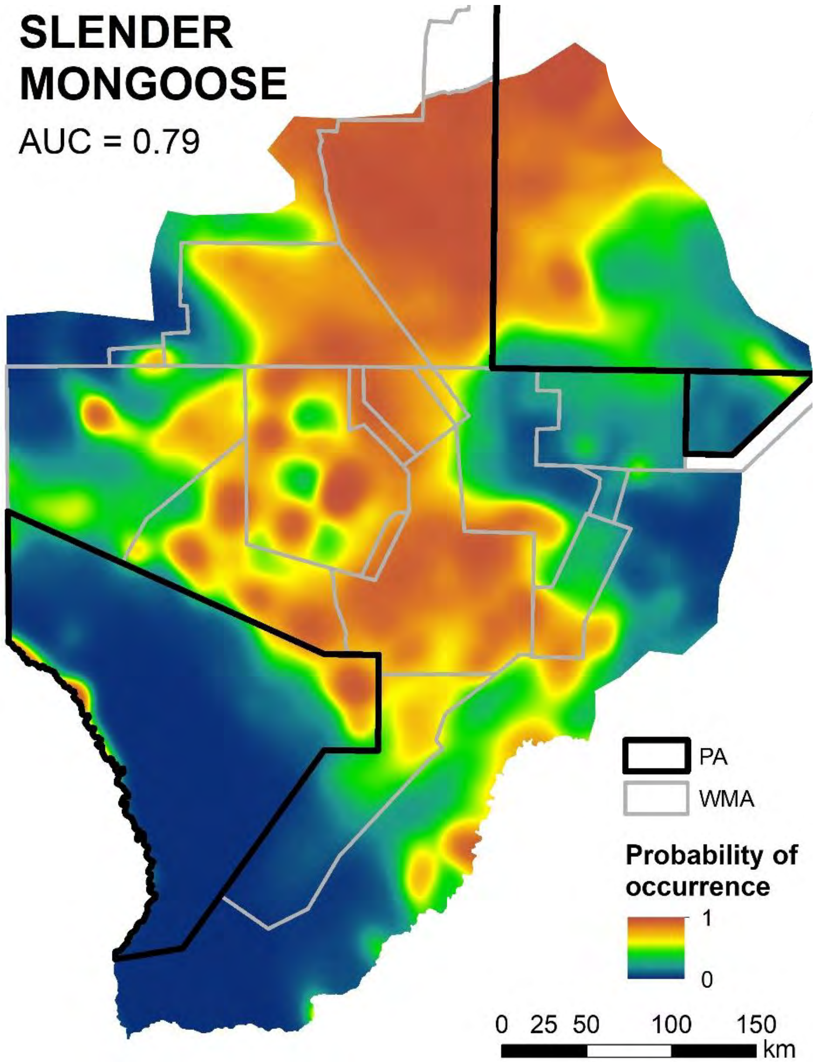
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.70	0.01
Kappa statistic	0.32	0.02
Sensitivity	0.76	0.11
Specificity	0.56	0.11
True skill statistic (TSS)	0.32	0.02
Percent correctly classified (PCC)	0.66	0.01
Deviance explained	0.90	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-32.44789	0.65680	**8.76E-09**	***	5.07E-09	
Pans	8	-13.95720	1.30575	0.34654		0.03798	0.461
Roads (sand)	8	-54.20551	1.17690	**0.00013**	***	9.19E-05	0.454
Soil (bulk density)	2	0.04317	0.00359	0.20234		0.02882	0.510
Soil (clay content)	8	0.01287	0.00061	**0.02644**	*	0.00757	0.630
Soil (Nitrogen)	8	0.00556	0.00018	**0.00280**	**	0.00088	0.558
Soil (sand content)	8	0.03534	0.00039	**5.63E-20**	***	2.46E-20	0.844
Soil (organic carbon)	8	-0.09444	0.00191	**3.33E-07**	***	1.98E-07	0.724
Vegetation continuous field	8	2.01154	0.04887	**0.00019**	***	8.28E-05	0.575
Land use (WMA)	8	-0.03720	0.01535	0.58955		0.03752	0.592



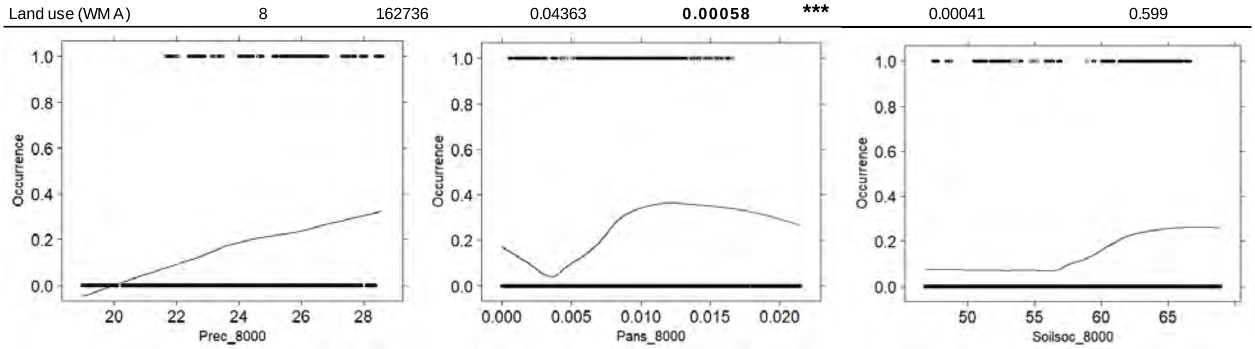


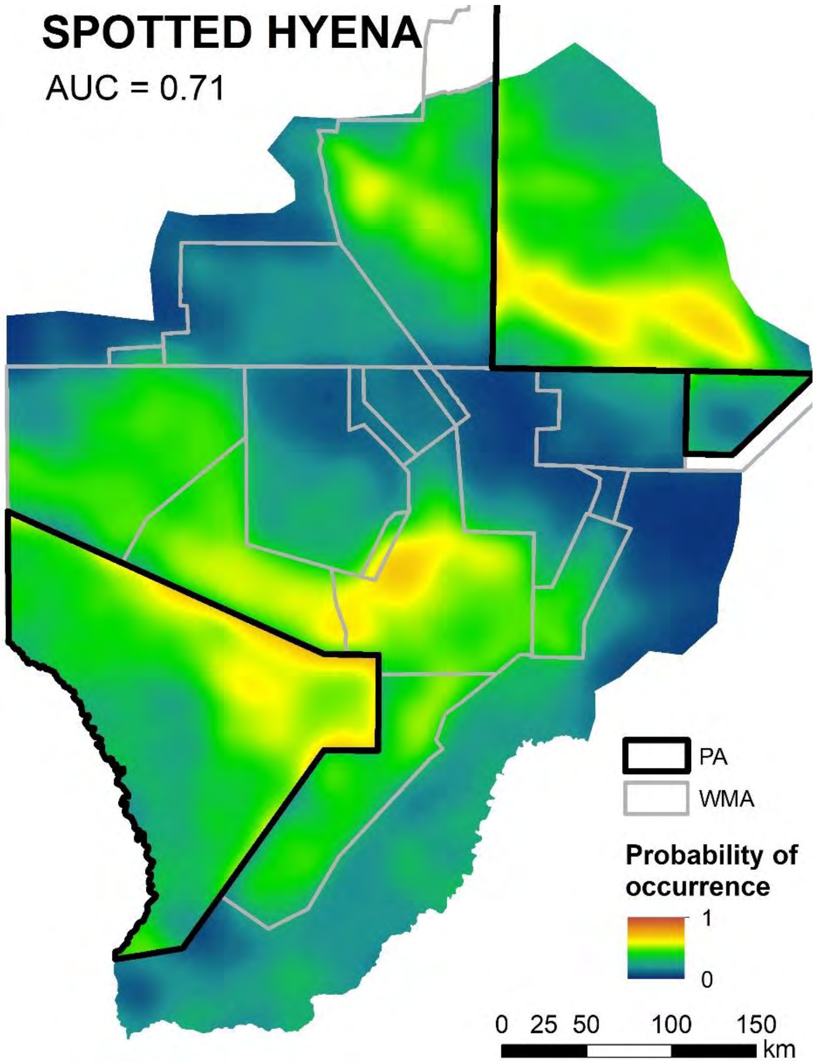
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.79	0.02
Kappa statistic	0.50	0.04
Sensitivity	0.94	0.06
Specificity	0.56	0.07
True skill statistic (TSS)	0.50	0.04
Percent correctly classified (PCC)	0.75	0.02
Deviance explained	0.74	0.03

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-1.18437	0.94915	0.47802		0.03888	
Kraals (inside fenced farms)	32	-12.65445	106946	0.21724		0.03708	0.440
Pans	8	206.53042	4.99676	**1.07E-07**	***	4.78E-08	0.929
Precipitation	8	0.79794	0.01320	**1.39E-08**	***	123E-08	1000
Roads (sand)	8	-58.87364	3.62582	0.11590		0.02732	0.414
Soil (clay content)	8	0.00050	0.00130	0.48756		0.03707	0.490
Soil (Nitrogen)	8	-0.01378	0.00039	**0.00035**	***	0.00023	0.666
Soil (sand content)	8	-0.01470	0.00131	0.24057		0.03609	0.543
Land use (WMA)	8	162736	0.04363	**0.00058**	***	0.00041	0.599



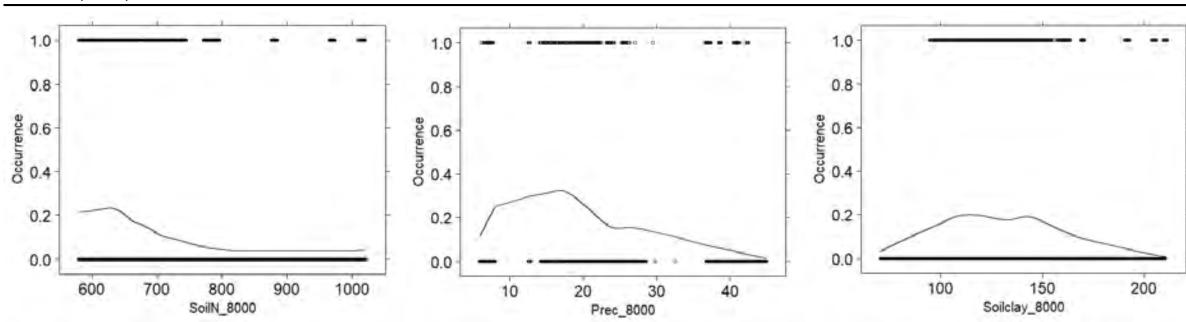


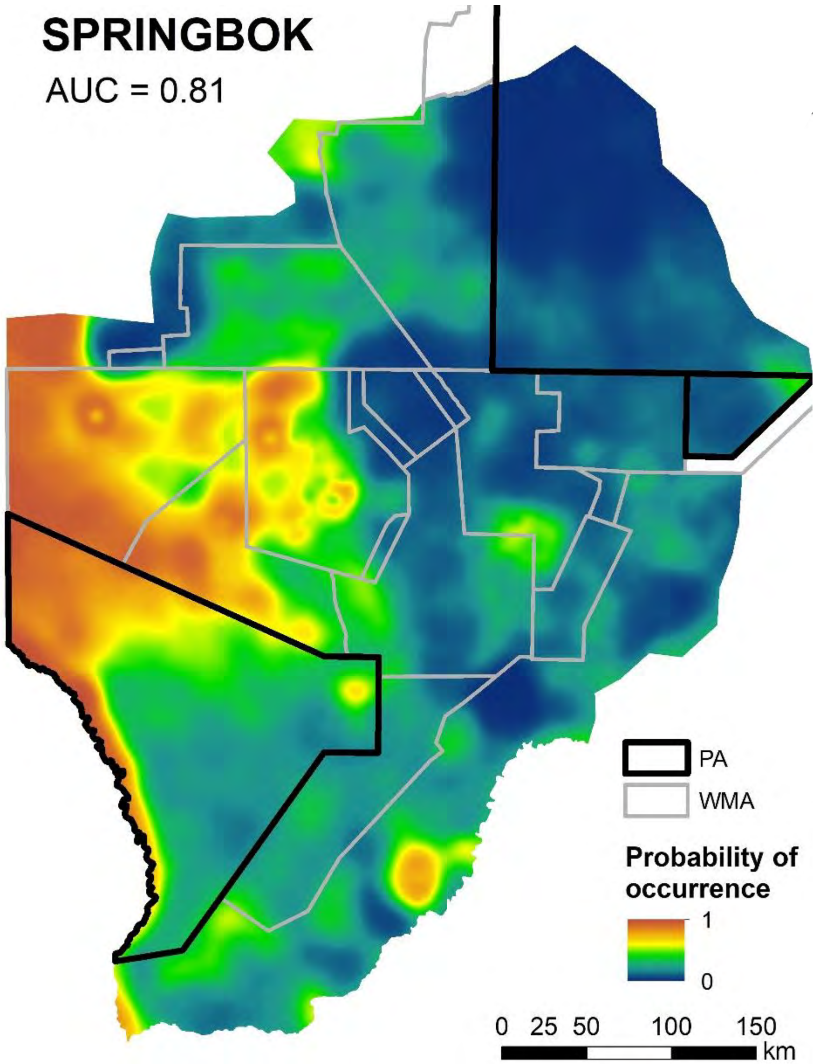
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.71	0.01
Kappa statistic	0.37	0.02
Sensitivity	0.78	0.02
Specificity	0.60	0.02
True skill statistic (TSS)	0.37	0.02
Percent correctly classified (PCC)	0.69	0.01
Deviance explained	0.89	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-3.87055	0.19764	0.07321		0.01945	
Vegetation greeness (NDVI)	8	8.10034	0.31521	**0.02451**	*	0.01591	0.712
Land use (PA)	0.25	152260	0.01863	**1.82E-19**	***	165E-19	0.594
Precipitation	8	-0.03449	0.00060	**6.61E-12**	***	4.43E-12	0.987
Soil (clay content)	8	0.00515	0.00032	0.15084		0.03097	0.781
Soil (Nitrogen)	8	-0.00833	0.00011	**3.34E-21**	***	2.28E-21	1000
Soil (sand content)	8	0.00901	0.00026	**0.00096**	***	0.00074	0.707
Soil (organic carbon)	8	-0.00175	0.00141	0.54519		0.03654	0.688
Vegetation continuous field	8	-0.64233	0.03434	0.06307		0.01458	0.744
Land use (WMA)	8	149819	0.01798	**1.68E-21**	***	138E-21	0.714



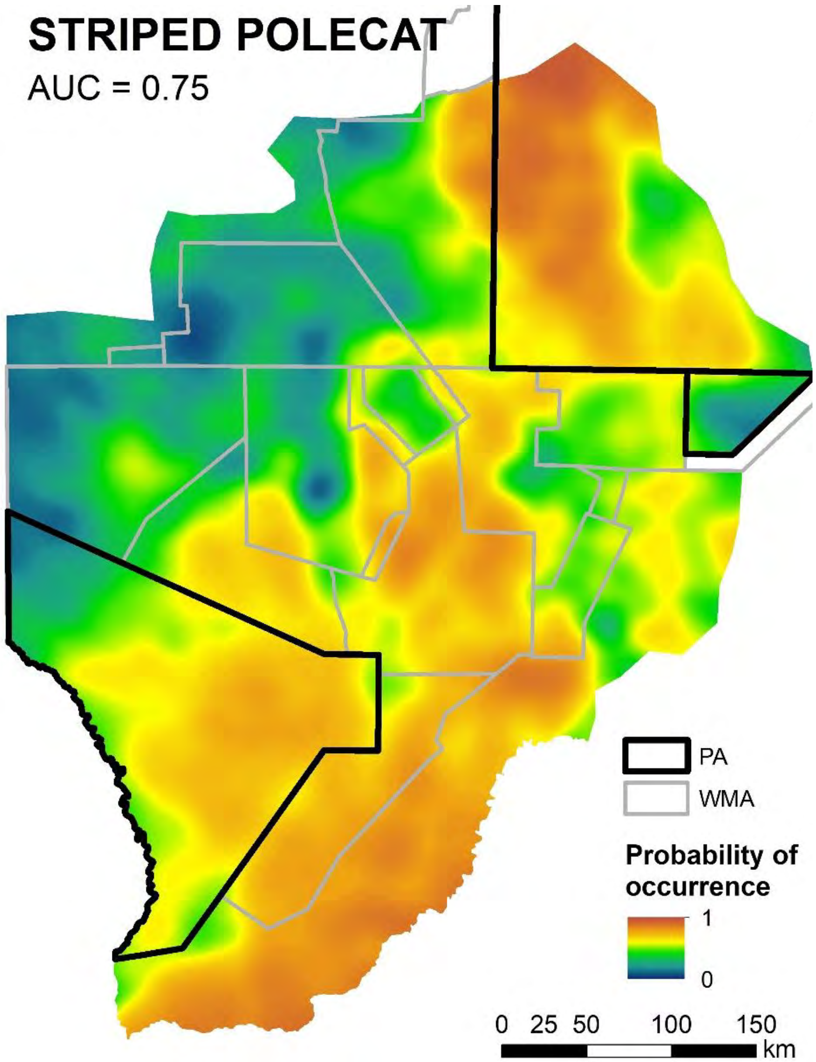


Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.81	0.01
Kappa statistic	0.58	0.03
Sensitivity	0.88	0.03
Specificity	0.70	0.04
True skill statistic (TSS)	0.58	0.03
Percent correctly classified (PCC)	0.79	0.01
Deviance explained	0.77	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		7.76692	136037	0.42230		0.04313	
Human population	32	0.07623	0.00398	**0.03875**	*	0.01114	0.335
Kraals (inside fenced farms)	32	-47.41727	130103	**8.81E-06**	***	7.59E-06	0.262
Land use (PA)	8	-1.16310	0.60200	0.29831		0.03736	0.252
Pans	8	125.06296	2.26222	**3.97E-07**	***	2.76E-07	0.652
Land use (CGA)	8	0.28542	0.61485	0.31738		0.03590	0.263
Precipitation	2	-0.05833	0.00098	**7.09E-11**	***	6.09E-11	1000
Roads (calcrete)	8	79.40684	3.54296	**0.03050**	*	0.00906	0.340
Roads (paved)	8	-6104229	4.90473	0.23181		0.03341	0.087
Roads (sand)	8	-25.41130	193155	0.25392		0.03481	0.460
Soil (bulk density)	4	0.17114	0.00813	**0.03479**	*	0.01221	0.350
Soil (clay content)	8	-0.05080	0.00073	**9.20E-14**	***	5.90E-14	0.721
Soil (sand content)	8	-0.02813	0.00052	**5.66E-09**	***	5.07E-09	0.567
Soil (organic carbon)	8	-0.06784	0.00244	**0.01238**	*	0.00451	0.756
Vegetation continuous field	8	-0.29547	0.05855	0.42758		0.04320	0.574
Land use (WMA)	8	-0.11385	0.60666	0.30783		0.03427	0.408

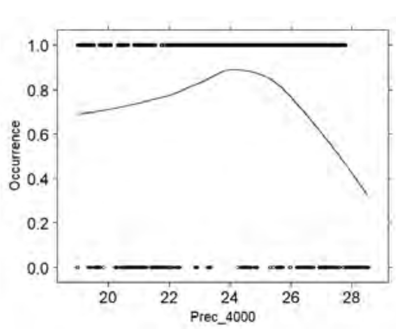


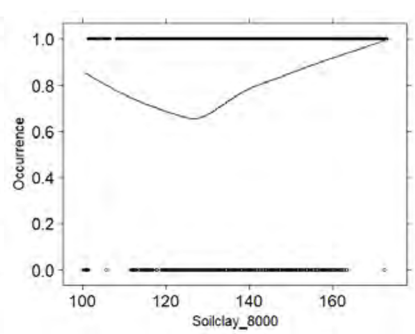
Model Prediction Performance

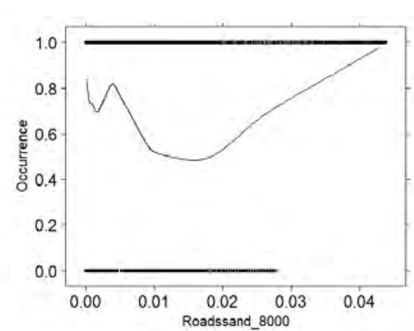
Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.75	0.01
Kappa statistic	0.44	0.02
Sensitivity	0.77	0.05
Specificity	0.67	0.05
True skill statistic (TSS)	0.44	0.02
Percent correctly classified (PCC)	0.72	0.01
Deviance explained	0.88	0.01

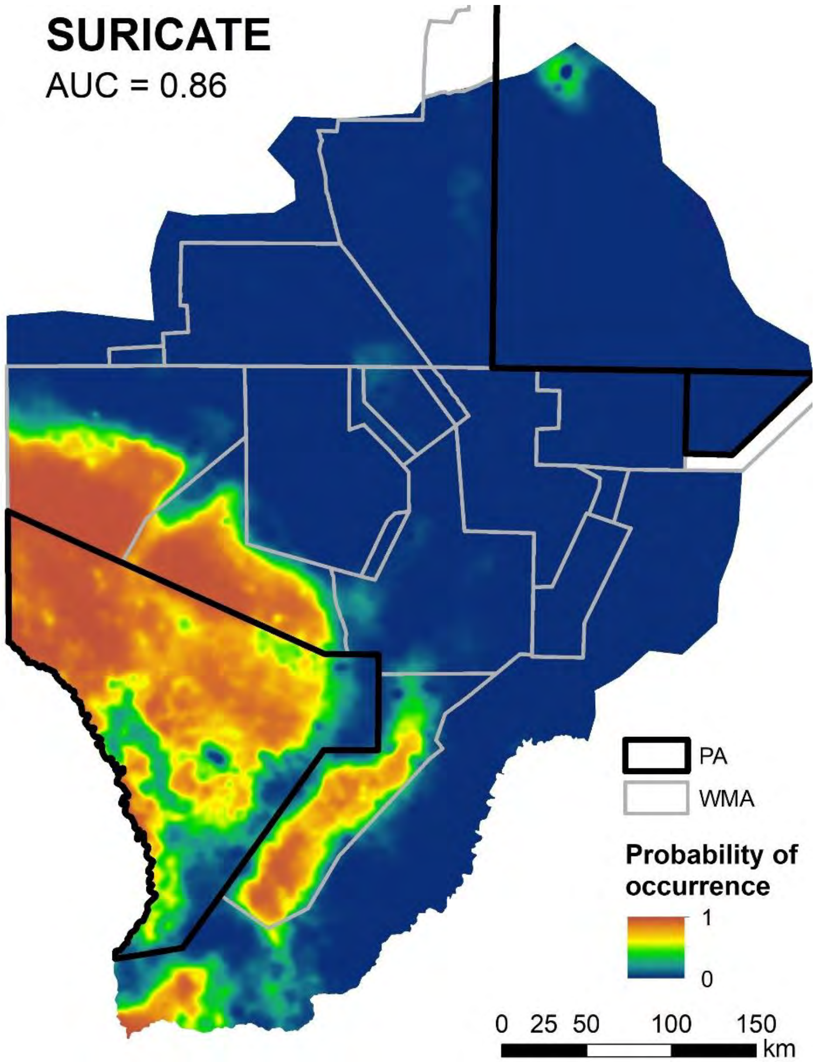
General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-22.55371	0.39030	**9.56E-10**	***	7.20E-10	
Roads (sand)	8	-56.20781	193010	**7.62E-05**	***	6.30E-05	0.929
Soil (clay content)	8	0.04369	0.00041	**3.37E-34**	***	2.56E-34	0.958
Soil (Nitrogen)	8	-0.00228	0.00020	0.26881		0.04223	0.845
Soil (sand content)	8	0.02565	0.00041	**4.02E-10**	***	3.75E-10	0.680
Vegetation continuous field	8	-2.61795	0.04709	**4.66E-07**	***	2.66E-07	0.852







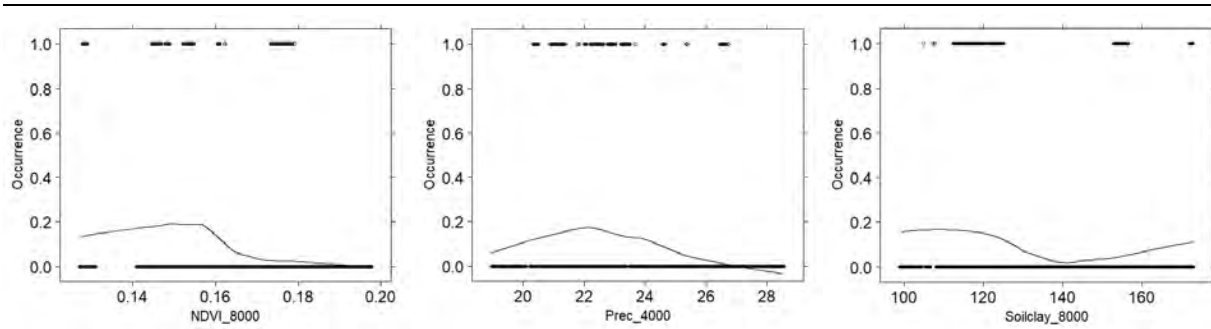


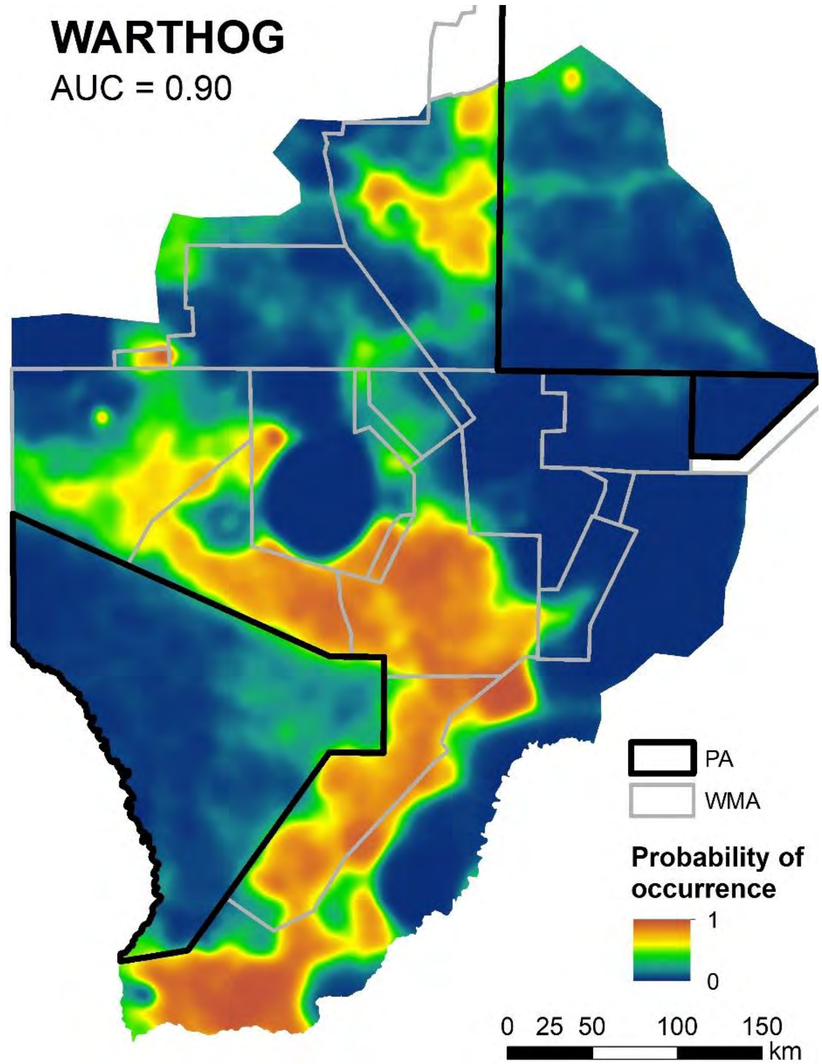
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.86	0.02
Kappa statistic	0.59	0.03
Sensitivity	0.82	0.11
Specificity	0.77	0.12
True skill statistic (TSS)	0.59	0.03
Percent correctly classified (PCC)	0.80	0.02
Deviance explained	0.63	0.04

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		60.80101	17.4145	**0.00057**	***	0.00020	
Vegetation greeness (NDVI)	8	-71.84951	2.11388	**0.00084**	***	0.00056	1.000
Roads (sand)	8	-536.68030	43.14487	**0.03914**	*	0.01954	0.549
Soil (clay content)	8	0.01865	0.00195	0.31573		0.04199	0.806
Soil (Nitrogen)	8	-0.01716	0.00079	0.05824		0.01021	0.547
Soil (sand content)	8	-0.04399	0.00199	**0.02679**	*	0.00909	0.730
Soil (organic carbon)	2	-0.18322	0.00638	**0.00219**	**	0.00058	0.757
Land use (WMA)	8	4.52135	0.08721	**1.98E-07**	***	1.34E-07	0.576



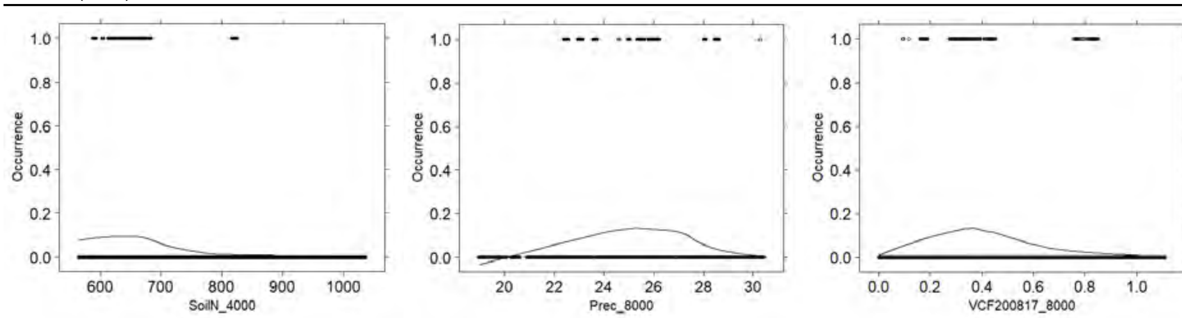


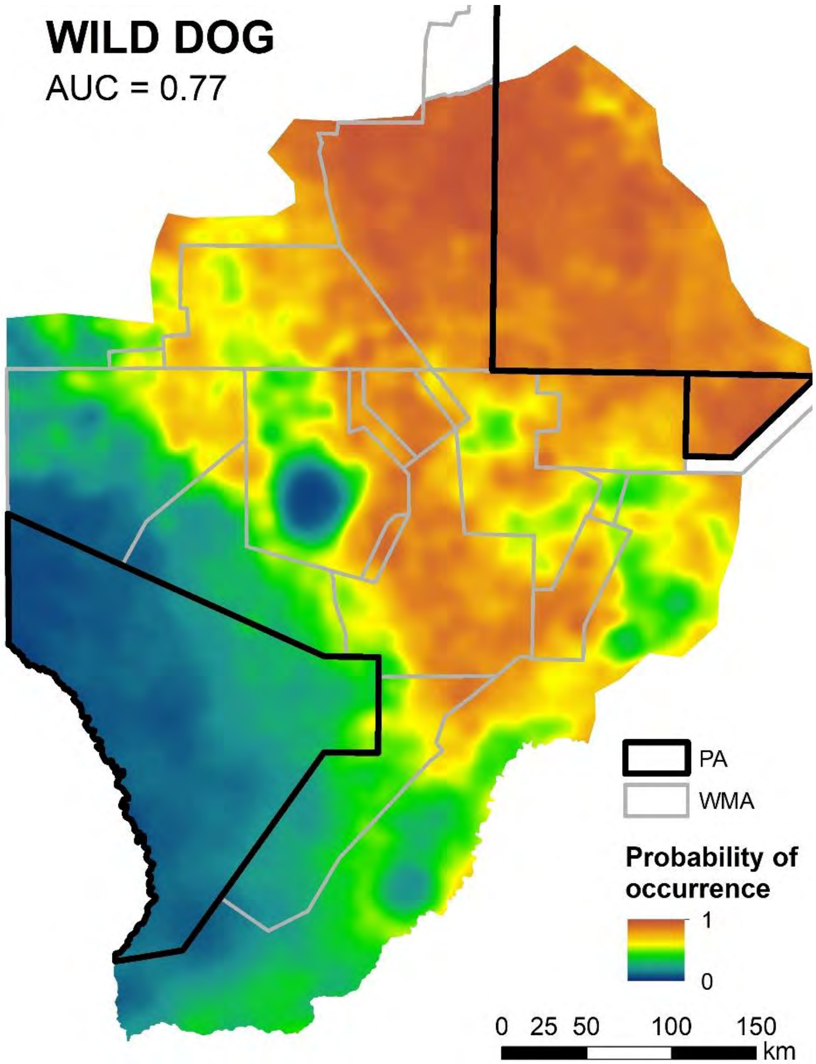
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.90	0.02
Kappa statistic	0.72	0.05
Sensitivity	0.93	0.04
Specificity	0.79	0.05
True skill statistic (TSS)	0.72	0.05
Percent correctly classified (PCC)	0.86	0.02
Deviance explained	0.53	0.05

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		35.33664	2.30080	0.11212		0.02491	
Kraals (inside fenced farms)	32	32.96182	172122	0.07646		0.01689	0.445
Kraals	32	-14.68426	0.69617	**0.01422**	*	0.00501	0.430
Land use (CGA)	8	4.58459	0.21958	**0.02244**	*	0.00519	0.497
Soil (bulk density)	4	-0.32835	0.01656	**0.04492**	*	0.01299	0.654
Soil (clay content)	4	0.03178	0.00145	**0.04604**	*	0.00966	0.739
Soil (Nitrogen)	4	-0.01716	0.00058	**9.35E-06**	***	3.26E-06	1000
Soil (sand content)	4	0.02108	0.00169	0.21849		0.03480	0.855
Soil (organic carbon)	4	0.02976	0.00522	0.46226		0.03729	0.487
Land use (WMA)	8	3.67180	0.09557	**0.00059**	***	0.00041	0.660



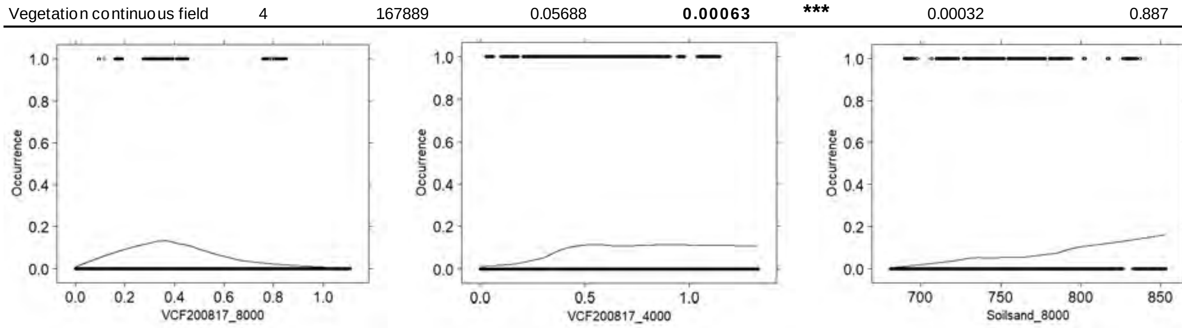


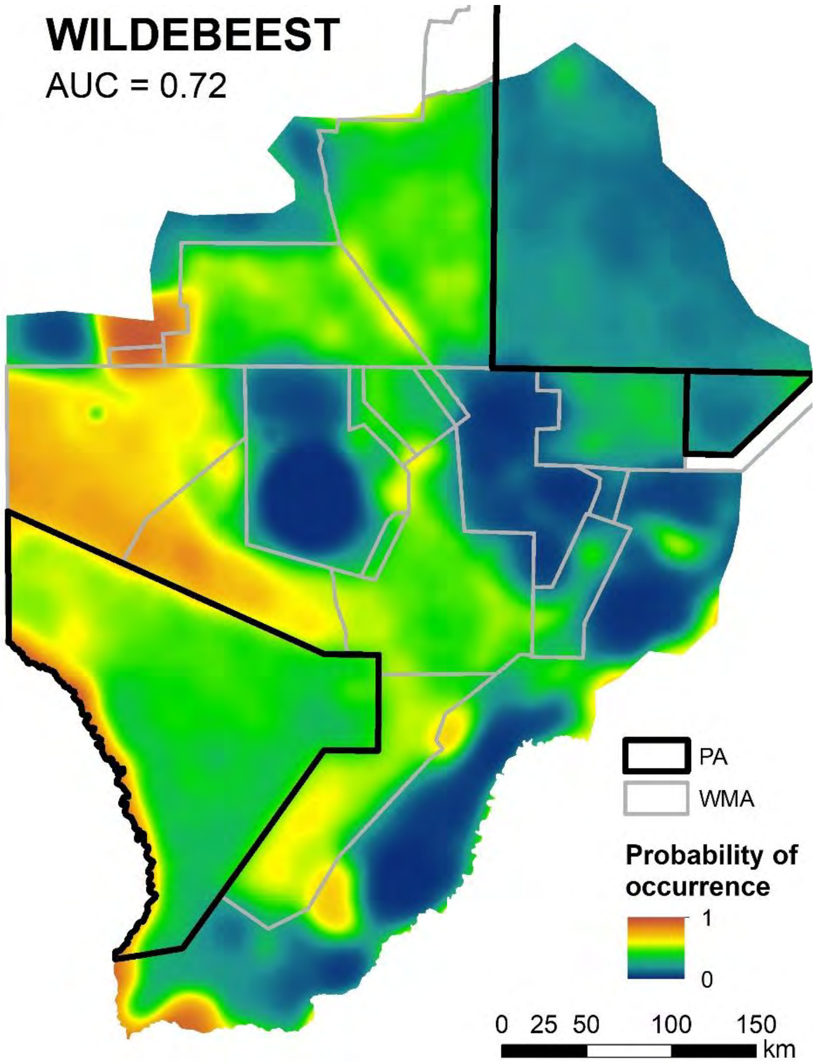
Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.77	0.01
Kappa statistic	0.43	0.03
Sensitivity	0.82	0.09
Specificity	0.61	0.09
True skill statistic (TSS)	0.43	0.03
Percent correctly classified (PCC)	0.72	0.01
Deviance explained	0.82	0.02

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-2101306	0.40336	**3.33E-09**	***	135E-09	
Kraals	32	-100825	0.06340	0.09518		0.02294	0.503
Vegetation greeness (NDVI)	8	2167117	0.76721	**0.00214**	**	0.00103	0.569
Pans	8	-9.60280	2.05900	0.44045		0.03749	0.464
Precipitation	4	-0.03521	0.00090	**0.00030**	***	0.00017	0.527
Soil (clay content)	4	0.02538	0.00060	**3.78E-05**	***	2.33E-05	0.678
Soil (Nitrogen)	8	-0.00430	0.00015	**0.01121**	*	0.00570	0.586
Soil (sand content)	8	0.02203	0.00046	**5.82E-06**	***	4.95E-06	0.718
Soil (organic carbon)	8	0.00609	0.00345	0.47206		0.04324	1000
Vegetation continuous field	4	167889	0.05688	**0.00063**	***	0.00032	0.887



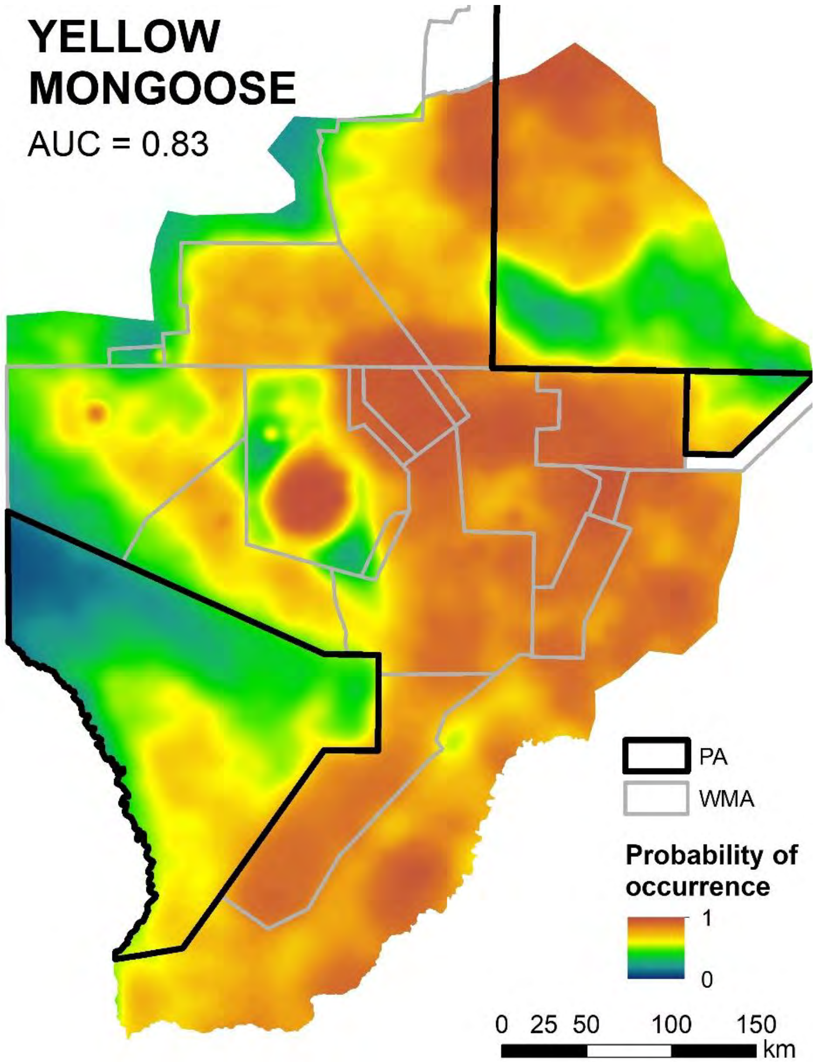


Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.72	0.01
Kappa statistic	0.33	0.02
Sensitivity	0.72	0.06
Specificity	0.61	0.06
True skill statistic (TSS)	0.33	0.02
Percent correctly classified (PCC)	0.66	0.01
Deviance explained	0.87	0.01

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-2.71119	0.65228	0.47414		0.03952	
Kraals (inside fenced farms)	32	1127611	0.44355	**0.02521**	*	0.00909	0.368
Kraals	32	-3.46168	0.05471	**6.56E-06**	***	4.42E-06	0.563
Vegetation greeness (NDVI)	8	-5.46623	0.23681	**0.02468**	*	0.00577	0.566
Land use (PA)	8	-2.69392	0.16815	0.08089		0.02112	0.257
Pans	8	42.08562	162633	**0.01371**	*	0.00496	0.403
Land use (CGA)	8	-3.31195	0.17382	**0.04833**	*	0.01970	0.389
Precipitation	4	-0.04118	0.00052	**2.30E-19**	***	142E-19	1000
Roads (calcrete)	8	-6.58551	2.63534	0.50847		0.04226	0.302
Roads (paved)	8	123.41092	190212	**1.99E-07**	***	1.19E-07	0.152
Roads (sand)	8	6.29185	126051	0.56495		0.04072	0.287
Soil (bulk density)	4	0.07997	0.00412	0.08993		0.02992	0.433
Soil (clay content)	8	-0.00315	0.00043	0.31371		0.03759	0.473
Soil (sand content)	8	-0.00336	0.00033	0.27925		0.03781	0.550
Soil (organic carbon)	8	-0.04358	0.00136	**0.00152**	**	0.00059	0.517
Land use (WMA)	8	-131347	0.16245	0.32375		0.04018	0.475

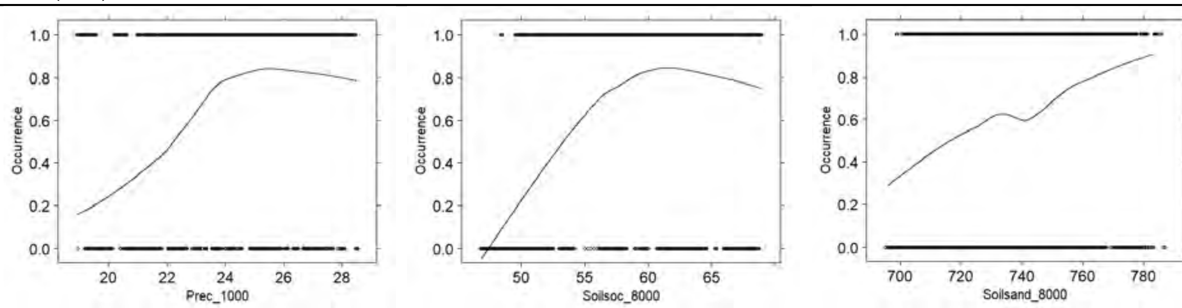


Model Prediction Performance

Performance Statistic [50 bootstrap runs]	Mean Value	Standard Deviation
Area under curve (AUC)	0.83	0.01
Kappa statistic	0.56	0.02
Sensitivity	0.87	0.05
Specificity	0.69	0.05
True skill statistic (TSS)	0.56	0.02
Percent correctly classified (PCC)	0.78	0.01
Deviance explained	0.73	0.02

General Linear Model from 50 bootstrap runs

Variable	Scale (km)	Mean Coefficient	Standard Error (Coeff)	p-value	Significance	Standard Error (p-value)	Model Improvement Ratio
(Intercept)		-0.54620	0.85592	0.54625		0.03898	
Kraals	32	2.02595	0.02992	**2.11E-14**	***	1.31E-14	0.406
Vegetation greeness (NDVI)	8	-9.76281	0.52725	0.09281		0.02031	0.599
Pans	8	-22.41091	1.98387	0.22051		0.03436	0.392
Soil (bulk density)	4	-0.15374	0.00523	**0.00513**	**	0.00225	0.415
Soil (clay content)	8	0.04018	0.00063	**1.02E-10**	***	8.43E-11	0.655
Soil (Nitrogen)	8	0.00449	0.00020	**0.03229**	*	0.01665	0.439
Soil (sand content)	8	0.01996	0.00052	**9.27E-05**	***	7.57E-05	0.680
Soil (organic carbon)	8	0.02083	0.00301	0.30965		0.03832	0.834
Land use (WMA)	8	1.40329	0.02187	**2.66E-13**	***	2.34E-13	0.400



3.1.1 Broad patterns of species occurrence in the AOI

Most smaller wildlife species' habitat suitability is not strongly predicted by variables of human influence (i.e. kraals, human population, land use type, roads), and some models contained no anthropogenic factors at all (bat-eared fox, cape fox, genet, kori bustard). Except for brown hyena and ostrich, large wildlife species are more negatively impacted by human-livestock disturbance than small wildlife species. This is among the strongest visible patterns in the large-bodied species maps; cool blue areas of low or zero occurrence probability mirroring kraals distribution. Among species, the impact of kraals occurred most strongly - almost consistently - at the largest variable scale measured (32 km kernel radius; see Figure 4). This shows that kraals have a large and, importantly, broad-scale effect on the occurrence patterns of the largest species. Of six large antelope species, KraalsNF (32 km scale) variable was selected and appeared as negative coefficient with high significance in the final models of five. Gemsbok differed only in 16 km being the strongest KraalsNF scale and again highly significant. Among large carnivores, cheetah and lion showed similar response to large antelope, while KraalsNF (32 km scale) also appeared with negative coefficient in the wild dog model, although not significant. Comparatively, among the 19 smaller species, KraalsNF appeared as a negative coefficient in the models of four, and significant in one of those (warthog), similar to large antelopes. Human-livestock activity at a broad spatial scale is the overwhelming determinant of large-bodied wildlife species distributions, especially herbivores, in the Kalahari landscape.

A second broad pattern is a disproportion of high habitat suitability in the southwest portion of the AOI relative to the northeast portion, among many species large and small. Notable exceptions include porcupine and wild dog, whose probability of occurrence distributions appear to mirror the declining rainfall gradient from northeast to southwest.

Note the HS models were developed without the inclusion of fences as a variable. Thus, large antelopes may have predicted occurrence in areas that are fenced ranching blocks where free-ranging populations are in fact excluded by the filter/barrier effect of fences. Fences with varying resistances to wildlife species movement will be introduced into Phase 2 models.

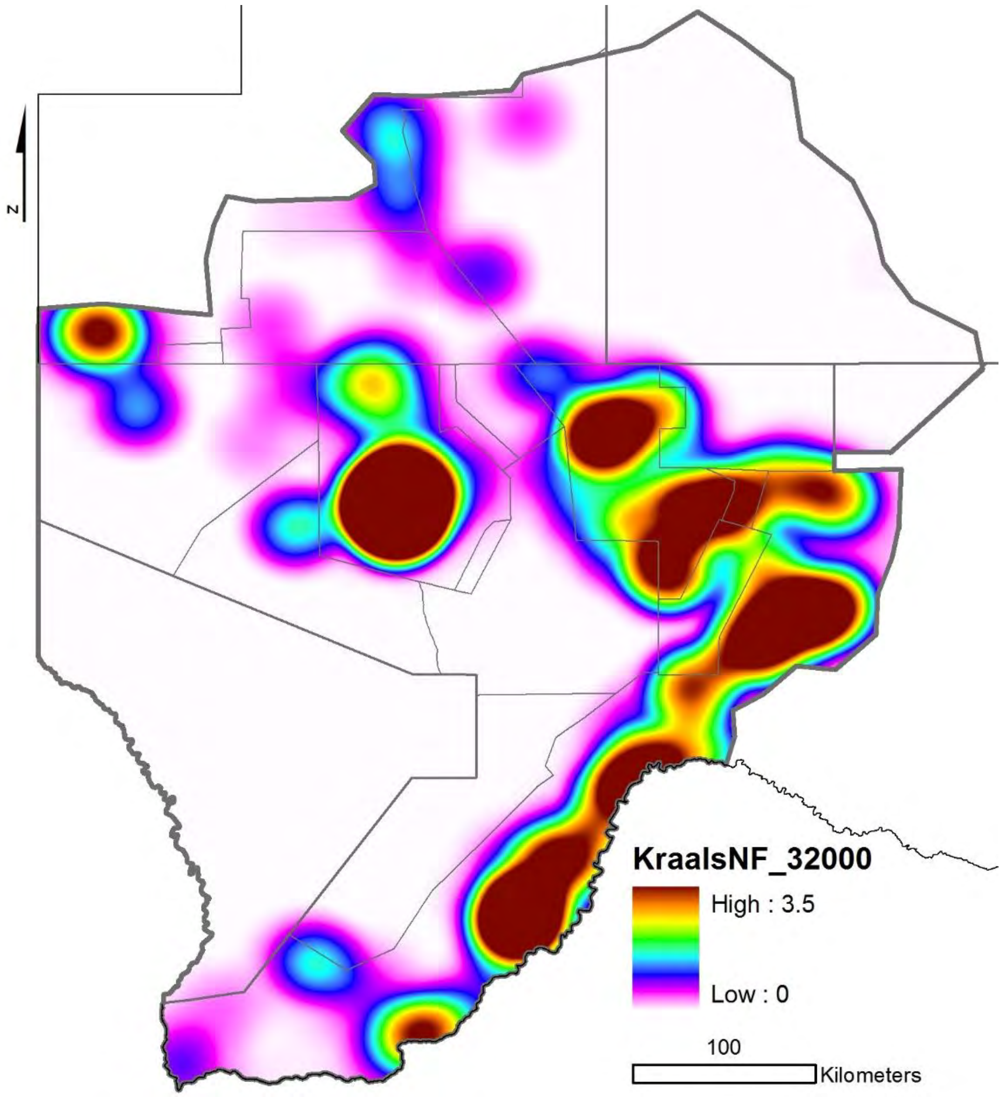


Figure 4 Kraals density (kraals/km^2) measured at the largest kernel density scale (32 km radius). This was the frequently selected scale for kraals density by multi-scale SDM model-building among species. Note encroachment into WMAs.

3.1.2 Discussion of select variables influencing habitat suitability

3.1.2.1 Soils

Soils were, somewhat profoundly, the most consistent predictor variables. Multiple soils variables appeared in every species model with only one exception (duiker). In a sense they are the most fundamental and foundational ecosystem attributes, determining, along with precipitation, the first trophic level (vegetation) upon which the wildlife community is then structured. Soils are directly ingested by ungulates via mineral licks (most prominently at pans) because direct mineral supplementation is essential during lactation and to reset their complex ruminant digestion during seasonal transitions between food type and quality. Besides this notable direct link between soils and wildlife, the manifest nebulous indirect effects can only be speculated. Other than greenness (NDVI) and tree cover (VCF), vegetation communities were unmodelled because we were unable to attain sufficiently precise data on vegetation. Unmodelled plant species distributions, compositions, and abundances almost certainly have large predictive power to wildlife species occurrences (e.g. Cushman and McGarigal 2003). The fact that soils are facilitating these plant communities must explain in part their huge contribution to the HS models. Vegetation type has obvious predictive value for herbivores, including small rodents (i.e. mice, gerbils), that scales up to carnivores and thus encompasses the entire wildlife community. One small/meso carnivore (Cape fox) and one insectivore (bat-eared fox) were predicted exclusively by soils. Certainly, grass and forb species are determined by soils which in turn determine termite species and their concentrations selected by bat-eared fox. Similarly, indirect effects may be speculated working from soils upward for every species modelled.

3.1.2.2 NDVI

NDVI is a measure of the density of photosynthetic vegetation covering a parcel of land, or more simply ‘greenness’. Areas with high NDVI value are characterized by both high biomass and high photosynthetic leaf area. Conversely, areas with low NDVI reflect areas that are dry and denuded. NDVI predicted positively for few species - only leopard, both hyenas and wild dog. For all large herbivore species for which NDVI appeared in the final model (gemsbok, hartebeest, kudu, ostrich, wildebeest), the coefficient was negative, and significant. This is counter to expectation, considering previous studies which related NDVI to broad patterns of antelope distribution measured by DWNP aerial surveys (Verlinden and Masogo 1997), and also for other bulk foragers (like cattle and buffalo (Kaszta et al. 2016) and elephants (Kaszta et al. 2021)). A possible explanation is that highly mobile Kalahari antelopes aggregate in their highest concentrations during early wet season green up, i.e. freshly growing vegetation including the smallest shoots of grasses. Such areas would have a low or at best modest NDVI signature, as vegetation has just started to sprout against a late dry season backdrop. By the time grasses are mature and go to seed their nutritional quality has typically dropped dramatically. Some prominent Kalahari grasses such as *Schmidtia kalahariensis* which commonly forms near monocultures (i.e. the ‘Kalahari cornfield’) are unpalatable when mature due to acidic secretions that can even injure antelope mouthparts. Ruminant herbivores select vegetation based on its quality, and therefore the greenest landscapes (i.e. where plants are their largest and most mature), may be less attractive than effectively browner landscapes of more palatable and higher quality early growth vegetation (e.g. Kaszta et al. 2016). The same pattern of attraction happens at earliest stages after veld fire. Furthermore, Kalahari antelopes tend to favor areas of lower grass stature and less shrub cover due to predation risk by ambush.

3.1.2.3 HP vs kraals vs land use

Human population might be expected to appear in more HS models than it did. Of 32 species it only appeared in the final model for springbok. However, at the 32 km scale human population was highly correlated with kraals (correlation coefficient typically 0.95-0.98), thus human population was frequently dropped during model building by the correlation filter. The density of kraals emerged as the superior explanatory variable of the two, perhaps in part because the variable captures finer resolution micro-settlements while HP was derived from major centres only and provided in coarse scale form without spatial precision. Furthermore, the density of kraals clustered around villages is proportional to the HP of those centres.

Land use is similarly correlated with and a coarse categorical proxy of kraals: PAR being largely occupied by villages and cattleposts, WMA having only limited areas of encroachment and isolated around few villages, and PA being devoid of kraals with the exception of a few areas in CKGR just outside the AOI.

3.1.2.4 KraalsNF vs KraalsF

Most species that responded negatively to kraals, also responded negatively to kraals that were separated from the wildlife space by a fenced boundary. This is important land planners may easily assume that fenced ranches do not impact wildlife beyond their fences, which are often assumed to represent a hard boundary containing potentially negative impacts therein. Notably, transect coverage did not sample inside of fenced farms and ranches. Therefore, the impact of the KraalsF (fenced kraals) variable on species occurrence was measured entirely from beyond the fenced boundaries of those farms, indicating diffuse, broad-scale effects beyond the fence into adjacent areas. The two species with the strongest aversion to kraals, also responded negatively to kraals behind fenced farm boundaries (Figure 5).

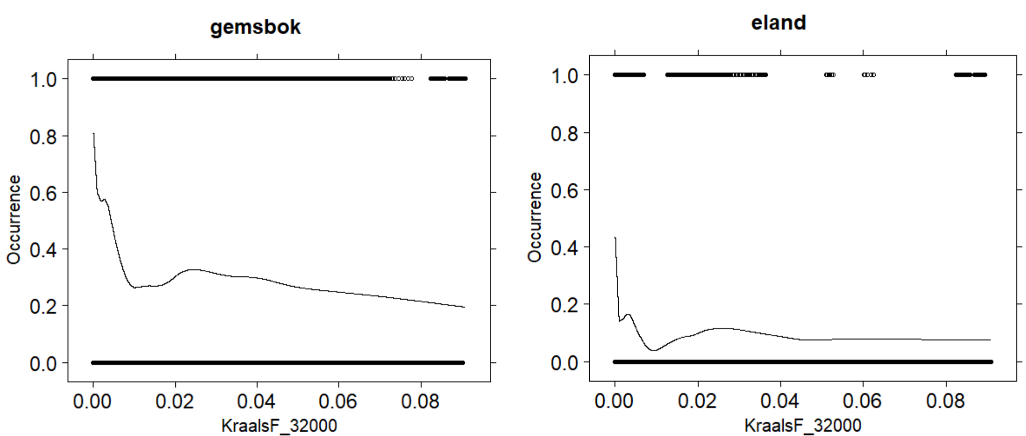


Figure 5 Effect of kraals enclosed inside fenced farms/ranches (KraalsF) on gemsbok and eland probability of occurrence. KraalsF was measured at the largest kernel density scale (32 km radius).

There are many designated ranching land use blocks in the AOI, especially Kgalagadi District, which are intended to be fenced and managed as farms, but have not yet been fenced since allocation. Instead, they function as typical unfenced cattleposts, presumably so awardees can exercise their dual grazing rights (i.e. benefit from grazing beyond their leased or titled boundaries). Importantly, the KraalsF variable does not include these cases. It only includes those areas which were truly fenced at their boundaries, as evidenced in satellite imagery. Thus, the measurable KraalsF effects are not muddied by intended land use, but rather reflect the true situation on the ground. However, some fenced farm boundaries established decades ago (e.g. Ncojane and other TGLP ranches) are now dilapidated, with major gaps and considerable stretches with no fence remaining at all. Moreover, in other areas, although fences are new, gates are sometimes installed to purposefully herd cattle outside and graze in the adjacent WMA. Even where fences are intact and maintained it does not preclude the people residing on the farm from accessing areas beyond for hunting (poaching). These phenomena likely explain the negative response beyond fenced farms exhibited by disturbance-sensitive species.

The take home point is that farms and ranches in the drylands ecosystem have large, broad-scale negative effects beyond proximity to their boreholes or kraals, and far beyond their fenced borders which negatively impact disturbance-sensitive wildlife species in particular. Fence lines cannot be conceptualized as hard boundaries separating agriculture from the wildlife space; kraals comprising the focal point of human-livestock activity inside fenced farms have attenuating impacts for substantial distances beyond those farms.

3.1.3 Select specific highlights

We do not comment on every species individually, but rather selectively highlight specific cases of management interest.

3.1.3.1 Eland and gemsbok

These two species had the highest performing models indicated by the model prediction performance measures displayed in their respective tables (e.g. AUC > 0.9). This is due largely to their exceptional predictability in relation to kraals. Results strongly reflect the fact that these two species, along with their most important predator (lion), are the most disturbance-sensitive Kalahari wildlife species and therefore most dependent on land uses free from pastoral/agricultural activity. These species are absent in areas with even low densities of kraals within a 32 km radius. One notable aberration in both eland and gemsbok HS models was the relatively high value predicted for areas directly south of KTP (KD27). Cattleposts within this communal grazing area are comprised of comparatively low numbers of kraals, thus gemsbok and eland were predicted to occur here based on kraals density variables, when in fact it is a dead zone. But there is more to this story in the context of landscape fragmentation and connectivity which can be explored in Phase 2.

3.1.3.2 Pangolin

Pangolin are rare in the landscape as evidenced by their limited records relative to other species (see Table 1). We are however confident that their rate of false absences (missed detections) on the prepared (dragged) comprehensive species track transects was very low and not differing from other species. Of those infrequent records several field observations occurred inside PAR land use types. Pangolin showed positive but non-significant occurrence probability responses to both KraalsNF and KraalsF at 32 km scale (see also Section 3.2 (e) below). Their apparent lack of avoidance of pastoral areas may put them into more frequent contact with people and facilitate their trafficking within illegal wildlife trade networks for traditional medicines regionally and in Asia. This habitat suitability pattern may be a hint of insight to their heightened extinction risk as a species and family.

3.1.3.3 Springbok

Results corroborate the DWNP aerial survey record showing springbok population decline and range contraction. Their predicted high probability of occurrence landscape is now confined to KD1,2 and KTP, as well as adjacent KD3, particularly in the northwest areas centered around Ohe and Hunhukwe pans.

The springbok map also serves as a good example highlighting the remarkable predictive power of the habitat suitability models generally. The final springbok model predicted isolated meta-populations in the vicinity of the intersection of KD12,13, and SO1 in between the villages of Inalegolo, Kokong and Morwamosu, as well in the Okwa valley near the Trans-Kalahari highway crossing. These two areas where either not sampled (latter case) or failed to reveal detections (former case). Small, isolated, and most probably declining meta-populations are known from both locations however, outside of the track transect sampling effort.

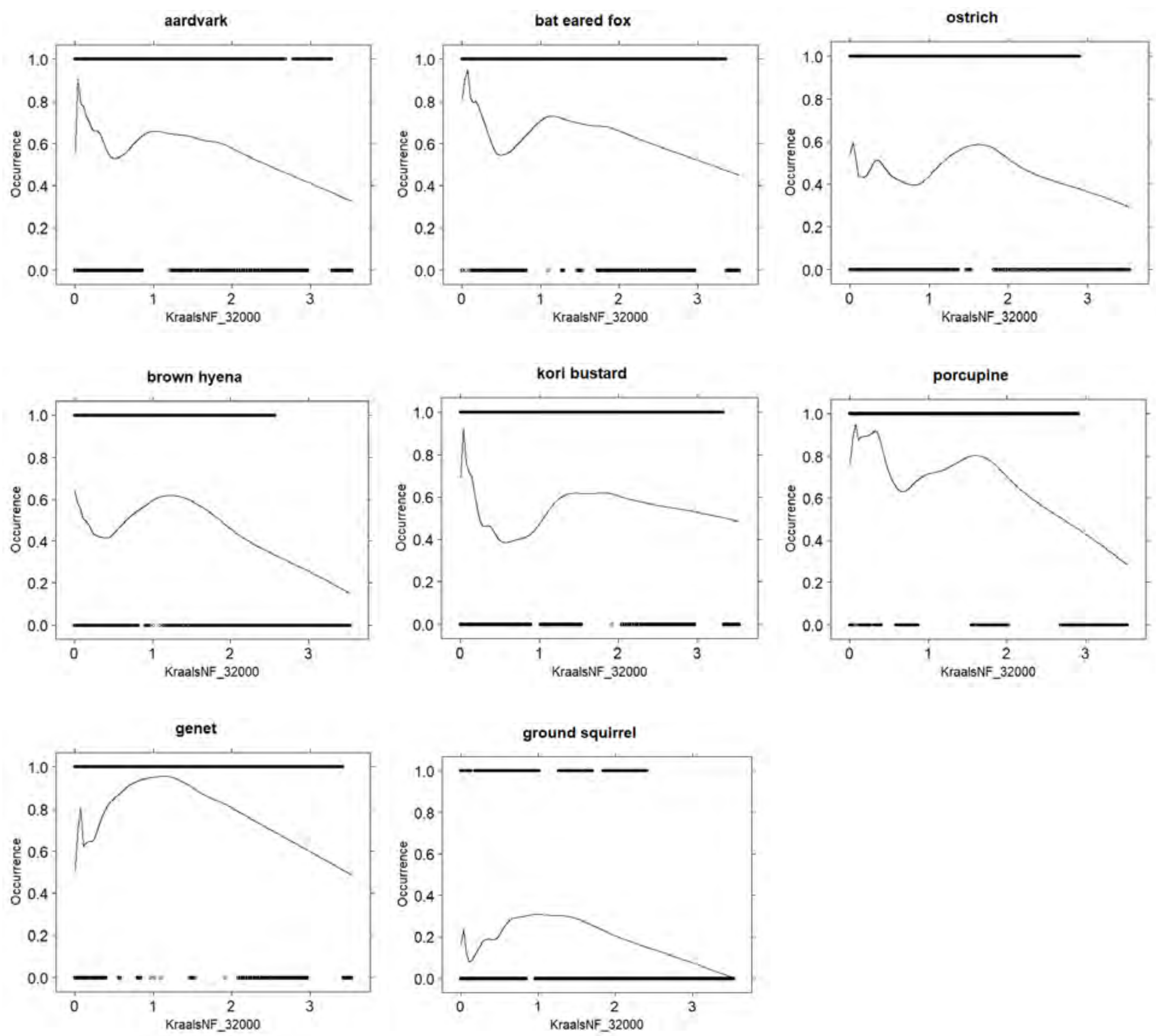
3.1.3.4 Wildebeest

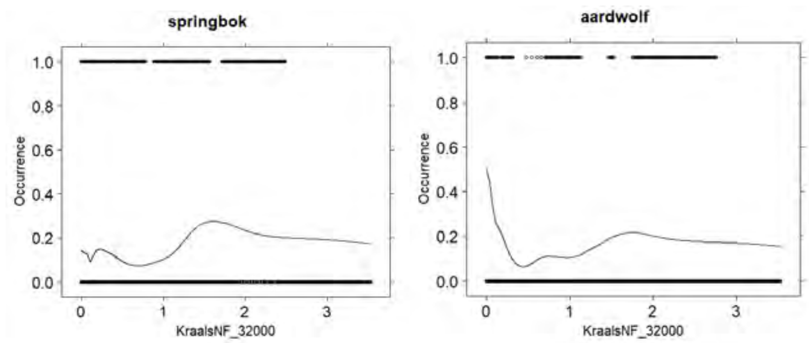
Among large antelopes in the under-researched AOI, wildebeest have had the most extensive research effort using geolocational telemetry collars. Foundational studies were carried out during 1989-92 (Bonifica 1992) and 1998-99 (Environment and Development Group 1999, DWNP 2000). These stressed the importance of southern GH11, and particularly the fidelity wildebeest showed to the area of intersection between GH11,13, and KD1. Remarkably, applying the track database collected 20-40 years after these radio-collaring studies, the final wildebeest model predicted this the highest probability occurrence hotspot throughout the entire Kalahari. It coincides with a narrowing pinchpoint of landscape connectivity threatened by proximity to Ncojane TGLP ranches, largely operated as unfenced cattleposts, the existence and potential expansion of the unofficial settlement of Ranyane, and most concerningly, the expansion of borehole allocations in the RAD development zone to the north and west of Ncaang. Wildebeest are known for their high wet season range fidelity in other regions (Morrison & Bolger 2012). This area centered around GH13 appears to be a remarkable example of such wet season range fidelity, despite the intervening decades when encroachment has intensified both in the surrounding vicinity as well as northeastwards towards Bere/Kacgae and Okwa valley, this latter fact probably increasing isolation of CKGR wildebeest (Selebatso 2017). Although connection to CKGR remains relevant, connectivity to the south in KD1,2 and KTP may now be more important for the bulk of remaining Kalahari wildebeest population to access this interdecadal hotspot centered on GH13, and this is imminently threatened by RAD cattlepost proliferation around Ukwi, Ngwatle and Ncaang. Spatial quantification of such risks will be measured in Phase 2 using the most disturbance-sensitive species.

3.2 Species responses to human-livestock disturbance

The following plots depict each species' occurrence probability response to the density of kraals (kraals/km^2) in the free-ranging landscape (KraalsNF) at the largest (32 km radius) kernel scale. Presence/Absence data points are also displayed from each logistic model (i.e. Occurrence 1 or 0), which can help in the interpretation of response in addition to the plsmo spline (line). Bear in mind the points themselves appear often as lines because there are hundreds or thousands each of presences and absences (see processed locational data in Table 1). Species are grouped a-e below according to commonality in their response. Figure 4 can assist in interpretation of the plots as the x-axis scale visualized in the landscape.

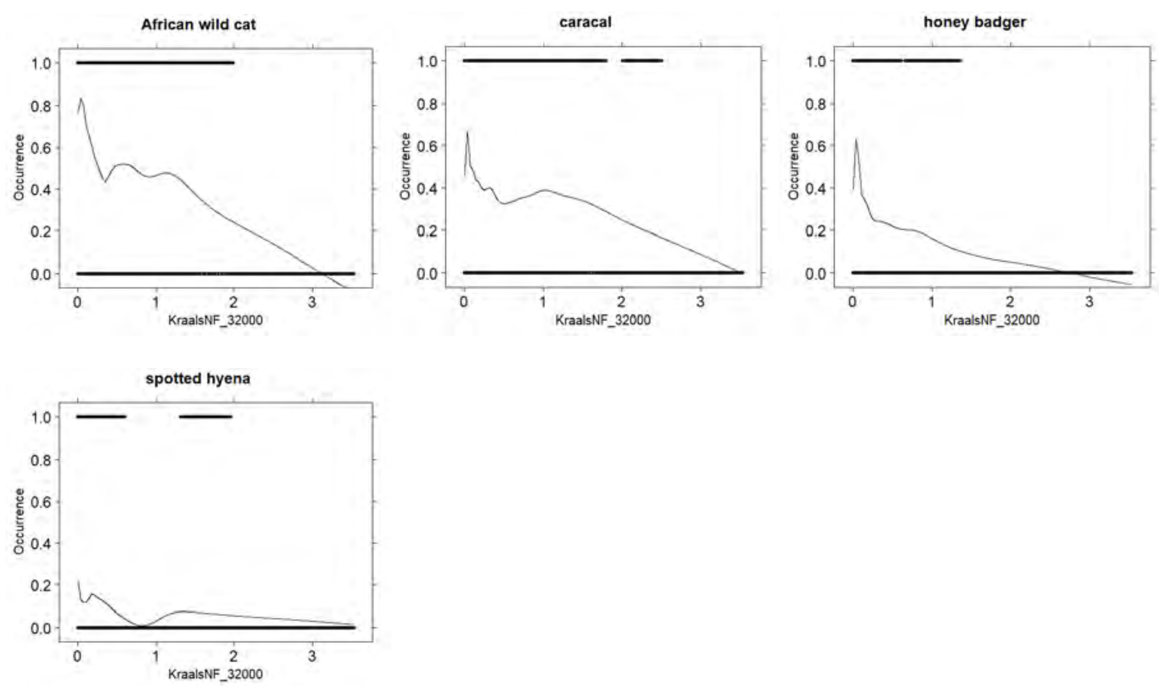
a) humped, decline at higher kraal density





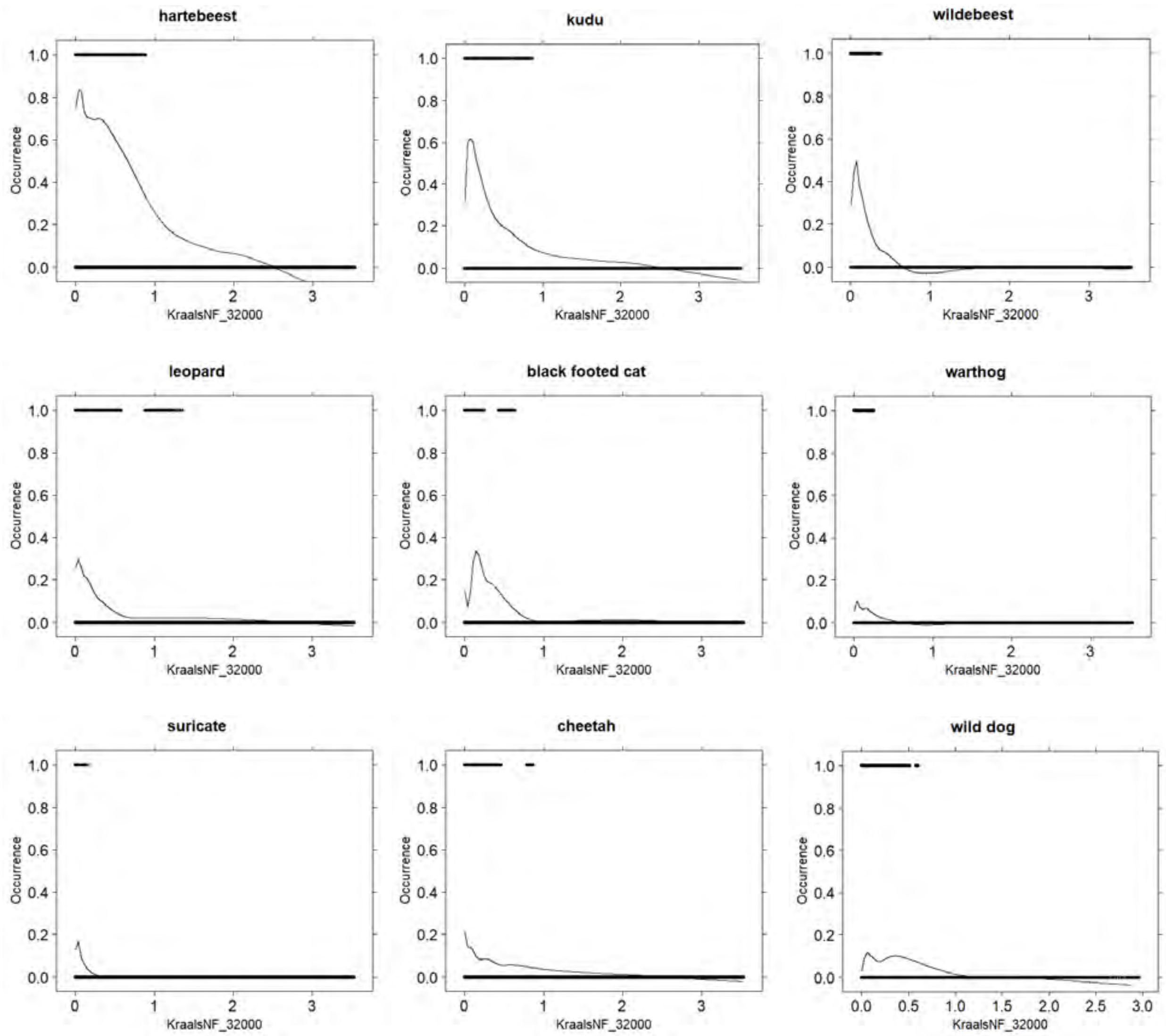
These species generally reach their highest occurrence probability near zero but have a secondary hump at intermediate kraals density. Genet, ground squirrel and springbok are exceptions that reach their highest occurrence probability at intermediate kraals density. Ostrich and brown hyena are nearly equivalent in highest probability both at zero and intermediate density. All these species show decline at highest density of kraals but appear tolerant of cattlepost country at the typical 6-8 km spacing of borehole allocations.

b) Negative, absent at highest density



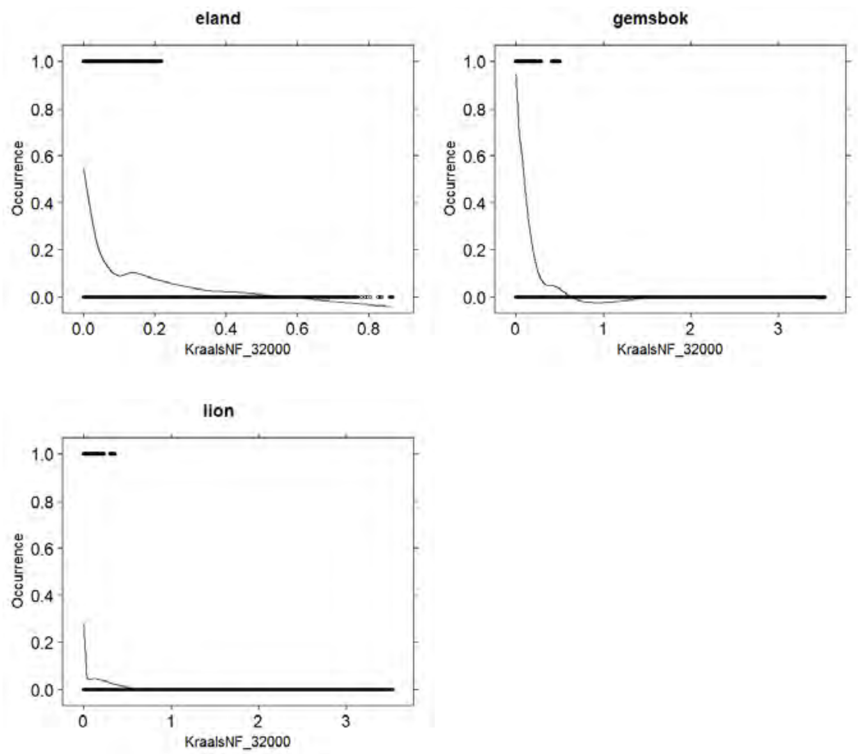
These species decline more steadily with increasing kraals density than the previous species and are absent at highest kraals density. African wild cat and caracal especially also show a minor hump or sub-peak at intermediate kraals density, but less prominently than the previous species.

c) Negative, absent before high density



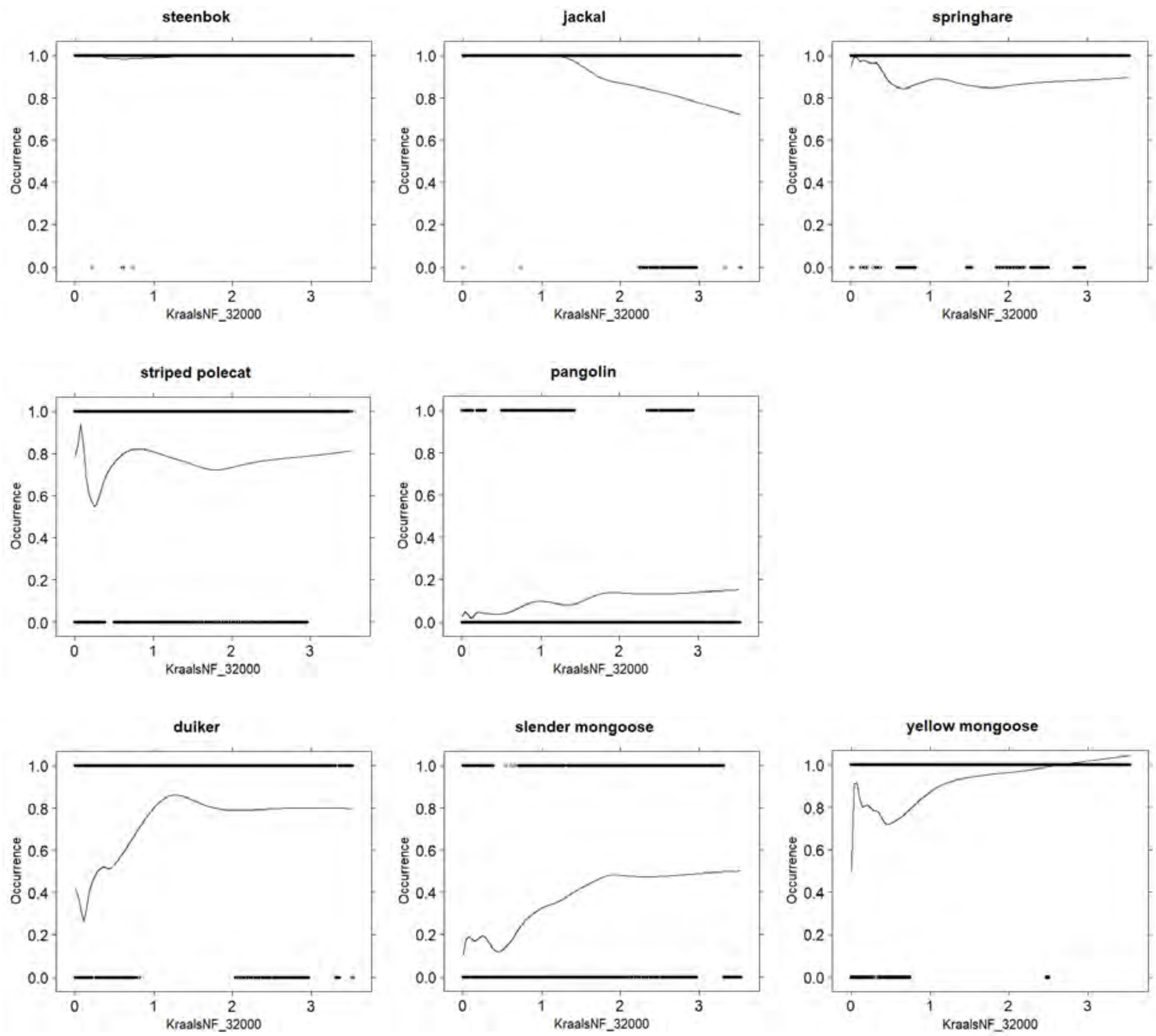
These species’ occurrence probabilities decline markedly as kraals density increases. They may occur at low probability in cattlepost country but are generally excluded from most areas of intermediate kraals density and absent before highest kraals densities are reached.

d) Strongly negative; absent even at relatively low kraal density



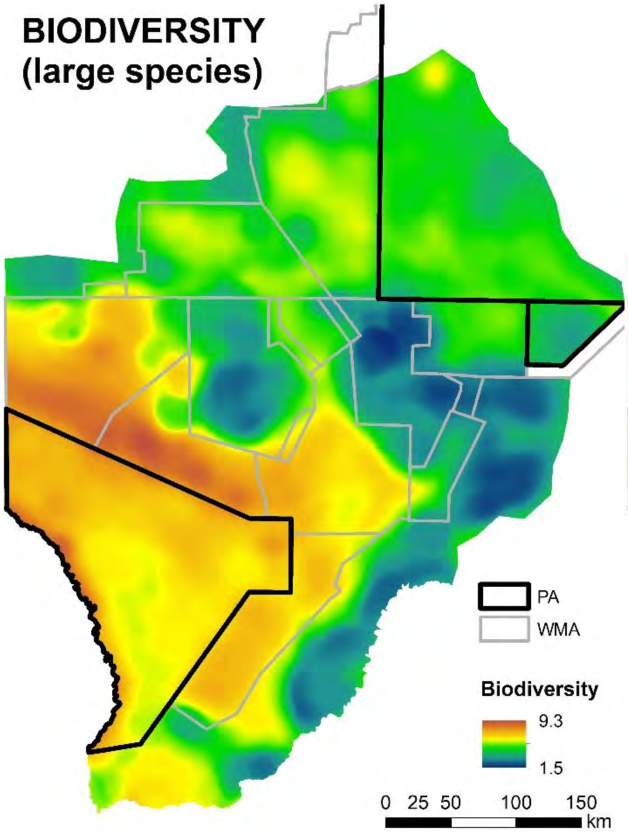
Note the differing x-axis scale for eland. These 3 species decline sharply at any density of kraals in the landscape at the largest 32 km radius scale. Their responses can also be interpreted in terms of proximity to nearest kraals/cattlepost, i.e. there are gaps in the landscape between where kraals are and where these species begin to occur at a probability above zero.

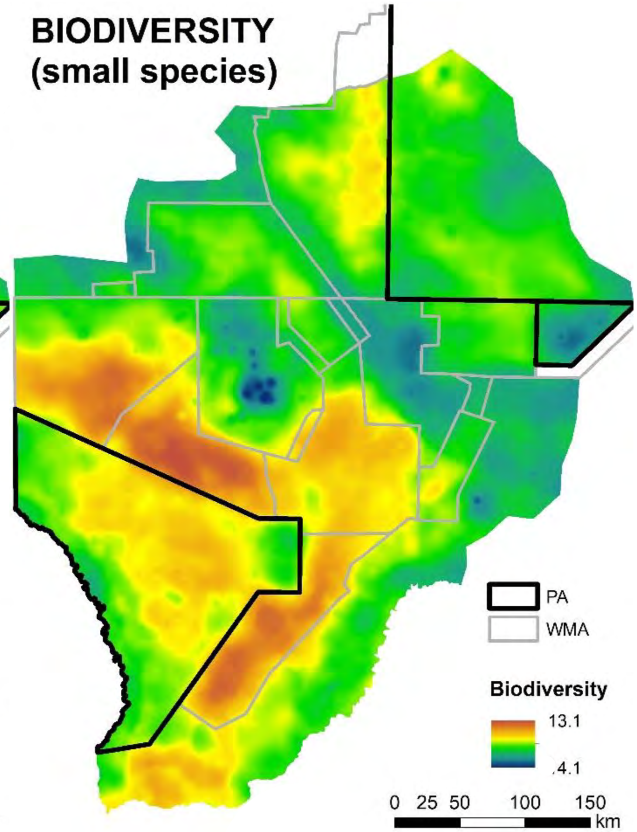
e) Neutral or positive

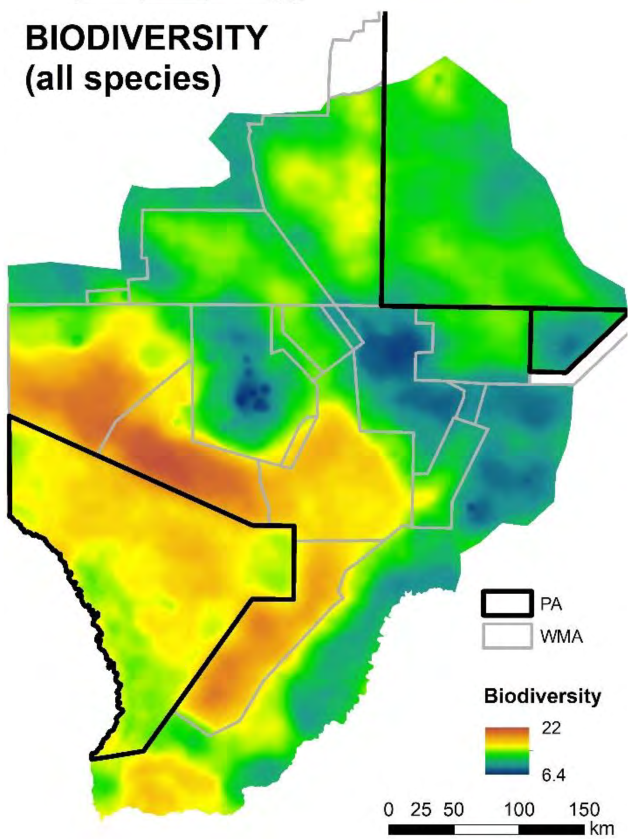


These species either have no apparent negative response to increasing kraals density, or in fact become more abundant in areas of higher kraals density. Hare had no absences and could therefore not be plotted. Had it been it would most resemble the steenbok plot i.e. a perfect horizontal line at occurrence probability 1 with no observations at 0. In other words, ubiquitous regardless of kraal density. Jackal is an exception among the group in showing a minor decline in areas of highest kraals density, although still common there. The remaining species: springhare, striped polecat, pangolin, duiker, slender and yellow mongooses all increase as landscape kraals density increases. Yellow mongoose shows an increasing trend even beyond the highest kraals density measured in the landscape at this scale.

3.3 Biodiversity maps







Once again, these biodiversity maps are the sums of the individual species occurrence probability surfaces (HS maps), which show the expected number of modelled species occurring at each pixel of the AOI (Penjor et al. 2021). The most prominent visual impression is the exceptional value of the ungazetted KD WMAs. Lest this be misinterpreted to the effect: "the KD WMAs are even more important than the KTP" - this is not a correct statement. KD WMAs continue to exhibit high value in large measure because they are contiguous with each other and the KTP. In the science and philosophy of landscape connectivity, their whole is greater than the sum of each part alone; that is, to remove, isolate or otherwise lose any piece of the puzzle - the KTP or any adjacent WMA - would predictably render whichever remaining smaller piece much less valuable. The KD WMAs would be immensely deprived if the KTP ceased to exist in its present functioning form, and vice versa. Shifting one step up to the larger AOI scale, if this KTP-KD WMA core were to become isolated from the CKGR core, it will predictably cause areas to decline in value. What this means specifically is wildlife populations decline, ranges contract, and extirpations increase in likelihood over shorter timeframes when habitat extent and connectivity among habitat core areas decline.

These phenomena underpinning the importance of extensive core habitats and connectivity among them have already been observed and documented at the next step up in scale - i.e. the isolation of the AOI from connection to Makgadikgadi and Okavango delta due to erection of game-proof fencing, and similar loss of connectivity to Namibia and South Africa. Wildebeest and hartebeest populations particularly experienced well-documented die-offs (Child 1972, Owens and Owens 1983, Parry 1987, Williamson and Mbano 1988, Spinage 1992, Thouless 1998) from which they have never recovered (DWNP 2015). A hard lesson of the racheting down effect of landscape fragmentation via human-livestock encroachment facilitated by uninformed planning, and reminder of the paramount goal of the present KGDEP to keep the two remaining cores within the AOI connected.

While KD WMAs would not stand on their own in any comparable measure of wildlife value in the absence of the KTP, it seems that their boundaries in fact may encompass richer ground than the KTP itself, providing critical limited resources to highly mobile wildlife that periodically retreat to the protected area. The Schwelle has long been noted the area of highest wildlife concentration in the Kalahari (DHV 1980, Bonifica 1992), speculatively due to the density of pans and relatively higher mineral concentration soils. The core of the Schwelle centred at the Matsheng villages (Hukuntsi, Tshane, Lehututu, Lokgwabe), and its extended axis (Kokong, Khakhea, Mabutsane, Sekoma in the southeast, Ncojane, Kule in the northwest) has been encroached and compromised by historical settlement and agricultural expansion diffusing from those centres.

Present day WMAs thus encompass the best of what remains of the Schwelle. Initial baseline assessments of the Schwelle stressed the importance of wildlife access to critical resources beyond the protected area boundaries if populations are to remain large and mobile as opposed to small and sedentary (Bonifica 1992). Massive seasonal redistributions of antelope populations of global note into the WMAs recorded in recent years (Keeping et al. 2018) attest to the importance of ungazetted KD WMAs effectively safeguarding what remains of the Schwelle via restrictions to borehole allocations for cattlepost expansion within their borders upheld by KD landboard.

3.4 Tentative inferences on core areas and connectivity from HS models

KTP plus the adjacent 4 WMAs (KD1,2,12,15) represent the largest and highest value core for wildlife within the AOI. In the context of landscape connectivity, this primary core can be viewed in relation to a second core of CKGR plus adjacent 2 WMAs (GH10, KW2). The habitat suitability maps, especially for large species, together with the cumulative (biodiversity) probability occurrence maps reflect 3 connections linking the two cores which mirror the configuration of interconnecting WMAs (GH11/13, KD5/6, SO2/KW6). Thus, the present rigorous analysis of wildlife habitat suitability using a novel comprehensive wildlife locational dataset corroborates earlier CI WKCC project findings highlighting these 3 potential corridors (Meyer and Meyer 2018).

Among the most disturbance-sensitive species, their habitat suitability maps indicate continuity through 2 of the 3 corridors previously highlighted by CI; what we will refer to hereafter as the 'central' corridor (KD5,6,12) and the 'western' corridor (GH11,13, KD1). These two corridors appear to have differing importance to different species. They are perhaps no better visualized than using the gemsbok HS map as example (see page 29). Gemsbok has continuous high habitat suitability through the central corridor linking core populations, while its habitat suitability declines and potentially fragments in the western corridor at the pinchpoint of encroachment between Ranyane and Ncaang. By contrast, eland and lion seem to have a stronger connection through the western corridor and face a large gap of nearly zero probability in the central corridor pinchpoint between KD12 and KD5,6.

The eastern corridor (SO2, KW6) appears relatively low value for movement of these three highly sensitive species due to its narrow width and high level of encroachment. KW6 has become compromised by encroachment of cattleposts at evenly spaced allocations as though Kweneng landboard has regarded it as PAR/CGA rather than WMA. The entire eastern corridor will become irrelevant for large wildlife anyway as SO2 is now intended to be carved into a fenced farm layout. The western corridor is immediately threatened by RAD allocations expanding from Ukwi, Ngwatle, Ncaang and by the location of Ranyane. The central corridor is threatened by misallocated boreholes inside KD12 and 6, by proposals to occupy KD5, and by a complete severing with new farms and fencing along the highway between Kang and Hukuntsi. The western and central corridors join in southeastern GH11 to form a single pinchpoint of connectivity which crosses the Trans-Kalahari highway. This singular remaining link between KTP and CKGR for disturbance-sensitive species is threatened by recent RAD allocation expansion south of Kacgae. All these issues will be explored and their impacts on wildlife connectivity quantified through the scenario modelling in Phase 2.

3.5 Look ahead to Phase 2 connectivity modelling

Phase 2 connectivity modelling is computationally intensive and time-consuming. This forces a trade-off between species examined and number of scenarios. There are a host of already allocated, planned and proposed agricultural developments (RAD borehole allocations, WMA dezonings, fenced ranch expansions) poised to affect precisely the remaining precarious corridors of wildlife connectivity. Ideally, each of these future changes to the landscape can be quantified to provide land planners with the most complete information with which to optimize their land planning decisions. To achieve this, we are opting to minimize the species considered in order to maximize the future scenarios information. We require a model species which is sensitive to the types of future changes coming to the landscape (i.e. more fences, farms, cattleposts).

3.5.1 Gemsbok as model, supplemented with eland and lion

Gemsbok model performance returned the highest AUC (0.93) among all species (although eland is a very close second). Arguably, gemsbok has the most reliable prediction of all models. Gemsbok distribution is explained by kraals (cattleposts) in the free-ranging landscape at exceedingly high probability (to the -170 decimal place p-value with zero standard error), and kraals on fenced farms at similarly ridiculously high probability (-43 decimal place p-value). The gemsbok HS map exhibits the sharpest visible gradient and contrast between high and low probability of occurrence areas. Essentially, gemsbok occurs at high probability throughout the Kalahari landscape in areas remote from human-livestock disturbance.

The pattern of attenuating avoidance of cattleposts by gemsbok detected by the multi-scale SDM model-building which related many thousands of gemsbok presence/absence locations to many thousands of kraal locations throughout the Kalahari landscape is remarkably quantitatively consistent. Previous studies have noted this striking pattern of gemsbok distribution in relation to human-livestock disturbance (e.g. DHV 1980, Bergstrom and Skarpe 1999) and some tentatively argued that hunting pressure is the cause (e.g. Verlinden et al. 1998). From a land use planning perspective, the cause is somewhat irrelevant. What is important rather is that the pattern exists, and it is highly predictable in relation to the planning that landboards have control over (i.e. borehole locations).

We therefore nominate gemsbok as an umbrella species in Phase 2 modelling for others less sensitive to disturbance, and a flagship species whose probability of occurrence reflects the remaining free-ranging Kalahari wilderness. Gemsbok are not yet a species of exceptional conservation concern. They are the most arid-adapted large-bodied Kalahari antelope and their population has fared well compared to other species more prone to long-distance drought-induced migrations like hartebeest, wildebeest, and springbok. However, the fact that Kalahari gemsbok do persist with high probability virtually everywhere that agricultural development is not – makes them an excellent candidate for connectivity modelling because it is exactly the distribution of livestock and people through the allocation of boreholes and ranching blocks that threatens conservation in this drylands ecosystem as recognized by the present UNDP project, and, most importantly, precisely that which land planning has control over. Noteworthy also is the fact that, while southern oryx (gemsbok) are still common in the Kalahari, 2 out of 4 global oryx species have been hunted (poached) to extinction in the wild. Gemsbok themselves, within their limited southern African range, are now more common in captivity (game ranches) than in the wild (East 1998, Relton et al. 2016). These facts, along with their clear high sensitivity to pastoral land use at broad scales, further support the application of gemsbok as a model umbrella species for the remaining free-ranging Kalahari wildlife community.

We will apply gemsbok to the full range of Phase 2 scenarios, but also model the other two highly disturbance sensitive species (eland and lion) against a limited subset of packaged scenario options to supplement the evidence base for land planners and policy makers to make informed choices regarding the future of the Kalahari landscape.

References

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

Andrew MH (1988) Grazing impact in relation to livestock watering points. *Trends in Ecology & Evolution* 3(12): 336-339.

Bergström R, Skarpe C (1999) The abundance of large wild herbivores in a semi-arid savanna in relation to seasons, pans and livestock. *African Journal of Ecology* 37(1): 12-26.

Bonifica (1992) *Technical assistance to the project: initial measures for the conservation of the Kalahari ecosystem*. Final Report under EDF project No. 6100.026.14.001 to the Department of Wildlife and National Parks, Government of Botswana. Bonifica, Rome, Italy.

Child G (1972) Observations of a Wildebeest die-off in Botswana. *Arnoldia* 31: 1-13.

Cushman SA, Macdonald EA, Landguth EL, Malhi Y, Macdonald DW (2017) Multiple-scale prediction of forest loss risk across Borneo. *Landscape Ecology* 32(8): 1581-1598.

Cushman SA, McGarigal K (2003) Landscape-level patterns of avian diversity in the Oregon Coast Range. *Ecological Monographs* 73(2): 259-281.

de Sousa L, Poggio L, Batjes NH, Heuvelink GBM, Kempen B, Ribeiro E, Rossiter D SoilGrids 2.0: producing quality-assessed soil information for the globe. Under submission to SOIL

DHV (1980) *Countrywide Animal and Range Assessment Project*. Final Report to Ministry of Commerce and Industry & Department of Wildlife and National Parks, Government of Botswana. DHV Consulting Engineers, Amersfoort, Netherlands.

DWNP (2015) Aerial census of animals in south-west Botswana. Dry season. Republic of Botswana.

DWNP (2000) Address to Gantsi District Council on Wildebeest Movements in Relation to Wildlife management areas in Gantsi District, 12 Dec 2000, Shimane W Makhabu, Senior Wildlife Biologist.

East R (1999) African Antelope Database, 1998. Gland Switzerland: IUCN SSC Antelope Specialist Group.

Elliot NB, Cushman SA, Macdonald DW, Loveridge AJ (2014) The devil is in the dispersers: predictions of landscape connectivity change with demography. *Journal of Applied Ecology* 51(5): 1169-1178.

Environment and Development Group (1999) Wildlife Conservation and Utilization in Central and Southern Botswana, Project No. 7-ACP-BT-022, Final Report Senior Wildlife Biologist (Population Monitoring) 1996-1999, Commission of the European Communities, The Environment and Development Group Oxford, for Republic of Botswana, Gaborone, June 1999.

Environmental Systems Resource Institute (2012) ArcGIS Desktop Sofware, Release 10. Redlands, CA, USA.

Kaszta Ż, Cushman SA, Slotow R (2021) Temporal Non-stationarity of Path-Selection Movement Models and Connectivity: An Example of African Elephants in Kruger National Park. *Frontiers in Ecology and Evolution* 9: p.207.

Kaszta Ż, Cushman SA, Sillero-Zubiri C, Wolff E, Marino J (2018) Where buffalo and cattle meet: modelling interspecific contact risk using cumulative resistant kernels. *Ecography* 41(10): 1616-1626.

Keeping D, Keitsile AO, Burger JH (2018) Eland in Botswana's Kalahari: now the largest free-ranging population in Africa? *Gnusletter* 35: 21-26.
Liebenberg L (1990) *A field guide to the animal tracks of southern Africa*. David Philip.
Macdonald DW, Chiaverini L, Bothwell HM, Kaszta Ż, Ash E, Bolongon G, Can ÖE, Campos-Arceiz A, Channa P, Clements GR, Hearn AJ (2020) Predicting biodiversity richness in rapidly changing landscapes: climate, low human pressure or protection as salvation? *Biodiversity and Conservation* 29(14): 4035-4057.
Macdonald DW, Bothwell HM, Kaszta Ż, Ash E, Bolongon G, Burnham D, Can ÖE, Campos-Arceiz A, Channa P, Clements GR, Hearn AJ (2019) Multi-scale habitat modelling identifies spatial conservation priorities for mainland clouded leopards (Neofelis nebulosa). *Diversity and Distributions* 25(10): 1639-1654.
Mateo-Sánchez MC, Cushman SA, Saura S (2014) Scale dependence in habitat selection: the case of the endangered brown bear (Ursus arctos) in the Cantabrian Range (NW Spain). *International Journal of Geographical Information Science* 28(8): 1531-1546.
McGarigal K, Wan HY, Zeller KA, Timm BC, Cushman SA (2016) Multi-scale habitat selection modeling: a review and outlook. *Landscape Ecology* 31(6): 1161-1175.
Meyer T, Meyer KC (2018) Integrating Multiplatform Remote Sensing and In Situ Datasets for Socio-Ecologically Sustainable Conservation Corridors: The Western Kalahari Conservation Corridor (WKCC) Project. In Liang (Ed), Comprehensive Remote Sensing 9: 220-235, Oxford: Elsevier.
Morrison TA, Bolger DT (2012) Wet season range fidelity in a tropical migratory ungulate. *Journal of Animal Ecology* 81(3):543-552.
Murphy MA, Evans JS, Storfer A (2010) Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. Ecology 91(1): 252–261
Owens M, Owens D (1983) Fences of death. *Wildlife* June: 214-217.
Parry D (1987) Wildebeest Connochaetes taurinus mortalities at Lake Xau, Botswana. *Botswana Notes and Records* 19: 95-101.
Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models using logistic regression. Ecological Modelling *133*(3): 225-245.
Penjor U, Kaszta Ż, Macdonald DW, Cushman SA (2021) Prioritizing areas for conservation outside the existing protected area network in Bhutan: the use of multi-species, multi-scale habitat suitability models. *Landscape Ecology* 36(5): 1281-1309.
Perkins JS (2018) Southern Kalahari piospheres: Looking beyond the sacrifice zone. *Land Degradation & Development* 29(9): 2778-2784.
R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna. Austria. URL https://www.R-project.org/
Relton C, Selier J, Strauss WM (2016) A conservation assessment of Oryx gazella. In Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT, editors. The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa
Selebatso M (2017) Wildebeest ecological dynamics in the CKGR: The implications for potential isolation from the wider Kalahari ecosystem (Doctoral dissertation, University of Botswana).
Spinage CA (1992) The decline of the Kalahari wildebeest. *Oryx* 26: 147-150.
Thouless CR (1998) Large mammals inside and outside of protected areas in the Kalahari. *Transactions of the Royal Society of South Africa* 53(2): 245-255.

Verlinden A, Perkins JS, Murray M, Masunga G (1998) How are people affecting the distribution of less migratory wildlife in the southern Kalahari of Botswana? A spatial analysis. *Journal of Arid environments* 38(1): 129-141.

Verlinden A, Masogo R (1997) Satellite remote sensing of habitat suitability for ungulates and ostrich in the Kalahari of Botswana. *Journal of Arid Environments* 35(3): 563-574.

Wan HY, Cushman SA, Ganey JL (2019) Improving habitat and connectivity model predictions with multi-scale resource selection functions from two geographic areas. *Landscape Ecology* 34(3): 503-519.

Wasserman TN, Cushman SA, Shirk AS, Landguth EL, Littell JS (2012) Simulating the effects of climate change on population connectivity of American marten (Martes americana) in the northern Rocky Mountains, USA. *Landscape Ecology* 27(2): 211-225.

Williamson DT, Mbano B (1988) Wildebeest mortality during 1983 at Lake Xau, Botswana. *African Journal of Ecology* 26: 341-344.