Bond, M. L., Bradley, C. M., Kiffner, C., Morrison, T. A. and Lee, D. E. (2017) A multi-method approach to delineate and validate migratory corridors. *Landscape Ecology*, 32(8), pp. 1705-1721. (doi:10.1007/s10980-017-0537-4)

A multi-method approach to delineate and validate migratory corridors

Monica L. Bond · Curtis M. Bradley · Christian Kiffner · Thomas A. Morrison · Derek E. Lee

M. L. Bond (Corresponding author) · D. E. Lee

Wild Nature Institute, 15 North Main St. Suite 208, Concord, New Hampshire 03302, USA

e-mail: monica@wildnatureinstitute.org

phone: +1-415-763-0348

C. M. Bradley

Center for Biological Diversity, P. O. Box 710, Tucson, Arizona 85701, USA

C. Kiffner

The School for Field Studies, Center for Wildlife Management Studies, P.O. Box 304, Karatu, Tanzania

T. A. Morrison

Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Graham Kerr Building, Glasgow G12 8QQ, UK

May 12, 2017

Manuscript word count: 8943

Abstract

Context Managers are faced with numerous methods for delineating wildlife movement corridors, and often must make decisions with limited data. Delineated corridors should be robust to different data and models.

Objectives We present a multi-method approach for delineating and validating wildlife corridors using multiple data sources, which can be used conserve landscape connectivity. We used this approach to delineate and validate migration corridors for wildebeest (*Connochaetes taurinus*) in the Tarangire Ecosystem of northern Tanzania.

Methods We used two types of locational data (distance sampling detections and GPS collar locations), and three modeling methods (negative binomial regression, logistic regression, and Maxent), to generate resource selection functions and define resistance surfaces. We compared two corridor detection algorithms (cost-distance and circuit theory), to delineate corridors. We validated corridors by comparing random and wildebeest locations that fell within corridors, and cross-validated by data type.

Results Both data types produced similar resource selection functions. Wildebeest consistently selected migration habitat in flatter terrain farther from human settlements. Validation indicated three of the combinations of data type, modeling, and corridor detection algorithms (detection data with Maxent modeling, GPS collar data with logistic regression modeling, and GPS collar data with Maxent modeling, all using cost-distance) far outperformed the other seven. We merged the predictive corridors from these three data-method combinations to reveal habitat with highest probability of use.

Conclusions The use of multiple methods ensures that planning is able to prioritize conservation of migration corridors based on all available information.

Keywords connectivity · *Connochaetes taurinus* · circuit theory · cost distance · land-use planning · Least Cost Path Analysis · Maxent · migration · resource selection functions · wildebeest

Introduction

A variety of animal taxa undertake seasonal long-distance migrations, defined as round-trip movements between discrete areas not used at other times of the year (Berger 2004), as a means to avoid predation and severe weather, and to maximize resource intake (Fryxell et al. 1988; Holdo et al. 2009; Poor et al. 2012). Long-distance migrations provide large-scale ecological benefits and services such as nutrient transfer, soil fertilization, and seed dispersal (Hamilton et al. 1998; Holdo et al. 2009; Estes 2014). However, migrations are in precipitous decline globally because of rapid environmental change in many of the landscapes in which they still occur (Bolger et al. 2008; Harris et al. 2009; Singh and Milner-Gulland 2011). The linear shape of many corridors (Sawyer et al. 2009) make migrations particularly sensitive to the effects of habitat loss and fragmentation, particularly in areas that act as natural bottlenecks such as valleys or passes (Morrison and Bolger 2014). Recent efforts to quantitatively delineate and validate corridors to conserve migratory species and their associated ecological services have significantly improved the resolution and accuracy of corridor planning, but relatively little effort has been made to compare amongst the many types of data and analyses that form the basis of corridor prediction (Berger 2004; Sawyer et al. 2009; Epps et al. 2011; Poor et al. 2012; McClure et al. 2016).

Typical methods to identify animal migration corridors use (1) spatially explicit animal location data, (2) a landscape ‘resistance surface’ as the inverse of a suitable habitat model, to provide a quantitative estimate of how environmental parameters impede or facilitate animal movement (reviewed by Zeller et al. 2012), and (3) a cost-based algorithm that defines the migratory pathways between seasonal core areas through the resistance surface (e.g., cost distance; Adriaensen et al. 2003 or circuit theory; McRae et al. 2008). The use of resistance surfaces to represent the landscape between migratory endpoints and algorithms to delineate movement pathways has increased over the last decade (Sawyer et al. 2009; Poor et al. 2012; LaPoint et al. 2013; Cushman et al. 2013; McClure et al. 2016). Beier et al. (2008) described many of the steps, choices, and assumptions involved in corridor (or linkage) delineation and design, and Zeller et al. (2012) provided guidance for defining the resistance surface(s). Both Beier et al. (2008) and Zeller et al. (2012) noted additional research gaps regarding methodological comparisons that we attempted to address in this study.

Land managers often must make conservation and management decisions in settings with limited data, particularly when events they are attempting to observe and protect, such as long-distance migrations, are unpredictable in timing and location (Sawyer et al. 2009; Singh et al. 2012). One data type, active individual tracking, directly observes animal movements but is expensive to collect and therefore often limited in sample size (Hebblewhite and Haydon 2010). Another data type, point observations, is less expensive and less invasive than collars and measures animal distributions well, but not their movements. Given the various data types and model structures in common use, there is a need to assess the sensitivity of the predicted corridor to data source and resistance surface model type, in order to ensure that the delineated corridor is robust and neither data- nor model-dependent. Here, we present an analytical framework for corridor delineation projects, using a multi-method approach with multiple data types, resistance surface models, corridor algorithms, and validation techniques.

The Tarangire Ecosystem (TE) in northern Tanzania supports one of the most diverse large-mammal communities in the world, involving long-distance migrations of eastern white-bearded wildebeest (*Connochaetes taurinus albojubatus*), plains zebra (*Equus quagga*), common eland (*Tragelaphus oryx*), fringe-eared oryx (*Oryx beisa*), Thomson's gazelle (*Eudorcas thomsonii*), and Grant's gazelle (*Nanger granti*) (Lamprey 1964; Morrison and Bolger 2012, 2014). The migration in the TE is one of only three long-distance migrations of wildebeest remaining in Africa (Estes 2014). Historically, ungulate populations in the TE migrated along at least 10 routes between the dry-season range near the Tarangire River and wet-season calving ranges on shortgrass plains (Lamprey 1964). Due to loss, degradation, and fragmentation of habitat from the expansion of farms, settlements, and mining, only two viable migration routes remain (Morrison and Bolger 2012, 2014; Morrison et al. 2016). These routes were highlighted by Bolger et al. (2008) and Caro et al. (2009) as requiring extremely urgent action to protect known animal movements. Morrison and Bolger (2014) outlined an approximate wildebeest migratory route through this landscape (Fig. 1), but a more explicit delineation of wildebeest migratory corridors is required to inform land-use planning for wildlife conservation.

#Figure 1 approximately here#

In this study, we compared (1) data types, (2) analytical processes to define the resistance surface, and (3) algorithms for delineating migration corridors, and used multiple methods to validate the predictive corridors. We utilized two types of animal locational data: detections from distance-sampling surveys, and GPS collar relocations. We used three methods of data analyses to generate resource selection functions (RSFs) and to define resistance surfaces: negative binomial regression (Nielson and Sawyer 2013), logistic regression (Manly et al. 2002), and maximum entropy (Maxent: Phillips et al. 2006). RSFs allowed us to test a set of habitat preferences and to compare the consistency of these preferences across the three modeling approaches. We tested the predictions that wildebeest selected flat, open grasslands with higher vegetative greenness that were close to water sources, and avoided permanent human settlements along the migration route. We compared two cost-based corridor detection algorithms: cost distance and circuit theory, in delineating predictive corridors. Finally, we validated our predictive corridors with locational data and cross-validated with the other data type based on methods suggested by McClure et al. (2016), by calculating whether corridors contained more or fewer data points than expected relative to random. This multi-method framework provides a rigorous, formalized process for delineating and validating migration corridors using multiple data sources, which can be utilized in land-use plans to effectively conserve landscape connectivity for wildlife.

Methods

Study area

The TE is in the eastern branch of the Great Rift Valley and encompasses roughly 30,000 km^2 (Borner 1985; Prins 1987). The approximate boundaries of the TE are Lake Natron to the north, Simanjiro plains to the southeast, and Irangi Hills to the southwest, with the Rift Valley escarpment forming the western boundary (Lamprey 1964; Morrison and Bolger 2012). Rain occurs almost exclusively from November–May. Rainfall in Tarangire National Park is variable (range = 312 to 1322 mm), with a mean total annual rainfall of 656 mm for years 1980–2004 (Foley and Faust 2010).

Agriculture in the TE increased fivefold from 1984 to 2000 causing substantial habitat loss, increasing fragmentation, and reducing connectivity (Msoffe et al. 2011). The TE contains a patchwork of reserves, including two national parks, Lake Manyara and Tarangire, as well as Manyara Ranch Conservancy (Fig. 1) that support high wildlife densities during the dry season, but all migratory routes and wet-season calving grounds lie outside formally protected areas (Morrison and Bolger 2014), and movement corridors in the ecosystem are considered to be at high risk of disappearing within the near future (Caro et al. 2009; Morrison and Bolger 2014; Morrison et al. 2016).

We delineated wildebeest migration corridors between Manyara Ranch and the Gelai Plains (Fig. 1). Our corridor analysis area covered approximately 1400 km^2 in village lands where both permanent and temporary settlements and domestic livestock are numerous. The Rift Valley escarpment (a steep cliff that prohibits wildebeest movements) formed the western boundary of the study area, and the town of Mto Wa Mbu, a region of high human population density and intensive agriculture, bounded the southwestern part of the study area. A high-traffic asphalt road between Manyara Ranch and Mto Wa Mbu defined the southern extent, and a line between the extinct volcanos of Losimingori, Kitumbeine, and Gelai formed the eastern boundary of the study area.

Wildebeest detection data

We collected wildebeest count data along 140 km of fixed-route, vehicle-based survey transects on the one double-lane gravel road and all single-lane gravel roads and dirt tracks (Fig. 1). We conducted daytime surveys for wildebeest between 0700 and 1800 hrs approximately every two months from October 2014 to September 2015. Two surveys were conducted during each of the three precipitation seasons: November–February (short rains), March–June (long rains), and July–October (long dry season).

Each sampling event consisted of driving all fixed-route transects at a speed of 15–20 kph, stopping only to count animals. The same two observers (DEL and MLB) counted all wildebeest visible out to 500 m along both sides of the track. For each singleton or animal group observed, we recorded vehicle location using GPS, number of individuals, perpendicular sighting distance measured with a laser rangefinder (Bushnell Scout 1000), and cardinal direction. If the sighting was a cluster of animals, we recorded the perpendicular distance from the transect to the center of the animal cluster. We plotted the animal locations on a GIS map using the distance from track and cardinal direction, and used those points as used habitat locations. We compared covariate data at the used habitat locations with 268 systematically placed pseudoabsence locations along the fixed-route transects to estimate binomial (count) and logistic (presence-absence) RSFs.

GPS collar relocation data

We obtained relocation data from two GPS-collared wildebeest (an adult male and adult female). TAM deployed the GPS collars (Telonics TGW-4780H) on 10 and 11 October 2011 in Manyara Ranch as described by Morrison and Bolger (2014). The male's collar collected data at intervals of 4 hours for 75 days, while the female's collar recorded locations at 12-hr intervals (noon and midnight) for 14 days.

Relocation data consist of two or more successive locations of the same animal, but not at frequent enough intervals to treat each sequence as a movement pathway. With relocation data, the focus is on the area between locations rather than the specific pathways between locations or the locations themselves (Zeller et al. 2012). We used Brownian bridge movement models

(BBMM) based on GPS locations during migration to estimate wildebeest movement habitat (Sawyer et al. 2009). The BBMM is a continuous-time stochastic model, where the probability of use is conditioned on the distance and elapsed time between successive locations, the location error, and an estimate of the animal's mobility (Horne et al. 2007; Sawyer et al. 2009). We defined used movement habitat as the 90% contour of the utilization distribution of the BBMM, and defined available habitat as the study area. We generated 718 points systematically every 500 m within the 90% utilization distribution and compared covariate data at these points with 4380 background pseudoabsence points systematically placed every 500 m throughout the entire study area to estimate RSFs from logistic regression. We estimated utilization distributions using the package BBMM (source code from Sawyer et al. 2009) for R (R Development Core Team, 2013). We used Maxent to model probability and correlates of wildebeest presences relative to the background pseudoabsence samples throughout the study area.

Habitat suitability and resistance surface modeling

Resource selection functions (RSFs) formed the basis for habitat suitability models and resistance surfaces. RSFs estimate the probability that animals select or avoid a particular habitat, given the availability of different habitats on the animal's landscape (Manly et al. 2002; Zeller et al. 2012). RSFs compare wildebeest locations (used points) with pseudoabsence data generated along the survey tracks and throughout the study area, depending on the analysis (see Supplementary Material 1). We used three analytical methods (generalized linear models with negative binomial [link = log] or binomial [link = logit] error distributions, and Maxent) to estimate RSFs from our two animal data types (detections and GPS collars). For detection data, we used all three methods: negative binomial regression, logistic regression, and Maxent. For GPS collar data, we used logistic regression and Maxent. Together, we compared among five unique data type–RSF combinations. Details of RSF model structures are provided in Supplementary Material 1.

Environmental and anthropogenic covariates and model selection

We developed a suite of *a priori* hypotheses about resource selection by wildebeest based on the published literature. Wildebeest favor shortgrass plains with green standing crop (McNaughton 1985) and must drink daily (Berry and Louw 1982). We hypothesized wildebeest would select open grasslands with greener vegetation, lower elevations, avoid steeper slopes, and select areas closer to permanent water sources (Hopcraft et al. 2014).

To calculate the environmental variables, we obtained a digital elevation model (DEM) of our study area from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model, Version 2, from the U.S. Geological Survey (USGS) Earth Explorer website (http://earthexplorer.usgs.gov/). This DEM had a resolution of approximately 30 m from which we calculated the elevation and percent slope for each cell in our study area using ArcMap 10.4 (ESRI 2016).

We were unable to find a land cover map for vegetation type that was satisfactory for our modeling purposes, so we created one using the standard terrain-corrected ‘Level 1T’ Landsat 8 OLI imagery taken on February 6th, 2015 and downloaded from the USGS Earth Explorer website at http://landsatlook.usgs.gov/viewer.html. We chose this date because it had the least cloud cover during the time of our study.

Landsat imagery is made up of individual images of spectral bands representing reflectance from different wavelengths of light that can be used to create a land cover map. We performed an unsupervised classification on bands 2–7 in ArcGIS using the Isocluster tool. The Isocluster tool groups clusters of similar pixels from the bands into a given number of classes. In our study we ran 20 iterations having a 20-cell minimum class size and a 10-cell sample window. The cell size of the imagery was 30 x 30 m and resulted in five discreet land cover classes.

Because the Isocluster tool creates its classes in an automated fashion we compared them with existing vegetation maps and our knowledge of the area to interpret our modeling results. We defined the land cover classifications as follows: Afromontane rain forest and Afromontane undifferentiated forest found in the higher elevations of the Kitumbeine Mountains and the Ngorongoro Conservation Area; deciduous wooded grasslands that occur on the mid-elevation slopes below Afromontane forests but above the valley bottoms; two classes of edaphic grasslands mixed with low densities of trees and shrubs; and lakes/seasonal wetlands that included Lake Magadi.

We used Normalized Difference Vegetation Index (NDVI) values as a metric of vegetation greenness. We downloaded NDVI maps from the USGS Famine Early Warning Systems Network data portal (http://earlywarning.usgs.gov/fews). This portal provides access to the Earth Resources Observation and Science (EROS) Center’s satellite-derived vegetation products generated from the Moderate Resolution Imaging Spectroradiometer where NDVI maps are composited in 10-day intervals. We selected six NDVI maps representing the approximate time periods during which we collected our field observations for wildebeest detections. Using bilinear interpolation in ArcMap, we downscaled the maps from their original ~260 m to 30 m to match the resolution of our other grids.

In order to analyze wildebeest resource selection according to the relative greenness of vegetation at the time of the survey, the spatial NDVI data for wildebeest locations were matched to the time period in which the location data were collected. We scaled NDVI values by subtracting the NDVI for each wildebeest location from the mean NDVI for all pseudoabsence locations for that month. Thus, the anomaly indicated relative greenness of vegetation available during each survey period.

We included distance to permanent water sources and two anthropogenic covariates representing distance to human structures in our models. Permanent settlements were clusters of concrete structures, whereas Masai homesteads (bomas) comprised one or more temporary structures made of natural materials such as wood, mud, and grass, and were encircled by fencing of cut thorny branches. We mapped all settlements and bomas using Google Earth (Mountain View, CA, USA) aerial imagery from June 2014. We created grids representing the Euclidian distance to bomas, permanent settlements, and water sources using the same 30-m spacing as our elevation and slope grids. We defined permanent settlements as polygons and calculated the distances from their edges, whereas bomas were smaller features approximately 100 m across that were defined as a point in the center of the boma. We hypothesized that wildebeest would avoid permanent settlements (Stabach et al. 2016), but be unaffected by bomas because pastoralist people from the Masai tribe have built such structures in this region over the past several hundred years and do not normally hunt wild ungulates (Reid 2012). We also included quadratic terms for each of the distance covariates to test for non-linear relationships.

Our global model of all candidate explanatory variables included: vegetation type (*veg*); NDVI anomaly (*NDVI anom*); elevation (*elev*); percent slope (*slope pct*); distance to water in km (*water dist*); distance to water squared (*water dist*^{2}); distance to settlement in km (*settlement dist*); distance to settlement squared (*settlement dist*^{2}); distance to boma in km (*boma dist*); and distance to boma squared (*boma dist*^{2}). Covariates were converted to a grid of 30-m^2 cells with each cell containing a value for the covariate at that location. All covariates were continuous with the exception of the categorical vegetation type (*veg*): we set deciduous wooded grasslands (the most common class in the study area), as the reference vegetation category. Small agricultural plots occurred only within permanent settlements, so that vegetation in the study area consisted of natural, uncultivated lands.

We did not analyze data by season because wildebeest counts were too sparse during the short rains and dry seasons, when most of the animals were on the northern calving grounds or southern dry-season range. However, some wildebeest were detected during these seasons, and we included these data in the modeled migration corridors, as our goal was to map important sites used by wildebeest in the analysis area.

We fitted and ranked regression models from our candidate set using information-theoretic model selection and Akaike's Information Criterion corrected for small sample sizes (AIC_c; Burnham and Anderson 2002). For details on model selection and inference, see Supplementary Material 2. For logistic and negative binomial regressions using wildebeest detections and GPS collar data, we reported the significant parameters and model weights, and presented final, weighted, averaged RSFs. We also reported odds ratios as calculated by ([exp(ß)-1] x 100%) for each of the parameters in the RSFs (see Supplementary Material 3).

For Maxent, model selection is integral with the algorithm's entropy maximizing machine learning process so only one final model is produced from the environmental covariates and animal presence/pseudoabsence data (Elith et al. 2011). Performance of the model is quantified using area under the receiver operating characteristic curve (AUC) where AUC = 0.5 represents a model that predicts habitat use no better than random, and where values closer to 1 represent better model prediction (Elith et al. 2011; Poor et al. 2012).

Landscape resistance

We used our RSFs to estimate the habitat suitability of each pixel on the GIS map based on landscape variables. We then created landscape resistance maps by calculating the inverse of the habitat suitability and scaling so that each cell in the grid was assigned a value from 1 to 1000, indicating the ‘cost’ (e.g., energy expenditure, mortality risk, or habitat avoidance) for the animals to move across it (Pullinger and Johnson 2010; McClure et al. 2016).

Migration corridor algorithms

We used two cost-based algorithms, least-cost path analysis (cost-distance) and circuit theory to define predicted corridors. Cost-distance minimizes the tradeoff between travel distance among habitat patches and exposure to unsuitable habitat, and provides the shortest cumulative cost-weighed distance between two endpoints (McClure et al. 2016). Circuit theory (McRae et al. 2008) treats cells in a landscape as a network of nodes connected to neighboring cells by resistors. Connectivity increases with multiple pathways in circuit networks, making this methodology useful for identifying multiple movement corridors (McRae et al. 2008). Cost-distance assumes individuals have complete knowledge of the entire landscape that they are traversing, whereas circuit theory is based on random-walk theory and assumes individuals only perceive the landscape within a 1-cell radius of their current location (McRae et al. 2008). The fact that resistance to current flow is the same in both directions suggests circuit theory may be less suitable for identifying a repeatedly used, directional migration corridor (McRae et al. 2008). However, migratory routes used by ungulates can vary by season and year (Bolger et al. 2008) and individuals may use a particular route multiple times within a single season or multiple routes might be used by different parts of the population (Sawyer et al. 2009). Thus, we presumed that both cost-distance and circuit theory would be useful and appropriate for delineating major and ancillary migration corridors for wildebeest in our study area.

Both cost-distance and circuit theory use resistance maps to predict the relative value of cells in the landscape for movement between two endpoints (McClure et al. 2016). We selected Manyara Ranch as the start point in the south and the shortgrass plains calving grounds on the Gelai Plains in the north as the end point (Fig. 1). We used the Linkage Mapper GIS tool (McRae and Kavanagh 2011) in ArcMap to run the cost-distance algorithm and CircuitScape (available on CircuitScape.org; McRae and Shah 2008) to run the circuit-theory algorithm. The Linkage Mapper tool identifies adjacent (neighboring) core areas and creates maps of least-cost corridors between them, and then mosaics the individual corridors to create a single composite corridor map. The resulting map reflects the relative value of each grid cell in providing connectivity between core areas, allowing users to identify which routes encounter more or fewer features that facilitate or impede movement between core areas. CircuitScape creates a resistance-based connectivity metric where each cell of the resistance map is converted into a node and, using circuit theory, the 'cost' of travelling to adjacent nodes is calculated. The path of adjoining cells that have the lowest resistance will thus have the highest conductance and denote the possible paths the animals would travel (McRae et al. 2008).

The results are predictive corridor maps from every data type–RSF–corridor algorithm combination where every cell is assigned a predictive corridor value and then ranked in order from highest to lowest.

Corridor model validation

We used slightly modified methods as outlined by McClure et al. (2016) to validate our corridor results. First, we delineated the 80th, 85th, 90th, and 95th percentile corridors for each predictive corridor map. A percentile corridor designates a corridor as the portion of the landscape that is predicted to experience the highest rates of movement. We scaled the CircuitScape cells from 0–

100 with the higher-value cells receiving a higher ranking. We scaled the Linkage Mapper cells from 0–100 and then calculated the inverse of this value, such that cells ranked at 20% were converted to 80%. We defined the 80th, 85th, 90th, and 95th percentile corridors as the most traversable 20%, 15%, 10%, and 5% of the landscape, respectively. We also delineated the 80th, 85th, 90th, and 95th percentile corridors in two null (distance-only) predictive corridor maps. Null predictive corridor maps assume that animals are most likely to simply travel in a straight line without regard to environmental features. We generated the null models by running cost-distance and circuit-theory algorithms on uniform resistance maps in which all cells have equal resistance (= 1).

To quantify how much of the target movement process (i.e. migration corridor) was included within each map, we calculated the percentage of the data points that was used to create that map which fell within the 80th, 85th, 90th, and 95th percentile corridors. We also generated 100 random points within the study area and calculated the number of random points within the 80th, 85th, 90th, and 95th percentile corridors. To determine whether our corridors included a significant number of the data points used to create them, we compared the percentage of data points observed within each percentile corridor with the number of random points using a chi-squared analysis. This validation method determines whether the corridors contained more or fewer data points than expected relative to a random spatial distribution of points, and assumes most wildebeest detections and relocations should occur within cells that have high predicted connectivity values. We also conducted a cross-validation procedure that used one data type (detection or GPS collar) to validate the corridors created using the other data type, therefore utilizing independent datasets for each validation. To delineate the portion of the landscape with the highest probability of use, we created a combined corridor map that merged the 95% predictive corridors from the best models that performed significantly better than random.

Results

Wildebeest detections

We detected 1202 wildebeest at 51 locations from six distance-sampling surveys. Counts included 11 locations of 65 wildebeest during the short rains, 35 locations of 1010 animals during the long rains, and 5 locations of 127 wildebeest during the dry season. The relatively low numbers of wildebeest counted in the short rains (Jan–Feb) and dry season (Sept–Oct) compared with the long rains was due to the fact that most animals in January and February were on the northern calving grounds outside our analysis area, and in September and October were at the southern dry-season range. During the long-rainy season, we recorded 441 wildebeest in May and 569 in June. May and June counts each reflected approximately 7–8% of the estimated total wildebeest population of approximately 7000 in the TE (Morrison et al. 2016).

GPS collar data

We obtained 19 relocations from GPS collar data from 2 wildebeest between Manyara Ranch and the Gelai Plains. The male wildebeest (9 GPS fixes) travelled 69 km north, and the female (10 GPS fixes) travelled north 52 km then turned back south and travelled 34 km.

Habitat suitability models

We considered the number of parameters in our final habitat suitability models to be adequate (i.e. models were not over-fitted) because parameter estimates were reasonable and standard errors were estimable.

Negative binomial regression model

For the negative binomial regression using detection data, four models were competitive ($\Delta AIC_c < 2$) and six models accounted for 95% of AIC_c weight (W) (Table 1). We found a significant positive correlation between number of wildebeest and NDVI anomaly (vegetation greenness), distance to permanent settlements, and elevation, and a significant negative correlation between number of wildebeest and steepness of slope and vegetation type of lakes or seasonal wetlands. These covariates had significant coefficient slope estimates ($P < 0.05$) and appeared in all models that carried any W (Table 2). The final weighted, averaged negative binomial regression RSF model describing wildebeest migration habitat selection from the detection data in the TE and odds ratios for parameters are presented in Supplementary Material 3.

Logistic regression models

For the logistic regression model using detections, four models were competitive and seven models carried > 95% of W (Table 1). Presence of wildebeest was significantly negatively correlated with steepness of slope and positively correlated with NDVI anomaly and distance to permanent settlements (Table 2). The final weighted, averaged logistic regression RSF model describing wildebeest migration habitat selection from detection data and odds ratios for parameters are provided in Supplementary Material 3.

For the logistic regression model using GPS collar locations in a BBMM utilization distribution, two models accounted for all W (Table 1). Model-averaged parameter estimates indicated wildebeests migrated within a moderate distance from permanent settlements and water, avoided bomas and steeper slopes, and preferred lower elevations (Table 2). This model suggested wildebeest avoided Afromontane forest and lakes/seasonal wetlands and used edaphic grasslands more than the reference category (deciduous wooded grasslands). In contrast to RSFs for detection data, RSFs for GPS collar data suggested no selection for higher NDVI values (Table 2). The final, weighted, averaged logistic RSF model for the GPS collar data and odds ratios for parameters are presented in Supplementary Material 3.

Maxent models

The AUC for the Maxent RSF model using detection data was 0.814, and for the BBMM using GPS collar data was 0.828, indicating both models performed well in predicting the distribution of wildebeest. The Maxent RSF model using detection data was similar to the logistic and binomial regressions described above in that probability of presence of wildebeest was positively correlated with distance to permanent settlements and flatter slopes. The most important parameters were distance to permanent settlements and bomas (Table 3). Distance to water as well as percent slope were also important parameters, but NDVI anomaly did not contribute to this model. For the Maxent RSF model using GPS collar data, the distance to permanent settlements was the most important parameter, contributing 66.9% to the model, while the other parameters all contributed < 7%.

Corridor delineation and validation

Predictive corridor maps from every data type–RSF–corridor algorithm combination are presented in Fig. 2 A-J. Null model predictive corridor maps are presented in Fig. 2 K-L. For both the data validation and the cross-validation processes, the same three predictive corridor maps (Maxent_detections_LinkageMapper [Fig 2E], logistic_regression_GPS_LinkageMapper [Fig 2G], and Maxent_GPS_LinkageMapper [Fig 2I]) contained significantly more data points than random in all four percentile corridors (Fig. 3), and these three were the only 95% predictive corridors that contained significantly more data points than random points. Figure 4 depicts the corridor with the highest probability of use from merging the three 95% predictive corridors that performed significantly better than random.

#Figures 2 and 3 approximately here#

Discussion

We present the first quantitative migration corridor modeling and validation results for long-distance migratory wildebeest in a system that is facing substantial habitat alteration and fragmentation. Our multi-method approach allowed us to utilize two data sources and address model uncertainty by using multiple analyses to examine variation in habitat suitability and resistance surface mapping to identify anthropogenic and environmental covariates affecting the probability of wildebeest use in the landscape. Combining models and datasets can be useful in settings where data are sparse and animals exhibit multiple behavioral states, such as migratory versus residency (Hopcraft et al. 2014). GPS collar data are useful for identifying routes used at low frequency, but are often limited by small sample sizes. Count data in contrast may not allow observation of areas used during rapid movements, but they may provide a larger sample size of individuals and spatial distribution of the population. In combination, the two data types permitted us to validate and cross-validate our predictive corridors, confirming their robustness.

Correlates of wildebeest habitat selection

Our five data type–RSF combinations revealed many consistent patterns, and some inconsistent ones. As predicted, all RSFs strongly identified wildebeest migration habitat to include flatter terrain farther from permanent human settlements, supporting recent findings from resident wildebeests in Kenya (Stabach et al. 2016). In contrast to our predictions, RSFs indicated migratory wildebeest preferred areas distant from bomas, suggesting avoidance even of very low-density human habitation. Also contrary to our predictions, we found some evidence that wildebeest avoided areas near fresh water sources, but we believe this was a consequence of permanent settlements and human uses that are themselves associated with year-round water sources, most of which have been appropriated for agriculture and livestock in this region.

The RSF models based on detection data indicated wildebeest selected areas with higher vegetation greenness values and used deciduous wooded grasslands more than other vegetation types, while the RSF models based on GPS collar data found no association with greenness, and indicated wildebeest used edaphic grasslands more than deciduous wooded grasslands. Furthermore, detection data suggested wildebeests preferred higher elevations while GPS collar data indicated selection for lower elevations. We believe these conflicting results may underscore behavioral differences underlying collar data collected during the northward migration, versus detection data collected throughout the year. Our GPS collar data indicated that animals moved quickly northwards from Manyara Ranch towards the breeding grounds once the rains began, often traveling at night through bottlenecks and across high-traffic roads (Morrison and Bolger 2014). Therefore, the collar data likely reflected wildebeest movements during an active migratory state, when animals were moving through open areas such as grasslands without necessarily stopping frequently to forage. Daytime detection sampling surveys captured animals throughout the year in various behavioral states including individuals migrating quickly northwards at the onset of the wet season and slowly southwards as the dry season progressed, as well as some individuals that may have resided within the corridor, because we detected wildebeest in the study area during all three precipitation seasons. Slower-moving southward migrants and potentially resident animals from our detection data were likely selecting foraging habitat represented by higher greenness (NDVI values), and seeking shade in wooded areas during the heat of the day.

There is growing recognition that ungulates have distinct migratory and sedentary phases, and that this difference is important to understanding the environmental context in which movements occur (Morales et al. 2004). Some corridors are used mainly for connectivity, while others have functional uses, such as habitat for short stop-overs and even longer foraging bouts (Sawyer and Kauffman 2011). Given the ecological significance of these different behavioral states, efforts to delineate corridors should ideally incorporate methods that can differentiate these behaviors and functions, though this can be difficult with ground-based observations because the timing of migration can vary substantially with the timing of rainfall, and is thus difficult to observe (Holdo et al. 2009; Singh and Milner-Gulland 2011). In our study, visual detections did not reveal which behavioral state (migratory versus sedentary) individuals expressed at the time of observation. However, wildebeest were detected in the corridor study area during all seasons, and our cross-validation procedure indicated similar corridor delineation results across both data types, suggesting that the area is likely used for both connectivity and foraging.

Delineation of migratory corridors

Ungulate migrations are typically believed to occur along one well-defined route, but Sawyer et al. (2009) suggested that multiple routes might be used by different parts of the population.

Furthermore, migratory routes used by ungulates can vary by season and year (Bolger et al. 2008). The predictive corridors from each data type–RSF combination were similar regardless of corridor detection algorithm, as evidenced by the similarities between pairs of predictive corridor maps in Fig. 2 (A-B), (E-F), (G-H), and (I-J) which included roughly the same areas. However, there were noticeable differences in predictive corridor maps due to variation in data types and RSF model structures. The three best predictive corridors for wildebeest migration in our study area were derived using cost-distance (i.e. Linkage Mapper) rather than circuit theory, similar to predictions for migratory elk (*Cervus elaphus*) pathways in Yellowstone National Park (McClure et al. 2016). Like McClure et al. (2016), we believe both corridor detection algorithms are useful. Linkage Mapper depicted a single, wider corridor for wildebeest, while CircuitScape offered a diversity of narrower corridors and revealed potentially important locations in need of further investigation, such as pinch points at a steep cliff that crossed the center of our study area. By merging the three best-performing 95% predictive corridor maps from our validation process, we delineated a combined corridor that was most strongly supported by our wildebeest sampling data (Fig. 4). We believe that combining high-performing corridor models is the best approach to ensure that land-use plans conserve sufficient movement pathways for migratory wildebeest.

#Figure 4 approximately here#

Conservation implications

Until the 1800s, grassland ecosystems around the world supported vast herds of migratory ungulates numbering in the millions of animals (Fryxell et al. 1988) and these moving herds structured entire ecosystems (Holdo et al. 2009). Human population growth and agricultural expansion have led to considerable encroachment on and eventual loss of many historical migratory routes (Bolger et al. 2008; Newmark 2008; Caro et al. 2009), including the elimination of most migrations throughout the range of *Connochaetes* (Estes 2014). East Africa still supports a high diversity and abundance of migratory ungulates, but most of the remaining populations are threatened (Bolger et al. 2008).

Bolger et al. (2008) noted that in obligate migratory populations, such as wildebeest in the TE, populations often respond suddenly and severely to the disruption of migratory routes. At the end of the 19th century the TE's wildebeest population likely numbered in the hundreds of thousands and likely connected with populations in the Amboseli basin and Athi-Kapiti Plains in southern Kenya (Estes 2014). More recently, the TE's wildebeest population decreased from an estimated 40,000 animals in 1988 to approximately 7000 today (Morrison et al. 2016). The TE's eastern white-bearded wildebeest is genetically distinct from the western white-bearded wildebeest (*C. t. mearnsi*) in the Serengeti-Ngorongoro Ecosystem (Georgiadis 1995), thus the extirpation of wildebeest in the TE would mean the loss of an evolutionarily significant population, with only small populations remaining in Kenya (Worden et al. 2010). The sustainability of the TE's wildebeest population is important to the ecological function (Lee et al. 2016) and economic value of Tarangire and Lake Manyara national parks, two of the most popular and profitable parks in the country. Conserving migratory habitat for wildebeest also will protect important rangelands for Masai pastoralists and their livestock, as both domestic cattle and wildebeest utilize similar habitats for grazing in the TE (Voeten and Prins 1999), and will benefit other wildlife species correlated with wildebeest presence in this ecosystem (Kiffner et al. 2015). Given the many demands on grazing lands in these landscapes, there is an important need to accurately identify core habitat used by migratory wildlife. Our analysis provides an illustration of ways to incorporate multiple data sources, models, and validation techniques to estimate a spatially explicit corridor for effective land-use planning and conservation.

Acknowledgements

This research was conducted with permission from Tanzania Commission for Science and Technology (COSTECH), Tanzania Wildlife Research Institute, and villages of Selela, Engaruka, and Mbaash, under COSTECH permits 2014-53-ER-90-172 and 2015-22-ER-90-172, and Wildlife Division permit NO.HA/403/563/0l/104. DEL and MLB received funding from Fulbright U.S. Scholar Program, ERM Group Foundation, Columbus Zoo, Rufford Foundation, and TRIAS. TAM received funds for GPS collaring from Wildlife Conservation Society Tarangire Elephant Project, Tembo Foundation, and Earthwatch Institute.

Table 1 Model selection results for resource selection functions of migratory habitat for eastern white-bearded wildebeest (*Connochaetes taurinus albojubatus*) in the Tarangire Ecosystem, Tanzania. The top-weighted 95% model set is shown. Bolded are models with < 2 ΔAIC_c.

Model	K	ΔAIC_c	W
Negative binomial detections			
slope pct + settlement dist + NDVI anom + elev + veg	6	**0.00**	**0.36**
settlement dist + NDVI anom + slope pct + veg + elev + settlement dist^2	7	**1.47**	**0.17**
slope pct + settlement dist + NDVI anom + elev + veg + water dist	7	**1.73**	**0.15**
slope pct + settlement dist + NDVI anom + boma dist + elev + veg	6	**1.78**	**0.15**
slope pct + settlement dist + NDVI anom + veg	5	3.16	0.07
boma dist + settlement dist + water dist + NDVI anom + slope pct + veg + elev + water dist^2	9	3.33	0.07
boma dist + settlement dist + NDVI anom + slope pct + boma dist^2 + settlement dist^2	6	5.71	0.02
Logistic regression detections			
boma dist + settlement dist + NDVI anom + slope pct	5	**0.00**	**0.28**
boma dist + settlement dist + NDVI anom + slope pct + water dist	6	**0.91**	**0.18**
boma dist + settlement dist + water dist + NDVI anom + slope pct + settlement dist^2	7	**1.37**	**0.14**
boma dist + settlement dist + NDVI anom + slope pct + boma dist^2 + settlement dist^2	7	**1.63**	**0.12**
boma dist + settlement dist + water dist + NDVI anom + slope pct + settlement dist^2	7	2.29	0.09
boma dist + settlement dist + water dist + NDVI anom + slope pct + boma dist^2 + settlement dist^2	8	2.59	0.08

boma dist + settlement dist + water dist + NDVI anom + slope pct + water dist^2	7	2.83	0.07
Logistic regression GPS collars			
slope pct + settlement dist + settlement dist^2 + NDVI anom + boma dist + boma dist^2 + water dist + water dist^2 + veg + elev	11	**0.00**	**0.98**
slope pct + settlement dist + settlement dist^2 + NDVI anom + boma dist + water dist + water dist^2 + veg	9	8.00	0.02

Pct = percent, dist = distance, anom = anomaly, elev = elevation, veg = vegetation type. Vegetation types were Afromontane forest; deciduous wooded grasslands, two classes of edaphic grassland, and lakes/seasonal wetlands.

Table 2 Model-averaged coefficient parameter estimates and *P*-values for estimating migratory habitat for eastern white-bearded wildebeest (*Connochaetes taurinus albojubatus*) in the Tarangire Ecosystem, Tanzania. Significant parameters are bolded.

Negative binomial regression with detection data

Variable	Coefficient	SE	*P*
(Intercept)	-4.673	1.811	0.042
Boma dist	0.081	0.112	0.210
Water dist	0.045	0.046	0.161
Water dist^2	-0.006	0.003	0.064
Settlement dist	**0.360**	**0.096**	**0.021**
Settlement dist^2	0.001	0.001	0.192
Slope pct	**-0.737**	**0.084**	**0.004**
NDVI anom	**0.209**	**0.049**	**0.017**
Elev	**0.455**	**0.189**	**0.047**
Afromontane forest	-12.718	2E+07	0.318
Edaphic grassland 1	-1.051	0.763	0.110
Lake/seasonal wetland	**-2.294**	**0.341**	**0.007**
Edaphic grassland 2	-2.026	1.001	0.063

Logistic regression with detection data

Variable	Coefficient	SE	*P*
(Intercept)	-3.535	0.838	0.017
Boma dist	1.384	0.637	0.056
Boma dist^2	-0.174	0.200	0.181
Water dist	-0.015	0.032	0.261
Water dist^2	0.000	0.002	0.296

Settlement dist	**0.257**	0.084	**0.031**
Settlement dist^2	-0.003	0.003	0.135
Slope pct	**-0.319**	0.121	**0.040**
NDVI anom	**0.108**	0.025	**0.016**
Elev	0.001	0.01	0.316

Logistic regression with GPS data

Variable	Coefficient	SE	*P*
(Intercept)	-12.522	0.744	0.001
Boma dist	**0.569**	**0.158**	**0.023**
Boma dist^2	0.024	0.047	0.250
Water dist	**0.357**	**0.067**	**0.011**
Water dist^2	**-0.033**	**0.005**	**0.008**
Settlement dist	**1.761**	**0.093**	**0.001**
Settlement dist^2	**-0.061**	**0.003**	**0.001**
Slope pct	**-0.162**	**0.020**	**0.005**
NDVI anom	0.011	0.007	0.096
Elev	**-0.142**	**0.041**	**0.025**
Afromontane forest	-1.792	1.030	0.079
Edaphic grassland 1	**0.618**	**0.183**	**0.026**
Lake/seasonal wetland	-0.018	0.124	0.311
Edaphic grassland 2	**0.798**	**0.134**	**0.009**

Pct = percent, dist = distance, anom = anomaly, elev = elevation, veg = vegetation type. Vegetation types were Afromontane forest; deciduous wooded grasslands, two classes of edaphic grassland, and lakes/seasonal wetlands.

Table 3 Percent contributions and permutation importance for parameters in Maxent models for migratory eastern white-bearded wildebeest (*Connochaetes taurinus albojubatus*) in the Tarangire Ecosystem, Tanzania.

Maxent with detection data		
Variable	Percent contribution	Permutation importance
Boma dist	21.5	26.8
Water dist	20.4	15.4
Settlement dist	28	26.9
Slope pct	18.8	18.1
NDVI anom	0	0
Elev	9.2	11.8
Veg	2.1	1
Maxent with GPS data		
Variable	Percent contribution	Permutation importance
Boma dist	1	1.4
Water dist	6.1	5
Settlement dist	66.9	73
Slope pct	4	3.5
NDVI anom	3.4	0.6
Elev	5.5	8
Veg	13	8.5

Pct = percent, dist = distance, anom = anomaly, elev = elevation, veg = vegetation type. Vegetation types were Afromontane forest; deciduous wooded grasslands, two classes of edaphic grassland, and lakes/seasonal wetlands.

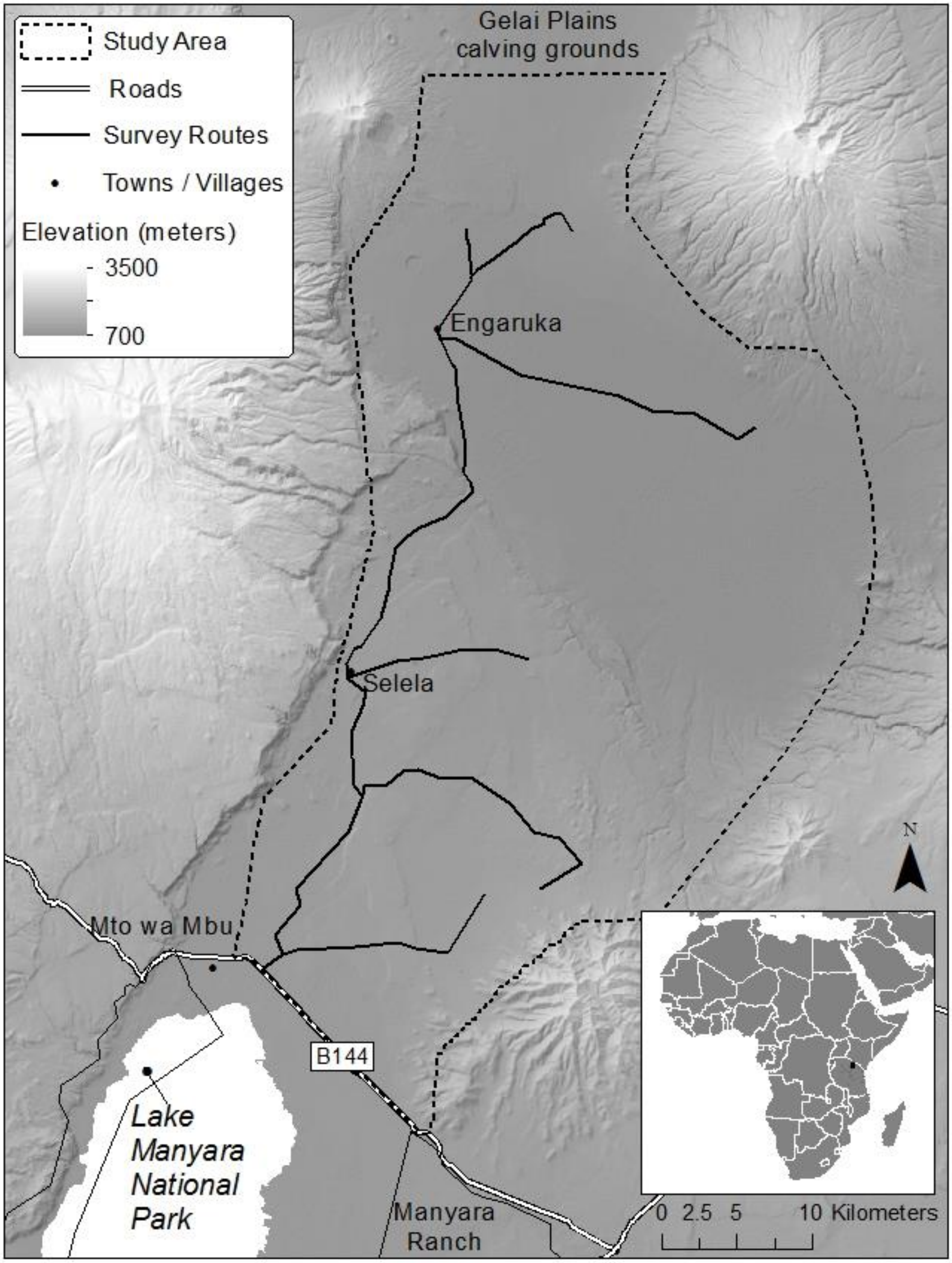
Figure Captions

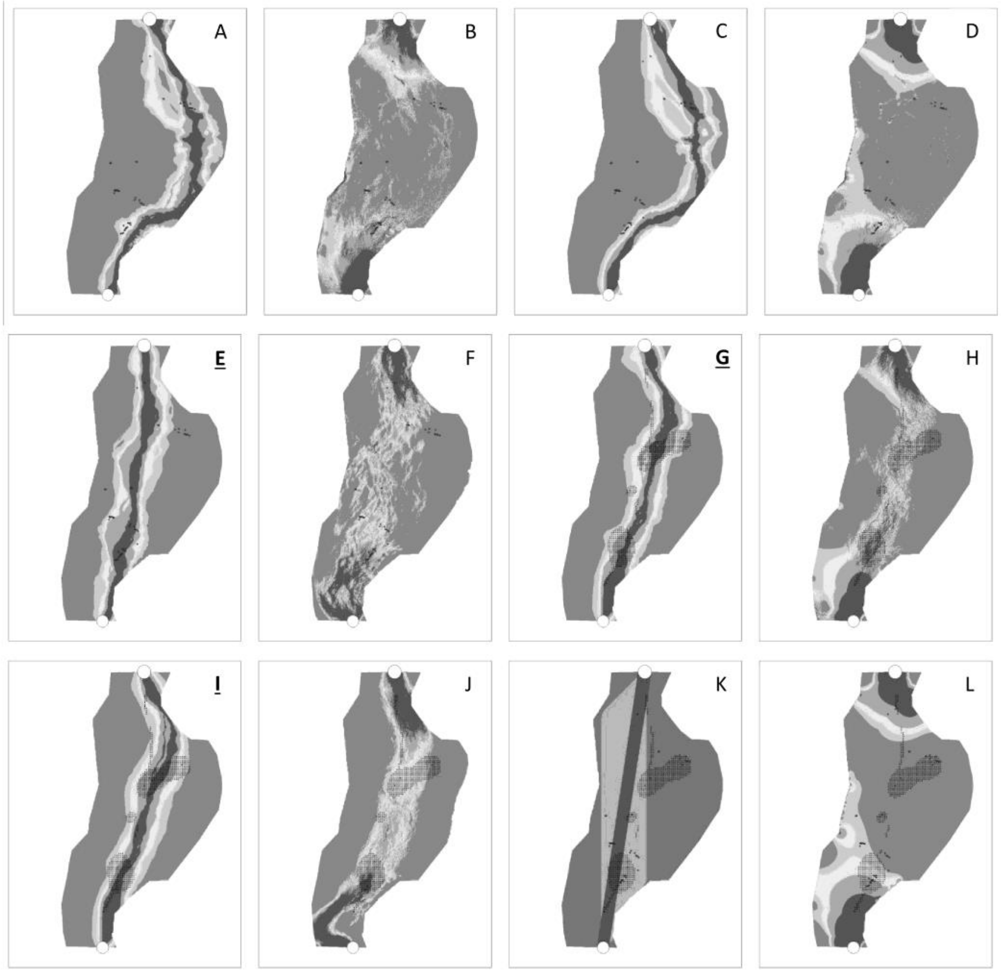
Fig. 1 Corridor study area (enclosed by dashed black line) for migratory eastern white-bearded wildebeest (*Connochaetes taurinus albojubatus*) in the Tarangire Ecosystem, Tanzania between Manyara Ranch and the calving grounds on the Gelai Plains. Lines are distance-sampling survey tracks, labelled towns are areas with high-density settlements and agricultural cultivation, and black square in inset map shows study area location in Africa.

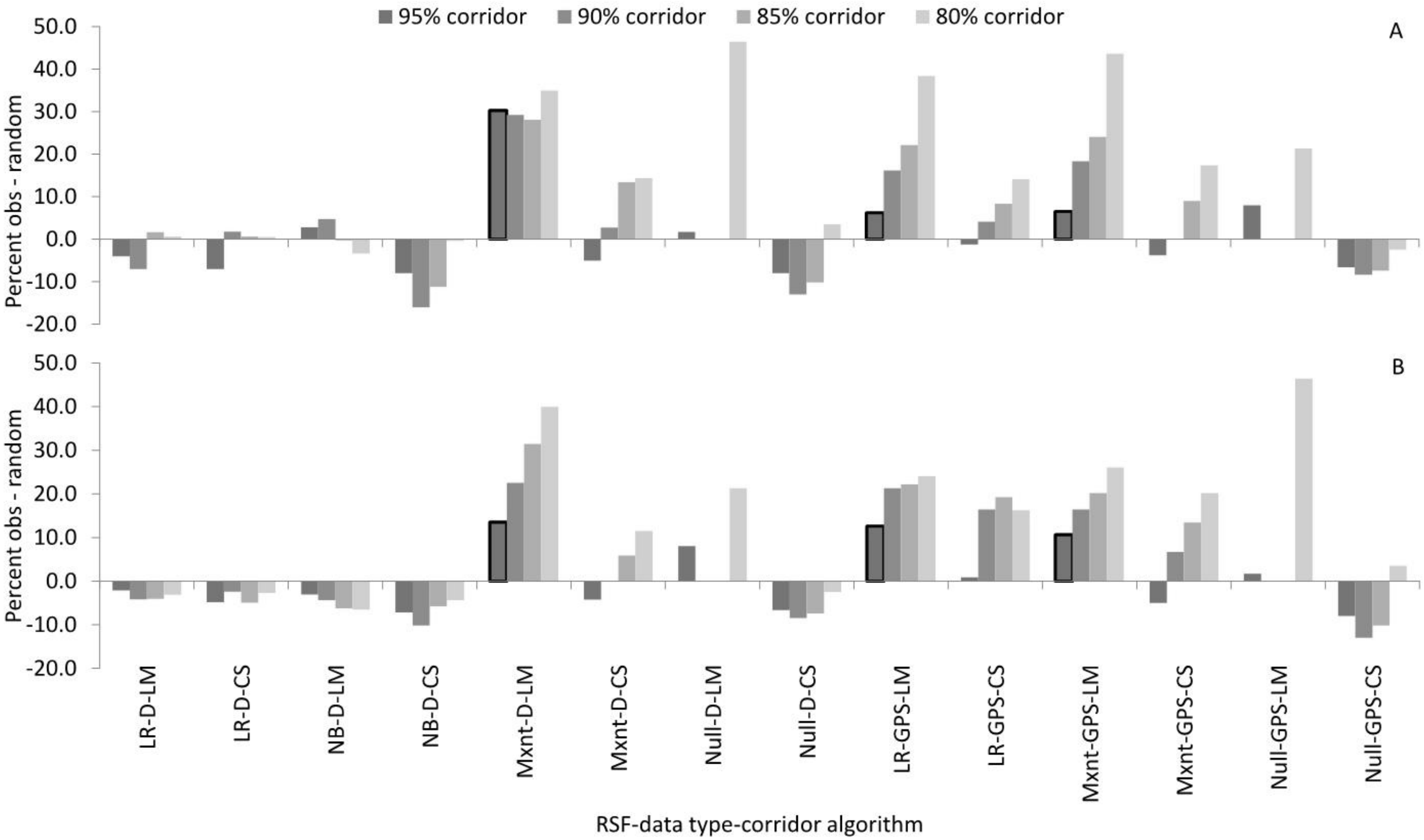
Fig. 2 Maps depicting the 80th, 85th, 90th, and 95th percentile predictive corridors for wildebeest migration habitat in the Tarangire Ecosystem, Tanzania derived from combinations of data types, Resource Selection Functions, and corridor detection algorithms: detection data, logistic regression, and LinkageMapper (**a**); detection data, logistic regression, and CircuitScape (**b**); detection data, negative binomial regression, LinkageMapper (**c**); detection data, negative binomial regression, and CircuitScape (**d**); detection data, Maxent, and LinkageMapper (**e**); detection data, Maxent, and CircuitScape (**f**); GPS collar data, logistic regression, and LinkageMapper (**g**); GPS collar data, logistic regression, GPS collar data, and CircuitScape (**h**); GPS collar data, Maxent, and LinkageMapper (**i**); GPS collar data, Maxent, and CircuitScape (**j**); Null LinkageMapper (**k**); and Null CircuitScape (**l**). Best models (**e**, **g**, **i**) from validation and cross-validation are underlined.

Fig. 3 Validation (**a**) and Cross-validation (**b**) of wildebeest migration predictive corridors in the Tarangire Ecosystem, Tanzania, done by comparing percent of data points observed in each corridor minus the percent of random points in each corridor. Validation used the same data as were used to create the corridors, cross-validation used the other data type e.g., GPS data were used to validate corridors that were defined using detections data. Outlined columns indicate 95% corridors that contained significantly more data points than random points (Chi-squared P value < 0.05).

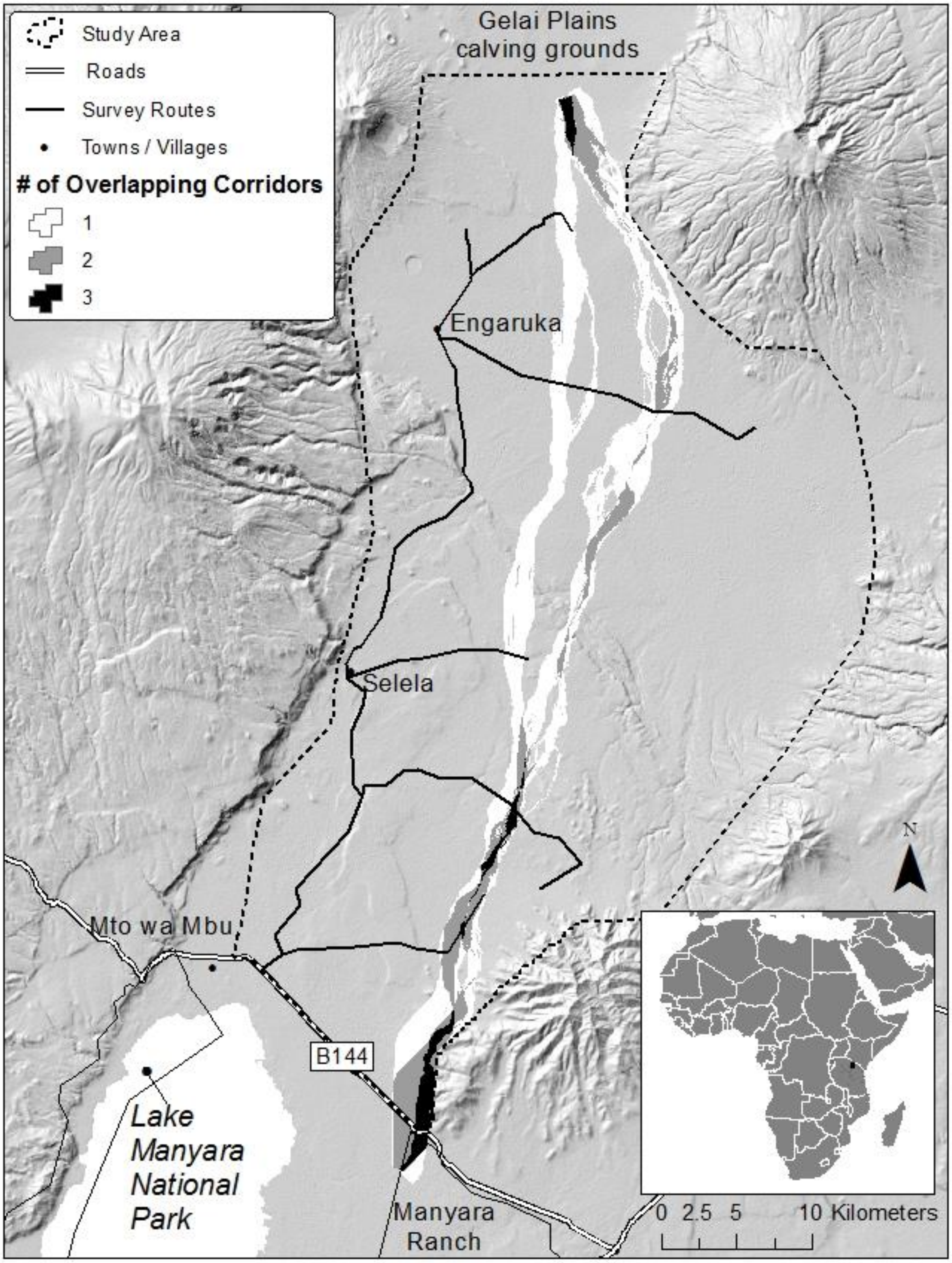
Fig. 4 Map combining top three predictive corridor models for migratory eastern white-bearded wildebeest (*Connochaetes taurinus albojubatus*) in the Tarangire Ecosystem, Tanzania between Manyara Ranch and the calving grounds on the Gelai Plains.


Study Area
Roads
Survey Routes
Towns / Villages
Elevation (meters)
3500
700
Gelai Plains
calving grounds
Engaruka
Selela
Mto wa Mbu
B144
Lake
Manyara
National
Park
Manyara
Ranch
N
0 2.5 5 10 Kilometers


A
B
C
D
E
F
G
H
I
J
K
L



95% corridor
90% corridor
85% corridor
80% corridor
A
B
Percent obs - random
50.0
40.0
30.0
20.0
10.0
0.0
-10.0
-20.0
LR-D-LM
LR-D-CS
NB-D-LM
NB-D-CS
Mxnt-D-LM
Mxnt-D-CS
Null-D-LM
Null-D-CS
LR-GPS-LM
LR-GPS-CS
Mxnt-GPS-LM
Mxnt-GPS-CS
Null-GPS-LM
Null-GPS-CS
RSF-data type-corridor algorithm


Study Area
Roads
Survey Routes
Towns / Villages
# of Overlapping Corridors
1
2
3
Gelai Plains
calving grounds
Engaruka
Selela
Mto wa Mbu
B144
Lake
Manyara
National
Park
Manyara
Ranch
N
0 2.5 5 10 Kilometers

585

References

Adriaensen F, Chardon JP, De Blust G, Swinnen E, Villalba S, Gulnick H, Matthysen E (2003) The application of “least cost” modelling as a functional landscape model. Landsc Urban Plan 64:233–247

Beier P, Majka DR, Spencer WD (2008) Forks in the road: choices in procedures for designing wildland linkages. Cons Biol 22:836–851

Berger J (2004) The last mile: how to sustain long-distance migration in mammals. Cons Biol 18:320–331

Berry HH, Louw GN (1982) Seasonal nutritive status of wildebeest in Etosha National Park. Madoqua 13:127–39

Bolger DT, Newmark WD, Morrison TA, Doak DF (2008) The need for integrative approaches to understand and conserve migratory ungulates. Ecol Letters 11:63–77

Borner M (1985) The increasing isolation of Tarangire National Park. Oryx 19:91–96

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretical approach. Springer-Verlag, New York

Caro T, Jones T, Davenport T (2009) Realities of documenting wildlife corridors in tropical countries. Biol Conserv 142:2807–2811

Cushman SA, McRae B, Adriaensen F, Beier P, Shirley M, Zeller K (2013) Biological corridors and connectivity. In: Key Topics in Conservation Biology 2, First Edition. Edited by David W. Macdonald and Katherine J. Willis, pp 384–404

Elith J, Phillips SJ, Hastie T, Dudík M, Yung En Chee, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. Diversity Distrib 17:43–57

Epps CW, Mutayoba BM, Gwin L, Brashares JS (2011) An empirical evaluation of the African elephant as a focal species for connectivity planning in East Africa. Diversity Distrib 17:603–6012

ESRI Environmental Systems Research Institute (2016) ArcGIS Desktop: Release 10.4. Redlands, CA, USA

Estes RD (2014) The Gnu’s World: Serengeti Wildebeest Ecology and Life History. Univ of California Press

Foley CA, Faust LJ (2010) Rapid population growth in an elephant *Loxodonta africana* population recovering from poaching in Tarangire National Park, Tanzania. Oryx 44:205–212

Fryxell JM, Greever J, Sinclair ARE (1988) Why are migratory ungulates so abundant? Amer Nat 131:781–798

Georgiadis N (1995) Population structure of wildebeest: implications for conservation. In Serengeti II: Dynamics, management and conservation of an ecosystem (eds A.R.E. Sinclair & P. Arcese). University of Chicago Press, Chicago and London

Hamilton EW, Giovannini MS, Moses SA, Coleman JS, McNaughton SJ (1998) Biomass and mineral element responses of a Serengeti short-grass species to nitrogen supply and defoliations: compensation requires a critical [N]. Oecologia 116:407–418

Harris G, Thirgood S, Hopcraft JGC, Cromsigt JPGM, Berger J (2009) Global decline in aggregated migrations of large terrestrial mammals. Endang Sp Res, 7:55–76

Hebblewhite M, Haydon DT (2010) Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. Phil Transac Royal Soc B: Biological Sciences 365:2303–2312

Holdo RM, Holt RD, Fryxell JM (2009) Opposing rainfall and plant nutritional gradients best explain the wildebeest migration in the Serengeti. Amer Nat 173:431–445

Hopcraft J, Morales J, Beyer H, Borner M, Mwangomo E, Sinclair A, Olff H, Haydon D (2014) Competition, predation and migration: individual choice patterns of Serengeti migrants captured by hierarchical models. Ecol Monographs 84:355–372

Horne JS, Garton EO, Krone SM, Lewis JS (2007) Analyzing animal movements using Brownian bridges. Ecology 88:2354–2363

Kiffner C, Albertini M, Ede A, Donnellan B, Hahn N, McGinnis MA, Nietlisbach NA, Tate J, Kioko J (2015) Performance of individual species as indicators of large mammal species richness in Northern Tanzania. Ecol Indic 53:70–77

Lamprey HF (1964) Estimation of the large mammal densities, biomass and energy exchange in the Tarangire Game Reserve and the Masai Steppe in Tanganyika. E Afr J Wildl 2:1–46

LaPoint S, Gallery P, Wikelski M, Kays R (2013) Animal behavior, cost-based corridor models, and real corridors. Landscape Ecol 28:1615–1630

Lee DL, Kissui BM, Kiwango YA, Bond ML (2016) Migratory herds of wildebeests and zebras indirectly affect calf survival of giraffes. Ecology and Evolution 6:8402–8411

Manly BF, McDonald LL, Thomas DL, McDonald DL, Erickson WP (2002) Resource selection by animals: statistical analysis and design for field studies. Kluwer, Dordrecht

McClure ML, Hanson AJ, Inman RM (2016) Connecting models to movements: testing connectivity model predictions against empirical migration and dispersal data. Landscape Ecol. DOI 10.1007/s10980-016-0347-0

McNaughton SJ (1985) Ecology of a grazing ecosystem: The Serengeti. Ecol Monogr 55:259–294

McRae B, Shah V (2008) Circuitscape User Guide. University of California, Santa Barbara

McRae BH, Kavanagh DM (2011) Linkage Mapper Connectivity Analysis Software. The Nature Conservancy, Seattle, WA. www.waconnected.org/habitat-connectivity-mapping-tools.php

McRae B, Dickson B, Keitt T, Shah V (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89:2712–2724

Morales JM, Haydon DT, Frair J, Holsiner KE, Fryxell JM (2004). Extracting more out of relocation data: Building movement models as mixtures of random walks. Ecology 85:2436–2445

Morrison TA, Bolger DT (2012) Wet season range fidelity in a tropical migratory ungulate. J Anim Ecol 81:543–52

Morrison TA, Bolger DT (2014) Connectivity and bottlenecks in a migratory wildebeest *Connochaetes taurinus* population. Oryx 48:613–621

Morrison TA, Link WA, Newmark WD, Foley CAH, Bolger DT (2016) Tarangire revisited: Consequences of declining connectivity in a tropical ungulate population. Biol Conserv 197:53–60

Msoffe FU, Said MY, Ogutu JO, Kifugo SC, de Leeuw J, van Gardingen P, Reid RS (2011) Spatial correlates of land-use changes in the Maasai-Steppe of Tanzania: Implications for conservation and environmental planning. Int J Biodiv Conserv. 3:280–290

Newmark WD (2008) Isolation of African protected areas. Front Ecol Environ 6: 321–328

Nielson RM, Sawyer H (2013) Estimating resource selection with count data. Ecol Evol 3:2233–2240

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions Ecol Modelling 190:231–259

Poor EE, Loucks C, Jakes A, Urban DL (2012) Comparing habitat suitability and connectivity modeling methods for conserving pronghorn migrations. PLoS ONE 7(11): e49390

Prins HHT (1987) Nature conservation as an integral part of optimal land use in East Africa: action plan for the Masai ecosystem of northern Tanzania. Biol Conserv 40:141–161

Pullinger MG, Johnson CJ (2010) Maintaining or restoring connectivity of modified landscapes: evaluating the least-cost path model with multiple sources of ecological information. Landscape Ecol 25:1547–1560

R Core Development Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Reid R (2012) Savannas of our Birth: People, Wildlife, and Change in East Africa. University of California Press

Sawyer H, Kauffman MJ (2011). Stopover ecology of a migratory ungulate. J Anim Ecol 80:1078–1087

Sawyer H, Kauffman MJ, Nielson RM, Horne JS (2009) Identifying and prioritizing ungulate migration routes for landscape-level conservation. Ecol App 19:2016–2025

Singh NJ, Milner-Gulland EJ (2011) Conserving a moving target: planning protection for a migratory species as its distribution changes. J Appl Ecol 48:35–46

Singh NJ, Börger L, Dettki H, Bunnefeld N, Ericsson G (2012) From migration to nomadism: movement variability in a northern ungulate across its latitudinal range. Ecol Appl 22:2007–2020

Stabach JA, Wittemyer G, Boone RB, Reid RS, Worden JS (2016) Variation in habitat selection by white-bearded wildebeest across different degrees of human disturbance. Ecosphere 7:1–17

Voeten MM, Prins HHT (1999) Resource partitioning between sympatric wild and domestic herbivores in the Tarangire region of Tanzania. Oecologia 120:287–294

Worden J, Mose V, Western D (2010) Aerial Census of Wildlife and Livestock in Eastern Kajiado. African Conservation Centre, Nairobi

Zeller KA, McGarigal K, Whiteley AR (2012) Estimating landscape resistance to movement: a review. Landscape Ecol, 27:777–797

A multi-method approach to delineate and validate migratory corridors

Monica L. Bond · Curtis M. Bradley · Christian Kiffner · Thomas A. Morrison · Derek E. Lee

M. L. Bond · D. E. Lee

Wild Nature Institute, P. O. Box 165, Hanover, New Hampshire 03755, USA

C. M. Bradley

Center for Biological Diversity, P. O. Box 710, Tucson, Arizona 85701, USA

C. Kiffner

The School for Field Studies, Center for Wildlife Management Studies, P.O. Box 304, Karatu, Tanzania

T. A. Morrison

Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Graham Kerr Building, Glasgow G12 8QQ, UK

Supplementary Material

1. RSF model structure

We utilized relevant spatial information for each location (counts, presences, pseudoabsences) to fit generalized linear or Maxent models that explained the spatial distribution of wildebeest in the corridor analysis area. For negative binomial and logistic RSF models from detections, we systematically placed pseudoabsence points every 500 m along the same survey transects that we used for documenting presences of wildebeest (Phillips et al. 2009). For the logistic RSF model from GPS collars and all Maxent RSF models, we systematically placed pseudoabsence points every 500 m throughout the entire study area. We compared environmental covariate data from detection locations with 268 systematically placed pseudoabsence locations along the fixed-route transects to estimate binomial (count) and logistic (presence-absence) RSFs. We compared the detection locations with 4380 background pseudoabsence points from the entire study area to estimate Maxent RSF.

We used negative binomial regression for RSF modeling of our count detections data, which is recommended for over-dispersed count data (Kabacoff 2011), and has an error structure that accepts a large number of zero count locations from pseudoabsences. Negative binomial regression can be considered as a generalization of Poisson regression since it has the same mean structure as Poisson regression and it has an extra parameter to model the over-dispersion. The negative binomial approach quantifies habitat use along a continuum that ranges from zero to the maximum number of observed animal locations in a sampling unit, and produces estimates of resource selection that are unbiased in the face of serial correlation in the location data (Nielson and Sawyer 2013). Negative binomial regression analyses (link = log) were performed in Program R (R Development Core Team 2013) using the glm.nb function (link=log) from the MASS package (Venables and Ripley 2002).

We used logistic regression to estimate RSFs using wildebeest detection data and GPS collar data as presences, with locations of wildebeest coded 1, and pseudoabsences coded 0. Logistic regression analyses were performed in Program R using the glm function (link = logit).

We used Maxent to estimate RSFs using wildebeest detection data and GPS collar data. Maxent uses presence data to model probability and correlates of the presence of a species relative to background pseudoabsence samples throughout the study area (Phillips et al. 2009). Using covariate data from presences and background samples, Maxent estimates the ratio of the probability density of covariates across the landscape of interest, typically a pixel of grids [denoted as $f(z)$], to the probability density of covariates at locations where the species was present [denoted as $f_1(z)$]. The Maxent model minimizes the distance, or relative entropy, between $f_1(z)$ with respect to $f(z)$ and is a log-linear model similar in form to a generalized linear model (Elith et al. 2011). One of the main critiques of Maxent, sample selection bias (Elith et al. 2011), was not germane to our sample because we had systematically collected presence data associated with temporal and spatial information.

2. Model Selection and Inference

We drew inference from models based on AICc weights (W) and strength of parameter estimates. We considered models with ΔAICc < 2 to be competitive, and used support from all weighted models to include model selection uncertainty and obtain robust predictions (Grueber et al. 2011). We developed our final predictive model by selecting the subset of best models that together carried > 95% of the total W (Burnham and Anderson 2002). We then used the zero method for model averaging to estimate the final model coefficients from the 95% model set (Burnham and Anderson 2002). We conducted model selection and model averaging using package MuMIn for R (Bartoń 2015; R Development Core Team, 2013).

3. Resource Selection Functions and odds ratios

Our suite of candidate explanatory variables to explain wildebeest migratory habitat selection included: vegetation type (*veg*); vegetation greenness indexed by Normalized Difference Vegetation Index anomaly (*NDVI anom*); elevation (*elev*); percent slope (*slope pct*); distance to permanent water sources in km (*water dist*); distance to water squared (*water dist*^{2}); distance to permanent settlement (*settlement dist*); distance to settlement squared (*settlement dist*^{2}); distance to Masai boma in km (*boma dist*); and distance to boma squared (*boma dist*^{2}).

Negative binomial regression

$$Number\ of\ wildebeests = \text{e}^{\wedge}(-4.673 + 0.081(\text{boma dist}) + 0.045(\text{water dist}) - 0.0006(\text{water dist}^2) + 0.360(\text{settlement dist}) - 0.001(\text{settlement dist}^2) - 0.737(\text{slope}) + 0.209(\text{NDVI anom}) + 0.455(\text{elev}) - 1.051(\text{edaphic grassland 1}) - 2.294(\text{lakes/seasonal wetlands}) - 2.026(\text{edaphic grassland 2})$$

Odds ratios suggest that the probability of greater numbers of wildebeest using a pixel should increase by 8% for every 1-km increase in distance to bomas; by 43% for every 1-km increase in distance to permanent settlements, and by 5% for every 1-km increase in distance to water, but then decrease by 1% beyond a threshold distance from water. For each 1-unit increase in the percent slope, the probability of use by more wildebeest decreased by 52%, and each unit increase in the NDVI anomaly resulted in a 23% increase in probability of use. Wildebeest were 65, 90, and 87% less likely to use edaphic grasslands and lakes and seasonal wetlands compared with the reference category (deciduous wooded grasslands).

Logistic regression detections

$Wildebeest\ Presence =$ e^(–3.535 +1.384(boma dist) – 0.174(boma dist^2) – 0.015(water dist) - 0.0004(water dist^2) + 0.257(settlement dist) – 0.003(settlement dist^2) – 0.319(slope) + 0.108(NDVI anom) + 0.001(elev)

Odds ratios indicate the probability of presence of migrating wildebeest in a pixel was expected to increase by 299% for every 1-km increase in distance to bomas and by 29% for every 1-km increase in distance to permanent settlement, and decrease by 1% for every 1-km increase in distance to water. Distance to bomas showed a quadratic relationship such that beyond a threshold distance, the probability of wildebeest presence began to decline. For each 1-unit increase in the percent slope, the probability of wildebeest presence decreased by 27%, and each unit increase in the NDVI anomaly resulted in an 11% increase in probability of wildebeest presence. All other parameters had negligible odds ratios (i.e. changed by < 1%).

Logistic regression GPS collar

$Wildebeest\ Presence =$ e^(–12.521 + 0.569(boma dist) + 0.024(boma dist^2) + 0.357(water dist) – 0.033(water dist^2) + 1.761(settlement dist) – 0.061(settlement dist^2) – 0.162(slope pct) + 0.011(NDVI anom) – 0.141(elev) – 1.792(Afromontane forest) + 0.618(edaphic grassland 1) – 0.018(lakes/seasonal wetlands) + 0.798(edaphic grassland 2)

Odds ratios suggested each 1-km increase in distance to bomas increased probability of presence of wildebeest by 77%; each 1-km increase in distance to water increased probability of presence by 27%, and each 1-km increase in distance to permanent settlements increased probability of presence by 482% but only to a threshold distance. Each unit increase in distance to water increased probability of presence by 43% to a threshold distance. For each 1-unit increase in the percent slope, the probability of wildebeest presence decreased by 15%. Each 1-unit increase in NDVI anomaly increased probability of presence by only 1%. Wildebeest were 83% less likely to use Afromontane forests, and 86 and 122% more likely to use edaphic grasslands more than the reference category (deciduous wooded grasslands).

References

Bartoń K (2015) MuMIN: Multi-model inference package in R. July 3 2015

Elith J, Phillips SJ, Hastie T, Dudík M, Yung En Chee, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. Diversity Distrib 17:43–57

Grueber CE, Nakagawa S, Laws RJ, Jamieson IG (2011) Multimodel inference in ecology and evolution: challenges and solutions. J Evol Biol, 24:699–711

Kabacoff RI (2011) R in Action: Data analysis and graphics with R. Manning Publications Co. Shelter Island, NY, USA

Nielson RM, Sawyer H (2013) Estimating resource selection with count data. Ecol Evol 3:2233–2240

Phillips SJ, Dudík M, Elith J, Graham CH, Lehman A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol Appl 19:181–197

R Core Development Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Venables WN, Ripley BD (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0