

Step by Step: Using Step Selection Functions to Simulate Dispersal and Assess Landscape Connectivity

David D. Hofmann^{1,2,§} Gabriele Cozzi^{1,2} John W. McNutt² Arpat Ozgul¹
Dominik M. Behr^{1,2}

September 3, 2021

¹ Department of Evolutionary Biology and Environmental Studies, University of Zurich,
Winterthurerstrasse 190, 8057 Zurich, Switzerland.

² Botswana Predator Conservation, Private Bag 13, Maun, Botswana.

§ Corresponding author (david.hofmann2@uzh.ch)

Running Title: Release the Dogs! Simulating Wild Dog Dispersal to Assess Landscape
Connectivity

Keywords: dispersal, simulation, movement, integrated step selection function,
Kavango-Zambezi Transfrontier Conservation Area, landscape connectivity, *Lycaon pictus*

Abstract

Dispersal of individuals is a crucial prerequisite for long-term species persistence but depends on a sufficient degree of landscape connectivity. To date, connectivity has primarily been quantified using least-cost analysis and circuit theory, albeit both methods make assumptions that are rarely met in reality. Least-cost analysis assumes that animals move towards a preconceived destination and choose a cost-minimizing route accordingly, whereas circuit theory presumes a complete random walk without directional persistence. While these shortcomings can be overcome by spatio-temporally explicitly simulating dispersal trajectories from individual-based movement models, a unified approach for such simulations is lacking.

Here, we present a simple three-step workflow to simulate dispersal and assess landscape connectivity starting from empirical GPS movement data. We exemplify the workflow using dispersal data collected on the endangered African wild dog (*Lycaon pictus*) in the Kavango-Zambezi Transfrontier Conservation Area, the world's largest transboundary conservation area. In step one, we use integrated step selection functions to parametrize a mechanistic movement model describing habitat and movement preferences of dispersers. In step two, we employ the parametrized model to simulate individual dispersal trajectories across the landscape. In step three, we interpret simulated trajectories by converting them into three complementary connectivity maps: a heatmap revealing frequently traversed areas, a betweenness-map delineating dispersal corridors, and a map of inter-patch connectivity depicting the presence and intensity of functional links between habitat patches.

We find that simulating dispersal from individual-based movement models fitted via step selection functions are readily interpretable and offer great flexibility to realistically render dispersal movements. Moreover, by not enforcing connections towards known endpoints but instead letting dispersers react to the environment as they go, such simulations overcome shortcomings inherent least-cost path analysis, while at the same time allowing for autocorrelation in simulated steps, which can not be incorporated using circuit theory. Inspite of these advantages, simulating dispersal entails numerous non-trivial modeling decisions, such as the number of simulated individuals and the duration of simulated dispersal events.

We show that simulations from step-selection functions offer a simple yet powerful alternative to traditional connectivity modeling techniques, although necessitating informed decisions about the number of simulated individuals and the duration of simulated dispersal events. Ultimately, our workflow permits a more mechanistic understanding of dispersal behaviour and landscape connectivity making it useful for a variety of applications in ecological, evolutionary, and conservation research.

Contents

1	Introduction	1
1.1	Importance of Connectivity & Connectivity Models	1
1.2	Issues with Traditional Connectivity Models	1
1.3	What about IBMMs?	2
1.4	Proposed Solution: Three-Step Workflow	2
1.5	Some Background on SSFs	3
1.6	Some Background on Connectivity “Metrics”	3
1.7	Case Study	4
2	Methods	5
2.1	Study Area	5
2.2	GPS Data	5
2.3	Covariates	6
2.4	Movement Model	6
2.5	Dispersal Simulation	8
2.6	Source Points	9
2.7	Convergence	9
2.8	Heatmap	10
2.9	Betweenness	10
2.10	Inter-Patch Connectivity	10
3	Results	11
3.1	Movement Model	11
3.2	Dispersal Simulation	12
3.3	Convergence	12
3.4	Heatmap	13
3.5	Betweenness	13
3.6	Inter-Patch Connectivity	14
4	Discussion	14
4.1	Short Summary	14
4.2	Movement Model	15
4.3	Simulation	15
4.4	Maps	16
4.5	Related Literature	17
4.6	Advantages of ISSF Simulations	17
4.7	Disadvantages of ISSF Simulations	18
4.8	Conclusion	20
5	Authors' Contributions	21
6	Data Availability	21
7	Acknowledgements	21

1 Introduction

1.1 Importance of Connectivity & Connectivity Models

Dispersal of individuals is a vital process that allows species to maintain genetic diversity (Perrin and Mazalov, 1999, 2000; Frankham et al., 2002; Leigh et al., 2012; Baguette et al., 2013), rescue non-viable populations (Brown and Kodric-Brown, 1977), and colonize or recolonize unoccupied habitats (Hanski, 1999; MacArthur and Wilson, 2001). However, the ability to disperse is contingent on a sufficient degree of landscape connectivity (Fahrig, 2003; Clobert et al., 2012), which is why the identification and protection of dispersal corridors that promote connectivity has become a task of fundamental importance (Nathan, 2008; Doerr et al., 2011; Rudnick et al., 2012). The identification of dispersal corridors not only necessitates a comprehensive understanding of the factors that limit dispersal, but also an appropriate model to estimate connectivity (Baguette et al., 2013; Vasudev et al., 2015; Hofmann et al., 2021). To date, the two most prominent connectivity models in the literature are least-cost path analysis (LCPA; Adriaensen et al., 2003) and circuit theory (CT; McRae, 2006; McRae et al., 2008), both graph-based approaches that quantify conductance of the landscape based on habitat permeability (Zeller et al., 2012).

1.2 Issues with Traditional Connectivity Models

Despite their intuitive nature and ease of use, both methods make rigorous assumptions that are hardly ever met in reality (Diniz et al., 2019). With LCPA, for instance, a least-costly route always exists, even if associated movement costs are unreasonably high and will never be incurred by dispersing individuals (Sawyer et al., 2011). The method also presumes that animals move towards a preconceived endpoint and choose a cost-minimizing route accordingly (Abrahms et al., 2017). While these assumptions may be fulfilled by migrating animals, they are unlikely to hold for dispersers, as dispersers typically move across unfamiliar territory and are therefore less aware of associated movement costs (Koen et al., 2014; Cozzi et al., 2020). CT relaxes these assumptions, but posits that animals move according to a random walk, entailing that autocorrelation between subsequent movements cannot be rendered (Diniz et al., 2019). In reality, however, autocorrelation in animal movements is regularly observed (Bovet and Benhamou, 1991; Schultz and Crone, 2001), especially during dispersal (Cozzi et al., 2020; Hofmann et al., 2021). Some of these deficiencies can be alleviated using less stringent graph-based methods, such as least-cost *corridors* (Pinto and Keitt, 2009), *thresholded* least-cost paths (Landguth et al., 2012), or *randomized* least-cost

paths (Panzacchi et al., 2016), yet a certain degree of arbitrariness remains. Besides this, graph-based methods are incapable of rendering the temporal dimension of dispersal, meaning that statements about the approximate duration required to move between habitats are impossible (Martensen et al., 2017; Diniz et al., 2019).

1.3 What about IBMMs?

To overcome the unrealistic assumptions of LCPA and CT, dispersal simulations from individual-based movement models (IBMMs) have been proposed and applied (Diniz et al., 2019). In these models, dispersal trajectories are simulated based on movement rules that govern how individuals move across and interact with the prevailing landscape (e.g. Gustafson and Gardner, 1996; Gardner and Gustafson, 2004; Graf et al., 2007; Kramer-Schadt et al., 2004; Revilla et al., 2004; Revilla and Wiegand, 2008; Kanagaraj et al., 2013; Clark et al., 2015; Allen et al., 2016; Hauenstein et al., 2019; Zeller et al., 2020; Vasudev et al., 2021). Once the trajectories are simulated, they are further processed and converted into measures of connectivity (Kanagaraj et al., 2013; Zeller et al., 2020). Because IBMMs do not enforce connections towards preconceived endpoints, simulated individuals are allowed to dynamically adjust their behavior as they disperse. This enables to render potential interactions between movement behavior and habitat conditions, shifting the focus from a structural to a more functional view on connectivity (Tischendorf and Fahrig, 2000; Kanagaraj et al., 2013; Hauenstein et al., 2019). Furthermore, IBMMs render movement sequentially, meaning that the temporal dimension of dispersal becomes explicit and that autocorrelation between movements can be incorporated (Diniz et al., 2019). In spite of these advantages, a unified framework to simulate dispersal and assess connectivity using IBMMs is lacking.

1.4 Proposed Solution: Three-Step Workflow

In this paper, we draw on recent innovations in the field of movement ecology and network theory to propose a simple three-step workflow for the simulation of dispersal and assessment of connectivity (Figure 1). In the first step, GPS data of the focal species is combined with relevant habitat covariates and analyzed in an integrated step selection framework to fit a movement model. In the second step, the parametrized model is employed as an IBMM to simulate dispersal trajectories across the study area. Finally, in the third step, simulated trajectories are converted into a set of three complementary connectivity maps, each focusing on a different aspect of connectivity. The set of maps includes a heatmap, revealing

frequently traversed areas, a betweenness-map, delineating dispersal corridors and bottlenecks, and a map of inter-patch connectivity, depicting the presence or absence, intensity, and dispersal duration of functional links between habitat patches.

1.5 Some Background on SSFs

To fit a movement model, we propose to employ the framework of step selection functions (SSFs; Fortin et al., 2005; Thurfjell et al., 2014; Avgar et al., 2016). Originally, SSFs were developed to learn about habitat preferences of the focal species using GPS data (Fortin et al., 2005). The method works by comparing covariates along realized steps (the straight-line connecting two consecutive GPS relocations; Turchin, 1998) to the same covariates along randomly generated alternative steps (Fortin et al., 2005; Thurfjell et al., 2014) using conditional logistic regression (Fortin et al., 2005; Muff et al., 2020). Recently, Avgar et al. (2016) have introduced *integrated* SSFs (ISSFs), which is a generalization of regular SSFs and requires that movement metrics are included in the corresponding regression model. This not only reduces biases in estimated preferences, but also permits simultaneous inference on a species' habitat kernel (i.e. habitat preferences), its movement kernel (i.e. movement preferences/capabilities), and potential interactions between the two kernels (Avgar et al., 2016; Fieberg et al., 2020). Importantly, a model that was fitted using ISSFs resembles a fully mechanistic IBMM, based on which movement can be simulated (Avgar et al., 2016; Signer et al., 2017). Previously, ISSFs have been applied to simulate utilization distributions of resident animals (Avgar et al., 2016; Signer et al., 2017), yet the method remains underutilized for the simulation of dispersal and detection of movement corridors.

1.6 Some Background on Connectivity “Metrics”

Irrespective of the chosen modeling framework, simulated dispersal trajectories need to be processed into measures of connectivity. To date, the most common practice to translate simulations into connectivity measures is to compute inter-patch connectivity between distinct habitats (Diniz et al., 2019). Inter-patch connectivity is a simple ratio between the number of dispersers successfully moving between habitat patches and the total number of simulated individuals from each patch (e.g. Kanagaraj et al., 2013). As an alternative, some researchers have tallied overlapping trajectories into what might be called “heatmaps”, i.e. maps that depict the frequency at which each spatial unit in the study area was traversed by simulated dispersers (e.g. Pe'er and Kramer-Schadt, 2008; Hauenstein et al., 2019; Zeller et al., 2020). More recently, Bastille-Rousseau et al. (2018) proposed a method to convert

trajectories into a network, so that metrics from network theory can be computed. With regards to connectivity, the betweenness metric seems promising, as it indicates the number of shortest paths going through each node in the network and therefore highlights areas of exceptional importance for connecting different regions in the study area. Because each of the three metrics focuses on a very different aspect of landscape connectivity, we propose to view them as complementary and to use them in concert when assessing connectivity.

1.7 Case Study

To exemplify the application of the proposed workflow in (Figure 1), we investigate landscape connectivity for the endangered African wild dog (*Lycon pictus*) as a case study. While once present across entire sub-Saharan Africa, this species has disappeared from a vast majority of its historic range, mainly due to persecution by humans, habitat fragmentation and destruction, and deadly diseases (Woodroffe and Sillero-Zubiri, 2012). Some of the biggest remaining populations reside near the Moremi Game Reserve in northern Botswana (McNutt, 1996; Cozzi, 2013; Behr et al., 2020), which is part of the Kavango-Zambezi Transfrontier Conservation Area (KAZA-TFCA), the world’s largest transboundary protected area. We previously assessed landscape connectivity for dispersing African wild dogs within this ecosystem using least-cost methods (Hofmann et al., 2021). More specifically, we used ISSFs to fit a habitat selection model and predicted a permeability surface based on which we computed least-cost paths and corridors between protected areas. We now expand on this knowledge and use ISSFs to develop a more detailed, mechanistic movement model of dispersing wild dogs (Figure 1). We then employ the model to simulate 80'000 dispersers moving 2'000 steps across the landscape of the KAZA-TFCA. Finally, we condense the simulated trajectories into a heatmap, betweenness-map, and inter-patch connectivity map. We thereby demonstrate that simulations from ISSFs provide a viable alternative to LCPA or CT and deliver a more comprehensive understanding of the emergence of connectivity. Nevertheless, we recognize that simulations from ISSFs entail several non-trivial modeling-decisions, such as the number of simulated dispersers and the duration of simulated dispersal events.

2 Methods

2.1 Study Area

The study area of our wild dog case study was defined by a bounding box centered at -17°13'9"S, 23°56'4"E (Figure 2a) and stretched over 1.3 Mio. km², encompassing the entire KAZA-TFCA (Figure 2b). The KAZA-TFCA comprises parts of Angola, Botswana, Namibia, Zimbabwe, and Zambia and hosts diverse landscapes, ranging from savannah to grassland and from dry to moist woodland habitats. In its center lies the Okavango Delta, a dominant hydrogeographical feature and the world's largest flood-pulsing inland delta. Although large portions of the KAZA-TFCA are designated national parks and other protected areas, considerable human influence remains from roads, agricultural sites and settlements.

2.2 GPS Data

GPS data of 16 dispersing wild dogs was obtained as described in Cozzi et al. (2020) and Hofmann et al. (2021). We collected data on 7 female and 9 male dispersers between 2011 and 2019 from a free-ranging population inhabiting the Moremi National Park in northern Botswana. Because behavior during dispersal is more pertinent to landscape connectivity than behavior during residence (Elliot et al., 2014; Abrahms et al., 2017), we discarded all data that was collected during residency. For some individuals, exact dispersal dates were known from field observations, whereas in other cases, we determined dispersal phases using net-squared displacement. Net-squared displacement is a metric that measures the squared Euclidean distance of a GPS relocation to a reference point (Börger and Fryxell, 2012), which in our case was set to the center of each individual's natal home range. Thus, dispersal was deemed to have started when an individual left its natal home range and ended once individuals became sedentary again. During dispersal, GPS collars recorded a fix every 4 hours and they regularly transmitted data over the Iridium satellite system. After collection, we converted GPS coordinates of dispersers ($n = 4'169$) to steps, where each step represented the straight-line connecting two consecutive GPS relocations (Turchin, 1998). To ensure a regular sampling interval, we removed fixes that were not successfully collected on the 4-hourly schedule (± 15 minutes) and identified bursts of regularly sampled GPS fixes.

2.3 Covariates

We represented the physical landscape in our study area using a set of habitat covariates, including water-cover, distance-to-water, woodland-cover, and shrub/grassland-cover. Because water-cover greatly changes within and between years in the Okavango Delta, we applied a remote sensing algorithm and obtained weekly updated water-cover layers and corresponding distance-to-water layers from MODIS satellite imagery (see Wolski et al., 2017 and Appendix A3 in Hofmann et al., 2021). Resulting layers thus temporally aligned with each dispersal event, ensuring minimal temporal lag between observed movements and corresponding water layers. We furthermore included a proxy for human influence, rendering anthropogenic pressures stemming from human-density, agricultural sites, and roads. All spatial layers were coarsened or interpolated to a target resolution of 250 m by 250 m. A detailed description of the preparation of each habitat covariate is provided in Hofmann et al. (2021).

Besides habitat covariates, we computed movement metrics that we used as movement covariates in our ISSF models (Avgar et al., 2016; Fieberg et al., 2020). Movement metrics were calculated for each step and included the step length (`sl`), its natural logarithm (`log(sl)`), and the cosine of the relative turning angle (`cos(ta)`). Because wild dogs follow a distinct activity pattern (Castelló, 2018), we also coded a binary variable (`LowActivity`) indicating whether a step was realized during periods of low wild dog activity (09:00 to 17:00 local time) or high wild dog activity (17:00 to 09:00 local time). Handling and manipulation of all data, as well as all models and simulations were implemented with the statistical software R, version 3.6.6 (R Core Team, 2020). Some helper functions were written in C++ and imported into R using the `Rcpp` package (Eddelbuettel and François, 2011; Eddelbuettel, 2013).

2.4 Movement Model

We used ISSFs to parametrize a mechanistic movement model for dispersing wild dogs (Avgar et al., 2016). To conduct ISSF analysis, we paired each realized step with 24 random steps, so that a realized step plus its 24 random steps formed a stratum that received a unique identifier. As suggested by Avgar et al. (2016), we generated random steps by sampling random turning angles from a von Mises distribution with concentration parameter $\kappa = 0$ (i.e. a uniform distribution on $(-\pi, +\pi)$) and step lengths from a gamma distribution that was fitted to realized steps (scale $\theta = 6'308$ and shape $k = 0.37$). Along each step, we extracted and averaged values from the habitat covariate layers using the `velox` package (Hunziker, 2021). We also calculated the movement metrics `sl`, `log(sl)`, and `cos(ta)` for each

realized and random step. To facilitate model convergence, we standardized all continuous covariates to a mean of zero and a standard deviation of one. Since correlation among covariates was low ($|r| < 0.6$; Latham et al., 2011), we retained all of them for modeling.

To contrast realized steps (scored 1) and random steps (scored 0), we assumed that animals assigned a selection score $w(x)$ of the exponential form to each step (Fortin et al., 2005). The selection score $w(x)$ of each step thus depended on the step's associated covariates (x_1, x_2, \dots, x_n) and on the animal's preferences (i.e. relative selection strengths; Avgar et al., 2017) towards these covariates $(\beta_1, \beta_2, \dots, \beta_n)$:

$$w(x) = \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n) \quad (\text{Equation 1})$$

The probability of a step being realized was then contingent on the step's selection score, as well as on the selection scores of all other step in the same stratum:

$$P(Y_i = 1 | Y_1 + Y_2 + \dots + Y_i = 1) = \frac{w(x_i)}{w(x_1) + w(x_2) + \dots + w(x_i)} \quad (\text{Equation 2})$$

To estimate preferences (i.e. the β 's), we used mixed effects conditional logistic regression analysis (Muff et al., 2020) that we implemented using the r-package `glmmTMB` (Brooks et al., 2017). We treated animal IDs as random effects and modeled random slopes for each covariate.

The structure of the movement model was based on a habitat selection model that we previously developed (Hofmann et al., 2021). In the original model (hereafter referred to as *base model*), no interactions between habitat covariates (i.e. Water, DistanceToWater^{0.5}, Woodland, Shrubs/Grassland, Human Influence) and movement covariates (i.e. sl, log(sl), cos(ta)) were considered. Hence, we slightly expanded the base model and proposed interactions between all movement covariates and habitat covariates, assuming that movement behavior depends on habitat conditions. More specifically, we started with the base model and incrementally increased model complexity by adding all possible two-way interactions between habitat covariates and movement covariates. For instance, for the covariate Water, we proposed the interactions Water:sl, Water:log(sl), and Water:cos(ta). Besides these interactions, we also allowed for correlations between turning angles and step lengths by proposing the interactions sl:cos(ta) and log(sl):cos(ta). Furthermore, we formed the interactions sl:LowActivity and log(sl):LowActivity to render that step lengths are likely to be shorter during periods of inactivity. To assess the most parsimonious movement model among model

candidates, we ran stepwise forward model selection based on Akaike’s Information Criterion (AIC, Burnham and Anderson, 2002).

We validated the predictive power of the most parsimonious movement model using k-fold cross-validation for case-control studies Fortin et al. (2009). Specifically, we randomly assigned 80% of the strata to a training set and the remaining 20% to a testing set. With the training data we parametrized a movement model and predicted selection scores $w(x)$ for all steps in the test data. Within each stratum, we then assigned ranks 1-25 to each step based on predicted selection scores, so that rank 1 was given to the step with the highest score $w(x)$. Within each strata, we determined the realized step’s rank and calculated rank frequencies of realized steps across all strata. Finally, we computed Spearman’s rank correlation between ranks and associated frequencies $r_{s,realized}$. We replicated this procedure 100 times and computed the mean correlation coefficient ($\bar{r}_{s,realized}$), as well as its 95% confidence interval across all replicates. For comparison, we repeated the same procedure 100 times assuming random preferences. Random preferences were implemented by discarding the realized step from all strata and identifying the rank of a random step in each stratum. Again, we calculated Spearman’s rank correlation coefficient ($r_{s,random}$), its mean across repetitions ($\bar{r}_{s,random}$), and its 95% confidence interval. Ultimately, this validation proves a significant prediction in case the confidence intervals of $\bar{r}_{s,realized}$ and $\bar{r}_{s,random}$ do not overlap (Fortin et al., 2009).

2.5 Dispersal Simulation

We used the most parsimonious movement model to simulate 80’000 virtual dispersers moving across the KAZA-TFCA. The simulation resembled an “inverted” ISSF and was set up as follows. (1) We defined a random source point and assumed a random initial orientation of the animal. (2) Departing from the source point, we generated 25 random steps by sampling turning angles from a uniform distribution $(-\pi, +\pi)$ and step lengths from our fitted gamma distribution. Similar to the input data, each random step represented the straight line movement possible within 4 hours. To prevent unreasonably large steps, we restricted sampled step lengths to a maximum of 35 km, which corresponded to the farthest distance ever traveled within 4 hours by one of our dispersers. (3) Along each random step, we extracted and averaged values from the habitat covariate layers and we calculated movement covariates. To ensure compatible scales with the fitted movement model, we standardized extracted values using the same scaling parameters that we applied to standardize our input data. (4) We applied the parametrized movement model to predict the selection score $w(x)$

for each step and we translated predicted scores into probabilities using Equation 2. (5) We sampled one of the random steps based on predicted probabilities and determined the animal’s new position. We then repeated steps (2) to (5) until 2’000 steps were realized.

To mitigate edge effects and to deal with random steps leaving the study area, we followed Koen et al. (2010) and artificially expanded all covariate layers by a 100 km wide buffer zone. Within the buffer zone, we randomized covariate values by resampling values from the original covariate layers. Through this buffer zone, simulated dispersers were able to leave and re-enter the main study area. In cases where proposed random steps transgressed the outer border of this buffer zone, we resampled transgressing steps until they fully lied within the buffer zone, forcing individuals to be repelled from these map borders.

2.6 Source Points

We released 80’000 virtual dispersers from 80’000 unique source points distributed across the study area. 50’000 virtual dispersers were released from randomly selected source points within contiguous protected areas that were large enough to sustain viable wild dog populations ($> 700 \text{ km}^2$; Pomilia et al., 2015), whereas another 30’000 dispersers were released at random locations inside the 100 km wide buffer to render potential immigrants into the study system. By distributing source points randomly, the number of source points per km^2 was approximately equal.

2.7 Convergence

Maybe this paragraph should go into another paper? To verify that the number of simulated individuals sufficed to ensure reliable estimates of connectivity, we evaluated how the relative traversal frequency across the landscape changed as we increased the number of simulated individuals. For this, we distributed 1’000 rectangular “checkpoints”, each with an extent of 5 km x 5 km at random locations inside the main study area (excluding the buffer). At each checkpoint, we then determined the relative traversal frequency by simulated dispersers as we increased the number of simulated individuals from 1 to 50’000. To assess variability in the relative traversal frequency, we repeatedly subsampled 100 times from all 50’000 and computed the mean traversal frequency across replicates, as well as its 95% confidence-interval. We deemed that a checkpoint converged as soon the width of the confidence-interval dropped below a value of 0.01.

2.8 Heatmap

To identify dispersal hotspots across our study area, we created a heatmap indicating the absolute frequency at which each raster-cell in our study area was visited by virtual dispersers (e.g. Pe'er and Kramer-Schadt, 2008; Hauenstein et al., 2019; Zeller et al., 2020). For this, we rasterized all simulated trajectories and tallied them into a single map. If the same trajectory crossed a raster-cell twice, we only counted it once, thereby mitigating biases from individuals that were trapped and moved in circles. To achieve high performance rasterization, we used the R-package `terra` (Hijmans, 2021).

2.9 Betweenness

To pinpoint discrete movement corridors and bottlenecks, we converted simulated trajectories into a network and calculated betweenness scores for each raster-cell in the study area (Bastille-Rousseau et al., 2018). Betweenness measures how often a specific network-node (i.e. raster-cell) lies on a shortest path between any other pair of nodes and is therefore pertinent to connectivity (Bastille-Rousseau et al., 2018). To transform simulated trajectories into a network, we overlaid the study area (including the buffer) with a raster resolved at 5 km x 5 km, where the center of each raster-cell served as node in the final network. To identify edges (connections) between the nodes, we used the simulated trajectories and determined all transitions occurring from one node to another, as well as the frequency at which those transitions occurred. This resulted in an edge-list that we translated into a weighted network using the r-package `igraph` (Csardi and Nepusz, 2006). The weight of each edge was determined by the frequency of transitions, yet because `igraph` handles edge weights (ω) as costs, we inverted the traversal frequency in each cell by applying $\omega = \frac{mean(TraversalFrequency)}{TraversalFrequency_i}$. Consequently, edges that were traversed frequently were assigned small weights and vice versa. Finally, we used the weighted network to calculate betweenness scores for all network nodes.

2.10 Inter-Patch Connectivity

To examine functional links between distinct patches in the KAZA-TFCA, we calculated inter-patch connectivity between national parks. The decision to focus on national parks was purely out of simplicity and does not imply that connections between other protected areas are impossible. To quantify inter-patch connectivity, we computed the relative frequency at which dispersers originating from one national park successfully moved into another national park. Successful movement was said to be achieved if the individuals' trajectory intersected

with the corresponding national park at least once. We also recorded the number of steps required until the first intersection with the respective national park, allowing us to compute average dispersal durations from one park to another. In summary, we determined *if* and *how often* dispersers moved between certain national parks, as well as *how long* dispersers had to move to realize those connections.

3 Results

3.1 Movement Model

Our most parsimonious movement model resulting from the ISSF analysis consisted of a movement kernel, a habitat kernel, and their interactions. Compared to the base model reported in (Hofmann et al., 2021), the movement model retained several additional interactions between habitat covariates and movement covariates (Figure 3 and ??). Although multiple models received an AIC weight above zero (Table 1 in Appendix S1), we only considered results from the most parsimonious model for simplicity. Since all models with positive AIC weight contained similar covariates (Table S1), this decision only marginally influenced subsequent steps. To aid with the interpretation of the final model, we followed Fieberg et al. (2020) and produced a collection of auxiliary plots that are provided in Appendix S2.

When all other covariates held constant at their means, the habitat kernel reveals that dispersing wild dogs avoid water but prefer its proximity, that dispersers avoid areas that are covered by woodlands, yet prefer regions covered by shrublands or grasslands, and that dispersers avoid movement through landscapes that are dominated by humans. All of these effects are strong and significant, except for the effect of `distance to water`, which is insignificant on the 5% significance level.

With regards to the movement kernel, the positive estimate for `cos(ta)` indicates that dispersers move with directional persistence, unlike what was proposed by the uniform turning angle distribution. This preference is particularly pronounced when dispersers realize large steps (i.e. move quickly), as indicated by the positive estimates for `cos(ta):sl` and `cos(ta):log(sl)`. Finally, the negative estimate for the interaction `sl:LowActivity` reveals that wild dogs realize shorter steps (i.e. move slower) between 09:00 and 17:00 o'clock than during the rest of the day. Aside from the interaction `sl:LowActivity`, which has a very strong effect, effects are moderate strength yet significant on the 5% significance level; only the interaction `cos(ta):sl` appears as insignificant on the 5% significance level.

Finally, several significant interactions between movement and habitat covariates suggest

that movement behavior is contingent on habitat conditions. For example, there's strong evidence that dispersers realize shorter steps in areas covered by water or in forested areas. Furthermore, it appears that dispersers realize larger steps in areas that are dominated by shrubs/grassland, but shorter steps in areas that are distant to water. Finally, it seems that the preference for directional persistence is less pronounced when dispersers cross human-dominated landscapes, yet more pronounced at great distance from water. Nevertheless, some of these effects are only weakly significant and exhibit small effect sizes.

The k-fold cross-validation of the movement model shows that the model substantially outperforms a random guess (Figure 3b) and correctly assigns high selection scores to realized steps. Confidence intervals of $\bar{r}_{s,realized}$ and $\bar{r}_{s,random}$ do not overlap and therefore prove a reliable prediction. Furthermore, the significant correlation between ranks and corresponding frequencies for realized steps indicates a good fit between predictions and observations (Figure 3b). In comparison to the base model ($\bar{r}_{s,realized} = -0.55$; Hofmann et al., 2021), the inclusion of interactions between movement and habitat covariates slightly improved model performance ($\bar{r}_{s,realized} = -0.65$).

3.2 Dispersal Simulation

On a machine with an octa-core AMD Ryzen 7 2700X processor (8 x 3.6 GHz, 16 logical cores) and 64 GB of RAM, a batch of 1'000 simulated dispersers moving over 2'000 steps required 90 minutes to compute ($\mu = 88.90$, $\sigma = 1.87$). Consequently, the simulation of all 80'000 dispersers (160 Mio. steps) terminated after 120 hours (i.e. five days). Comparable simulations will be substantially faster for smaller study areas and lower resolution covariates, as the covariate extraction from large and high-resolution rasters was computationally the most demanding task. Out of the 50'000 dispersers initiated inside the main source area, only 4.5% were eventually repelled by a map boundary, suggesting that biases due to boundary effects should be minimal.

3.3 Convergence

Our examination of the traversal frequencies across checkpoints shows that we reached convergence across all checkpoints after 10'500 simulated individuals (Figure 4a and b). Even though variability kept decreasing as we increased the number of simulated individuals, the marginal benefit of each additional disperser diminished quickly (Figure 4c).

3.4 Heatmap

Figure 5 depicts the heatmap from all simulated individuals after 2'000 steps (additional maps in Appendix SX). The map shows that large portions of land beyond the borders of the KAZA-TFCA are only infrequently visited by dispersers (dark blue areas), whereas within the KAZA-TFCA several extensive regions are regularly traversed (bright yellow and red areas). Most notably, the region in northern Botswana south of the Linyanti swamp stands out as a highly frequented dispersal hotspot. Still, the presence of several massive water bodies, such as the Okavango Delta, the Makgadikgadi Pan, and the Linyanti swamp, restricts dispersal movements and limits realized connectivity within the KAZA-TFCA. Similarly, the heatmap shows that dispersal across Zambia's and Zimbabwe's part of the KAZA-TFCA is limited, which can largely be attributed to substantial human influences resulting from high human density, roads, and agricultural activities in these areas. Outside the KAZA-TFCA, the most heavily used regions include the areas inside the Central Kalahari National Park in Botswana, the area south-west of the Khaudum National Park in Namibia, and the area around the Liuwa Plains National Park in Zambia. Even though the heatmap facilitates the identification of areas that are intensely used by virtual dispersers, it is not straightforward to spot dispersal corridors. Consequently, we turn to the betweenness map.

3.5 Betweenness

Betweenness scores emerging after 2'000 simulated steps are presented in Figure 6 and reveal a set of discrete dispersal corridors (additional maps in Appendix SX). Again, the region in northern Botswana stands out as a central dispersal hub that connects more remote regions in the study system. Towards east, the extension of this corridor runs through the Chobe National Park into the Hwange National Park. From there, a further extension connects to the distant Matusadona National Park in Zimbabwe. Northwest of the Linyanti ecosystem, a major corridor expands into Angola, where it splits and finally traverses over a long stretch of unprotected area into the Kafue National Park in Zambia. Several additional corridors with slightly lower betweenness scores exist, yet most of them run within the boundaries of the KAZA-TFCA. In general, only a few corridors directly link the peripheral regions of the KAZA-TFCA. For instance, there are few corridors between the Matusadona National Park in Zimbabwe and the Kafue National Park in Zambia. Similarly, there are no direct links between the Zimbabwean and Angolan “spikes” of the KAZA-TFCA. In comparison to the heatmap, the betweenness map facilitates the identification of linear corridors and bottlenecks between habitat patches.

3.6 Inter-Patch Connectivity

Results from the analysis of inter-patch connectivity are given in Figure 7, which shows all realized links by simulated dispersers between national parks. The map furthermore indicates the relative frequency at which dispersers originating from one national park successfully reached another national park, as well as the average duration dispersers had to move to realize those links. It is again worth pointing out that the figure is only intended as an illustration, as for clarity we only considered inter-patch connectivity between national parks (NPs), albeit plenty of links between other protected areas exist. Overall, inter-patch connectivity between NPs in Angola, Namibia, and Botswana appears to be high, with relatively short dispersal durations between national parks. In contrast, we observe that connections from the central region into the Kafue NP in Zambia require rather long dispersal events and are rather infrequent. Similarly, relatively few connections lead into Zimbabwe's Chizarira and Matusadona NP. In some cases, we also find imbalances between ingoing and outgoing links, hinting at potential source-sink dynamics that occur due to asymmetries in disperser's willingness to cross an area, depending on the direction in which the area is traversed. For instance, while a fair share of dispersers originating from the Chizaria NP in Zimbabwe manages to move into the Hwange NP, there are comparably few dispersers that succeed in the opposite direction.

4 Discussion

4.1 Short Summary

We used ISSFs to analyse data of dispersing wild dogs and to parametrize a fully mechanistic movement model describing how dispersers move through the available landscape. We employed the parametrized model as an individual-based movement model to simulate 80'000 dispersing wild dogs moving 2'000 steps across the extent of the KAZA-TFCA, the world's largest transboundary conservation area. Based on simulated dispersal trajectories, we derived three complementary maps, each geared towards a better understanding of dispersal and landscape connectivity. The set of maps included a heatmap, revealing frequently traversed areas, a betweenness-map, delineating critical dispersal corridors, and a map of inter-patch connectivity, indicating the presence or absence of functional links between national parks as well as the average dispersal duration required to realize those links. We thereby showcase that ISSFs offer a simple, yet powerful framework to parametrize movement models and simulate dispersal to assess landscape connectivity. Importantly,

individual-based simulations from ISSFs overcome several conceptual shortcomings inherent to more traditional connectivity modeling techniques, such as LCPA and CT.

4.2 Movement Model

Because our movement model of dispersing wild dogs comprised two interacting kernels, it effectively rendered habitat and movement preferences of dispersers, as well as how preferences depended on habitat conditions. Results from the habitat kernel were largely in concert with previous studies that investigated habitat selection by dispersing wild dogs (Davies-Mostert et al., 2012; Masenga et al., 2016; Woodroffe et al., 2019; O’Neill et al., 2020; Hofmann et al., 2021). However, by also incorporating a movement kernel, we were able to model several additional complexities inherent to dispersal. For instance, we could accommodate that dispersers move with directional persistence (Cozzi et al., 2018; Hofmann et al., 2021) and exhibit step lengths that are correlated with turning angles (Morales et al., 2004; Börger and Fryxell, 2012) by including the appropriate interactions in the movement model. While correlations between step lengths and turning angles could also be rendered by sampling them from copula probability distributions (?), the ISSF framework allowed us to directly incorporate them in the movement model. In addition, by forming interactions between habitat covariates and movement covariates, we could render potential dependencies between movement and habitat preferences. For example, our final model contained an interaction between water-cover and step length, showing that dispersers realize shorter steps in areas covered by water. Likewise, we found that dispersers move more tortuously across water bodies than over dryland, which is clearly to be expected, given that wild dogs have to wade or swim when traversing waterbodies. The ability of accompanying such effects in a single model is one of the great strengths of ISSFs (Avgar et al., 2016; Fieberg et al., 2020), which is why we believe the method offers a suitable framework for simulating movement and assessing connectivity.

4.3 Simulation

Our simulation of 80'000 dispersers moving 2'000 steps across the landscapes of the KAZA-TFCA required five days of computation on a modern desktop machine. The long simulation time was primarily caused by the massive extent considered (ca. 1.8 Mio. km² when including the buffer) and the large number of dispersers simulated. Most connectivity studies are limited to much smaller extents (e.g. Kanagaraj et al., 2013; Clark et al., 2015; McClure et al., 2016; Abrahms et al., 2017; Zeller et al., 2020) and will therefore achieve faster

simulation times. We also believe that fewer simulated dispersers will often suffice, as the relative traversal frequency by simulated individuals through randomly placed checkpoints in the study area converged already after 10'500 simulated individuals. However, the required number of simulated individuals to achieve reliable estimates of connectivity will vary depending on the structure of the landscape and the dispersal ability of the focal species.

4.4 Maps

The heatmap generated from simulated dispersal trajectories highlighted that numerous individuals traversed the Moremi NP and the Chobe NP in northern Botswana. We previously uncovered the same area as potential dispersal hotspot using least-cost methods (Hofmann et al., 2021), yet it has been questioned whether this was a consequence of the region being located in the center of the study area and connections being enforced between predefined start and endpoints. With the current simulations, connections were no longer enforced and simulated individuals were able to leave the main study area. Nonetheless, a majority of simulated individuals traversed the area in northern Botswana. This suggests that the dispersal hotspot is not merely an artifact of the applied method but truly results from landscape characteristics. The same region is also pronounced on the betweenness map, implying that it facilitates the relocation of individuals into more remote regions of the KAZA-TFCA. While this is an example of an area where both the heatmap and the betweenness map attest great importance, there are other instances where this is not true. For example, while the area between the Lengue-Luiana NP in Angola and the Kafue NP in Zambia receives a high betweenness-score, the heatmap shows that the area is only rarely traversed by dispersers. Hence, despite the region's importance for linking Angola's NPs to Zambia's NP, only few simulated dispersers actually used the corridor. Conversely, we find that the Central Kalahari NP receives a low betweenness score, despite being highly frequented by simulated dispersers. Besides highlighting frequently traversed areas and dispersal corridors, we also also consulted inter-patch connectivity between NPs, depicting the presence or absence of functional links between national parks. Because dispersal movements were rendered sequentially, we could also derive the average dispersal duration that was required to realize those links. This enabled us to show that movements from Angola into Zambia's Kafue NP are rare and require a large number of steps , whereas dispersal between the Moremi NP and Chobe NP occurs frequently and requires relatively few steps. All in all, the rich inference that can be drawn from the ensemble of proposed connectivity maps.

4.5 Related Literature

Our approach of simulating movement to assess connectivity is related to a series of previously published papers. Clark et al. (2015), for instance, collected GPS data on American black bears (*Ursus americanus*) and used SSFs to fit a habitat selection model. They then used the fitted model to simulate movement and identify likely movement corridors between four distinct habitat patches. For the same species, Zeller et al. (2020) used SSFs and forecasted seasonal habitat connectivity under changing land-use. Because both of these studies relied on *regular* SSFs, rather than *integrated* SSFs, they could not account for interdependencies among habitat and movement preferences (Avgar et al., 2016). In addition, both studies employed data collected on resident individuals instead of dispersers, although evidence suggests that residents are more reluctant to cross areas that are readily traversed by dispersers (Elliot et al., 2014; Gastón et al., 2016; Abrahms et al., 2017; Keeley et al., 2017). The application of data collected on resident animals may therefore result in an underestimate of connectivity (Elliot et al., 2014). Two further studies that used SSFs to simulate animal movement have been conducted by Potts et al. (2013) and Signer et al. (2017), yet the primary purpose here was to estimate steady-state utilization distributions of resident animals and not on the analysis of dispersal and connectivity.

4.6 Advantages of ISSF Simulations

A simulation-based approach as proposed in this article offers several advantages over LCRA and CT. In contrast to LCRA, for instance, an individual-based simulation does not require to assume known endpoints. Instead, each endpoint emerges naturally from a simulated dispersal trajectory. The ability of not needing to provide endpoints is particularly valuable for dispersal studies, because dispersers often venture into unfamiliar territory and are therefore unlikely to know the destination of their journey (Elliot et al., 2014; Abrahms et al., 2017; Cozzi et al., 2020). Moreover, LCRA always enforces a connection towards the pre-defined endpoints, even if associated movement costs are unreasonably high. With simulations from ISSFs this is no longer the case. A connectivity model that does not require pre-defined endpoints also ensures that movement corridors are not enforced between certain start- and endpoints, which permits to detect potential routes that do not lead into suitable habitats but into ecological traps (Dwernychuk and Boag, 1972; Van der Meer et al., 2014) or areas with a high susceptibility for human-wildlife conflicts (Cushman et al., 2018).

In contrast to LCRA and CT, simulations from ISSFs furthermore yield the advantage of an explicit representation of time, which enables to answer questions such as: “*How long*

will it take a disperser to move from A to B?” or “*Is it possible for a disperser to move from A to B within X days?*”. An explicit representation of time also yields opportunities for studying how seasonality affects connectivity and to investigate whether some dispersal corridors are only available temporarily (*dynamic connectivity*; Zeller et al., 2020). With LCPA or CT, incorporating seasonality is currently impractical, as both methods require a static permeability surface as inputs. Hence, the only possibility to study seasonality effects is to repeat the same analysis using different permeability surfaces, each rendering the environment at a different point in time (e.g. Benz et al., 2016; Osipova et al., 2019). With simulations from ISSFs, on the other hand, the environment can be rendered dynamically “as the dispersers move”, such that simulated individuals can respond to seasonal factors directly within the simulation. Hence, rather than employing a set of static habitat layers, each layer would be updated as the dispersers move, thus correctly rendering seasonal changes in the environment.

While an explicit representation of time provides several benefits, it requires that step lengths and turning angles are modeled properly (Kanagaraj et al., 2013), so that dispersal durations between areas can be estimated reliably. Correctly rendering step lengths and turning angles under varying environmental conditions is one of the key strengths of ISSFs (Avgar et al., 2016; Prokopenko et al., 2017; Fieberg et al., 2020), which is why we believe that the framework is exceptionally well suited for simulating dispersal and assessing landscape connectivity. In addition, the framework enables to model autocorrelation between step lengths and turning angles, thereby incorporating directional persistence. Here, we only considered first order autocorrelation, i.e. correlation between two consecutive steps. Although higher order autocorrelation is conceivable and might be desirable to model, this requires vast amounts of GPS data that is not intercepted by missing fixes and is therefore often impractical to model in reality.

4.7 Disadvantages of ISSF Simulations

Despite the benefits that simulations from ISSFs offer, we also want to confer some of the non-trivial modeling decisions involved.

In particular, it is worth pointing out five modeling decisions: (1) number of simulated individuals, (2) location of source points, (3) dispersal duration, (4) boundary behavior, and (5) how to handle uncertainty and individual variability.

(1) When simulating dispersal, the modeler needs to decide on the number of simulated individuals. A higher number is always desirable, as each additional disperser provides novel

information about landscape connectivity. However, this comes at the cost of computational efficiency, implying that a trade-off needs to be managed. As noted by Signer et al. (2017), the trade-off can be handled by simulating additional individuals only until estimated metrics converge. Here, we employed the relative traversal frequency across checkpoints as target metric and found that convergence across all checkpoints was already achieved after 10'500 simulated individuals, yet we recognize that this strongly depends on the focal species dispersal ability and landscape characteristics.

(2) Aside from specifying the absolute number of simulated individuals, one also needs to define a source point within a suitable source area for each individual. Here, we placed source points within protected areas large enough to sustain viable wild dog populations. Given that wild dogs primarily survive in formally protected areas (Woodroffe and Ginsberg, 1999; Davies-Mostert et al., 2012; Woodroffe and Sillero-Zubiri, 2012; Van der Meer et al., 2014) we considered this decision to be appropriate. Due to a lack of precise knowledge about wild dog abundances in the different protected areas, we distributed source points randomly within them. If, however, corresponding data is available, source points can be distributed accordingly, reflecting the fact that source areas do not necessarily produce an identical number of dispersers. Alternatively, source points can be distributed homogeneously, but be weighted afterwards according to estimated densities in the respective source area. In cases where knowledge about suitable source areas is lacking, these could also be delineated using habitat suitability models (e.g. Squires et al., 2013). After all, the challenge of selecting meaningful source areas and source points is not unique to individual-based simulations and also applies to LCPA or CT.

(3) When employing ISSFs to simulate dispersers, it is also required to decide on meaningful dispersal durations (i.e. number of simulated steps). When sufficient dispersal data of the focal species is available, dispersal durations can be sampled from observed dispersal events. Due to the low number of observed dispersal events and due to the great variability in wild dogs' dispersal distances (Davies-Mostert et al., 2012; Masenga et al., 2016; Cozzi et al., 2020) we opted against this approach. Instead, we simulated individuals for 2'000 steps, which is at the upper end of observed dispersal durations. Once the trajectories have been simulated, it is straightforward to subsample them and investigate the sensitivity of derived results with regards to different dispersal durations.

(4) Unless simulated individuals are strongly drawn towards a point of attraction (e.g. Signer et al. (2017)), dispersers eventually approach a map boundary, so that one or several of the proposed random steps leave the study area. In theoretical applications, this issue can

be circumvented by simulating movement on a torus (?). For real data, however, alternative solutions are needed. One approach would be to terminate the simulation, assuming that the simulated animal left the study area forever. This can be problematic for individuals that are initiated in areas that are located close to map boundaries, especially since already a single random step leaving the study area breaks the simulation. Here, instead of breaking the simulation loop, we simply resampled transgressing random steps until they fully lied within the study area. This enforced simulated dispersers to be repelled by map boundaries and to remain within the main study area. Additionally, we artificially increased the study area using a buffer zone with randomized covariate values. This enabled virtual dispersers to leave and re-enter the main study area. The same method has been shown to effectively mitigate edge effects for graph-based connectivity models (Koen et al., 2010).

(5) We simulated dispersal using point estimates from our most parsimonious movement model. Depending on the amount of data and individual variability, these estimates can be subject to substantial uncertainty. For dispersal studies in particular, the low amount of data typically results in model coefficients with large confidence intervals (Wiegand et al., 2003; Kramer-Schadt et al., 2007). In these cases, point estimates may lead to biased connectivity estimates, which is why we urge future studies to investigate the sensitivity of ISSF simulations with respect to employed model parameters and to simulate dispersal based on parameters that encapsulate model uncertainty.

4.8 Conclusion

To this end, we proposed and applied a simple three-step workflow that uses ISSFs to parametrize an individual-based movement model from which dispersal can be simulated with the purpose of informing about landscape connectivity. By explicitly simulating dispersal trajectories, simulations enable a more mechanistic understanding of connectivity, and overcome several of the unrealistic assumptions inherent to graph-based connectivity models, such as least-cost analysis or circuit theory. We exemplified the application of the proposed workflow using data of dispersing wild dogs to assess landscape connectivity for the species within the KAZA-TFCA ecosystem. With this, we hope to have sparked interest in the powerful framework of step selection functions for investigating dispersal behavior and landscape connectivity. Nevertheless, we propose to view simulations from ISSF-models as complementary and not as substitutes to traditional connectivity modeling techniques.

5 Authors' Contributions

D.D.H., D.M.B., A.O. and G.C. conceived the study and designed methodology; D.M.B., G.C., and J.W.M. collected the data; D.D.H. and D.M.B. analysed the data; G.C. and A.O. assisted with modeling; D.D.H., D.M.B., and G.C. wrote the first draft of the manuscript and all authors contributed to the drafts at several stages and gave final approval for publication.

6 Data Availability

GPS movement data of dispersing wild dogs will be made available on dryad at the time of publication. Access to R-scripts that exemplify the application of the proposed framework to simulated data are provided through Github.

7 Acknowledgements

We thank the Ministry of Environment and Tourism of Botswana for granting permission to conduct this research. We thank C. Botes, I. Clavadetscher, and G. Camenisch for assisting with wild dog immobilizations. We also thank B. Abrahms for sharing her data of three dispersing wild dogs. Furthermore, we are indebted to Johannes Signer for assisting with the simulation algorithm. This study was funded by Basler Stiftung für Biologische Forschung, Claraz Foundation, Idea Wild, Jacot Foundation, National Geographic Society, Parrotia Stiftung, Stiftung Temperatio, Wilderness Wildlife Trust Foundation, Forschungskredit der Universität Zürich, and a Swiss National Science Foundation Grant (31003A_182286) to A. Ozgul.

References

- Abrahms, B., Sawyer, S. C., Jordan, N. R., McNutt, J. W., Wilson, A. M., and Brashares, J. S. (2017). Does Wildlife Resource Selection Accurately Inform Corridor Conservation? *Journal of Applied Ecology*, 54(2):412–422.
- Adriaensen, F., Chardon, J., De Blust, G., Swinnen, E., Villalba, S., Gulinck, H., and Matthysen, E. (2003). The Application of Least-Cost Modelling as a Functional Landscape Model. *Landscape and Urban Planning*, 64(4):233–247.
- Allen, C. H., Parrott, L., and Kyle, C. (2016). An Individual-Based Modelling Approach to Estimate Landscape Connectivity for Bighorn Sheep (*Ovis canadensis*). *PeerJ*, 4:e2001.
- Avgar, T., Lele, S. R., Keim, J. L., and Boyce, M. S. (2017). Relative Selection Strength: Quantifying Effect Size in Habitat- and Step-Selection Inference. *Ecology and Evolution*, 7(14):5322–5330.
- Avgar, T., Potts, J. R., Lewis, M. A., and Boyce, M. S. (2016). Integrated Step Selection Analysis: Bridging the Gap Between Resource Selection and Animal Movement. *Methods in Ecology and Evolution*, 7(5):619–630.
- Baguette, M., Blanchet, S., Legrand, D., Stevens, V. M., and Turlure, C. (2013). Individual Dispersal, Landscape Connectivity and Ecological Networks. *Biological Reviews*, 88(2):310–326.
- Bastille-Rousseau, G., Douglas-Hamilton, I., Blake, S., Northrup, J. M., and Wittemyer, G. (2018). Applying Network Theory to Animal Movements to Identify Properties of Landscape Space Use. *Ecological Applications*, 28(3):854–864.
- Behr, D. M., McNutt, J. W., Ozgul, A., and Cozzi, G. (2020). When to Stay and When to Leave? Proximate Causes of Dispersal in an Endangered Social Carnivore. *Journal of Animal Ecology*, 89(10):2356–2366.
- Benz, R. A., Boyce, M. S., Thurfjell, H., Paton, D. G., Musiani, M., Dormann, C. F., and Ciuti, S. (2016). Dispersal Ecology Informs Design of Large-Scale Wildlife Corridors. *PLOS ONE*, 11(9):e0162989.
- Bovet, P. and Benhamou, S. (1991). Optimal Sinuosity in Central Place Foraging Movements. *Animal Behaviour*, 42(1):57–62.
- Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., and Bolker, B. M. (2017). glmmTMB Balances Speed and Flexibility among Packages for Zero-Inflated Generalized Linear Mixed Modeling. *The R Journal*, 9(2):378–400.
- Brown, J. H. and Kodric-Brown, A. (1977). Turnover Rates in Insular Biogeography: Effect of Immigration on Extinction. *Ecology*, 58(2):445–449.
- Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Science & Business Media, Ney York, NY, USA.
- Börger, L. and Fryxell, J. (2012). Quantifying Individual Differences in Dispersal Using Net Squared Displacement. In Clobert, J., Baguette, M., Benton, T. G., and Bullock, J. M., editors, *Dispersal Ecology and Evolution*. Oxford University Press, Oxford, UK.
- Castelló, J. R. (2018). *Canids of the World: Wolves, Wild dogs, Foxes, Jackals, Coyotes, and their Relatives*. Princeton Field Guides. Princeton University Press, Princeton, NJ.
- Clark, J. D., Laufenberg, J. S., Davidson, M., and Murrow, J. L. (2015). Connectivity among Subpopulations of Louisiana Black Bears as Estimated by a Step Selection Function. *The Journal of Wildlife Management*, 79(8):1347–1360.

- Clobert, J., Baguette, M., Benton, T. G., and Bullock, J. M. (2012). *Dispersal Ecology and Evolution*. Oxford University Press, Oxford, UK.
- Cozzi, G. (2013). *Patterns of Habitat Use and Segregation among African Large Carnivores*. PhD Thesis, University of Zurich, Zurich.
- Cozzi, G., Behr, D., Webster, H., Claase, M., Bryce, C., Modise, B., McNutt, J., and Ozgul, A. (2020). The Walk of Life: African Wild Dog Dispersal and its Implications for Management and Conservation across Transfrontier Landscapes. In press.
- Cozzi, G., Maag, N., Börger, L., Clutton-Brock, T. H., and Ozgul, A. (2018). Socially Informed Dispersal in a Territorial Cooperative Breeder. *Journal of Animal Ecology*, 87(3):838–849.
- Csardi, G. and Nepusz, T. (2006). The igraph Software Package for Complex Network Research. *InterJournal*, Complex Systems:1695.
- Cushman, S. A., Elliot, N. B., Bauer, D., Kesch, K., Bahaa-el din, L., Bothwell, H., Flyman, M., Mtare, G., Macdonald, D. W., and Loveridge, A. J. (2018). Prioritizing Core Areas, Corridors and Conflict Hotspots for Lion Conservation in Southern Africa. *PLOS ONE*, 13(7):e0196213.
- Davies-Mostert, H. T., Kamler, J. F., Mills, M. G. L., Jackson, C. R., Rasmussen, G. S. A., Groom, R. J., and Macdonald, D. W. (2012). Long-Distance Transboundary Dispersal of African Wild Dogs among Protected Areas in Southern Africa. *African Journal of Ecology*, 50(4):500–506.
- Diniz, M. F., Cushman, S. A., Machado, R. B., and De Marco Júnior, P. (2019). Landscape Connectivity Modeling from the Perspective of Animal Dispersal. *Landscape Ecology*, 35:41–58.
- Doerr, V. A. J., Barrett, T., and Doerr, E. D. (2011). Connectivity, Dispersal Behaviour and Conservation under Climate Change: A Response to Hodgson et al.: Connectivity and Dispersal Behaviour. *Journal of Applied Ecology*, 48(1):143–147.
- Dwernychuk, L. W. and Boag, D. A. (1972). Ducks Nesting in Association with Gulls - An Ecological Trap? *Canadian Journal of Zoology*, 50(5):559–563. Publisher: NRC Research Press.
- Eddelbuettel, D. (2013). *Seamless R and C++ Integration with Rcpp*. Springer, New York. ISBN 978-1-4614-6867-7.
- Eddelbuettel, D. and François, R. (2011). Rcpp: Seamless R and C++ Integration. *Journal of Statistical Software*, 40(8):1–18.
- Elliot, N. B., Cushman, S. A., Macdonald, D. W., and Loveridge, A. J. (2014). The Devil is in the Dispersers: Predictions of Landscape Connectivity Change with Demography. *Journal of Applied Ecology*, 51(5):1169–1178.
- Fahrig, L. (2003). Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):487–515.
- Fieberg, J., Signer, J., Smith, B., and Avgar, T. (2020). A ‘How-to’ Guide for Interpreting Parameters in Resource- and Step-Selection Analyses. preprint, Ecology.
- Fortin, D., Beyer, H. L., Boyce, M. S., Smith, D. W., Duchesne, T., and Mao, J. S. (2005). Wolves Influence Elk Movements: Behavior Shapes a Trophic Cascade in Yellowstone National Park. *Ecology*, 86(5):1320–1330.
- Fortin, D., Fortin, M.-E., Beyer, H. L., Duchesne, T., Courant, S., and Dancose, K. (2009). Group-Size-Mediated Habitat Selection and Group Fusion–Fission Dynamics of Bison under Predation Risk. *Ecology*, 90(9):2480–2490.

- Frankham, R., Briscoe, D. A., and Ballou, J. D. (2002). *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK.
- Gardner, R. H. and Gustafson, E. J. (2004). Simulating Dispersal of Reintroduced Species within Heterogeneous Landscapes. *Ecological Modelling*, 171(4):339–358.
- Gastón, A., Blázquez-Cabrera, S., Garrote, G., Mateo-Sánchez, M. C., Beier, P., Simón, M. A., and Saura, S. (2016). Response to Agriculture by a Woodland Species Depends on Cover Type and Behavioural State: Insights from Resident and Dispersing Iberian Lynx. *Journal of Applied Ecology*, 53(3):814–824.
- Graf, R. F., Kramer-Schadt, S., Fernández, N., and Grimm, V. (2007). What you see is where you go? Modeling Dispersal in Mountainous Landscapes. *Landscape Ecology*, 22(6):853–866.
- Gustafson, E. J. and Gardner, R. H. (1996). The Effect of Landscape Heterogeneity on the Probability of Patch Colonization. *Ecology*, 77(1):94–107.
- Hanski, I. (1999). *Metapopulation Ecology*. Oxford University Press.
- Hauenstein, S., Fattebert, J., Grüebler, M. U., Naef-Daenzer, B., Pe'er, G., and Hartig, F. (2019). Calibrating an Individual-Based Movement Model to Predict Functional Connectivity for Little Owls. *Ecological Applications*, 29(4):e01873.
- Hijmans, R. J. (2021). *terra: Spatial Data Analysis*. R package version 1.2-10.
- Hofmann, D. D., Behr, D. M., McNutt, J. W., Ozgul, A., and Cozzi, G. (2021). Bound within Boundaries: How Well Do Protected Areas Match Movement Corridors of their Most Mobile Protected Species? *Journal of Applied Ecology*. (in 2nd review).
- Hunziker, P. (2021). *velox: Fast Raster Manipulation and Extraction*. R package version 0.2.1.
- Kanagaraj, R., Wiegand, T., Kramer-Schadt, S., and Goyal, S. P. (2013). Using Individual-Based Movement Models to Assess Inter-Patch Connectivity for Large Carnivores in Fragmented Landscapes. *Biological Conservation*, 167:298 – 309.
- Keeley, A. T., Beier, P., Keeley, B. W., and Fagan, M. E. (2017). Habitat Suitability is a Poor Proxy for Landscape Connectivity during Dispersal and Mating Movements. *Landscape and Urban Planning*, 161:90–102.
- Koen, E. L., Bowman, J., Sadowski, C., and Walpole, A. A. (2014). Landscape Connectivity for Wildlife: Development and Validation of Multispecies Linkage Maps. *Methods in Ecology and Evolution*, 5(7):626–633.
- Koen, E. L., Garraway, C. J., Wilson, P. J., and Bowman, J. (2010). The Effect of Map Boundary on Estimates of Landscape Resistance to Animal Movement. *PLOS ONE*, 5(7):e11785.
- Kramer-Schadt, S., Revilla, E., Wiegand, T., and Breitenmoser, U. (2004). Fragmented Landscapes, Road Mortality and Patch Connectivity: Modelling Influences on the Dispersal of Eurasian lynx. *Journal of Applied Ecology*, 41(4):711–723.
- Kramer-Schadt, S., Revilla, E., Wiegand, T., and Grimm, V. (2007). Patterns for Parameters in Simulation Models. *Ecological Modelling*, 204(3-4):553–556.
- Landguth, E. L., Hand, B. K., Glassy, J., Cushman, S. A., and Sawaya, M. A. (2012). UNICOR: A Species Connectivity and Corridor Network Simulator. *Ecography*, 35(1):9–14.
- Latham, A. D. M., Latham, M. C., Boyce, M. S., and Boutin, S. (2011). Movement Responses by Wolves to Industrial Linear Features and Their Effect on Woodland Caribou in Northeastern Alberta. *Ecological Applications*, 21(8):2854–2865.

- Leigh, K. A., Zenger, K. R., Tammen, I., and Raadsma, H. W. (2012). Loss of Genetic Diversity in an Outbreeding Species: Small Population Effects in the African Wild Dog (*Lycaon pictus*). *Conservation Genetics*, 13(3):767–777.
- MacArthur, R. H. and Wilson, E. O. (2001). *The Theory of Island Biogeography*, volume 1. Princeton University Press, Princeton, New Jersey, USA.
- Martensen, A. C., Saura, S., and Fortin, M. (2017). Spatio-Temporal Connectivity: Assessing the Amount of Reachable Habitat in Dynamic Landscapes. *Methods in Ecology and Evolution*, 8(10):1253–1264.
- Masenga, E. H., Jackson, C. R., Mjingo, E. E., Jacobson, A., Riggio, J., Lyamuya, R. D., Fyumagwa, R. D., Borner, M., and Roskaft, E. (2016). Insights into Long-Distance Dispersal by African Wild Dogs in East Africa. *African Journal of Ecology*, 54(1):95–98.
- McClure, M. L., Hansen, A. J., and Inman, R. M. (2016). Connecting Models to Movements: Testing Connectivity Model Predictions against Empirical Migration and Dispersal Data. *Landscape Ecology*, 31(7):1419–1432.
- McNutt, J. (1996). Sex-Biased Dispersal in African Wild Dogs (*Lycaon pictus*). *Animal Behaviour*, 52(6):1067–1077.
- McRae, B. H. (2006). Isolation by Resistance. *Evolution*, 60(8):1551–1561.
- McRae, B. H., Dickson, B. G., Keitt, T. H., and Shah, V. B. (2008). Using Circuit Theory to Model Connectivity in Ecology, Evolution, and Conservation. *Ecology*, 89(10):2712–2724.
- Morales, J. M., Haydon, D. T., Frair, J., Holsinger, K. E., and Fryxell, J. M. (2004). Extracting More out of Relocation Data: Building Movement Models as Mixtures of Random Walks. *Ecology*, 85(9):2436–2445.
- Muff, S., Signer, J., and Fieberg, J. (2020). Accounting for Individual-Specific Variation in Habitat-Selection Studies: Efficient Estimation of Mixed-Effects Models Using Bayesian or Frequentist Computation. *Journal of Animal Ecology*, 89(1):80–92.
- Nathan, R. (2008). An Emerging Movement Ecology Paradigm. *Proceedings of the National Academy of Sciences*, 105(49):19050–19051.
- Osipova, L., Okello, M. M., Njumbi, S. J., Ngene, S., Western, D., Hayward, M. W., and Balkenhol, N. (2019). Using Step-Selection Functions to Model Landscape Connectivity for African Elephants: Accounting for Variability across Individuals and Seasons. *Animal Conservation*, 22(1):35–48.
- O'Neill, H. M. K., Durant, S. M., and Woodroffe, R. (2020). What Wild Dogs Want: Habitat Selection Differs across Life Stages and Orders of Selection in a Wide-Ranging Carnivore. *BMC Zoology*, 5(1).
- Panzacchi, M., Van Moorter, B., Strand, O., Saerens, M., Kivimäki, I., St. Clair, C. C., Herfindal, I., and Boitani, L. (2016). Predicting the Continuum Between Corridors and Barriers to Animal Movements Using Step Selection Functions and Randomized Shortest Paths. *Journal of Animal Ecology*, 85(1):32–42.
- Perrin, N. and Mazalov, V. (1999). Dispersal and Inbreeding Avoidance. *The American Naturalist*, 154(3):282–292.
- Perrin, N. and Mazalov, V. (2000). Local Competition, Inbreeding, and the Evolution of Sex-Biased Dispersal. *The American Naturalist*, 155(1):116–127.
- Pe'er, G. and Kramer-Schadt, S. (2008). Incorporating the Perceptual Range of Animals into Connectivity Models. *Ecological Modelling*, 213(1):73–85.
- Pinto, N. and Keitt, T. H. (2009). Beyond the Least-Cost Path: Evaluating Corridor Redundancy Using a Graph-Theoretic Approach. *Landscape Ecology*, 24(2):253–266.

- Pomilia, M. A., McNutt, J. W., and Jordan, N. R. (2015). Ecological Predictors of African Wild Dog Ranging Patterns in Northern Botswana. *Journal of Mammalogy*, 96(6):1214–1223.
- Potts, J. R., Bastille-Rousseau, G., Murray, D. L., Schaefer, J. A., and Lewis, M. A. (2013). Predicting local and non-local effects of resources on animal space use using a mechanistic step selection model. *Methods in Ecology and Evolution*, 5(3):253–262.
- Prokopenko, C. M., Boyce, M. S., and Avgar, T. (2017). Characterizing Wildlife Behavioural Responses to Roads Using Integrated Step Selection Analysis. *Journal of Applied Ecology*, 54(2):470–479.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Revilla, E. and Wiegand, T. (2008). Individual Movement Behavior, Matrix Heterogeneity, and the Dynamics of Spatially Structured Populations. *Proceedings of the National Academy of Sciences*, 105(49):19120–19125.
- Revilla, E., Wiegand, T., Palomares, F., Ferreras, P., and Delibes, M. (2004). Effects of Matrix Heterogeneity on Animal Dispersal: From Individual Behavior to Metapopulation-Level Parameters. *The American Naturalist*, 164(5):E130–E153.
- Rudnick, D., Ryan, S., Beier, P., Cushman, S., Dieffenbach, F., Epps, C., Gerber, L., Harter, J., Jenness, J., Kintsch, J., Merenlender, A., Perkl, R., Perziosi, D., and Trombulack, S. (2012). The Role of Landscape Connectivity in Planning and Implementing Conservation and Restoration Priorities. *Issues in Ecology*.
- Sawyer, S. C., Epps, C. W., and Brashares, J. S. (2011). Placing Linkages among Fragmented Habitats: Do Least-Cost Models Reflect How Animals Use Landscapes? *Journal of Applied Ecology*, 48(3):668–678.
- Schultz, C. B. and Crone, E. E. (2001). Edge-Mediated Dispersal Behavior in a Prairie Butterfly. *Ecology*, 82(7):1879–1892.
- Signer, J., Fieberg, J., and Avgar, T. (2017). Estimating Utilization Distributions from Fitted Step-Selection Functions. *Ecosphere*, 8(4):e01771.
- Squires, J. R., DeCesare, N. J., Olson, L. E., Kolbe, J. A., Hebblewhite, M., and Parks, S. A. (2013). Combining Resource Selection and Movement Behavior to Predict Corridors for Canada Lynx at their Southern Range Periphery. *Biological Conservation*, 157:187–195.
- Thurfjell, H., Ciuti, S., and Boyce, M. S. (2014). Applications of Step-Selection Functions in Ecology and Conservation. *Movement Ecology*, 2(4).
- Tischendorf, L. and Fahrig, L. (2000). On the Usage and Measurement of Landscape Connectivity. *Oikos*, 90(1):7–19.
- Turchin, P. (1998). *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Plants and Animals*. Sinauer Associates, Sunderland, MA, USA.
- Van der Meer, E., Fritz, H., Blinston, P., and Rasmussen, G. S. (2014). Ecological Trap in the Buffer Zone of a Protected Area: Effects of Indirect Anthropogenic Mortality on the African Wild Dog (*Lycaon pictus*). *Oryx*, 48(2):285–293.
- Vasudev, D., Fletcher, R. J., Goswami, V. R., and Krishnadas, M. (2015). From Dispersal Constraints to Landscape Connectivity: Lessons from Species Distribution Modeling. *Ecography*, 38(10):967–978.
- Vasudev, D., Goswami, V. R., and Oli, M. K. (2021). Detecting Dispersal: A Spatial Dynamic Occupancy Model to Reliably Quantify Connectivity across Heterogeneous Conservation Landscapes. *Biological Conservation*, 253:108874.

- Wiegand, T., Jeltsch, F., Hanski, I., and Grimm, V. (2003). Using Pattern-Oriented Modeling for Revealing Hidden Information: A Key for Reconciling Ecological Theory and Application. *Oikos*, 100(2):209–222.
- Wolski, P., Murray-Hudson, M., Thito, K., and Cassidy, L. (2017). Keeping it Simple: Monitoring Flood Extent in Large Data-Poor Wetlands Using MODIS SWIR Data. *International Journal of Applied Earth Observation and Geoinformation*, 57:224–234.
- Woodroffe, R. and Ginsberg, J. R. (1999). Conserving the African Wild Dog *Lycaon pictus*. Diagnosing and Treating Causes of Decline. *Oryx*, 33(2):132–142.
- Woodroffe, R., Rabaiotti, D., Ngatia, D. K., Smallwood, T. R. C., Strelak, S., and O'Neill, H. M. K. (2019). Dispersal Behaviour of African Wild Dogs in Kenya. *African Journal of Ecology*.
- Woodroffe, R. and Sillero-Zubiri, C. (2012). *Lycaon pictus*. *The IUCN Red List of Threatened Species*, 2012:e. T12436A16711116.
- Zeller, K. A., McGarigal, K., and Whiteley, A. R. (2012). Estimating Landscape Resistance to Movement: A Review. *Landscape Ecology*, 27(6):777–797.
- Zeller, K. A., Wattles, D. W., Bauder, J. M., and DeStefano, S. (2020). Forecasting Seasonal Habitat Connectivity in a Developing Landscape. *Land*, 9(7):233.

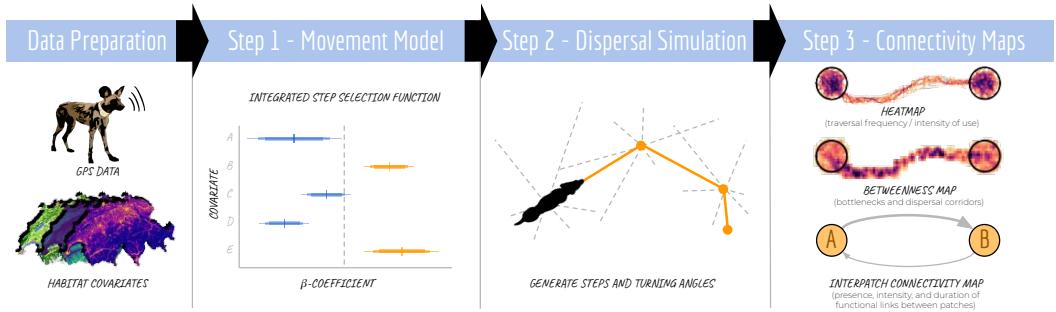


Figure 1: Flowchart of the simulation-based connectivity analysis as proposed in this article. First, GPS data and habitat covariates have to be collected. The combined data is then analyzed in an integrated step selection model, which enables the parametrization of the focal species' habitat and movement kernels and results in a mechanistic movement model. The parametrized model is then treated as an individual-based movement model and used to simulate dispersal trajectories. Ultimately, simulated trajectories serve to produce a set of maps that are pertinent to landscape connectivity. This includes a heatmap, indicating the traversal frequency across each spatial unit of the study area, a betweenness map, highlighting movement corridors and bottlenecks, and, finally, an inter-patch connectivity map, where the frequency of connections and their average duration can be depicted.

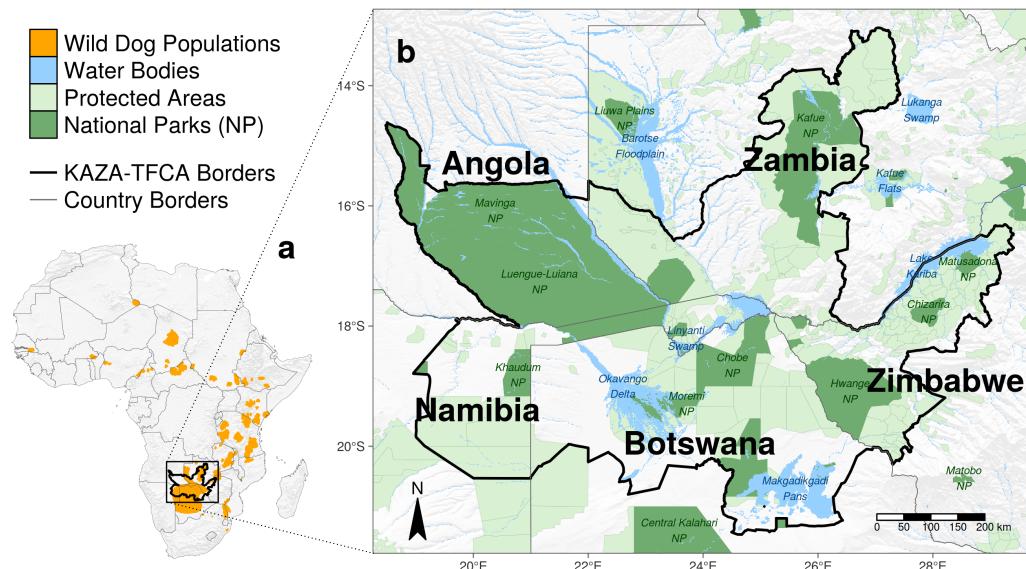


Figure 2: Illustration of the study area located in southern Africa. (a) The study area was confined by a bounding box spanning the entire KAZA-TFCA and comprised parts of Angola, Namibia, Botswana, Zimbabwe, and Zambia. (b) The KAZA-TFCA represents the world's largest terrestrial conservation area and covers a total of 520'000 km². Its main purpose is to re-establish connectivity between already-existing national parks (dark green) and other protected areas (light green).

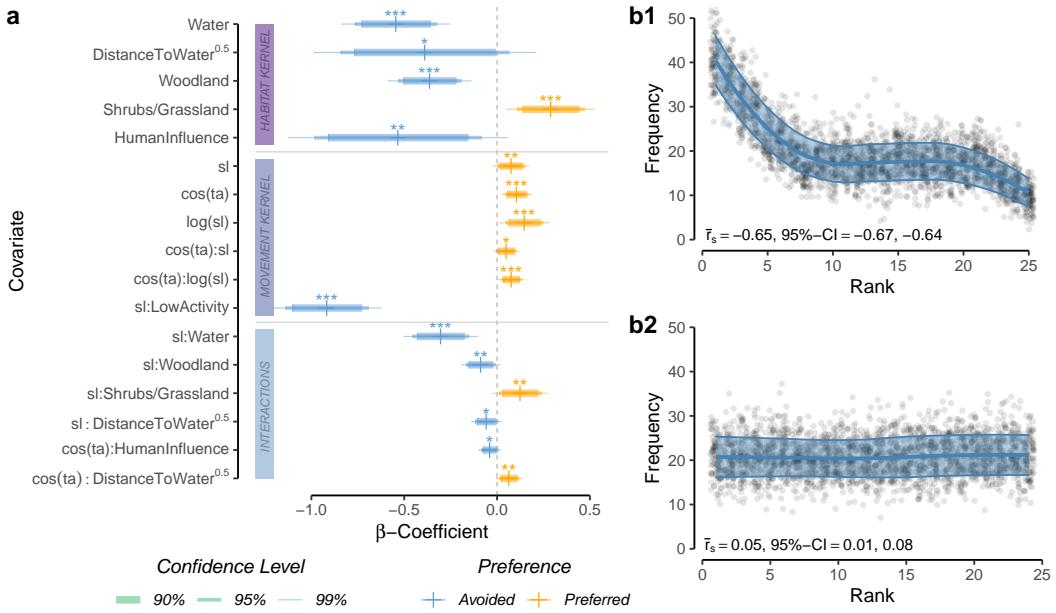


Figure 3: (a) Most parsimonious movement model for dispersing wild dogs. The model comprises a habitat kernel, a movement kernel, as well as their interactions. The horizontal line segments delineate the 90%, 95%, and 99% confidence-intervals for the respective β -coefficients. Significance codes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (b) Results from the k-fold cross validation procedure. The upper plot shows rank frequencies of realized steps according to model predictions with known preferences, whereas the lower plot shows rank frequencies of realized steps when assuming random preferences. The blue ribbon shows the prediction interval around a loess smoothing regression that we fitted to ease the interpretation of the plots. The significant correlation between rank and associated frequency in (b1) highlights that the most parsimonious model successfully outperforms a random guess (b2) and assigns comparably high selection scores to realized steps.

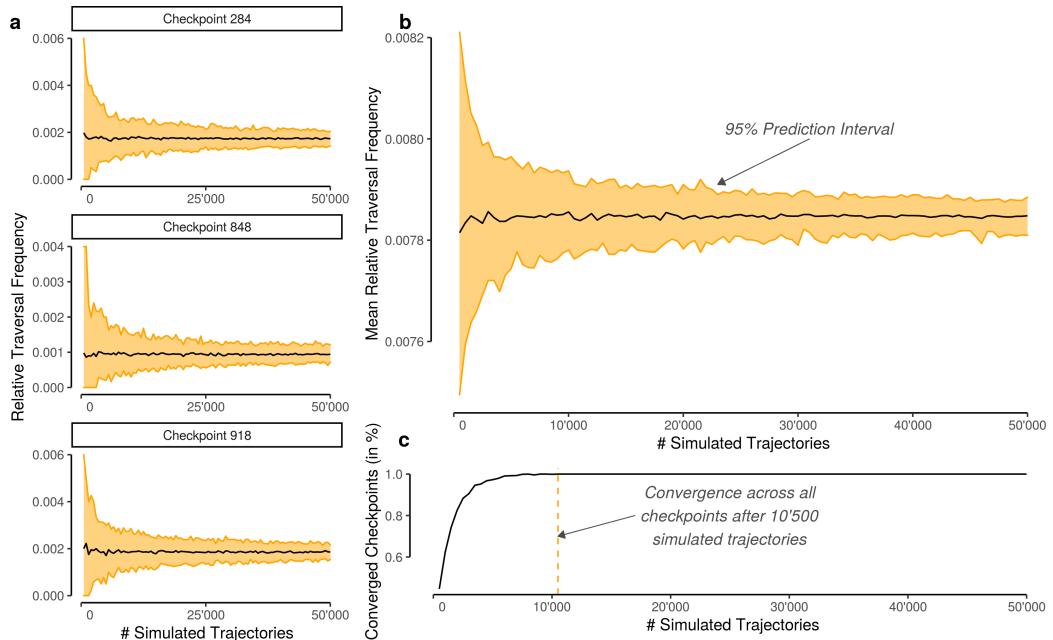


Figure 4: Relative traversal frequency through 1'000 checkpoints (5 km x 5 km) distributed randomly across the study area. The relative traversal frequency is plotted against the number of simulated individuals to visualize how quickly the metric converges to a steady state. (a) Replicated (100 times) relative traversal frequencies across three randomly chosen checkpoints as well as the corresponding 95% prediction interval (PI). (b) Averaged relative traversal frequency across all checkpoints and replicates including a 95% PI. (c) Width of the PI in relation to the number of simulated dispersers.

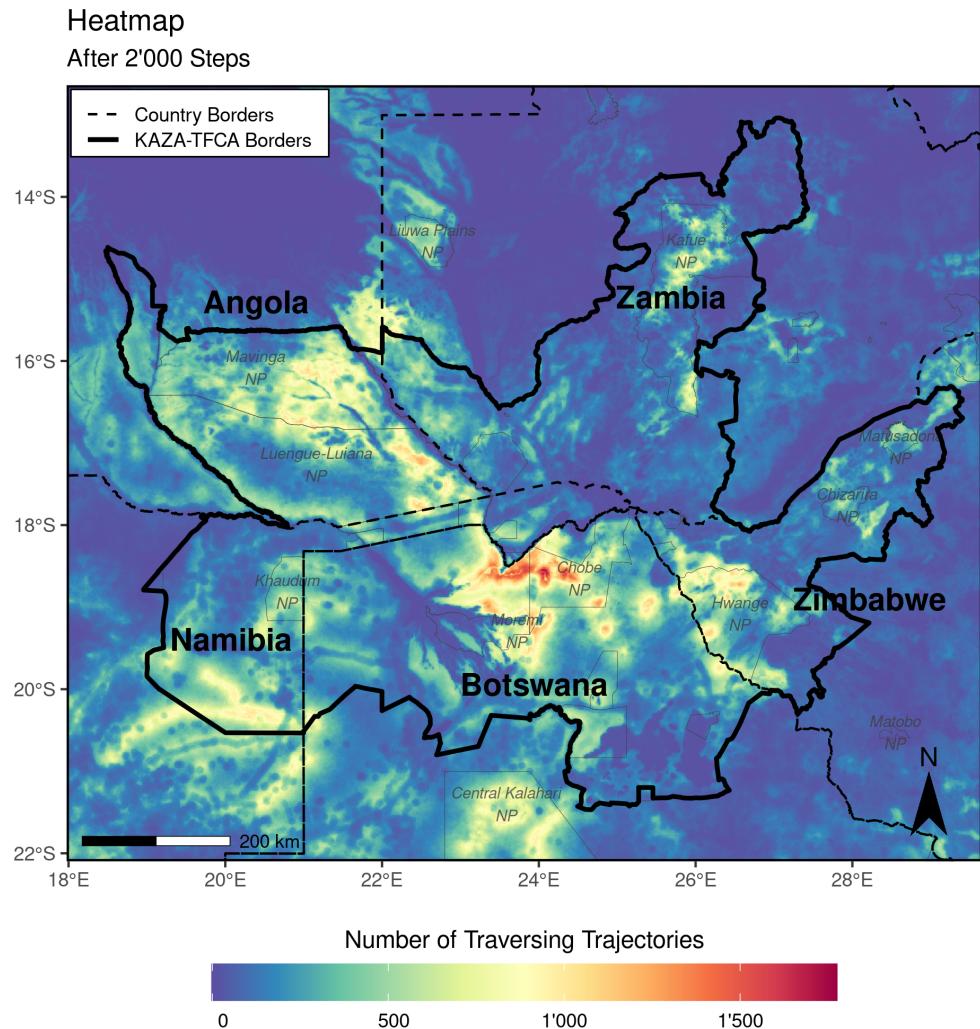


Figure 5: Heatmap showing traversal frequencies of 80'000 simulated dispersers moving 2'000 steps across the KAZA-TFCA. Simulations were based on an integrated step selection model that we fitted to the movement data of dispersing African wild dogs. To generate the heatmap, we rasterized and tallied all simulated trajectories. Consequently, the map highlights areas that are frequently traversed by virtual dispersers. Additional heatmaps showing the traversal frequency when individuals move fewer than 2'000 steps are provided in Appendix S3. For spatial reference we plotted a few selected national parks (dark gray).

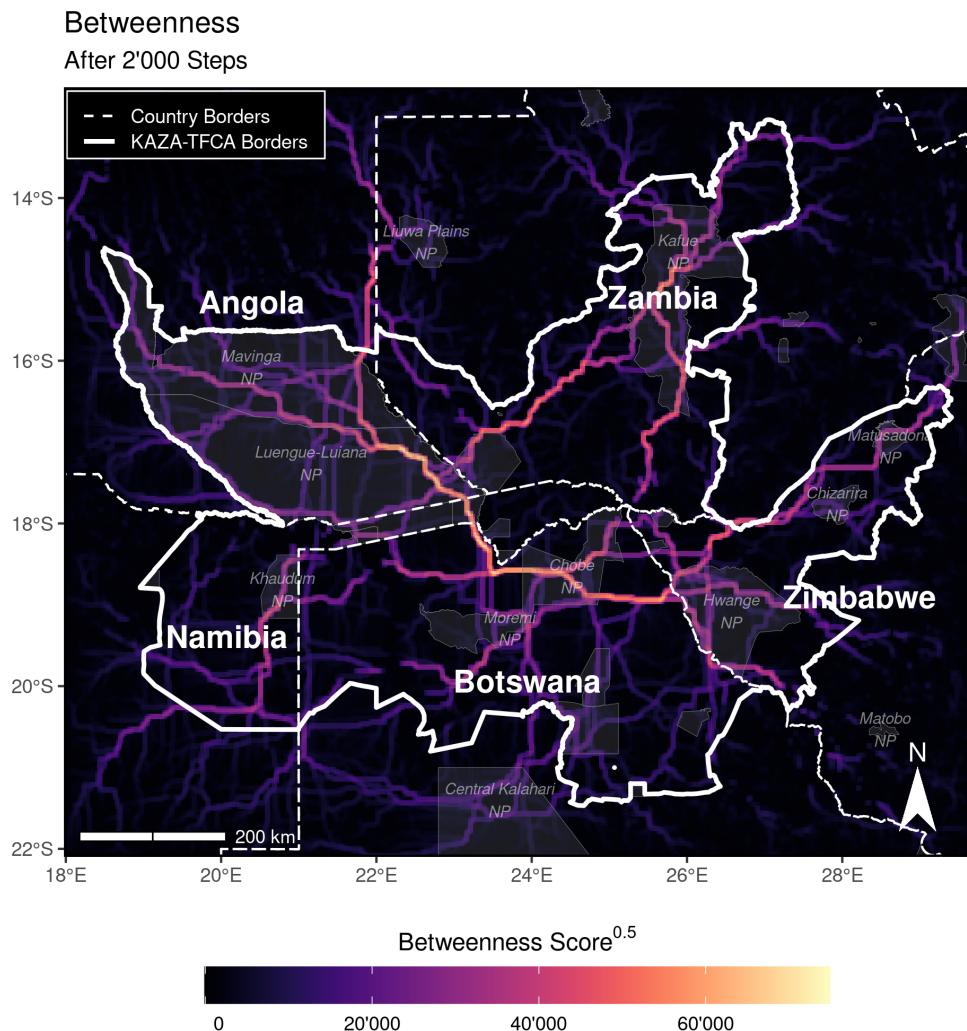


Figure 6: Map of betweenness scores, highlighting distinct dispersal corridors and potential bottlenecks across the extent of the KAZA-TFCA. A high betweenness score indicates that the respective area is exceptionally important for connecting different regions in the study area. In this sense the metric can be used to pinpoint discrete movement corridors (Bastille-Rousseau et al., 2018). Note that we square-rooted betweenness scores to improve visibility of corridors with low scores.

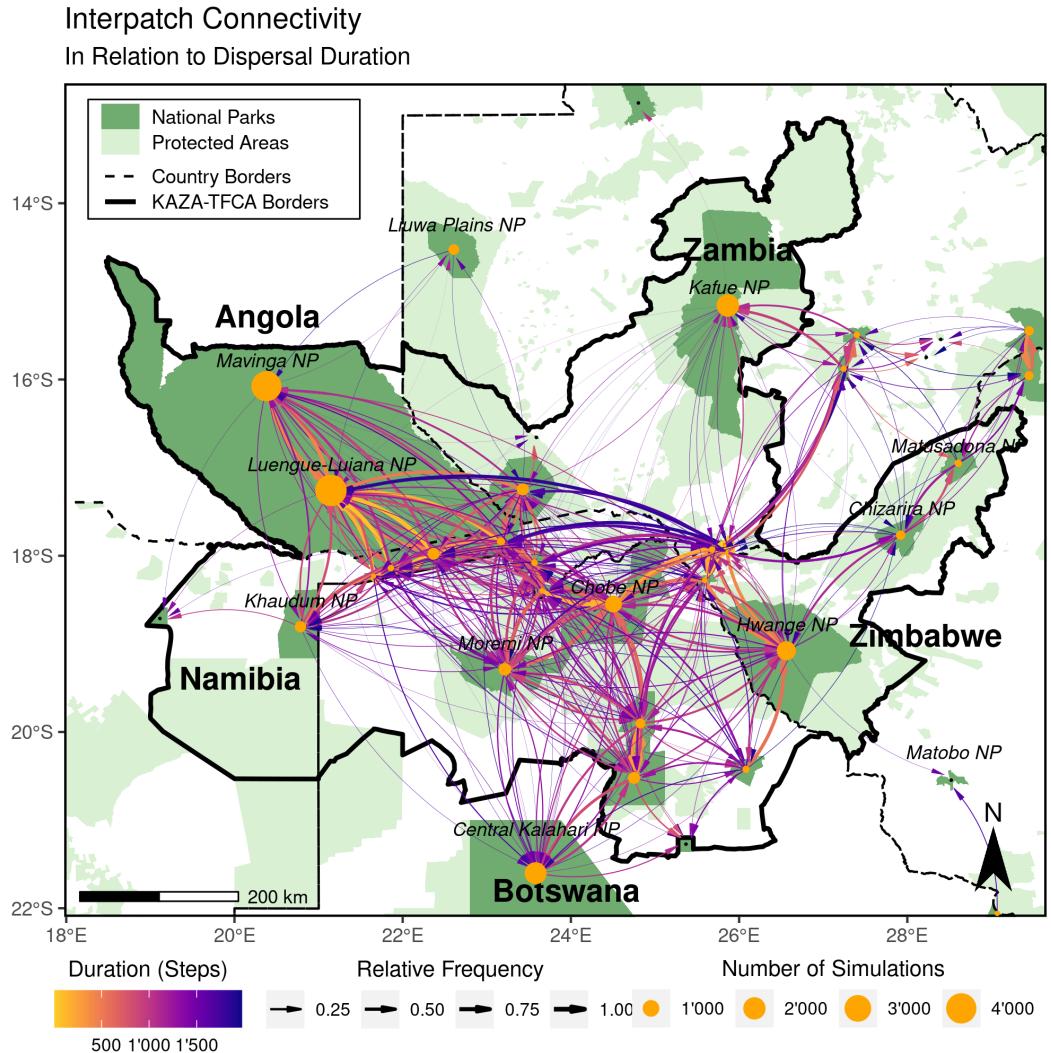


Figure 7: Map of inter-patch connectivity, highlighting connections between national parks (dark green). Yellow bubbles represent the center of the different national parks and are sized in relation to the number of simulated dispersers originating from each park. Black dots represent national parks that were smaller than 700 km^2 and therefore did not serve as source areas. Arrows between national parks illustrate between which national parks the simulated dispersers successfully moved and the color of each arrow shows the average number of steps (4-hourly movements) that were necessary to realize those connections. Additionally, the line thickness indicates the relative number of dispersers originating from a national park that realized those connections. Note that a similar network view could be adopted to investigate connectivity between other protected areas need not to be restricted to national parks.