

Appendix

Bound within Boundaries: How Well Do Protected Areas Match Movement Corridors of Their Most Mobile Protected Species?

David D. Hofmann^{1,§,*} Dominik M. Behr^{1,2,*} John W. McNutt²

Arpat Ozgul¹ Gabriele Cozzi^{1,2}

January 15, 2021

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

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

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

* Shared first authorship

Running Title: Connectivity across a Transfrontier Conservation Area.

Keywords: dispersal, habitat selection, integrated step selection function,
Kavango-Zambezi Transfrontier Conservation Area, landscape connectivity, least-cost
corridors, *Lycaon pictus*, permeability surface, protected areas, wildlife management

¹ A.1 Net Squared Displacement

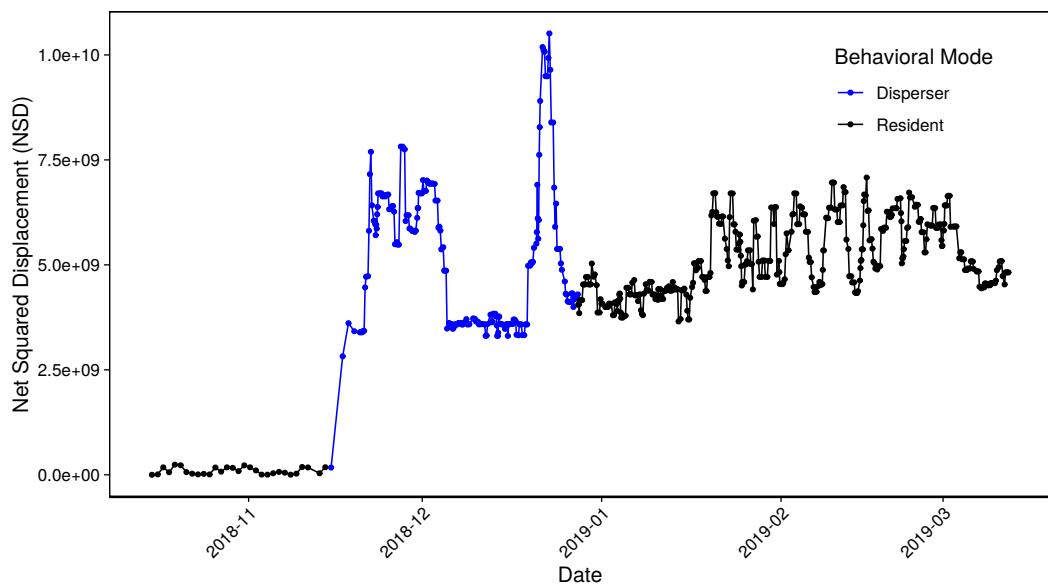


Figure S1: NSD displacement through time for one of our dispersers. The blue line indicates the period during which we classified the individual as dispersing.

2 A.2 GPS Data

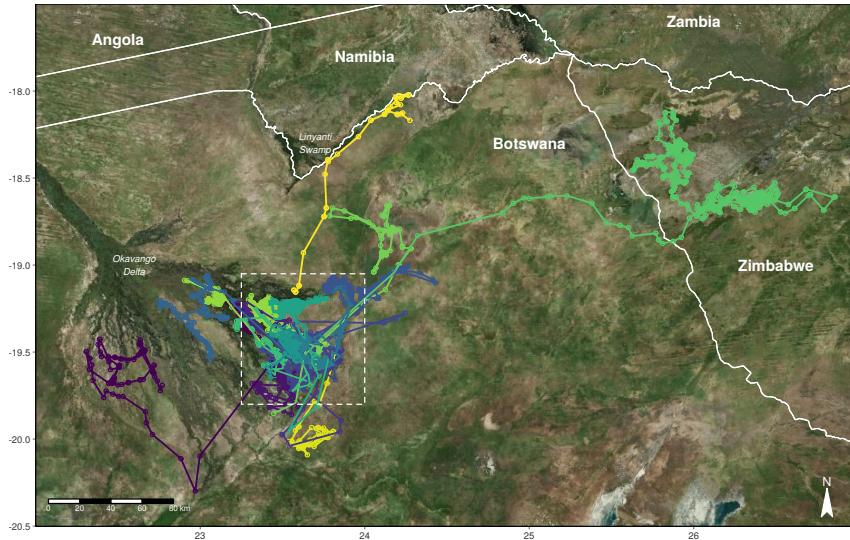


Figure S2: Illustration of all trajectories that we recorded. Each color represents a different dispersing coalition. All coalitions departed from the area which is indicated by the white dashed rectangle. The coalition dispersing towards the far east of the map covered over 360 km in under 10 days. Satellite background imagery was provided by Microsoft Bing.

Table S1: Summary statistics of all GPS relocations that have been recorded on dispersing coalitions. The coalition ID refers to the individual in the dispersal coalition that was equipped with a GPS collar. All coalitions consisted of same-sex individuals.

Coalition ID	Sex	Pack Affiliation	# Fixes Total	# Fixes During Dispersal	# Days Dispersing	Euclidean Dispersal Distance (in km)	Cumulative Dispersal Distance (in km)
Abel	M	MU	894	45	9	131	205
Amacuro	F	MB	954	583	137	23	1'090
Belgium	M	ZU	1'097	158	28	18	319
Dalwhinnie	F	PA	545	62	22	50	243
Denali	F	MN	1'096	173	33	11	528
Everest	M	MN	389	123	38	67	572
Kalahari	F	HT	1'753	467	130	20	1'963
Karisimbi	M	MN	438	141	34	45	251
Liuwa	F	AP	946	92	19	144	451
Lupe	M	KW	2'209	396	34	8	436
MadameChing	F	AP	776	729	136	263	1'560
Mirage	M	HT	814	182	36	7	435
Odzala	M	AP	1'410	205	42	53	412
Scorpion	M	KB	2'676	393	34	4	471
Stetson	M	MT	384	383	33	3	481
Taryn	F	AP	896	37	9	10	130
<i>Mean</i>		-	1'080	261	48	54	597
<i>(SD)</i>		-	(649)	(207)	(44)	(71)	(508)

³ **A.3 Spatial Covariates**

⁴ To investigate habitat preferences of dispersing wild dogs, we used a set of geo-referenced
⁵ covariates that we aggregated in the categories *land cover*, *protection status*, and *human*
⁶ *influence*. We did not include any terrain features due to the absence of noteworthy ele-
⁷ vational gradients in our study area. For each covariate, we prepared spatial raster layers
⁸ from freely available online services or from remotely sensed satellite imagery. To ensure a
⁹ consistent resolution (i.e. cell-size or grain) across covariates, we coarsened or interpolated
¹⁰ all layers to match a resolution of 250m x 250m. We performed processing and manipulation
¹¹ of data as well as all spatial and statistical analyses using R, version 3.6.1 (R Core Team,
¹² 2019).

¹³ **A.3.1 Land Cover**

¹⁴ **A.3.1.1 Water**

¹⁵ The covariate water included rivers, wetlands, and swamps. Because the inundation extent
¹⁶ of the flood in the Okavango Delta is highly variable within and between years, we created
¹⁷ dynamic “flood maps” that were updated every 8th day following a remote sensing algorithm
¹⁸ developed by the Okavango Research Institute (ORI; Wolski et al., 2017). To implement
¹⁹ the algorithm, we defined two sets of polygons located in the region of the Okavango Delta
²⁰ (Figure S3). The first set consisted of areas known to be permanent dryland, whereas the
²¹ second set consisted of permanent waters. Since we were unable to retrieve the original
²² polygons used in Wolski et al. (2017), we geo-referenced and digitized the polygons reported
²³ in their publication. After recreating the polygons, we used the R-package *getSpatialData*
²⁴ (Schwalb-Willmann, 2018) to download and pre-process all relatively cloud-free MODIS
²⁵ Terra images (MCD43A4; Schaaf and Wang, 2015) available for the period of our dispersal
²⁶ events. Assessment of cloud cover was based on visual inspection of MODIS images on
²⁷ ORI’s website (www.okavangodata.ub.bw/ori/monitoring/flood_maps). After download,
²⁸ we classified each MODIS image into a binary map of water (flood) and dryland using a
²⁹ threshold that was identified as follows. First, we extracted all reflectance values of MODIS
³⁰ Terra Band 7 within the water- and dryland-polygons. Second, we computed histograms
³¹ of water-reflectances and dryland-reflectances and empirically verified that reflectances of
³² the two groups were sufficiently distinct. More specifically, we checked if superimposing the
³³ histograms of water-reflectances and dryland-reflectances resulted in a bimodal histogram.
³⁴ This was said to be achieved if the 99th percentile of water-reflectances did not severely

35 exceed the 1st percentile of dryland-reflectances ($p_{0.99,water} - \frac{10}{255} < p_{0.01,dryland}$). Third, if
 36 bimodality was achieved, we calculated a threshold (t) using Equation S1:

$$t = \tilde{p}_{water} + 0.3 * (\tilde{p}_{dryland} - \tilde{p}_{water}) \quad (\text{Equation S1})$$

37 where \tilde{p}_{water} and $\tilde{p}_{dryland}$ were the median reflectances of water and dryland, respectively.
 38 We then classified all pixels of MODIS Terra Band 7 with a value greater than t as dryland
 39 and all pixels with a value smaller than t as water.

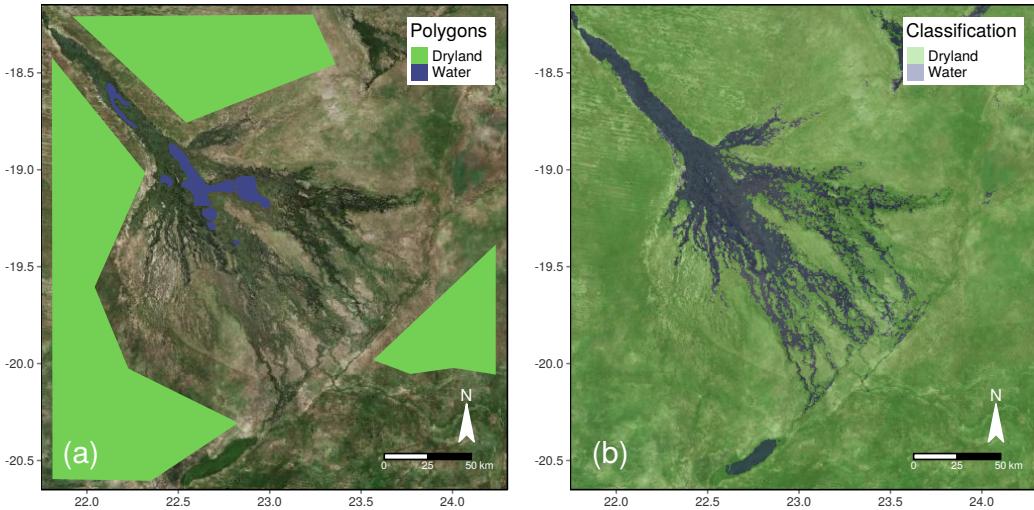


Figure S3: Images describing the flood mapping algorithm. (a) The colored polygons indicate permanent waters (blue) and permanent dryland (green). Below these polygons we extracted reflectance values of MODIS Terra Band 7 and used their respective medians to calculate a classification threshold t . (b) Example of a classified MODIS Terra Band 7 image after application of the threshold. The satellite image in the background was provided by Microsoft Bing.

40 Importantly, bimodality was not always achieved and in some cases no flood map could be
 41 calculated. In fact, it appears that non-bimodality caused the ORI algorithm to fail since
 42 the end of 2018, which is why no flood maps have been generated since then (ORI, personal
 43 comm.). We hypothesized that this was caused by the application of static water-polygons
 44 that did not cover permanent waters correctly anymore. Therefore, we revised the algorithm
 45 and allowed for a more dynamic polygonization of water. That is, for each MODIS image
 46 we calculated new water-polygons comprising areas that were covered by the flood in 99% of
 47 the flood maps from the previous five years. All of the necessary flood maps from previous
 48 years were kindly provided to us by ORI. Using this slightly amended approach, we were
 49 able to address some of the bimodality issues and to classify several additional flood maps
 50 for the period of our study. Because MODIS Terra Band 7 had a resolution of 500m x 500m,
 51 we interpolated all maps to 250m x 250m.

52 To validate and compare the performance of our own algorithm to the original ORI-
 53 algorithm, we randomly sampled 48 dates for which ORI prepared classified images. To
 54 make sure that months were equally represented in the sampled dates, we employed stratified
 55 sampling based on months (regardless of the year) and randomly sampled four maps for each
 56 month. For the sampled dates we downloaded and classified MODIS Terra Band 7 images
 57 and compared our classified images to those provided by ORI (Figure S4). For each pair of
 58 maps we created a difference map indicating false positives and false negatives and computed
 59 the relative number of wrongly classified pixels. We achieved an overall accuracy of 97%,
 60 which presumably is an underestimate of the true performance, as we introduced some errors
 61 when resampling the ORI-maps to our reference grid.

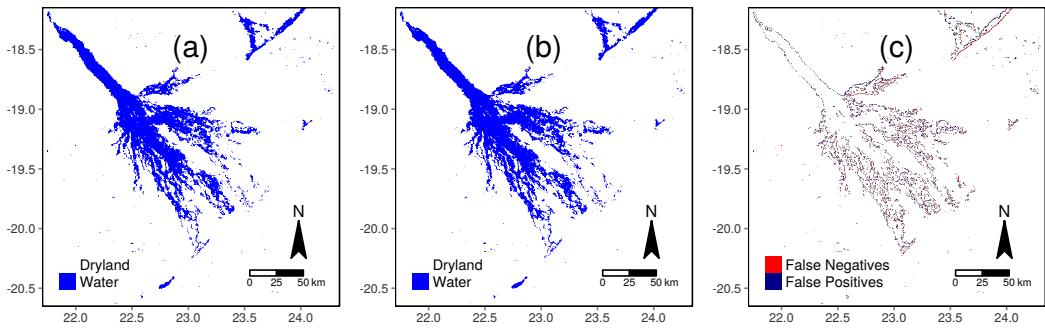


Figure S4: Validation procedure of our flood mapping algorithm. (a) Classified image that was provided to us by ORI. (b) Image for the same date but now classified using our own algorithm. (c) Difference image indicating false positives and false negatives in our own classification.

62 While we created dynamic flood maps for the Okavango Delta, we assumed the extent of
 63 all other water bodies (e.g. Chobe river, Zambezi river) to be static within and between
 64 years. This static representation was based on Globeland's land cover dataset (Chen et al.,
 65 2015), from which we only retained the categories *wetland* and *water bodies* and collectively
 66 reclassified them to *water*. Globeland had an original resolution of 30m x 30m, so we
 67 coarsened the layer to 250m x 250m using the mode of each 250m x 250m cell. We further
 68 improved river representation by employing the rasterized MERIT Hydro dataset (Yamazaki
 69 et al., 2019) from which we added all rivers with a width of over 10m to our Globeland layer.
 70 We merged dynamic and static water maps into a large rasterstack, covering the entire study
 71 area. We also created a rasterstack rendering the covariate *distance to water* by calculating
 72 the Euclidean distance of each raster cell in the study area to the nearest source of water.

73 A.3.1.2 Dryland

74 We subdivided dryland into three layers as derived from the MODIS Terra Vegetation Con-
 75 tinuous Fields dataset (MOD44B; Dimiceli et al., 2015). The three layers depicted percent-

76 age cover of tree-vegetation (henceforth *trees*), non-tree-vegetation (henceforth *shrubs/grassland*),
77 and non-vegetated (henceforth *bare land*) and added up to 100% of dryland coverage. We
78 used our flood map that aligned with the creation date of these MODIS layers and defined
79 anything covered by water as 0% vegetated. The MODIS vegetation layers had a resolution
80 of 250m x 250m and no coarsening or interpolation was required.

81 A.3.2 Protection Status

82 We created a binary layer separating protected from unprotected land. We downloaded cor-
83 responding data on protection status in shapefile format from the Peace Parks Foundation
84 (www.peaceparks.org; Peace Parks Foundation, 2019). Protected areas included forest re-
85 serves, game reserves, wildlife management areas, and national parks. We classified anything
86 not covered by these categories as unprotected (e.g. communal pastoral land, private land).
87 We rasterized the two categories to the binary raster *protection status* (1 = protected, 0 =
88 unprotected) with a resolution of 250m x 250m.

89 A.3.3 Anthropogenic

90 We created a raster layer representing human influence by integrating information on (1)
91 human density, (2) farming, and (3) roads.

92 (1) We obtained spatial human density estimates through a publicly available 30m x
93 30m high-resolution population density dataset (www.dataforgood.fb.com; Facebook,
94 2019). We coarsened the layer to 250m x 250m by summing up human density values
95 within each 250m x 250m cell.

96 (2) We sourced spatial information on farms from the Globeland (Chen et al., 2015) and
97 Cropland (Xiong et al., 2017) land cover datasets from which we retained areas that
98 were classified as either *cultivated land* or *croplands*. Any other land cover class was
99 not pertinent to farming and therefore omitted. Because both layers had a resolution
100 of 30m x 30m we coarsened them to 250m x 250m by assigning a value of 1 to any
101 250m x 250m cell that covered farmland and a value 0 otherwise. Thus, the final layer
102 depicted presence (= 1) or absence (= 0) of farms within each 250m x 250m cell.

103 (3) We obtained geo-referenced data on roads from Open Street Map (Open Street Map,
104 2019), downloaded through Geofabrik (www.geofabrik.de). We only retained main
105 tarmac roads and omitted smaller roads (Table S2) as these are scarcely frequented
106 and do not represent an obstacle to wild dog movements (Abrahms et al., 2016). We

107 rasterized main tarmac roads to the binary raster *roads* (1 = roads, 0 = no roads)
 108 with 250m x 250m resolution. Finally, we created the covariate *distance to roads* by
 109 calculating the Euclidean distance of each raster cell in the study area to the nearest
 110 road.

Table S2: Description of road types, as sourced from Open Street Map's mapping guide (<https://wiki.openstreetmap.org/wiki/Key:highway>). Roads types that were considered for the purpose of this study are shaded in light gray.

Group	Subgroup	Description
Roads	motorway	A restricted access major divided highway, normally with 2 or more running lanes plus emergency hard shoulder. Equivalent to the Freeway, Autobahn, etc.
Roads	trunk	The most important roads in a country's system that aren't motorways. Need not necessarily be a divided highway.
Roads	primary	The next most important roads in a country's system. Often link larger towns.
Roads	secondary	The next most important roads in a country's system. Often link towns.
Roads	tertiary	The next most important roads in a country's system. Often link smaller towns and villages
Roads	unclassified	The least important thorough roads in a country's system, i.e. minor roads of a lower classification than tertiary, but which serve a purpose other than access to properties. Often link villages and hamlets.
Roads	residential	Roads which serve as an access to housing, without function of connecting settlements. Often lined with housing.
Roads	service	For access roads to, or within an industrial estate, camp site, business park, car park etc.
Link roads	motorway_link	The link roads (sliproads/ramps) leading to/from a motorway from/to a motorway or lower class highway. Normally with the same motorway restrictions.
Link roads	trunk_link	The link roads (sliproads/ramps) leading to/from a trunk road from/to a trunk road or lower class highway.
Link roads	primary_link	The link roads (sliproads/ramps) leading to/from a primary road from/to a primary road or lower class highway.
Link roads	secondary_link	The link roads (sliproads/ramps) leading to/from a secondary road from/to a secondary road or lower class highway.
Link roads	tertiary_link	The link roads (sliproads/ramps) leading to/from a tertiary road from/to a tertiary road or lower class highway.
Special road types	living_street	For living streets, which are residential streets where pedestrians have legal priority over cars, speeds are kept very low and where children are allowed to play on the street.
Special road types	pedestrian	For roads used mainly/exclusively for pedestrians in shopping and some residential areas which may allow access by motorised vehicles only for very limited periods of the day.
Special road types	track	Roads for mostly agricultural or forestry uses.
Special road types	bus-guideway	A busway where the vehicle is guided by the way (though not a railway) and is not suitable for other traffic.
Special road types	escape	For runaway truck ramps, runaway truck lanes, emergency escape ramps, or truck arrester beds. It enables vehicles with braking failure to safely stop.
Special road types	raceway	A course or track for racing
Special road types	road	A road/way/street/motorway/etc. of unknown type. It can stand for anything ranging from a footpath to a motorway.

111 Because layers (1), (2), and (3) depicted features that are typically spatially clustered and
 112 because not all dispersing coalitions moved within meaningful distance to each of these
 113 features, we totaled values from the layers describing *human density* (continuous), *farming*
 114 (binary), and *roads* (binary). This approach implied that *roads* and *farms* entered the final
 115 layer with a value of 1, whereas human density entered the final layer with a value ≥ 0
 116 and potentially unbound. To reduce the influence of outliers in human density estimates,
 117 totaled values were limited to a maximum of 50, which visually resulted in a good balance

118 between high and low anthropogenic influence and was therefore considered appropriate for
 119 our analysis. To render the fact that humans influence their surroundings beyond their
 120 presence, we followed Elliot et al. (2014) and applied to each raster-cell a 5km focal buffer
 121 within which we summed up and log-transformed human-influence values.

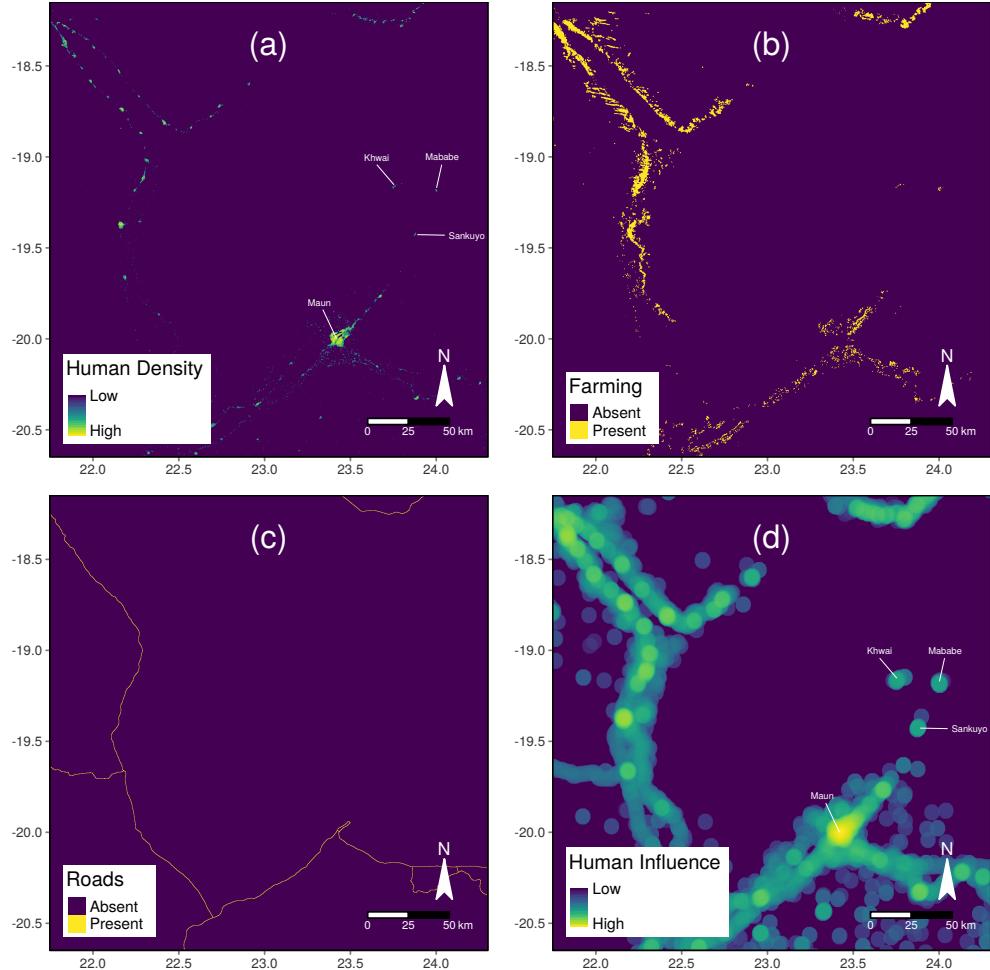


Figure S5: Sequence of figures that exemplifies how we combined the layers (a), (b), and (c) into a single layer for human influence (d). For better visibility we show the procedure only for the extent of the Okavango Delta. The layer in (a) is based on Facebook's high resolution human density dataset (www.dataforgood.fb.com; Facebook, 2019) and depicts the estimated number of humans living in each 250m x 250m raster-cell (coarsened from 30m x 30m). The layer in (b) is a binary layer and shows whether raster-cells are cover any sort of agricultural fields. Corresponding data was obtained through the Globeland and Cropland land cover datasets (Chen et al., 2015; Xiong et al., 2017). The layer in (c) shows the presence or absence of roads and is based on data from Open Street Map (Open Street Map, 2019). We merged the layers in (a), (b), and (c) by summing up their values, truncating the summed values to a maximum of 50. We then log-transforming the values and applied to each raster cell a focal buffer of 5km within which we totaled human influence values. The layer in (d) depicts the final human influence layer that entered our habitat selection model.

₁₂₂ A.4 Typical Flood Pulse

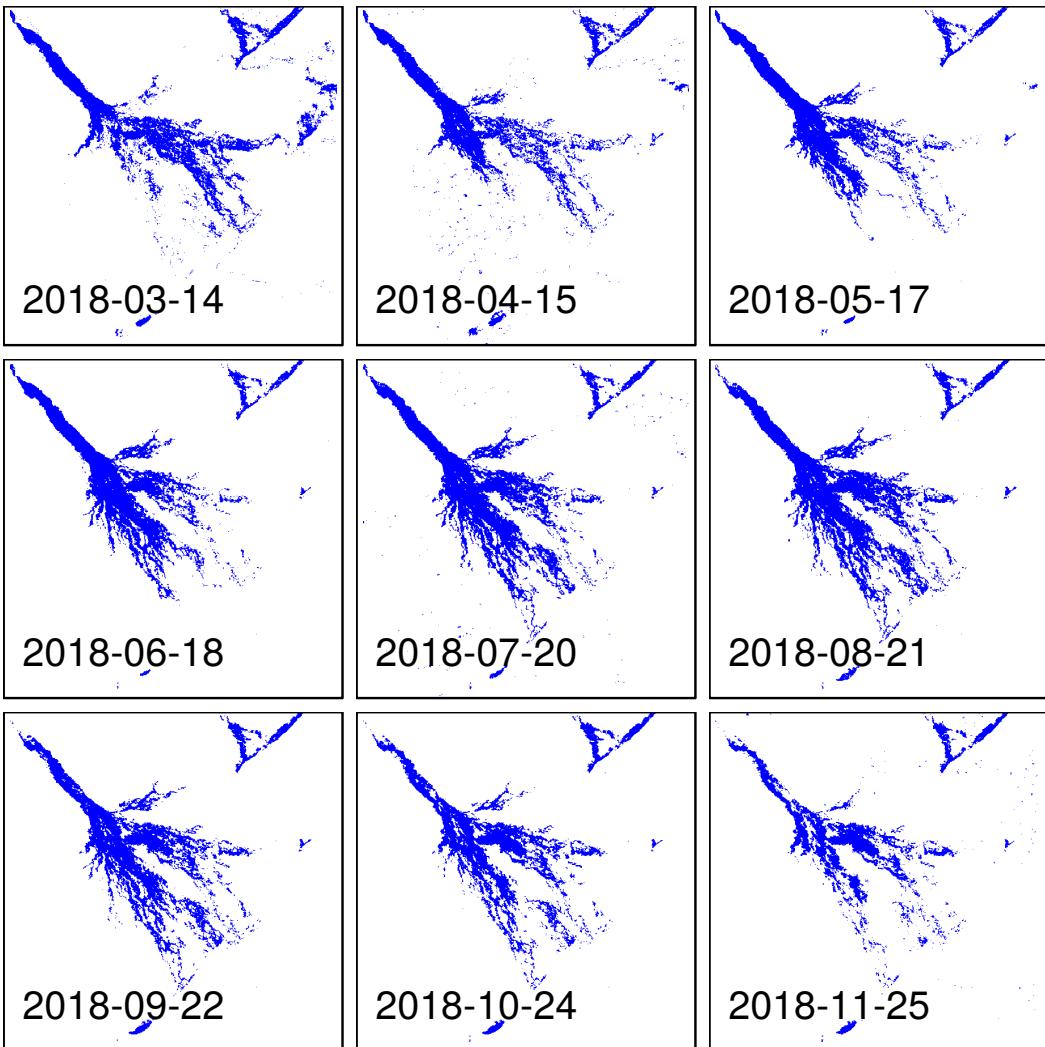


Figure S6: Sequence of flood maps showing a typical flood pulse throughout the year. The flood arrives from the north-western corner (so called “pan-handle”) of the Okavango Delta and slowly descends through the delta in south-eastern direction, where it nourishes several tributaries. The extent of the flood peaks around August or September and then slowly retracts. Between December and March the reflectance properties of water and dryland change, which is why often no accurate flood maps can be obtained for these months using remote sensing techniques (Wolski et al., 2017).

123 A.5 Integrated Step Selection Function

124 We used an integrated step selection function (iSSF; Avgar et al., 2016) to investigate
125 dispersers' selection or avoidance of spatial covariates. In the iSSF framework, covariates
126 experienced along realized steps are contrasted with covariates experienced along alternative
127 random steps that the animal could have taken but decided not to. A step in this framework
128 is defined as the connecting line between two consecutive GPS relocations (Turchin, 1998).
129 In contrast to regular SSFs, iSSFs require to include movement metrics as covariates in
130 the corresponding conditional logistic regression model. Their inclusion, in turn, allows
131 simultaneous inference on habitat and movement preferences, as well as to reduce potential
132 biases in estimated habitat preferences (Forester et al., 2009; Warton and Aarts, 2013; Avgar
133 et al., 2016).

134 To conduct iSSF analysis, we followed the recommendations described in Appendix S1
135 of the publication by Avgar et al. (2016). We prepared our GPS relocation data for iSSF-
136 analysis using the R-package *amt* (Signer et al., 2019) and coerced relocations recorded
137 during dispersal to steps that were regularly spaced four hours apart. Steps that were sepa-
138 rated by more than four hours (e.g. due to GPS failure) were omitted from further analysis
139 (allowing for a minor mismatch of up to 15 minutes). Each remaining step was paired
140 with 24 random steps, generated by sampling turning angles from a uniform distribution
141 $U(-\pi, \pi)$ and step lengths from a gamma distribution that was fitted using realized step
142 lengths (shape = 0.3677, scale = 6'302). Together, a realized and its 24 associated random
143 steps formed a stratum of 25 steps that received a unique identifier.

144 We extracted spatial covariates along realized and random steps (Table S3). For contin-
145 uous covariates, we calculated the average value, for categorical covariates the percentage
146 cover along the step. We further derived a binary variable indicating whether a step crossed
147 a road. We square-rooted extracted values to render a decreasing marginal impact of dis-
148 tance. We scaled covariates using a z-score transformation and screened for correlation using
149 Pearson's Correlation Coefficient. None of the covariates were overly correlated ($|r| > 0.6$;
150 Latham et al., 2011) and we retained all of them for modeling. Despite the covariates men-
151 tioned in Table S3, we included two movement metrics, namely the cosine of the turning
152 angle ($\cos(ta)$) and the logarithm of the step length ($\log(sl)$) in our regression model (Avgar
153 et al., 2016). The movement metric $\cos(ta)$ serves to describe the directionality of a step, as
154 it transforms the circular measure of $(-\pi$ to $\pi)$ into a linear measure (-1, 1). Thus, positive
155 values indicate forward movements, whereas negative values indicated backward movements
156 (Turchin, 1998). The movement metric $\log(sl)$, on the other hand, is as an indicator of the

¹⁵⁷ preferred step length. Since in our case steps were spaced by four hours, $\log(sl)$ can also be
¹⁵⁸ interpreted as movement rate.

Table S3: Overview of spatial covariates and their sources. We extracted covariates along realized and random steps. For continuous covariates we calculated average values along steps, for categorical covariates the percentage coverage along steps. We also prepared two covariates indicating the distance to water and distance to roads, respectively. We square-rooted the values for these two covariates to render a decreasing marginal impact of the effect of distance. Finally, we derived a binary indicator of whether a step crossed a road or not.

Category	Covariate	Description	Values	Source
Land Cover	Water	Percentage cover of water	0-100%	(1) (2) (3)
	Dryland*	Percentage cover of dryland	0-100%	(1) (2) (3)
	DistanceToWater	Average distance to nearest water source	$\geq 0\text{m}$	(1) (2) (3)
	Shrubs/Grassland	Average non-tree vegetation	0-100%	(4)
	Trees	Average tree-vegetation	0-100%	(4)
Protection Status	Bareland*	Average non-vegetated area	0-100%	(4)
	Protected	Percentage cover of protected area	0-100%	(5)
	Unprotected*	Percentage cover of unprotected area	0-100%	(5)
Anthropogenic	HumanInfluence	Average human influence	≥ 0	(1) (6) (7) (8)
	DistanceToRoads	Average distance to nearest road	$\geq 0\text{m}$	(1) (6) (7) (8)
	RoadCrossing	Binary; whether a step crossed a road	0, 1	(1) (6) (7) (8)

Sources: (1) Chen et al. (2015) (2) Schaaf and Wang (2015) (3) Yamazaki et al. (2019) (4) Dimiceli et al. (2015)
(5) Peace Parks Foundation (2019) (6) Facebook (2019) (7) Open Street Map (2019) (8) Xiong et al. (2017)

* Note: The covariates *Water* and *Dryland* added up to 100%, which is why only *Water* was included as explanatory variable in our models. The same applied for the group *Shrubs/Grassland*, *Trees*, and *Bareland*, where we omitted *Bareland* for modeling. Finally, from the group *Protected* and *Unprotected*, we only included *Protected* in our models.

¹⁵⁹ We then used the iSSF framework to parameterize a habitat selection model that further
¹⁶⁰ served to predict landscape permeability. This habitat selection model operated under
¹⁶¹ the assumption that dispersing wild dogs assigned a selection score $w(x)$ of the following
¹⁶² exponential form to each realized and random step:

$$w(x) = \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n) \quad (\text{Equation S2})$$

¹⁶³ That is, the selection score $w(x)$ of a step depended on its associated covariates (x_1, x_2, \dots, x_n) ,
¹⁶⁴ as well as on the animal's preferences for these covariates $(\beta_1, \beta_2, \dots, \beta_n)$. The probability
¹⁶⁵ that a step i was realized $P(Y_i = 1)$ was then contingent on the step's selection score, as
¹⁶⁶ well as on the selection scores of all alternative steps in the stratum:

$$P(Y_i = 1 | Y_1 + Y_2 + \dots + Y_i = 1) = \frac{w(x_i)}{w(x_1) + w(x_2) + \dots + w(x_i)} \quad (\text{Equation S3})$$

¹⁶⁷ Habitat and movement preferences of interest, i.e. the β 's, were then estimated by compar-
¹⁶⁸ ing realized (scored 1) and random (scored 0) steps in a conditional logistic regression model
¹⁶⁹ (Fortin et al., 2005). In this model, positive β -coefficients indicate selection of a covariate,
¹⁷⁰ negative β -coefficients avoidance of a covariate. To deal with multiple individuals, we ap-

plied mixed effects conditional logistic regression analysis following Muff et al. (2020). We implemented their method using the R-package *glmmTMB* (Brooks et al., 2017) and used dispersing coalition ID to model random intercepts and slopes.

We defined the movement metrics $\cos(ta)$ and $\log(sl)$ as core covariates and ran forward model selection based on Akaike's Information Criterion (AIC; Burnham and Anderson, 2002) for all other covariates. We ranked models according to AIC, assessed relative model weights, and identified the most parsimonious model. Due to convergence issues, we were unable to model interactions between covariates.

To validate the predictive power of the most parsimonious habitat selection model, we ran k-fold cross-validation for case-control studies as described in Fortin et al. (2009). Using 80% of randomly selected strata, we parameterized a habitat selection model and predicted selection scores $w(x)$ for all steps in the remaining 20% of strata. According to predicted selection scores we assigned ranks 1-25 within each stratum, with rank 1 indicating the highest selection score. We identified the realized step's rank in each stratum and tallied rank frequencies of realized steps across all strata. Finally, we carried out a Spearman-rank correlation analysis between ranks and associated frequencies and we recorded the correlation coefficient ($r_{s,realized}$). We repeated this procedure 100 times with replacement and computed the mean correlation coefficient ($\bar{r}_{s,realized}$), as well as its 95% confidence interval. For comparison, we also repeated the same procedure 100 times assuming completely randomized preferences. We implemented randomized preferences by omitting the realized step from each stratum and identifying the rank of a randomly chosen random step within each stratum (now only ranks 1-24). 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, the validation proved a significant prediction in case the confidence intervals of $\bar{r}_{s,realized}$ and $\bar{r}_{s,random}$ did not overlap.

¹⁹⁶ **A.6 Identification of Least-Cost Paths & Corridors**

¹⁹⁷ **A.6.1 Least-Cost Paths**

¹⁹⁸ We implemented factorial LCP analysis between source points using the R-package *gdistance*
¹⁹⁹ (Figure S.7; van Etten, 2017). The package translated the (unscaled) permeability surface
²⁰⁰ into a network of nodes to find shortest effective distances between source points based on
²⁰¹ probabilities of moving from cell to cell. In our case, the transition probability of moving
²⁰² between two adjacent cells depended on their averaged permeability. We allowed individuals
²⁰³ to move from each cell to the cell's eight surrounding neighbors (i.e. Moores neighborhood)
²⁰⁴ and applied a geographic correction to account for the fact that diagonal neighbors were
²⁰⁵ more remote than orthogonal neighbors. Because African wild dogs have been observed
²⁰⁶ to cover large dispersal distances (Davies-Mostert et al., 2012; Masenga et al., 2016; Cozzi
²⁰⁷ et al., 2020), we did not limit LCPs to a maximal effective cost. After computation, we
²⁰⁸ tallied overlapping LCPs and identified high-frequency routes.

²⁰⁹ **A.6.2 Least-Cost Corridors**

²¹⁰ We calculated factorial LCCs (Pinto and Keitt, 2009; Sawyer et al., 2011; Elliot et al., 2014),
²¹¹ again using the R-package *gdistance* (Figure S.7; van Etten, 2017). To identify LCCs, we first
²¹² computed for each source point a cumulative cost map, which indicated the total minimal
²¹³ costs required to get from the source point to any other location in the study area. We then
²¹⁴ obtained an LCC between two source points by adding up their cumulative cost maps and
²¹⁵ masking out all cell-values exceeding the lowest cell-value by more than 5% (Pinto and Keitt,
²¹⁶ 2009). We repeated this procedure for each possible unique pairwise combination of source
²¹⁷ points and thereby identified LCCs between all 68 selected source points. We normalized the
²¹⁸ resulting corridor-maps to range from zero to one and tallied them into a single connectivity
²¹⁹ map.

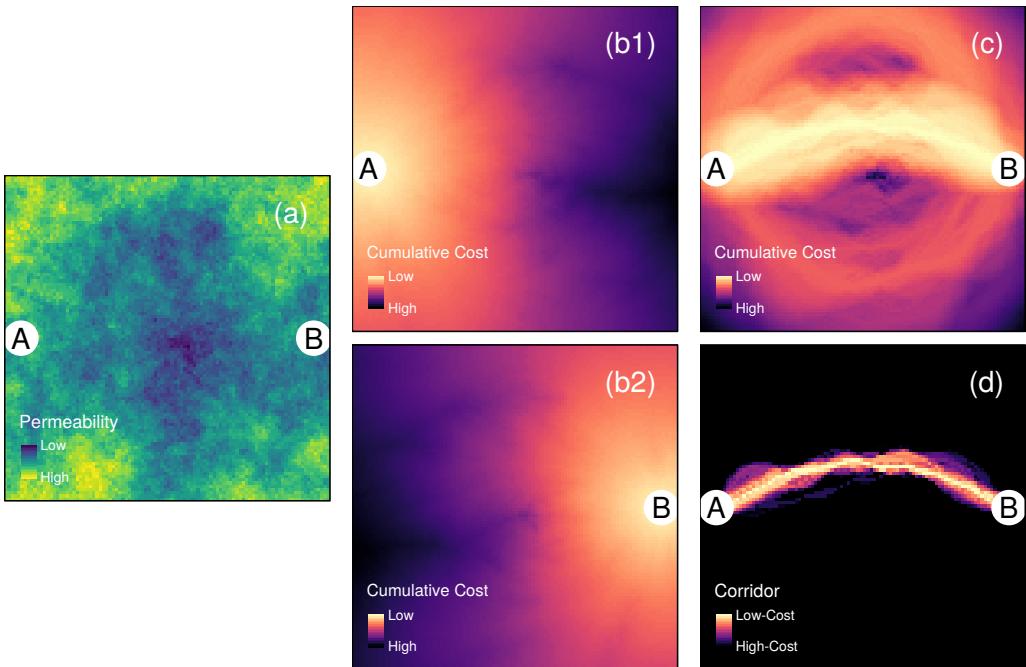


Figure S7: Images illustrating the process of identifying a least-cost corridor between source points A and B following Pinto and Keitt (2009). (a) Example of a permeability surface, which determines the costs of movement. (b1) Cumulative cost map for point A, depicting the total minimal costs necessary to get from point A to every other location. (b2) Cumulative cost map for point B, depicting the total minimal costs to get from point B to every other location. (c) Summed cost maps of points A and B. (d) Masked out corridor containing pixels that do not exceed the cheapest pixel by more than 5%.

220 A.7 Model Selection Results

Table S4: Results from the forward model selection procedure based on Akaike's Information Criterion (AIC; Burnham and Anderson, 2002) for the habitat selection model. The most parsimonious model outperformed all other models ($\Delta AIC > 2$) and received a weight of one.

Covariates	AIC	ΔAIC	Weight	LogLik
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{S}$	90068.15	0.00	1.00	-45017.08
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI}$	90071.84	3.69	0.00	-45020.92
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{S} + \text{DTR}$	90071.94	3.79	0.00	-45016.97
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{S} + \text{P}$	90071.94	3.79	0.00	-45016.97
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{S} + \text{RC}$	90073.46	5.30	0.00	-45015.73
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{DTR}$	90075.66	7.50	0.00	-45020.83
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{P}$	90075.66	7.51	0.00	-45020.83
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{S} + \text{DTR} + \text{P}$	90075.79	7.64	0.00	-45016.89
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{S}$	90076.71	8.56	0.00	-45023.36
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{RC}$	90076.84	8.69	0.00	-45019.42
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{S} + \text{DTR} + \text{RC}$	90077.20	9.05	0.00	-45015.60
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW}$	90080.08	11.92	0.00	-45027.04
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{HI} + \text{S} + \text{DTR} + \text{P} + \text{RC}$	90080.96	12.81	0.00	-45015.48
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{DTR}$	90082.95	14.79	0.00	-45026.47
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{P}$	90083.31	15.16	0.00	-45026.66
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{HI}$	90103.28	35.13	0.00	-45038.64
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T}$	90109.40	41.25	0.00	-45043.70
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{S}$	90110.35	42.20	0.00	-45042.17
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTR}$	90112.55	44.40	0.00	-45043.27
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{P}$	90113.11	44.96	0.00	-45043.56
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{RC}$	90113.60	45.45	0.00	-45041.80
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{DTW}$	90118.55	50.40	0.00	-45048.28
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{HI}$	90128.70	60.54	0.00	-45053.35
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{S}$	90132.22	64.06	0.00	-45055.11
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W}$	90134.85	66.69	0.00	-45058.42
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{DTR}$	90138.31	70.16	0.00	-45058.16
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{P}$	90138.50	70.35	0.00	-45058.25
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{RC}$	90139.30	71.15	0.00	-45056.65
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{S}$	90141.98	73.83	0.00	-45061.99
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{DTW}$	90225.64	157.49	0.00	-45103.82
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{T}$	90271.73	203.58	0.00	-45126.86
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{HI}$	90273.18	205.02	0.00	-45127.59
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{P}$	90285.24	217.08	0.00	-45133.62
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{DTR}$	90285.33	217.18	0.00	-45133.67
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{RC}$	-	-	-	-
$\cos(\text{ta}) + \text{sl} + \log(\text{sl}) + \text{W} + \text{T} + \text{DTW} + \text{RC}$	-	-	-	-

Note: W = Water, DTW = Distance To Water, S = Shrubs/Grassland, T = Trees, P = Protected, HI = Human Influence, RC = Road Crossing, DTR = Distance To Roads. The two models at the bottom failed to converge, which is why no AIC value could be obtained.

²²¹ **A.8 Habitat Selection Model: Random Effects**

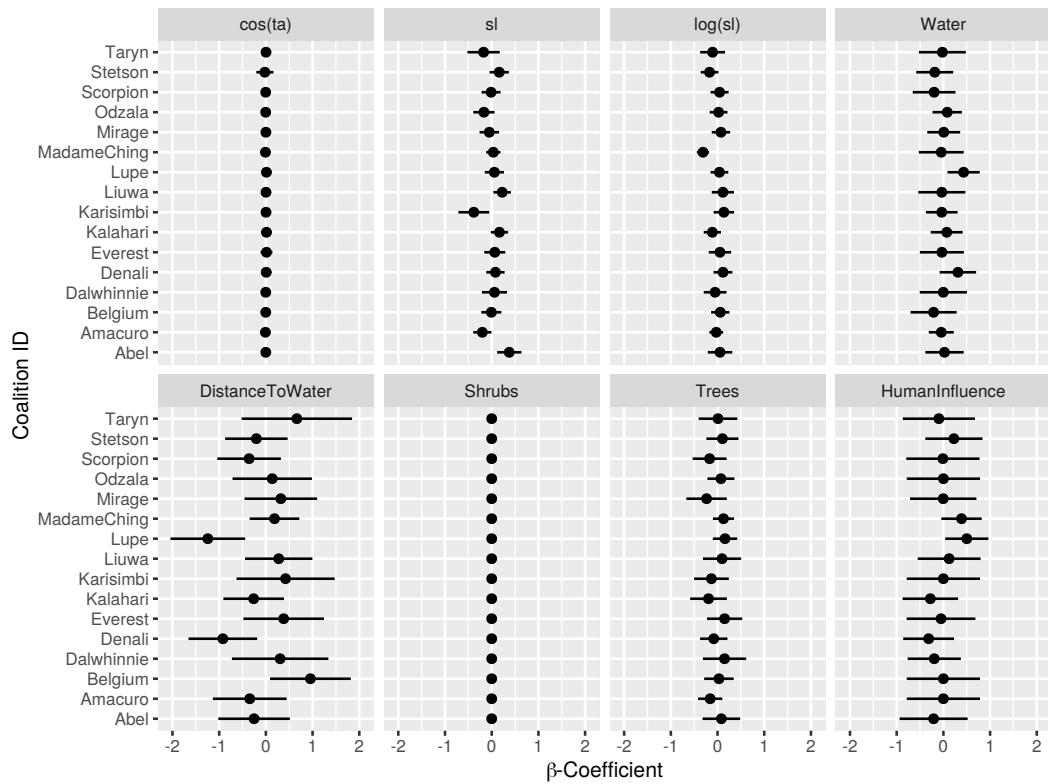


Figure S8: Plot of random effects showing variability across dispersal coalitions. The coalition ID refers to the individual in the dispersal coalition that was equipped with a GPS collar.

A.9 Alternative Source Points

Because we were interested in assessing the sensitivity of our connectivity maps with respect to the location of source points, we implemented two competing approaches to select source points. In a first run, we applied an omnidirectional *go-through* approach as proposed by Koen et al. (2014) and applied by Pitman et al. (2017). That is, we distributed 68 equally spaced source points along the map border and we calculated least-cost paths and least-cost corridors for all pairwise combinations of them. In contrast to other methods, this approach has been shown to reduce unwanted biases arising from the selection of unreasonable source points. Furthermore, such an approach enables the identification of dispersal routes that do not necessarily run between occupied patches of suitable habitat. In a second run, we followed the approach described in Elliot et al. (2014) and overlayed the study area with a regular grid of points spaced 100 km apart. From this regular grid we only considered those points that fell within protected areas $> 700 \text{ km}^2$, which conforms with home-range requirements of African wild dogs (Pomilia et al., 2015). Finally, we defined centroids as source points for those protected areas $> 700 \text{ km}^2$ that were not assigned any source points from the regular grid. Because wild dogs residing outside of protected areas are rare and unviable (Van der Meer et al., 2014), we consider such a selection of source points within protected areas to be reasonable. The second approach also resulted in 68 source points, implying 2'278 unique pairwise combinations and therefore 2'278 unique LCPs and LCCs per approach. Figure S9 illustrates the resulting least-cost paths and least-cost corridors and suggests that both methods produce qualitatively similar results. In fact, map values of Figure S9 (b1) and (b2) correlate with $r = 0.86$. Nevertheless, one can see that connectivity towards the map borders is more pronounced using the first approach and that the final connectivity map in (b1) appears to be a “stretched” version of the map in (b2). This is an expected result given that all source points were located at the map borders. Regardless of the method, we find that the major dispersal corridors run within the borders of the KAZA-TFCA and that mainly impenetrable landscape remains beyond its borders.

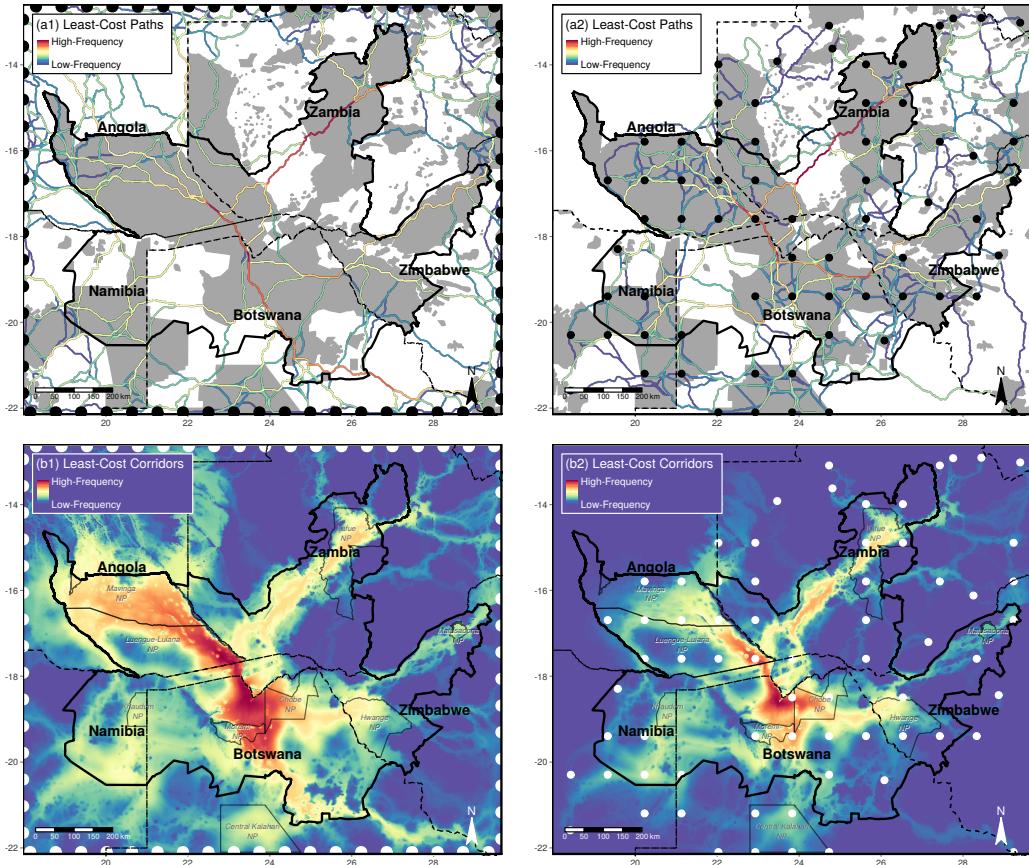


Figure S9: Comparison of least-cost paths (a1 and a2) and least-cost corridors (b1 and b2) using different source points. In subfigures (a1) and (b1) 68 source points were placed along the map border (black and white semicircles). In subfigures (a2) and (b2), on the other hand, source points (black and white circles) were placed in protected areas (gray) that were larger than 700 km^2 , which conforms to home range requirements for viable wild dog populations (Pomilia et al., 2015). Continuous black lines indicate the borders of the KAZA-TFCA, whereas dashed black lines delineate country-borders. For ease of spatial reference, we also labeled some national parks (NPs, in dark-grey).

249 References

- 250 Abrahms, B., Jordan, N. R., Golabek, K. A., McNutt, J. W., Wilson, A. M., and Brashares,
251 J. S. (2016). Lessons from Integrating Behaviour and Resource Selection: Activity-Specific
252 Responses of African Wild Dogs to Roads. *Animal Conservation*, 19(3):247–255.
- 253 Avgar, T., Potts, J. R., Lewis, M. A., and Boyce, M. S. (2016). Integrated Step Selection
254 Analysis: Bridging the Gap Between Resource Selection and Animal Movement. *Methods
in Ecology and Evolution*, 7(5):619–630.
- 255 Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen,
256 A., Skaug, H. J., Maechler, M., and Bolker, B. M. (2017). glmmTMB Balances Speed and
257 Flexibility among Packages for Zero-Inflated Generalized Linear Mixed Modeling. *The R
Journal*, 9(2):378–400.
- 258 Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A
Practical Information-Theoretic Approach*. Springer Science & Business Media, Ney York,
259 NY, USA.
- 260 Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., and Lu,
261 M. (2015). Global Land Cover Mapping at 30m Resolution: A POK-Based Operational
262 Approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103:7–27.
- 263 Cozzi, G., Behr, D. M., Webster, H. S., Claase, M., Bryce, C. M., Modise, B., Mcnutt, J. W.,
264 and Ozgul, A. (2020). African Wild Dog Dispersal and Implications for Management. *The
Journal of Wildlife Management*, 84(4):614–621.
- 265 Davies-Mostert, H. T., Kamler, J. F., Mills, M. G. L., Jackson, C. R., Rasmussen, G. S. A.,
266 Groom, R. J., and Macdonald, D. W. (2012). Long-Distance Transboundary Dispersal
267 of African Wild Dogs among Protected Areas in Southern Africa. *African Journal of
Ecology*, 50(4):500–506.
- 268 Dimiceli, C., Carroll, M., Sohlberg, R., Kim, D., Kelly, M., and Townshend, J. (2015).
269 MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid
270 v006. *NASA EOSDIS Land Processes DAAC*. Accessed 2019-11-12 from <https://doi.org/10.5067/MODIS/MOD44B.006>.
- 271 Elliot, N. B., Cushman, S. A., Macdonald, D. W., and Loveridge, A. J. (2014). The Devil
272 is in the Dispersers: Predictions of Landscape Connectivity Change with Demography.
273 51(5):1169–1178.
- 274 Facebook (2019). High Resolution Population Density Maps. Accessed 2019-11-12 from
275 <https://data.humdata.org/dataset/highresolutionpopulationdensitymaps>.
- 276 Forester, J. D., Im, H. K., and Rathouz, P. J. (2009). Accounting for Animal Movement
277 in Estimation of Resource Selection Functions: Sampling and Data Analysis. *Ecology*,
278 90(12):3554–3565.
- 279 Fortin, D., Beyer, H. L., Boyce, M. S., Smith, D. W., Duchesne, T., and Mao, J. S. (2005).
280 Wolves Influence Elk Movements: Behavior Shapes a Trophic Cascade in Yellowstone
281 National Park. *Ecology*, 86(5):1320–1330.
- 282 Fortin, D., Fortin, M.-E., Beyer, H. L., Duchesne, T., Courant, S., and Dancose, K. (2009).
283 Group-Size-Mediated Habitat Selection and Group Fusion–Fission Dynamics of Bison
284 under Predation Risk. *Ecology*, 90(9):2480–2490.
- 285 Koen, E. L., Bowman, J., Sadowski, C., and Walpole, A. A. (2014). Landscape Connectivity
286 for Wildlife: Development and Validation of Multispecies Linkage Maps. *Methods in
287 Ecology and Evolution*, 5(7):626–633.
- 288 Latham, A. D. M., Latham, M. C., Boyce, M. S., and Boutin, S. (2011). Movement Re-
289 sponses by Wolves to Industrial Linear Features and Their Effect on Woodland Caribou
290 in Northeastern Alberta. *Ecological Applications*, 21(8):2854–2865.

- 297 Masenga, E. H., Jackson, C. R., Mjingo, E. E., Jacobson, A., Riggio, J., Lyamuya, R. D.,
298 Fyumagwa, R. D., Borner, M., and Røskuft, E. (2016). Insights into Long-Distance
299 Dispersal by African Wild Dogs in East Africa. *African Journal of Ecology*, 54(1):95–98.
- 300 Muff, S., Signer, J., and Fieberg, J. (2020). Accounting for Individual-Specific Variation in
301 Habitat-Selection Studies: Efficient Estimation of Mixed-Effects Models Using Bayesian
302 or Frequentist Computation. *Journal of Animal Ecology*, 89(1):80–92.
- 303 Open Street Map (2019). Planet dump retrieved from <https://planet.osm.org>. Accessed
304 2019-11-12 from <https://www.openstreetmap.org>.
- 305 Peace Parks Foundation (2019). SADC Protected Areas. Accessed 2019-11-12 from <http://new-ppfmaps.opendata.arcgis.com/datasets/ppf-protected-areas-detailed?geometry=-13.87,-25.558,69.846,-11.001>.
- 306 Pinto, N. and Keitt, T. H. (2009). Beyond the Least-Cost Path: Evaluating Corridor
307 Redundancy Using a Graph-Theoretic Approach. *Landscape Ecology*, 24(2):253–266.
- 308 Pitman, R. T., Fattebert, J., Williams, S. T., Williams, K. S., Hill, R. A., Hunter, L.
309 T. B., Robinson, H., Power, J., Swanepoel, L., Slotow, R., and Balme, G. A. (2017).
310 Cats, Connectivity and Conservation: Incorporating Data Sets and Integrating Scales for
311 Wildlife Management. *Journal of Applied Ecology*, 54(6):1687–1698.
- 312 Pomilia, M. A., McNutt, J. W., and Jordan, N. R. (2015). Ecological Predictors of African
313 Wild Dog Ranging Patterns in Northern Botswana. *Journal of Mammalogy*, 96(6):1214–
314 1223.
- 315 R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foun-
316 dation for Statistical Computing, Vienna, Austria.
- 317 Sawyer, S. C., Epps, C. W., and Brashares, J. S. (2011). Placing Linkages among Fragmented
318 Habitats: Do Least-Cost Models Reflect How Animals Use Landscapes? *Journal of
319 Applied Ecology*, 48(3):668–678.
- 320 Schaaaf, C. and Wang, Z. (2015). MCD43A4 MODIS/Terra + Aqua BRDF/Albedo Nadir
321 BRDF Adjusted RefDaily L3 Global - 500m v006. *NASA EOSDIS Land Processes DAAC*.
322 Accessed 2019-11-12 from <https://doi.org/10.5067/MODIS/MCD43A4.006>.
- 323 Schwalb-Willmann, J. (2018). *getSpatialData: Get Different Kinds of Freely Available Spatial
324 Datasets*. R package version 0.0.4.
- 325 Signer, J., Fieberg, J., and Avgar, T. (2019). Animal Movement Tools (amt): R Package
326 for Managing Tracking Data and Conducting Habitat Selection Analyses. *Ecology and
327 Evolution*, 9:880–890.
- 328 Turchin, P. (1998). *Quantitative Analysis of Movement: Measuring and Modeling Population
329 Redistribution in Plants and Animals*. Sinauer Associates, Sunderland, MA, USA.
- 330 Van der Meer, E., Fritz, H., Blinston, P., and Rasmussen, G. S. (2014). Ecological Trap in
331 the Buffer Zone of a Protected Area: Effects of Indirect Anthropogenic Mortality on the
332 African Wild Dog (*Lycaon pictus*). *Oryx*, 48(2):285–293.
- 333 van Etten, J. (2017). *R Package gdistance: Distances and Routes on Geographical Grids*.
- 334 Warton, D. and Aarts, G. (2013). Advancing our Thinking in Presence-Only and Used-
335 Available Analysis. *Journal of Animal Ecology*, 82(6):1125–1134.
- 336 Wolski, P., Murray-Hudson, M., Thito, K., and Cassidy, L. (2017). Keeping it Simple:
337 Monitoring Flood Extent in Large Data-Poor Wetlands Using MODIS SWIR Data. *In-
338 ternational Journal of Applied Earth Observation and Geoinformation*, 57:224–234.

³⁴¹ Xiong, J., Thenkabail, P., Tilton, J., Gumma, M., Teluguntla, P., Oliphant, A., Congalton,
³⁴² R., Yadav, K., and Gorelick, N. (2017). Nominal 30m Cropland Extent Map of Continental
³⁴³ Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and
³⁴⁴ Landsat-8 Data on Google Earth Engine. *Remote Sensing*, 9(10):1065.

³⁴⁵ Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., and Pavelsky, T. M.
³⁴⁶ (2019). MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest
³⁴⁷ Topography Dataset. *Water Resources Research*, 55(6):5053–5073.