

Appendix

Individual-Based Dispersal Simulations to Study Dynamic Connectivity in Seasonal Landscapes

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

May 12, 2024

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

² Botswana Predator Conservation Program, Wild Entrust, Private Bag 13, Maun,
Botswana.

§ Corresponding author: david.hofmann2@uzh.ch

Running Title: Simulating Dispersal in Seasonal Landscapes

Keywords: African wild dog, Connectivity, Dispersal, Dynamism, Individual-based,
Lycaon pictus, Okavango Delta, Seasonality

Contents

A.1Pan Mapping	2
A.1.1Satellite Product Overview	2
A.1.2Evaluation of Landsat 8 and Sentinel 2	4
A.1.2.1 Training Polygons	4
A.1.2.2 Satellite Data Download	4
A.1.2.3 Land Cover Classifiers	6
A.1.2.4 Validation and Comparison	6
A.1.3Bulk Download	8
A.2Moon Illumination	11
A.3Light-Type	12
A.4Movement Model Results	13
A.5Number of Random Steps	17
A.6Model Interpretation	18
A.7Spearman's Rank Correlation in Relation to the Number of Random Steps	19

A.1 Pan Mapping

A.1.1 Satellite Product Overview

We dynamically mapped the spatial distribution of ephemeral water bodies that span only a few meters in diameter (Figure S1) using a custom remote sensing algorithm. These waterbodies are typically referred to as “pans” and mainly inundate during the rainy season, after which the cumulated water slowly evaporates. Previously, we developed an algorithm to detect large scale floodwaters across the extent of the Okavango Delta. This algorithm made use of MODIS MCD34A4 satellite imagery, which is a 8-day composite of daily updated MODIS satellite data. The benefit of MODIS’ high temporal resolution is that monthly composites are almost guaranteed to be cloud free, even during Botswana’s rainy season (mid October to mid May). However, due to its coarse spatial resolution of 250 meters, MODIS satellite data was unsuitable to detect finer-scaled waterbodies on satellite images. In an attempt to overcome this limitation, we evaluated alternative satellite products that provided better spatial resolution than MODIS, while retaining a sufficiently high temporal resolution to render seasonal patterns. Candidate products comprised Landsat 7, Landsat 8, and Sentinel 2 satellite imagery. Data associated with each of these satellites are freely accessible and provide spatial resolutions between 10 and 60 meters, depending on the respective bands. Furthermore, the satellites have revisit-times between 5 and 16 days, thus allowing to generate frequently updated composite images (Table S1).

Table S1: Overview of the spatial and temporal resolutions of the candidate satellite products.

Satellite	Availability	Spatial Resolution (m) [*]	Temporal Resolution (days)
Landsat 7	1999 – Present	15 – 60	16
Landsat 8	2013 – Present	15 – 100	16
Sentinel 2	2015 – Present	10 – 60	5 – 10 [†]

^{*} *Spatial resolutions of the same product can differ depending on the band.*

[†] *Sentinel 2’s revisit duration decreased to 5 days after March 2017.*

Our goal was to produce dynamically updated pan-maps that we could overlay with GPS data of dispersing African wild dogs. Since GPS data was collected between 2015 and 2022, this implied that satellite data needed to be available for the same period. At first glance, all satellite products appeared to meet this criteria. However, Landsat 7’s scan-line detector had been failing since May 31, 2003, introducing data gaps between adjacent tiles, thus rendering the product virtually unusable for our purposes (but see Storey et al., 2005). Sentinel 2 and Landsat 8, by contrast, appeared as viable alternatives, as both of them spanned the desired period and did not exhibit any device failures.



Figure S1: A pan as it can be found in Botswana right after the rainy season. This particular pan is comparably small at around 10 meters in diameter. Some of these pans dry out quickly, whereas others provide a source of water for the entire dry-season. Here, two spotted hyenas (*Crocuta crocata*) take a quick bath.

A.1.2 Evaluation of Landsat 8 and Sentinel 2

A.1.2.1 Training Polygons

With Landsat 8 and Sentinel 2 imagery remaining, we conducted a preliminary investigation to compare the two products and evaluate their suitability for our needs. Specifically, we investigated how well we could remote sense pans using either of the two products, utilizing supervised learning methods. For this, we prepared a set of training polygons, consisting of the land cover classes dryland, water, and wet-pans. Although our primary goal was to detect wet-pans, we included the two other classes to facilitate the categorization of reflectance values into distinct groups. Due to the large study area considered, *in situ* ground-truthing of the training polygons was impossible, and we instead opted for *on-screen selection* of training data. More specifically, we utilized Google Earth to digitize areas that were clearly identifiable as either dryland, water, or wet-pans. At the highest zoom level in Google Earth, these categories are visually easy to tell apart (Figure S2). To ensure a sharp distinction between wet-pans and dryland, we traced several dryland polygons in areas that were seasonally covered by water (Figure S2). Google Earth provides the date of any satellite map, and so we could assign a timestamp to each training-polygon. This was necessary to later match training polygons with Landsat or Sentinel data of the same date. Google Earth's ability to display imagery from different dates furthermore allowed us to generate a training dataset that comprised polygons that belonged to different classes depending on the season. For example, a dryland polygon obtained in the dry season could overlap with a wet-pan polygon obtained during the rainy season. Overall, we generated 268 polygons of varying size and shape (104 on dryland, 56 in water, and 108 in wet-pans) for 3 distinct dates (August 18, 2018, March 25, 2019, and July 24, 2021).

A.1.2.2 Satellite Data Download

We downloaded Landsat 8 imagery through Google Earth Engine using the `rggee` R-package (Aybar et al., 2024) and used the associated quality band to mask pixels classified as clouds or shadows. We also computed several normalized difference indices (NDs), as listed in Table S2. To download Sentinel 2 data, we used the `sen2r` R-package (Ranghetti et al., 2020) which provides a standardized interface to download and process Sentinel 2 imagery from the European Space Agency's data hub. Sentinel 2 data are available as either top-of-the-atmosphere (TOA, level 1C) or bottom-of-the-atmosphere (BOA, level 2A) reflectance values. In analyses where temporal trends are considered, the use of BOA reflectances is recommended, as differences in reflectance properties due to atmospheric conditions are ac-



Figure S2: Example of a *wet pan* and *dryland* training polygon digitized on Google Earth. The extent of water can easily be gauged at the maximum zoom level in Google Earth. While the dryland polygon in this case covered an area that is seasonally covered by water, we placed other dryland polygons in areas that were never inundated. However, to ensure reliable differentiation between wet and dry pans, we included several dryland polygons located in dry pans.

counted for (Gilabert et al., 1994; Vermote and Kotchenova, 2008; Chraibi et al., 2022). However, a large portion of the publicly available Sentinel 2 data has not yet been processed from level 1C to level 2A, requiring researchers to apply the corrections themselves. Fortunately, the European Space Agency provides the correction tool `sen2cor` for free (Main-Knorn et al., 2017), and this tool has been integrated into the R-package `sen2r` (Ranghetti et al., 2020). We therefore applied the necessary corrections after download for any product that hadn't already been corrected. Furthermore, we used Sentinel 2's "Scene Classification Map", to remove pixels classified as either clouds or shadows. Finally, we computed the same ND-indices as for Landsat 8 (Table S2).

Table S2: We computed normalized difference indices between certain bands, hoping they would improve the land-cover classifiers. Depending on the satellite, we used different bands to compute similar indices. The function to compute a normalized difference between bands b_1 and b_2 is given by $\frac{b_1 - b_2}{b_1 + b_2}$.

Index	Bands	
	Landsat 8	Sentinel 2
NDWI	B3, B5	B3, B8
NDMI	B5, B6	B8, B11
NDSI	B3, B6	B3, B11
BEST	B3, B7	B3, B12

A.1.2.3 Land Cover Classifiers

To train a land-cover classifier, we generated 3'000 random points within training polygons following a stratified equal random sampling scheme (Shetty et al., 2021). That is, we ensured that 1'000 random points were sampled per training category. Stratified random sampling was necessary to ensure that wet-pans, our category of interest, was sampled frequently enough, as it represented a minority class, only making up of (0.27%) the total training area. Stratified equal random sampling has been shown to provide good class-level accuracy for minority classes (Shetty et al., 2021), so we deemed this approach suitable for our purposes. We also kept track of the dates of the underlying polygons from which random points were generated. At each random point, we then extracted reflectance values of Sentinel 2 and Landsat 8 data that temporally aligned with the date assigned to the random point (Figure S3a). For instance, if a random point fell into a polygon that was digitized using a Google Earth image generated on Aug 18, 2018, we extracted values from the Sentinel 2 and Landsat 8 layers that were closest to that date. Finally, we parameterized a Random Forest (RF) classifier using the R-package `randomForest` (Cutler and Wiener, 2024) and a Classification and Regression Tree (CART) classifier using the R-package `rpart` (Therneau et al., 2024). In both cases, we included all bands, as well as derived ND-indices (Table S2) as explanatory covariates. We visualized the decision trees for the CART classifiers (Figure S3b) and variable importance for the RF models (Figure S3c) to investigate the importance of different bands or indices in separating the three classes. To compute variable importance we used the R-package `caret` (Kuhn, 2008).

A.1.2.4 Validation and Comparison

To validate the predictive power of the two classifiers across the two datasets, we employed 5-fold cross-validation. For this, we randomly split the data into 5 groups and repeatedly fitted both classification models using 80% of the data, to then predict land cover categories in the remaining 20%. We generated confusion matrices to contrast true and predicted labels and computed estimates of the classifier's specificity, sensitivity, and overall accuracy. The results from this validation show that both classifiers achieve $\geq 90\%$ accuracy across both satellite products (Figure S4). In fact, the RF classifier resulted in an overall accuracy of 99% on Sentinel 2 data.

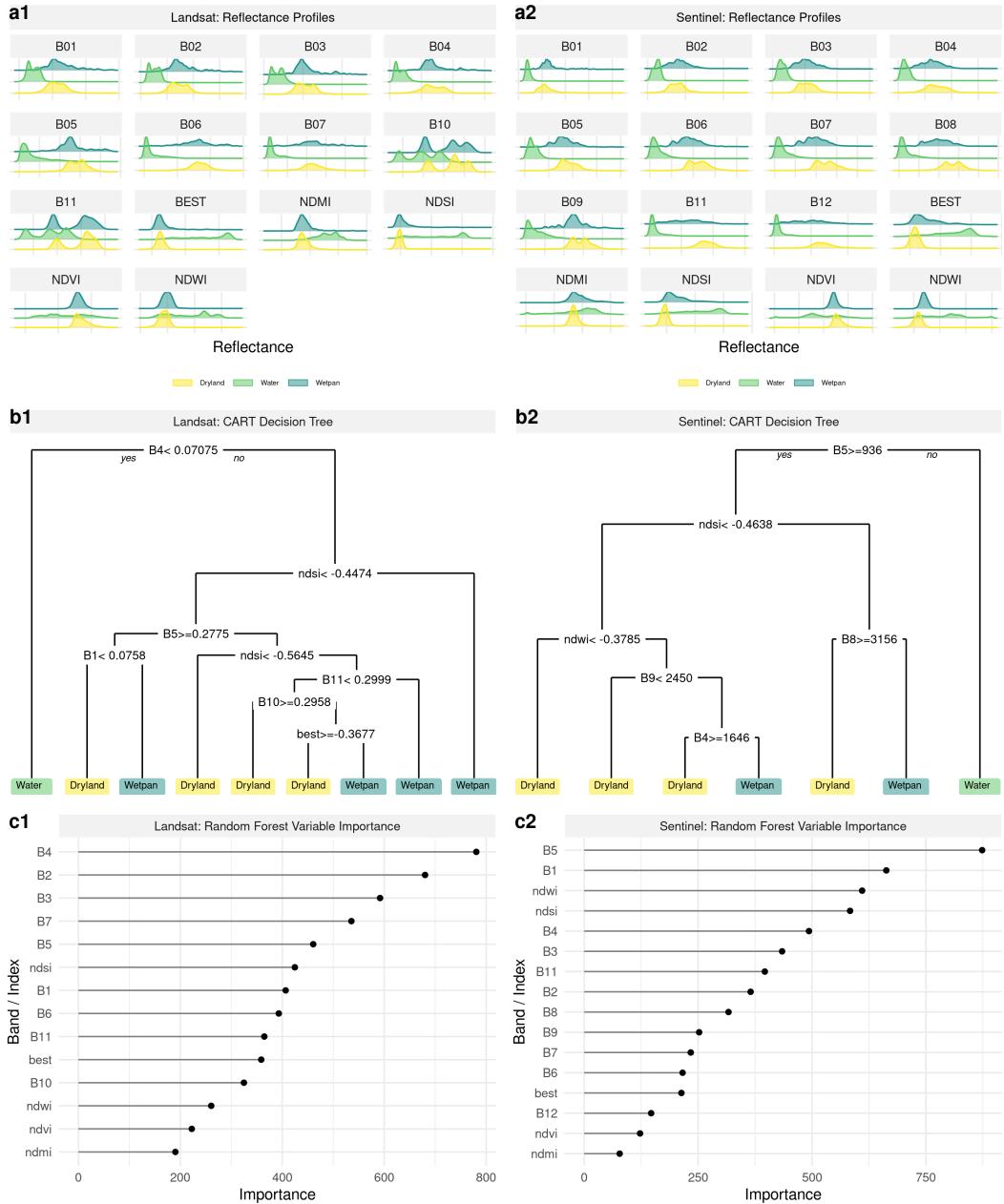


Figure S3: Reflectance properties of the Landsat 8 (a1) and Sentinel 2 (a2) bands and NDs at the extracted training points for each of the three categories (colored). Based on extracted reflectance values we parametrized Classification and Regression Tree models (b1 and b2) as well as Random Forest models (c1 and c2).

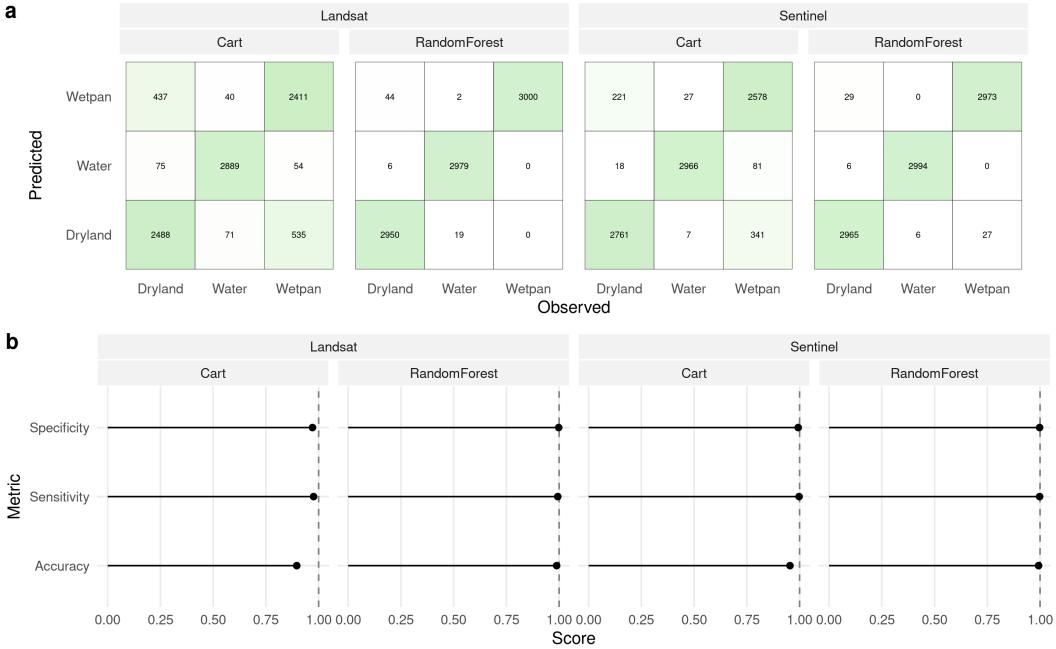


Figure S4: Confusion matrices (a) and derived performance metrics (b) for the CART and RF classifiers for both the Landsat 8 and Sentinel 2 datasets.

A.1.3 Bulk Download

Although both Landsat 8 and Sentinel 2 provided very good results, we considered Sentinel 2 data to be marginally superior, mainly due to Sentinel’s higher temporal and spatial resolution. A higher temporal resolution was key to compensate missing data from satellite images obtained on cloudy days and therefore pivotal in achieving cloud free monthly composites. We therefore decided to download Sentinel 2 data for the period during which we also collected GPS data of dispersing individuals. Instead of downloading all Sentinel 2 tiles overlapping with our study area and matching the study-period, we created a spatio-temporal moving window (Figure S5). This window was updated each month and comprised all GPS locations collected during that month, buffered by a 100 km radius. We then identified Sentinel 2 tiles intersecting with each of the monthly updated moving windows (Figure S6). This resulted in a list of 2,226 tiles that needed to be downloaded, 1,373 of which were already corrected to BOA reflectances, while the remaining 853 tiles still needed to be corrected. The download followed the same procedure and included the same corrections (from TOA to BOA, and cloud masking) as outlined in Appendix A.1.2.2 using the `sen2r` package in R (Ranghetti et al., 2020). Upon completion of the download and pre-processing, we applied the trained RF classifier to obtain binary “pan-maps”, showing the spatial distribution of ephemeral water bodies. To generate monthly composites, we merged all pan-maps falling into the same month and retained pans if they were detected in at least 50% of the overlap-

ping tiles. Finally, we generated “distance-to-pan” maps using the `distance` function from the `terra` package in R (Hijmans et al., 2024).

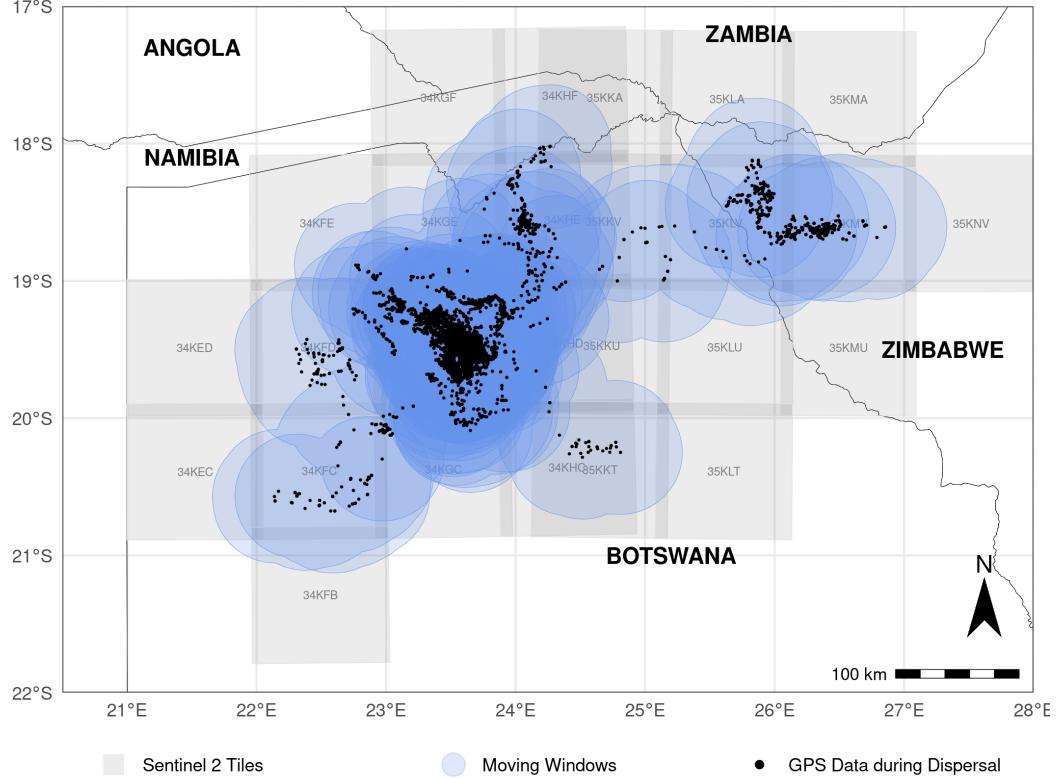


Figure S5: To reduce the number of Sentinel 2 tiles that we needed to download, we generated a monthly updated spatio-temporal moving window that comprised all GPS data of dispersing wild dogs (black dots) during the respective month, buffered by a 100 km radius. Based on the so created moving windows (in blue), we identified all overlapping Sentinel 2 tiles (tiles in gray) that we needed to download each month, which resulted in a total of 2,226 tiles.



Figure S6: Moving windows and associated tiles faceted by months.

A.2 Moon Illumination

We used the R-package `moonlit` (Śmielak, 2023) to obtain estimates of moon illumination at 5 minute intervals at the average location of dispersal GPS data (lon = of 23.5, lat = -19.0; Figure S7). The `moonlit` package is currently not on CRAN, but can be installed from GitHub (<https://github.com/msmielak/moonlit>). Besides providing information on the moon cycle, this package also allows estimating the amount of moonlight illumination on the ground. For instance, even during full-moon nights, the moon may only appear at a low angle in the sky, thus providing minimal illumination. Estimates from the `moonlit` package can therefore be viewed as biologically more relevant. To match the temporal resolution of our GPS data during dispersal, we calculated four hourly moonlight summaries. Specifically, we computed if an interval fell into nighttime and the average amount of moon-illumination during the four hours.

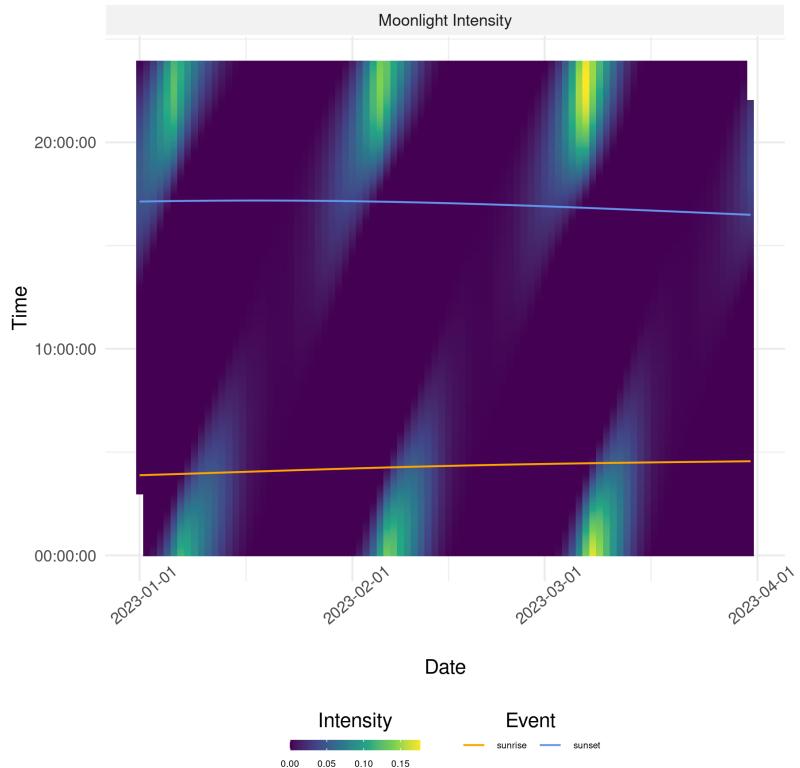


Figure S7: Moonlight intensity (illumination) as estimated from the `moonlit` package (Śmielak, 2023) in R calculated for a couple of example timestamps in 2023.

A.3 Light-Type

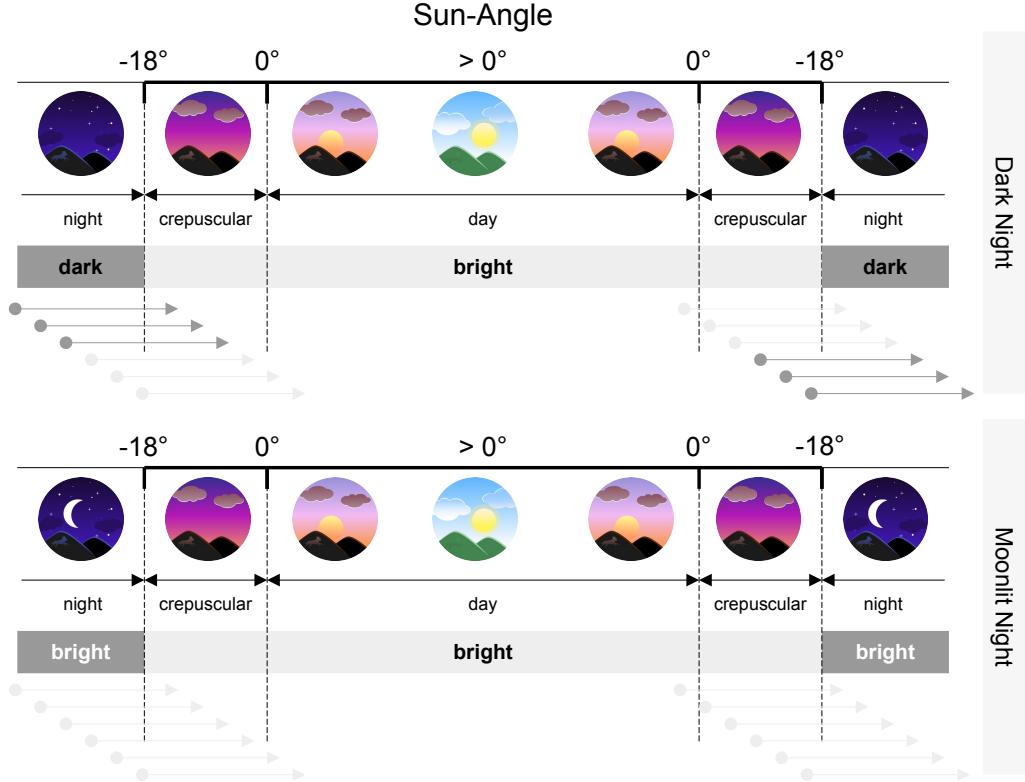


Figure S8: Schematic illustration of how we categorized steps into *bright* and *dark* steps. First, we used the `suncalc` and `moonlit` packages to obtain estimates of the sun-angle and moon illumination at 5 minute intervals for each date of interest. Whenever the sun angle was $\leq -18^\circ$, we deemed the respective period to be at night. A night was either bright (moon illumination > 0.02 of maximum illumination) or dark (moon illumination ≤ 0.02 of maximum illumination). Any other period was considered to provide enough illumination for wild dogs to move and therefore considered as bright. Since each step covered a time-span of approximately four hours, we defined a step as bright if at least 25% (i.e., one hour) of it occurred during a bright period. The gray arrows represent some example steps that are categorized as bright (light gray) or dark (dark gray).

A.4 Movement Model Results

Table S3: Estimates obtained from the **simple** (no interactions) integrated step-selection models fitted using **static** covariates. GPS Data was either pooled across seasons (labeled “all”) or split into dry and wet season. Only underlined covariates differed between the static and dynamic configurations.

Season	Covariate	Fixed Effects			Random Effects
		Coefficient	p-value	Significance	
All	sl	0.055	0.223		0.199
	log(sl)	0.060	0.056	*	0.110
	cos(ta)	0.113	0.000	***	0.059
	Humans	-0.289	0.015	**	0.346
	<u>Trees</u>	-0.284	0.007	***	0.363
	<u>Shrubs</u>	0.115	0.559		0.148
	<u>Water</u>	-0.488	0.012	**	0.229
	<u>DistanceToWater</u> ^{0.5}	-0.754	0.000	***	0.671
Dry	sl	0.046	0.377		0.170
	log(sl)	0.087	0.001	***	0.036
	cos(ta)	0.121	0.000	***	0.062
	Humans	-0.238	0.089	*	0.318
	<u>Trees</u>	-0.323	0.030	**	0.442
	<u>Shrubs</u>	-0.154	0.558		0.311
	<u>Water</u>	-0.825	0.001	***	0.063
	<u>DistanceToWater</u> ^{0.5}	-0.903	0.000	***	0.720
Wet	sl	0.008	0.908		0.244
	log(sl)	0.033	0.482		0.131
	cos(ta)	0.117	0.001	***	0.089
	Humans	-0.222	0.167		0.295
	<u>Trees</u>	-0.066	0.441		0.000
	<u>Shrubs</u>	0.557	0.076	*	0.128
	<u>Water</u>	0.162	0.591		0.002
	<u>DistanceToWater</u> ^{0.5}	-0.526	0.034	**	0.738

Table S4: Estimates obtained from the **simple** (no interactions) integrated step-selection models fitted using **dynamic** covariates. GPS Data was either pooled across seasons (labeled “all”) or split into dry and wet season. Only underlined covariates differed between the static and dynamic configurations.

Season	Covariate	Fixed Effects			Random Effects
		Coefficient	p-value	Significance	SD
All	sl	0.047	0.287		0.196
	log(sl)	0.066	0.037	**	0.112
	cost(ta)	0.114	0.000	***	0.058
	Humans	-0.255	0.029	**	0.333
	<u>Trees</u>	-0.230	0.006	***	0.333
	<u>Shrubs</u>	-0.001	0.993		0.234
	<u>Water</u>	-0.668	0.000	***	0.263
	<u>DistanceToWater</u> ^{0.5}	-0.237	0.093	*	0.544
Dry	sl	0.035	0.509		0.176
	log(sl)	0.085	0.000	***	0.000
	cos(ta)	0.120	0.000	***	0.059
	Humans	-0.189	0.159		0.292
	<u>Trees</u>	-0.371	0.003	***	0.412
	<u>Shrubs</u>	-0.157	0.263		0.321
	<u>Water</u>	-0.861	0.000	***	0.001
	<u>DistanceToWater</u> ^{0.5}	-0.284	0.135		0.596
Wet	sl	0.007	0.918		0.239
	log(sl)	0.049	0.318		0.141
	cos(ta)	0.123	0.000	***	0.089
	Humans	-0.204	0.205		0.306
	<u>Trees</u>	-0.077	0.295		0.159
	<u>Shrubs</u>	0.281	0.081	*	0.254
	<u>Water</u>	-0.320	0.082	*	0.199
	<u>DistanceToWater</u> ^{0.5}	-0.183	0.348		0.552

Table S5: Estimates obtained from the **full** (with interactions) integrated step-selection models fitted using **static** covariates. GPS Data was either pooled across seasons (labeled “all”) or split into dry and wet season. Only underlined covariates differed between the static and dynamic configurations.

Season	Covariate	Fixed Effects			Random Effects
		Coefficient	p-value	Significance	
All	sl	-0.021	0.602		0.142
	log(sl)	0.489	0.000	***	0.126
	cos(ta)	0.127	0.000	***	0.061
	sl:Dark	-0.276	0.000	***	-
	log(sl):Dark	-0.895	0.000	***	-
	sl: <u>Temperature</u>	-0.216	0.000	***	-
	log(sl): <u>Temperature</u>	-0.207	0.000	***	-
	Humans	-0.304	0.019	**	0.400
	<u>Trees</u>	-0.244	0.028	**	0.392
	Shrubs	0.148	0.456		0.096
	<u>Water</u>	-0.527	0.007	***	0.249
	<u>DistanceToWater</u> ^{0.5}	-0.865	0.000	***	0.752
	sl: <u>Trees</u>	0.013	0.741		-
	sl: <u>Shrubs</u>	0.059	0.654		-
	sl: <u>Water</u>	-0.200	0.116		-
Dry	cos(ta): <u>DistanceToWater</u> ^{0.5}	0.059	0.000	***	-
	sl	-0.075	0.161		0.151
	log(sl)	0.554	0.000	***	0.127
	cos(ta)	0.130	0.000	***	0.056
	sl:Dark	-0.104	0.097	*	-
	log(sl):Dark	-1.018	0.000	***	-
	sl: <u>Temperature</u>	-0.213	0.000	***	-
	log(sl): <u>Temperature</u>	-0.253	0.000	***	-
	Humans	-0.253	0.100		0.373
	<u>Trees</u>	-0.273	0.078	*	0.465
	<u>Shrubs</u>	-0.124	0.634		0.243
	<u>Water</u>	-0.888	0.000	***	0.186
	<u>DistanceToWater</u> ^{0.5}	-1.059	0.000	***	0.801
	sl: <u>Trees</u>	0.037	0.452		-
	sl: <u>Shrubs</u>	-0.046	0.777		-
	sl: <u>Water</u>	-0.343	0.029	**	-
	cos(ta): <u>DistanceToWater</u> ^{0.5}	0.040	0.057	*	-

Table S6: Estimates obtained from the **full** (no interactions) integrated step-selection models fitted using **dynamic** covariates. GPS Data was either pooled across seasons (labeled “all”) or split into dry and wet season. Only underlined covariates differed between the static and dynamic configurations.

Season	Covariate	Fixed Effects			Random Effects SD
		Coefficient	p-value	Significance	
All	sl	-0.188	0.000	***	0.208
	log(sl)	0.448	0.000	***	0.127
	cos(ta)	0.110	0.000	***	0.059
	sl:Dark	-0.278	0.000	***	-
	log(sl):Dark	-0.788	0.000	***	-
	sl: <u>Temperature</u>	-0.172	0.000	***	-
	log(sl): <u>Temperature</u>	-0.087	0.000	***	-
	Humans	-0.255	0.047	**	0.380
	<u>Trees</u>	-0.222	0.009	***	0.340
	<u>Shrubs</u>	-0.004	0.973		0.225
	<u>Water</u>	-0.775	0.000	***	0.175
	<u>DistanceToWater</u> ^{0.5}	-0.313	0.027	**	0.551
	sl: <u>Trees</u>	-0.025	0.167		-
	sl: <u>Shrubs</u>	0.038	0.550		-
Dry	sl:Water	-0.442	0.000	***	-
	cos(ta): <u>DistanceToWater</u> ^{0.5}	0.048	0.003	***	-
	sl	-0.281	0.000	***	0.170
	log(sl)	0.475	0.000	***	0.086
	cos(ta)	0.116	0.000	***	0.056
	sl:Dark	-0.098	0.111		-
	log(sl):Dark	-0.884	0.000	***	-
	sl: <u>Temperature</u>	-0.170	0.000	***	-
	log(sl): <u>Temperature</u>	-0.086	0.000	***	-
	Humans	-0.204	0.191		0.375
	<u>Trees</u>	-0.367	0.003	***	0.406
	<u>Shrubs</u>	-0.140	0.320		0.305
	<u>Water</u>	-0.940	0.000	***	0.001
Wet	<u>DistanceToWater</u> ^{0.5}	-0.353	0.056	*	0.570
	sl: <u>Trees</u>	-0.019	0.365		-
	sl: <u>Shrubs</u>	0.019	0.806		-
	sl:Water	-0.340	0.002	***	-
	cos(ta): <u>DistanceToWater</u> ^{0.5}	0.017	0.416		-
	sl	-0.112	0.173		0.248
	log(sl)	0.443	0.000	***	0.168
	cos(ta)	0.119	0.001	***	0.099
	sl:Dark	-0.886	0.000	***	-
	log(sl):Dark	-0.656	0.000	***	-
	sl: <u>Temperature</u>	-0.201	0.000	***	-
	log(sl): <u>Temperature</u>	-0.099	0.013	**	-
	Humans	-0.162	0.327		0.315
	<u>Trees</u>	-0.064	0.382		0.154
	<u>Shrubs</u>	0.295	0.057	*	0.208
	<u>Water</u>	-0.434	0.010	**	0.001
	<u>DistanceToWater</u> ^{0.5}	-0.221	0.276		0.596
	sl: <u>Trees</u>	-0.058	0.122		-
	sl: <u>Shrubs</u>	0.128	0.217		-
	sl:Water	-0.516	0.000	***	-
	cos(ta): <u>DistanceToWater</u> ^{0.5}	0.097	0.000	***	-

A.5 Number of Random Steps

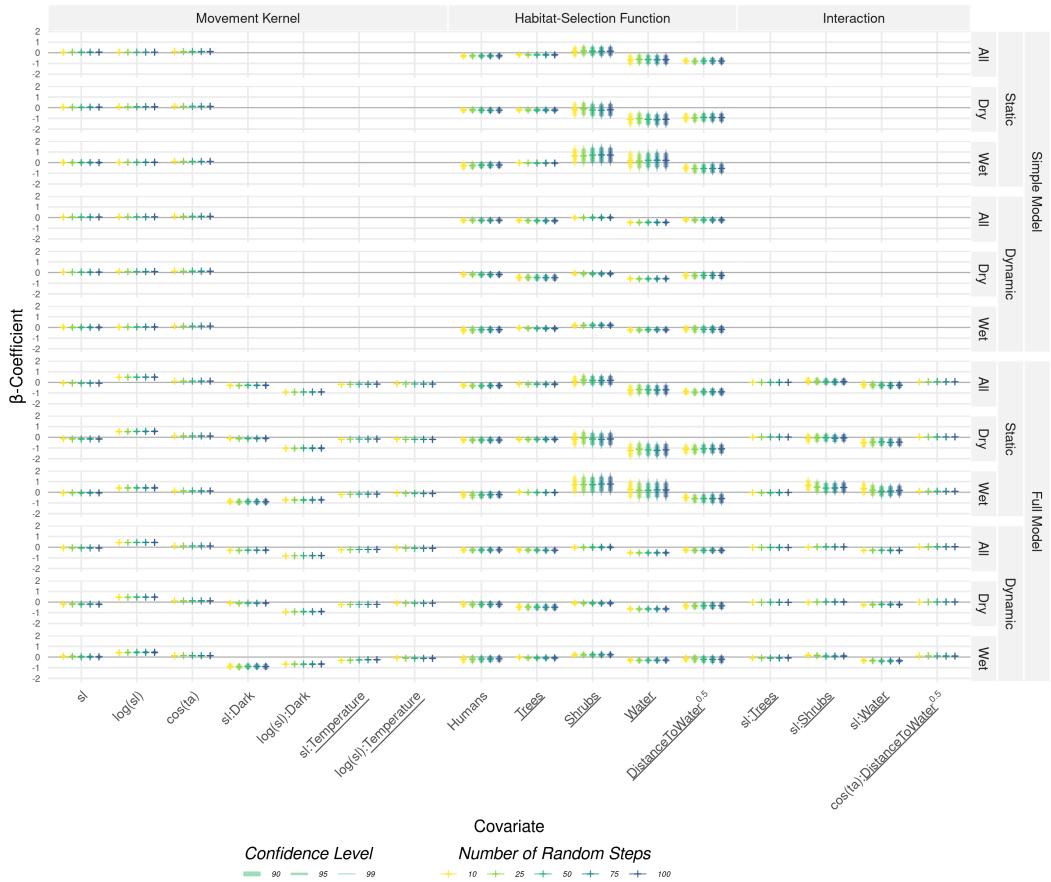


Figure S9: Results from the integrated step-selection models when considering different numbers of random steps. Results are shown for the simple and full model for both all configurations of fitting covariates (static vs dry) and model seasons (all vs. wet + dry).

A.6 Model Interpretation

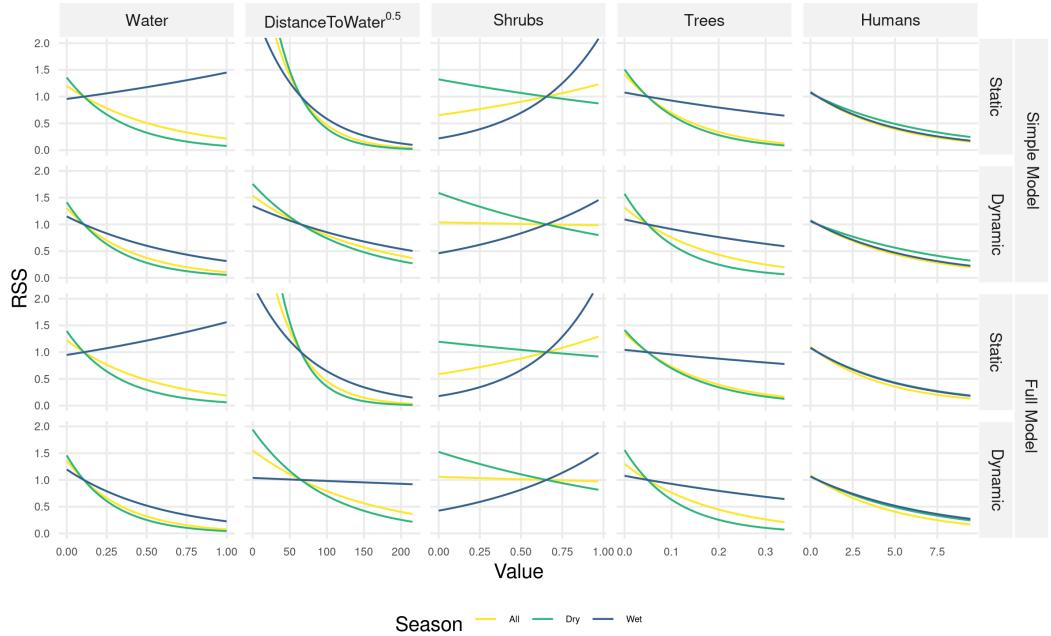


Figure S10: Relative selection scores (RSS) with respect to environmental covariates. RSS scores were computed by comparing a location x_1 with average conditions to a location x_2 with average conditions but the respective covariate varied between its minimum and maximum observed value. Confidence intervals overlapped substantially, hence we omitted them from the figure to simplify with the interpretation of the differences among configurations.

A.7 Spearman's Rank Correlation in Relation to the Number of Random Steps

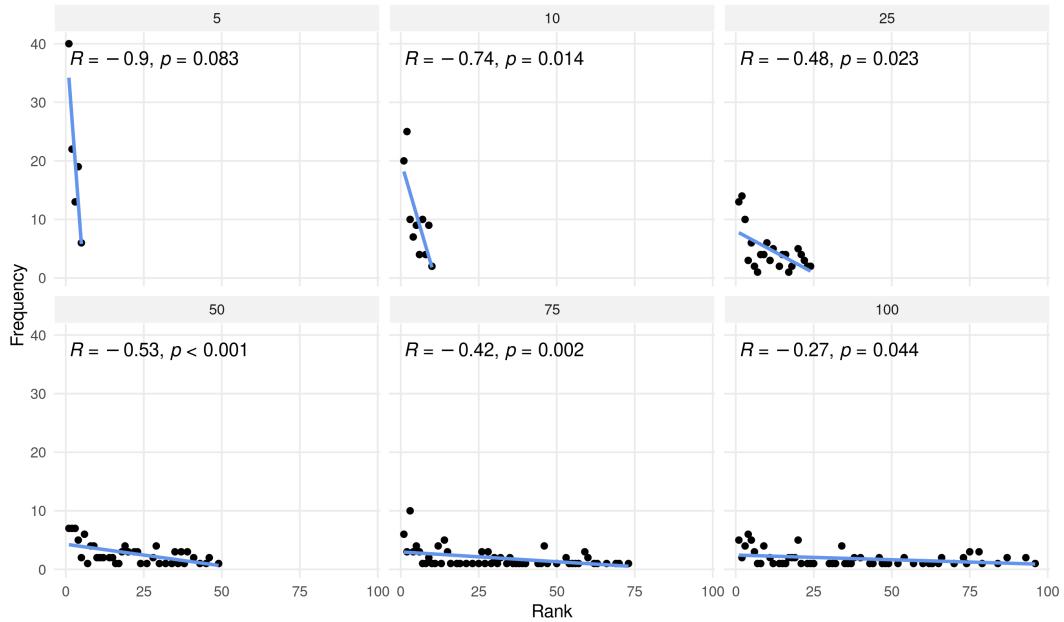


Figure S11: Illustration how the Spearman's rank correlation changes when the number of random steps per stratum is varied from 5 to 100 (facettes). The y-axis shows the frequency at which the observed step was assigned the rank indicated on the x-axis. A low rank indicates that the observed step was assigned a high probability of being chosen. The better the prediction, the more frequently the observed step should be assigned a low rank, thus resulting in negative Spearman's rank-correlation. However, it is obvious that the metric heavily depends on the number of random steps per stratum. If there are only a few random steps, the correlation is more likely to be negative. The show patterns are based on simulated data.

References

- Aybar, C., Qiusheng, W., Bautista, L., Yali, R., Barja, A., Ushey, K., Ooms, J., Appelhans, T., Allaire, J. J., Tang, Y., Roy, S., Adauto, M., Carrasco, G., Bengtsson, H., Hollister, J., Donchyts, G., and JOSS, M. (2024). Rgee: R Bindings for Calling the 'Earth Engine' API.
- Chraibi, E., de Boissieu, F., Barbier, N., Luque, S., and Féret, J.-B. (2022). Stability in Time and Consistency Between Atmospheric Corrections: Assessing the Reliability of Sentinel-2 Products for Biodiversity Monitoring in Tropical Forests. *International Journal of Applied Earth Observation and Geoinformation*, 112:102884.
- Cutler, F. o. b. L. B. a. A. and Wiener, R. p. b. A. L. a. M. (2024). randomForest: Breiman and Cutler's Random Forests for Classification and Regression.
- Gilabert, M. A., Conese, C., and Maselli, F. (1994). An Atmospheric Correction Method for the Automatic Retrieval of Surface Reflectances from TM Images. *International Journal of Remote Sensing*, 15(10):2065–2086.
- Hijmans, R. J., Bivand, R., Pebesma, E., and Sumner, M. D. (2024). Terra: Spatial Data Analysis.
- Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, 28(5).
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., and Gascon, F. (2017). Sen2Cor for Sentinel-2. In Bruzzone, L., Bovolo, F., and Benediktsson, J. A., editors, *Image and Signal Processing for Remote Sensing XXIII*, page 3, Warsaw, Poland. SPIE.
- Ranghetti, L., Boschetti, M., Nutini, F., and Busetto, L. (2020). "sen2r": An R Toolbox for Automatically Downloading and Preprocessing Sentinel-2 Satellite Data. *Computers & Geosciences*, 139:104473.
- Shetty, S., Gupta, P. K., Belgiu, M., and Srivastav, S. K. (2021). Assessing the Effect of Training Sampling Design on the Performance of Machine Learning Classifiers for Land Cover Mapping Using Multi-Temporal Remote Sensing Data and Google Earth Engine. *Remote Sensing*, 13(8):1433.
- Śmielak, M. K. (2023). Biologically Meaningful Moonlight Measures and Their Application in Ecological Research. *Behavioral Ecology and Sociobiology*, 77(2):21.
- Storey, J., Scaramuzza, P., Schmidt, G., and Barsi, J. (2005). Landsat 7 Scan Line Corrector-Off Gap-Filled Product Development.
- Therneau, T., Atkinson, B., and Ripley, B. (2024). Rpart: Recursive Partitioning and Regression Trees.
- Vermote, E. F. and Kotchenova, S. (2008). Atmospheric Correction for the Monitoring of Land Surfaces. *Journal of Geophysical Research: Atmospheres*, 113(D23).