



A Multiverse Approach to Assessing the Impacts of Analysis Choice on Estimates of Habitat Selection

Submitted for Doctor of Philosophy
University of Stirling

Benjamin Michael Marshall

November 2024

Contents

Prerequisites	5
General Abstract	5
Authorship Statement	6
Acknowledgements	6
List of Figures	7
List of Tables	26
Introduction	27
1 abmAnimalMovement: An R package for simulating animal movement using an agent-based model	31
1.1 Abstract	31
1.2 Introduction	32
1.3 Methods	36
1.4 Results	51
1.5 Discussion	55
1.6 Future directions	58
1.7 Use cases	59
2 A Habitat Selection Multiverse Reveals Largely Consistent Results Despite a Multitude of Analysis Options	73
2.1 Abstract	73
2.2 Introduction	74
2.3 Methods	79
2.4 Results	90
2.5 Discussion	102
2.6 Conclusions	105
2.7 Supplementary Material	108
3 Population-level Habitat Selection Analyses Produce Consistent Effect Direction, but Strength is Sensitive to Researcher Choice	115
3.1 Abstract	115
3.2 Introduction	116
3.3 Methods	119
3.4 Sampling and Analysis Options	120
3.5 Results	126
3.6 Discussion	134
3.7 Conclusions	138
3.8 Supplementary Material	140

4 Snake Habitat Selection Findings are not Undermined by Researcher Analysis Choice	143
4.1 Abstract	143
4.2 Introduction	144
4.3 Methods	147
4.4 Results	158
4.5 Discussion	169
4.6 Conclusions	171
4.7 Supplementary Material	178
Conclusions	209

Prerequisites

General Abstract

Possibly the best way to determine an animal's needs is to follow their movements. Once we have learnt of the animal's movements we can infer habitat requirements, behaviour, and potential threats. Combined, the movement data and inferences can inform decisions on animal conservation. However, extracting useful information from animal movement requires many steps, from sampling to analysis. Other studies have shown that a single dataset can result in many different answers in the hands of different researchers, so how can we be confident the results from animal movement are leading to the correct decisions in animal conservation? One approach to explore this question is to perform a wide range of alternative analyses on conservation-relevant datasets and examine whether the resulting conclusions would differ. The combination of all these alternative analyses can be called a multiverse. We used a multiverse approach to explore thousands of ways of extracting the habitat selection estimates from the movement of simulated animals with a pre-defined selection. We found that despite different sampling approaches, and completely different analysis methods, the results agree and correctly identify habitat selection. The agreement between different habitat selection estimates tended to be better with more data, and when using more modern analysis methods. We applied this same multiverse of analysis methods to four case studies of snake movement data from Thailand, with the aim of determining the robustness of the originally published estimates of habitat selection. The spread of results from the multiverse indicated that the original conclusions were in agreement with the vast majority of analysis endpoints, and not the product of a particularly peculiar combination of analysis decisions. However, it did reveal the species-specific impacts of certain analysis decisions indicating that it would be unwise to discount analysis choice as a potential source of disagreement between studies examining the same or similar questions.

Authorship Statement

I hereby confirm that this thesis is an original piece of work conducted independently by the undersigned and was done so with the ethical approval of the Ethical Review Body. All research material has been duly acknowledged and cited; the contributions of others are listed on a per chapter basis. No generative Artificial Intelligence was used during this research and the writing of this thesis.

A handwritten signature in black ink, appearing to read 'Ben Marshall', written in a cursive style.

Benjamin Michael Marshall

Acknowledgements

I would like to thank the IAPETUS2 Doctoral Training Program (grant reference NE/S007431/1) for providing the funding and opportunities for training during this PhD. I would like to thank my excellent supervisory team, particularly Brad Duthie for their support from the very start of this project. I would also like to thank Lucy Hawkes and Thiago Silva for all the comments and suggestions on the thesis during the viva process.

Chiefly, I would like to thank Aubrey Alamshah for the unrelenting support and encouragement throughout this project.

List of Figures

- 1.1 A diagram showing the transition probability matrix between the three behavioural states. Remain values indicate the probabilities of the animal remaining in the same behavioural state. Values correspond directly to those in the transition matrix used as a simulation input. 37
- 1.2 A diagram showing the simulated animal's decision processes operating at two different time frames. A - the time frame of the destination choice, B - the time frame of movements at each time step. 38
- 1.3 The three resulting landscape layers to be fed into the simulation for the badger example: shelter quality, foraging resources, movement ease. 45
- 1.4 The distribution of step lengths used during the badger simulation example. Parametrising distributions are shown as solid areas using the same shape and scale parameters used as input into the simulation function. The lower plot shows the distribution used to generate potential foraging destinations. The realised resulting simulated step lengths (scaled to the input units) are shown by the solid lines. The realised resulting turn angles are shown in left hand sub-plots. Inset pie charts show the number of step lengths that were below the shelter site size. Note that x axis is not consistent between the three plots. 52
- 1.5 The observed locations of the three simulated species (A - Badger, B - Vulture, C - King Cobra). Black points show the observed location at each time step. The orange squares show the dynamically selected foraging destinations. Circles with an interior S show the shelter site locations, and circles with an interior A show the avoidance points. Note that the size represented by each unit on the x and y axis differs depending on the species. 53
- 1.6 The simulated locations of the three simulated species (A - Badger, B - Vulture, C - King Cobra) over the first month of time steps. Grey path shows the overall movement during that month, overlaid points indicate where the animal was in a given behavioural mode. Circles with an interior S show the shelter site locations, and circles with an interior A show the avoidance points, black squares show the dynamically selected foraging destinations. Note that the size represented by each unit on the x and y axis differs depending on the species. . 54

1.7	The simulated states (top half) and parametrised activity cycles (bottom half) governing sheltering behaviour in the three example species (A - Badger, B - Vulture, C - King Cobra). Point colour and position describe the behavioural state at each simulated time step, whereas the purple waves indicated the input values. Note that the input waves acting on the simulated animal in conjunction with the behavioural transition matrix.	56
1.8	The percentage of time steps spent in state 0 - resting per day, and how it varies over the entire simulated year.	57
1.9	The three resulting landscape layers to be fed into the simulation for the vulture example: shelter quality, foraging resources, movement ease.	63
1.10	The three resulting landscape layers to be fed into the simulation for the king cobra example: shelter quality, foraging resources, movement ease. Movement ease is blocked by bars of -99 weighting.	67
1.11	The distribution of step lengths used during the vulture simulation example. Parametrising distributions are shown as solid areas using the same shape and scale parameters used as input into the simulation function. The lower plot shows the distribution used to generate potential foraging destinations. The realised resulting simulated step lengths (scaled to the input units) are shown by the solid lines. The realised resulting turn angles are shown in left hand sub-plots. Inset pie charts show the number of step lengths that were below the shelter site size. Note that x axis is not consistent between the three plots.	68
1.12	The distribution of step lengths used during the king cobra simulation example. Parametrising distributions are shown as solid areas using the same shape and scale parameters used as input into the simulation function. The lower plot shows the distribution used to generate potential foraging destinations. The realised resulting simulated step lengths (scaled to the input units) are shown by the solid lines. The realised resulting turn angles are shown in left hand sub-plots. Inset pie charts show the number of step lengths that were below the shelter site size. Note that x axis is not consistent between the three plots.	69
2.1	Diagram showing how multiverse analysis differs from other approaches. Each branch node represents a choice made during analysis. Based on Figure 1 from Dragicevic et al. (2019).	78

2.2	An example of the underlying environmental matrices used during the Badger species simulation, and the subsequent habitat classes to be used in the habitat selection analysis.	81
2.3	A diagram of the decisions creating a multiverse examining animal movement. Grey choices pertain to data simulation and sampling. Species: the three different movement archetypes the simulations were based on. Individual: the number of individuals simulated per each species. Tracking duration: the length in days of the tracking data sampled from the overall simulated year. Tracking frequency: the frequency of the sampling of data points from the 1-per-minute simulated data. Selection scenarios: the two classified habitat landscapes used during analysis. Purple choices concerned with analysis pathways using area-based estimations of availability. Area method: the method used to generate the available area polygon. Contour: the threshold used to convert the utilisation/occurrence distribution into a polygon. Available points: the number of generated points per data point to extract available habitat within the polygon. Points pattern: whether the random points were generated in a stratified or random pattern. Weighting: the weighting of the available points in relation to the used points in the RSF formula. Light orange choices pertain to the SSF analysis pathways. Model formula: whether the step and turn angles were added as interactions to the habitat variable. Available points: the number of randomly generated available points per used point. Turn angle: the distribution used to generate the turn angles of the available points. Step distribution: the distribution used to generate the step distances of the available points. Dark orange is the WRSF analysis. Decisions that required discrete choices are illustrated with circles, whereas continuous are depicted as triangles. . . .	83

- 2.4 A specification curve showing all habitat selection estimates from the Wides analysis pathways applied to the simulated data. Each point is an estimate of habitat selection for habitat 2. Top plot shows all estimates arranged by estimated selection strength, with an arbitrary sorted y-position. The distribution of the selection estimates is also shown as a grey density plot. Lower plot shows the same estimates in relation to the decisions that lead to them. Left-hand numbers describe the number of NA results connected to each decision, indicating a failure somewhere in the calculations. Solid circles are the medians for each choice. 91
- 2.5 A specification curve showing all habitat selection estimates from the RSF analysis pathways applied to the simulated data. Each point is a estimate of habitat selection for habitat 2. Top plot shows all estimates arranged by estimated selection strength, with an arbitrary sorted y-position. The distribution of the selection estimates is also shown as a grey density plot. Lower plot show the same estimates in relation to the decisions that lead to them. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. Colours illustrate deviation from the median, with purple indicating lower levels of selection and orange showing higher selection estimates. 93
- 2.6 All standard errors associated with estimates of habitat selection from RSF, SSF, and wRSF analysis pathways for habitat 2. More extreme estimates of selection are associated with dramatically larger confidence intervals, N.b., y axis is log scaled. 95
- 2.7 A specification curve showing all habitat selection estimates from the SSF analysis pathways applied to the simulated data. Each point is a estimate of habitat selection for habitat 2. Top plot shows all estimates arranged by estimated selection strength, with an arbitrary sorted y-position. The distribution of the selection estimates is also shown as a grey density plot. Lower plot show the same estimates in relation to the decisions that lead to them. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. Colours illustrate deviation from the median, with purple indicating lower levels of selection and orange showing higher selection estimates. 96

2.8	A specification curve showing all habitat selection estimates from the wRSF analysis pathways applied to the simulated data. Each point is a estimate of habitat selection for habitat 2. Top plot shows all estimates arranged by estimated selection strength, with an arbitrary sorted y-position. The distribution of the selection estimates is also shown as a grey density plot. Lower plot show the same estimates in relation to the decisions that lead to them. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. Colours illustrate deviation from the median, with purple indicating lower levels of selection and orange showing higher selection estimates.	97
2.9	Estimated betas for each choice's impact on the deviation from the median habitat selection estimate. Deviation from the median estimate was examined as the absolute deviation (left) and the raw deviation (right). All eight independent models are displayed together, split between analysis method and deviation type. Each point represents the point estimate derived from the 95% median highest density continuous interval, and is plotted alongside a line showing the extent of the 95% credible interval. The y-axis shows all the choices explored in the different analysis methods, grouped by the methods they apply to.	98
2.10	The resulting interactions from the Bayesian regression model looking to explain absolute deviation from the median in the RSF analysis pathway. Left-hand plot shows how scaled tracking duration interacted with area method chosen, while the right-hand shows the interaction between scaled tracking frequency and area method chosen.	100
2.11	The posterior distributions of the betas from the Bayesian regression model looking to explain absolute deviation from the median in the Wides analysis pathway.	109
2.12	The posterior distributions of the betas from the Bayesian regression model looking to explain absolute deviation from the median in the RSF analysis pathway.	109
2.13	The posterior distributions of the betas from the Bayesian regression model looking to explain absolute deviation from the median in the SSF analysis pathway.	110

2.14	The posterior distributions of the betas from the Bayesian regression model looking to explain absolute deviation from the median in the wRSF analysis pathway.	110
2.15	The posterior distributions of the betas from the Bayesian regression model looking to explain raw deviation from the median in the Wides analysis pathway.	111
2.16	The posterior distributions of the betas from the Bayesian regression model looking to explain raw deviation from the median in the RSF analysis pathway.	111
2.17	The posterior distributions of the betas from the Bayesian regression model looking to explain raw deviation from the median in the SSF analysis pathway.	112
2.18	The posterior distributions of the betas from the Bayesian regression model looking to explain raw deviation from the median in the wRSF analysis pathway.	112

- 3.1 A diagram of the different decisions creating a multiverse examining animal movement. Species: three different movement archetypes simulations were based on. Sample Size: the number of simulated individuals per each species, with each random sample being repeated twice. Tracking duration: the length (days) of the tracking data sampled from the overall simulated year. Tracking frequency: the frequency of sampled data points from the 1-per-minute simulated data. Selection scenarios: the two classified habitat landscapes used during analysis. Purple choices highlight the analysis pathways using area-based estimations of availability. Area method: the method used to generate the available area polygon. Contour: the threshold used to convert the utilisation/occurrence distribution into a polygon. Available points: the number of generated points per data point to extract available habitat within the polygon. Points pattern: whether the random points were generated in a stratified or random pattern. Light orange choices pertain to the Two-Step analysis pathways. Medium orange-brown choices pertain to the SSF analysis pathways. Dark brown choices pertain to the Poisson analysis pathways. Model formula: whether the step and turn angles were added as interactions to the habitat variable. Available points: the number of randomly generated available points per step/used point. Turn angle: the distribution used to generate the turn angles of the available points. Step distribution: the distribution used to generate the step distances of the available points. Dark orange is the WRSF analysis. Averaging: the method of model averaging used for the SSF models to gain an overall estimate. Decisions that required discrete choices are illustrated with circles, whereas those that continuous are depicted as triangles. 121
- 3.2 A specification curve showing habitat selection estimates resulting from the Compana (area-based) analysis pathways when applied to all simulated animal movement datasets. Every point represents a separate estimate of habitat selection. The top plot shows all estimates organised by estimate strength, with the purple vertical line depicting the median estimate and grey distribution describing the overall pattern of the estimates. Instances of outlying high or low estimates are labelled instead of plotted. The lower plot shows the same estimates, but split by the analysis or sampling decisions. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. . . . 127

- 3.3 A specification curve showing habitat selection estimates resulting from the Step-selection Function (SSF) analysis pathways when applied to all simulated animal movement datasets. Every point represents a separate estimate of habitat selection. The top plot shows all estimates organised by estimate strength, with the purple vertical line depicting the median estimate and grey distribution describing the overall pattern of the estimates. Instances of outlying high or low estimates are labelled instead of plotted. The lower plot shows the same estimates, but split by the analysis or sampling decisions. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. 129
- 3.4 A specification curve showing habitat selection estimates resulting from the Two-step analysis pathways when applied to all simulated animal movement datasets. Every point represents a separate estimate of habitat selection. The top plot shows all estimates organised by estimate strength, with the purple vertical line depicting the median estimate and grey distribution describing the overall pattern of the estimates. Instances of outlying high or low estimates are labelled instead of plotted. The lower plot shows the same estimates, but split by the analysis or sampling decisions. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. 130
- 3.5 A specification curve showing habitat selection estimates resulting from the Poisson analysis pathways when applied to all simulated animal movement datasets. Every point represents a separate estimate of habitat selection. The top plot shows all estimates organised by estimate strength, with the purple vertical line depicting the median estimate and grey distribution describing the overall pattern of the estimates. Instances of outlying high or low estimates are labelled instead of plotted. The lower plot shows the same estimates, but split by the analysis or sampling decisions. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. 131

3.6	Point estimates and 95% credible intervals associated with all population effects from the Bayesian Regression models aiming to predict deviation from the median estimate using the information on analysis choices. Negative effects indicate that the decisions tend to decrease the deviation from the median estimate, whereas positive effects suggest those decisions are increasingly connected to estimates far from the median estimate. The effects are grouped into those pertaining to sampling choices and tracking, those associated with the step-based analysis approaches (Two-Step, SSF, and Poisson), those connected just to the SSF approach, and those connected to the area-based Compana approach.	134
4.1	The study site locations for all four case studies. Top left shows the location within Thailand, both the Sakaerat Biosphere Reserve and the Suranaree University of Technology campus are located within Nakhon Ratchasima Province. Bottom left shows the overall study location for King Cobra, Burmese Python, and Banded Krait (circular points). The purple areas depict the land uses that are included in any of the species hypotheses. Gray lines indicate roads. Bottom right is the same region, but focused on the Banded Krait locations to better show the field margin data used for their assessment of habitat selection. Top right is the study location for the Malayan Krait on the Suranaree University of Technology campus.	148
4.2	A summary of the species and hypotheses derived from the original studies regarding the habitat selection of each snake species. Alongside are the sample sizes for each species and a visualisation describing the two landscape classifications used in the multiverse.	150
4.3	A visual of all the decisions contained in the multiverse and colour coded to which approach that decision applies. Subtitle indicates the name of the decision, and the values or categories below denote the choices explored in the multiverse. Each decision is combined with all others pertaining to the same analysis method, in unique combinations, and applied to each species to create the full count of habitat selection estimates.	154

- 4.4 All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot show the estimates relative to each analysis choice. The colours depict the species, and shape separate hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. 160
- 4.5 All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot shows the estimates relative to each analysis choice. The colours depict the species, and shape separates hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 161
- 4.6 All estimates of habitat selection derived from the step-based Two-Step analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot shows the estimates relative to each analysis choice. The colours depict the species, and shape separates hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 163

- 4.7 All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot shows the estimates relative to each analysis choice. The colours depict the species, and shape separates hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 164
- 4.8 All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot shows the estimates relative to each analysis choice. The colours depict the species, and shape separate hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 166
- 4.9 All estimates of habitat selection derived from the weighted Resource Selection Function (wRSF) analysis. Circles and numbers above show the point estimate, and the horizontal bars depict the 95% confidence interval. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 167
- 4.10 Posterior estimates of the intercept for all meta-analysis models, with 95% HDICs, indicating whether the overall meta-analysed habitat selection estimate supports the original conclusions. Positive values indicate support. Colour denote species, and shape separates hypotheses for King Cobras. Continuous habitat classification was incompatible with the Compana method, and so is not shown. 168

4.11	The distribution between subsequent data points for each species. Circles below the distributions show the median time lag. Reported alongside is the mean, standard error ($\pm$), and range of time lags. Large time lags are usually the result of prolonged missing periods or transmitter failures. N.b., the x axis is square-rooted to aid with visualisation.	179
4.12	The timeline of all tracked individuals, split by species. Each recorded data point is depicted by a 	180
4.13	The landscape covariates used in the King Cobra models for hypothesis 1, overlayed with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat. .	181
4.14	The landscape covariates used in the King Cobra models for hypothesis 2, overlayed with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat. .	182
4.15	The landscape covariates used in the Burmese Python models for hypothesis 1, overlayed with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat.	183
4.16	The landscape covariates used in the Malayan Krait models for hypothesis 1, overlayed with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat. .	184
4.17	The landscape covariates used in the Banded Krait models for hypothesis 1, overlayed with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat. .	185

- 4.18 All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. Shape separates hypothesis 1 and 2 (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and hypothesis medians are connected with appropriated coloured lines. 186

- 4.19 All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). . . 187

- 4.20 All estimates of habitat selection derived from the step-based Two-Step analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates, split by hypothesis, with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate for each hypothesis. Lower plot shows the estimates relative to each analysis choice. Shape separates hypothesis 1 and 2 (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 188

- 4.21 All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates, split by hypothesis, with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate for each hypothesis. Lower plot shows the estimates relative to each analysis choice. Shape separates hypothesis 1 and 2 (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 189

- 4.22 All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates, split by hypothesis, with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate for each hypothesis. Lower plot shows the estimates relative to each analysis choice. Shape separates hypothesis 1 and 2 (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 190
- 4.23 All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. 191
- 4.24 All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). . . 192

- 4.25 All estimates of habitat selection derived from the step-based Two-Step analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 193
- 4.26 All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). . . 194
- 4.27 All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). . . 195

- 4.28 All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. 196
- 4.29 All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). . . 197
- 4.30 All estimates of habitat selection derived from the step-based Two-Step analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 198

- 4.31 All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). . . 199
- 4.32 All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). . . 200
- 4.33 All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. 201

4.34 All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 202

4.35 All estimates of habitat selection derived from the step-based Two-Step analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 203

4.36 All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right). 204

4.37	All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis for Banded Kraits (<i>Bungarus fasciatus</i>). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).	205
4.38	Conditional R^2 values for all meta-analysis models, with 95% HDCIs.	206
4.39	Posterior estimates of the population effects for all meta-analysis models, with 95% HDCIs.	207

List of Tables

1	The potential sources leading to variation in results (Hoffmann et al., 2021). . .	28
2.1	All R^2 values for Bayesian regression models.	108
2.2	Analysis results when applied directly to un-resampled movement data at both the decision timescale and movement timescales.	113
3.1	An overall summary of all Compana analyses and associated decisions in relation to the frequency of significant results.	140
3.2	An overall summary of all SSF analyses and associated decisions in relation to the frequency of significant results.	141
3.3	An overall summary of all Poisson analyses and associated decisions in relation to the frequency of significant results.	141
3.4	An overall summary of all Two-Step analyses and associated decisions in relation to the frequency of significant results.	142
4.1	Summary of tracking effort, tracking time lags (hours), and step lengths (metres) for each species.	178

Introduction

Scientific progress requires building upon past results, and therefore requires confidence in past results. Issues arise when subsequent research is based upon weak foundations –i.e., studies with a limited capacity to be replicated because of questionable practices or over-generalisation. Early significant results can dictate the direction of research and grow resistant to later contradictory results (Barto & Rillig, 2012); therefore, early diagnoses of overly confident results or previously unknown sources of variation becomes a priority.

In medical fields, a lack of replicability comes with direct monetary and well-being costs (Freedman, Cockburn & Simcoe, 2015), as well as stoking fears that the lack of replicability feeds distrust of science more generally (Anvari & Lakens, 2018). Like the medical field, ecological studies can come with well-being costs to the study subjects (e.g., direct surgery/marketing of the animal Reinert & Cundall, 1982; Winne et al., 2006), as well as impacts on stakeholders stemming from management decisions (Acebes et al., 2024). The welfare costs are heightened when animals are required to be captured and have some bio-telemetry devices attached (sometimes surgically, Weaver, Westphal & Taylor, 2021). Therefore, maximising replicability in ecology is key to minimising research waste (Grainger et al., 2019) and the negative impacts on study subjects and systems.

The reasons leading to a replication not matching the results of an original study can be numerous. The variation between final results can be considered to originate from six sources of uncertainty/variation (Table 1, Hoffmann et al., 2021). While natural random variation between samples is a candidate, sampling variation and even biases in samples is well recognised (Webster & Rutz, 2020; Farrar & Ostojić, 2021). In other cases, the cause behind the variability in results may be hard to diagnose (Brezna et al., 2021), or less likely to be questioned because of the agreement with existing theory (Gelman & Loken, 2013). However, what remains clearer is what researchers see as the cause. Many point to flaws in the scien-

tific environment that researchers work in, potentially leading to non-ideal decisions in the post-data-collection stages of research (Baker, 2016). The key detrimental aspect of the research environment identified is the pressure to publish, and the resulting secondary pressures to prioritise novelty and selectively report to achieve a clean narrative (Jennions & Møller, 2002b; Cassey et al., 2004; O’Boyle, Banks & Gonzalez-Mulé, 2014; Vinkers, Tijdink & Otte, 2015; Forstmeier, Wagenmakers & Parker, 2017; Brembs, 2019).

There are a suite of questionable options for researchers to massage results to something more publishable (Fraser et al., 2018b). They chiefly come in three forms: HARKing, cherry-picking, and p-hacking. Hypothesising After Results Known (HARKing) is where the research can present the results as a confirmatory result, despite originally there being no or contrary hypotheses. HARKing can sometimes be further enabled and rationalised by hindsight bias, where unexpected results are perceived as more likely once they have been observed (Gelman & Loken, 2013; Forstmeier, Wagenmakers & Parker, 2017). Cherry-picking is the removal or non-reporting of data points or (co-)variables, that did not yield significant results. P-hacking is the repeated use of statistical tests, with different settings, to achieve statistical significance. Arguably, the existence of p-hacking is enabled by an over-reliance on p-value thresholds, rather than flexible p-value thresholds that are predefined based on effect size of interest, sample size, and desired accuracy of estimation (Lakens et al., 2018).

For any of these questionable practices to be of use there must be adequate flexibility in how data is processed and analysed. Decisions (or “researcher degrees of freedom”) concerning data preprocessing, parameter, model, and method selection are likely to be particularly key in questionable research practices. Therefore, we may expect that disciplines with more study flexibility during those stages of research contain greater scope for questionable practices. Quantifying the potential for researchers’ choice (i.e., degrees of freedom) would be a valuable step in identifying areas for heightened awareness of questionable practices.

A subset of ecology that may be particularly difficult to execute direct replications of is movement ecology. On one hand researchers should want to avoid the repeated well being costs of

Table 1: The potential sources leading to variation in results (Hoffmann et al., 2021).

Origin	Details	Example
Measurement	Randomness from the act of measuring	Measurement error stemming from imperfect precision or accuracy
Data preprocessing	Decisions on data inclusion/exclusion and transforming	Differing thresholds for what counts as outlying
Parameter	Decisions on which parameters used as covariates/predictors	Differing thresholds on acceptable covariance of predictors
Model	Decisions on model structure and specification	Model formulation and specification of random effects
Method	Decisions on method choice and parameterisation	Choosing an overarching GLMM over a series of T-tests
Sampling	Randomness as a result of sampling a wider population	Sampling from a wider population results in different samples

bio-telemetry based tracking of animals (e.g., weight compensation in tagged birds Portugal & White (2022)). On the other hand, movement ecology, specifically animal movement tracking, has exploded in terms of technology and methods in the past 20 years leaving researchers with a new wealth of devices and approaches to choose from (Joo et al., 2022). As the field charges forward to new exciting horizons, it would be prudent to re-examine the past studies to assess how much can be built upon and what should be re-assessed. While re-assessments can be somewhat limited by resistance to data sharing (Miyakawa, 2020; Tedersoo et al., 2021) and data loss as it ages (Vines et al., 2014), there are visible improvements in movement ecology with movement data specific repositories enabling data protection, re-use and sharing (Kays et al., 2022). Such retroactive assessment is an attractive option when other methods to explore false positive rates (Hoffmann et al., 2021), such as preregistration and registered reports (Kaplan & Irvin, 2015; Scheel, Schijen & Lakens, 2021), will require more time to yield results.

One approach to retroactively explore the impacts of research choices is to construct a multiverse of analysis choices. A multiverse is a structured tree of possible analysis routes or pathways with branches formed from different analytical decisions, all aimed at answering the same research question (Dragicevic et al., 2019; Simonsohn, Simmons & Nelson, 2020; Rijnhart et al., 2021). Multiverse analysis presents an avenue to explore causes of variation between studies without the additional costs of practical studies, while also being capable of exploring data processing decisions that may not have immediately apparent impacts on final results.

To provide a baseline for multiverse explorations of real animal movement data, we first constructed a simulated scenario to determine the effectiveness of habitat selection analyses to recover a known “true” habitat selection. For the data of this simulated scenario, we created an agent-based animal model that simulates movement patterns, with a known underlying preference for a certain habitat (Chapter 1). Following the generation of a simulated suite of animals of several species archetypes, we applied a multiverse of habitat selection analyses looking at both individual (Chapter 2) and population level habitat selection (Chapter 3). These multiverses contained various different approaches to determining habitat selection and varied the choices within each approach. Using specification curves and Bayesian regression models, we examined the thousands of resulting habitat selection estimates to determine whether particular analytical decisions were more likely to lead to incorrect or outlying estimates. Finally, we applied this same multiverse to determine whether research analysis

choices would be sufficient to change the conclusions of previously published studies (Chapter 4). We selected four snake movement case studies from Northeast Thailand, where all the snake species are involved in some form of human-snake conflict.

Chapter 1

abmAnimalMovement: An R package for simulating animal movement using an agent-based model

Benjamin M. Marshall¹, Alexander B. Duthie¹

¹Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling,
Stirling, FK9 4LA, UK

Chapter is available as a reviewed preprint, awaiting revision: Marshall, B. M., & Duthie, A. B. (2022). abmAnimalMovement: An R package for simulating animal movement using an agent-based model. F1000Research, 11, 1182. <https://doi.org/10.12688/f1000research.124810.1>

1.1 Abstract

Animal movement datasets are growing in number and depth, and researchers require a growing number of analytical approaches to adequately answer questions using movement datasets. As the complexity of questions and analyses increase, deciding on the best approach both in terms of study design and analysis can become more difficult. A potential solution is to sim-

ulate an array of synthetic datasets under varying study designs and simulation parametrisations to gain insight into the impact of analysis choice(s) in different contexts. The `abmAnimalMovement` R package provides the means of simulating animal movement for this purpose. The `abmAnimalMovement` simulations use a discrete time agent-based model and do not require previous movement data as an input. The simulations include a number of key internal and external movement influences, as well as parameters for navigation and mobility capacity of the animal. Internal influences include three predefined behavioural states (e.g., rest, explore, forage) and any number of activity cycles (e.g., diel, seasonal). External influences are implemented via matrices describing landscape characteristics (e.g., shelter quality, foraging resources, movement ease), and predefined points describing shelter sites and points the animal avoids. Navigation capacity is defined by the range the animal can dynamically choose a foraging location to which it is subsequently attracted. Mobility capacity is implemented by user defined distributions, from which step length and turn angles are drawn at each time step, governing the possible subsequent locations of the animal. Critically, the navigation capacity (the choice of destination) operates on a different time scale to the mobility capacity, allowing the internal state of the animal to differ from the observed movements. When combined with other emergent properties, such as site fidelity generated via repeated shelter site use, the simulations offer opportunities to test whether movement analyses can accurately recover hidden mechanisms, states, and drivers.

1.2 Introduction

As the biodiversity crisis deepens, we are in constant need of new refined methods of mitigating human impacts on animals. Developing mitigation strategies is aided by a strong understanding of animals lives and needs. How animals move offers a window into their decision making process (Bastille-Rousseau et al., 2017; Sridhar et al., 2021), as well as how they prioritise resources (Prange, Gehrt & Wiggers, 2004; Christy et al., 2017; Vogt et al., 2018), avoid threats (Loveridge et al., 2017), and react to anthropogenic modification of the environment (Tucker et al., 2018; Arrondo et al., 2018; Tucker et al., 2019). Examinations of animal movement are frequently integrated into conservation plans (Fraser et al., 2018a) because of the insights they provide on habitat and space requirements, but also because they offer avenues to explore fundamental questions about ecology. For example, space use has been tied directly to body size connected via the energy requirements for animals (Noonan

et al., 2020), and movement has provided incredible insights into behaviour (Nathan et al., 2012; Studd et al., 2021).

The information gathered from animal movement studies is only as robust as the methods and analyses applied. In the past decade, the volume of animal data has exploded (Wild et al., 2022; Joo et al., 2022), and the infrastructure for collating these data has improved drastically (Kays et al., 2022). There are stand out examples of novel methods (Gurarie, Andrews & Laidre, 2009; Saunders et al., 2022), and exceptionally detailed insights into animal behaviour using new bio-logging technology (Nathan et al., 2012; e.g., Studd et al., 2021). However, on the whole, analysis approaches to maximise the information extracted have arguably lagged behind the technological and data improvements (Joo et al., 2022).

The physical cost to tracked animals carrying bio-telemetry devices (across taxonomic groups reptiles (Weatherhead & Blouin-Demers, 2004); mammals (Robstad et al., 2021); and birds (Portugal & White, 2022)), paired with the possible conservation sensitivity of those animals, means researchers have an ethical duty to maximise the value/impact of previously and currently collected data.

Maximising the value of data should include ensuring that results gained are robust and replicable. Meta-analyses have revealed that several textbook examples in ecology are not as universally replicable as once thought (Wang et al., 2018; Sánchez-Tójar et al., 2018; Roche et al., 2020; Clark et al., 2020). Whereas lab studies can more feasibly be repeated (although such replications are far from cheap), movement studies conducted on free-ranging wild animals are difficult to repeat while satisfactorily meeting the conditions of true replication (Fraser et al., 2020). It is hard to justify placing bio-telemetry devices on more animals to repeat a study, especially when they are of conservation concern. Studies tend to focus on larger, less vulnerable species (Crane et al., 2021; Tam et al., 2021), but as the benefits of detailed movement data grow, we can expect greater demand for tracking species of conservation concern (Fraser et al., 2018a). With the limited scope for repeating wild movement studies, and the desire to maximise study-to-benefit ratio for sensitive species, it is imperative that we maximise the quality of the initial studies.

Expanding pre-study planning, and developing better approaches to guide study design, may provide a partial solution (Williams et al., 2020; Silva et al., 2023). Study scenarios where researchers have control over the environment, randomisation, and controls can satisfy the assumptions of many statistical approaches (Christie et al., 2019). Instead, movement studies

conducted in the wild are fraught with uncontrollable confounding variables and often further restricted by sample sizes limited by ethics, capture rates, and cost of bio-telemetry equipment. As the complexity of the research environment increases, so must the analytical approach to account for the uncontrollable –but potentially extraneous to the question– variables. Testing these analysis approaches, and their limitations therefore becomes more difficult to assess *a priori*.

A potential solution is virtual ecology: where synthetic data are simulated to cover a range of possible scenarios allowing researchers to explore different study approaches/designs with the aim of identifying the best for a given question (Gupta et al., 2019). Exploration of analysis methods using simulations can yield useful insights even for broadly understood techniques (e.g., GLMMs (DeBruine & Barr, 2021)). Such an approach has the potential to adequately recreate the complex emergent properties that could hamper analyses (DeBruine & Barr, 2021). The suite of tools for simulating aspects of ecological systems (e.g., landscapes (Sciaini et al., 2018b), genetics (Petr et al., 2022), conservation choices (Gupta et al., 2019), ecological community comparisons (Guerra-Castro et al., 2021), resource selection (Street et al., 2021)) is growing, and some are explicitly focusing on aiding study design. There are also examples of simulation approaches being directly tied to applied issues like human-wildlife conflict (e.g., snakes (Goldstein et al., 2021), tigers (Ahearn et al., 2001)).

Population monitoring studies offer a great example of a mature relationship between simulated data, study design, and applied research, implemented in a complex study environment (Theng et al., 2022). There are examples of simulated datasets being used to validate field protocols (e.g., camera trap effort Howe et al., 2017; Cappelle et al., 2021) and analyses (e.g., accounting for species interactions (Kellner et al., 2022), or difficult to account for variation in populations (Milleret et al., 2022)). More recently, explicitly simulating animal movement processes rather than the population distribution has helped validate spatially-explicit capture-recapture methods, while also prompting deeper questions into findings’ generalisability to drastically different movement processes (Theng et al., 2022). Once validated with real-world data, simulated explorations of study design presents opportunities to understand the trade-offs between effort and precision (Cappelle et al., 2021). But similarly, disagreements on the implementation of surveying methods have led to further explorations of how we conceptualise and simulate animal processes (Abolaffio, Focardi & Santini, 2019).

As illustrated by the spatially-explicit-capture-recapture example mentioned, different simu-

lation approaches and different analyses accounting for different additional complexities offer different suggestions on the optimal study design. There are a number of approaches for simulating animal movement including (but not limited to):

1. *Continuous movement processes*, best exemplified by the *ctmm* package (Calabrese, Fleming & Gurarie, 2016) and worked examples such as Silva et al. (2022) and Theng et al. (2022). These methods rely on a mathematically defined movement process (such as Ornstein-Uhlenbeck foraging process Fleming et al., 2014). A key aspect of animal movement they simulate is correlated speeds and central tendency (i.e., home range).
2. *Hidden markov models*, exemplified by the `simData` function from the *moveHMM* package (Michélot, Langrock & Patterson, 2016). Their biggest contribution to simulating animal movement is the inclusion of state switching (i.e., different behaviours). They are best used in conjunction with existing movement data for parametric bootstrapping to explore estimator uncertainty (Michélot, Langrock & Patterson, 2016).
3. *Discrete agent-based approach*, is a bottom-up approach where simulations of an agent (e.g., an animal) are based upon a number of rules evaluated at discrete time-steps (Tang & Bennett, 2010), and through repeatedly following those rules a complex output emerges. In the model we present here, this output represent the movement pathways in a landscape. Previous examples of agent-based movement models have targeted specific key components of animal movements such as movement heavily constrained by landscape features (Quaglietta & Porto, 2019), simulating memory/home range (Van Moorter et al., 2009), or migratory behaviour (Bennett & Tang, 2006).

Here we present an agent-based approach, *abmAnimalMovement*, to supplement those existing simulations, which provides a unique combination of features. The agent-based approach is suited to explore a combination of the complexities of animal movement, and the emergent properties. Our model does not require fitting to prior movement data (only general definitions of movement characteristics), allows for multiple predefined points of attraction/avoidance, multiple-levels of activity cycling, three behavioural states, and multiple spatial environmental covariates. A new independently developed approach to simulating animal movement will provide a additional routes to test the robustness of movement analyses.

1.3 Methods

1.3.1 Overview of the agent-based model

The *abmAnimalMovement* package provides the functionality to simulate animal movements via an agent-based model. We wrote the model using *C++* via the *Rcpp* v.1.0.13.1 package (Eddelbuettel & François, 2011; Eddelbuettel, 2013; Eddelbuettel & Balamuta, 2018) to ensure it runs efficiently while allowing for easier manipulation of model inputs and outputs via R (R is the most used analysis tool in movement ecology with numerous supporting packages (Joo et al., 2020, 2022)).

The model simulates animal locations over a given period of time at discrete time steps. At each time step, the agent (i.e., simulated animal) is presented with a range of movement options in the form of new locations (Figure 1.2), and will choose from amongst these (i.e., sum-based model where the animal is considering all options simultaneously, as opposed to facing a series of sequential binary decisions, see Sridhar et al. (2021) for an example of the latter). The possible movement options, and how the new location is selected, are influenced by several factors: behavioural state of the animal, environmental quality, and proximity to points of attraction/avoidance. By simulating these drivers of animal movement, we hope to capture aspects of the internal state (i.e., motivation to move via behaviour); motion capacity (i.e., the individual’s varying ability to move); navigation capacity (i.e., ability to plan ahead beyond the immediate movement distance); and external factors (i.e., the landscape that steers and limits movement), all of which are defined as key components of animal movement (Nathan et al., 2008; Tang & Bennett, 2010).

The animal has three behavioural states, broadly defined here as state 0 - sheltering (or resting), state 1 - exploring, and state 2 - foraging. Each behavioural state modifies the movement characteristics of the animal; for example, exploring comprises of more random movements with longer distances between subsequent chosen locations, whereas resting behaviour is largely defined by stationary or very low discrete movements after returning to a shelter site. The movement characteristics are defined by two distributions: a Gamma distribution that step lengths are drawn from, and a Von-Mises that turn angles are drawn from (both widely used and frequently good representatives for step lengths and turn angles respectively). The resulting step length, combined with a turn angle describes the change in

animal location between each time step (as selected from a given number of options). We will define these based on the movement capacity of the animal over a minute, as a reasonable starting point for study design explorations.

As the simulation will be running with three behavioural states, we need to define how likely an animal is to switch between the behaviours. We achieved this by creating a transition matrix that describes the probability at each time step of the animal changing to another behaviour (Figure 1.1), where each value describes the probability of the animal transitioning from the current behaviour (row: b0, b1, b2), to the behaviour for the next time step (columns: b0, b1, b2). The diagonal describes the probability of the animal remaining in the same behavioural state. These probabilities can be kept very high to introduce autocorrelation in the behavioural state; a high autocorrelation is required if we are modelling time steps as one minute as there is a high likelihood the animal will remain in the same behavioural state minute-to-minute.

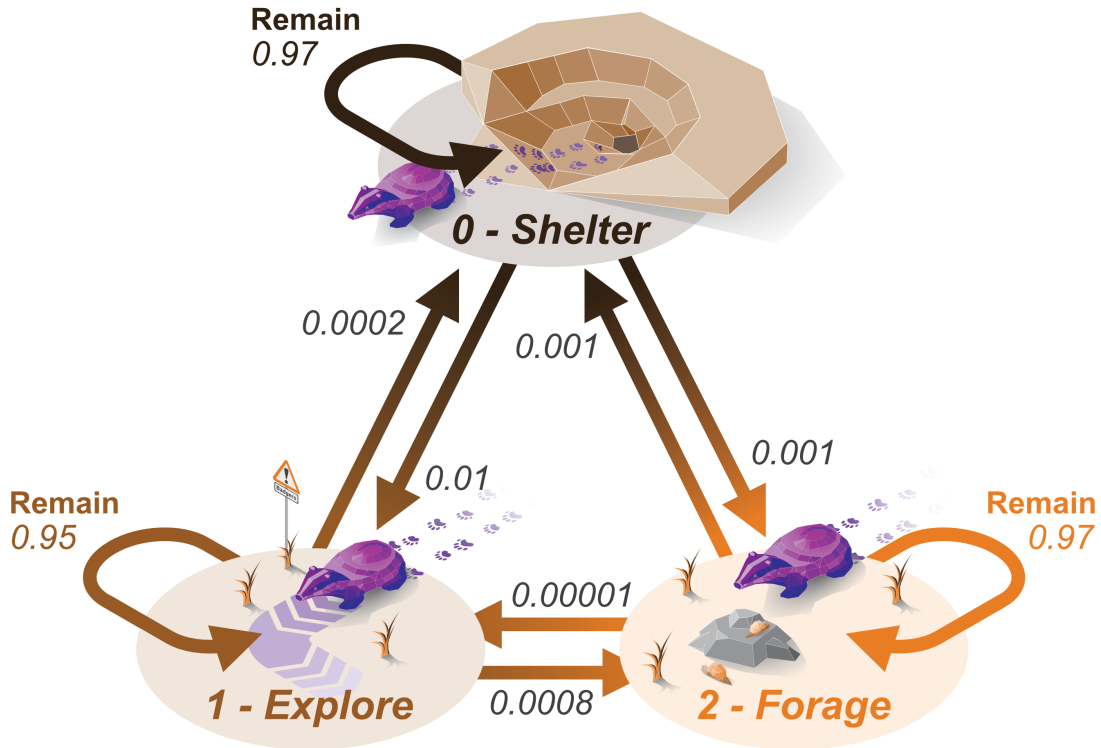


Figure 1.1: A diagram showing the transition probability matrix between the three behavioural states. Remain values indicate the probabilities of the animal remaining in the same behavioural state. Values correspond directly to those in the transition matrix used as a simulation input.

Animals' behaviour is frequently expressed via cycles, such as day/night or diel cycles (Rivera et al., 2022); therefore, we want the transition matrix to vary over time. We describe a number of cycles (or waves) that can be applied to the core transition matrix impacting the probability of entering resting behaviour. For example, a diel cycle can be expressed via a Sine wave with a frequency of 12 hours and an amplitude of 0.1. At each time step we draw a value from the wave, and add that weighting (from -0.05 to 0.05) to the transition probability of entering resting state (row 1, column 1 of the matrix); therefore, the probability of resting will increase and decrease following the Sine wave at a approximately 12 hour activity pattern. We can define as many cycles as needed, and they impact the resting probability additively.

When entering the resting state the animal will seek out a shelter site. In the *abmAnimalMovement* package, we can supply a number of shelter sites to simulate this need and create site fidelity. As these shelter sites act as points of attraction for the animal for each rest, the animal occupies a consistent area, or something approximating a home range. Home ranges are meant to represent areas in which the animal can source all resources required for a given life stage (Silva et al., 2022). The predefined and steady state of these shelter sites provides the stability we look for in a home range, as well as a means of predefining a level of site fidelity. As the *abmAnimalMovement* model can have multiple attraction (rest) sites supplied, we can simulate home ranges with unequal and behaviourally influenced space use.

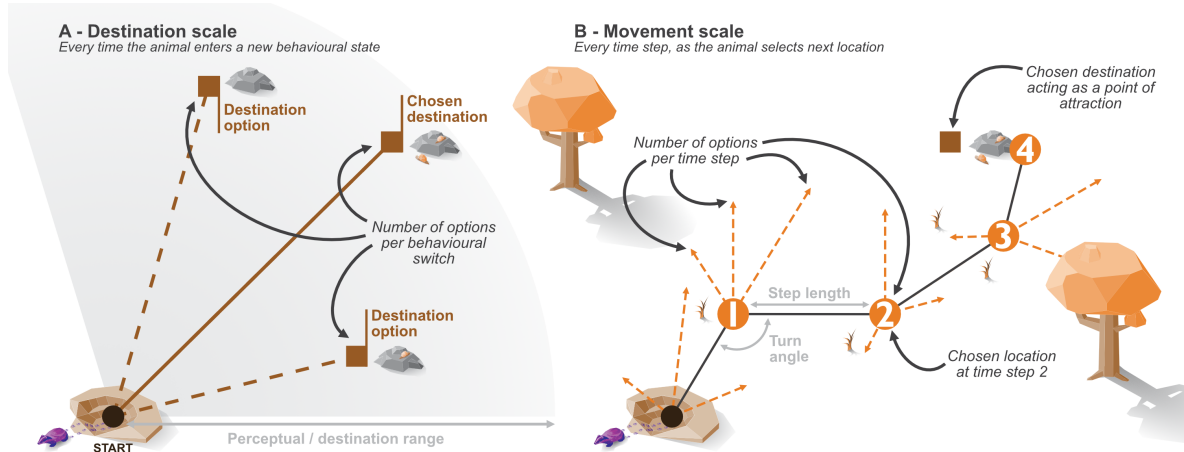


Figure 1.2: A diagram showing the simulated animal's decision processes operating at two different time frames. A - the time frame of the destination choice, B - the time frame of movements at each time step.

In addition to the predefined attraction to shelter sites, the *abmAnimalMovement* package allows for a more dynamic attraction to areas of high resource quality. Movements cannot be directly translated to animal preference; for example, habitat preference may miss key move-

ment corridors (Scharf et al., 2018); or the habitat decisions may occur at scales different to observed movement (Bastille-Rousseau et al., 2017). Therefore, the simulation required a mechanism that somewhat detaches the preference/choice for observed movements. When entering foraging mode, the animal randomly selects foraging destinations as a point of attraction (but weighted towards areas of higher quality) (Figure 1.2). This attraction impacts the movement choices made at each time step, with the animal more likely to choose (and therefore move to) options closer to the foraging destination. Therefore, foraging destination choice operates on a different time frame to the movement and allows movements through low-quality foraging areas. The broader distance at which this destination can be chosen acts as a proxy for the animal’s familiarity with the landscape more broadly (i.e., memory and planning), and creates a tacit assumption that the simulated individual is resident. This presents a critical benefit of the simulation approach, as we define the internal state decision making process of the animal. The two time frames also allow exploration of assumptions connecting observed movement choices to choices grounded in resource use, and whether downstream analyses can accurately recover a hidden decision making process.

At all times, the subsequent locations of the animal are impacted by a movement resistance matrix. Differing values represent different environmental conditions that could change how easily/likely an animal is to use that area to move towards a destination. For example, rivers or hard barriers within a landscape could present very high movement resistance and an animal would aim to avoid traversing them. Alternatively, areas of lower movement resistance could aid movement, and potentially form movement corridors. The movement matrix has values that describe the weighting of movement probability (lower values indicating a lower probability of entering a cell). Whereas the resource environmental layer and shelter locations interact with destination/goal decisions, the movement resistance affects the step-by-step movement decisions.

The interplay between shelter sites, foraging quality, and movement resistance simulates the site fidelity required for home ranges, while allowing the environment to help shape the size and diffusion of that home range. Simulating a more dynamic and messy array of movements can help test how far assumptions of uniform circular animal movement are practical.

1.3.2 Generating three species

To demonstrate the variation that movement ecology methods needs to wrangle, as well as the scope of movements the agent-based model can recreate, we provide three example species. The examples cover a range of movement capacities, site fidelity, and resting/sheltering patterns. Unlike alternative methods that are built to fit and simulate movements from existing data, our examples are loosely based on summary statistics from previous published results (cited below in each species section). The choice of using species specific examples is intended to provide some biological context while demonstrating a range of simulation parametrisations. Therefore, we are only interested in whether the simulated data returns the intended traits we are attempting to simulate for each example species (e.g., avoidance, site fidelity, expected step lengths). We show one example via a more detailed walk-through (badger), and the other example parametrisations are included as uses cases at the end of the manuscript (vulture, king cobra).

1.3.3 Primary example

Ecology and objectives - badger

Our primary example is based on a badger. Badgers occupy setts, in our example two home/shelter sites, to where they routinely return (Feore & Montgomery, 1999; Kowalczyk, Zalewski & Bogumiła, 2006). When not at the sett, the badger forages, and the foraging distances are impacted by resource position and movement capacity of the badger (i.e., how far can a badger feasibly travel for food from the sett). The badger therefore expresses very high site fidelity, a range dictated by the spatial positioning of resources, and is subject to the movement resistance of the terrestrial environment.

We can draw on studies such as Kowalczyk, Zalewski & Bogumiła (2006) and Rosalino, Macdonald & Santos-Reis (2005) to roughly gauge the speed of badgers (i.e., step length per minute). How this speed differs between behaviours is more difficult to estimate from existing literature, but we can use Kowalczyk, Zalewski & Bogumiła (2006) reported maximum speed to guide more direct movements (e.g., state 0 - sheltering and state 1 - exploring). In particular Kowalczyk, Zalewski & Bogumiła (2006) mentioned the heightened speeds moving from and to the setts. We also want to allow the exploratory (state 1) movements to be great enough

to occasionally exceed the normal home range or territory (Kelly et al., 2020). We can draw on statements regarding maximum distance travelled in a night from papers such as Rosalino, Macdonald & Santos-Reis (2005), Kowalczyk, Zalewski & Bogumiła (2006), and Loureiro et al. (2007) to approximate how distant foraging locations could be from a sett, while also confirming the nocturnal activity cycle for the badger. For the example, we will consider badgers as fully nocturnal, and with active periods lasting for around 8 hours each night (Rosalino, Macdonald & Santos-Reis, 2005; Magowan et al., 2022). The 8 hour active periods are not consistent throughout the year. Badgers occupying temperate areas are impacted by seasonal shifts that modify daylight hours and available resources (Rosalino, Macdonald & Santos-Reis, 2005; Magowan et al., 2022). Badger movements are also impacted by their territoriality, avoiding areas occupied by other badger groups, but also displaying occasional extra-territorial movements (Feore & Montgomery, 1999; Kelly et al., 2020). This suggests a general, but not complete avoidance of areas occupied by other badger groups.

We can broadly summarise the badger ecology we want to parametrise as follows:

1. Exhibits site fidelity via the use of two shelter sites (Feore & Montgomery, 1999; Kowalczyk, Zalewski & Bogumiła, 2006)
2. Movement speed approximated by summary statistics from previous studies (Rosalino, Macdonald & Santos-Reis, 2005; Kowalczyk, Zalewski & Bogumiła, 2006), while constrained by terrestrial environment and territoriality (Feore & Montgomery, 1999; Kelly et al., 2020)
3. An 8-12 hour activity cycle, that shifts over the year (Rosalino, Macdonald & Santos-Reis, 2005; Magowan et al., 2022)

1.3.4 Operation

The *abmAnimalMovement* package requires links to the *Rcpp* ($\geq 1.0.8.3$) (Eddelbuettel & François, 2011; Eddelbuettel, 2013; Eddelbuettel & Balamuta, 2018) and *BH* ($\geq 1.78.0.0$) packages (Eddelbuettel, Emerson & Kane, 2021), and therefore requires a version of *R* $\geq 3.5.0$ (R Core Team, 2021) (available from www.r-project.org). Currently the *abmAnimalMovement* package can be installed from Github using `install_github` provided by the *devtools* package (Wickham et al., 2021).

1.3.5 Implementation

While only the *abmAnimalMovement* package is required for running core simulations (*abm_simulate* function), to generate, organise, and review the inputs and outputs of simulations, we make use of a number of *R* packages. An alternative simpler implementation with no dependencies is provided as a package vignette.

```
# core package
library(abmAnimalMovement)

# data manipulation
library(dplyr)
library(reshape2)

# visualisation
library(ggplot2)
library(ggforce)
library(ggtext)
library(ggribes)
library(patchwork)

# environmental matrix generation
library(raster)
library(NLMR)
```

We used *R* v.4.4.2 (R Core Team, 2021) via *RStudio* v.2024.9.0.375 (R Studio Team, 2021), and used *rmarkdown* v.2.29 (Xie, Allaire & Golemund, 2018; Xie, Dervieux & Riederer, 2020; Allaire et al., 2022), *bookdown* v.0.42 (Xie, 2016, 2022a), *tinytex* v.0.54 (Xie, 2019, 2022b), and *knitr* v.1.49 (Xie, 2014, 2015, 2021) packages to generate type-set outputs. We used the *here* v.1.0.1 package (Müller, 2020a) to help with relative file path definition. We used the *dplyr* v.1.1.4 and *reshape2* v.1.4.4 packages for data manipulation (Wickham, 2007a; Wickham et al., 2022b). We used *ggplot2* v.3.5.1 for creating figures (Wickham, 2016a; Wickham et al., 2022a), with the expansions: *ggribes* v.0.5.6 (Wilke, 2021), *ggtext* v.0.1.2 (Wilke, 2020), *ggforce* v.0.4.2 (Pedersen, 2021), and *patchwork* v.1.3.0 (Pedersen, 2020). We used the *raster* v.3.6.30 (Hijmans, 2022), *sp* v.2.1.4 (Pebesma & Bivand, 2005a; Bivand, Pebesma & Gomez-Rubio, 2013; Pebesma & Bivand, 2022), and *NLMR* v.1.1.1 (Sciaini et al., 2018a, 2021) for generating environmental matrices. None of the above packages are dependencies of the *abmAnimalMovement* package.

Generating environmental matrices

To ensure that the simulation completed during this example is repeatable, we set a seed. For the sake of simplicity, the year the examples were written (2022) is used.

```
seed <- 2022
set.seed(seed)
```

Before starting to simulate animal movement, we needed to generate a landscape. The landscapes in this case will take the form of matrices, where each cell is describing the quality of foraging, shelter, and movement ease. The highest quality locations/cells are coded as 1, with quality decreasing as the values range down to 0. The 1 to 0 quality values in each cell are later used to help the animal to choose how and where it moves throughout the landscape. Depending on the behavioural state, the weighing of which matrix/layer used will change (e.g., when in a resting behavioural state the shelter site quality layer is used).

For most applications, a landscape would be known *a priori*, but for the convenience of this demonstration and testing of methods we use a selection of randomly generated landscape matrices (Figure 1.3), using the *NLMR* package (neutral landscape models).

We used a Gaussian field with an autocorrelation range of 40, a magnitude of variation across the landscape of 5, a magnitude of variation in the scale of autocorrelation range of 0.2, and a mean of 0.5. For foraging, we built on the Gaussian field already produced, allocating all values lower than 0.4 as 0 (i.e., no value for foraging), then normalising the remaining values between 0 and 1.

For a baseline movement resistance layer, we take the original Gaussian field and increase areas greater than 0.6 by 0.5 to allow areas of high resource quality to be easily accessible. We also greatly increase the movement ease in “edge” habitat (+1), where values fall between 0.6 and 0.3. Again we normalise between 0 and 1, where 1 are areas easily traversed. The resulting environment is one with easily traversable edge areas, surrounding better quality foraging locations than can be moved into easily.

Shelter quality is intermediate; we increased weighting by +1 in areas where cell values were greater than 0.5 and lower than 0.7. Thereby shelter sites are more likely to occur in the areas of higher foraging quality, but less likely to exist in the core (i.e., >0.7) or edge areas (<0.5). This provides some balance between accessibility and proximity to resources. The


```

rowcol = TRUE)

BADGER_shelterLocs <- data.frame(
  "x" = sampledShelters[,2],
  "y" = sampledShelters[,1])

BADGER_shelterSize <- 8

```

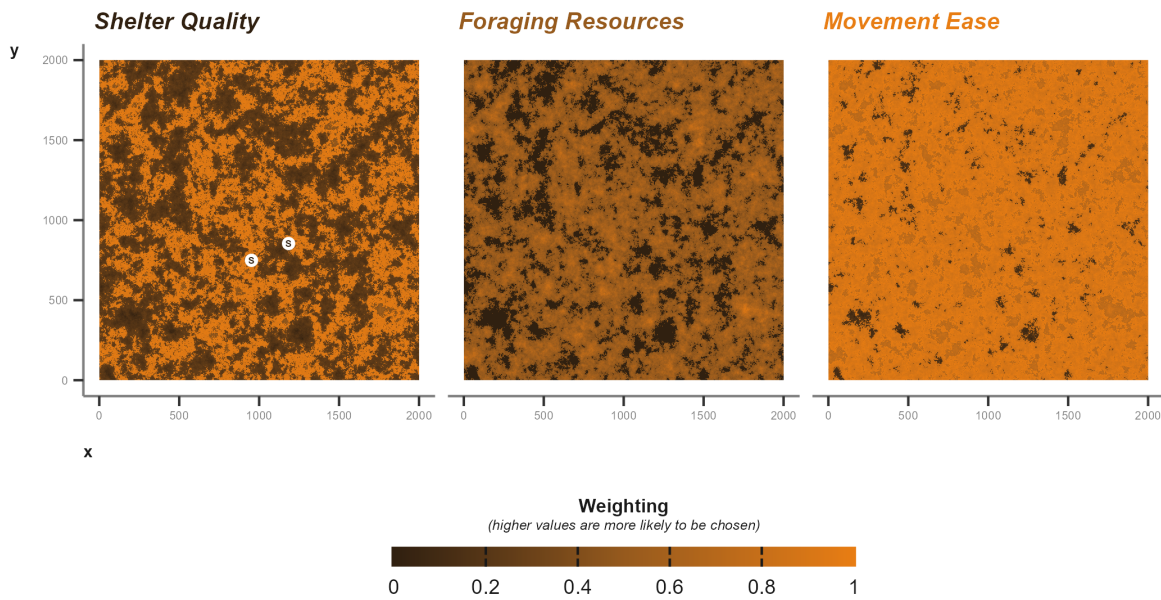


Figure 1.3: The three resulting landscape layers to be fed into the simulation for the badger example: shelter quality, foraging resources, movement ease.

Badger movement is set via the definition of three pairs of Gamma and Von Mises distributions. Each behavioural state is provided with a Gamma distribution to describe the step lengths between locations, and a Von Mises to describe the turn angles. When combined, these two distributions provide the animal with a number of locations to choose from at each time step (Figure 1.2). As we have three behavioural states, each movement parameter takes the form of a vector of length three. The step lengths require predefined shape (k) and scale (θ) parameters to describe the Gamma distributions, and we provide these to the simulate function via the `k_step` and `s_step` arguments. For the turn angles we require the mean (μ) and concentration (κ) to define the Von Mises distributions, and provide these via the `mu_angle` and `k_angle` arguments.

The perceptual range, in other words, where the badger decides to forage is set in a similar

fashion. Where **destinationRange** provides the shape (k) and scale (θ) for the Gamma distribution describing distance of possible foraging locations, and **destinationDirection** provides the mean (μ) and concentration (κ) the Von Mises distribution describing the angle at which those locations can fall (Figure 1.4).

We can also alter the strength of attraction to the foraging destinations. At each time step, the animal is offered a number of location options to move to next time step. The simulation calculates the distance between all options and the destination (i.e., a foraging location or shelter site), then normalises the distances between 0 and 1. The normalised values serve as weighting to increase the chances the animal will move towards its destination. We modify the weighting with the **destinationTransformation** and **destinationModifier** arguments. The **destinationTransformation** allows for weightings to be transformed (where 0 = no transformation, 1 = weights are square-rooted, 2 = weights are squared). Values supplied to **destinationModifier** apply a linear multiplicative effect to all the weightings. When these weightings increase in value the animal will exhibit a stronger attraction and choose more direct movements to its current destination.

We also store a value ready for the **rescale_step2cell** argument. The rescale value is a simple means of adjusting the environment layers to fit with the scale/unit in which the step lengths are provided. In this case, we treat each cell as 5m by 5m, thereby ensuring that our 2000x2000 matrix is sufficient for the badger to traverse without leaving. The rescale value therefore allows high resolution movements on lower resolution environments, offering a crucial memory saving optimisation.

```
BADGER_k_step <- c(0.3*60, 1.25*60, 0.25*60)
```

```
BADGER_s_step <- c(0.8, 0.25, 0.5)
```

```
BADGER_mu_angle <- c(0, 0, 0)
```

```
BADGER_k_angle <- c(0.6, 0.99, 0.6)
```

```
BADGER_destinationRange <- c(3, 120)
```

```
BADGER_destinationDirection <- c(0, 0.01)
```

```
BADGER_destinationTransformation <- 2
```

```
BADGER_destinationModifier <- 2
```

```
BADGER_rescale <- 5
```

The territoriality is not simulated directly via other individuals, instead we approximate it by supplying a number of point locations the badger will avoid. An alternative way of simulating this behaviour would be to create areas of high movement resistance to prevent the badger from entering. We provide a set of x, y coordinates of locations to be avoided. Alongside the locations, we need to define how strongly the badger will avoid them using `avoidTransformation` (distance to point is weighting squared) and `avoidModifier` (distance to point is weighting multiplied by 4). Both `avoidTransformation` and `avoidModifier` operate in the same way as the destination attraction arguments, but inverted; therefore, the avoidance behaviour will directly counteract the attraction to destination behaviour.

```
BADGER_avoidLocs <- data.frame(
  "x" = c(1205, 1500, 1165),
  "y" = c(980, 1090, 1250))

BADGER_avoidTransformation <- 2
BADGER_avoidModifier <- 4
```

We implement two cycles to capture the daily and seasonal cycles the badger experiences. Both cycles are created by describing a Sine wave defined by its amplitude, midline, offset (ϕ) and frequency (τ ; via the `cycle_draw` function). The `abm_simulate` function has two arguments for wave parameters: `rest_Cycle` is a mandatory input to describe diel activity cycling, and `additional_Cycles` to define any number of additive additional cycles. For `rest_Cycle` we will store a vector of length 4, with values defining the amplitude, midline, offset, and frequency in that order (Frequency is supplied in hours, i.e., 60 times the time step). We set the amplitude to 0.12 and midline as 0, meaning the probability of resting will be modified by values ranging from -0.06 to 0.06 depending on the location in the wave. Only a low amplitude wave is required to modify the balance between the behavioural state transition probabilities (Figure 1.1). Our badger example requires a 24 hour cycle where the probability of resting will peak and nadir, hence offset and frequency are set to 24 hours. When combined with the base transition probability to rest, the above parametrisation creates a scenario of approximately 8 hour periods of activity every 24 hours.

To add a seasonal cycle we use `additional_Cycles`. We predefine a data frame where each row contains four elements describing the amplitude, midline, offset, and frequency. Whereas our diel activity is defined by a 24 hour cycle, a seasonal cycle is 365 days and we shift the peak

half a year (offset = $365/2$). The seasonal wave increases the probability of resting behaviour during a portion of the year, while not also overwhelming a relatively stable diel cycle (i.e., badger will sleep longer during one part of the year) as we have set a lower amplitude for the seasonal cycle (0.075). A more extreme implementation of the second cycle could introduce hibernation behaviour.

We also pull forward the behaviour transition matrix we showed during the overview, and rename it so all simulation input objects have the `BADGER_` prefix.

```
BADGER_rest_Cycle <- c(0.12, 0, 24, 24)

# additional cycle
c0 <- c(0.075, 0, 24* (365/2), 24* 365) # seasonal

BADGER_additional_Cycles <- rbind(c0)
BADGER_additional_Cycles

##      [,1] [,2] [,3] [,4]
## c0 0.075    0 4380 8760

BADGER_behaveMatrix <- Default_behaveMatrix
BADGER_behaveMatrix

##      b0      b1      b2
## b0 0.9700 0.01000 0.0010
## b1 0.0002 0.95000 0.0008
## b2 0.0010 0.00001 0.9900
```

Running the simulation

We need to initially define a start location for our animal. To introduce some more individual variation, we can randomly vary this starting location, but we restrict the start locations to be proximal to the centre of the environment. For the simulation, we need a vector of length 2, where the first value is the x location, the second is y.

```
startLocation <- sample(900:1100, 2, replace = TRUE)
```

We finally have some parameters that describe the scope and intensity of the simulation.

The length of the simulation is provided via the `timesteps` argument; in our examples we are operating under the assumption that one time step is equal to one minute. Equally `timesteps` can be thought of as the number of movement choices simulated. We also need to specify how many options will be drawn and considered by the animal at each time step. The `options` argument is where this is input. Similarly `des_options` is where we specify the number of dynamically selected foraging attraction/destinations to choose from when it enters the foraging behaviour mode (state 1).

We can then call all our settings describing badger behaviour and movement settings we stored as objects previously, and run the simulation using `abm_simulate`.

```
simSteps <- 24*60 *365

simRes <- abm_simulate(
  # a data frame with x and y coordinates
  start = startLocation,
  # an integer describing the length of the simulation
  timesteps = simSteps,
  # an integer describing the number of foraging destination options an animal
  # is offered
  des_options = 10,
  # an integer describing the number of movement options an animal is offered
  options = 12,
  # a data frame providing x and y coordinates of the shelter locations
  shelterLocations = BADGER_shelterLocs,
  # a value describing the radius around shelter sites that movement step
  # lengths are reduced
  shelterSize = BADGER_shelterSize,
  # a data frame providing x and y coordinates of the avoidance locations
  avoidPoints = BADGER_avoidLocs,
  # a numeric vector of length two that contains the shape and scale values
  # that describe the Gamma distribution for potential foraging destinations
  destinationRange = BADGER_destinationRange,
  # a numeric vector of length two that contains the mean and concentration values
  # that describe the Von Mises distribution for potential foraging destinations
```

```

destinationDirection = BADGER_destinationDirection,
# a value to choose the type of transformation applied to the animal's
# attraction to a chosen destination
destinationTransformation = BADGER_destinationTransformation,
# a value modifying the animal's attraction to a chosen destination
destinationModifier = BADGER_destinationModifier,
# a value to chose the type of transformation applied to the animal's
# avoidance to avoidance locations
avoidTransformation = BADGER_avoidTransformation,
# a value modifying the animal's avoidance of avoidance locations
avoidModifier = BADGER_avoidModifier,
# a vector of three numbers describing the three behavioural states' Gamma
# distributions' shape parameter for step lengths
k_step = BADGER_k_step,
# a vector of three numbers describing the three behavioural states' Gamma
# distributions' scale parameter for step lengths
s_step = BADGER_s_step,
# a vector of three numbers describing the three behavioural states' Von Mises
# distributions' mean parameter for turn angles
mu_angle = BADGER_mu_angle,
# a vector of three numbers describing the three behavioural states' Von Mises
# distributions' concentration parameter for turn angles
k_angle = BADGER_k_angle,
# a numeric value to specify the size of the environmental matrices cells
rescale_step2cell = BADGER_rescale,
# a 3x3 numeric matrix describing the transition probabilities between the
# three behavioural states
behave_Tmat = BADGER_behaveMatrix,
# A vector of length 4 for amplitude, midline, offset and frequency to define the
# resting/active cycle
rest_Cycle = BADGER_rest_Cycle,
# A data.frame 4 columns wide for amplitude, midline, offset and frequency to
# define any additional activity cycles, where each row is another cycle

```



```

additional_Cycles = BADGER_additional_Cycles,
# Three arguments for the three matrices describing the landscape the
# simulated animal occupies
shelteringMatrix = BADGER_shelter,
foragingMatrix = BADGER_forage,
movementMatrix = BADGER_move)

```

1.4 Results

1.4.1 Output format

The simulation function (`abm_simulate`) outputs a list containing: (1) a dataframe of realised movements, (2) a dataframe of options the animal had available, and (3) a list of the inputs used to simulate the movement. The realised movement dataframe (`locations`) describes all realised locations the animal occupied, step length information, behavioural state, and current point of attraction, where each row is equal to a time step. Note the simulation is scale agnostic, so each row can represent a different time step to the example. Other than the lack of timestamps and location error, the format largely mirrors a typical movement dataset. Our example is minute by minute; changing the time step would require a reparametrisation of step length and behavioural transition probabilities. The `options` dataframe describes all the options available to the animal over the entire simulation duration, where each row is equal to an option repeated for each time step.

1.4.2 Outputs review

Once simulated, we can review the movement characteristics of the three species. The simplest to examine is the movement speeds or step lengths. During the simulation parametrisation we provided rescale values for the size of the cells describing environmental information (see `rescale_step2cell` argument). The simulated movement will return the scaled values, so to plot something comparable to the input values, we must rescale the simulated step lengths.

We can compare the simulation's inputs displayed in Figure 1.4 to the simulated/observed outputs and see that they largely agree as expected. The turn angles are more variable, as the

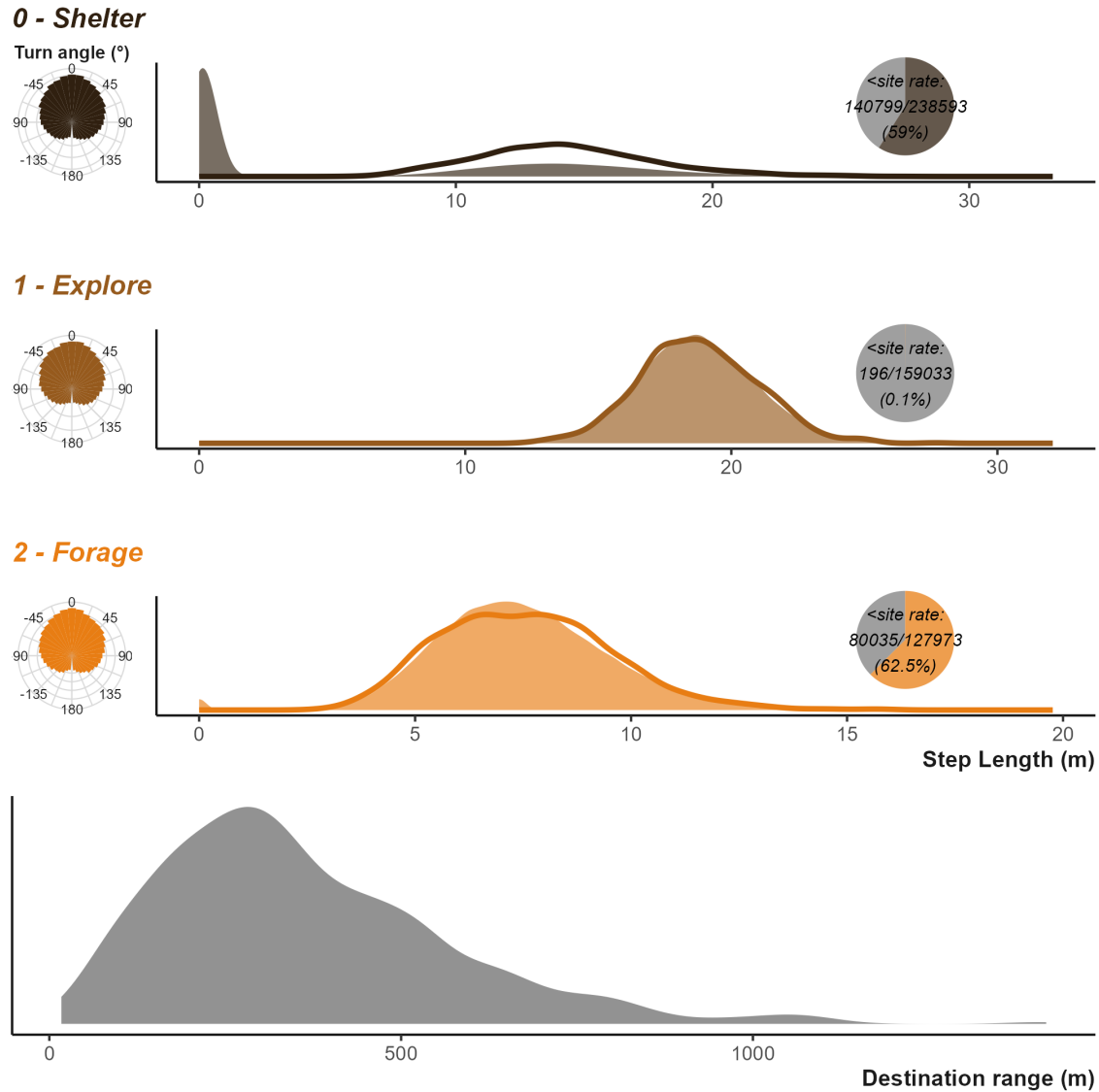


Figure 1.4: The distribution of step lengths used during the badger simulation example. Parametrising distributions are shown as solid areas using the same shape and scale parameters used as input into the simulation function. The lower plot shows the distribution used to generate potential foraging destinations. The realised resulting simulated step lengths (scaled to the input units) are shown by the solid lines. The realised resulting turn angles are shown in left hand sub-plots. Inset pie charts show the number of step lengths that were below the shelter site size. Note that x axis is not consistent between the three plots.

destination decisions and attraction to locations heavily influence the distribution overriding the initial parametrisation of the Von Mises distribution. Figure 1.11 and Figure 1.12 in the use cases section show the outputs for vulture and king cobra simulations.

We can review how the movements appear in space. With the badger example, the impact of the avoidance points is clearly visible (Figure 1.5a), whereas the lack of or weaker avoidance in vulture (Figure 1.5b) and king cobra (Figure 1.5c) makes the avoidance less influential. The vulture's movements and chosen foraging locations are largely to the east, demonstrating the impact of the underlying foraging quality environmental layer. The king cobra plot shows the impact of a near impermeable barrier truncating the southward movements.

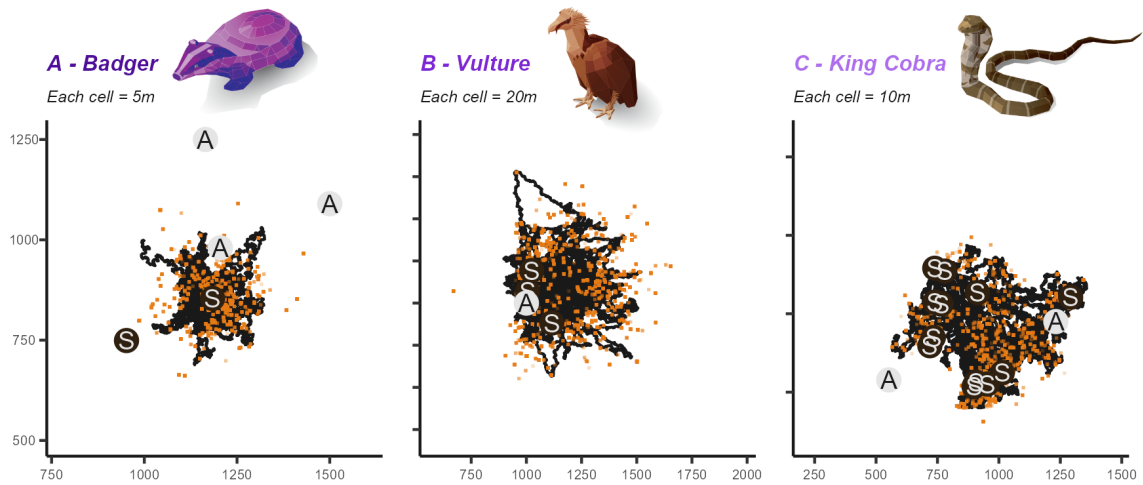


Figure 1.5: The observed locations of the three simulated species (A - Badger, B - Vulture, C - King Cobra). Black points show the observed location at each time step. The orange squares show the dynamically selected foraging destinations. Circles with an interior S show the shelter site locations, and circles with an interior A show the avoidance points. Note that the size represented by each unit on the x and y axis differs depending on the species.

The activity cycles and the timing of behavioural shifts is a key component in the simulations. We can examine that the predefined cycles result in expected patterns. A year of data makes observing all but the broadest cycles difficult, so in Figure 1.7 we can look at several months of data as well as the daily cycle. Badger and vulture daily cycles are largely the same (Figure 1.7a & b), with a consistent daily activity cycle differing only slightly in the time spent active and balance between shifts from sheltering to foraging and exploring. Some of the differences are a result of the different behavioural transition matrix provided to simulated the two species, where both had different baseline probabilities of shifting between behavioural states. By contrast, the king cobra example demonstrates the interaction between the daily and

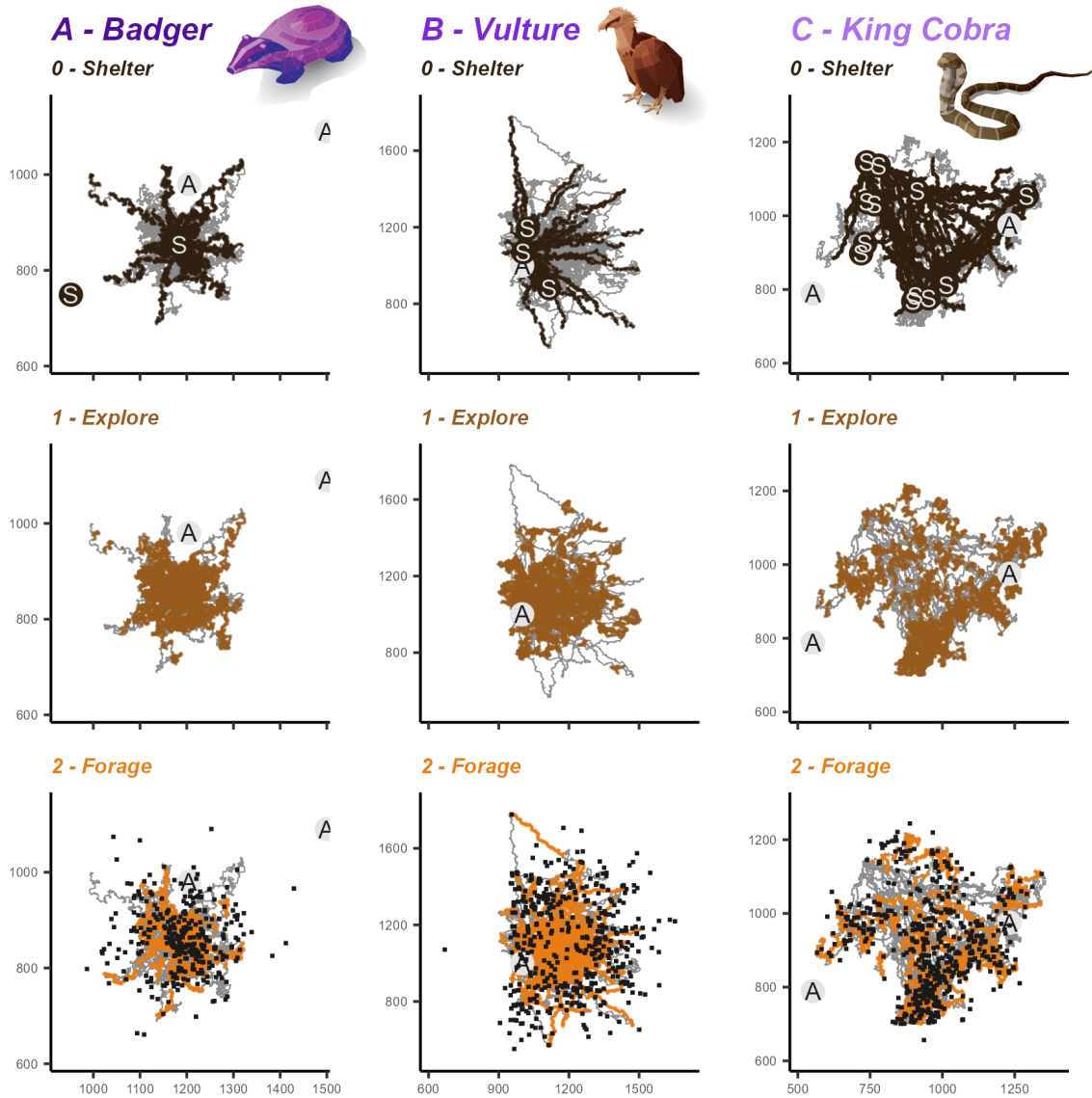


Figure 1.6: The simulated locations of the three simulated species (A - Badger, B - Vulture, C - King Cobra) over the first month of time steps. Grey path shows the overall movement during that month, overlaid points indicate where the animal was in a given behavioural mode. Circles with an interior S show the shelter site locations, and circles with an interior A show the avoidance points, black squares show the dynamically selected foraging destinations. Note that the size represented by each unit on the x and y axis differs depending on the species.

weekly cycles (Figure 1.7). We can see intermittently extended sheltering periods, punctuated by short exploratory or foraging bouts.

At the daily and monthly scale, we cannot see the impact of the broad scale seasonal cycles. Instead we can look at the percentage of time steps per day the animal was in sheltering behaviour (Figure 1.8). Again the badger and vulture sheltering rates are similar, differing in intensity, but with both demonstrating a seasonal decrease in the middle of the simulation. The king cobra cycles reveal a decrease in the number of days spent entirely sheltering and an overall impression that the seasonal cycle is less influential (as it is only one of three activity cycles).

1.5 Discussion

Animal movement datasets are complex and require a suite of analytical approaches to tackle satisfactorily. Efforts to develop new, and test existing, analyses would be aided by access to a range of diverse datasets. While *ideal* simulations often accompany new analysis methods and provide superb validation for the method in question, reaffirming the method’s robustness and pushing them to new limits with a messier, more stochastic, simulation approach could greatly strengthen our confidence in results. The *abmAnimalMovement* package provides an independent route to test new methods that covers a range of interacting movement features, not necessarily directly tied to a single analytical method.

By including a range of features linked to movement and behaviour, the *abmAnimalMovement* package can be implemented to investigate a suite of questions commonly asked of movement data –from habitat selection, to behaviour detection (for examples of common themes see Joo et al., 2022). While some of the features can appear simplistic, there remains ample flexibility to simulate a wide range of useful scenarios. For example, we conceptualise the three movement states as sheltering, exploring, and foraging. However, only several aspects are immutable: one state exhibits site fidelity (state 0), one state is free from all attraction (state 1), and one state is driven by an underlying environmental layer (state 2).

The *abmAnimalMovement* package has an advantage over data-driven simulation methods in scenarios where data are scarce, as much of the animal world is untracked (Crane et al., 2021; Joo et al., 2022). For the untracked animals, we may be limited to basic information of speed, activity, and resources, or such information may even need to be inferred from ecologically

A - Badger

Observed
behaviour
&
Rest prob.
modifier

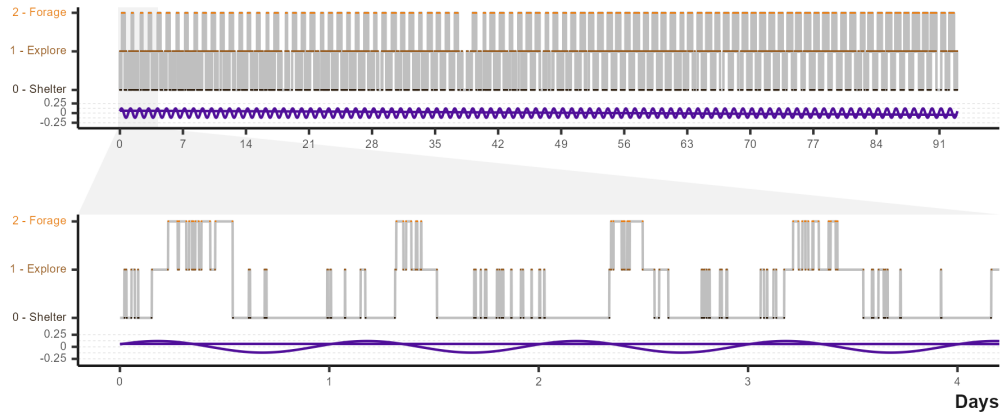
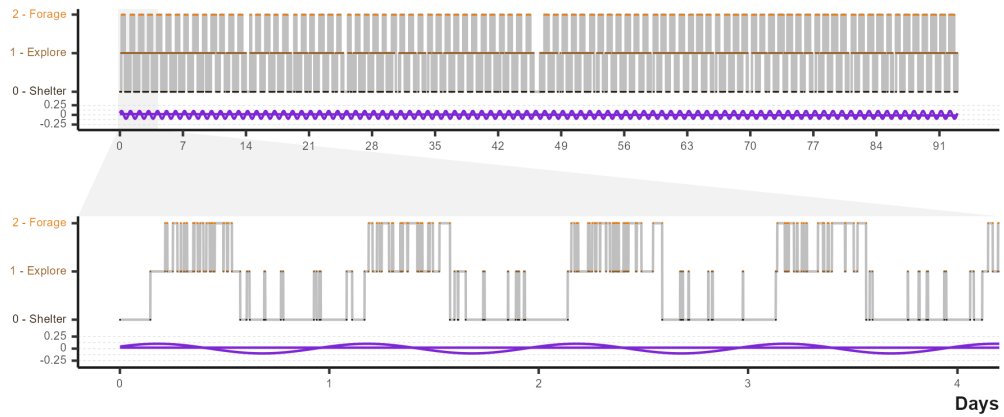
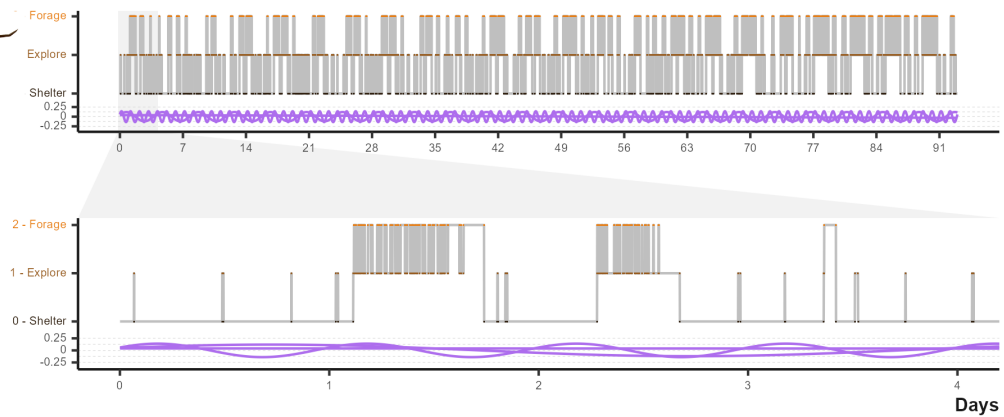
**B - Vulture****C - King Cobra**

Figure 1.7: The simulated states (top half) and parametrised activity cycles (bottom half) governing sheltering behaviour in the three example species (A - Badger, B - Vulture, C - King Cobra). Point colour and position describe the behavioural state at each simulated time step, whereas the purple waves indicated the input values. Note that the input waves acting on the simulated animal in conjunction with the behavioural transition matrix.



Figure 1.8: The percentage of time steps spent in state 0 - resting per day, and how it varies over the entire simulated year.

similar species. In such data starved situations, the *abmAnimalMovement* package's low computational cost and minimal data requirements allows for a large number of alternative parametrisations to be explored. Via the explorations of different parametrisations researchers can help build a picture of the study and analysis methods best suited for their questions, with the opportunity to test those analyses on synthetic simulated data.

In cases where data cannot be shared (e.g. due to concerns over species sensitivity), using a synthetic dataset could allow researchers to provide peer reviewers a dataset to test and error-check analysis code (Quintana, 2020). Providing data is a key component to ensuring computational reproducibility (Gerstner et al., 2017), and synthetic datasets provide an avenue to limit the reproducibility loss from data sharing restrictions. Additionally, such synthetic datasets could be integrated into preregistration as a means of demonstrating the validity of an analysis plan prior to undertaking a study.

Producing a range of simulated datasets that cover alternate scenarios may present researchers opportunities to test real data against a null model. For example, researchers looking to investigate whether an animal was avoiding a certain landscape feature could calibrate a number

of simulations covering a range of differing avoidance strengths. The simulated results could then be compared to the real data to gauge how different the real data was from simulations exhibiting zero avoidance (i.e., a null model scenario). This approach could complement current analysis methods, akin to sensitivity analysis.

1.6 Future directions

The *abmAnimalMovement* package provides adequate functionality to simulate a range of scenarios and movements. However, there are several aspects that will bear updating in future versions.

1. Dynamic state 0. While state 0 and the steady state of the attraction locations is key to simulating range stability (i.e., home range), there may be scenarios where this stability is not desired. For example, when simulating dispersal behaviour of juvenile or sub-adult animals there may be a desire to have shelter site dynamically chosen for a time. Currently such behaviour could be simulated, but it would require the dispersal to occur immediately, and the dispersal destination to be predefined (i.e., the sites for state 0 attraction). Therefore, in the current state the package may be limited in its ability to help predict possible dispersal destinations, but potentially capable of informing dispersal routes.
2. Autocorrelated speed. We may need to improve the autocorrelation of the animal's speed. Currently the speeds are non-independent based on the behavioural mode of the animal. The need to implement a more aggressive movement momentum/autocorrelative structure may be felt more acutely at different time frames, and for animals with a great variation in step lengths (i.e., a larger θ for the Gamma distribution). Explorations of simulated data using methods that measure autocorrelation in animal speed will reveal how much of a priority this should be.
3. Dynamic environment. All the environmental matrices are static; currently there is no system to update values during the simulation. This prevents shifts in the landscape such as seasonal variation in resources, or the development of trails. Currently, the closest solution is to run multiple simulations where the end location of simulation₁ is the start location for simulation₂, where simulation₂ is provided with a new season-appropriate resource layer. This solution would be inadequate for trail development, as

trail development would require a system within the simulation to update previously used cells for the animal at each time step.

1.7 Use cases

This section provides alternative parametrisations for the secondary examples to better demonstrate the range of movement types and scenarios the *abmAnimalMovement* package can replicate. We provide example implementations of vulture and king cobra -like movement.

1.7.1 Ecology and objectives - vulture

Unlike badgers and other terrestrially moving animals, vultures can move great distances with minimal obstruction (Figure 1.9). Vultures can also move greater distances more rapidly (Hribsek et al., 2021), resulting in a more variable and distribution of step lengths (Margalida et al., 2016; García-Jiménez, Pérez-García & Margalida, 2018; Subedi et al., 2020) (Figure 1.11).

Similar to badgers, vultures exhibit significant site fidelity, re-using roosting and nesting sites (Bracis, Bildstein & Mueller, 2018). Such shelter sites could be predefined; for example, if shelter sites were known *a priori* discovered via the capture and tagging of animals, which in the case of birds is more likely.

Vultures also offer an opportunity to demonstrate how the underlying resource availability impacts the movements of animals. In the case of vultures, their moments have been seen to follow carcasses, creating starkly contrasting areas where vultures will and will not travel (Arrondo et al., 2018) (Figure 1.9).

Vultures have a very similar cycle pattern to the badgers (just inverted), one defined by a standard 12 hours of activity during the day, and 12 hours of increased resting behaviour during the night. The importance of simulating such cycles is made clear by vultures studies demonstrating that day-night cycles can impact rates of location collection (Silva et al., 2017). How the the animals' activity cycle and the probability of collected data interact could be a key consideration for some research questions. Again, similar to badgers we would expect seasonal shifts in the form of an increase and decrease in activity depending on the time of year (García-Jiménez, Pérez-García & Margalida, 2018; Hribsek et al., 2021; Peshev et al.,

2021).

We can broadly summarise the vulture ecology we want to parametrise as follows: 1. Medium site fidelity via the use of multiple roosting/resting sites (Bracis, Bildstein & Mueller, 2018), 2. Movement speed approximated by summary statistics from previous studies (Margalida et al., 2016; García-Jiménez, Pérez-García & Margalida, 2018; Subedi et al., 2020; Hribsek et al., 2021), with minimal landscape derived resistance, 3. A 8-12 hour activity cycle, that shifts over the year.

1.7.2 Ecology and objectives - king cobra

Whereas the previous two species have a very limited or single shelter sites, king cobras make use of a wider range of shelter sites distributed more widely over their home ranges (Marshall et al., 2019, 2020a). These sites can also be larger, comprising of burrow systems or rock complexes.

What more dramatically sets king cobras, and other snakes, apart is a vastly differing rest-forage cycle. While snakes still exhibit a diel cycle, the intermittent depredation of large prey items and the time required sheltering to digest large meals results in a second broader activity cycle operating over a more widely observed diel cycle (Siers, Yackel Adams & Reed, 2018; Marshall et al., 2020a; D'souza et al., 2021; Smith et al., 2021). We can conceptualise this pattern as two additive cycles, one that will describe the daily activity cycle, and a second that describes the foraging and digestion cycle. Seasonality also impacts king cobra activity. As king cobras occupy tropical regions, the seasonality they experience is not as pronounced as the badger or vulture examples. Overall king cobras have three activity cycles acting on three different scales.

In this example we can also demonstrate the movement resistance dramatically impacting movement possibilities. King cobra movement can be limited by roads (Marshall et al., 2018; Jones et al., 2022), where unless provided with crossing structures king cobras are vulnerable to vehicle hits (Figure 1.10). In addition to roads presenting linear barriers across the landscape, king cobras face persecution when near or in human settlements (Shankar et al., 2013a; Marshall et al., 2018). Despite the risks, avoidance of such areas appears weak (Marshall et al., 2018).

Finally, king cobra movement characteristics will be dramatically reduced compared to the

vulture's, but with similar shape to the badger with greater variability (Figure 1.12) as king cobras are known to range over large areas (Silva et al., 2018; Marshall et al., 2019, 2020a).

We can broadly summarise the king cobra ecology we want to parametrise as follows: 1. Lower site fidelity via the use of many shelter sites (Marshall et al., 2019, 2020a), 2. Movement speed approximated by summary statistics from previous studies (Silva et al., 2018; Marshall et al., 2019, 2020a), with examples of very high landscape derived resistance (Marshall et al., 2018; Jones et al., 2022), 3. A 8-12 hour activity cycle, with a approximately weekly forage-digest cycle, and weak seasonality (Siers, Yackel Adams & Reed, 2018; Marshall et al., 2020a; D'souza et al., 2021; Smith et al., 2021).

Simulation inputs

Vulture inputs Largely the vulture parametrisation matches the badger. We specified a number of predefined resting/shelter sites, assuming that bird capture and bio-tagging is more likely to occur at known sites (i.e., nests or roosts). The shelter site size is smaller than the badger sett, so is only 5 m. Step length parameters are all larger and more variable, along with destination ranges (Figure 1.11). We set a larger rescale value also to ensure the vulture has a larger landscape to operate across. The rest cycle is broadly similar, with a 12 hour diel cycle and gentle seasonality. We made two small modifications to the transition matrix allowing for more frequent switches between foraging and exploring. We also alter two of the environmental matrices. We maximise movement ease across the entire landscape, while also lowering the foraging quality for an area to the West reflecting lower resources in that area (Figure 1.9). Finally, we specify that the vulture will not be avoiding any locations; we still require inputs for the simulation, but we can provide zeroes to negate the effect.

```
# predefined shelter sites
VULTURE_shelterLocs <- data.frame(
  "x" = c(1024, 1005, 1115),
  "y" = c(1193, 1070, 882))
# shelter site radius to reduce movements
VULTURE_shelterSize <- 5

# parameters defining Gamma (step length) and Von Mises (turn angle) for each
# behaviour state
```

```
# Gamma shape
VULTURE_k_step <- c(2, 2.2*60, 1.5*60)

# Gamma scale
VULTURE_s_step <- c(40, 1.2, 1)

# Von Mises mean
VULTURE_mu_angle <- c(0, 0, 0)

# Von Mises concentration
VULTURE_k_angle <- c(0.6, 0.99, 0.6)


# parameters defining Gamma (shape and scale) and Von Mises (mean and
# concentration) for foraging destination options
VULTURE_destinationRange <- c(50, 120)
VULTURE_destinationDirection <- c(0, 0.01)
# the transformation (2 = squared) and strength of attraction to destinations
VULTURE_destinationTransformation <- 2
VULTURE_destinationModifier <- 2


# each cell of the environmental matrix is 20x20
VULTURE_rescale <- 20


# the diel resting/active cycle: amplitude, midline, offset, frequency
VULTURE_rest_Cycle <- c(0.1, 0, 24, 24)
# additional cycles as a dataframe: amplitude, midline, offset, frequency
c0 <- c(0.025, 0, 24* (365/2), 24* 365) # seasonal
VULTURE_additional_Cycles <- rbind(c0)


# select modifications to the default behavioural transition matrix
VULTURE_behaveMatrix <- Default_behaveMatrix
VULTURE_behaveMatrix[2,3] <- 0.0002
VULTURE_behaveMatrix[3,2] <- 0.000015


# maximising movement ease by changing the default movement matrix
VULTURE_movementMatrix <- landscapeLayersList$movement
```

```

VULTURE_movementMatrix[] <- 1

# reducing foraging quality in the West of the landscape
VULTURE_forageMatrix <- landscapeLayersList$forage
VULTURE_forageMatrix[1:950,1:2000] <- VULTURE_forageMatrix[1:950,1:2000] - 0.6
VULTURE_forageMatrix[VULTURE_forageMatrix[] < 0] <- 0

# place holder avoidance location
VULTURE_avoidLocs <- data.frame(
  "x" = c(1000),
  "y" = c(1000))

# vulture has zero avoidance of the place holder point
VULTURE_avoidTransformation <- 0
VULTURE_avoidModifier <- 0

```

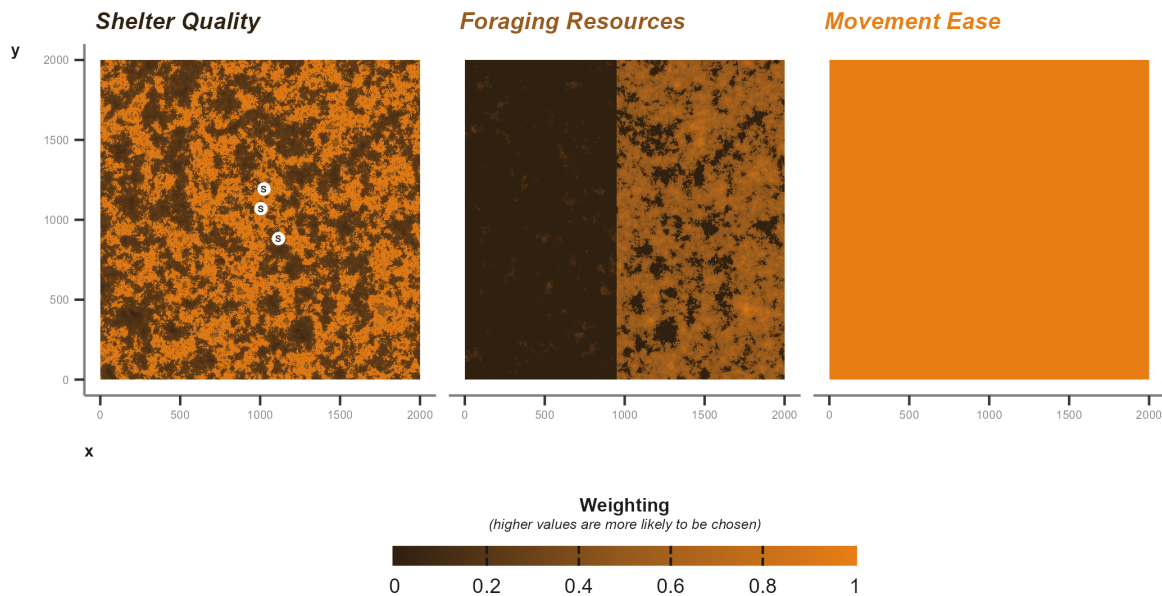


Figure 1.9: The three resulting landscape layers to be fed into the simulation for the vulture example: shelter quality, foraging resources, movement ease.

King cobra inputs Compared to the vulture, the king cobra movements are smaller, but king cobras can still occupy large areas (Figure 1.12). We provide a rescale value of 10 m to provide a large enough landscape to capture a potentially large range. The activity cycle provides an opportunity to demonstrate three additive cycles. The diel cycle is largely

similar to the other examples, but we also define an approximately weekly cycle to cover the forage-digestion cycle as well as weak seasonal cycle. Combined the three cycles create an activity pattern quite different from either the badger or vulture. To balance the three cycles and their influence on the behaviour transitions, we make some minor adjustments to the underlying behavioural transition matrix. As mentioned in the ecology and objectives, we want to simulate a barrier limiting movement. We construct this by altering the values in the environmental matrices, minimising the weighting in cells following two perpendicular lines (Figure 1.10). We also provide a set of avoidance locations and a relatively weak avoidance weighting coefficient. Finally, we can set the shelter sites. We draw 12 sites randomly based on shelter site quality, and limit the area that they can be drawn from to ensure they are not on the far side of the barrier. We use a larger shelter site size, as king cobras are known to occupy large rock complexes at times.

```
# parameters defining Gamma (step length) and Von Mises (turn angle) for each
# behaviour state
# Gamma shape
KINGCOBRA_k_step <- c(30, 40, 20)
# Gamma scale
KINGCOBRA_s_step <- c(0.75, 1.2, 1.75)
# Von Mises mean
KINGCOBRA_mu_angle <- c(0, 0, 0)
# Von Mises concentration
KINGCOBRA_k_angle <- c(0.6, 0.99, 0.6)

# parameters defining Gamma (shape and scale) and Von Mises (mean and
# concentration) for foraging destination options
KINGCOBRA_destinationRange <- c(50, 10)
KINGCOBRA_destinationDirection <- c(0, 0.01)
# the transformation (2 = squared) and strength of attraction to destinations
KINGCOBRA_destinationTransformation <- 2
KINGCOBRA_destinationModifier <- 1.5

# each cell of the environmental matrix is 10x10
KINGCOBRA_rescale <- 10
```

```

# the diel resting/active cycle: amplitude, midline, offset, frequency
KINGCOBRA_rest_Cycle <- c(0.14, 0, 24, 24)

# additional cycles as a dataframe: amplitude, midline, offset, frequency
c0 <- c(0.12, 0, 24, 24*4) # digestion
c1 <- c(0.05, 0, 24 * (365/2), 24* 365 ) # seasonal
KINGCOBRA_additional_Cycles <- rbind(c0, c1)


# select modifications to the default behavioural transition matrix
KINGCOBRA_behaveMatrix <- Default_behaveMatrix
KINGCOBRA_behaveMatrix[1,1] <- 0.95
KINGCOBRA_behaveMatrix[1,2] <- 0.005
KINGCOBRA_behaveMatrix[3,1] <- 0.00025
KINGCOBRA_behaveMatrix[3,2] <- 0.000001
KINGCOBRA_behaveMatrix[3,3] <- 0.999


# extracting the default matrices ready for changes
KINGCOBRA_shelteringMatrix <- landscapeLayersList$shelter
KINGCOBRA_forageMatrix <- landscapeLayersList$forage
KINGCOBRA_movementMatrix <- landscapeLayersList$movement

# defining the start and end of strong intersections hampering movement
roadMin_x <- 1360
roadMax_x <- roadMin_x + 40
roadMin_y <- 660
roadMax_y <- roadMin_y + 40


# applying the change, dramatically reducing the weighting for all matrices
KINGCOBRA_shelteringMatrix[roadMin_x:roadMax_x,1:2000] <-
  KINGCOBRA_shelteringMatrix[roadMin_x:roadMax_x,1:2000] - 90
KINGCOBRA_shelteringMatrix[1:2000,roadMin_y:roadMax_y] <-
  KINGCOBRA_shelteringMatrix[1:2000,roadMin_y:roadMax_y] - 90
KINGCOBRA_shelteringMatrix[!KINGCOBRA_shelteringMatrix >= -99.9] <- -99

```

```

KINGCOBRA_forageMatrix[roadMin_x:roadMax_x,1:2000] <-
  KINGCOBRA_forageMatrix[roadMin_x:roadMax_x,1:2000] - 90
KINGCOBRA_forageMatrix[1:2000,roadMin_y:roadMax_y] <-
  KINGCOBRA_forageMatrix[1:2000,roadMin_y:roadMax_y] - 90
KINGCOBRA_forageMatrix[!KINGCOBRA_forageMatrix >= -99.9] <- -99

KINGCOBRA_movementMatrix[roadMin_x:roadMax_x,1:2000] <-
  KINGCOBRA_movementMatrix[roadMin_x:roadMax_x,1:2000] - 90
KINGCOBRA_movementMatrix[1:2000,roadMin_y:roadMax_y] <-
  KINGCOBRA_movementMatrix[1:2000,roadMin_y:roadMax_y] - 90
KINGCOBRA_movementMatrix[!KINGCOBRA_movementMatrix >= -99.9] <- -99

# defining a number of avoidance points
KINGCOBRA_avoidLocs <- data.frame(
  "x" = c(552, 1232, 1587),
  "y" = c(789, 975, 1356))

# and specifying a weak avoidance of the points
KINGCOBRA_avoidTransformation <- 2
KINGCOBRA_avoidModifier <- 1

# after the matrices have been altered, drawing 12 shelter sites
# based on shelter quality
sampledShelters <- sampleRandom(raster(KINGCOBRA_shelteringMatrix), 12,
                                ext = extent(0.35, 0.65, 0.42, 0.65),
                                rowcol = TRUE)

# and storing those shelter sites as a dataframe ready for the
# simulation function
KINGCOBRA_shelterLocs <- data.frame(
  "x" = sampledShelters[,2],
  "y" = sampledShelters[,1])

# specifying a larger shelter site radius
KINGCOBRA_shelterSize <- 10

```

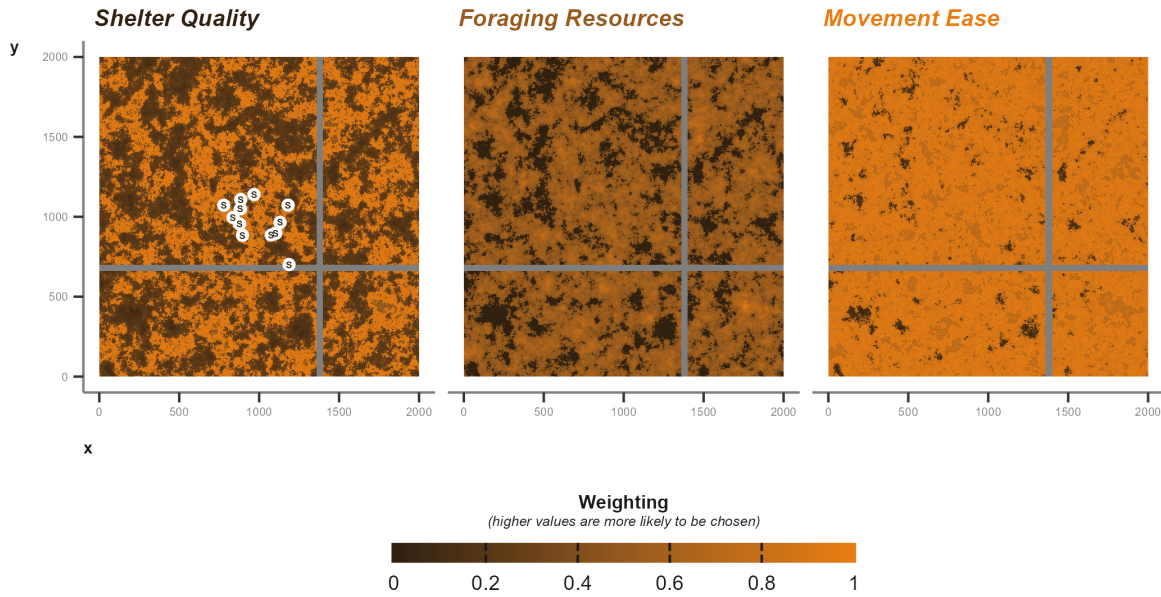



Figure 1.10: The three resulting landscape layers to be fed into the simulation for the king cobra example: shelter quality, foraging resources, movement ease. Movement ease is blocked by bars of -99 weighting.

Simulation outputs

Figures 1.11 and 1.12 describe the inputs for the vulture and king cobra example respectively, and can be compared to the realised movements. Figures 1.11, and 1.12 also provide information on the rates of stationary behaviour, defined in the plot as step lengths less than the shelter site size. The king cobra example in particular highlights the prolonged near weekly resting periods.

1.7.3 Data availability

All data underlying the results are generated as part of the article, no source data are required.

The example badger, vulture, and king cobra data simulated for this manuscript are available on Zenodo: “Simulated data from abmAnimalMovement: An R package for simulating animal movement using an agent-based model” <https://doi.org/10.5281/zenodo.6992496>. Files include:

- BADGER_locations.csv: A csv file that contains the realised locations of the example

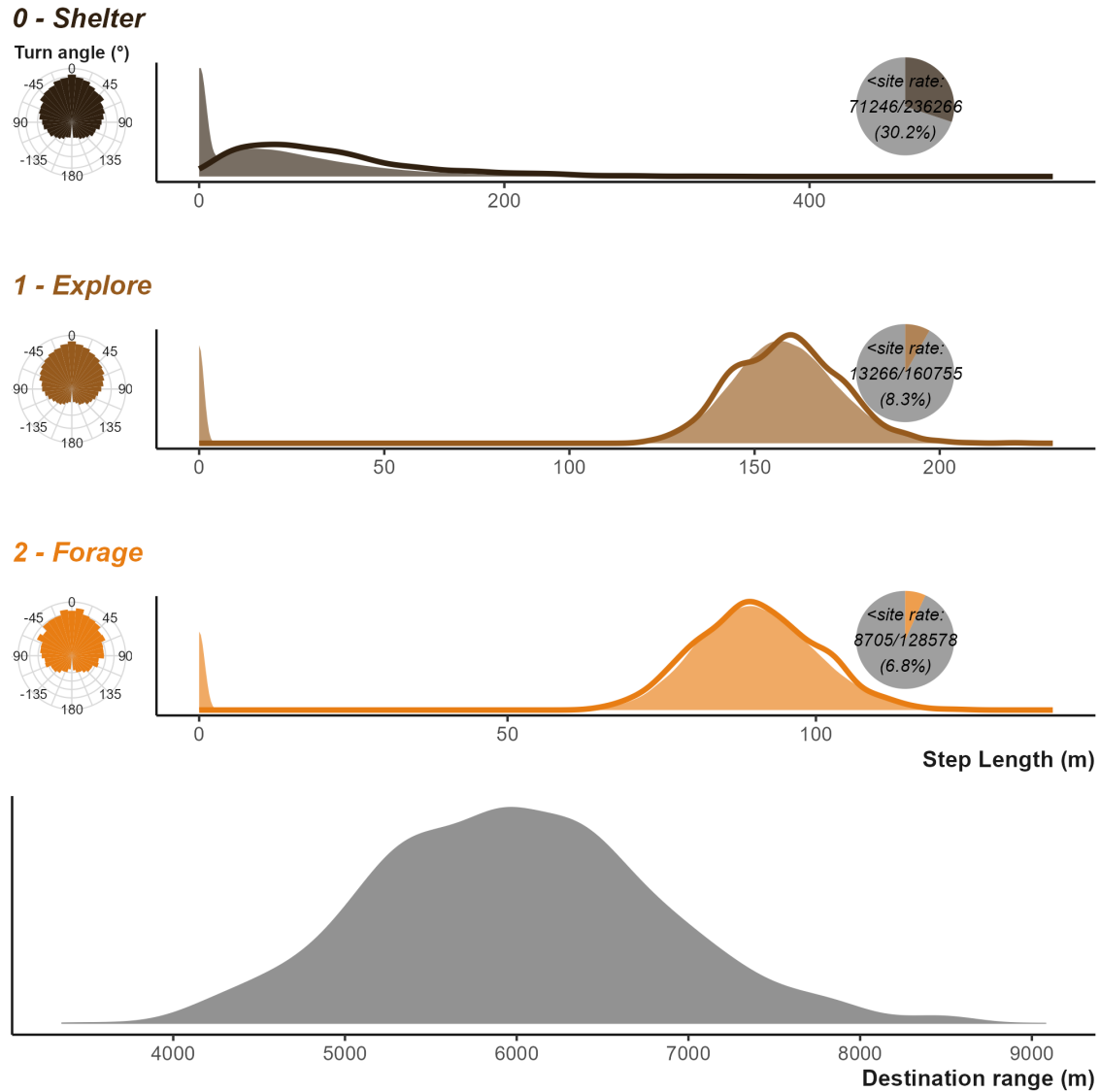


Figure 1.11: The distribution of step lengths used during the vulture simulation example. Parametrising distributions are shown as solid areas using the same shape and scale parameters used as input into the simulation function. The lower plot shows the distribution used to generate potential foraging destinations. The realised resulting simulated step lengths (scaled to the input units) are shown by the solid lines. The realised resulting turn angles are shown in left hand sub-plots. Inset pie charts show the number of step lengths that were below the shelter site size. Note that x axis is not consistent between the three plots.

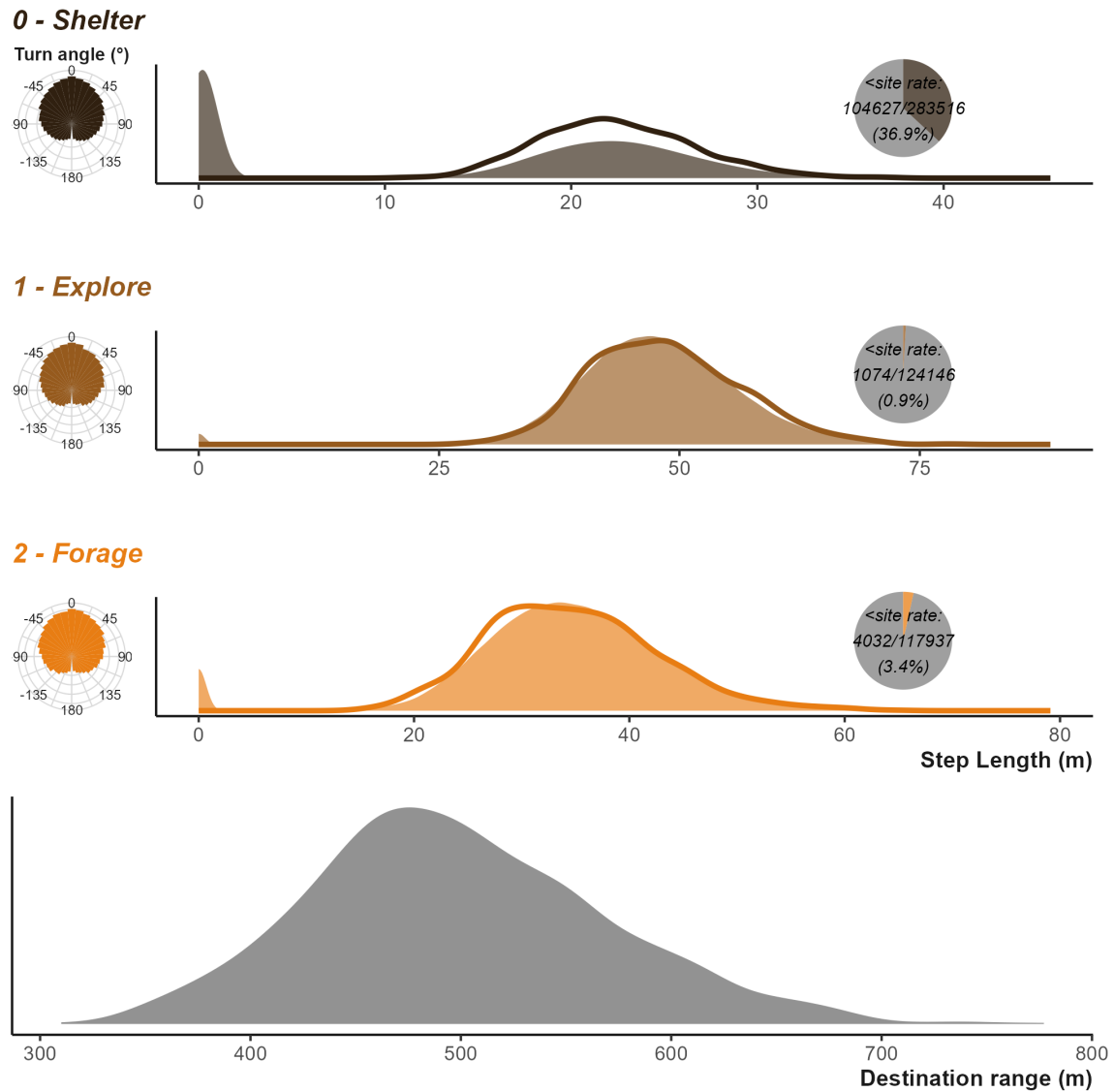


Figure 1.12: The distribution of step lengths used during the king cobra simulation example. Parametrising distributions are shown as solid areas using the same shape and scale parameters used as input into the simulation function. The lower plot shows the distribution used to generate potential foraging destinations. The realised resulting simulated step lengths (scaled to the input units) are shown by the solid lines. The realised resulting turn angles are shown in left hand sub-plots. Inset pie charts show the number of step lengths that were below the shelter site size. Note that x axis is not consistent between the three plots.

badger simulation, where each row is equal to a timestep. Columns include: timestep, the timestep as an integer; x, the x coordinate of the animal; y, the y coordinate of the animal; sl, the step length between locations used during the simulation; sl_rescale the rescale factor required to return step lengths back to the input scale; ta, turning angle between locations in degrees; behave, the behavioural mode the animal was in at a given timestep; chosen, the location chosen out of the number of options available; destination_x and destination_y the point the animal was attracted to at that time (note exploratory behaviour is not subject attraction).

- `BADGER_options.csv`: A csv file that contains the options available to the example badger simulation over the entire simulation duration, where each row is equal to an option repeated for each timestep. Columns include: timestep, the timestep as an integer; oall_x, and oall_y show the x and y coordinates of all the options available to an animal at a timestep; oall_steplengths are the step lengths from the current location compared to all the options.
- `completelist.RDS`: This RDS file contains a list object of length three, where the full simulation outputs from each three examples are stored. Each species slot contains the “locations” dataframe (see description of `locations.csv`), the “options” dataframe (see description of `options.csv`), and a nested list containing all the “inputs” used to generate the simulated results (split into subsections: `inputs_basic` that contains inputs linked to simulation duration and intensity, `inputs_destination` that contains inputs linked to destination and attraction aspects, `inputs_movement` that contains inputs linked to movement capacity and behavioural switching, `inputs_cycle` that contains inputs linked to activity cycling, `inputs_layerSeed` that contains the environmental matrices and seed). A fourth object is returned called “others” that captures all other outputs, mainly used internally for debugging and checking.
- `KINGCOBRA_locations.csv`: A csv file that contains the realised locations of the example king cobra simulation, where each row is equal to a timestep. The file structure follows the same as the `BADGER_options.csv` file. `KINGCOBRA_options.csv`: A csv file that contains the options available to the example king cobra simulation over the entire simulation duration, where each row is equal to an option repeated for each timestep. The file structure follows the same as the `BADGER_options.csv` file.
- `VULTURE_locations.csv`: A csv file that contains the realised locations of the example

vulture simulation, where each row is equal to a timestep. The file structure follows the same as the `BADGER_locations.csv` file. `VULTURE_options.csv`: A csv file that contains the options available to the example vulture simulation over the entire simulation duration, where each row is equal to an option repeated for each timestep. The file structure follows the same as the `BADGER_locations.csv` file.

Source code to reproduce the results and manuscript are available alongside the software, archived at: <https://doi.org/10.5281/zenodo.6951937>. The parameters used to generate the examples are within: `notebook/manuscript/Agent-based_model_walkthrough.Rmd`.

1.7.4 Software availability

Archived versions of the software are available on Zenodo: “BenMMarshall/abmAnimalMovement” <https://doi.org/10.5281/zenodo.6951937>.

The development version of the package can be found at [BenMMarshall/abmAnimalMovement](https://github.com/BenMMarshall/abmAnimalMovement).

License: GPL (≥ 3)

1.7.5 Competing interests

No competing interests were disclosed.

1.7.6 Grant information

BMM was funded by the Natural Environment Research Council (NERC) via the IAPETUS2 Doctoral Training Partnership.

1.7.7 Author contributions

Conceptualization: BMM; Software: BMM; Supervision: ABD; Visualization: BMM; Writing - Original Draft: BMM; Writing - Review & Editing: BMM, ABD.

Chapter 2

A Habitat Selection Multiverse Reveals Largely Consistent Results Despite a Multitude of Analysis Options

Benjamin M. Marshall¹, Alexander B. Duthie¹

¹Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, FK9 4LA, UK

Chapter is available as a preprint: Marshall, B. M., & Duthie, A. B. (2024). A Habitat Selection Multiverse Reveals Largely Consistent Results Despite a Multitude of Analysis Options. bioRxiv, 2024-06. <https://doi.org/10.1101/2024.06.19.599733>

2.1 Abstract

Researchers are intrinsically part of the research process. While we may strive for objectivity, there are always judgement calls required during research. When you ask ten researchers to answer the same question with the same dataset, you will likely receive ten different answers. This variation in answers has been linked to several disciplines' replication crises. Here, we explore whether answers from movement ecology, specifically habitat selection, vary as a result

of differing analytical choices. We conducted a multiverse analysis on around 400 synthetic animal movement datasets, exploring a multitude of analysis pathways to determine habitat selection, resulting in approximately half a million unique estimates of selection. By using simulated virtual animals with a known preference, we were able to show which decisions during analysis could lead to more variable estimates of habitat selection. The multiverse revealed that data quantity (i.e., tracking frequency and duration) was more important to obtaining consistent answers than any analysis choice. Overall, the pattern of estimates shows the majority of analysis pathways provide similar final results, particularly for modern analysis methods. The pattern reflects findings from other disciplines, indicating that while movement ecology is not immune to issues of non-replicability stemming from researcher choice, it is also not at any greater risk than other disciplines.

2.2 Introduction

Researchers are inextricably part of the research process (Levins & Lewontin, 1985; Tang-Martínez, 2020), and our expectations can shape our conclusions (Holman et al., 2015). While we can strive to conduct research objectively, research commonly requires judgement calls (Steege et al., 2016). Such choices occur throughout the research process, encompassing everything from study design (e.g., sample size, sampling intensity, sample stratification) to analysis (e.g., Bayesian or frequentist, handling of outliers). We draw on our own experience, and the input of peers, to try and ensure the best decisions are made to produce robust and reliable results, but we also contend with data, expertise, and interpretability constraints (Liu, Althoff & Heer, 2020).

Researchers are also influenced by the incentive system around them (Anderson et al., 2007; Receveur et al., 2024). We cannot undertake research in a vacuum; we often require institutions, funders, and scientific journals to produce and share research. These bodies can influence what research is conducted; they are in a position to incentivise or disincentivise research of certain topics or methodologies (Fanelli, 2010a; Ware & Munafò, 2015; Smaldino & McElreath, 2016). The use of impact factor (and similar metrics) is an example of citations being used as a measure of quality or research worth. But when examined closer, the impact factor appears detached from the robustness or reliability of the research (Brembs, 2018). A decrease in robustness can be seen in the increases in effect size inflation, p-value misreporting, and other measures of quality (Brembs, 2018). Similarly, novelty has been promoted

by journals to the detriment of replication (Vinkers, Tjldink & Otte, 2015; Forstmeier, Wagenmakers & Parker, 2017; Brembs, 2019), despite widespread agreement on the importance of replications (Fraser et al., 2020). There is a bias towards positive, statistically significant results (Jennions & Møller, 2002b; Cassey et al., 2004). Due to the nature of statistical significance, a prioritisation of significant results can elevate underpowered studies and boost false positive rates (Forstmeier & Schielzeth, 2011; Albers, 2019; Clements et al., 2022).

Unfortunately, there is evidence that the systems of incentives trickle down to impact the decisions of researchers while undertaking research and publishing results. The more detrimental of these decisions have been termed questionable research practices (Fraser et al., 2018b; Bishop, 2019). Questionable research practices can be broadly viewed as methods to achieve a neat, statistically significant, and publishable narrative (O’Boyle, Banks & Gonzalez-Mulé, 2014). In the worst cases, narratives can be prioritised over transparently reporting results.

There is a fear that questionable research practices, and the broader incentives they are connected to, are responsible for the replicability crisis (also referred to as the reproducibility crisis). Across many disciplines, there are examples of replication studies being unable to replicate prior research (Freedman, Cockburn & Simcoe, 2015; Open Science Collaboration, 2015; Kelly, 2019; Clements et al., 2022). Often these replication efforts are conducted with larger sample sizes, or rely on the consolidation of many independent studies (often in the form of meta-analyses). The implication is not that the original studies were necessarily flawed; but, in the absence of questionable research practices, sufficient variation exists in the study subjects or methods to obscure a consistent effect (i.e., variation other than that stemming from sampling alone Simonsohn, 2015).

However, variation can also stem from analysis flexibility; i.e., the presence of many reasonable ways to analyse the same data to answer the same question. This flexibility helps enable questionable research practices by providing a suite of answers to choose between (Fraser et al., 2018b), and is potentially steered by publication bias (Jennions & Møller, 2002b; Cassey et al., 2004) if more publishable results are prioritised or rewarded over less exciting but robust results. Given the prominence of questionable research practices and publication bias, the inconsistencies between initial and replication studies warrant investigation (especially when analysis flexibility is also implicated in potentially flawed replications Bryan, Yeager & O’Brien, 2019). Even in the absence of any undesirable incentives, analysis flexibility can still lead to variable results simply as the result of researchers considering approaches of differing

validities for a given dataset (Gelman & Loken, 2013; Gould et al., 2023).

Replications are a key tool for identifying potential sources of variation that lead to differing conclusions (e.g., subtle differences in methodologies or sampling). For some disciplines, such as ecology, replications can be difficult to justify due to the impacts on the study subjects and high monetary costs. Ecological systems are also complex and in constant flux, often frustrating perfect replications due to changes in space and time (Nakagawa & Parker, 2015; Schnitzer & Carson, 2016), leaving researchers with a distinct lack of direct replications in ecology (Kelly, 2019). The low prevalence of ecological replication studies makes it difficult to assess the overall rate of replication in ecology (Kelly, 2019), but there are several examples that suggest irreproducibility is something ecologists should be wary of (Wang et al., 2018; Sánchez-Tójar et al., 2018; Roche et al., 2020; Clark et al., 2020). The potential for irreproducibility is further supported by evidence of positive publication bias (Fanelli, 2010b, 2012), and links between smaller sample sizes and inflated effect sizes (Lemoine et al., 2016).

As direct replications can be difficult to conduct, ecology is often left to assess replicability via conceptual replications (Fraser et al., 2020) or efforts broadly referred to as quasi-replications (Palmer, 2000). Replications range in intensity. Direct (or exact) replications are attempts to replicate a tightly defined concept or hypothesis while duplicating all characteristics of the original study. Partial replications are a step looser, where the concept/hypothesis tested is less clearly defined (e.g., applicable to a broader scale), but efforts are made to repeat the same methodology. The most general category is conceptual replications, where the subject and method of study varies from the original study, but the replication targets the same concept or hypothesis (Nakagawa & Parker, 2015; Kelly, 2019). Both partial and conceptual can be classed as quasi-replications if the concept and scale are broadly defined (Nakagawa & Parker, 2015).

Conceptual replications are valuable, but rely on researchers' ability to compare replication efforts to previous findings. An important aspect of those comparisons is accounting for factors differing between the studies that are not salient to the effect of interest (Forstmeier, Wagenmakers & Parker, 2017), e.g., those linked to sampling differences (Simonsohn, 2015). An example of sampling differences leading to differences in results can be seen in the case of reptile space use. Silva et al. (2020) showed how the frequency at which a reptile was located (via radio-telemetry) by a researcher interacted with the space-use estimation method, leading to differences in area estimates even when using the same estimation method. This reveals

how the choices during analysis (e.g., choice of area estimation method) impact results, and how the uncertainty introduced by those choices change depending on the sampling approach. It presents a scenario where the *correct* choice is dependent on preceding decisions; thereby, highlighting the need to explore the impacts of multiple decisions simultaneously.

As seen in the reptile space use example, the choices available to the researcher (i.e., “researcher degrees of freedom,” Simmons, Nelson & Simonsohn, 2011) is a key source of variation among studies. Research degrees of freedom (or flexibility in analysis, Forstmeier, Wagenmakers & Parker, 2017) have been elegantly demonstrated by a number of “many analysts” studies (e.g., Silberzahn et al., 2018; Huntington-Klein et al., 2021). In these studies, a number of researchers, or research groups, are tasked with answering the same question. Naturally each participant takes a slightly different approach, both in how the question is interpreted (Auspurg & Brüderl, 2021), and the analysis approach chosen (Gelman & Loken, 2013; Bastiaansen et al., 2020; Gould et al., 2023), resulting in different final results. Such variation in results can be sufficient to change the final conclusions (Salis, Lena & Lengagne, 2021), or at least alter the strength of an estimated effect (Desbureaux, 2021).

2.2.1 Multiverse analysis

An approach to address the unknown impacts of undisclosed researcher degrees of freedom is to fully explore all plausible or reasonable analysis choices open to researchers, i.e., to explore a multiverse of design choices (Steege et al., 2016). This multiverse analysis—closely linked to vibration of effects (Patel, Burford & Ioannidis, 2015), multi-model analysis (Young & Holsteen, 2017), and specification curve analysis (Simonsohn, Simmons & Nelson, 2020)—entails running a multitude of different study and analysis designs. By comparing the various study pathways, a multiverse exploration has the potential to quantify the variation stemming from researchers’ analyses choices (Rijnhart et al., 2021). Choices can include everything from sample sizes and splits (e.g., Webb & Demeyere, 2021) to measurement and summary statistics (e.g., Parsons, 2020), but crucially should only include options that are reasonable (Simonsohn, Simmons & Nelson, 2020; Del Giudice & Gangestad, 2021). What counts as “reasonable” is not necessarily obvious, and inclusion of irrelevant choices can easily mask important choices because of the multiplicative nature of a branching path network (Del Giudice & Gangestad, 2021) (Fig. 2.1). Construction of a multiverse requires justification for which decisions are treated as variable, and why there is not an *a priori* and defensible

single solution (Del Giudice & Gangestad, 2021). A multiverse populated with well-justified decisions allows the exploration of which choices inflate variation among analysis universes, while also offering insights into how to deflate variation (e.g., refinement of initial study design, the removal of ambiguities like tightening categories definitions, Steegen et al., 2016).

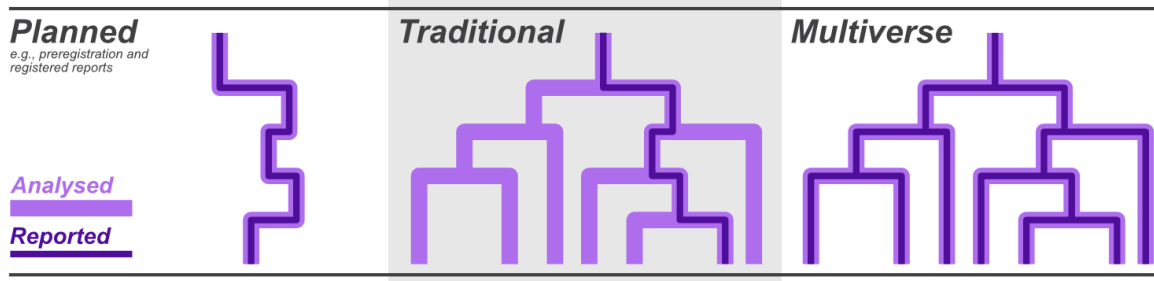


Figure 2.1: Diagram showing how multiverse analysis differs from other approaches. Each branch node represents a choice made during analysis. Based on Figure 1 from Dragicevic et al. (2019).

However, even if all choices are equally valid, multiverse analysis cannot simply provide a correct answer (Steegen et al., 2016). The “average” result is not necessarily the closest to the truth. If we were to undertake a multiverse analysis in a scenario with a “known truth”, i.e., using a simulated dataset (Bastiaansen et al., 2020), we may be able to identify the amount of variation attributable to different sources (e.g., biological variation vs study design variation, Breznau et al., 2021), and potentially the systematic biases stemming from specific choices. The use of simulated data is an established way to explore the robustness of methodologies (Minchin, 1987; Silva et al., 2020), and here we harness the benefits of simulated data to assess the impacts of researcher degrees of freedom on the results garnered from animal movement datasets.

As mentioned, ecological systems are complex to study and frustrate replication efforts (Nakagawa & Parker, 2015; Schnitzer & Carson, 2016); this can be especially true in the case of movement ecology and animal movement. Studying animal movement is often conducted using bio-telemetry/bio-logging devices that when attached to an animal record their location (Joo et al., 2022). From the resulting dataset of times and locations, researchers can derive a multitude of insights about the animal’s life, ranging from more superficial, such as movement characteristics (e.g., speed, turn angle), to more in-depth or derived such as home range or habitat selection. For the deeper insights a greater number of decisions are required during the analysis process, opening up the possibility that different researchers will take different pathways to answer the same question. With many forking pathways, analysis of movement

data to determine an animal's habitat selection presents a perfect setting for a multiverse exploration of how the researcher decisions impact the final results.

Here we construct a multiverse of analyses to explore how decisions concerning the design and analysis of animal tracking data can impact the findings on habitat selection. Using a branching set of sampling and analysis options, we examine which decisions lead to the most variation in estimated habitat selection for three distinct simulated animal species.

2.3 Methods

2.3.1 Simulating the Scenarios

We simulated three scenarios each represented by a different species, on different landscapes. We simulated the landscapes using the NLMR v.1.1.1 package (Sciaini et al., 2018c), and the animal movement using `abmAnimalMovement` v.0.1.3.0 package [Marshall & Duthie (2022); Chapter 1]. The `abmAnimalMovement` simulations are a discrete time, agent-based modelling approach for simulating animal movement. Further details of how the scenarios were parametrised can be found in Marshall & Duthie (2022), and also in the `abmAnimalMovement` github repository along with associated documentation: <https://github.com/BenMMarshall/abmAnimalMovement>, also archived: <https://doi.org/10.5281/zenodo.6951937>.

The simulated species are as follows:

- Species 1, Badger: site fidelity to two shelter sites, a low movement speed constrained by terrestrial environment and territoriality, an 8-12 hour activity cycle with seasonal shifts.
- Species 2, Vulture: medium site fidelity via the use of multiple roosting/resting sites, a high and variable movement speed with minimal landscape resistance, an 8-12 hour activity cycle with seasonal shifts.
- Species 3, King Cobra: lower site fidelity making use of many shelter sites, a medium movement speed through a landscape with high resistance barriers, an 8-12 hour activity cycle combined with an approximately weekly forage-digest cycle and a weak seasonal cycle.

Each simulation generated a year's worth of movement data with a data point recorded every minute, and we simulated 4 individuals per species. The full parametrisation of the simulated species can be found in the code associated with this paper, specifically the `simulate_individual.R` function (available at: <https://github.com/BenMMarshall/multiverseHabitat>, and archived at: <https://doi.org/10.5281/zenodo.12169335>).

The agent-based animal movement simulations require an environment to guide movement. Our landscapes are defined as three stacked matrices, where the rows and columns of the matrix represent the X and Y locations in the landscape, and the value in each cell numerically describes the quality of a given resource the animal can consider when deciding where to move. Each of the three matrices' describe one resource: foraging quality, shelter site quality, or movement ease (Fig. 2.2), which are evaluated by the animal differently depending on the behavioural state the animal is in (resting, exploring, foraging). All three are based on a single initial random generation using a Gaussian field.

From the initial independent Gaussian fields, we altered the values to exaggerate the difference between high resource areas and low value areas; thereby generating non-independence of resources. Broadly, we created: core resource areas with higher foraging quality but lower movement ease, edge areas that overlap with the core resource area with better movement ease, and more barren areas with high movement ease but with minimal foraging value. All shelter sites occurred within the higher quality core resource areas.

While our simulation landscapes used to generate the simulated animal movement comprise of continuous values that affect the probability of a simulated animal using a cell or not, many habitat selection studies use categorical habitat metrics or land use types. Therefore, to simplify the interpretation of the multiverse, we simplified our landscape into three distinct categories for analysis (Fig. 2.2): a low quality area where resources are low but movement is easy (class 0 in the figure), an edge habitat with middling resources and high movement ease (class 1), and a high resource habitat with low movement ease (class 2). This results in a classified matrix/raster of habitat types usable for downstream analyses (Fig. 2.2). Class/habitat 2 was our target for all downstream analysis as it represents a preferred foraging area for the simulated species.

While the categorisation of landscapes is a major decision during the analysis of habitat selection, we elected to keep the landscape classification constant to maintain a feasible multiverse analysis. However, to provide a point of comparison and to begin to explore the impact

of selection strength on the analysis choices and estimates, we re-ran the entire multiverse with a scrambled habitat raster. The scrambled habitat raster contained the same three classifications, but their positions were randomised, resulting in a habitat raster detached from the underlying simulated selection thereby representing low or no selection (i.e., random independent cell assignment, Fig. 2.2).

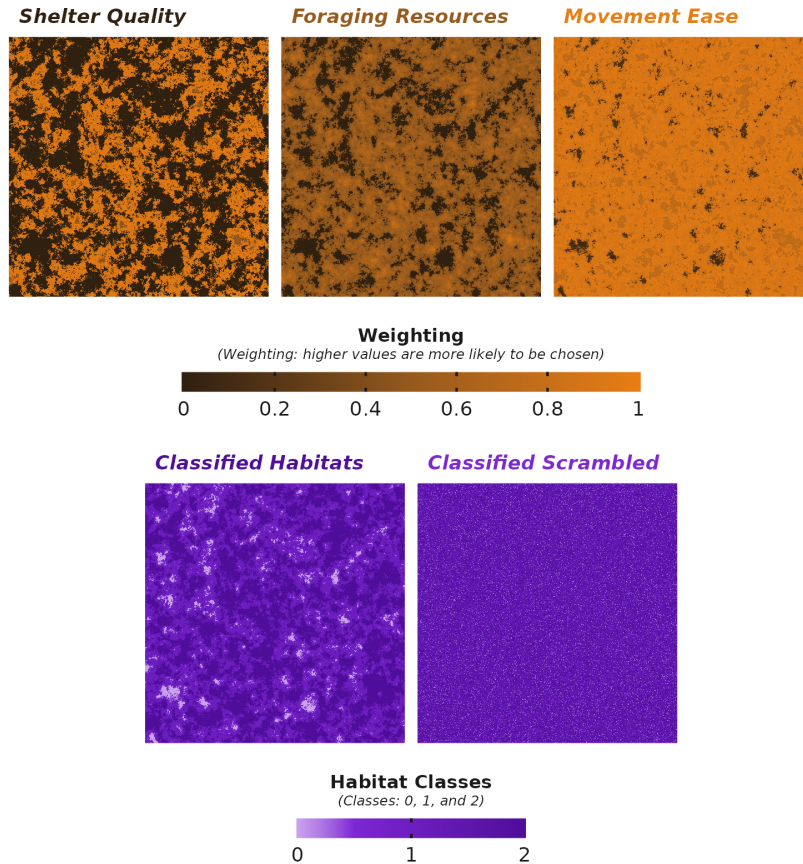


Figure 2.2: An example of the underlying environmental matrices used during the Badger species simulation, and the subsequent habitat classes to be used in the habitat selection analysis.

2.3.2 Sampling and Analysis Options

The number of variations and choices we were able to explore was ultimately dictated by computational costs and time. No multiverse can ever be entirely comprehensive, but we endeavoured to include the major decisions that are required during analysis and may not have immediately apparent single answers.

Sampling

The first decisions concern data quantity. We varied tracking duration and tracking frequency, while keeping tracking consistency the same (i.e., the time lag between data points was consistent for each combination of duration and frequency). While tracking consistency, or data loss during tracking, is an element affecting data quantity, the numerous ways of defining consistency and splitting tracking data into bursts led us to avoid exploring tracking consistency. Tracking timing in relation to time of day is also an important consideration; for example, recording the locations of a diurnal animal only at night is highly unlikely to be a viable way of determining foraging choices. For this exploration, we assumed that the researcher has sufficient knowledge about the animals' ecology to prioritise tracking during active hours. Timing only becomes a consideration when tracking frequency lowers to the point where the tracks performed during a day could occur entirely outside the animal's active period. Therefore, for less frequent tracking we ensured that daily tracking was centred on the animals' active period (as all our simulated species had daily activity patterns). We varied tracking duration from 7 to 240 days, and tracking frequency from 0.5 to 168 hours between subsequent data points (i.e., 2 to 0.01 points per hour). Where the combination of sampling frequency and duration resulted in fewer than 30 data points, we did not conduct analysis as this could prohibit accurate available area creation.

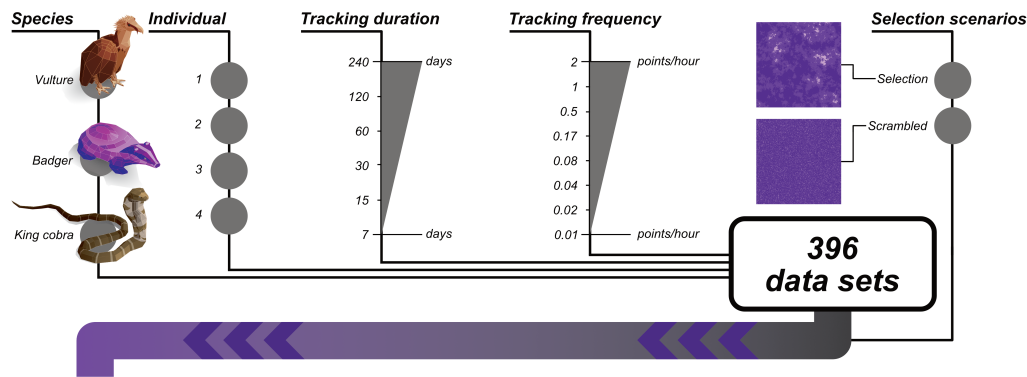
Data or location accuracy is another aspect of sampling variation. Different bio-logging equipment, terrain, animal behaviour, and weather can all impact the location error when tracking an animal. The causes of (Silva et al., 2017; Hofman et al., 2019; Vales, Nielson & Middleton, 2022), and potential solutions to (Frair et al., 2010; e.g., Hofmann, Cozzi & Fieberg, 2024), location error are numerous ; therefore, we did not explore the impact of error, instead prioritising more variation in other decision nodes.

In summary, the sampling decisions covered were: variation in three species represented by 4 independent individuals, tracking frequency, and tracking duration (Fig. 2.3).

Analysis

For the analysis, we focused entirely on decisions that result from an R analysis workflow (Fig. 2.3). R is the most common tool for analysing animal movement data (Joo et al., 2022). Joo et al. (2020) provide a review of the packages available to analyse movement data in

Sampling



Analysis

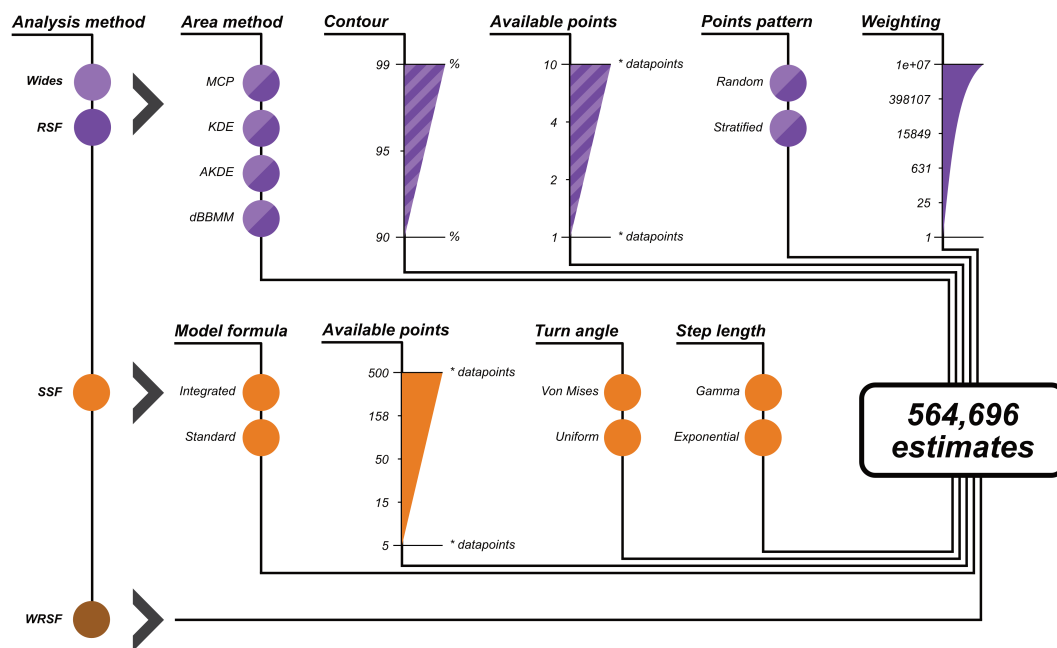


Figure 2.3: A diagram of the decisions creating a multiverse examining animal movement. Grey choices pertain to data simulation and sampling. Species: the three different movement archetypes the simulations were based on. Individual: the number of individuals simulated per each species. Tracking duration: the length in days of the tracking data sampled from the overall simulated year. Tracking frequency: the frequency of the sampling of data points from the 1-per-minute simulated data. Selection scenarios: the two classified habitat landscapes used during analysis. Purple choices concerned with analysis pathways using area-based estimations of availability. Area method: the method used to generate the available area polygon. Contour: the threshold used to convert the utilisation/occurrence distribution into a polygon. Available points: the number of generated points per data point to extract available habitat within the polygon. Points pattern: whether the random points were generated in a stratified or random pattern. Weighting: the weighting of the available points in relation to the used points in the RSF formula. Light orange choices pertain to the SSF analysis pathways. Model formula: whether the step and turn angles were added as interactions to the habitat variable. Available points: the number of randomly generated available points per used point. Turn angle: the distribution used to generate the turn angles of the available points. Step distribution: the distribution used to generate the step distances of the available points. Dark orange is the WRSF analysis. Decisions that required discrete choices are illustrated with circles, whereas continuous are depicted as triangles.

R, and we used this review as a resource for determining the options for habitat analysis. Specifically, we explored a subset of decisions made during the workflows using *adehabitatHS* v.0.3.18 (Calenge & Mathieu Basille, 2023), and *amt* v.0.2.2.0 (Signer, Fieberg & Avgar, 2019) R packages. During the writing of this paper, a new method was described as part of the *ctmm* package v.1.2.1 (Fleming & Calabrese, 2023), which we also preliminarily investigated. Combined, the packages offer many options for habitat analysis. We focus on four that tackle individual habitat selection:

- From the *adehabitatHS* package: resource selection ratios (Wides)
- From the *amt* package: Resource Selection Function (RSF)
- From the *amt* package: Step Selection Function (SSF)
- From the *ctmm* package: the newly described Weighted Resource Selection Function (wRSF Alston et al., 2023).

Each of the methods require downstream decisions resulting in a “garden of forking paths” to a final estimate of habitat selection (Fig. 2.3).

Area based approaches The Wides and RSF methods share many analysis decisions. The Wides approach aims to calculate selection ratios of used habitat compared to those available (Manly et al., 2007); whereas the RSF approach uses a logistic regression to compare the used and available habitats. Therefore, a principal concern in both approaches is how available habitat is defined.

The first decision is whether to approach the habitat selection analysis as a type I (population habitat use compared to landscape-level habitat availability), II (individual habitat use compared to landscape-level habitat availability) or III design (individual habitat use compared to individual-level habitat availability). We are concerned with an individual’s habitat selection, so we ignored type I. While type II habitat selection can identify individual selection, it requires habitat availability to be defined on a population or landscape scale (i.e., a universal availability). As movement data presents an opportunity to estimate individual selection, and we are not simulating a true interacting population, we focus on type III. Type III and the need to individually define available habitat opens up many different analytical decisions concerning how that availability is defined.

We used a range of methods to build areas that could be considered available to the animal

from the recorded locations of the animal: Minimum Convex Polygons (MCPs), Kernel Density Estimations (KDEs), Autocorrelated Kernel Density Estimations (AKDEs), and dynamic Brownian Bridge Movement Models (dBBMMs).

Minimum convex polygons (MCPs) simply create a polygon based on the outermost locations, where the interior angles sum to less than 180 degrees.

Kernel Density Estimations (KDEs) use kernel smoothing to build a heat map of use or utilisation. Critical to KDEs' operation is a bandwidth or smoothing factor (h) that alters the resulting area treated as used by the animal (Silva et al., 2020). For this analysis, we use the reference bandwidth (h_{ref}) as it is commonly used (Crane et al., 2021), while also quickly and consistently calculable on an individual-by-individual basis. An alternative common method to determine smoothing factor is Least Squares Cross Validation (LSCV). However, we avoided the LSCV method for two reasons: 1) the LSCV method can fail to converge on a smoothing factor, thereby introducing an unpredictable fail state in any potential multiverse, and 2) the more limited area estimate it produces makes it less suitable for defining habitat availability.

The limitations of MCPs and KDEs have prompted the development of newer methods of area use estimation (Silva et al., 2020) that better account for the non-independence and autocorrelative structures within animal movement data (Noonan et al., 2019). We used two of these methods to define availability.

First is the Autocorrelated Kernel Density Estimations (AKDEs) (Fleming et al., 2015; Fleming & Calabrese, 2017) from the *ctmm* package (Calabrese, Fleming & Gurarie, 2016; Fleming & Calabrese, 2023). The *ctmm* package provides a workflow for creating an area estimate based on a number of movement models. The movement models include different levels of autocorrelative structure: Ornstein-Uhlenbeck (OU) models account for central tendency; the Ornstein-Uhlenbeck Foraging (OUF) models build on this by also accounting for autocorrelation in movement speed; while the Independent Identically Distributed (IID) models assume location independence, meaning they are similar to traditional kernel density approaches. Each model type is also fitted in isotropic (circular) or anisotropic (elliptical) forms. The movement model used would constitute a decision during analysis; however, the *ctmm* package allows for model comparisons (using AICc) to choose a single best model. Therefore, we use the guidance from (Silva et al., 2022) to generate weighted AKDEs using the perturbative hybrid residual maximum likelihood method (pHREML), and select the best performing

by AICc for inclusion in further analysis. Using weighted AKDEs means the area estimates better address gaps between locations (Silva et al., 2022), with no apparent drawbacks nor prohibitively large computational cost.

The second movement-specific method we used was dynamic Brownian Bridge Movement Models (dBBMMs) (Kranstauber et al., 2012) from the move package (Kranstauber, Smolla & Scharf, 2023), which estimates movement capacity of the animal to calibrate repeated random walks between known locations. DBBMMs require a window and margin size that defines the number of data points over which movement capacity (motion variance) is calculated (Kranstauber et al., 2012). Window and margin are defined by a number of data points; therefore, to keep the time they represent the same between different tracking frequencies, we changed their values for each tracking frequency. As our most infrequent tracking is 168 hours (7 weeks), we set the window to the number of data points collected over 168 hours, and a margin of 48 hours. The broader window and margin sizes helped reduce computational costs, a consideration when many dBBMMs require calculation.

Unlike the AKDEs, dBBMMs do not produce utilisation distributions. As dBBMMs are estimating the uncertainty surrounding potential movement pathways, the resulting distribution is better described as an occurrence distribution. This occurrence distribution describes the within-sample uncertainty rather than the potential areas available to the animal beyond the sampling period (i.e., it possess little to no predictive capabilities). Therefore, using dBBMMs to define habitat availability means rather than comparing the habitat use (derived from the movement data) to availability (broader area predicted by the movement data), we are more comparing habitat use (derived from the movement data) to an alternative measure of habitat use (an area estimating uncertainty surrounding the movement pathways). We included dBBMMs to examine whether this arguably easy-to-make conceptual mistake impacts the final results (Alston et al., 2022).

What is considered available to an animal is not cleanly translated from a biological concept to a statistical one. While the area estimation methods (MCPs, KDEs, AKDEs, dBBMMs) described above make –to varying degrees– use of the movement information contained within the dataset, they remain abstractions of the underlying drivers of animal movement (i.e., selection and perception). Therefore, the researcher must make a somewhat arbitrary choice regarding what is included/excluded and counted as available; each of these area estimation methods requires a choice to be made regarding the outermost boundary. We explored the

impact of using a 90, 95, and 99% contour for all the methods. MCPs areas are generated based on a percentage of the location points, whereas the KDE, AKDE, and dBBMM methods require a contour to be extracted from the utilisation distribution (occurrence distribution for dBBMMs). AKDEs also provide a 95% confidence interval surrounding any chosen contour; for the purposes of simplicity, we only use the point estimate for each % contour.

Once we had created a defined availability area, we generated points at which the habitat type is recorded to estimate the relative availability. How many, and in what alignment, these availability points are generated poses two additional points of analytical choice. For this study, we varied the number of availability points as a multiple of the number of animal locations in the dataset (from 1 to 10, Fig. 2.3). For the alignment of the points, points were either random or stratified within the available area previously defined.

As mentioned, resource selection functions (RSF) share the above decisions on availability with the Wides methods. In addition, for the RSF we also explored the impact of varying the weighting of the available points when the RSF (i.e., generalized linear model / logistic regression) model is run, which impacts the fitting process. There are suggestions that altering the weighing can improve model convergence and decrease the uncertainty surrounding habitat selection estimates (Fieberg et al., 2021); we included this decision to examine whether the pursuit of a more confident answer is biasing the point estimate (from 1 to $1e+07$).

Step selection functions In addition to the area based methods of Wides and RSF, we explored of Step Selection Functions (SSF) and Weighted Resource Selection Functions (wRSF).

Instead of estimating habitat availability using area, SSFs use observed step lengths and turn angles to generate available locations at each time step (in our case for each data point in the dataset; strata). Consequently, SSFs have a number of decisions not shared with the area methods.

We explored three decisions associated with the generation of random (available but unused) locations for each step. The first is similar to the area methods; we varied the number of random locations generated per step from 5 to 500. The second and third decisions concern the distributions used to generate the random step and turn angles. Step lengths were tested with a Gamma and Exponential distributions, while turn angle was tested with Von-Mises and Uniform distributions.

The other decisions we explored in SSFs were whether to run the model as a standard step

selection or an integrated step selection function. The standard step selection model is a conditional logistic regression, where we aim to identify the association between whether a location is used/unused and the habitat values at that location ($\text{case_} \sim \text{values} + \text{strata}(\text{step_id_})$). The integrated step selection model is very similar, but the predictors also included step length and turn angle as interactive terms with the habitat values. The addition of these two components is meant to reduce bias by better accounting for the selection that may simply be an artefact of the movements of the animal rather than active habitat selection (Avgar et al., 2016).

We have neglected to explore an expansive aspect of SSFs regarding high resolution data –the use of “bursts”. Bursts are used when the researcher wants to look at selection over a given timeframe or tackle issues stemming from inconsistent tracking frequencies (Thurfjell, Ciuti & Boyce, 2014; Hofmann, Cozzi & Fieberg, 2024); therefore, locations are grouped into bursts and the bursts become the strata in the model rather than timestep (or timestep is nested within burst). Such grouping of data deserves investigation, but there are myriad of ways bursts can be defined, and their interaction with tracking duration, frequency, and consistency would warrant a separate multiverse analysis. As such, we have restricted this exploration to step selection functions where the datapoints are equal to the strata in the model.

Weighted Resource Selection Functions Compared to the other methods, Weighted Resource Selection Functions do not have many associated choices. Part of the motivation behind their development was to integrate autocorrelation and other critical structures of movement data into the modelling of habitat selection (Alston et al., 2023). This contrasts to the other methods that either ignore the confounding effect of these structures, or rely on potentially arbitrary decisions for how to mitigate any biasing effects the structures produce. As a result, we have no exploration of analysis decisions for this method because aspects such as an available area or points no longer apply, and availability is estimated as part of the model (Alston et al., 2023). Instead we explore only variations in data quantity (i.e., tracking duration and frequency).

2.3.3 Assessing the multiverse

To generate the multiverse, we used a meta-programming approach via the targets v.1.9.0 and tarchetypes v.0.11.0 R packages (Landau, 2021a,b). The targets package was created

to facilitate reproducible workflows, but is also paired with functionality enabling multiple analysis routes and parallel processing. Targets’ ability to track analysis task progress and intermediate R objects reduces repeat computation in a complex workflow that may not be able to be run in a single sitting. We created a branching analysis workflow based on the decisions we wished to explore, then used targets to iterate over all combinations of all decisions and compile all analysis end points ready for further examination. We set a seed during the running of the multiverse to further aid reproduction of these results. The single seed means that all simulated individuals, and subsequent analysis using randomness are unique within the multiverse.

Further examination was complicated by the four methods described above producing habitat selection values on different scales with different decisions associated with them. We elected to analyse the impact of the decisions separately. For each method, we ran two Bayesian regression models, using the *brms* v.2.22.0 (Bürkner, 2021), to explore: 1) which decisions best explain the absolute deviation from the median estimated selection, 2) which decisions best explain the raw deviation from the median estimated selection. The first model provides us with an idea of the decisions that can lead to random, but potentially unbiased variation surrounding the median estimate. Positive beta coefficients associated with the decisions would indicate that the decision leads to estimates further from the median estimate; whereas negative coefficients would suggest that decision is associated with estimates close to the overall median estimate. The second model highlights decisions that lead to systematic over or under estimation of habitat selection. Positive beta coefficients associated with the decisions would indicate that the decision leads to higher/stronger estimates of habitat selection; whereas negative coefficients would suggest more conservative of selection and a greater chance the selection is either not detected or negative (i.e., avoidance). All models included individual and species as nested group effects, as well as a group effect accounting for whether the classified landscape was scrambled or not. For example, the formula for the SSF model looking at absolute deviation: “absolute delta from median estimate” $\sim 1 +$ “tracking duration scaled” + “tracking frequency scaled” + “integrated formula” + “step distribution” + “turn distribution” + “available points per step scaled” + “(1|landscape scrambled)” + “(1|species/individual)”. We normalised all predictor variables because they all existed on vastly different scales, as we wanted to assess relative impact. We assessed Bayesian model convergence using R-hat values, acf, and trace plots. Based on these assessments we modified the running parameters (RSF 500 iterations, 150 warmup; Wides 1500 iterations, 500

warmup; SSF 1500 iterations, 250 warmup; wRSF 2000 iterations, 500 warmup).

We elected to have the model try to predict the deviation from the median estimate of habitat selection, because the simulation cannot provide a direct analogous value to the outputs of all four methods. The analysis closest to mirroring the agent-based model of the simulated data is the SSF. The simulation is a discrete time model, where the animal makes decisions based upon a number of options making it similar to a SSF step-wise model structure. Despite this similarity, the decision-making of the animal is made on two time frames; thereby, undermining our ability to directly compare simulation values with step selection outputs. Additionally, the simulated selection of the animal is balanced against other demands such as site fidelity, movement resistance, and avoidance of certain locations; therefore, any input habitat selection into the simulation becomes directly incomparable with a derived answer (i.e., analysed) from the simulated movement. The closest analogous values we could generate came from running RSF and iSSF models directly on the simulated decisions made by the simulated animal. We ran these models for both time scales: movement decisions every time step, and destination decisions every behavioural state switch. These model results, while not strictly comparable to the other outputs, help confirm the animal was correctly preferring the chosen habitat matching the parametrisation provided to the simulation. The estimates calculated this way likely underestimate selection further as the rest sites were only generated within the preferred habitat (a simulated strict selection). The lack of direct comparability between simulated parameters and analysed estimates of selection supports the use of the median estimate to explore the cause behind the most deviant estimates.

2.4 Results

The completion of the multiverse resulted in 396 datasets representing an individual animal's movements sampled via a distinct tracking regime (i.e., a combination of tracking frequency and duration; Fig. 2.3). The datasets covered 3 virtual species, 4 individuals per species, frequencies from 0.006 to 2 locations per hour, and durations from 7 to 240 days. Each of these datasets underwent various combinations of analysis choices resulting in 564,696 estimates of habitat selection (Wides: 76,032; RSF: 456,192; SSF: 31,680; wRSF: 792).

2.4.1 General patterns

We can get an initial impression of the pattern of the final selection estimates using a specification curve (Simonsohn, Simmons & Nelson, 2020; Hall et al., 2022). In a scenario where the decisions make little difference, all estimates should group together around a common answer, with the majority of deviation from that common answer being the result of sampling or individual variation.

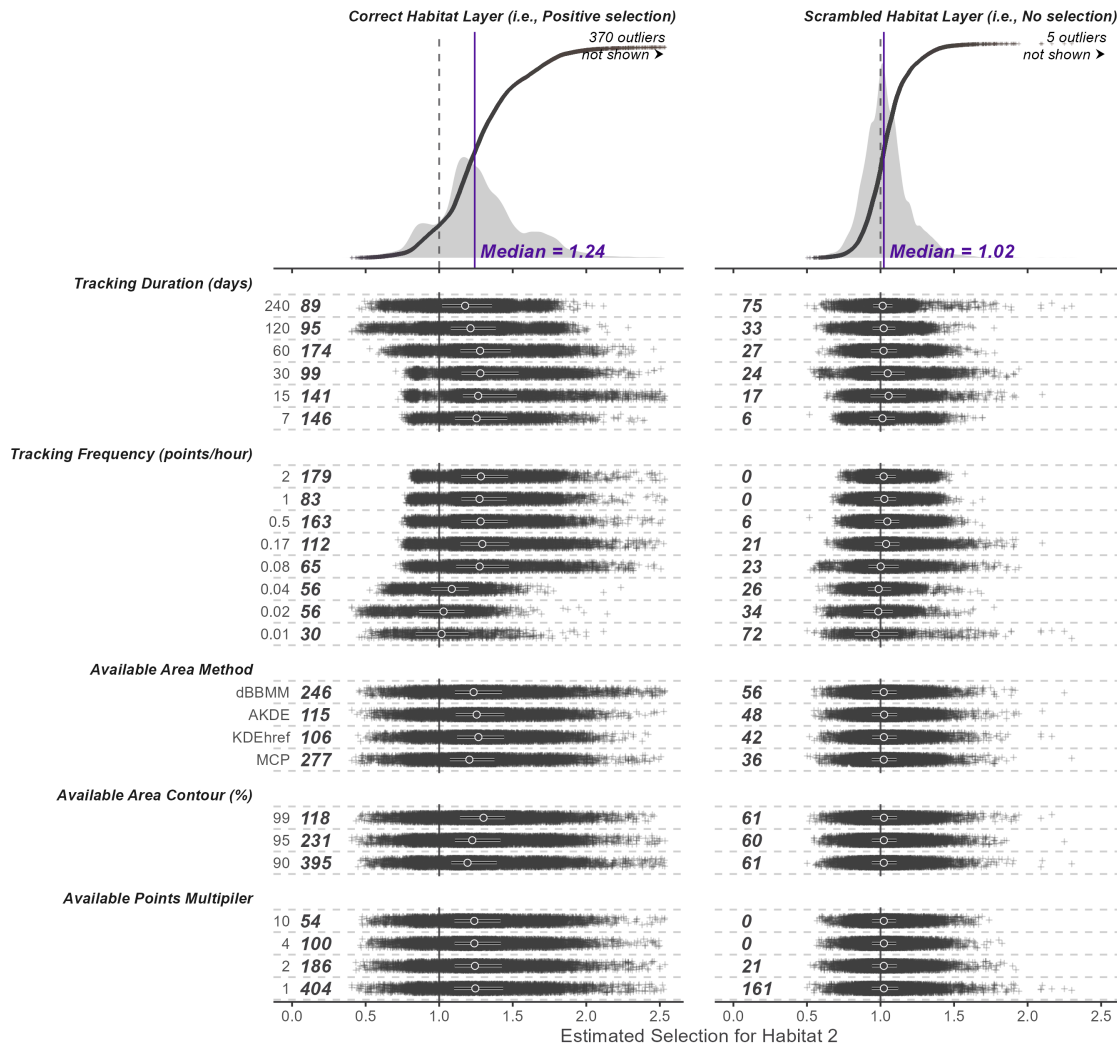


Figure 2.4: A specification curve showing all habitat selection estimates from the Wides analysis pathways applied to the simulated data. Each point is an estimate of habitat selection for habitat 2. Top plot shows all estimates arranged by estimated selection strength, with an arbitrary sorted y-position. The distribution of the selection estimates is also shown as a grey density plot. Lower plot shows the same estimates in relation to the decisions that lead to them. Left-hand numbers describe the number of NA results connected to each decision, indicating a failure somewhere in the calculations. Solid circles are the medians for each choice.

Starting with the Wides results, we see a quite gentle *S* shape (Fig. 2.4), with most estimates falling in the range of 0.5 to 2. There is an agreement of positive selection for habitat 2 as expected (the simulated habitat type that is preferred for foraging), and a stronger selection identified in the selection scenario compared to the scrambled raster (no selection) scenario. While the positive selection appears in reasonable agreement that selection is occurring (more estimates toward the higher values than lower), the exact estimate appears relatively sensitive to the choices made. Looking at the specific choices, we can see several weak patterns. The first is in tracking duration where we see a slight reduction in the variation of estimates with longer tracking durations, when selection is present. The same is not clear in tracking frequency because the lower tracking frequency could not be combined with lower tracking durations resulting in fewer estimates for low frequency scenarios. Despite this, we do see a slight reduction in estimate variation from frequencies of 0.17 locations per hour and higher, as well as more consistent positive selection estimates. Similar to the tracking intensity choices, the contour used during the area generation appears to affect the spread of estimates, with a larger contour leading to a slightly greater agreement between estimates. So too does the multiplier of available points, with more points slightly reducing the spread of estimates, although we see no movement in the median estimate. Higher numbers of available points, and a higher tracking intensity, did appear to lead to a greater chance of a successful estimate being produced (see numbers to the right of the 0 line in Fig. 2.4). The choice of area method (MCP = Minimum Convex Polygon, AKDE = Autocorrelated Kernel Density Estimate based on the best performing movement model, KDEhref = Kernel Density Estimate using reference bandwidth selection, dBBMM = dynamic Brownian Bridge Movement Model) appears to suggest more consistently similar estimates with AKDE and KDEhref methods compared to MCPs and dBBMMs, with MCPs leading to slightly lower estimates in the selection scenario (Fig. 2.4). This difference disappears under the scrambled habitat classification, and in that scenario all analytical choices appear to make little impact on the median estimate.

The Resource Selection Function (RSF) specification curves show a more ideal *S* shape, with the majority of estimates clearly converging on a common answer (Fig. 2.5). Similar to the Wides analysis, the majority of the answers and medians show selection for habitat 2 in the selection scenario. The no selection scenario still overall shows selection, but it is considerably lower. While there are answers incorrectly suggesting avoidance of habitat 2, all but two medians connected to a specific choice (the lowest two tracking frequencies) suggest selection. To best see the changes in variation connected to a given choice, we can look at

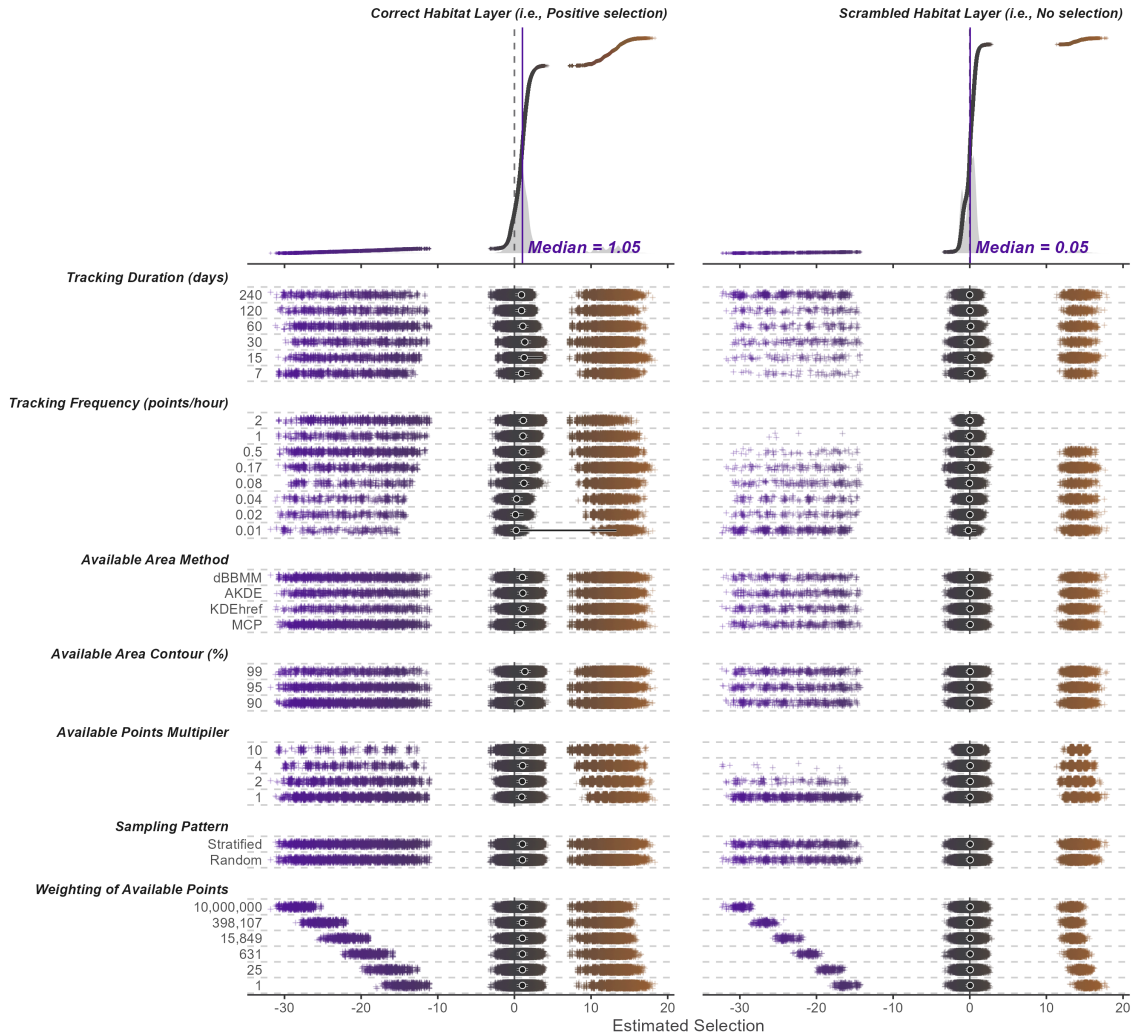


Figure 2.5: A specification curve showing all habitat selection estimates from the RSF analysis pathways applied to the simulated data. Each point is a estimate of habitat selection for habitat 2. Top plot shows all estimates arranged by estimated selection strength, with an arbitrary sorted y-position. The distribution of the selection estimates is also shown as a grey density plot. Lower plot show the same estimates in relation to the decisions that lead to them. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. Colours illustrate deviation from the median, with purple indicating lower levels of selection and orange showing higher selection estimates.

the rates and position of the outliers (highlighted in purple and orange). Although weak patterns, tracking frequency, and multiplier of available points see slight reductions in the number of estimates that are very high or very low. This is particularly apparent in the no selection scenario, where higher tracking frequencies led to very similar answers surrounding zero. Other patterns are difficult to discern, except for the weighting of used points that has a clear impact on the outlying low estimates. The distinct step-like pattern is a result of the logarithmic scale the weighting varies over (required to capture such a wide range of possible values), which can be seen reflected in other choices' groupings of estimates. The weighting does not appear to impact the overall median, with weighting only appearing to dramatically impact the outlying results. Comparing the selection scenario to the no selection scenario, we can see that the selection scenario is more likely to produce larger outlying results, both positive and negative.

For the RSF analyses, we also recovered the range of the 95% confidence intervals surrounding each estimate. When plotted against the estimate, we see clear groupings where those most extreme estimates are also, by a significant margin, the most uncertain (Fig. 2.6).

The Step Selection Function (SSF) specification curves appear even more successful than the RSF. There is a reasonably clear *S* shape indicating many analysis end points resulting in similar estimates of selection; with a distribution that reveals less variation in final estimates of selection (Fig. 2.7). There are also considerably fewer estimates that indicate avoidance of habitat 2. The stand out patterns are more easily seen in the outlying high estimates. Once again, there are indications of fewer outlying high estimates when data quantity increases, both with duration and frequency. The number of points per step (or strata) appears to intuitively follow that more available steps leads to more consistent estimations, and also the outliers that exist appear to become less extreme. The decisions concerning the distributions used to model step lengths and turn angles, as well as whether the model is integrated or not, appear to make little difference to the spread of final estimates. The impact of analysis choice on the SSF method appears to be more consistent between the two selection scenarios.

Very much like the RSF explorations of the raw data, the 95% confidence intervals are dramatically larger (between 100 and 1000 fold) for the outlying estimates than those nearer the median estimate (Fig. 2.6).

The Weighted Resource Selection Function approach has very few decisions associated with it currently; however, we still explored the impact of tracking duration and frequency on

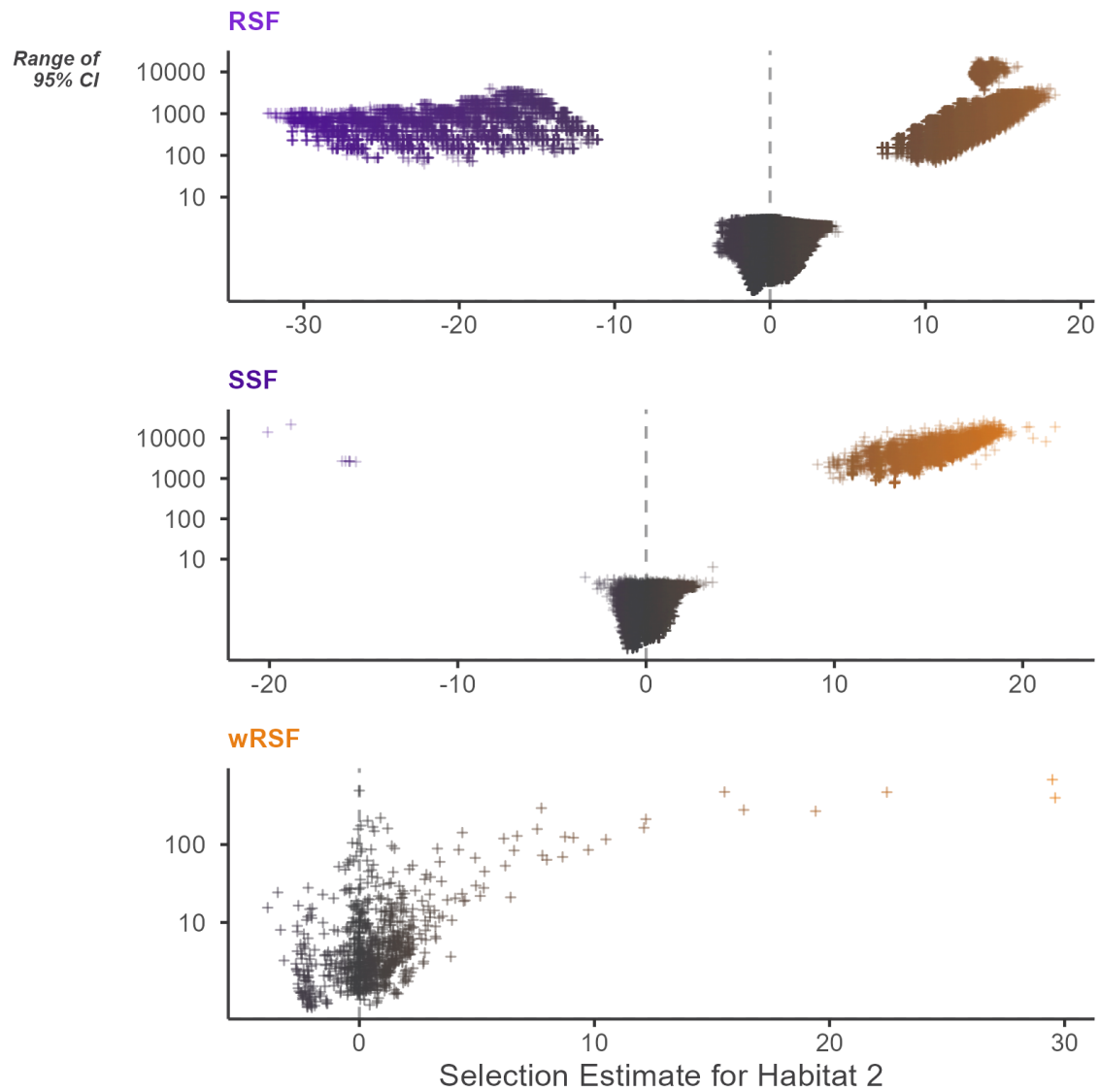


Figure 2.6: All standard errors associated with estimates of habitat selection from RSF, SSF, and wRSF analysis pathways for habitat 2. More extreme estimates of selection are associated with dramatically larger confidence intervals, N.b., y axis is log scaled.

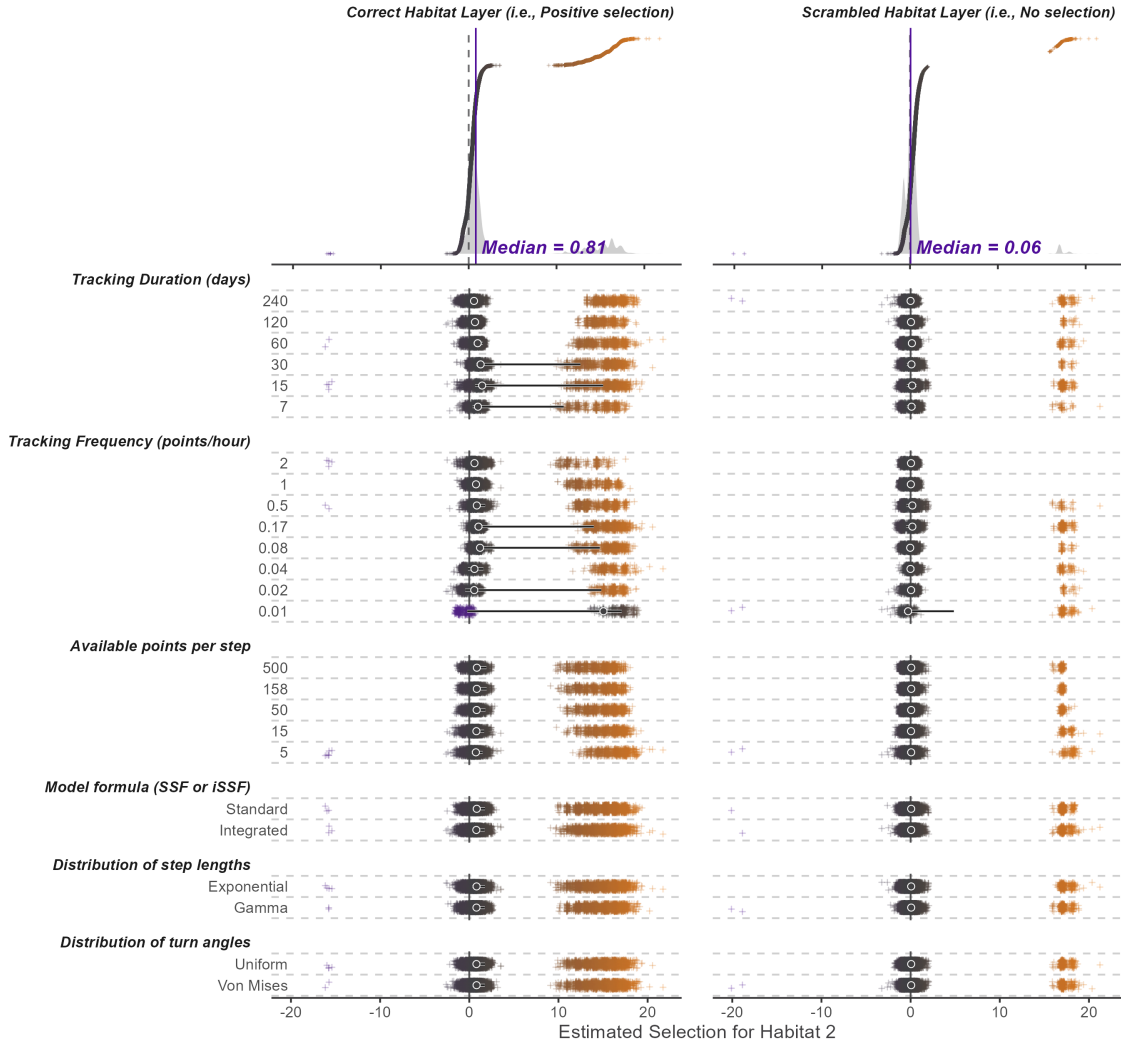


Figure 2.7: A specification curve showing all habitat selection estimates from the SSF analysis pathways applied to the simulated data. Each point is a estimate of habitat selection for habitat 2. Top plot shows all estimates arranged by estimated selection strength, with an arbitrary sorted y-position. The distribution of the selection estimates is also shown as a grey density plot. Lower plot show the same estimates in relation to the decisions that lead to them. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. Colours illustrate deviation from the median, with purple indicating lower levels of selection and orange showing higher selection estimates.

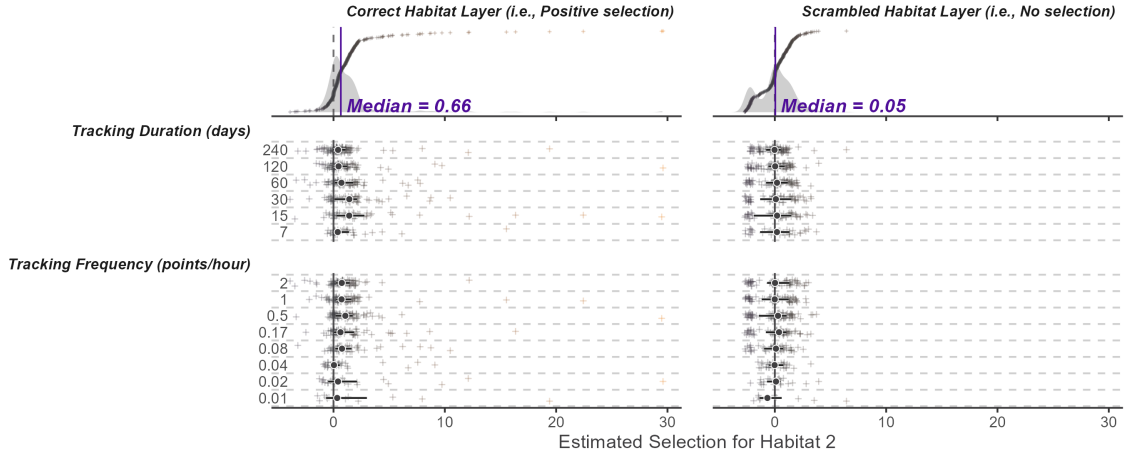


Figure 2.8: A specification curve showing all habitat selection estimates from the wRSF analysis pathways applied to the simulated data. Each point is a estimate of habitat selection for habitat 2. Top plot shows all estimates arranged by estimated selection strength, with an arbitrary sorted y-position. The distribution of the selection estimates is also shown as a grey density plot. Lower plot show the same estimates in relation to the decisions that lead to them. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range. Colours illustrate deviation from the median, with purple indicating lower levels of selection and orange showing higher selection estimates.

the estimates (Fig. 2.8). The highest estimates appear connected to the lower tracking frequencies, but like the SSF results, those high estimates also tend to have large confidence intervals (Fig. 2.6). The bulk of the estimates agree and are tightly concentrated around the median with very few underestimating. As with RSF and SSF methods, the wRSF is very likely to correctly identify the positive selection. Also the median selection under the no selection scenario is correctly estimated as lower, particularly when more data are available.

2.4.2 Model exploration

While the specification curves are a good way of getting an intuitive sense of variation, the scale and range of estimates can skew the interpretation of which choices have the largest impact. Our Bayesian models cut through the noise displayed in specification curves to reveal more generalised patterns and absorb variation stemming from the stochastic nature of the animal simulations and the variation between species.

All models showed that the variation stemming from the analysis choices was less than that explained by the group effects of individual, species, and whether the landscape classification was scrambled (Table. 2.1). The conditional R^2 values ranged from 0.08 to 0.3, with the RSF

models performing the worst. Largely the marginal R^2 values were very low (often below 0.1), meaning that while the models reveal clear effects, the overall variation they ascribe to given analysis choices is reasonably limited compared to the overall variation stemming from the different species, individuals, landscape classification, and stochastic nature of the movement data simulation.

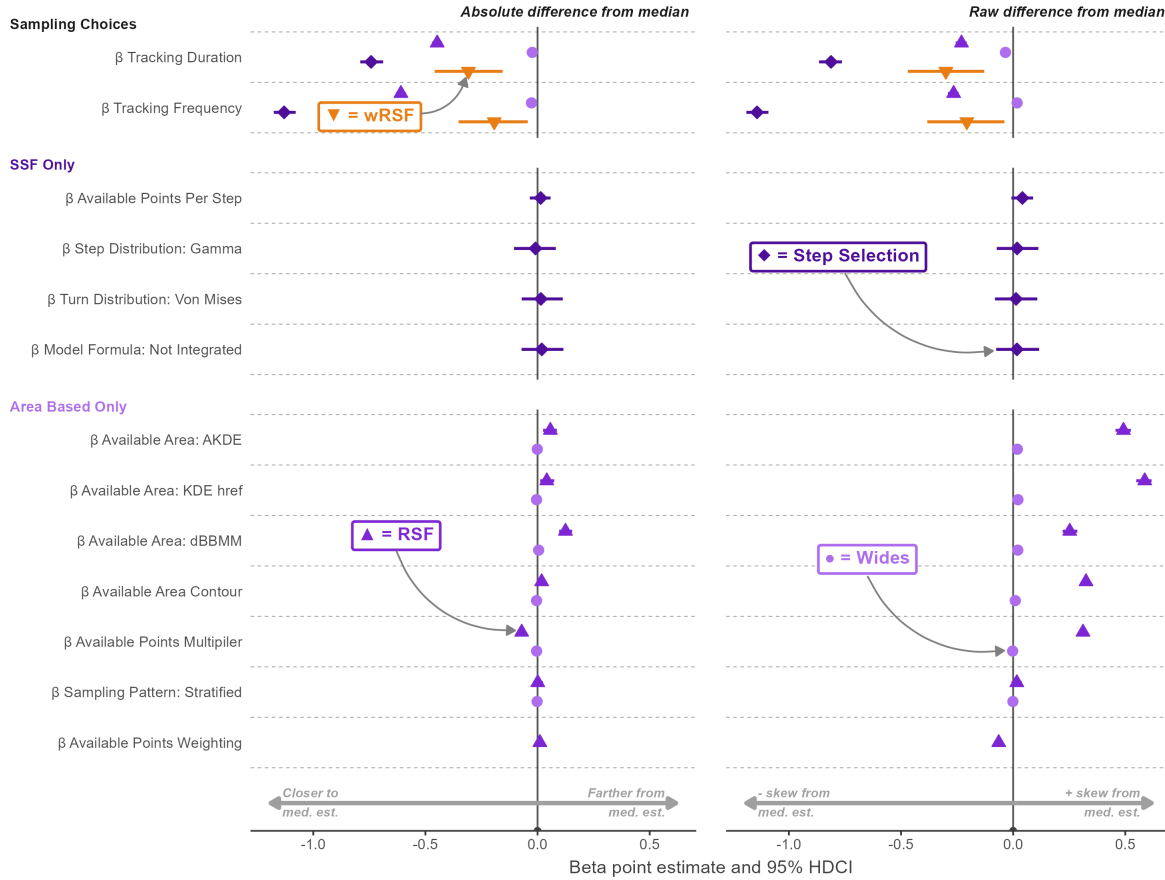


Figure 2.9: Estimated betas for each choice’s impact on the deviation from the median habitat selection estimate. Deviation from the median estimate was examined as the absolute deviation (left) and the raw deviation (right). All eight independent models are displayed together, split between analysis method and deviation type. Each point represents the point estimate derived from the 95% median highest density continuous interval, and is plotted alongside a line showing the extent of the 95% credible interval. The y-axis shows all the choices explored in the different analysis methods, grouped by the methods they apply to.

For the decisions connected to Wides results, the absolute deviation was curtailed most effectively by increases in data quantity (Fig. 2.9; Fig. 2.11), with increases in tracking duration (β : -0.024, 95% CrI -0.026 – -0.021) being marginally more effective than tracking frequency (β : -0.027, 95% CrI -0.029 – -0.024). Of the other decisions, only the multiplier of available points appeared to decrease absolute deviation unambiguously (β : -0.004, 95% CrI -0.005 – -0.003), and its effect was much smaller than those concerning data quantity. Choice of area

method as well as the pattern used to generate random points seem to make little clear difference, with all credible intervals connected to those choices overlapping zero. Contour choice increases the absolute deviation from the median estimate β : -0.004, 95% CrI -0.006 – -0.003. That appears counter to the initial specification curve plot and highlights the difficulty of identifying patterns from the raw data alone.

The model examining raw deviation from the median estimate reveals that the Wides approach is contrastingly impacted by tracking frequency β : 0.018, 95% CrI 0.015 – 0.021 and tracking duration β : -0.034, 95% CrI -0.037 – -0.031, with the former leading to greater estimates of selection (Fig. 2.9; Fig. 2.15). The choice of stratified versus random available point positioning β : -0.001, 95% CrI -0.004 – 0.002 and the available points multiplier β : -0.002, 95% CrI -0.004 – -0.001 have less of an impact, mirroring the smaller effects seen in the absolute deviation model. All area related choices do not move relative to each other, and reveal that MCPs as an area method lead to lower estimates of selection in this scenario.

Every decision examined by the RSF absolute effect model appears to make a difference when estimating selection (Fig. 2.9; Fig. 2.12). Like the Wides analysis, data quantity is most important. The two largest effects are from tracking duration (β : -0.447, 95% CrI -0.47 – -0.425) and frequency (β : -0.609, 95% CrI -0.631 – -0.588), showing that more data leads to more stable estimates of habitat selection (i.e., more likely to get an answer closer to the overall median estimate).

However, unlike with Wides, the other analysis decisions also matter, although to varying extents. It appears that smaller contours (β : 0.018, 95% CrI 0.007 – 0.028) and a greater multiplier of available points (β : -0.071, 95% CrI -0.08 – -0.06) decreases the deviation from the median estimate. Whereas, using a stratified available points pattern led to very similar deviation from the median than a purely random available points pattern (β : 0.001, 95% CrI -0.019 – 0.025). Choice of area method can alter results in differing ways. Compared to using MCPs, both AKDE (β : 0.057, 95% CrI 0.025 – 0.086) and KDEhref (β : 0.041, 95% CrI 0.012 – 0.075) produce estimates marginally less likely to be closer to the median estimate, and to similar degrees. By contrast, dynamic Brownian Bridge Movement Models are more likely to lead to greater variation in estimates than those analyses that use MCPs (β : 0.124, 95% CrI 0.096 – 0.154).

Examining the interactions between tracking duration and frequency with the area methods further reveals the connection between dBBMMs and more variable estimates. While the

other area methods interact with tracking duration in a similar fashion, dBBMMs are more sensitive, meaning that at shorter tracking durations the variation created by choosing dBBMMs over another method is greater (Fig. 2.10). When interacting with tracking frequency, the sensitivity is less pronounced; instead we find MCPs being the most sensitive to a lowering of tracking frequency (Fig. 2.10).

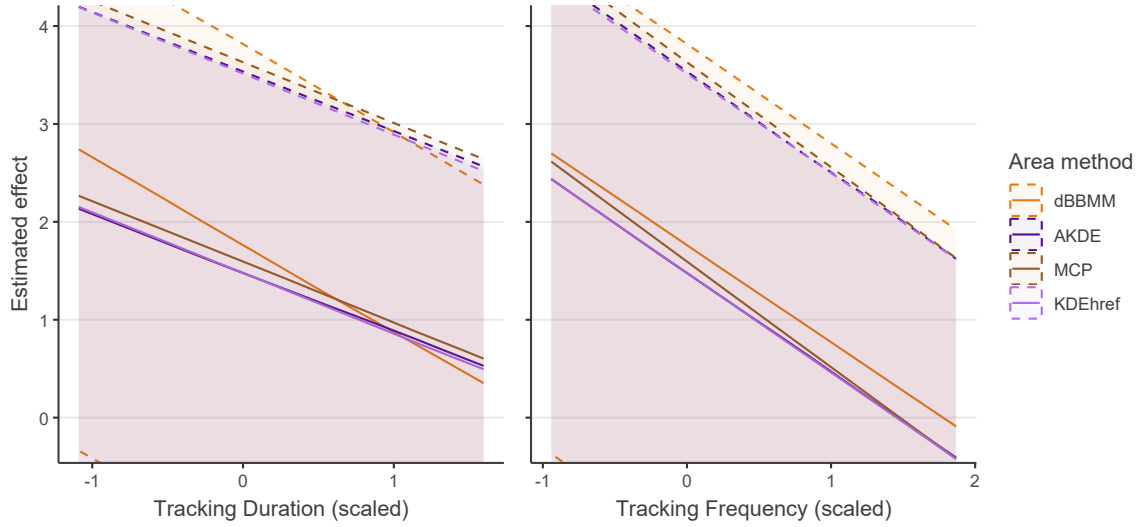


Figure 2.10: The resulting interactions from the Bayesian regression model looking to explain absolute deviation from the median in the RSF analysis pathway. Left-hand plot shows how scaled tracking duration interacted with area method chosen, while the right-hand shows the interaction between scaled tracking frequency and area method chosen.

The RSF model examining raw deviation from the median estimate reveals a general tendency for higher data quantity to lead to more conservative estimates of selection (tracking duration β : -0.231, 95% CrI -0.256 – -0.207 and tracking frequency β : -0.265, 95% CrI -0.291 – -0.243; Fig. 2.9; Fig. 2.16). While the way the movement data is simulated makes it impossible to definitively determine whether this conservatism is correct, the direct estimates from the simulations using an RSF approach suggest the low positive estimates (also near the median; see associated data for directly estimated selection values, Table. 2.2) are more likely to be the correct estimation of simulated selection via RSF. This matches the absolute deviation model where more data leads to estimations closer to the overall median estimate.

Stratified available point sampling was associated with more positive estimates of selection (β : 0.016, 95% CrI -0.007 – 0.038), but also those closer to the median (as indicated by the previous absolute deviation model). This suggests that random point sampling, by contrast, leads to marginally more variable estimates that tend to be lower. The same pattern is visible

in available points multiplier, where more points lead to a more consistent estimate but with a tendency for that estimate to be higher (β : 0.312, 95% CrI 0.301 – 0.324).

The area method choices reveal a similar pattern. Dynamic Brownian Bridge Movement models were connected to the more variable estimates, and that variability tended towards higher estimates (β : 0.252, 95% CrI 0.221 – 0.286). The other methods are connected to comparatively lower deviations from the median estimate, but higher raw estimates. The area contour selected also reflects this (β : 0.325, 95% CrI 0.312 – 0.336). Taken together, it appears that methods capturing a spatially broader picture of availability have a tendency to produce more stable (in relation to other analysis decisions) but higher estimates of selection. The dBBMMs, by contrast, are more vulnerable to variation in other decisions, a trait reflected in the interaction with data quantity measures.

Overall there does not appear to be a decision that leads to great deviation from the median estimate in one direction.

The SSF absolute deviation model very closely matches the previous two models (Fig. 2.9; Fig. 2.13), where tracking quantity both in terms of duration (β : -0.742, 95% CrI -0.791 – -0.688) and frequency (β : -1.13, 95% CrI -1.176 – -1.079) dominate in regards to explaining why estimates tend to be closer to the median estimate. Unlike the other analyses, all other decisions made a negligible impact on an estimate's absolute proximity to the median estimate.

Due to the one-sided variability of the SSF estimates, the raw deviation model mostly mirrors the absolute deviations (Fig. 2.9; Fig. 2.17). The only noticeable difference is in the model formula decision that changes the direction of the beta point estimate from negative to positive (β : 0.017, 95% CrI -0.076 – 0.116). However, the large overlap with zero means that in both cases the step distribution decision is not making a considerable impact on the estimates relative to the median estimate.

The results from the wRSF models are largely similar to the SSF models, although with fewer decisions (Fig. 2.9; Fig. 2.14). Both tracking duration and tracking frequency lead to lower deviation from the median estimate (tracking duration β : -0.308, 95% CrI -0.459 – -0.156 and tracking frequency β : -0.193, 95% CrI -0.352 – -0.044), and the near-exclusively positively skewed outliers mean that lower raw deviation is associated with lower estimates (tracking duration β : -0.3, 95% CrI -0.47 – -0.129 and tracking frequency β : -0.207, 95% CrI -0.383 – -0.039; Fig. 2.9; Fig. 2.14). The only difference (although both effects overlap considerably) is that increased tracking duration tended to be linked more strongly with lower selection

estimates.

2.5 Discussion

Our multiverse exploration of individual habitat selection analysis methods reveals the importance of data quantity in generating consistent results. Regardless of the approach used, data quantity increases resulted in estimates of habitat selection closer to the median estimate and more consistently in line with the underlying simulated habitat selection. There appeared to be a reduction in deviation from the median estimate when using newer methods (e.g., SSF and wRSF). This suggests that advancements in habitat selection analysis methods are not only worthwhile, but actively counter reproducibility issues that may stem from researcher choice.

The overall pattern in estimates mirrors those seen in other fields, where the bulk of answers converge but outliers remain (Silberzahn et al., 2018; Simonsohn, Simmons & Nelson, 2020; Gould et al., 2023). While the incentives for finding positive results remain in the publishing system (Fanelli, 2010b, 2012), in the case presented here, using analysis choices to game or mine for a given exciting result appears time-consuming because the proportion of extreme estimates is relatively low compared to the majority of estimates that agree near the median. For example, the 10 and 90% percentiles easily exclude outlying estimates, particularly those lower estimates, and remain reasonably close to the median estimates Wides selection: 10% 0.913, 90% 1.68; Wides scrambled: 10% 0.867, 90% 1.23; RSF selection: 10% -0.52, 90% 11.3; RSF scrambled: 10% -1.12, 90% 0.881; SSF selection: 10% -0.0184, 90% 15.6; SSF scrambled: 10% -0.811, 90% 0.763; wRSF selection: 10% -0.474, 90% 2.97; wRSF scrambled: 10% -2.21, 90% 1.75. If the past habitat selection studies report extreme findings, there is a general tendency for these to be paired with very wide confidence intervals; if the studies have reported the confidence intervals, this provides an additional safeguard against over interpretation of potentially incorrect extreme results.

While the goal of this study was not to review and provide direct guidance for the most effective ways of determining habitat selection, our findings show several key patterns that reflect previous guidance in the literature. In particular the benefits of larger datasets and the importance of defining availability (Northrup et al., 2013; Street et al., 2021); the latter is particularly key because of the potential for subjective definitions (Beyer et al., 2010). One

specific example would be the tendency for larger more generous area methods (e.g., KDEhref, AKDE, and wider contours) leading to more generous estimates of selection that are closer to the median compared methods that grow more tightly with data quantity (e.g., dBBMMs); this mirrors previously revealed patterns connecting broader spatial scales and transferability (Paton & Matthiopoulos, 2015).

The overall pattern shows data quality supremacy, but this cannot be used as an excuse to ignore the decisions during analysis. Two aspects of the results highlight that complacency is unwise. First RSF, SSF, and wRSF approaches appear to be more consistent than the more basic and older Wides ratio based approach. Both RSF, SSF, and wRSF have improved capacity to account for the peculiarities of movement data structure (e.g., via random effects) than the Wides approach. Second is the evidence for conceptual mistakes to have impacts on final results, implying that in some cases there are *more correct* answers to certain analysis decisions. For example, the use of dBBMMs to define availability is not strictly correct as the (confidence) area generated by dBBMMs reflects uncertainty surrounding use. This renders the down stream analysis of habitat selection a comparison of use (measured by points) to use (measured by area of uncertainty surrounding movement trajectories); as opposed to the desired use versus availability. The dBBMM results exemplify the need to maintain a connection between abstractions (e.g., an area) of the data and the intended use. A similar argument can be made for data quantity, where mismatches between frequency and desired behaviours or an inadequate tracking duration will obscure answers regardless of the analysis applied. Our simulations contained daily cycles and the tracking spacing was skewed towards those active hours in those cycles for lower tracking frequency regimes. Failure to do this, instead allowing locations to be recorded equally at night and day, would have skewed estimates towards the habitat used for shelter rather than foraging. Matching question to analysis approach remains important, but the cost of mistakes may be less dramatic than feared for general assessments of habitat selection.

Further good news is that some decisions connected to more stable estimates have very little reason not to be maximised. For example, the points multiplier or points per step decision used to generate the number of available points can be maximised to whatever extent allowed by the computational resources available to the researcher. Increases in the number of available points will improve accuracy when defining available habitats; Street et al. (2021) indicate that the systematic sampling with high numbers of points is preferable to capture a complete unbiased availability. Fieberg et al. (2021) states that available points should be high enough

to ensure stable coefficient estimates, that is also aided by increased weighting of available points in the logistic regression. While we did not see the improvements in points per step in the SSF analysis, the benefits of high available steps is more pronounced when working with rare habitat types (Thurfjell, Ciuti & Boyce, 2014). More available points reduces the chances of missing small areas of rare habitat; our analysis would not have revealed this because of the abundance of the preferred habitat type.

We did not see large improvements using different step length or turn angle distributions; despite that this has been shown to play an increasingly important role in mitigating biased estimates in scenarios with stronger selection strengths (Forester, Im & Rathouz, 2009). The limited cost of integrating step length and turn angle into SSF formulas would suggest iSSF formulation is likely always preferable for estimation of individual habitat selection.

2.5.1 Limitations of the multiverse approach and future directions

While this study helps demonstrate the impact of researcher choice in habitat selection studies, the true number of choices is impractical to explore and ever-growing. Time and computational limits compound with the infinite granularity of decisions connected to continuous variables (e.g., tracking frequency) to ensure that all choices cannot be completely explored.

The relationship between tracking regime, data quantity produced, and the resulting habitat selection results is a complex area to assess; this becomes increasingly complex as variation stemming from species and individual plasticity is integrated. We focused on systematic controlled changes to tracking frequency, but in field scenarios data quantity can be impacted by practical and logistical constraints that result in stochastic data loss. Newer methods help alleviate the problems with inconsistent tracking (Fleming et al., 2018; Silva et al., 2022; Alston et al., 2023; Hofmann, Cozzi & Fieberg, 2024), providing an additional reason why analysis decisions should not be ignored.

Ultimately the multiverse approach cannot reveal the true answer; the multiverse’s mean result describes the mean of the analysis pathways and does not necessarily converge towards a true effect. In much the same way as replication does not provide a direct proxy for truth, the chosen analysis pathways, and their relation to an underlying truth, is not necessarily improved by conducting more. Improvement will only come with the additions being fundamentally more accurate, and better reflective of the study system (Devezer et al., 2021).

This unclear relationship between consistency of results and the study system is why we tried to capture a range of variability in the simulations. We attempted to make our findings more generalisable by simulating three distinct animal species that move in different fashions and that face differing levels of habitat connectivity. We cannot discount scenarios with different movement characteristics that interact with particular analysis choices to bias habitat selection estimates. Scale remains a key consideration for any study (Thurfjell, Ciuti & Boyce, 2014; Northrup et al., 2022), and context-specific selection remains difficult to generalise (Avgar, Betini & Fryxell, 2020). The difficulty in generalising a “one-size fits all” approach highlights the benefits of methods that can be parametrised by the underlying characteristics of the data (e.g., wRSF Alston et al., 2023), and where the uncertainty stemming from sampling can be carried forward to the final estimates.

Despite the inherent limitations of multiverses, the exploration of analysis choice presents an important consideration for any researcher looking to know the reliability of their findings (Hoffmann et al., 2021). More formal and mathematically based estimates of reliability are available for specific approaches, for example RSFs (Street et al., 2021). These present an important component when reporting RSFs. However, they still require some level of *a priori* decision making [e.g., landscape and availability definition; Northrup et al. (2022)]; therefore, researcher choice remains.

Beyond the aforementioned limitations to the multiverse approach, there is scope to expand the exploration. Our next step is to reapply the same meta-programming pipeline approach to population targeting habitat analyses (Chapter 3). A focus on these approaches will allow us to explore the third key component of data quantity –sample size– as well as how well analyses tackle individual variability, either in the analysis itself or while summarising individual habitat selection estimates. There are already methods for assessing required sample size to detect a given selection for RSFs (Street et al., 2021).

2.6 Conclusions

We show, using a multiverse of analysis pathways, that data quantity is an important safeguard against extreme outlying results in individual habitat selection analyses. The indications that analysis choices are less important than data quantity improves our confidence in previously published habitat selection studies as there is no more scope to mine (via ques-

tionable research practices) for a desired result than in other disciplines. Newer methods of habitat selection analysis appear to reduce variation from researcher choice further, indicating that future studies may be more reproducible than those before.

2.6.1 Acknowledgements

This work was supported by the Natural Environment Research Council (NERC) via the IAPETUS2 Doctoral Training Partnership held by Benjamin Michael Marshall (grant reference NE/S007431/1). We would also like to thank Aubrey Alamshah for their support and manuscript review.

2.6.2 Software availability

In addition to packages already mentioned in the methods we also used the following.

We used *R* v.4.4.2 (R Core Team, 2022) via *RStudio* v.2023.12.0.369 (RStudio Team, 2022). We used *here* v.1.0.1 (Müller, 2020b), *readr* v.2.1.5 (Wickham, Hester & Bryan, 2024), and *qs* v.0.27.2 (Ching, 2023) to manage directory addresses and saved objects.

We used *raster* v.3.6.30 (Hijmans, 2023a) and *RandomFields* v.3.3.14 (Schlather et al., 2015) to aid landscape raster creation alongside *NLMR* v.1.1.1 (Sciaini et al., 2018c).

We used *ggplot2* v.3.5.1 for creating figures (Wickham, 2016b), with the expansions: *patchwork* v.1.3.0 (Pedersen, 2024), *ggridges* v.0.5.6 (Wilke, 2024), *ggtext* v.0.1.2 (Wilke & Wiernik, 2022), and *ggdist* v.3.3.2 (Kay, 2023a).

We used *brms* v.2.22.0 (Bürkner, 2021) to run Bayesian models, with diagnostics generated using *bayesplot* v.1.11.1 (Gabry & Mahr, 2024), *tidybayes* v.3.0.7 (Kay, 2023b), and *performance* v.0.12.4 (Lüdtke et al., 2021).

We used the *dplyr* v.1.1.4 (Wickham et al., 2023), *tibble* v.3.2.1 (Müller & Wickham, 2023), *reshape2* v.1.4.4 (Wickham, 2007b), and *stringr* v.1.5.1 (Wickham, 2022) packages for data manipulation.

We used *sp* v.2.1.4 (Bivand, Pebesma & Gomez-Rubio, 2013), *adehabitatHR* v.0.4.22 (Calenge & Scott Fortmann-Roe, 2023), *move* v.4.2.6 (Kranstauber, Smolla & Scharf, 2023) for manipulation of spatial data and estimation of space use not otherwise mentioned in the methods.

We used *rmarkdown* v.2.29 (Xie, Allaire & Golemund, 2018; Xie, Dervieux & Riederer, 2020; Allaire et al., 2023), *bookdown* v.0.42 (Xie, 2016, 2022a), *tinytex* v.0.54 (Xie, 2019, 2023a), *kableExtra* v.1.4.0 (Zhu, 2021), and *knitr* v.1.49 (Xie, 2014, 2015, 2023b) packages to generate type-set outputs.

We generated R package citations with the aid of *grateful* v.0.2.10 (Francisco Rodríguez-Sánchez, Connor P. Jackson & Shaurita D. Hutchins, 2023).

2.6.3 Data availabilty

The code used to simulate the movement data, perform the multiverse, and analyse the results is available at: <https://github.com/BenMMarshall/multiverseHabitat>, and archived at <https://doi.org/10.5281/zenodo.12169335>. These locations also include model outputs and results not otherwise displayed here.

2.6.4 Author contributions

Conceptualization: BMM; Software: BMM; Supervision: ABD; Visualization: BMM; Writing - Original Draft: BMM; Writing - Review & Editing: BMM, ABD.

2.7 Supplementary Material

Table 2.1: All R^2 values for Bayesian regression models.

R^2	Standard error	Confidence interval	CI Lower	CI Upper	Component	Response	Model Method
0.097	0.001	0.95	0.095	0.099	Conditional	Absolute Deviation	rsf
0.034	0.000	0.95	0.033	0.035	Marginal	Absolute Deviation	rsf
0.078	0.001	0.95	0.077	0.080	Conditional	Raw Deviation	rsf
0.029	0.000	0.95	0.028	0.030	Marginal	Raw Deviation	rsf
0.279	0.002	0.95	0.274	0.284	Conditional	Absolute Deviation	wides
0.023	0.001	0.95	0.021	0.025	Marginal	Absolute Deviation	wides
0.302	0.002	0.95	0.297	0.306	Conditional	Raw Deviation	wides
0.030	0.001	0.95	0.029	0.032	Marginal	Raw Deviation	wides
0.225	0.004	0.95	0.218	0.232	Conditional	Absolute Deviation	ssf
0.064	0.002	0.95	0.059	0.068	Marginal	Absolute Deviation	ssf
0.226	0.004	0.95	0.219	0.233	Conditional	Raw Deviation	ssf
0.062	0.002	0.95	0.057	0.066	Marginal	Raw Deviation	ssf
0.088	0.019	0.95	0.051	0.125	Conditional	Absolute Deviation	wrsf
0.021	0.010	0.95	0.005	0.043	Marginal	Absolute Deviation	wrsf
0.168	0.022	0.95	0.125	0.212	Conditional	Raw Deviation	wrsf
0.016	0.008	0.95	0.003	0.032	Marginal	Raw Deviation	wrsf

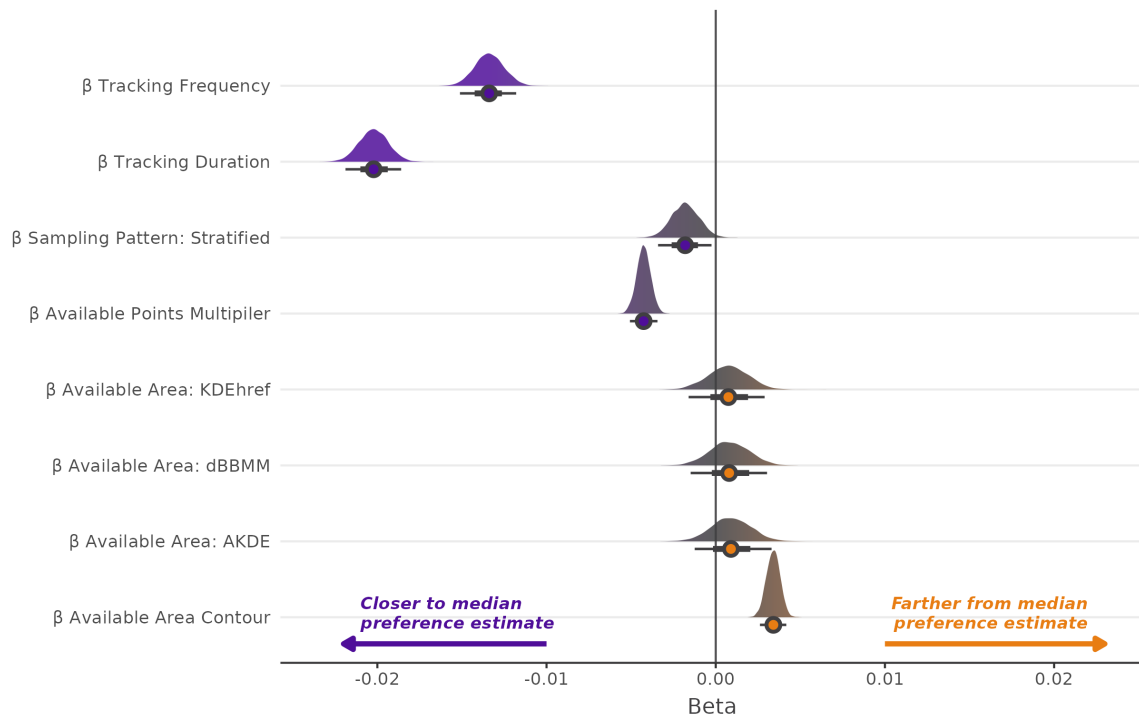


Figure 2.11: The posterior distributions of the betas from the Bayesian regression model looking to explain absolute deviation from the median in the Wides analysis pathway.

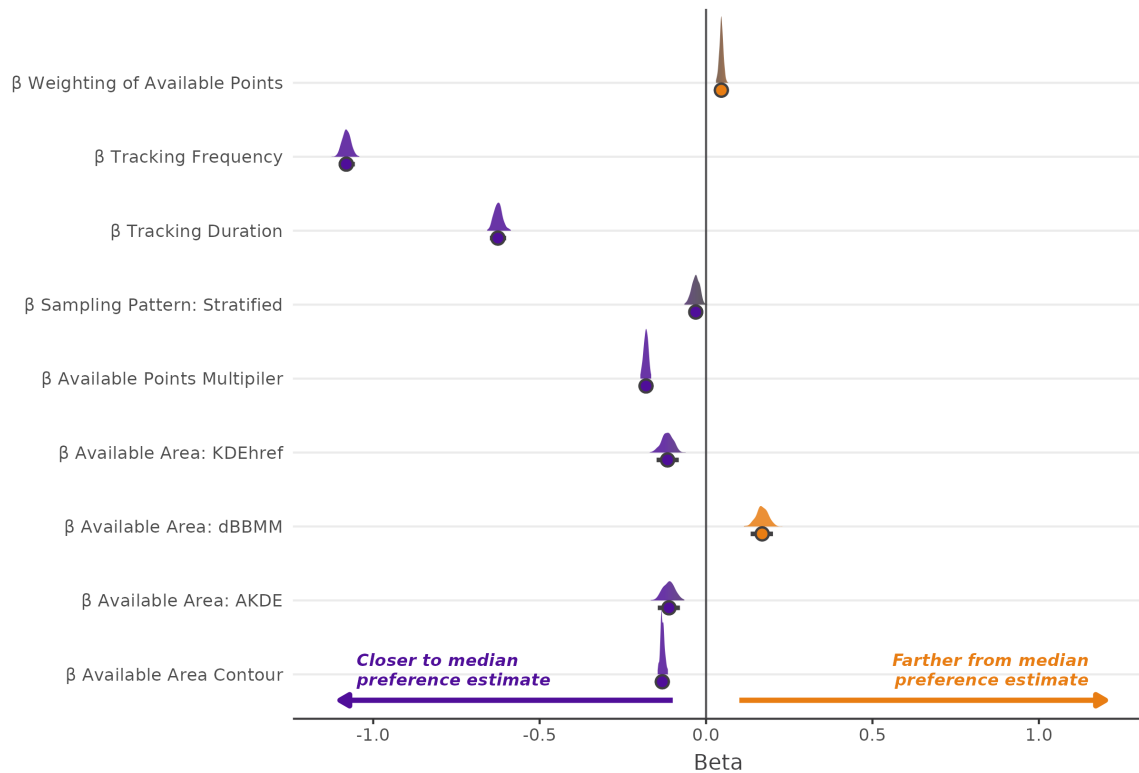


Figure 2.12: The posterior distributions of the betas from the Bayesian regression model looking to explain absolute deviation from the median in the RSF analysis pathway.

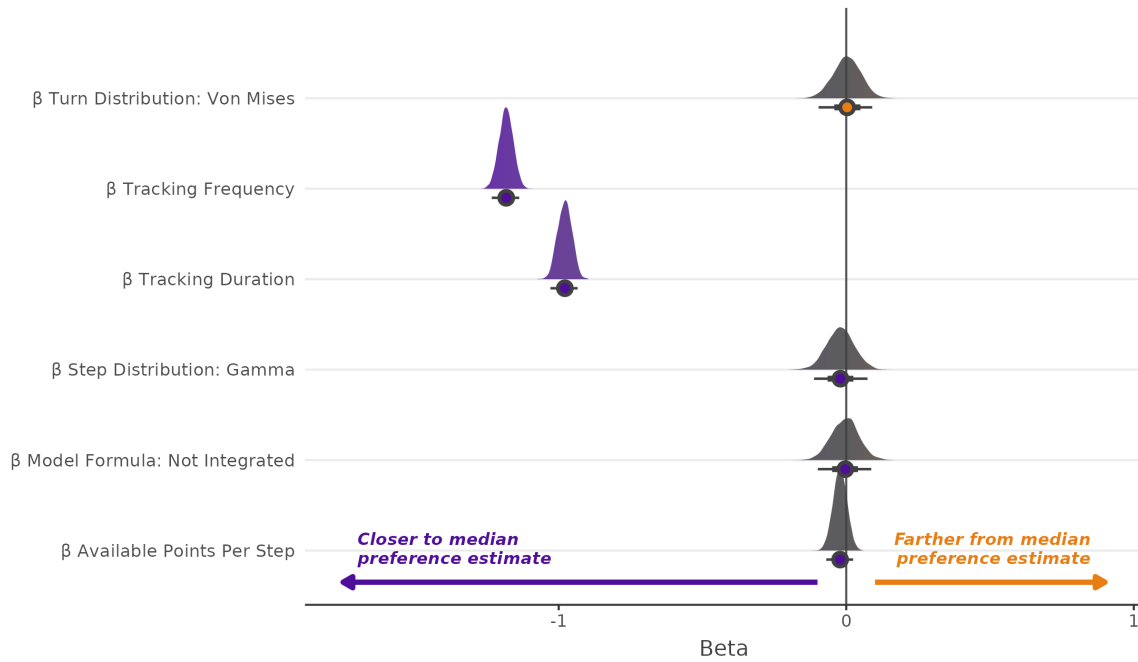


Figure 2.13: The posterior distributions of the betas from the Bayesian regression model looking to explain absolute deviation from the median in the SSF analysis pathway.

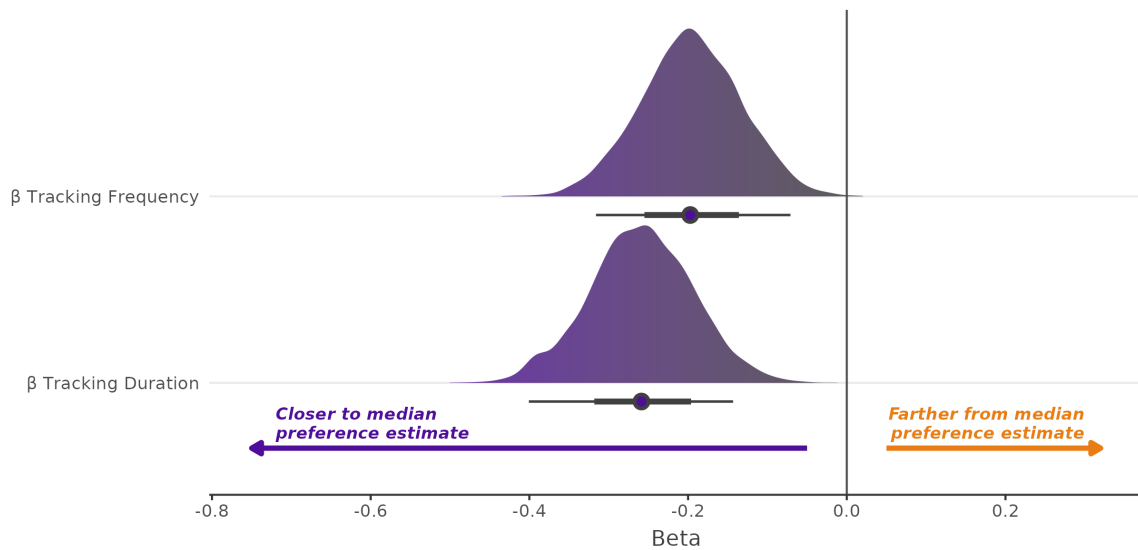


Figure 2.14: The posterior distributions of the betas from the Bayesian regression model looking to explain absolute deviation from the median in the wRSF analysis pathway.

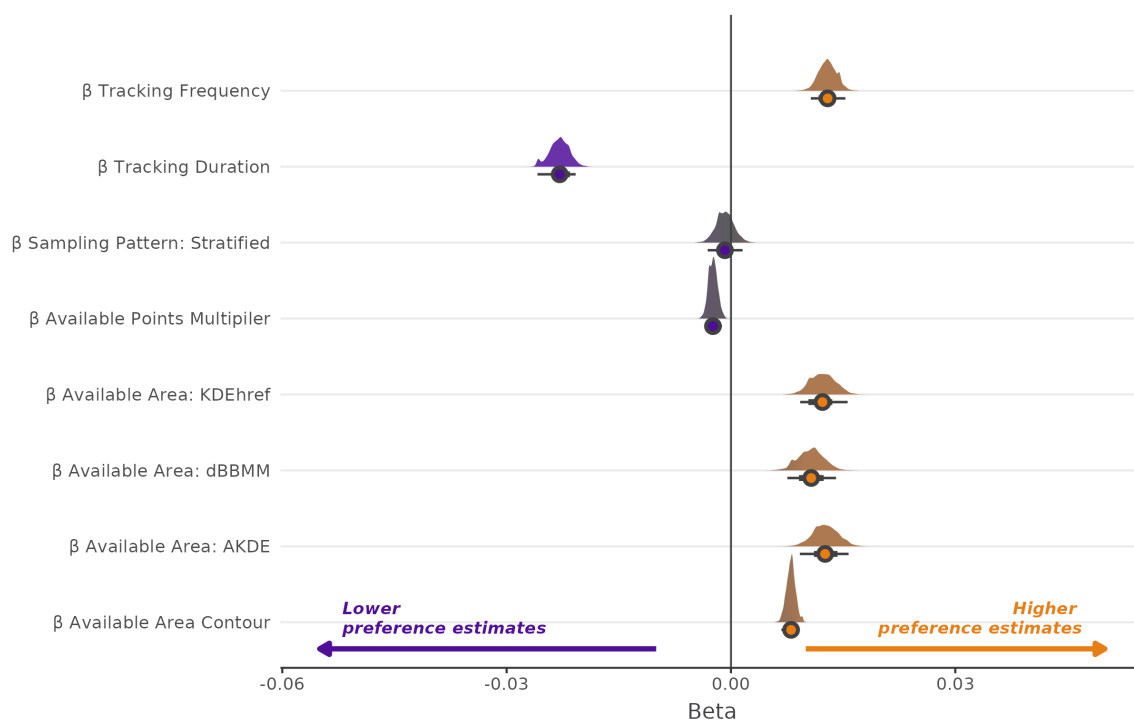


Figure 2.15: The posterior distributions of the betas from the Bayesian regression model looking to explain raw deviation from the median in the Wides analysis pathway.

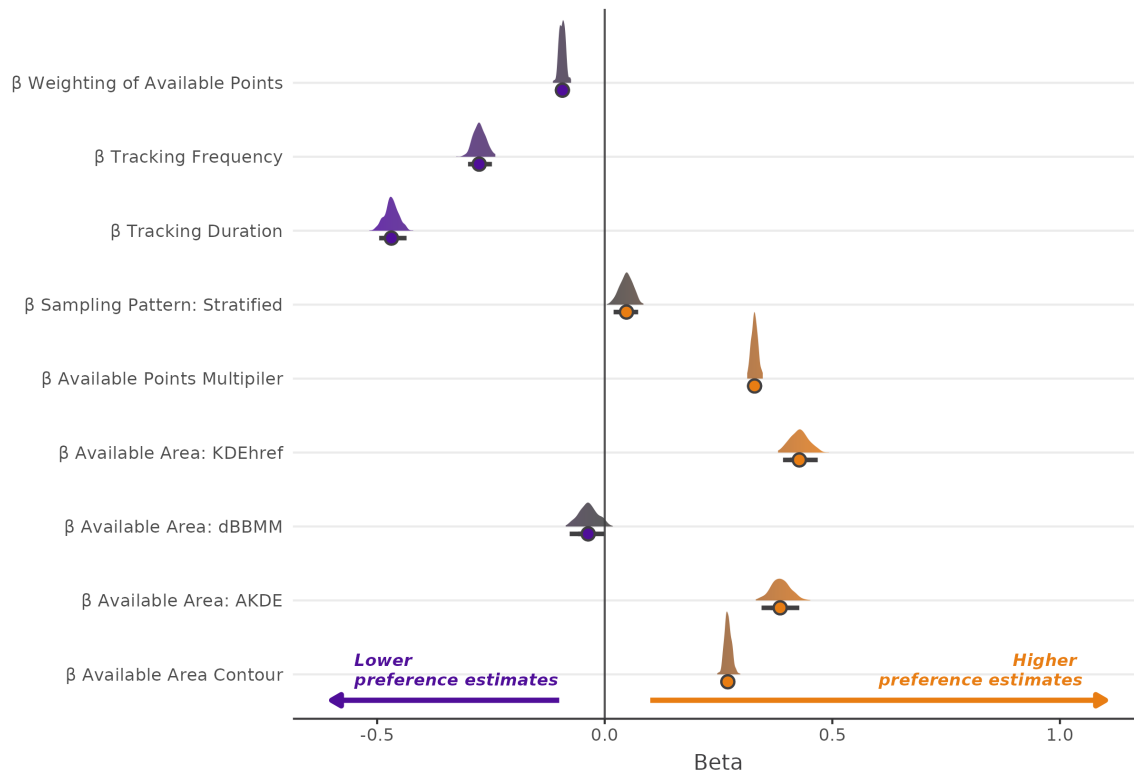


Figure 2.16: The posterior distributions of the betas from the Bayesian regression model looking to explain raw deviation from the median in the RSF analysis pathway.

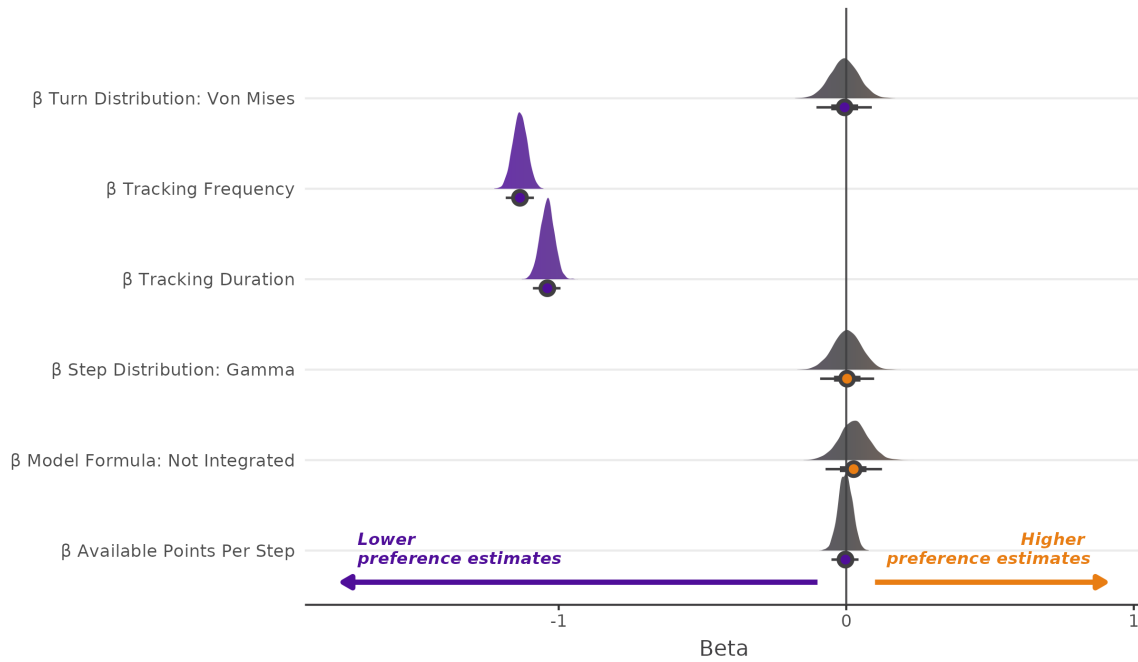


Figure 2.17: The posterior distributions of the betas from the Bayesian regression model looking to explain raw deviation from the median in the SSF analysis pathway.

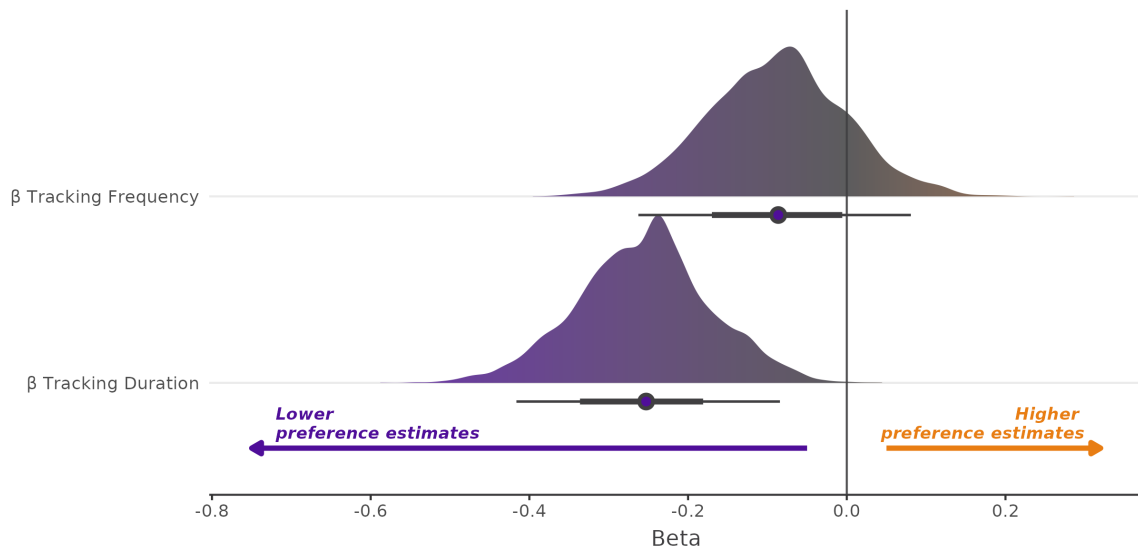


Figure 2.18: The posterior distributions of the betas from the Bayesian regression model looking to explain raw deviation from the median in the wRSF analysis pathway.

Table 2.2: Analysis results when applied directly to un-resampled movement data at both the decision timescale and movement timescales.

Estimate	Standard error	Analysis method	Simulation scale	Individual	Species
0.019	0.059	rsf	Movement	1	BADGER
1.000	NA	wides	Movement	1	BADGER
0.078	0.076	ssf	Movement	1	BADGER
0.281	0.223	rsf	Destination	1	BADGER
1.087	NA	wides	Destination	1	BADGER
-0.002	0.223	ssf	Destination	1	BADGER
-0.002	0.011	rsf	Movement	1	VULTURE
1.002	NA	wides	Movement	1	VULTURE
0.005	0.016	ssf	Movement	1	VULTURE
0.344	0.065	rsf	Destination	1	VULTURE
1.202	NA	wides	Destination	1	VULTURE
-0.114	0.066	ssf	Destination	1	VULTURE
-0.009	0.010	rsf	Movement	1	KINGCOBRA
1.001	NA	wides	Movement	1	KINGCOBRA
-0.022	0.017	ssf	Movement	1	KINGCOBRA
0.081	0.048	rsf	Destination	1	KINGCOBRA
1.082	NA	wides	Destination	1	KINGCOBRA
-0.172	0.051	ssf	Destination	1	KINGCOBRA
0.014	0.008	rsf	Movement	2	BADGER
1.002	NA	wides	Movement	2	BADGER
0.046	0.014	ssf	Movement	2	BADGER
0.357	0.054	rsf	Destination	2	BADGER
1.179	NA	wides	Destination	2	BADGER
-0.072	0.056	ssf	Destination	2	BADGER
0.004	0.013	rsf	Movement	2	VULTURE
1.006	NA	wides	Movement	2	VULTURE
0.027	0.019	ssf	Movement	2	VULTURE
0.293	0.065	rsf	Destination	2	VULTURE
1.158	NA	wides	Destination	2	VULTURE
-0.183	0.066	ssf	Destination	2	VULTURE
-0.001	0.018	rsf	Movement	2	KINGCOBRA
1.001	NA	wides	Movement	2	KINGCOBRA
0.017	0.028	ssf	Movement	2	KINGCOBRA
0.322	0.090	rsf	Destination	2	KINGCOBRA
1.057	NA	wides	Destination	2	KINGCOBRA
0.150	0.094	ssf	Destination	2	KINGCOBRA
0.021	0.017	rsf	Movement	3	BADGER
1.001	NA	wides	Movement	3	BADGER
0.027	0.025	ssf	Movement	3	BADGER
0.789	0.086	rsf	Destination	3	BADGER
1.152	NA	wides	Destination	3	BADGER
0.446	0.087	ssf	Destination	3	BADGER
0.000	0.015	rsf	Movement	3	VULTURE
1.002	NA	wides	Movement	3	VULTURE
0.001	0.022	ssf	Movement	3	VULTURE
0.306	0.067	rsf	Destination	3	VULTURE
1.173	NA	wides	Destination	3	VULTURE
-0.181	0.068	ssf	Destination	3	VULTURE
0.011	0.012	rsf	Movement	3	KINGCOBRA
1.002	NA	wides	Movement	3	KINGCOBRA
0.038	0.019	ssf	Movement	3	KINGCOBRA
0.194	0.056	rsf	Destination	3	KINGCOBRA
1.094	NA	wides	Destination	3	KINGCOBRA
-0.017	0.059	ssf	Destination	3	KINGCOBRA
0.079	0.061	rsf	Movement	4	BADGER
0.999	NA	wides	Movement	4	BADGER
0.124	0.084	ssf	Movement	4	BADGER
-0.117	0.125	rsf	Destination	4	BADGER
1.081	NA	wides	Destination	4	BADGER
-0.417	0.127	ssf	Destination	4	BADGER
0.007	0.011	rsf	Movement	4	VULTURE
1.007	NA	wides	Movement	4	VULTURE
0.032	0.017	ssf	Movement	4	VULTURE
0.348	0.066	rsf	Destination	4	VULTURE
1.186	NA	wides	Destination	4	VULTURE
-0.139	0.067	ssf	Destination	4	VULTURE
0.004	0.018	rsf	Movement	4	KINGCOBRA
1.001	NA	wides	Movement	4	KINGCOBRA
0.006	0.028	ssf	Movement	4	KINGCOBRA
0.212	0.085	rsf	Destination	4	KINGCOBRA
1.069	NA	wides	Destination	4	KINGCOBRA
0.049	0.089	ssf	Destination	4	KINGCOBRA

Chapter 3

Population-level Habitat Selection Analyses Produce Consistent Effect Direction, but Strength is Sensitive to Researcher Choice

Benjamin M. Marshall¹, Alexander B. Duthie¹

¹Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, FK9 4LA, UK

3.1 Abstract

Science strives to be repeatable and reproducible, but we know that different researchers will reach different conclusions even when presented with the same data and questions. This variation due to different choices researchers make during study and analysis may play a role in increasing variation in results. Variation in results may be particularly pronounced in disciplines with high task uncertainty and difficult-to-control experimental environments, such as ecology. Ecological research is often concerned with conservation, and a key component of animal conservation is determining animals' habitat requirements that can steer effective habitat protection. Insights into habitat requirements can be gained via the examination of animal movement, but the translation of movement data to usable estimates of habitat

selection can be approached in many ways. We use a multiverse approach to explore a suite of possible habitat selection approaches and determine the chances of incorrectly identifying habitat selection in two simulated scenarios. We find that the majority of different analysis pathways correctly identify selection, but the strength of that selection is susceptible to analysis choice. We identify several choices that have disproportionate impacts on selection estimates, with some choices having contrasting impacts depending on the analysis approach. Overall our results suggest that while we can be broadly confident in the conclusions drawn from habitat selection studies, we cannot ignore analysis choices, especially when the strength of selection is of particular concern.

3.2 Introduction

The successful application of science to policy relies on results being repeatable, reliable, and generalisable (Community, 2025). Efforts to achieve this often consist of removing idiosyncrasies of studies; we remove confounding variables, control for non-target variables; and conduct replicants. Despite our best efforts, we can never create a fully generalisable study.

The vast majority of the time our efforts to account for study-specific variation appear adequate, resulting in answers that can form the foundations for further studies. However, the re-examination of foundational work reveals the need to repeatedly question and retest the assumptions subsequent work relies upon (Open Science Collaboration, 2015). For example, this need for reassessment can be seen in studies of bird plumage and its role in mate selection, where false positive rates in published studies may have elevated the importance of plumage colour and pattern size in mate selection (Seguin & Forstmeier, 2012; Parker, 2013; Sánchez-Tójar et al., 2018). Subsequent meta-analyses have highlighted the need to be more cautious with interpreting initial findings, and that future hypotheses require refinement.

While other disciplines' replication rates have been more closely examined (Freedman, Cockburn & Simcoe, 2015; Open Science Collaboration, 2015; see Kelly, 2019 for ecological replication rates), there is evidence that ecology and evolution studies require similar scrutiny. Ecology and evolution not uniquely free of the same deleterious incentive structures that are suggested to be responsible for concerns surrounding replication in other disciplines, such as a prioritisation of novel findings and clear significant research narratives (Fanelli, 2010a; Ware & Munafò, 2015; Smaldino & McElreath, 2016; Brembs, 2018). Replication studies and

meta-analyses have shown that results in ecology and evolution may need further scrutiny (Seguin & Forstmeier, 2012; Parker, 2013; Sánchez-Tójar et al., 2018; Clements et al., 2022). In general, it appears that initial findings are more extreme than subsequent findings, and that their initial momentum in regards to research direction is slow to reorient (Jennions & Møller, 2002a; Barto & Rillig, 2012).

For disciplines where direct replications are more common and feasible, replication can play a larger feedback role (Freedman, Cockburn & Simcoe, 2015; Open Science Collaboration, 2015; Peterson & Panofsky, 2021). Ecological systems are complex and impossible to fully control, leading to ecologists' frequent reliance on "natural experiments".

Natural experiments by their nature are impossible to replicate perfectly, forcing ecology to rely on conceptual or quasi replications (Palmer, 2000; Fraser et al., 2020). These replications may have subjects and methods that differ from the original study, or where the research questions are only conceptually similar (Nakagawa & Parker, 2015; Kelly, 2019). These replication efforts are still valuable, and valued by the scientific community (Fraser et al., 2020), as they aid gauging the reliability of proposed patterns, theories, or laws.

Ecology and evolution, due to the high task uncertainty and limited options for direct replication in the wild, tends towards a system of organic self-correction, where the complexity of study systems can explain away the differences between studies (i.e., inconsistencies in replication), leaving the integration of past and subsequent findings more open to interpretation (Peterson & Panofsky, 2021).

Animal movement, like many aspects of ecology, is a field of study in which replications are difficult to conduct. Tracking animals is costly and comes with serious ethical considerations, especially for animals where the tracking equipment can impact their health, mobility, and behaviour (Sperry et al., 2009; Homberger et al., 2021; Portugal & White, 2022). Therefore, there is a strong incentive to extract the maximum value for every tracked animal and to ensure that all results are as reliable as possible.

The findings from animal movement studies can be linked with landscape management and integrated into conservation plans (Fraser et al., 2018a; Doherty & Driscoll, 2018). A key component for protecting species is identifying that species' requirements; the examination of movement can reveal those requirements (Mueller et al., 2011; Doherty, Fist & Driscoll, 2019). Suitable habitat (as defined in various ways) is a fundamental requirement, and the protection/restoration of suitable habitat is an actionable conservation response (Intern-

tional Union for Conservation of Nature and Natural Resources & International Union for Conservation of Nature and Natural Resources, 2001). Reliable determinations of suitable habitat are therefore required for correct conservation decisions.

As demonstrated by “many analysts” projects, different answers can originate from the same dataset, and even the same question (e.g., Silberzahn et al., 2018; Huntington-Klein et al., 2021). Such “researcher degrees of freedom” [or flexibility in analysis; Forstmeier, Wagenmakers & Parker (2017)] can be the result of differences in question interpretation (Auspurg & Brüderl, 2021) and the analysis approach taken (Gelman & Loken, 2013; Bastiaansen et al., 2020). When drawing information to be used in a conservation plan, it would be advisable to know whether decisions would be different had the data been examined in a different way. Previous explorations have suggested that researcher degrees of freedom are sufficiently large to alter effect strength and the final conclusions (Salis, Lena & Lengagne, 2021; Desbureaux, 2021).

While many analysts projects represent an ideal, actually capable of capturing the variation in results stemming from the researchers themselves, the logistics of conducting them limits their everyday practicality [although the Many* initiatives further supports the feasibility of integrating the efforts of researchers; e.g., ManyDogs Project et al. (2023); Many Primates et al. (2019); Lambert et al. (2022)]. An alternative, albeit with its own set of limitations, would be to conduct a multiverse analysis (Steege et al., 2016). Multiverse analyses consist of a suite of analysis pathways, where diverging analysis decisions create a compounding set of possible answers to a given question. Multiverses cannot provide evidence for the correct answer, but can allow researchers to explore the sensitivity of their conclusions to any number of analytical variables (Rijnhart et al., 2021). Part of this exploration can disentangle the variation stemming from different sampling or analysis choices, compared to variation stemming directly from the effect of interest (Simonsohn, 2015; Steege et al., 2016; Forstmeier, Wagenmakers & Parker, 2017). This can be particularly true when we use the multiverse in a synthetic or simulated environment where we are able to hold the effect of interest constant.

Previously we explored how individual habitat selection estimates were impacted by decisions regarding sampling and analysis [Marshall & Duthie (2024); Chapter 2]. Broadly we found that data quantity was more important than analysis decisions when estimating habitat selection, and that more modern analysis methods tend to produce less variable results. Here we build upon those findings to target population level habitat selection analyses; i.e., an

habitat selection estimate that summarises across a number of sampled individuals. We use simulated animal data with a known habitat preference to explore the impacts of sampling and analysis choice in recovering the direction of that predefined preference.

3.3 Methods

3.3.1 Simulating the Scenarios

We used the `abmAnimalMovement` v.0.1.3.0 (Marshall & Duthie, 2022) R package to simulate the movement data of an animal with a predefined (i.e., known) habitat preference. The `abmAnimalMovement` package uses an agent-based approach to simulate terrestrial animal movement using raster environmental data to guide the animals' decisions. We used the `NLMR` v.1.1.1 package (Sciaini et al., 2018c) to generate the three required resource/environmental rasters: movement resistance, foraging quality, and shelter site quality. The `abmAnimalMovement` package has systems for simulating activity cycles, three separate behavioural states (resting, foraging, exploring; differing in movement characteristics and resource prioritisation), and site fidelity. For the purposes of this study, we used the pre-created pseudo-species presented in Marshall & Duthie (2022).

1. Badger. The badger is a terrestrial species subject to differing movement resistance across the landscape occupying several shelter sites, with an 8-12 hour activity cycle and minor seasonal variation.
2. Vulture. The vulture is treated as a terrestrial species but does not experience any movement resistance across the landscape. It uses three shelter sites, with an 12 hour activity cycle and minor seasonal variation. Unlike the randomly uniform resource landscape of the badger, the vulture's landscape is divided into two distinct areas of higher and lower foraging quality.
3. King Cobra. The king cobra is a terrestrial species subject to differing movement resistance across the landscape similar to the badger. It uses 12 shelter sites, and has a more complex activity cycle with an 8-12 hour diel cycle, an approximately weekly cycle of resting to mimic long digestion periods, and a weak seasonal cycle all working additively to define resting behaviour. The landscape is universally random similar to the badger, but contains two lines of heightened movement resistance serving as movement barriers.

For the purposes of the analysis we simplified the landscape information into categories – akin to the sort of land-use information more frequently available to researchers of animal movement. We focused on foraging quality because it influences the greatest amounts of movement (in terms of time) compared to sheltering or exploratory movements. We converted the continuous foraging quality raster into a binary, where higher quality areas (greater than 0.5) are classed as 2, and lower quality areas as 0. This classified landscape raster served as the basis for assessing habitat selection, in a scenario with a known preference for habitat 2. We also ran all described multiverse analyses (see below) with a second landscape classification. In this second scenario we scrambled the values of the cells, thereby breaking the connection between the underlying raster layers used to simulate the movement and the classification used in the analysis. This scrambled habitat classification allowed us to examine the variation in estimates in a zero-selection scenario, while serving as a direct comparison for the selection scenario.

We used that simulated landscape and `abmAnimalMovement` to simulate a population of 25, which was later sampled from for collecting movement data. All individuals in this population had the same simulation settings apart from starting location. Therefore, any variation between individuals was due to stochasticity rather than variation in the predefined habitat preference.

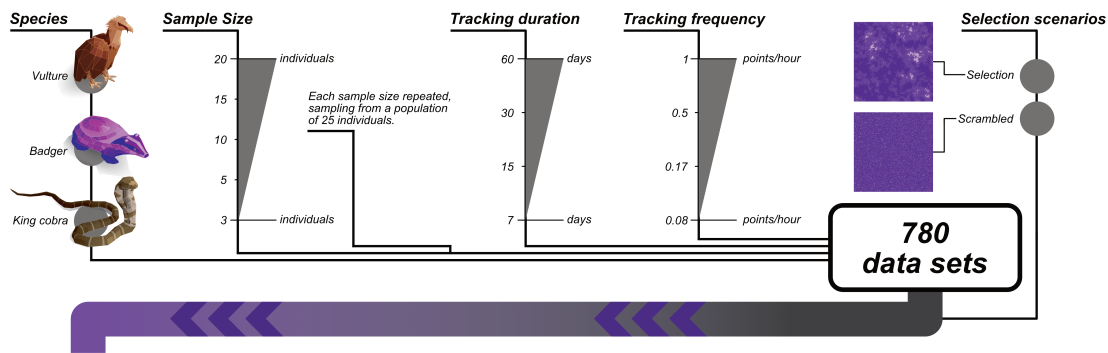
3.4 Sampling and Analysis Options

To manage the sampling of the population and the compounding growth of subsequent analysis decisions, we used the `targets` v.1.9.0 and `tarchetypes` v.0.11.0 R packages (Landau, 2021a,b). These packages allowed a branching workflow pipeline, while keeping track of object creation thereby optimising the compute time required to explore the multiverse of analysis choice (Fig. 3.1).

Sampling

The first decision in most animal movement studies will concern tracking regime, consisting of decisions regarding tracking duration (the overall time the animal’s movements are tracked) and tracking frequency (how frequently the animal is relocated). This decision is often dictated by more practical considerations such as anatomy and behaviour of the animal, cost of the

Sampling



Analysis

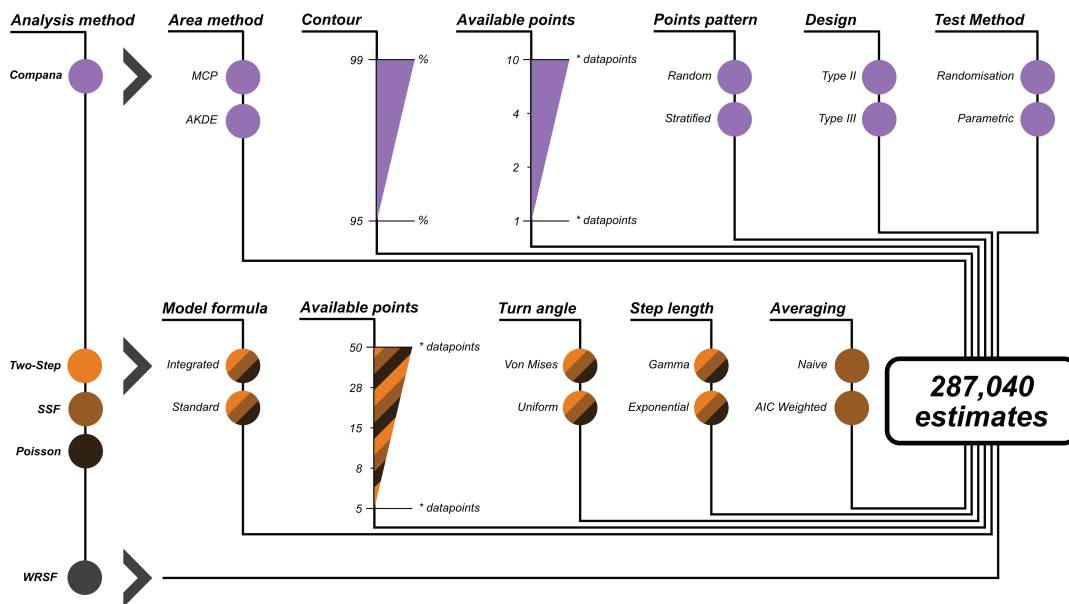


Figure 3.1: A diagram of the different decisions creating a multiverse examining animal movement. Species: three different movement archetypes simulations were based on. Sample Size: the number of simulated individuals per each species, with each random sample being repeated twice. Tracking duration: the length (days) of the tracking data sampled from the overall simulated year. Tracking frequency: the frequency of sampled data points from the 1-per-minute simulated data. Selection scenarios: the two classified habitat landscapes used during analysis. Purple choices highlight the analysis pathways using area-based estimations of availability. Area method: the method used to generate the available area polygon. Contour: the threshold used to convert the utilisation/occurrence distribution into a polygon. Available points: the number of generated points per data point to extract available habitat within the polygon. Points pattern: whether the random points were generated in a stratified or random pattern. Light orange choices pertain to the Two-Step analysis pathways. Medium orange-brown choices pertain to the SSF analysis pathways. Dark brown choices pertain to the Poisson analysis pathways. Model formula: whether the step and turn angles were added as interactions to the habitat variable. Available points: the number of randomly generated available points per step/used point. Turn angle: the distribution used to generate the turn angles of the available points. Step distribution: the distribution used to generate the step distances of the available points. Dark orange is the WRSF analysis. Averaging: the method of model averaging used for the SSF models to gain an overall estimate. Decisions that required discrete choices are illustrated with circles, whereas those that continuous are depicted as triangles.

tracking devices, and environmental factors. Here we covered a range of tracking regimes that vary in the frequency of location fixes (1 to 0.08 points per hour), and the total duration of tracking (7 to 60 days). We created sub-sampled datasets based on every combination of tracking frequencies and durations, provided they would result in greater than 30 data points per individual.

An important component of assessing population-level habitat selection is the number of individuals included in analysis. Therefore, we randomly generated a number of samples from our population of 25 simulated individuals. We varied these sample sizes from 3 to 20 individuals, and ran 2 repeats for each size. A sample never mixed tracking regimes.

Analysis

Building on the decisions concerning tracking regime and population sampling, our multiverse expanded dramatically by exploring four primary analysis routes. These routes included an area-based approach, using Compana analysis (Compositional Analysis of Habitat Use), and three step-based approaches including averaged individual Step selection models (SSF), Two-Step conditional regression models (Two-Step) and a Poisson model.

We ran the Compana analysis using the *adehabitatHS* v.0.3.18 (Johnson, 1980; Aebischer, Robertson & Kenward, 1993; Calenge & Mathieu Basille, 2023) R package. Compana allows for the assessment of habitat selection of multiple animals in defined habitat types; it also allows habitat selection to be estimated at different scales. We explored two scales in the multiverse: selection within the home range (Type III), and selection within an overall area available to the population (Type II). The former design (Type III) requires the available habitats to be defined on an individual-by-individual basis; therefore, we are comparing individual use to an individual's defined available habitats. While the latter (Type II) design uses a summarised population level availability; therefore, we are comparing individual use to the habitats available to the population overall.

To define availability, we created home range polygons, then sampled points within, recovering the corresponding habitat at those points. The home range (AKA availability) polygons can be generated via many different processes. We explored MCP and AKDE (Minimum Convex Polygons, Autocorrelated Kernel Density Estimators). We used the *ctmm* v.1.2.1 (Fleming & Calabrese, 2023) package for the creation of AKDEs, and the *adehabitatHR*

v.0.4.22 (Calenge & Scott Fortmann-Roe, 2023) package for the MCPs. The AKDEs required the fitting of a movement model; we fitted multiple models and selected the top scoring model by AIC. For the best model, we followed the recommendations of Silva et al. (2022), opting to run AKDEs as weighted and using the perturbative hybrid residual maximum likelihood method (pHREML), as both these options tend to be the most robust in scenarios of lower tracking frequencies. There are many options involved in the creation of AKDEs that could shape habitat availability and therefore estimations of selection. We avoided a deeper exploration of the impact of the internal AKDE choices, as they are likely smaller than those between area methods (e.g., compared to MCPs; Silva et al. (2020)), and the relatively greater computational cost to generate many AKDEs. By comparison the MCPs have very few options impacting their creation (Calenge & Scott Fortmann-Roe, 2023). Mainly MCP size varies based on the percentage of outliers excluded from the polygon area. For all area methods, we vary this percentage: 95%, 99%, that correspond to increasingly large areas included in the availability polygon (i.e., covering areas that are less used by the individual). We selected these contour values as they are the most frequently used in spatial studies (Crane et al., 2021).

Once we had created all area polygons for each individual, for every combination of tracking regime (duration x frequency, total regimes: 13), we created the population available areas for the type II design analysis by combining all polygons into a single polygon. We used this combined polygon for the type II design, where all individuals have the same availability.

We used each availability polygon to define the extent within which we sampled habitat characteristics. Within those defined polygons we generated a number of random points per animal location to determine the available habitats. We varied both the number of points generated relative to the number of animal locations (1 to 10 times the number of known animal locations), as well as the pattern of how they were distributed within the polygon (random or stratified). Once the points were generated, we used them to extract the underlying habitat from the classified landscape raster.

The final decision in the area-based method approach was whether the Compana analysis was tested using randomisation or in a parametric fashion. The former uses randomisation tests to estimate the habitat selection, while the latter uses a chi-squared distribution.

The other analysis pathways do not rely on an available area; instead they focus on randomly generated steps that mirror the observed movement of the animal. As such, the decisions

involved in running the step-selection, two-step conditional regression, and Poisson models are largely the same.

The first decision is the number of random steps generated per observed step; we ranged this from 5 to 50. The generation of random steps is guided by two distributions, one describing the step lengths from the last observed location, and another describing the turn angle from the last observed direction of travel. We chose to explore the impacts of different distributions. For the step length, we used Gamma and exponential distributions; for the turn angle we used Von Mises and uniform distributions.

All the step-based methods require the definition of a formula, where use/availability is predicted by the habitat characteristics at those locations. Model formulation opens up many possible alternative routes of analysis, but for this exploration we focused on the impacts of integrating step and turn angle interactions with habitat characteristics. We ran two versions of all models, one with no interactions and one where both step lengths and turn angles interact with the habitat classification. Previous work has suggested that the integrated formulation (i.e., model with interactions) performs better, producing less biased estimates of habitat selection (Forester, Im & Rathouz, 2009).

A key component of all of the population approaches is how they translate a highly structured dataset into an overall summary of habitat selection. We ran individual step selection models using the *amt* v.0.2.2.0 (Signer, Fieberg & Avgar, 2019) package, to explore summary methods (i.e., averaging regression coefficients Thurfjell, Ciuti & Boyce (2014)). The individual step selection models need to be averaged in some manner to extract a population mean selection. We explored two simple options for achieving this: a naive mean of the final estimates (i.e., a mean of the estimated betas), and a model average using the *MuMin* v.1.48.4 package (Bartoń, 2023). The *MuMin* model average approach is weighted based on AICc. However, AICc is not directly comparable between models fitted with different datasets, so this does not necessarily present a solution for real-world scenarios as it is unlikely that multiple animals would have identical amounts of locations recorded. In the simulated scenario here, as the model formulas are identically complex and the datasets equal in data quantity, the differences in AIC are purely describing the goodness of fit to the different individuals. Therefore, the model average produced by *MuMin* is weighted by model goodness of fit.

The difficulties in model averaging based on different data have led to the creation of models that account for individual variability in the model formula. A Two-Step Conditional Regression

is one such solution, where strata (time steps) and clusters (individual animals) are accounted for (Craiu et al., 2011). We implemented these models using the *TwoStepCLogit* v.1.2.5 (Craiu et al., 2016) package, again using the same decisions as applied in the SSF approach (except model averaging). The Two-Step implementation in *TwoStepCLogit* required no NAs in landscape covariates (i.e., all locations had associated classified habitat information). In the rare instances where a simulated animal reached the edge of the landscape and available points were generated beyond the boundary (i.e., return a habitat class of NA), we repeated the random generation of points until fewer than 10 NAs existed in the data that were then excluded.

More recently Muff, Signer & Fieberg (2020) have suggested a reformulation of the step selection models to provide faster, reliable estimations of population habitat selection. The suggestion is to reformulate the model as a Poisson model, but critically with stratum-specific fixed effects. As this would result in a large number of fixed effects, the computational cost would be large; so instead they are treated as a random effect, but to avoid excessive shrinkage to the mean a large variance is pre-set. To avoid the slow convergence of MCMC estimation, Muff, Signer & Fieberg (2020) made use of integrated nested Laplace approximation (INLA) to approximate the Bayesian inference required for such a reformulation. We followed the code provided by Muff, Signer & Fieberg (2020) and made use of *INLA* v.24.6.27 (Rue, Martino & Chopin, 2009; Lindgren, Rue & Lindström, 2011; Martins et al., 2013a; Rue et al., 2017; Kourounis, Fuchs & Schenk, 2018) to run Poisson models to estimate habitat selection. We applied the same suite of choices as applied to the SSF and Two-Step approaches.

3.4.1 Assessing the multiverse

Specification curves provide an overview of the estimates of a given range of analyses (Simonsohn, Simmons & Nelson, 2020). Tighter more steep curves suggest greater agreement between all the analysis end points. Here we use specification curves and plot the estimates against the different decisions that result in those estimates, allowing direct comparison on how the sampling and analysis decisions impact the variation in the estimates.

To better detect the impact of decisions, while accounting for the random variation stemming from the differences in individuals/samples, we ran a number of Bayesian Regression Models. The Bayesian Regression Models aimed to describe how much of the deviation from a median answer could be explained by the various sampling and analysis decisions. For each analysis

route, we ran a model that included tracking frequency, tracking duration, sample size, and all the corresponding analysis choices. All continuous variables (x) were scaled to help determine their relative importance to each other: $(x - \text{mean}(x))/\text{sd}(x)$.

For the area-based Compana approach, population effects included: the continuous variable contour (continuous); and the cataric predictors sampling pattern (random, stratified), test (randomisation, parametric), area method (AKDE, MCP), and design type (II, III). For the step-based approaches they all included: model formula (integrated, non-integrated), step distribution (gamma, exponential), turn distribution (Von Mises, uniform). The step-selection model approach also included: averaging method (naive, MuMIn average).

We ran the Bayesian Regression Models using the *brms* v.2.22.0 (Bürkner, 2021) package. Once complete, we checked model convergence using R-hat values, ACF, and trace plots. After assessment of the model convergence measures, we modified the running parameters to 4000 iterations, a warmup of 750, and a thinning rate of 4 for all models bar the Poisson models that instead used 20000 iterations, a warmup of 8000, and a thinning rate of 20. We also examined the R^2 values, to determine how much of the variation in the results could be explained by the decisions.

3.5 Results

3.5.1 Specification Curves

The specification curves provide an overview of all the analysis end points ($n = 287,040$; Fig. 3.1). The area-based method – Compositional Analysis (Compana) – returned a median estimate that agreed with the underlying known simulation preference (i.e., preference for habitat 2; Fig. 3.2), and a near zero effect for the scrambled selection scenario. In the positive selection scenario very few estimates suggested negative selection, and those that did appeared to be extreme outliers. When broken down by decision, all decisions except higher sample sizes and the test method, appear to have a large range in final estimates. The sample size decisions had a clear reduction in possible answers when the sample was 15 or 20 individuals, but also a potential interaction with the test method when including 20 individuals. Unexpectedly, the lowest sample size of 3 also led to visibly reduced estimate variation, and the higher sample sizes had a lower rate of significant positive estimates (e.g., 6902/9984 significant results with

3 individuals, compared to only 254/9984 with 20 individuals). The test method presented two different outcomes, with the randomisation method presenting a far tighter grouping of estimates in both selection scenarios, and with higher rates of significance in the positive selection scenario (11663/24960 compared to 5432/24960). In the scrambled scenario, the grouping was reduced, but rates of significance were very similar (13985/24960 compared to 13175/24960; Table. 3.1).

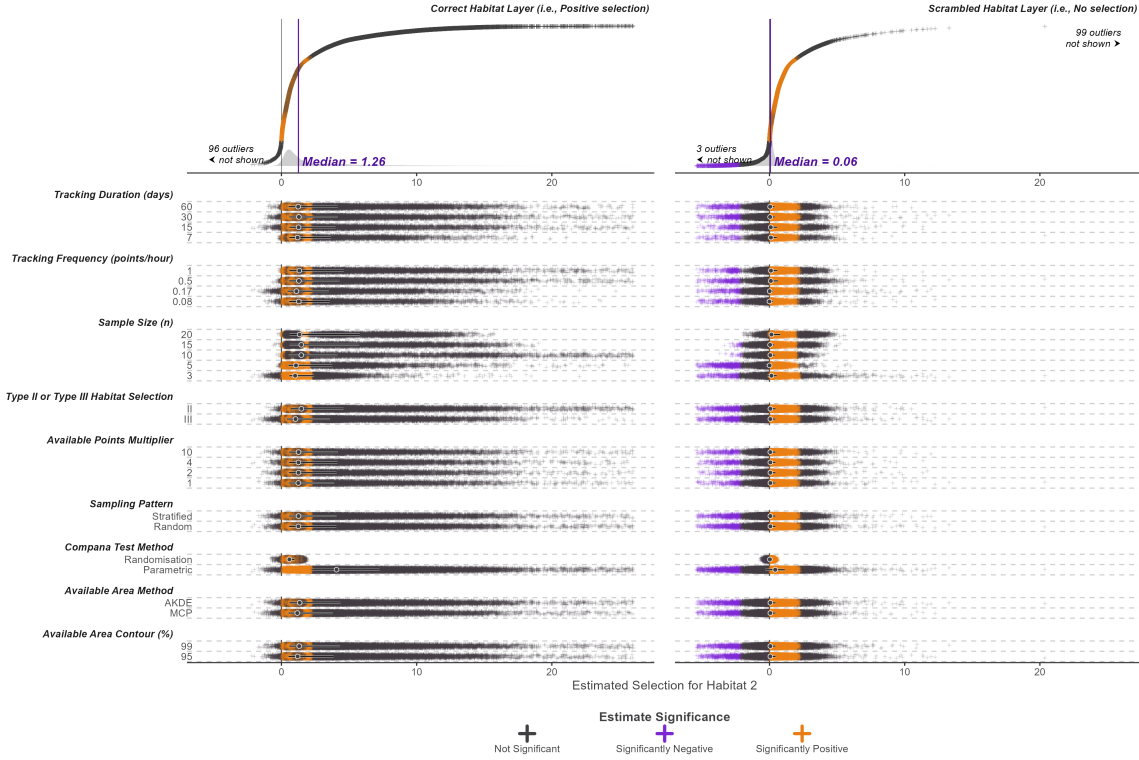


Figure 3.2: A specification curve showing habitat selection estimates resulting from the Compans (area-based) analysis pathways when applied to all simulated animal movement datasets. Every point represents a separate estimate of habitat selection. The top plot shows all estimates organised by estimate strength, with the purple vertical line depicting the median estimate and grey distribution describing the overall pattern of the estimates. Instances of outlying high or low estimates are labelled instead of plotted. The lower plot shows the same estimates, but split by the analysis or sampling decisions. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range.

The specification curve generated from summarised Step Selection Functions (SSFs) similarly revealed a median estimate that agrees well with the simulated parametrisation for both selection and scrambled scenarios (Fig. 3.3). Also like the Compans curve, the SSF curve had far fewer results that indicated very low habitat preference in the positive selection scenario. The majority of these lower estimates of habitat selection appeared connected to the decision to use a non-integrated (standard) model formulation, and the decision to average models

via the AIC weighted approach. The naively-model-averaged estimates were also much more tightly grouped with fewer estimates failing to correctly detect positive preference; the tighter grouping was also reflected in the scrambled selection scenario.

Outside the lower estimates, the clearest structure in the estimates results from different decisions in sample size and tracking regime. The tracking regime decisions, frequency and duration, appeared to shift the overall range of estimates but with limited impact on the variation. By comparison, sample size increases appeared to reduce the spread of estimates produced.

Compared to the Compana analysis there was a cleaner pattern of significance, with the selection scenario having much higher rates of significant selection compared to the scrambled scenario (Table. 3.2). We also saw a clear increase in significant answers when we increased the sample size in the selection scenario (e.g., sample size of 3 1687/6240 compared to 20 4322/6240).

The Two-Step model approach appeared to have had the greatest difficulty consistently detecting the positive selection, but the median selection in the positive selection scenario was still greater than the median from the scrambled scenario (Fig. 3.4). Tracking frequency appeared to be a critical decision in the two-step model, with the 1 point per hour choice resulting in more detections of positive selection, and a tighter grouping of estimates. The decision to integrate step and turn angle into the model formula appeared similarly critical. The choice not to include the step and turn angle resulted in estimates more consistently positive in the positive selection scenario, and close to zero in the scrambled selection scenario. The integrated model formulation appeared to be largely responsible for the majority of the variation in estimates. Unlike the SSF results, the Two-Step approach appeared to provide significant answers (both positively and negatively) very readily (Table. 3.4). The balance of positively significant estimates did appear more positive skewed in the selection scenario (128429/327600: 39%) compared to the scrambled scenario (41069/327600: 13%), but all choices resulted in numbers of false positives and negatives bar the non-integrated model formulation choice.

The Poisson based approach, like those before, showed that the vast majority of analysis pathways result in the correct identification of positive preference for habitat 2 (Fig. 3.5), and centre on zero for the scrambled scenario. Unlike the previous analyses, the tracking duration did not have as clear an impact on the spread of estimates. Sample size and tracking frequency

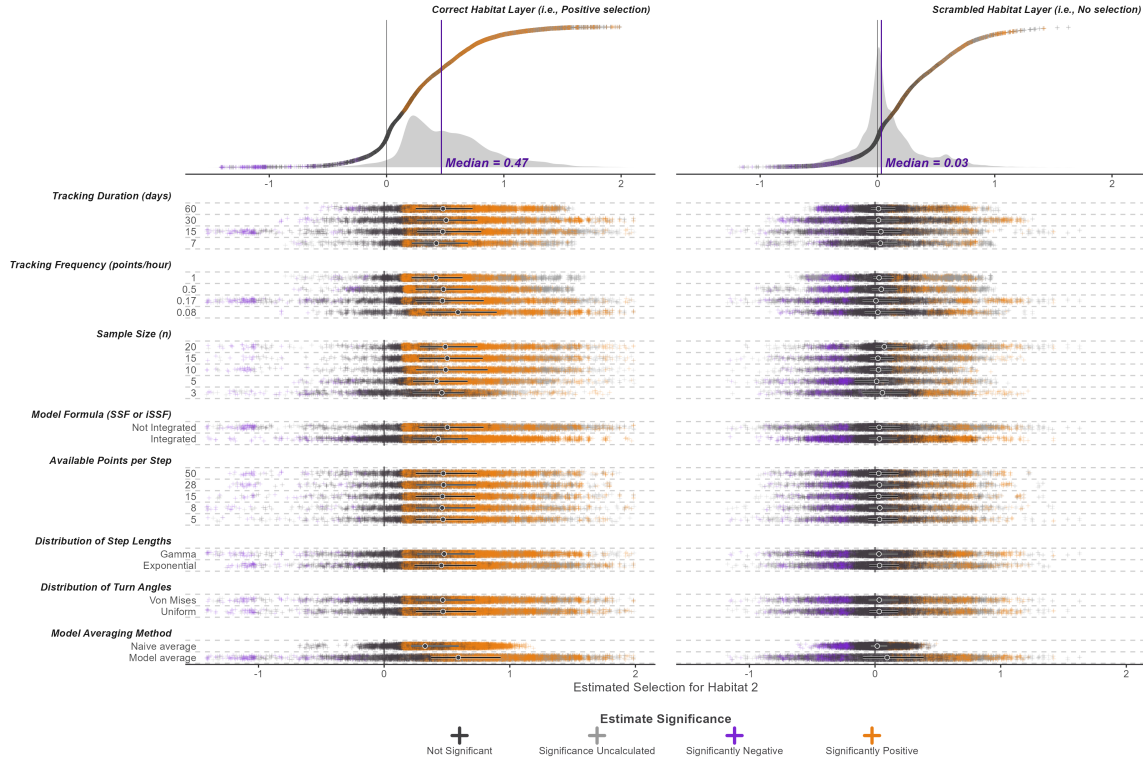


Figure 3.3: A specification curve showing habitat selection estimates resulting from the Step-selection Function (SSF) analysis pathways when applied to all simulated animal movement datasets. Every point represents a separate estimate of habitat selection. The top plot shows all estimates organised by estimate strength, with the purple vertical line depicting the median estimate and grey distribution describing the overall pattern of the estimates. Instances of outlying high or low estimates are labelled instead of plotted. The lower plot shows the same estimates, but split by the analysis or sampling decisions. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range.

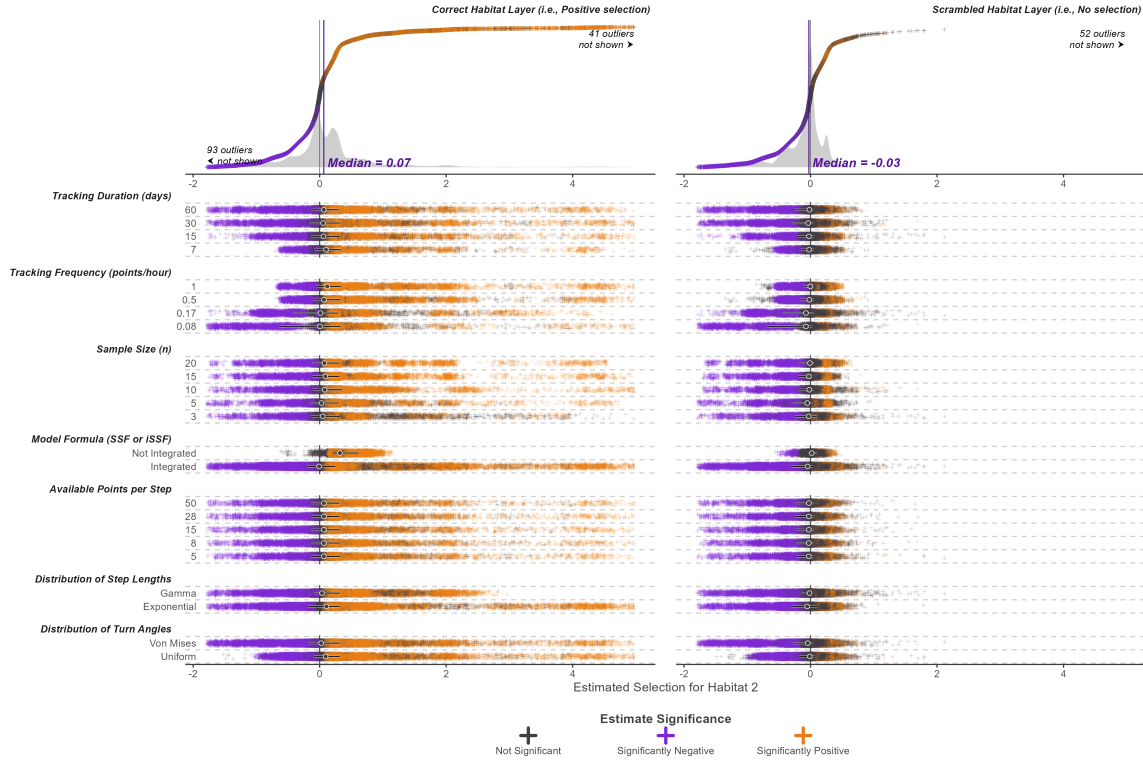


Figure 3.4: A specification curve showing habitat selection estimates resulting from the Two-step analysis pathways when applied to all simulated animal movement datasets. Every point represents a separate estimate of habitat selection. The top plot shows all estimates organised by estimate strength, with the purple vertical line depicting the median estimate and grey distribution describing the overall pattern of the estimates. Instances of outlying high or low estimates are labelled instead of plotted. The lower plot shows the same estimates, but split by the analysis or sampling decisions. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range.

did; larger sample sizes and higher frequency led to overall decreases in estimate spread. Of particular note is that the majority of failures to detect positive preference occurred when the sample size was three and at lower tracking frequencies. The analysis decisions appeared largely inconsequential, particularly the number of points per step. Model formulation did have an impact, with the integrated formulation being more consistent at retrieving positive preference in the positive selection scenario, but the spread of estimates was much larger. This spread increase associated with the integrated model formulation was similarly seen in the scrambled selection scenario, but resulted in near identical median estimates across model formulations. Unexpectedly, despite quite clear positive selection in many estimates in the selection scenario, all estimates (regardless of scenario) were deemed not to be significant, with 95% confidence intervals overlapping zero (Table. 3.3).

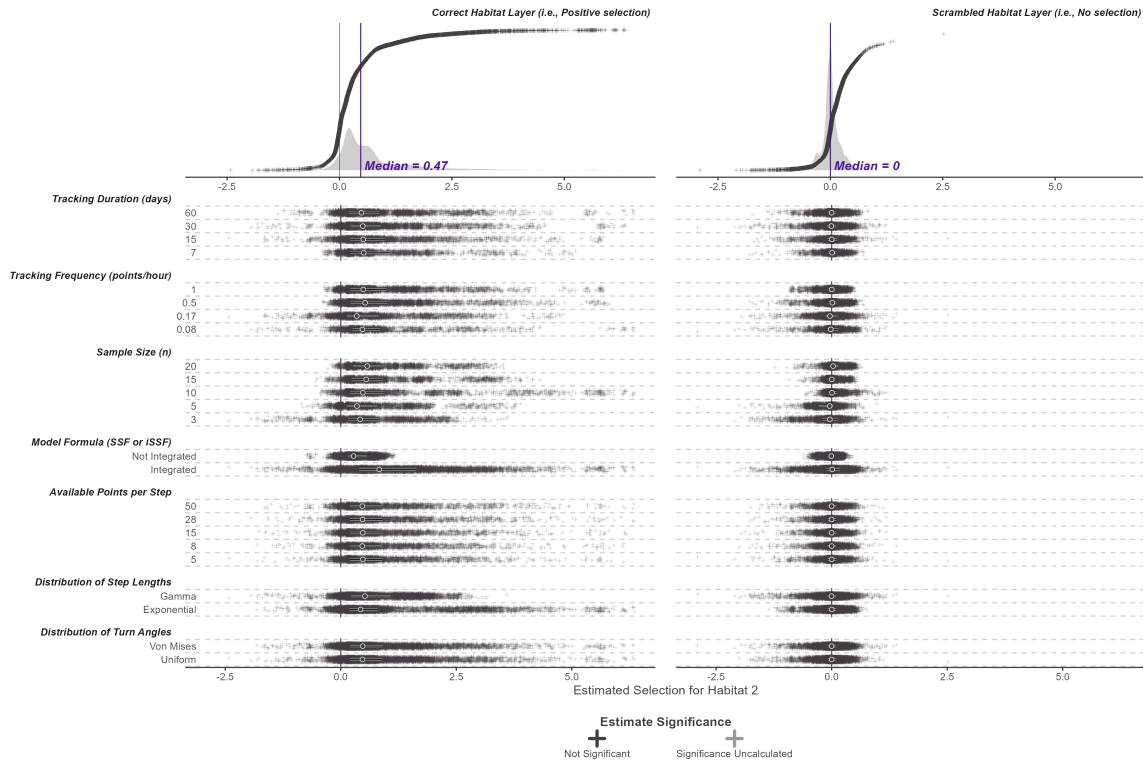


Figure 3.5: A specification curve showing habitat selection estimates resulting from the Poisson analysis pathways when applied to all simulated animal movement datasets. Every point represents a separate estimate of habitat selection. The top plot shows all estimates organised by estimate strength, with the purple vertical line depicting the median estimate and grey distribution describing the overall pattern of the estimates. Instances of outlying high or low estimates are labelled instead of plotted. The lower plot shows the same estimates, but split by the analysis or sampling decisions. Solid circles are the medians for each choice and the lines show the 75% inter-quartile range.

3.5.2 Model Results

We also used four Bayesian Regression Models to explore the impact of different decisions (Fig. 3.6), and whether those decisions could predict deviation from the median estimate (i.e., as a proxy for seeing which decisions lead to outlying extreme results). Greater variation would indicate greater scope for researcher choice to have a significant impact on the final results. The models provided a means to account for the variation that originates from stochasticity in the simulation and sampling captured by the group effects and reflected in the difference between marginal and conditional R^2 values, as the sheer number of final estimates can obscure patterns in specification curves.

The conditional R^2 (overall model R^2 ; fixed/population effects and random/group effects) values differed for the four models, ranging from 0.14 to 0.32, indicating a sizeable amount of variation in estimates being unaccounted for by the analysis decisions or random/group effects of sample ID, selection scenario, and species. The Compana results model had a conditional R^2 of 0.32; whereas the SSF model returned 0.29; the Two-Step model returned 0.14; and the Poisson model returned 0.3.

The marginal R^2 (fixed/population effects only) represents the bulk of the R^2 for Compana and SSF models, suggesting an important role for the fixed/population effects. The Compana results model had a marginal R^2 of 0.18; whereas the SSF model returned 0.18. The Two-Step model returned a marginal R^2 of 0.05; and the Poisson model returned 0.11. The low marginal R^2 in the Two-Step model highlights the greater sensitivity to variation between species, individuals and samples, compared to the variation resulting from analysis decisions. The same is true for the Poisson model, but to a lesser degree.

The Poisson model approach showed the most consistent benefits from more data (Fig. 3.6), where all increases in sampling intensity led to marked reductions in deviation from the median estimate: sample size ($\beta = -0.02$; 95% HD CI -0.06 - 0.02), tracking duration ($\beta = -0.02$; 95% HD CI -0.02 - -0.01), and tracking frequency ($\beta = -0.02$; 95% HD CI -0.03 - -0.02). Similarly SSF methods benefited from increasing sampling duration ($\beta = -0.03$; 95% HD CI -0.03 - -0.02) and frequency, but the impact on increasing sample size was not as apparent ($\beta = 0.01$; 95% HD CI -0.01 - 0.02). Two-Step methods appeared to benefit from increases in tracking frequency ($\beta = -0.08$; 95% HD CI -0.09 - -0.08), but not duration ($\beta = 0.00$; 95% HD CI 0.00 - 0.01) or sample size ($\beta = 0.00$; 95% HD CI -0.02 - 0.01). Potentially the benefits from some

of these decisions are masked by the extreme variation in the integrated model formulation decision. The area-based Compana approach appeared to vary more in its estimates when provided more data. Sample size ($\beta = 0.28$; 95% HD CI 0.10 - 0.47), tracking duration ($\beta = 0.06$; 95% HD CI 0.04 - 0.07) and tracking frequency ($\beta = 0.12$; 95% HD CI 0.10 - 0.13) all led to increased variation from the median estimate.

The points per step decision made little impact on the variation of estimates in any of the step-based approaches (Fig. 3.6), potentially due to the simplified nature of the landscape used (i.e., no rare habitat that would require high numbers of random points to detect). Use of the gamma distribution, in contrast to the exponential distribution, in all the step-based approaches led to reduced deviation from the median (Poisson: ($\beta = -0.07$; 95% HD CI -0.08 - -0.06); SSF: ($\beta = -0.01$; 95% HD CI -0.01 - -0.01); Two-Step: ($\beta = -0.09$; 95% HD CI -0.10 - -0.09)). Whereas the use of Von Mises, in contrast to a uniform distribution, failed to have a sizeable impact in Poisson ($\beta = 0.00$; 95% HD CI -0.01 - 0.01) and SSF ($\beta = 0.00$; 95% HD CI 0.00 - 0.00) methods and led to greater variation in Two-Step estimates ($\beta = 0.09$; 95% HD CI 0.08 - 0.10). The underlying simulation used a gamma distribution to generate step lengths and Von Mises for turn angles, which may partially explain this result.

The most mixed decision was the use of a model formulation that included the step and turn angles (i.e., integrated versus not integrated; Fig. 3.6). Here we saw that including the step and turn angles in the model formula for the summarised SSF approach led to greater variation from the median estimate ($\beta = 0.04$; 95% HD CI 0.04 - 0.05), but a contrasting effect for both the Two-step ($\beta = -0.17$; 95% HD CI -0.17 - -0.16) and Poisson approaches ($\beta = -0.37$; 95% HD CI -0.38 - -0.36) where the non-integrated formula tended to reduce the spread of estimates.

The greater deviation in averaged SSFs when using the integrated formulation may be a result of its interaction with the model averaging approach. The naive model averaging approach led to a dramatically lower spread of estimates ($\beta = -0.18$; 95% HD CI -0.18 - -0.17), and avoided underestimating selection in the positive selection scenario (unlike the AIC model average approach; (Fig. 3.3).

Finally, the area-based approach had a number of unique decisions. Largely the decisions associated with defining the available area had a larger impact than those linked to generating the random points (Fig. 3.6). In brief, the use of MCPs ($\beta = -0.16$; 95% HD CI -0.19 - -0.13) and type III designs ($\beta = -0.29$; 95% HD CI -0.32 - -0.26) tended to lead to less variable results.

Larger contours (areas) ($\beta = 0.05$; 95% HDI 0.03 - 0.06) and more points ($\beta = 0.02$; 95% HDI 0.01 - 0.04) unexpectedly increased variation of estimates from the median. By far, the clearest result in the area-based approach was the Compana test method, where randomisation dramatically reduced the spread of estimates ($\beta = -2.32$; 95% HDI -2.35 - -2.29).

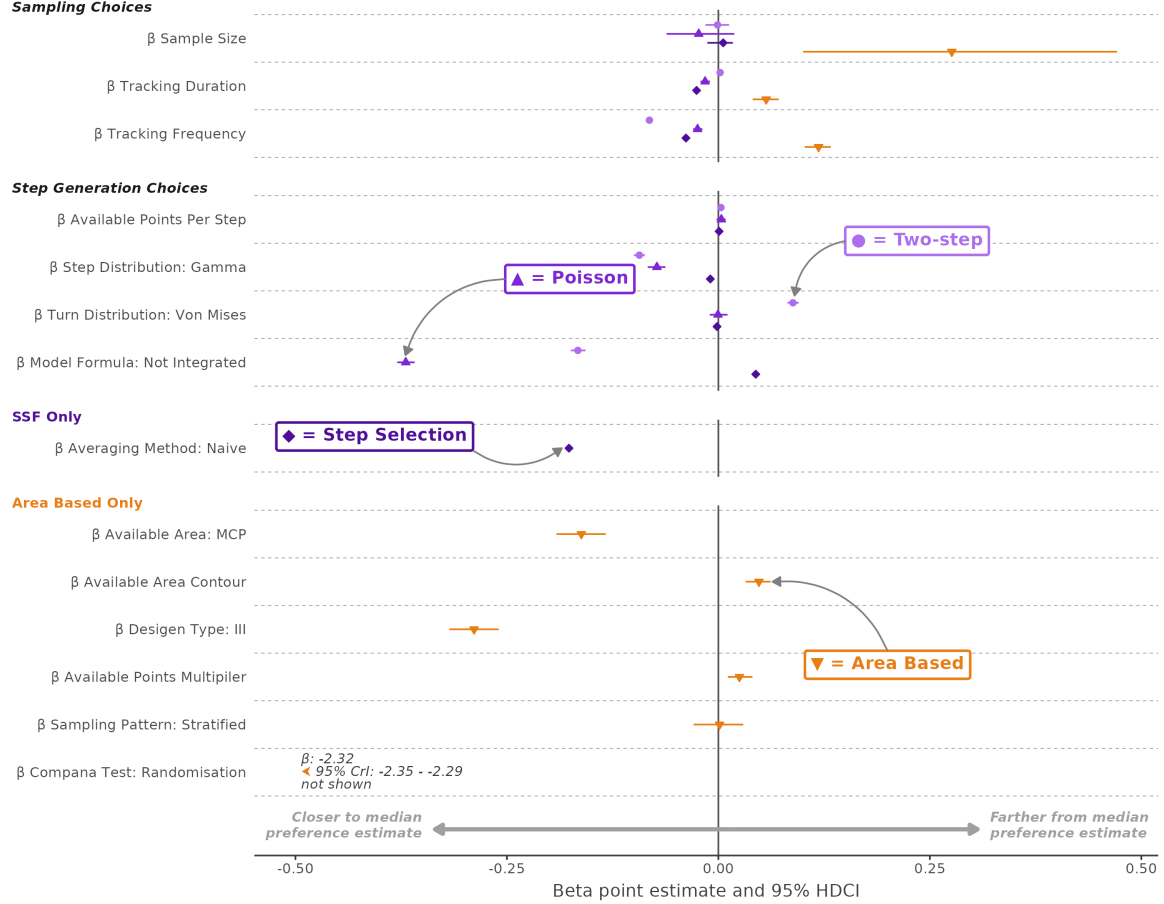


Figure 3.6: Point estimates and 95% credible intervals associated with all population effects from the Bayesian Regression models aiming to predict deviation from the median estimate using the information on analysis choices. Negative effects indicate that the decisions tend to decrease the deviation from the median estimate, whereas positive effects suggest those decisions are increasingly connected to estimates far from the median estimate. The effects are grouped into those pertaining to sampling choices and tracking, those associated with the step-based analysis approaches (Two-Step, SSF, and Poisson), those connected just to the SSF approach, and those connected to the area-based Compana approach.

3.6 Discussion

The multiverse approach highlights the diverse array of answers one can obtain even when analysing the same data, simulated using the same parameters. Fortunately the vast majority of the answers from the 287,040 analysis end points agree, and correctly recover the positive

habitat selection programmed into the simulated animals in the positive selection and correctly recover the absence of habitat selection in the no selection scenarios. This broad agreement is cause for optimism, suggesting that the potentially deviating analytical choices made by researchers largely converge to an agreement –albeit within this given simulated scenario and given habitat preference strength. Looking back at past studies reporting habitat preference, we can have considerable confidence that different (or reported) choices during analysis likely would not have changed the overall conclusions.

Unlike previous multiverse investigations into individual selection [Marshall & Duthie (2024); Chapter 2], the largest effect on whether an estimate was close to the median estimate did not come as a result of sampling approach (either tracking regime or number of individuals). This is possibly due to the overall increase in data quantity that came from running a multi-individual sample, perhaps past the point of benefiting as dramatically from more data. There was no pre-programmed individual variation (which is a key consideration in real-world scenarios Stuber, Carlson & Jesmer, 2022), therefore any variation between individuals of the same pseudo-species was the result of stochasticity (future expansions could easily integrate varying selection in the simulations). This lack of individuality may dampen the variation, reducing effects of increased sample sizes. However, a failure to reduce the deviation from the median estimate does not mean that increased sample sizes were not beneficial. Increased sample sizes may have actually led to estimates that were closer to the “true” selection of the animal, and Craiu et al. (2011) suggests greater stability in estimates with more data.

When examining the impact of individual analysis choices, we see a few that appear to be key. In the area-based (Compana) analysis, the decision to test selection using randomisation led to dramatically more consistent estimates. We suspect this is due to the parametric method making more stringent assumptions regarding the data’s distribution, resulting in a greater frequency of unstable outlying estimates. For the SSF and Poisson models, the decision on integrating step and turn angle into the model formulation appears critical. For SSF, it is suggested that integrating step and turn is ideal as it limits bias by accounting for the movement tendencies of the animal, particularly for high selection scenarios (Forester, Im & Rathouz, 2009). We saw reduced extreme high and low estimates when using the integrated formulation that supports this.

Poisson models are similarly affected, but in the opposite direction. Muff, Signer & Fieberg (2020) previously highlighted the potential contrasting impact of integration in the Poisson

models, and our multiverse mirrors this with the integrated formulations leading to estimates of greater strengths and that varied more. In the context of Muff, Signer & Fieberg (2020), we suspect the more outlying variable estimations are more biased results and ultimately further from the true selection strength. The most deviated estimations appear at low or intermediate sample sizes, and low tracking frequencies for the outlying low estimates. This could indicate that any bias resulting from the integrated formula could be mitigated with more data, but further investigation would be required to confirm this as we do not see greater tracking durations eliminating outlier estimates.

Outside of approach-specific decisions we can compare the different approaches. While we should be cautious comparing the different approaches, as they target and interpret selection in different ways, we can see that SSF, Poisson, and Compana approaches do better than the Two-Step approach when correctly identifying selection and no-selection. The Two-Step approach appears to be overly keen to consider selection significant (in either direction), and could frequently encounter convergence issues. The latter is likely caused by data inadequacies in either the strata or clusters (Craiu et al., 2011). Compana did appear to be more consistent than Two-Step; and the SSF and Poisson approaches appear to react more predictably to improvements in data quantity and sample size. The SSF and Poisson approaches appear to be the stronger choices and offer increased flexibility with regards to habitat classification (i.e., allow for continuous habitat variables).

3.6.1 Multiverse in context

While multiverses can explore the possible answers that can come from a given dataset, they cannot provide guidance on which of those answers is correct. The creation of the multiverse is subject to the same decision-making procedure that any single given pathway is: what decisions to include, and how to vary the choices (Simonsohn, Simmons & Nelson, 2020; Del Giudice & Gangestad, 2021). Therefore, the overall median or mean answer is heavily dependent on the construction of the multiverse. Additionally, the choices and resulting analysis pathways are not equally as valid, nor equally likely to be undertaken by a researcher. This means the multiverse of answers may provide some insight on spread of estimates (Steege et al., 2016), but should not be directly interpreted as providing more or less support for a given estimate based on the distribution of estimates. We can take this issue further, reminding ourselves that agreement is not necessarily a reliable proxy for accuracy (Devezer et

al., 2021). A better model reflecting the mechanistic reality of the study system is preferable to a suite of models that agree but converge on a less accurate answer. An example in the conducted multiverse: we used AKDEs and MCPs as two alternative definitions of availability for the area-based method. Arguably we should have excluded the MCP choice in favour of the AKDE choice, as the latter does a better job at capturing the movement processes/patterns underlying animal space-use (Fleming et al., 2015). Judgement calls such as this are unavoidable, and determining whether a model better reflects the underlying system can be exceptionally difficult outside of simulated “known-truth” scenarios. It very well may be ideal to conduct multiple analyses as standard procedure to explore selection at different scales (Aebischer, Robertson & Kenward, 1993).

When we are uncertain of the mechanism, or unable to truly model the vast complexity of the system (as is very frequent in ecology), then informed agreement may be our best option. Recent efforts looking at how many researchers answer the same question from the same dataset reveals that these judgement calls result in different answers (Gould et al., 2023). In many ways the more organic real-world results from Gould et al. (2023) reflect the multiverse results presented here –a general agreement on direction, but variation in uncertainty and magnitude. What differs, is that here we know that the simulated correct answer is a positive effect, while in Gould et al. (2023), the truth is unknown, so we rely on the good agreement of many independent informed analyses. The logistics of organising 246 researchers to assess uncertainty in analyses is infeasible on a broader scale. The multiverse approach presents a complementary method for assessing the scope of researcher choice on final results –albeit more vulnerable to the individual researcher’s biases and potentially limited to analyses with lower computational costs.

We have avoided suggesting rules or guides based on these results, instead reporting broad trends. The diverse array of species tracked and the habitats in which they reside makes specific guidance nigh impossible without informed assumptions tailored to a given study system. We simulated three species to try and broaden the generalisability of our findings. However, they all shared core assumptions stemming from the agent-based simulation, namely: site fidelity, consistent circadian cycles, and distinct behavioural states (Marshall & Duthie, 2022). For example, the routine use of shelter sites in preferred habitat would mean that missing those sites/locations is unlikely with a reasonably frequent tracking regime. This may have dampened the benefits of tracking frequency in our scenarios, and comparatively in scenarios where the use of preferred habitat is more fleeting or temporally inconsistent,

tracking frequency may be more crucial. Other studies have highlighted that certain decisions may have increased importance in different scenarios. The number of available points in SSF analysis becomes more important when the analysis is required to identify the use of smaller/rarer habitat types (Thurfjell, Ciuti & Boyce, 2014). Some methods, such as Resource Selection Functions, have the means of estimating the required data quantity to detect a given strength of selection in a given landscape, and provide a guide to how consistent results would be from that dataset (Street et al., 2021). These methods still require informed decisions to be made concerning the strength of selection that is of interest as well as how to define the landscape, but complement the goal of the multiverse assessment with a more rigorously mathematical approach.

Habitat selection analyses continue to expand and develop further, for example, to better tackle the non-independence of observed animal locations while building agreement between movement (i.e., local selection) and overall habitat selection (Michelot, Blackwell & Matthiopoulos, 2019; Michelot et al., 2020), integrate non-linear relationships (Klappstein et al., 2024), and develop methods less influenced by data missingness (Michelot et al., 2019; Hofmann, Cozzi & Fieberg, 2024). Repeating multiverse-like assessments may present a valuable context for these analyses, which is independent of the specific modelling approach.

3.7 Conclusions

Overall our multiverse exploration of habitat selection analyses shows that the majority of analysis pathways correctly identify the pre-programmed selection present in the simulated scenarios. While there was broad agreement, which was enhanced in most cases with increased data, several analysis choices highlight the continued need for attention when analysing animal movement data as the impact of those choices will not be the same under different analysis approaches.

3.7.1 Acknowledgements

This work was supported by the Natural Environment Research Council (NERC) via the IA-PETUS2 Doctoral Training Partnership held by Benjamin Michael Marshall (grant reference NE/S007431/1).

3.7.2 Software availability

In addition to packages already mentioned in the methods we also used the following.

We used *R* v.4.4.2 (R Core Team, 2022) via *RStudio* v.2024.09.0 (RStudio Team, 2022). We used *here* v.1.0.1 (Müller, 2020b) and *qs* v.0.27.2 (Ching, 2023) to manage directory addresses and saved objects.

We used *raster* v.3.6.30 (Hijmans, 2023a) and *RandomFields* v.3.3.14 (Schlather et al., 2015) to aid landscape raster creation alongside NLMR v.1.1.1 (Sciaini et al., 2018c).

We used *ggplot2* v.3.5.1 for creating figures (Wickham, 2016b), with the expansions: *patchwork* v.1.3.0 (Pedersen, 2024), *ggridges* v.0.5.6 (Wilke, 2024), *ggdist* v.3.3.2 (Kay, 2023a), and *ggtext* v.0.1.2 (Wilke & Wiernik, 2022).

We used *brms* v.2.22.0 (Bürkner, 2021) to run Bayesian models, with diagnostics generated used *bayesplot* v.1.11.1 (Gabry & Mahr, 2024), *tidybayes* v.3.0.7 (Kay, 2023b), and *performance* v.0.12.4 (Lüdecke et al., 2021).

We used the *dplyr* v.1.1.4 (Wickham et al., 2023), *tibble* v.3.2.1 (Müller & Wickham, 2023), and *stringr* v.1.5.1 (Wickham, 2022) packages for data manipulation.

We used *sp* v.2.1.4 (Pebesma & Bivand, 2005b), *move* v.4.2.6 (Kranstauber, Smolla & Scharf, 2023) for manipulation of spatial data and estimation of space use not otherwise mentioned in the methods.

We used *rmarkdown* v.2.29 (Xie, Allaire & Golemund, 2018; Xie, Dervieux & Riederer, 2020; Allaire et al., 2023), *bookdown* v.0.42 (Xie, 2016, 2022a), *tinytex* v.0.54 (Xie, 2019, 2023a), and *knitr* v.1.49 (Xie, 2014, 2015, 2023b) packages to generate type-set outputs.

We generated R package citations with the aid of *grateful* v.0.2.10 (Francisco Rodríguez-Sánchez, Connor P. Jackson & Shaurita D. Hutchins, 2023).

3.7.3 Data availability

All code used to simulate the movement data, implement the multiverse of analysis, and model the variation in resulting estimates is available at <https://github.com/BenMMarshall/multiverseHabitatPop>, and archived at <https://doi.org/10.5281/zenodo.14034792>. We have

included further model diagnostic plots not included here in the repository.

3.7.4 Author contributions

Conceptualization: BMM; Software: BMM; Supervision: ABD; Visualization: BMM; Writing - Original Draft: BMM; Writing - Review & Editing: BMM, ABD.

3.8 Supplementary Material

Table 3.1: An overall summary of all Compans analyses and associated decisions in relation to the frequency of significant results.

Analysis	Selection Scenario	Decision	Value	Significantly Negative	Not Significant	Significantly Positive	Total Estimates	Significance Uncalculated
Compans	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	7	0	4852	2828	7680	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	15	0	7654	3866	11520	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	30	0	10090	5270	15360	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	60	0	10229	5131	15360	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.08	0	4871	2809	7680	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.17	0	7198	4322	11520	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.5	0	10288	5072	15360	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	1	0	10468	4892	15360	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	3	0	3082	6902	9984	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	5	0	3716	6268	9984	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	10	0	7039	2945	9984	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	15	0	9258	726	9984	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	20	0	9730	254	9984	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Type II or Type III Habitat Selection	III	0	15353	9607	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Type II or Type III Habitat Selection	II	0	17472	7488	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Available Points Multiplier	1	0	8093	4387	12480	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Available Points Multiplier	2	0	8231	4249	12480	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Available Points Multiplier	4	0	8213	4267	12480	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Available Points Multiplier	10	0	8288	4192	12480	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Sampling Pattern	Random	0	16376	8584	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Sampling Pattern	Stratified	0	16449	8511	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Compans Test Method	Parametric	0	19528	5432	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Compans Test Method	Randomisation	0	13297	11663	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Available Area Method	MCP	0	15950	9010	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Available Area Method	AKDE	0	16875	8085	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Available Area Contour (%)	95	0	16144	8816	24960	0
Compans	Correct Habitat Layer (i.e., Positive selection)	Available Area Contour (%)	99	0	16681	8279	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	7	147	2868	4665	7680	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	15	206	4885	6429	11520	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	30	406	7020	7934	15360	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	60	338	6890	8132	15360	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.08	283	4394	3003	7680	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.17	250	6353	4917	11520	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.5	294	5730	9336	15360	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	1	270	5186	9904	15360	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	3	310	3978	5696	9984	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	5	677	4359	4948	9984	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	10	46	4036	5902	9984	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	15	64	4589	5331	9984	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	20	0	4701	5283	9984	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Type II or Type III Habitat Selection	III	528	10757	13675	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Type II or Type III Habitat Selection	II	569	10906	13485	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Available Points Multiplier	1	259	5383	6838	12480	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Available Points Multiplier	2	269	5388	6823	12480	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Available Points Multiplier	4	295	5435	6750	12480	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Available Points Multiplier	10	274	5457	6749	12480	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Sampling Pattern	Random	557	10791	13612	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Sampling Pattern	Stratified	540	10872	13548	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Compans Test Method	Parametric	1027	10758	13175	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Compans Test Method	Randomisation	70	10905	13985	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Available Area Method	MCP	543	10859	13558	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Available Area Method	AKDE	554	10804	13602	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Available Area Contour (%)	95	571	10704	13685	24960	0
Compans	Scrambled Habitat Layer (i.e., No selection)	Available Area Contour (%)	99	526	10959	13475	24960	0

Table 3.2: An overall summary of all SSF analyses and associated decisions in relation to the frequency of significant results.

Analysis	Selection Scenario	Decision	Value	Significantly Negative	Not Significant	Significantly Positive	Total Estimates	Significance Uncalculated
SSF	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	7	5	1663	1966	4800	1166
SSF	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	15	84	2104	3494	7200	1518
SSF	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	30	22	2600	5178	9600	1800
SSF	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	60	44	2014	5946	9600	1596
SSF	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.08	4	1397	2782	4800	617
SSF	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.17	102	2576	3481	7200	1041
SSF	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.5	34	2942	5141	9600	2083
SSF	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	1	15	2066	5180	9600	2339
SSF	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	3	62	3176	1687	6240	1315
SSF	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	5	28	2066	2845	6240	1301
SSF	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	10	28	3509		6240	1199
SSF	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	15	12	852	4221	6240	1155
SSF	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	20	25	783	4322	6240	1110
SSF	Correct Habitat Layer (i.e., Positive selection)	Model Formula (SSF or iSSF)	Integrated	95	5480	10025	15600	0
SSF	Correct Habitat Layer (i.e., Positive selection)	Model Formula (SSF or iSSF)	Not Integrated	60	2901	6559	15600	6080
SSF	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	5	17	1740	3160	6240	1323
SSF	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	8	28	1689	3246	6240	1277
SSF	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	15	30	1667	3322	6240	1221
SSF	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	28	41	1640	3407	6240	1152
SSF	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	50	39	1645	3449	6240	1107
SSF	Correct Habitat Layer (i.e., Positive selection)	Distribution of Step Lengths	Exponential	87	4437	8036	15600	3040
SSF	Correct Habitat Layer (i.e., Positive selection)	Distribution of Step Lengths	Gamma	68	3944	8548	15600	3040
SSF	Correct Habitat Layer (i.e., Positive selection)	Distribution of Turn Angles	Uniform	64	4146	8329	15600	3061
SSF	Correct Habitat Layer (i.e., Positive selection)	Distribution of Turn Angles	Von Mises	91	4235	8255	15600	3019
SSF	Correct Habitat Layer (i.e., Positive selection)	Model Averaging Method	Model average	155	3440	5925	15600	6080
SSF	Correct Habitat Layer (i.e., Positive selection)	Model Averaging Method	Naive average	0	4941	10659	15600	0
SSF	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	7	81	3178	375	4800	1166
SSF	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	15	191	4777	714	7200	1518
SSF	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	30	415	6440	945	9600	1800
SSF	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	60	410	6250	1344	9600	1596
SSF	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.08	174	3508	501	4800	617
SSF	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.17	211	5283	665	7200	1041
SSF	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.5	340	6116	1061	9600	2083
SSF	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	1	372	5738	1151	9600	2339
SSF	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	3	253	4050	622	6240	1315
SSF	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	5	529	4140	270	6240	1301
SSF	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	10	91	3661	389	6240	1199
SSF	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	15	129	4237	719	6240	1155
SSF	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	20	95	3657	1378	6240	1110
SSF	Scrambled Habitat Layer (i.e., No selection)	Model Formula (SSF or iSSF)	Integrated	237	12437	2326	15600	0
SSF	Scrambled Habitat Layer (i.e., No selection)	Model Formula (SSF or iSSF)	Not Integrated	260	8208	1052	15600	6080
SSF	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	5	210	4120	587	6240	1323
SSF	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	8	211	4111	641	6240	1277
SSF	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	15	218	4121	680	6240	1221
SSF	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	28	229	4125	734	6240	1152
SSF	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	50	229	4168	736	6240	1107
SSF	Scrambled Habitat Layer (i.e., No selection)	Distribution of Step Lengths	Exponential	495	10407	1658	15600	3040
SSF	Scrambled Habitat Layer (i.e., No selection)	Distribution of Step Lengths	Gamma	602	10238	1720	15600	3040
SSF	Scrambled Habitat Layer (i.e., No selection)	Distribution of Turn Angles	Uniform	485	10405	1649	15600	3061
SSF	Scrambled Habitat Layer (i.e., No selection)	Distribution of Turn Angles	Von Mises	612	10240	1729	15600	3019
SSF	Scrambled Habitat Layer (i.e., No selection)	Model Averaging Method	Model average	680	7019	1821	15600	6080
SSF	Scrambled Habitat Layer (i.e., No selection)	Model Averaging Method	Naive average	417	13626	1557	15600	0

Table 3.3: An overall summary of all Poisson analyses and associated decisions in relation to the frequency of significant results.

Analysis	Selection Scenario	Decision	Value	Significantly Negative	Not Significant	Significantly Positive	Total Estimates	Significance Uncalculated
Poisson	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	7	0	2400	0	2400	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	15	0	3600	0	3600	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	30	0	4800	0	4800	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	60	0	4799	0	4800	1
Poisson	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.08	0	2400	0	2400	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.17	0	2400	0	2400	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.5	0	4800	0	4800	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	1	0	4800	0	4800	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	3	0	3120	0	3120	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	5	0	3120	0	3120	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	10	0	3120	0	3120	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	15	0	3119	0	3120	1
Poisson	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	20	0	3120	0	3120	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Model Formula (SSF or iSSF)	Integrated	0	7799	0	7800	1
Poisson	Correct Habitat Layer (i.e., Positive selection)	Model Formula (SSF or iSSF)	Not Integrated	0	7800	0	7800	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	5	0	3119	0	3120	1
Poisson	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	8	0	3120	0	3120	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	15	0	3120	0	3120	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	28	0	3120	0	3120	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	50	0	3120	0	3120	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Distribution of Step Lengths	Exponential	0	7800	0	7800	0
Poisson	Correct Habitat Layer (i.e., Positive selection)	Distribution of Step Lengths	Gamma	0	7799	0	7800	1
Poisson	Correct Habitat Layer (i.e., Positive selection)	Distribution of Turn Angles	Uniform	0	7799	0	7800	1
Poisson	Correct Habitat Layer (i.e., Positive selection)	Distribution of Turn Angles	Von Mises	0	7800	0	7800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	7	0	2400	0	2400	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	15	0	3600	0	3600	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	30	0	4800	0	4800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	60	0	4800	0	4800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.08	0	2400	0	2400	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.17	0	3600	0	3600	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.5	0	4800	0	4800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	1	0	4800	0	4800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	3	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	5	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	10	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	15	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	20	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Model Formula (SSF or iSSF)	Integrated	0	7800	0	7800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Model Formula (SSF or iSSF)	Not Integrated	0	7800	0	7800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	5	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	8	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	15	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	28	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	50	0	3120	0	3120	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Distribution of Step Lengths	Exponential	0	7800	0	7800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Distribution of Step Lengths	Gamma	0	7800	0	7800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Distribution of Turn Angles	Uniform	0	7800	0	7800	0
Poisson	Scrambled Habitat Layer (i.e., No selection)	Distribution of Turn Angles	Von Mises	0	7800	0	7800	0

Table 3.4: An overall summary of all Two-Step analyses and associated decisions in relation to the frequency of significant results.

Analysis	Selection Scenario	Decision	Value	Significantly Negative	Not Significant	Significantly Positive	Total Estimates	Significance Uncalculated
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	7	1699	2716	2785	7200	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	15	3115	3636	4049	10800	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	30	4687	4252	5461	14400	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Tracking Duration (days)	60	4859	3489	6052	14400	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.08	2692	2421	2087	7200	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.17	3853	3748	3199	10800	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	0.5	4156	4301	5943	14400	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Tracking Frequency (points/hour)	1	3659	3623	7118	14400	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	3	2274	4967	2519	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	5	2869	3603	2888	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	10	2953	2362	4045	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	15	3069	1883	4408	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Sample Size (n)	20	3195	1678	4487	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Model Formula (SSF or ISSF)	Integrated	14360	12421	12219	39000	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Model Formula (SSF or ISSF)	Not Integrated	0	1672	6128	7800	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	5	2802	2927	3631	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	8	2851	2866	3643	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	15	2876	2809	3675	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	28	2910	2753	3697	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Available Points per Step	50	2921	2738	3701	9360	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Distribution of Step Lengths	Exponential	7409	6017	9974	23400	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Distribution of Step Lengths	Gamma	6951	8076	8373	23400	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Distribution of Turn Angles	Uniform	6080	7618	9702	23400	0
TwoStep	Correct Habitat Layer (i.e., Positive selection)	Distribution of Turn Angles	Von Mises	8280	6475	8645	23400	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	7	1928	4663	609	7200	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	15	3565	6016	1219	10800	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	30	5230	7397	1773	14400	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Tracking Duration (days)	60	5297	6837	2266	14400	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.08	2921	3560	719	7200	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.17	4221	5235	1344	10800	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	0.5	4679	7997	1724	14400	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Tracking Frequency (points/hour)	1	4199	8121	2080	14400	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	3	2535	5642	1183	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	5	3392	5188	780	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	10	3263	5068	1029	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	15	3310	4903	1147	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Sample Size (n)	20	3520	4112	1728	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Model Formula (SSF or ISSF)	Integrated	15737	18348	4915	39000	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Model Formula (SSF or ISSF)	Not Integrated	283	6565	952	7800	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	5	3118	5083	1159	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	8	3162	5047	1151	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	15	3211	4958	1191	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	28	3257	4915	1188	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Available Points per Step	50	3272	4910	1178	9360	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Distribution of Step Lengths	Exponential	9593	9498	4309	23400	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Distribution of Step Lengths	Gamma	6427	15415	1558	23400	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Distribution of Turn Angles	Uniform	6825	13267	3308	23400	0
TwoStep	Scrambled Habitat Layer (i.e., No selection)	Distribution of Turn Angles	Von Mises	9195	11646	2559	23400	0

Chapter 4

Snake Habitat Selection Findings are not Undermined by Researcher Analysis Choice

Benjamin M. Marshall¹, Matthew Crane², Cameron W. Hodges³, Max D. Jones⁴, Tyler K. Knierim³, Samantha N. Smith⁵, Inês Silva⁶, Colin T. Strine⁷, Alexander B. Duthie¹

¹Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, FK9 4LA, UK; ²Unaffiliated; ³Western Ecosystems Technology, Inc, 415 W. 17th Street, Suite 200, Cheyenne, WY 82001, USA; ⁴Department of Fish and Wildlife Conservation, Virginia Tech, 310 W. Campus Drive, Blacksburg, VA, 24061, USA; ⁵Department of Forestry and Environmental Conservation, 261 Lehotsky Hall, Clemson University, Clemson, SC 29634, USA; ⁶Center for Advanced Systems Understanding (CASUS), Helmholtz-Zentrum Dresden-Rossendorf (HZDR), Görlitz, Germany; ⁷Department of Natural Sciences, Dickinson State University, Dickinson, North Dakota, 58601, United States

4.1 Abstract

Possibly the best way to determine a snake's needs is to follow their movements. Once we have learnt of the snake's movements we can infer habitat requirements, behaviour, and potential threats. Combined, the movement data and inferences can inform decisions on snake conservation and human-snake conflict. Extracting useful information from snake movement requires

many steps, from sampling to analysis. Other studies have shown that a single dataset can result in many different answers in the hands of different researchers, so how can we be confident the results from snake movement studies are providing the correct guidance for snake conservation? We used a multiverse approach to explore several thousand ways of extracting habitat selection estimates from four snake movement datasets, using categorical and continuous classifications of the landscape. The resulting habitat selection estimates showed an overall agreement with regards to direction, indicating general support for the original conclusions on the snakes' selected habitats. However, the strength of the selection appeared more sensitive to choices made during analysis, with several analysis choices appearing only to be impactful with a particular species' data. This leads us to suggest that generalisations on the impact of specific analysis decisions may be difficult to predict a priori for other species and study systems. The broad agreement shown here should bolster confidence in past studies' conclusions, whereas the variation displayed in the selection strengths highlights how analysis choices could inflate the differences between studies.

4.2 Introduction

A key component of science is the continual reassessment of past work and findings (Alberts et al., 2015), whether that reassessment takes the form of direct replications aiming to evaluate the reliability of previous work, or a more integrative approach testing the boundaries of previous findings' generalisability by retesting questions in different study systems [e.g., quasi-replications; Nakagawa & Parker (2015); Peterson & Panofsky (2021)].

Reassessments and replications can aid the formal and organic self-correcting process of science (Peterson & Panofsky, 2021). Initial findings set the stage for subsequent work, building momentum that can accelerate progress, but this momentum can be difficult to redirect after the initial impetus (Jennions & Møller, 2002a; Barto & Rillig, 2012). In ecology and evolution, research on mating systems shows exactly this momentum. Early studies on *Anolis spp.* mating systems concluded that territoriality played a key role in mating success by controlling access to breeding females (Kamath & Losos, 2017, 2018b). The assumption of a polygynous mating system coloured and shaped subsequent research, and was only adequately challenged when genetic investigation revealed the importance of female mate choice and multiple partners [Kamath & Losos (2018a)]. The new research highlighted that the initial assumptions of territoriality were overly strong and had misdirected research and obscured an important

part of *Anolis* ecology (Kamath & Losos, 2017, 2018b). Similar perspective changes are seen in studies of bird mating systems, where earlier research focused on mating competition and potentially lost sight of explanations connected to mating cooperation (Griffith, 2019). In this case, over emphasis on mate choice via differential ornamentation could obscure the roles of social compatibility (i.e., pair behavioural compatibility Ihle, Kempenaers & Forstmeier, 2015; Wang et al., 2017; and the imperfect information communicated by ornamentation Wang, Forstmeier & Kempenaers, 2017). These examples emphasise the importance of checking and confirming results early, before assumptions are entrenched and can steer the focus of future studies; we can see this principle recognised in the peer review system aiming to catch errors prior to broad dissemination.

Checking previous findings through replication is more difficult in systems with high task uncertainty. High task uncertainty systems lead to high levels of uncontrollable stochasticity, make direct diagnostic replications impractical or impossible, while rendering the evidence from quasi-replications weaker (Peterson & Panofsky, 2021). Ecological systems often generate high task uncertainty, with many interconnected elements, and when studying wild systems many of those elements can be uncontrollable.

With ecological systems, such complexity and the difficulties in controlling experiments makes direct replications costly, potentially explaining their rarity (Kelly, 2019). When studying wild animals with a level of direct intervention, repeating experiments might be unethical due to the well-being costs (Weatherhead & Blouin-Demers, 2004; Robstad et al., 2021; Tomotani et al., 2021; Portugal & White, 2022; Altobelli et al., 2022). When faced with limited options for direct replications, an alternative approach would be to re-examine existing datasets. Pooling old and new datasets and re-analysing them may provide opportunities for broader generalisations. Good examples exist in movement ecology, where the re-examination of many animal movement datasets revealed general patterns of reduced mammal movement with increased human footprint (Tucker et al., 2018), and increased bird movements in increasingly homogeneous landscapes (Tucker et al., 2019). These findings, based on many species, across various regions, are better placed to be the building blocks of future hypotheses.

In some cases older data may have been collected and recorded in ways that enables completely fresh analysis [e.g., Movebank; Kays et al. (2022)]. As methodologies develop, conceptualisations of how systems operate change, and computational power increases, new avenues for examining the same data may materialise (e.g., Noonan et al., 2020). As these new methods

are developed and applied, we may see the conclusions based upon those data change. There are a growing number of examples demonstrating that analysis approach can alter results (Salis, Lena & Lengagne, 2021; Desbureaux, 2021), and that researchers themselves can be a key source of variation in analysis and results (Silberzahn et al., 2018; Huntington-Klein et al., 2021; Gould et al., 2023). Recently for ecology and evolution, Gould et al. (2023) showed that when ~250 researchers were asked to answer the same question with the same dataset (either bird nestling success, or eucalyptus growth) the analysis approaches and results varied, and in some cases indicated opposing effects. These examples elegantly show the possible extent of technical uncertainty present in some systems (i.e., the uncertainty resulting from differences in methodology). In practice researchers often do not have the time to explore many possible analysis pathways; therefore, we may be underestimating the uncertainty surrounding published results.

Not all disciplines have explored the sources of uncertainty in results equally. Prudence would push for examination of uncertainty in all its forms, in particular for fields that already tackle high levels of uncertainty originating from wild study systems. Movement ecology is such a field. Animals are complex, existing in complex ecosystems with individuality and personality (Stuber, Carlson & Jesmer, 2022). Depending on the research question, controls in movement ecology can be difficult to achieve, and replications difficult to justify given the strict ethical considerations surrounding interventionist study (Hawkins, 2004; Bodey et al., 2018). Movement ecology has also seized the opportunities presented by technological developments (Wilmers et al., 2015), enabling higher resolution tracking of animal movements (e.g., GPS tracking) and more sophisticated analysis that can integrate the high dimensional data (Joo et al., 2020; e.g., x-y coordinates, time, acceleration, individual, other covariates of interest, Joo et al., 2022).

Personality and the repeatability of behaviours presents a key component of the variation when attempting to generalise in movement ecology (Stuber, Carlson & Jesmer, 2022). However, here we focus on technical uncertainty, the uncertainty originating from the researcher and how they approach the data. “Many analyst” projects highlight the potential for researcher-created variation (Silberzahn et al., 2018; Huntington-Klein et al., 2021; Gould et al., 2023), and multiverse explorations of movement ecology methods highlight the variation potentially presented within a synthetic movement dataset [Marshall & Duthie (2024); also see Chapter 3]. Here we take the multiverse approach further by applying it to a number of real case studies to explore whether different analysis approaches alter general conclusions.

We selected four separate but connected movement ecology studies that attempt to determine the habitat selection exhibited by snakes in north-eastern Thailand. All four cases focus on snakes that come into conflict with humans to varying degrees, either because of the risks posed by their venom: King Cobra (Marshall et al., 2019, 2020a), Malayan Krait (Hodges et al., 2022), Banded Krait (Knierim et al., 2019), or because of their appetite for domestic livestock: Burmese python (Smith et al., 2021). In all cases, the habitat selection results could be used to guide snake conservation efforts, as well as interventions into human behaviour to mitigate human-snake conflict by highlighting what the snakes are drawn to. With these general goals in mind, we aim to re-examine the movement datasets using a multiverse of habitat selection analysis pathways to reveal whether the same data could lead to different conclusions.

4.3 Methods

4.3.1 Summary of Field Methods

All four case studies took place in north-eastern Thailand, within Nakhon Ratchasima province. Three case studies (King Cobra, Burmese Python, Banded Krait) were conducted within the Sakaerat Biosphere Reserve. The reserve consists of three zones of management: core, buffer, and transitional. The core is largely primary forest; the buffer surrounds the core and is comprised of forest regeneration efforts, whereas the transitional zone allows more development resulting in a mix of agriculture, settlements, and plantation forest. Bisecting the transitional zone, and running adjacent to the protected forest areas, is a four-lane highway connecting the city of Nakhon Ratchasima to Bangkok. The Malayan Krait case study was undertaken nearer to Nakhon Ratchasima proper, on the Suranaree University of Technology campus. The university campus is a mix of scrub forest, open lawn, university buildings, and homes. Further details on the study sites' characteristics can be found in the original publications (Marshall et al., 2019, 2020a; Knierim et al., 2019; Smith et al., 2021; Hodges et al., 2022). For land use classification, we relied on data supplied by the Land Development Department of Thailand (Land Development Department, 2017).

Snakes can be difficult to detect in wild scenarios (Durso & Seigel, 2015; Boback et al., 2020), requiring a wider and more opportunistic suite of methods to gather adequate sample sizes. In all the chosen case studies snakes were obtained for study using trapping arrays,

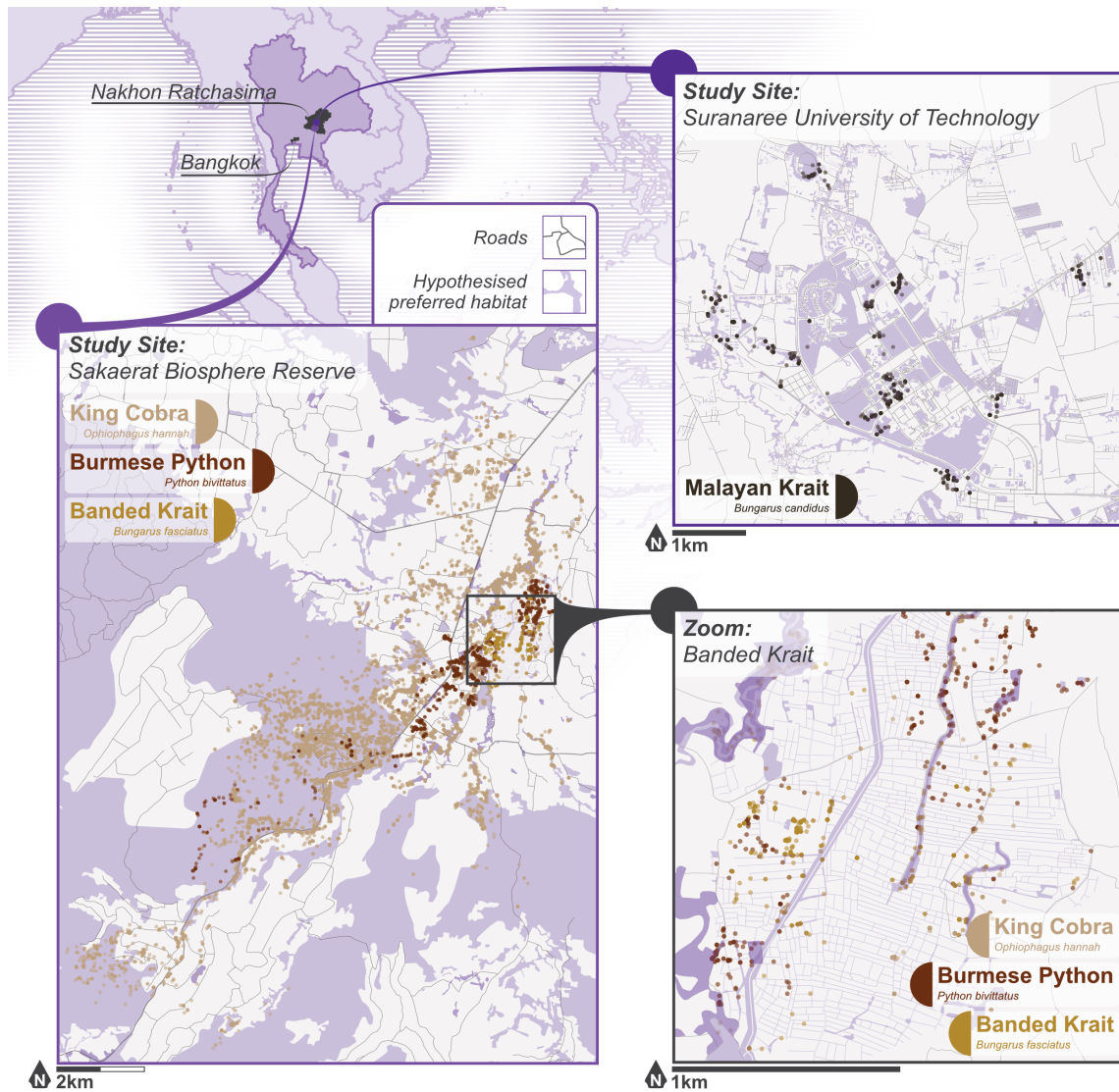


Figure 4.1: The study site locations for all four case studies. Top left shows the location within Thailand, both the Sakaerat Biosphere Reserve and the Suranaree University of Technology campus are located within Nakhon Ratchasima Province. Bottom left shows the overall study location for King Cobra, Burmese Python, and Banded Krait (circular points). The purple areas depict the land uses that are included in any of the species hypotheses. Gray lines indicate roads. Bottom right is the same region, but focused on the Banded Krait locations to better show the field margin data used for their assessment of habitat selection. Top right is the study location for the Malayan Krait on the Suranaree University of Technology campus.

active surveying, and notifications from locals. The local notifications often arose from snakes entering human settlements, and a desire for the snake to be removed.

In all cases, we implanted snakes with Holohil radio transmitters (either AI-2T, SI-2T, SB-2, or BD-2, Holohil Inc., Ontario, Canada)). The implant procedure was conducted by trained team members, or later by independent trained veterinarians, following Reinert & Cundall (1982). We anaesthetised snakes with isoflurane and implanted transmitters in the coelomic cavity. The size of the transmitter varied between species and was calibrated to body size. At no time did we implant transmitters exceeding 5% of the body mass of the animal; in the vast majority of cases, the transmitter weighed considerably less than 1% (e.g., King Cobra $0.55 \pm \text{SD } 0.47\%$; Banded Krait $0.01 \pm \text{SD } 0.09\%$; Burmese Python 0.26%). Inspections of longer-term tracked individuals indicated that transmitter presence was likely not detrimental to their health (Marshall et al., 2020a). Greater details on the implantation procedure, and supplemental snake processing, can be found in the original publications, alongside details concerning ethical approval and permitting.

Once implanted, we released snakes back into the wild as near as feasibly possible to their capture location; human-snake conflict was the main reason that exact return to capture location was not achieved. Following release, we routinely relocated the snakes using radio receivers, triangulating their locations then recording coordinates with handheld GPS devices. The four species had varying tracking regimes, which did not necessarily remain consistent across the entire study period. However, broadly speaking the tracking regime for King Cobras was three to four locations a day, whereas the other species we located once per day. The main exception is for the first three tracked King Cobra, who we tracked more frequently before changing tracking methods to minimise disturbance.

4.3.2 Study Species and Hypotheses

We simplified the conclusions of the original studies into hypotheses targeting the preferred habitat for each species. We registered these hypotheses prior to undertaking the multiverse exploration (<https://osf.io/kv8rg/>), they remained unchanged and are detailed below.

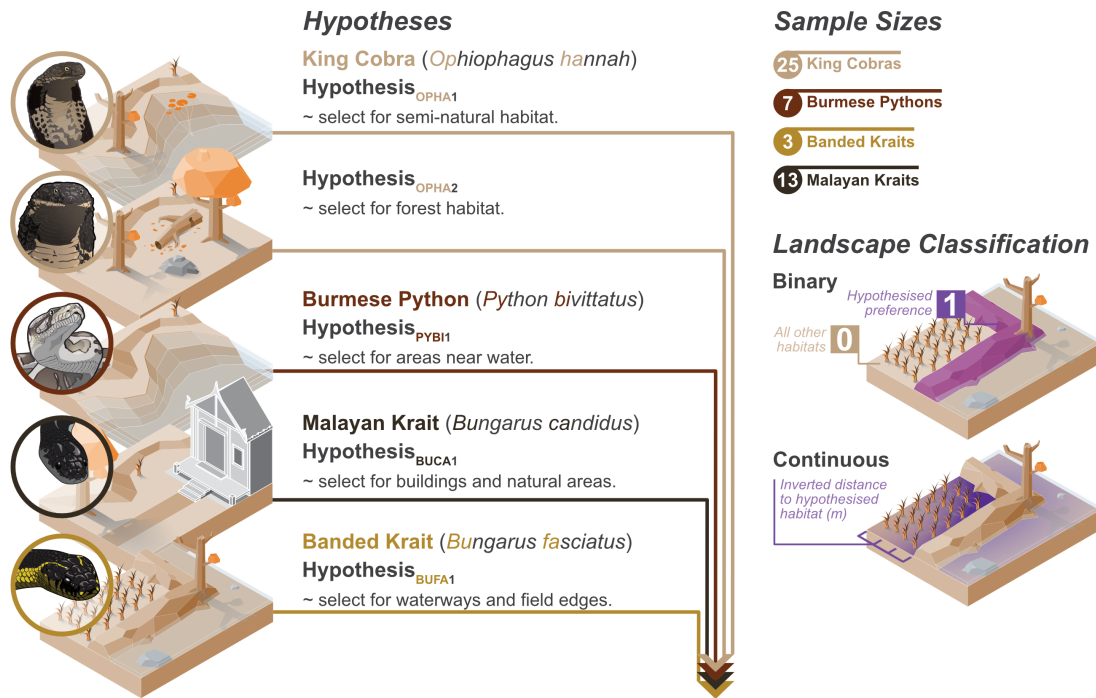


Figure 4.2: A summary of the species and hypotheses derived from the original studies regarding the habitat selection of each snake species. Alongside are the sample sizes for each species and a visualisation describing the two landscape classifications used in the multiverse.

King Cobra

King Cobras (*Ophiophagus hannah*) are a large (tracked individuals 1.40 to 3.71m snout to vent length), diurnal, active foraging snake species that depredate snakes and monitor lizards (Marshall et al., 2019, 2020a; Jones et al., 2020). While considered a predominately forest-dwelling species (Stuart et al., 2012), they are known to make use of more human-altered areas (Whitaker & Captain, 2004; Rao et al., 2013; Jones et al., 2022), which can lead to frequent human-snake conflict (Shankar et al., 2013b; Marshall et al., 2018). Compared to the other species, the habitat selection of King Cobras is relatively more well studied. However, as of 2024, King Cobras have been split into four species (Shankar et al., 2021; Das et al., 2024), meaning that the information gathered on King Cobras in other areas may be less generalisable to the King Cobras in Thailand. Additionally, the split means that the IUCN Redlist classification of Vulnerable may be optimistic due to the large original range of King Cobras. The extremely low occurrence of King Cobra bites in Thailand means that instances of human-snake conflict are primarily a conservation concern as opposed to a human health concern (Viravan et al., 1992; Pochanugool et al., 1998). Therefore, the split heightens the

importance of being confident in these case studies conducted within the range of the now Northern King Cobra.

Marshall et al. (2019) do not conclude on an actual habitat selection, instead highlighting the King Cobras' excursions out of the protected forest. Marshall et al. (2020a) look more specifically at selection, highlighting the importance of semi-natural areas that occupy the banks of irrigation canals and intersect the agricultural matrix surrounding the protected forest. Therefore, we pooled both datasets and examine two non-mutually exclusive hypotheses:

King Cobra Hypothesis 1: King Cobras select for semi-natural habitat.

King Cobra Hypothesis 2~: King Cobras select for forest habitat.

Burmese Python

Burmese Pythons are large (tracked individuals between 2.21 and 3.09m snout to vent length), ambush predators capable of tackling prey over 100% their own body mass (Bartoszek et al., 2018), and whose predation pressure can play a role in limiting mammal populations sizes (Dorcas et al., 2012). Flexibility in prey size means snakes of this size are inevitably drawn into conflict with humans over livestock, a pattern mirrored across the globe for large snakes (Miranda, Ribeiro- & Strüssmann, 2016).

Smith et al. (2021) describe Burmese Python (*Python bivittatus*) habitat selection and movement. The conclusions of Smith et al. (2021) on python habitat selection are similar to those made on King Cobras, with an active selection for areas near water. The land classification used in Smith et al. (2021) was slightly different to Marshall et al. (2020a), grouping semi-natural areas with larger water bodies (e.g., agricultural ponds).

Burmese Python Hypothesis 1~: Burmese Pythons select for areas near water.

Malayan Krait

The Malayan Kraits (*Bungarus candidus*) are a smaller species, with tracked individuals between 0.65 and 1.46m snout to vent, and all lived on a university campus. Malayan Kraits, like many elapids, have a potent and medically significant venom; bites of Malayan Kraits can be fatal (Looareesuwan, Viravan & Warrell, 1988; SEARO, 2016). They are (mostly) nocturnal and actively foraging (Hodges et al., 2021a), known to depredate a wide range of

prey (Kuch, 2001; Hodges, D’souza & Jintapirom, 2020).

Unlike the other case studies, Hodges et al. (2022) is undertaken in a more urban environment. The scale of the Malayan Krait movements meant the study was conducted at a finer spatial scale; habitat types are therefore more finely separated (e.g., buildings vs settlements). The overall conclusions highlight two habitat types comprising of buildings and natural vegetation are potentially being selected for, in contrast to the lack of selection for open areas.

Malayan Krait Hypothesis 1~: Malayan Kraits select for buildings and natural areas.

Banded Krait

Like its smaller cousin, the Banded Krait (*Bungarus fasciatus*) is also a nocturnal active forager, with a potent venom. The Banded Krait is heavier-bodied and grows to longer lengths, tracked individuals ranging from 1.13 and 1.58 m snout to vent length. However, unlike the Malayan Krait, the Banded Krait appears less tolerant of human disturbance in this region of Thailand and tends to have a more ophiophagus diet (Knierim, Barnes & Hodges, 2017).

Knierim et al. (2019) looked at Banded Kraits located in agricultural land, and like the other krait, Banded Kraits had movements more conducive to finer habitat classifications. For example, field margins were found as a key nesting site (Knierim et al., 2019). Knierim et al. (2019) shows the importance of field margins is reflected in movement and habitat selection, as Banded Kraits follow the linear water or field margin features as opposed to the wider more exposed field areas.

Banded Krait Hypothesis 1~: Banded Kraits select for waterways and field edges.

4.3.3 Tracking Data

We downloaded tracking data from Movebank or linked data repositories associated with the original studies (see data availability statement). The land use data were largely included with the associated data repositories; in cases where they were not we retrieved this from the original authors. We collated all tracking and land use data, unifying the data formats, naming conventions, and converting files to more accessible formats. We were left with two files for every study: a csv containing the movements of all individuals, and geoJSON file con-

taining the polygon data describing land use. The movement dataset contains seven columns: “species” = the scientific species name, “id” = the individual animal ID, “sex” = whether the individual is male (M) or female (F), “x” = the UTM Easting coordinate (m), “y” = the UTM Northing (m), “UTMzone” = the UTM zone either 47N or 48N, “datetime” = the date and time formatted as ISO 8601 “YYYY-MM-DD hh:mm:ss” recorded in “Asia/Bangkok”. The land use data contain geometry data in UTM projection and data on the “landuse” category as well as a “habitat” column denoting whether the habitat is part of a hypothesis.

To check for any potential issues that could hinder habitat selection analysis, we used code from Crane et al. (2021) to generate summaries of all the tracking datasets (archived code see reference for Crane et al., 2020). During this process, we identified duplicated data points in the King Cobra data that required removal. We suspect that these escaped detection during the initial publication (Marshall et al., 2019) because the methods used did not use time in estimation of home range.

4.3.4 Multiverse Construction

To explore the variation that could arise when testing the above hypotheses, we constructed a multiverse of analytical choices. A multiverse is a tree of branching paths, where decisions made during analysis spawn multiple branches, each with subsequent analysis decisions. Previous work has revealed the importance of data quantity over analysis decisions in determining habitat selection [Marshall & Duthie (2024); Chapter 3], but we are still left with sizeable variation in final answers –and perhaps more importantly, answers that offer contradictory conclusions. The variation in results are similarly present in more practical demonstrations, such as “many analysts” projects (Gould et al., 2023).

Our multiverse starts with two alternative ways of defining each species hypothesised preferred habitats. Each case study has a map of available habitats, defined as categories. One approach we used is to examine the habitat selection by retaining these categories, only simplifying them to fit the hypotheses. This simplification took the habitat(s) of interest and coded them as 1, whereas all other habitats previously indicated in the studies to be avoided/not-selected for are coded as 0 (Fig. 4.13; 4.14; 4.15; 4.16; 4.17). For example, in the case of Malayan Kraits, buildings and natural areas were classed as 1, and all other types classed as 0. We then used the 0/1 classification as a predictor in the habitat selection analyses. We undertook this simplification to facilitate the repeated use of generic code in the multiverse. The second

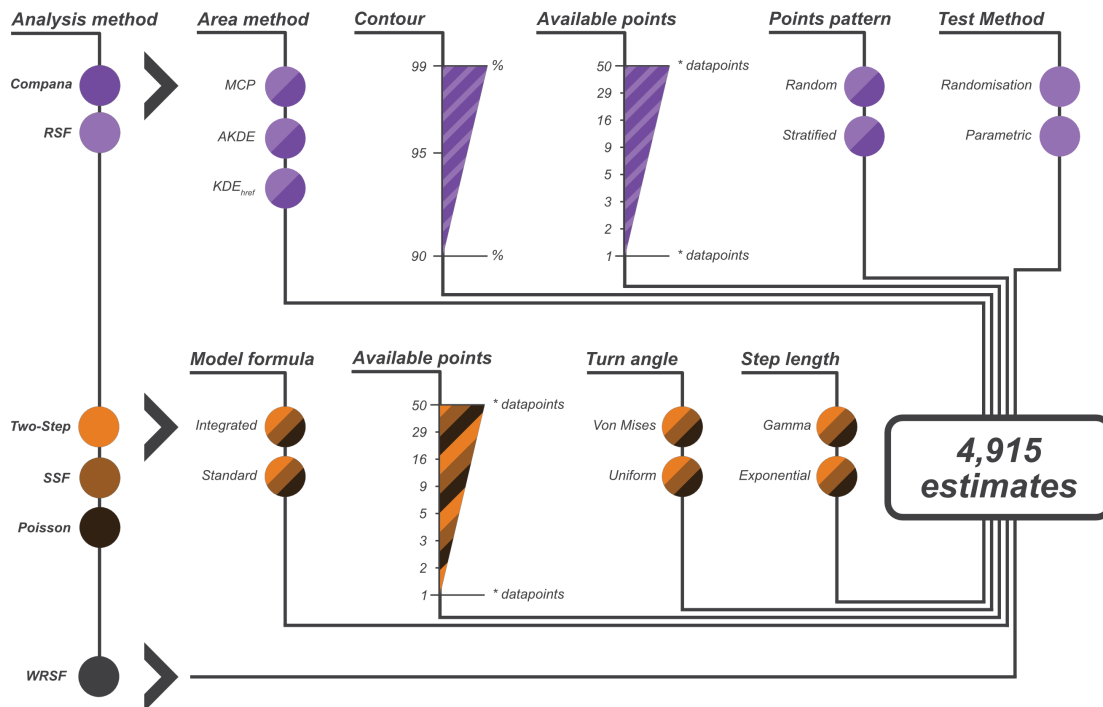


Figure 4.3: A visual of all the decisions contained in the multiverse and colour coded to which approach that decision applies. Subtitle indicates the name of the decision, and the values or categories below denote the choices explored in the multiverse. Each decision is combined with all others pertaining to the same analysis method, in unique combinations, and applied to each species to create the full count of habitat selection estimates.

approach involved converting the simplified categorical habitat types into continuous rasters, where each cell described the distance to a given habitat type. We inverted the distances (and scaled and centred), so that a positive effect in the model meant a positive selection towards a habitat type (the greater the effect the greater the selection), resulting in a more intuitive final output (Fig. 4.13; 4.14; 4.15; 4.16; 4.17). All analyses (see below) were run on both classifications, with the exception of the Compositional Analysis that required categorical inputs.

The habitat selection methods we explored consisted of the following:

1. Resource Selection Functions (RSF) (Gillies et al., 2006) and Composite Analysis (Compana) (Manly et al., 2007) that we called area-based methods, which used an estimated available area to help define availability for the snakes.
2. Step Selection Functions (SSF) (Forester, Im & Rathouz, 2009; Thurfjell, Ciuti & Boyce, 2014), Two-Step (Two-Step) (Craiu et al., 2011), and Poisson models (Poisson) (Muff, Signer & Fieberg, 2020) that we refer to as step-based methods, which used randomly generated alternative locations based on the movement and sinuosity of the individual being examined.
3. weighted Resource Selection Functions (wRSF) (Alston et al., 2023) a newly described method, which used a fitted movement model to calibrate the available habitat.

The area-based category made use of a defined area of availability, for which random available points can be drawn. To explore the impact of the definition of available, we created a number of different polygons surrounding the tracking data locations. We created available areas using Minimum Convex Polygons (MCP), Kernel Density Estimates (KDE: using the reference smoothing bandwidth [href]), and Autocorrelated Kernel Density Estimates (aKDE). We used the `adehabitatHR` package for MCP and KDE creation (Calenge & Scott Fortmann-Roe, 2023), and the `ctmm` package for AKDEs (Fleming et al., 2015; Calabrese, Fleming & Gurarie, 2016; Fleming & Calabrese, 2017, 2023). For the AKDEs, we fitted a number of movement models: Ornstein-Uhlenbeck (OU), Ornstein-Uhlenbeck Foraging (OUF), and Independent Identically Distributed (IID); we selected the best performing by AICc. We elected to follow guidance from (Silva et al., 2022), opting for weighted AKDEs and the perturbative hybrid residual maximum likelihood method (pHREML). The weighted AKDEs are better suited for tracking data with gaps. All three methods require an outermost contour to be selected to define the edge of the available habitat; we varied this outer edge including

analysis of 90, 95, and 99% contours. For the AKDE areas, we selected the point estimate connected to the contour percentage, ignoring the 95% confidence intervals associated with the estimate.

Once an available area was defined, we generated points within the area to extract the available habitat (also referred to as available points). We varied the point generation process, either spatially random or stratified across the overall area. In addition to the point generation method, we varied the the number of available points by multiplying the number of known/used data points by 1 to 50. This provided a reasonable variety of options, while keeping computational costs down. As each individual had their own available area, we explored Type III habitat selection (i.e., individual use compared to individually defined availability); this kept this analysis closer to the intent of the original publications.

Instead of using an available area, the step-based models use available steps randomly generated for each time step. For these step-based methods a different suite of choices was explored. First is the number of random steps generated per known location; we ranged this from 1 to 50. To generate those steps, we draw values from distributions, the choice of those distributions make up the next two choices. For the random step lengths we looked at the impact of using Gamma and Exponential distributions; whereas for the turn angles, we looked at Von Mises and Uniform distributions. Once the random available locations had been generated, we explore an additional choice regarding the model formula: whether to have the step lengths and turn angles interact with the habitat (Avgar et al., 2016). Termed “integrated step-selection”, the inclusion of the step lengths and turn angles is meant to aid the acquisition of less biased estimates of habitat selection. However, the impacts of the inclusion differ when using step-selection versus a Poisson model (see Chapter 3); we include it here to explore how dramatic that difference can be. While the Poisson models explicitly target population level selection, the SSF models required summation to a single value. To generate that value, we conducted a simple mean of the point estimates and calculated corresponding 95% confidence intervals. This process likely underestimates the overall uncertainty surrounding the SSF methods, as the uncertainty surrounding the initial estimates is lost.

We also used the newly developed wRSF methods (Alston et al., 2023). This approach dramatically reduces the choices during analysis; instead of requiring a defined area of availability, the wRSF weights individual animal locations (their impact penalised by repeated information resulting from autocorrelation) and compares them against a range distribution taking

the form of a weighted AKDE (Alston et al., 2023). Once the movement models had been generated, and wRSF models fitted, we used the mean functionality within the ctmm package to generate overall weighted population selection.

We handled the running of all analyses and their associated choices using the targets v.1.9.0 and tarchetypes v.0.11.0 R packages (Landau, 2021a,b). These packages allowed for efficient parallel processing and tracking of intermediate objects while avoiding errors in one analysis pathway from delays in the completion of others.

During the running of the multiverse, there were several instances of model fitting issues. King Cobra 002, 004, and 005 had issues with step lengths fitting to an exponential distribution; we addressed this by excluding the top 25% quantile of observed steps when the exponential fit failed. One Banded Krait individual did not have sufficient relocations to generate any area estimates; therefore we excluded them. There were infrequent instances where individuals did not enter both habitat types (or available points did not cover both types made more likely with fewer available points); this caused issues with model fitting in Two-Step and SSF approaches. We re-ran random step generation for those individuals to maximise the chances that available points captured both habitat types. If the convergence failure persisted, we excluded those individuals from that habitat selection estimate end-point.

4.3.5 Bayesian meta-analysis

We ran a series of 45 Bayesian meta-analysis models to determine an overall estimate of habitat selection for each analysis approach and hypothesis, while accounting for as much as the multiverse structure as possible (Auspurg & Brüderl, 2024). We used the following model formula: “estimate”|se(“estimate confidence”) ~ 1 + “choices” + (1|“index of analysis end point”). The estimate confidence measure varied from method to method: area-based Compana did not have a confidence interval, so we excluded that term; for the RSF and Two-Step estimates we used standard error; for the SSF models, we used standard errors of the model averaging; and for the Poisson models, we used the standard deviation. Choices varied per method, and included a term for every choice made to reach a given analysis end point. For the area-based methods (Compana, RSF), the choice terms included area method MCP, AKDE, KDEhref, contour percentage (90, 95, 99, scaled), available points generation method (random, stratified), and the number of available points generated per animal location (1 to 50, scaled). The Compana analysis had the additional choice of test method (randomisation,

parametric). For the step-based methods (Two-Step, SSF, Poisson), the choice terms included model formula (integrated, non-integrated), step length distribution (Gamma, Exponential), turn angle distribution (Von Mises, Uniform), and the number of random steps generated per animal location (1 to 50, scaled). The inclusions of “index of analysis end point” allows us to estimate τ^2 , a measure of between-study heterogeneity. We used weakly informative priors. For the intercept prior, we used a normal distribution with a mean of 0 and a standard deviation of 2, and for the choices’ priors we used Cauchy distributions with a location of -0.1 and scale as 3, except for the number of available points generated per animal location where we used a location of 0.1. We ran Bayesian models using the brms package (Bürkner, 2021), running 4 MCMC chains for 1200 iterations each with a warmup of 400 and a thinning rate of 2. For reporting, we used 95% median highest density continuous intervals when describing posteriors. We used R^2 and τ^2 to explore how much variation is explained by the analysis decisions, and the between-study heterogeneity (in this case analysis end points).

4.4 Results

4.4.1 Tracking Summary

Sample sizes and data density varied between species and individuals. We tracked 25 King Cobras, 7 Burmese Pythons, 13 Malayan Kraits, and 3 Banded Kraits, which had adequate data for analysis (Fig. 4.12). Tracking duration ranged widely from 37 to 1176 days (Table. 4.1), being cut short by snake deaths, transmitter failure, and studies ending. On average, we tracked an individual snake for $281.37 \pm \text{SD } 246.33$ days. Despite best efforts to maintain a consistent tracking regime, snakes evade rediscovery, weather prevents tracking, and equipment fails. Therefore, all species tend to have mean tracking lag times (i.e., the time between subsequent relocations) a little larger than intended (Fig. 4.11; Table. 4.1). Overall tracking effort resulted in 26,778 data points, across 48 snakes, resulting in a mean of $557.88 \pm \text{SD } 657.82$ per snake.

Activity and propensity to move varied between the species. We recorded the greatest number of moves for King Cobras ($273.28 \pm \text{SE } 47.86$), due to their more intensive tracking regime. The other species had on average fewer recorded moves per individual, with Burmese Pythons making on average $112.43 \pm \text{SE } 33.79$ moves, followed by Banded Kraits with $86.67 \pm \text{SE } 38.18$, and finally the more stationary Malayan Kraits with $28.54 \pm \text{SE } 3.95$.

4.4.2 Summarising Hypothesis Support

The area-based Compositional analysis appeared to provide majority support for three out of five of our hypotheses: King Cobra H1 (semi-natural areas), Malayan Krait H1 (buildings and natural areas), and Banded Krait H1 (field edges; Fig. 4.4). In contrast, King Cobra H2 (forests) and Burmese Python H1 (water bodies) hypotheses appeared almost entirely unsupported, with nearly zero analysis end points resulting in significant support (Fig. 4.18; 4.23). The lack of support does not appear to be connected to any single analysis decision. King Cobra H2 and Burmese Python H1 hypotheses appear more likely to be supported when using MCPs to define available area, and least likely when defined via AKDEs, but even using MCPs there is not consistent support in the point estimates. High outlying support in Burmese Python H1 estimates are somewhat mitigated when larger numbers of available points are used. Also, the other species revealed that available area method and contour can have varying impacts. Only the Compana test method (randomisation or parametric) appears to have a consistent effect, with both Krait hypotheses seeing dramatically higher estimates of selection when parametric testing is used. For Malayan Kraits, this does not change any conclusions (all estimates are significantly supportive; Fig. 4.28); however, for Banded Kraits this single decision is responsible for the half of estimates providing support (Fig. 4.33).

Unlike the Compositional analysis, Resource Selection Functions (RSF) estimate medians provided support for all hypotheses (Fig. 4.5); although, King Cobra H2's support is marginal. The more limited support for King Cobra H2 (for both habitat classification methods) appears driven by estimates based on MCP available areas and using larger contours (Fig. 4.19). All other hypotheses are universally supported, and significantly so (with the exception of only six of 144 Burmese Python estimates). The chosen available area method does appear to impact the variability of results, even among the positive selection estimates. Those based on MCPs appear to be less variable regardless of other analysis decisions, but this pattern only exists when a continuous habitat classification was used. The continuous habitat classification does appear to open the possibility for more dramatic variation in estimated selection strength. The number of available points impacts the estimates in both Krait species, but only when not using MCPs, and this impact appears to diminish above 29 available points per observed snake location.

The Two-Step approach was by far the least effective and most problematic when running within a multiverse, as the models had frequent convergence issues. This failure rate is

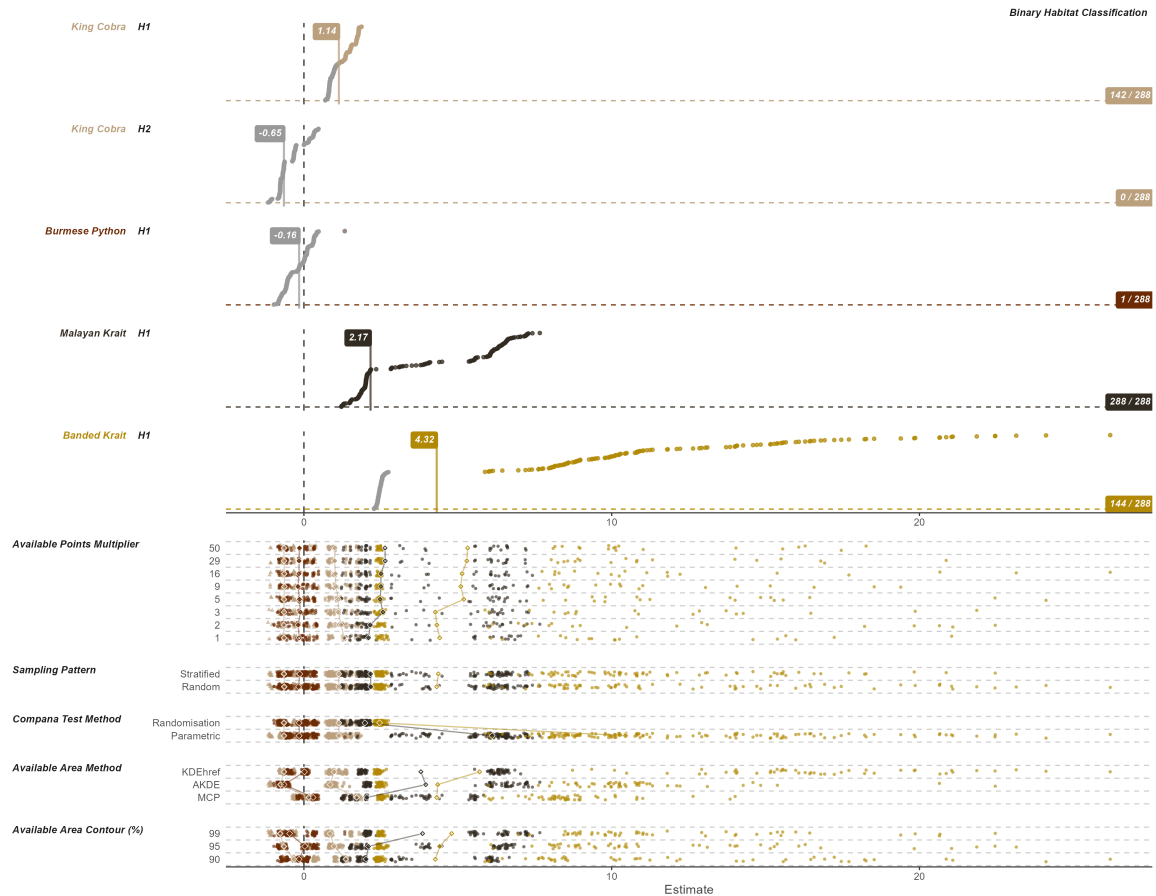


Figure 4.4: All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot show the estimates relative to each analysis choice. The colours depict the species, and shape separate hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines.

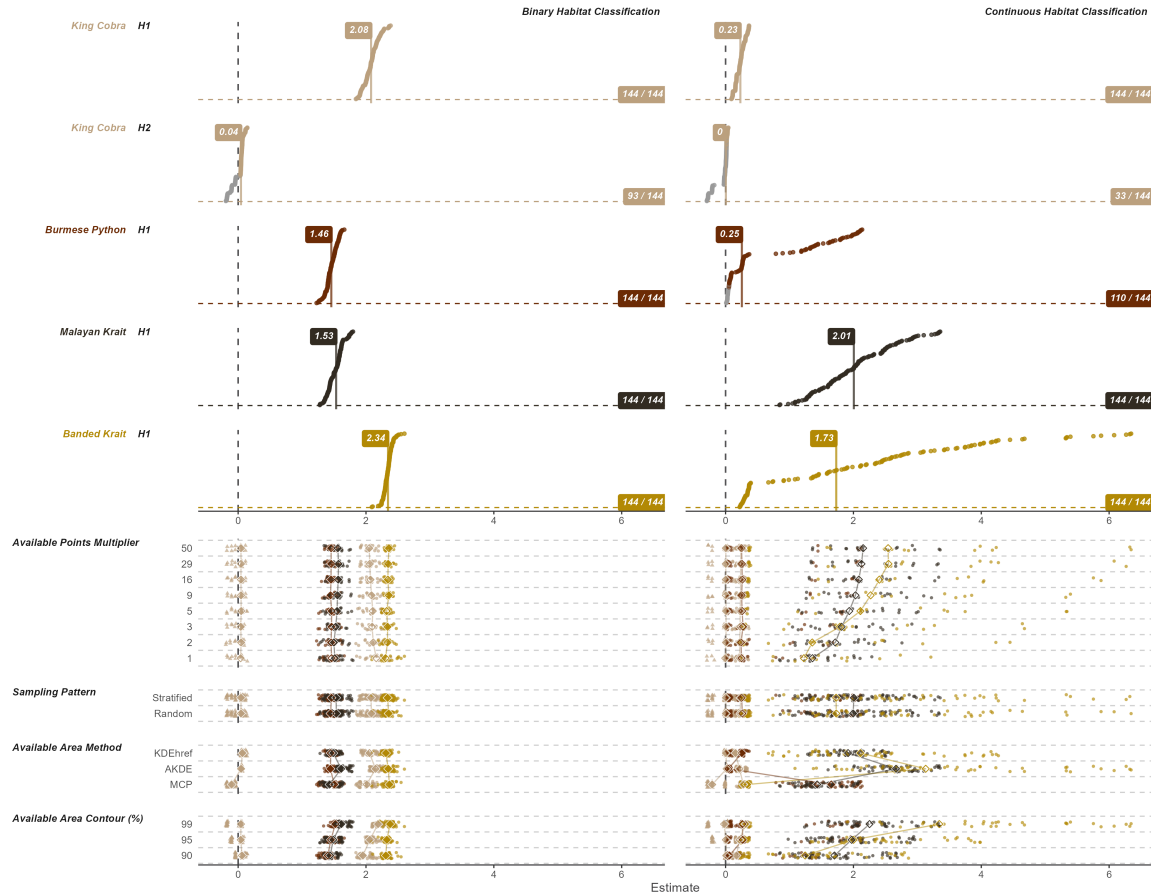


Figure 4.5: All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot shows the estimates relative to each analysis choice. The colours depict the species, and shape separates hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

particularly visible in the King Cobra data when run with low numbers of available points or with attempts to integrate step and turn angle into the model formula (Fig. 4.6). In addition to the failure rate, we see a more even spread of estimates that contrasts with the ideal “S” shape that would indicate more agreement between the analysis end points. The lack of agreement in selection strength appears more prevalent when using the continuous habitat classification, but unlike with RSFs, increased variability does not appear connected to any single choice. The most consistent impact of the choice is whether to integrate step and turn angle into the model formula. As mentioned, this led to convergence issues for King Cobras, but for other species, the integrated formulations tended to result in more conservative estimates of selection. Only for Malayan Kraits with a binary classification did the median estimate swap direction due to this analysis choice, and this explains the lower rates of significant selection (32/64; Fig. 4.30).

The Step Selection Function approach was impacted by the habitat classification method (Fig. 4.7). When using a binary habitat classification, only the Kraits results showed broad support for the hypotheses. Whereas when using the the continuous habitat classification, we see more universal support across species. This lack of support when using the binary classification was not clearly connected to any particular analysis decision. The only decision that makes a consistent change when paired with binary classification was the number of available steps. Increased steps tended to lower the overall variability of selection estimates, especially for Malayan Kraits. While we did see near universal support for hypotheses in the continuous classification, the estimates for Burmese Pythons were more marginal with only 22 of 64 considered significantly supported (Fig. 4.26). We should be cautious with this support, as the SSF model averaging method may be additionally reducing the uncertainty present in individual SSF models. The impact of decisions are most visible in Malayan Kraits, whose estimates’ strength is impacted by the step and turn angle distributions, as well as whether the formula was integrated. The distribution choices make less of an impact, with the spread of estimates largely overlapping, but using an integrated model formula led to more generous and variable estimates of preference. However, despite this impact, all bar one of Malayan Kraits’ estimates significantly support the hypothesis regardless of habitat classification method (Fig. 4.31).

The Poisson model approach appears similar to the SSF results, with universal support for the hypotheses when examining the median estimates (Fig. 4.8). Once again the most marginal support was seen in King Cobra H2, where both classification methods resulted in low median

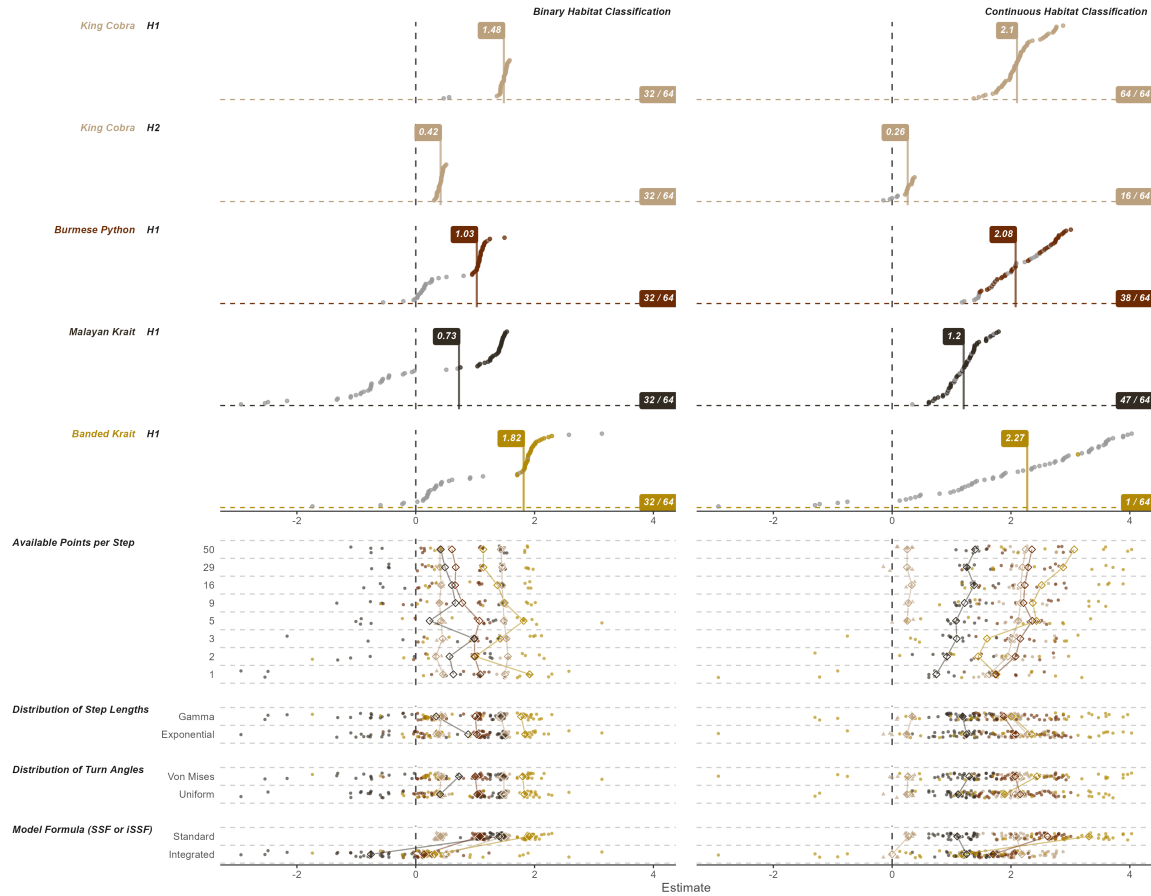


Figure 4.6: All estimates of habitat selection derived from the step-based Two-Step analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot shows the estimates relative to each analysis choice. The colours depict the species, and shape separates hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

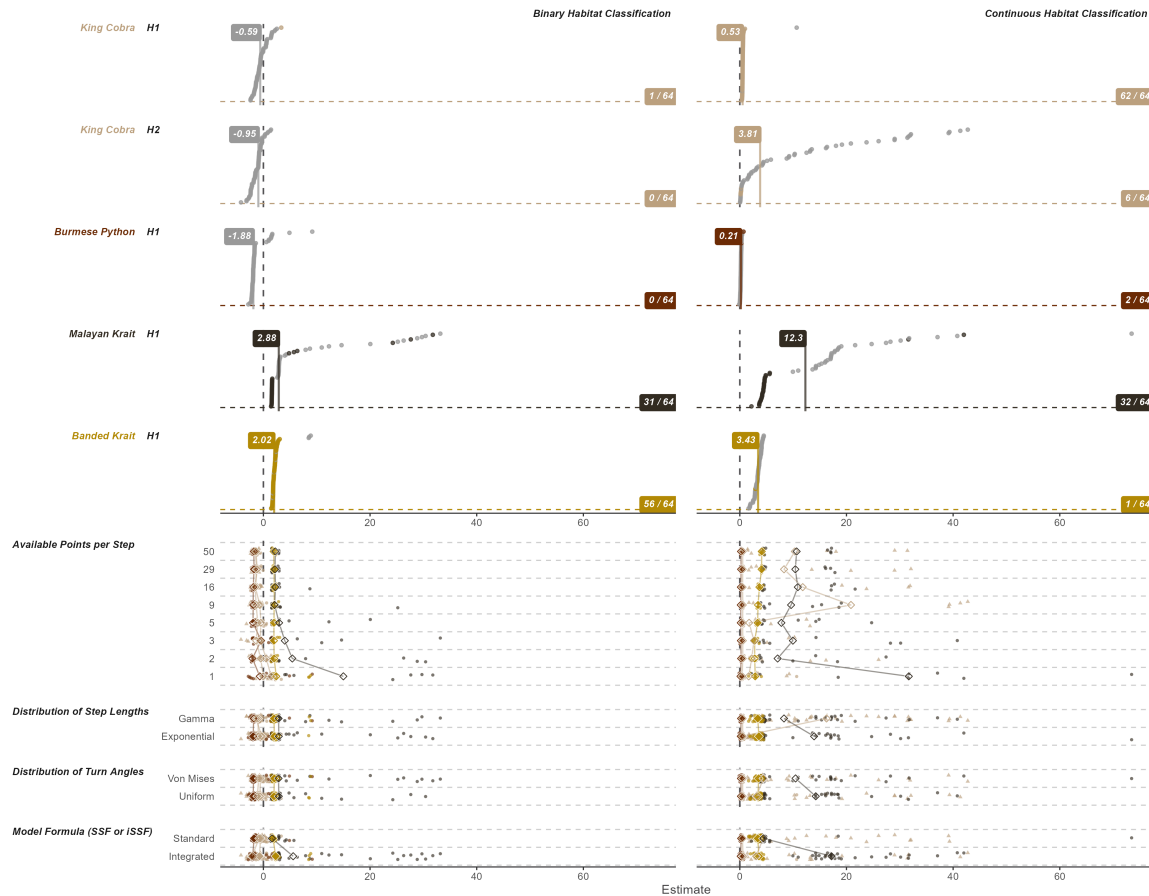


Figure 4.7: All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot shows the estimates relative to each analysis choice. The colours depict the species, and shape separates hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

estimates of preference. The marginal medians seem to be largely due to a single decision: whether or not to run an integrated model formulation. In this case, the integrated models provide far more conservative estimates of selection, and for King Cobra H2 when using binary classification the median falls below zero (Fig. 4.22). The impact of model formulation appears for all species, but the most dramatic impacts are seen when paired with binary habitat classification. What is missed from the point estimates, and median summaries, is the large uncertainty surrounding estimates when using the binary classification. Zero estimates when using the binary habitat provided significant support. This contrasts with the continuous habitat classification, where all species-hypothesis combinations show complete or majority significant support. The impacts of other decisions are considerably smaller than model formulation. There appear to be advantages to generating more available points, as estimates with few points tend to explain the upper and lower outliers. The mitigation of extreme estimates is best seen in Banded Kraits when using a continuous habitat classification (Fig. 4.37).

Our final analysis approach was using weighted Resource Selection Functions (wRSF). All wRSF results provided support for the hypotheses (Fig. 4.9). Although, once again, support for King Cobra H2 when using a continuous classification is borderline and ultimately rounded to zero. The King Cobra H1 saw no support when using the continuous habitat classification but clear support when the landscape was binary classified. All other species saw unambiguous support when the habitat was binary classed, but when classed continuously as inverted distance, the 95% confidence intervals for Burmese Pythons and Banded Kraits approximated zero.

4.4.3 Meta-analysis Estimates

The Bayesian meta-analyses tended to perform very well, with the population effects producing high R^2 values ($22/45 > 0.6$, median: $0.6 \pm \text{SD } 0.31$). The RSF models had the highest R^2 values, with the continuous classification estimates being almost entirely described by the predictors (Fig. 4.38). The step-based approaches tended to have more variable R^2 , with step-selection models having the lowest values.

τ^2 values (between analysis end point heterogeneity) ranged from very low (0.0061) for the binary classed RSF models examining Banded Kraits, to 1.6 for the binary classed Poisson models examining the same species. Mostly τ^2 values fell below 0.1 $27/45$ (overall median of

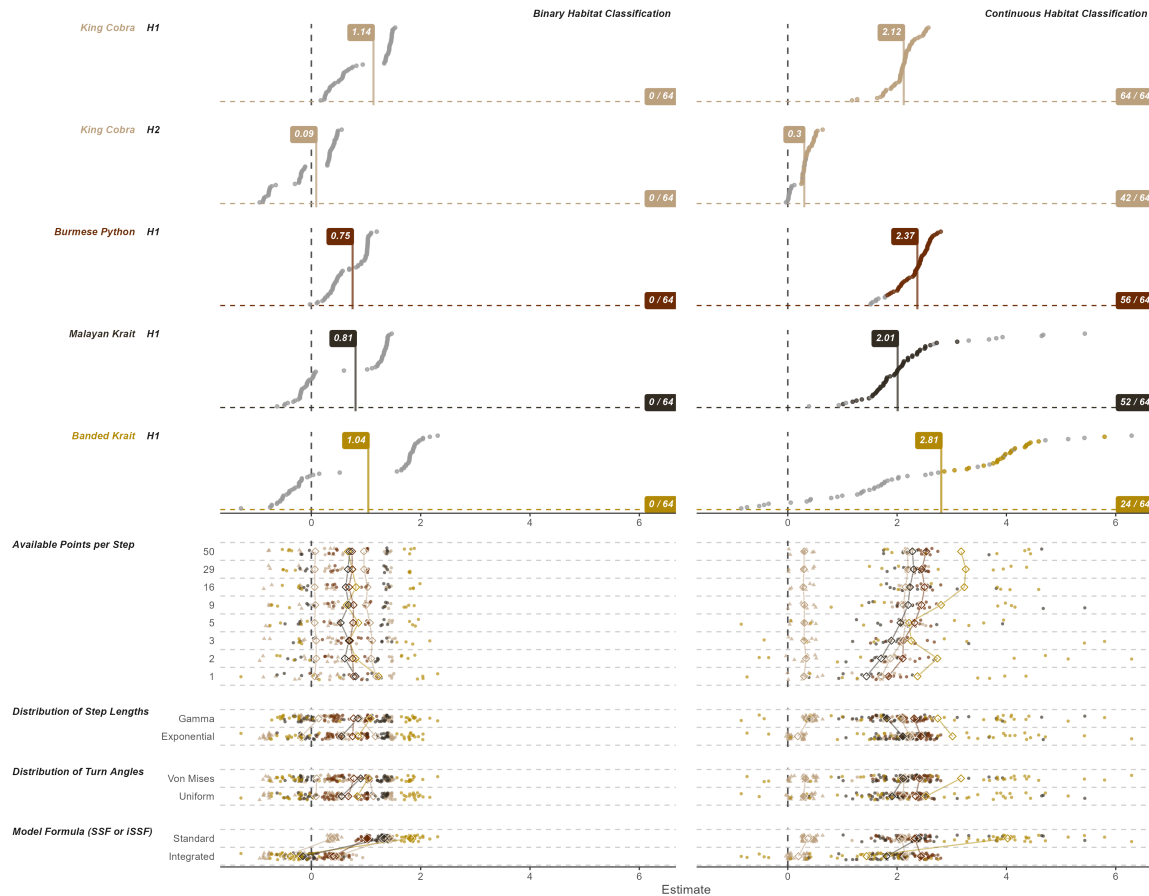


Figure 4.8: All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis. Top curves show all estimates, split by species and hypothesis, with coloured points indicating those with point estimates supporting the hypothesis (i.e., > 0). Labelled vertical lines show the median estimate for each species-hypothesis combination. Lower plot shows the estimates relative to each analysis choice. The colours depict the species, and shape separate hypothesis 1 and 2 for the King Cobras (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and species-hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

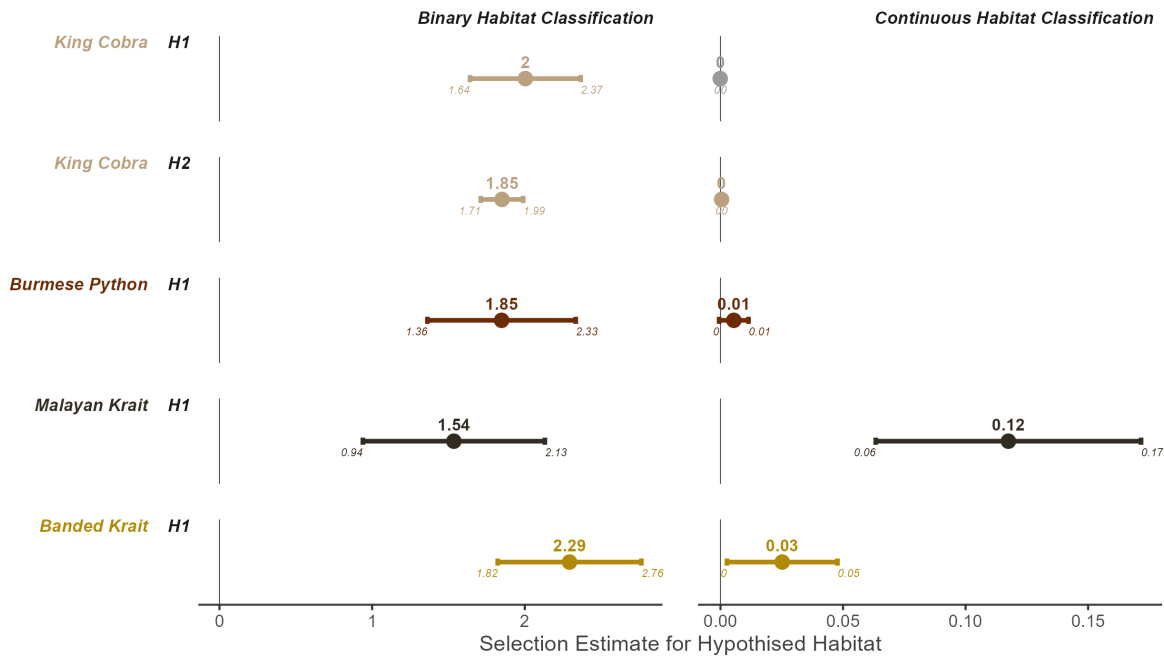


Figure 4.9: All estimates of habitat selection derived from the weighted Resource Selection Function (wRSF) analysis. Circles and numbers above show the point estimate, and the horizontal bars depict the 95% confidence interval. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

0.05 $\pm$ SD 0.42).

The intercepts of the meta-analysis models reveal that 38/42 hypotheses are supported overall, with the only exceptions existing when the habitat was binary classified (Fig. 4.10): King Cobra H2 and Burmese Python H1 using Compana, Malayan Krait H1 using Two-Step, King Cobra H2 and Burmese Python H1 using SSF.

Overall the patterns, of meta-analysis results mirror the specification curves, and the similarities reappear in the analysis choice impacts. Model formulation appeared to have the most dramatic effect on the meta-analysis estimates, and that effect is seen in all three step-based approaches (Fig. 4.39). The effects themselves were not consistent, with different directions and strengths depending on approach and species. Only Banded Kraits were impacted consistently the same way, with the non-integrated estimates tending to have a positive impact on the estimates (Fig. 4.39). Another choice that appears to have considerable impact was the available area method. Like the model formulation, the impact varied by species and landscape classification. It reveals that impact of available area method is not necessarily very predictable, and likely interacts with the specific movement characteristics of the study species. Available points, either steps or points within an area, tended to be positively related

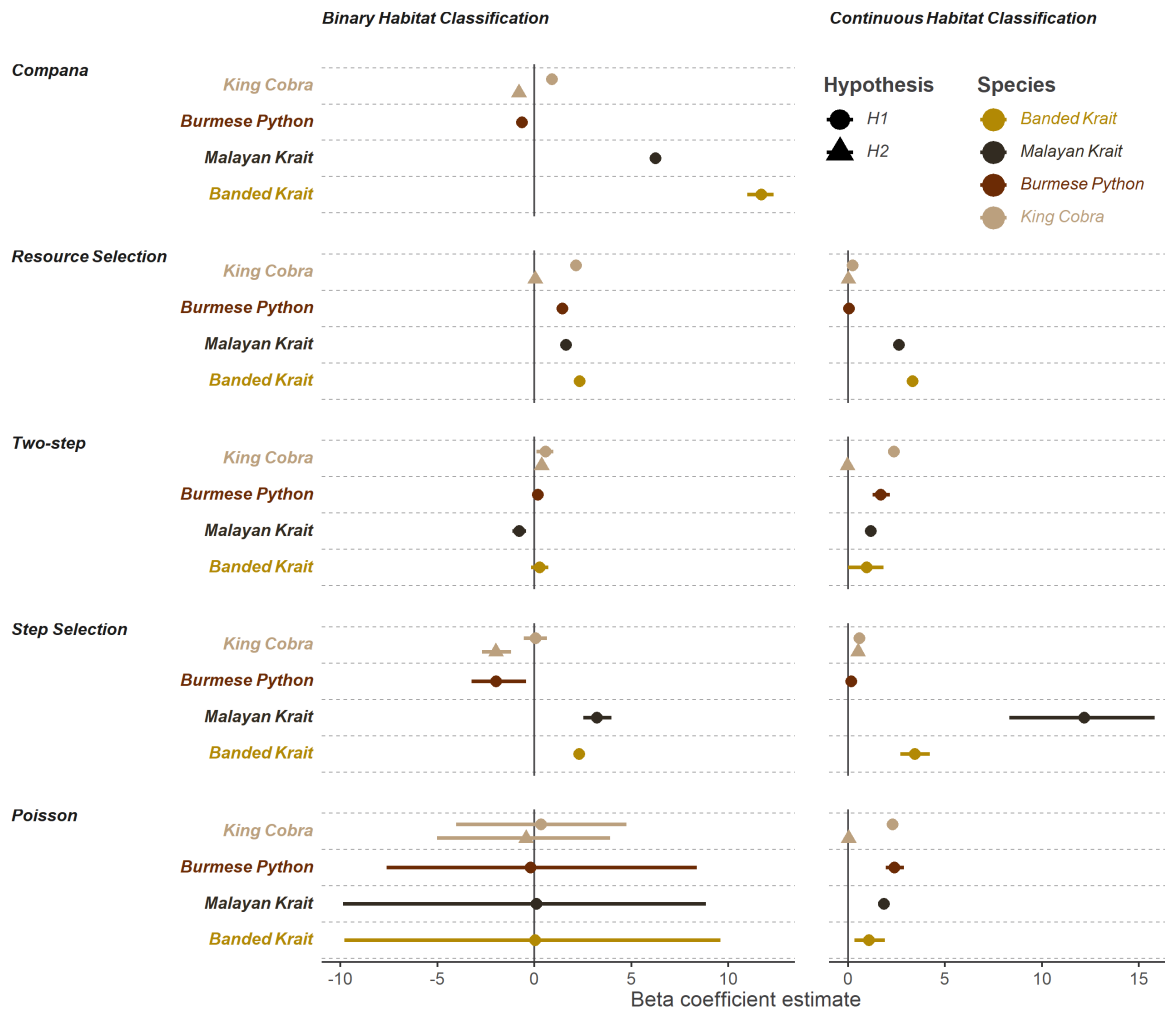


Figure 4.10: Posterior estimates of the intercept for all meta-analysis models, with 95% HDICs, indicating whether the overall meta-analysed habitat selection estimate supports the original conclusions. Positive values indicate support. Colour denote species, and shape separates hypotheses for King Cobras. Continuous habitat classification was incompatible with the Compana method, and so is not shown.

to estimate strength. This was reasonably universal, but the larger effects were seen in the Krait species, potentially suggesting greater impacts when the habitat's are defined at a finer spatial scale.

4.5 Discussion

Our re-testing of snake habitat selection studies using a multiverse of over 4915 analysis endpoints suggests that the original conclusions remain broadly supported. - King Cobras in a mixed forested and agricultural landscape select for semi-natural areas near irrigation canals and forests. - Burmese Pythons in agricultural land select for water bodies. - Malayan Kraits on a university campus select for buildings and natural vegetated areas. - Banded Kraits in agricultural land select for field margins. More specifically, it appears that there is a reasonably low chance that the analysis choices made in the original publications led to strange or outlying results. While the multiverse revealed broad support for the direction of the selection (i.e., agreeing that the hypothesised habitat is selected for), the strength of the effect/selection was more sensitive, and we did identify a few analysis choices capable of contradicting the originally published results.

Of the four species we examined, Burmese Pythons' preference for water bodies appears the most marginal, but the preference was clear in all but six RSF analyses and the wRSF analysis. All these analysis approaches have different assumptions, and treat the data in different ways. Therefore, the apparent disagreement between RSF and other methods may tell us something about the preference that Burmese Pythons are demonstrating. Burmese Pythons are ambush hunters; therefore, their movements may be less informative compared to their overall location in the landscape; particularly if they remain at the same location for extended periods, for example during nesting (Smith et al., 2024). RSFs' de-emphasis of movement compared to step-based methods, may explain the differences in support. The support from the more modern wRSF may also be ascribable to a broader selection for water bodies.

We only identified two decisions that could clearly alter conclusions, and even those did not appear universal. Namely the choice to use randomisation or parametric tests in the Compositional analysis, and whether to integrate step lengths and turn angles in the Step Selection and Poisson models. Banded Kraits' selection for field margins was supported by the vast majority of analyses, but the weakest support was when using Compositional analysis.

While the point estimates still suggested selection, a decision to use randomisation rather than parametric tests for significance would render all analysis non-significant. Previous simulated scenarios have also revealed the sensitivity of estimates to this choice (Chapter 3). Other species did not see the same impact; randomisation did lead to reduced strength estimates for Malayan Kraits, but all remained significant. We suspect the weaker support seen in Compositional Analysis may be the result of location error. Compositional analysis could only be used with the binary habitat classification. While every effort was made during tracking to ensure a correct GPS location was retrieved, error is unavoidable especially in denser forested areas or near buildings. Even small errors may lead to mismatches between the true location (and habitat) of an animal and the land-use data we had access to. This mismatch may be more pronounced when habitats are rarer or when animals tend to remain near its edges. Banded Kraits' focus on very thin field margins may have made it particularly susceptible to misclassified habitat use, and similarly Burmese Pythons' selection for the banks of water bodies. The rarity of these habitats may have also played a role, as we see a clear narrowing of the confidence intervals surrounding the estimates as we increased the number of available points sampled. We suspect that location error resulting in binary habitat mis-classification may explain the large confidence intervals of the Poisson models and the limited support in Burmese Python SSF analyses, as this uncertainty and lack of support is not visible in the continuous habitat models.

A second decision that could alter estimate strength dramatically was whether we used an integrated model formulation (i.e., include step and turn angle in the model formula). The work of Muff, Signer & Fieberg (2020) and Forester, Im & Rathouz (2009), as well as chapter 3 showed that integrated formulas have contrasting impacts on estimates in step-selection approaches and poisson models. The step-selection models tend to produce less biased estimates when using an integrated formula, whereas the poisson models produced the opposite. With those findings in mind, we should weight the findings using the non-integrated poisson models more heavily. The non-integrated Poisson models show clearer support for the hypotheses in all cases with either classification, bar King Cobra H1 and Burmese Python H1 whose estimates seem negligibly impacted when using the continuous classification. Unexpectedly the choice of model formulation did not have as noticeable impact in the step-selection approach, where only Malayan Kraits' estimates were markedly impacted. Despite this, the strength of the selection exhibited by Malayan Kraits meant that the more conservative estimates from the non-integrated models did not render the selection estimates non-significant.

Overall, we found that the final estimates could not be completely predicted by the analysis choices, despite there existing no sampling variation in the primary data (i.e., the same dataset was analysed each time). In some cases, such as RSF binary, the analysis decisions explained the vast majority of variation and τ^2 indicated that there was very little heterogeneity between end points. Whereas others such as the Poisson binary meta-analysis models left most of the variation unexplained, and indicated much higher levels of between end point heterogeneity, which appear comparable to meta-analysis on multiple real studies that would also need to be contending with sampling variation. The propensity for the step-based methods to show higher τ^2 values could be due to random sampling of step and turn angles in the available points. The variation in τ^2 and R^2 values from the meta-analysis models indicate that analysis choices are interacting with the species data and hypothesis, suggesting variation in habitat selection estimates may be difficult to predict in other systems and limited the generalisability of patterns shown here.

While the average or summation of multiverse findings cannot be used as to estimate a true effect, they do highlight the high likelihood of repeated analyses agreeing. Their agreement with the other results (in terms of direction), and the fact they largely provide significant support for the original hypothesis is reason to be confident in the original conclusions.

4.6 Conclusions

King Cobra

The previously published conclusions that King Cobras make heavy use of semi-natural areas [i.e., banks of irrigation canals, and scrub between fields; Marshall et al. (2020a)] is nearly universally supported regardless of approach, especially when classified as inverted continuous distance. Why King Cobras use these habitats is not immediately clear, but it could be driven by prey availability (e.g., other snake species Barnes et al., 2017; Strine et al., 2018), the avoidance of threats, and density of shelter sites. Other reptiles are known to make use of linear habitats (Kay et al., 2016; Doherty, Fist & Driscoll, 2019), including other snakes at this study site [i.e., King Cobra's prey; Jones et al. (2020)]. Both prey species and King Cobras could be using these linear connected semi-natural areas as a means of avoiding the more exposed and dangerous fields and anthropogenic areas. Human activity in fields, such as ploughing, will kill snakes (Knierim, Barnes & Hodges, 2017), and we have documented many fatalities

of King Cobras as a direct result of human actions [e.g., persecution, road collision; Marshall et al. (2020a)]. We can be confident that suggestions regarding the importance of edge or scrub habitats amongst a human-modified landscape are robust in the face of a multitude of analysis choices.

The second hypothesis regarding their selection for forest was not nearly as widely supported and only appears to be supported in certain analysis approaches. The original paper highlighted the propensity for King Cobras to leave the forest and the heightened movement particularly during breeding season (Marshall et al., 2019), potentially mate searching behaviour. Mate searching has been indicated as a potential seasonal driver of heightened movements in other snake species (Duvall & Schuett, 1997; Bauder et al., 2016). The King Cobra dataset used in the first paper (Marshall et al., 2019) was skewed towards King Cobras found in/near protected forest, hence the focus on forest use. The data used here pooled old and new data that presents a more complete picture of King Cobra selection, likely weakened that association with protected forest. However, data collected after the original paper showed individuals using the forest for nesting; to maintain access to the forest these individuals travelled via semi-natural areas and were required to cross roads (Marshall et al., 2020a; Jones et al., 2022). The nest construction of King Cobras requires adequate vegetated material to maintain the requisite internal nest temperatures (Dolia, Das & Kelkar, 2023). This highlights the potential importance of forested areas during certain times of year, even if it is less likely to be identified by the broad (rather generalised) habitat selection multiverse-analysis here. Other studies have revealed King Cobra potentially favour less disturbed areas for nesting, even if a clear preference of forests is not apparent (Koirala & Dawa Tshering, 2021).

Burmese Python

The conclusions regarding Burmese Pythons' habitat selection appear susceptible to differing analysis approaches, but overall we can be fairly confident in the assertion that Burmese Pythons ideally remain closer to water bodies. The patterns of selection in Burmese Pythons appear more clearly with broader landscape level approaches compared to those using steps. It is likely that Burmese Pythons are prioritising being close to water bodies for food and shelter. The use of water bodies and aquatic landscapes is well documented in Burmese Python's invasive range in Florida, USA (Mutascio et al., 2018), and in Thailand evidence of Burmese Pythons capitalising on both shelter and prey provided by the water bodies is apparent. As

mentioned for the King Cobras, edge areas in agricultural land in this area of Thailand are considerably less trafficked and modified compared to the agricultural matrix of crops. Less disturbance likely boosts shelter site options, as well as ambush sites and associated prey (e.g., aquatic birds). Shelter sites along edge habitats, such as field margins and banks, appear to be suitable for nesting [also evidenced by Banded Krait nest site choice; Smith et al. (2024); Knierim et al. (2019)]. Nesting requires prolonged periods of non-disturbance, evidencing the potential stability of some of these edge habitat areas. Whereas for food, rice paddies can house suitable sized water birds (Fujioka, Don Lee & Kurechi, 2010), and Burmese Pythons in this area have been documented depredating domestic waterfowl (Smith et al., 2021), together supporting suggestions that the water bodies are a resource.

Malayan Krait

The previously published conclusions that Malayan Kraits prefer buildings and natural areas appears to be the most clearly supported and one of the least susceptible to analysis choices. Broadly the analysis would support their selection for buildings and natural areas on the university campus. We suspect the support is not due to a particular movement pattern, but instead the strength of the selection, which is visible in the wRSF and meta-analysis intercepts. Malayan Kraits are likely avoiding open areas, instead choosing to make use of the presumably safer micro-habitats available in buildings and natural areas. Hodges et al. (2022) suggests these environments may be better for thermoregulation, and other evidence reveals buildings are used for foraging (Hodges, D'souza & Jintapirom, 2020; Hodges et al., 2021a). The selection for buildings presents a clear potential for human-snake conflict. Malayan Krait are highly medically significant (Looareesuwan, Viravan & Warrell, 1988; SEARO, 2016), and apparent overlap between their preference and human movements could heighten the chances of human-snake interactions resulting the harm of either human or snake. Fortunately, Malayan Krait are mostly nocturnal, presenting the chance for human and snake to avoid conflict temporally. The robustness of the selection analyses means we can safely support the suggestions in Hodges et al. (2022) for the university to encourage the use of handheld lights when moving in the evening and night, wearing of closed footwear, and ensuring buildings are adequately sealed (Samuel et al., 2020). The selection for natural areas may further present an opportunity to draw snakes away from buildings, if more ideal shelter sites are available farther from concentrations of human activity. Studies elsewhere in Thailand have suggested

that other medically significant snakes will avoid buildings, preferring more vegetated areas (Barnes et al., 2024).

Banded Krait

Banded Kraits' conclusions fall mostly in the middle; they are very well supported by the overall meta-analysis intercepts but did clearly reveal some scenarios where researcher choice could have obscured what appears to be a selection for field margins. As mentioned for other species, field margins and other edge habitats appear key to snakes' lives in agricultural areas in northeastern Thailand (Knierim et al., 2018), providing shelter and movement corridors. The movements of Banded Kraits support this. Their use of edge habitats is not limited to sheltering, as like Burmese Pythons, the edges offer suitable (possibly communal) nest sites (Knierim et al., 2019). When paired with other studies indicating the importance of field margins for mammals and amphibians (Brown & Singleton, 2001; Groffen, Borzée & Jang, 2018), particularly rice paddy edges, it is clear that Banded Kraits likely find all the resources required in these edge habitats. The concentration to the field margins may indicate a human-snake interaction hotspot, the identification of which could aid refinement of agent-based snakebite risk maps (e.g., Goldstein et al., 2021).

Overall conclusions

Looking forward, the multiverse examinations here provide support for the broad conclusions drawn by the original studies and that the analysis choices of the original publications did not result in peculiar conclusions. The support originating from more modern methods is additionally reassuring. The wRSFs should be best placed to provide us with a broad estimate of habitat selection, while simultaneously tackling the issues arising from inconsistent tracking (Alston et al., 2023). However, the use of exact selection strengths should be examined in more detail if they are to be adapted to future purposes such as human-snake conflict prediction [as supported by simulation habitat multiverses; Marshall & Duthie (2024); Chapter 3]. In those cases, reassessment of the selection using newer methods more acutely tailored to prediction may be required, either by greater integration of mechanistic drivers of snake movement, or via methods that are scale independent enabling better beyond-sample extrapolation (e.g., Michelot et al., 2019).

Summarised together, this study reveals the robustness of snake habitat selection studies to researcher analysis choice, while also reaffirming the conclusions regarding the important role that habitat patches (edges, banks, canals, etc.) play for snakes in an otherwise human dominated unsuitable landscape.

4.6.1 Acknowledgements

BMM was funded by the Natural Environment Research Council (NERC) via the IAPETUS2 Doctoral Training Partnership.

4.6.2 Software availability

All code used in the creation of this manuscript is available at <https://github.com/BenMMarshall/multiverseSnakes>; and archived at <https://doi.org/10.5281/zenodo.14034844>.

In addition to packages mentioned in the methods, a complete list of software used follows.

We used *R* v.4.4.2 (R Core Team, 2022) via *RStudio* v.2023.6.2.561 (RStudio Team, 2022). We used *here* v.1.0.1 (Müller, 2020b), *readr* v.2.1.5 (Wickham, Hester & Bryan, 2024), *inborutils* v.0.4.0 (Van Calster, Oldoni & Van Hoey, 2023), *targets* v.1.9.0 and *tarchetypes* v.0.11.0 (Landau, 2021a,b), and *qs* v.0.27.2 (Ching, 2023) to manage directory addresses and saved objects.

We used *ggplot2* v.3.5.1 for creating figures (Wickham, 2016b), with the expansions: *patchwork* v.1.3.0 (Pedersen, 2024), *ggridges* v.0.5.6 (Wilke, 2024), *ggtext* v.0.1.2 (Wilke & Wiernik, 2022), *ggnewscale* v.0.5.0 (Campitelli, 2024), and *ggdist* v.3.3.2 (Kay, 2023a).

We used *brms* v.2.22.0 (Bürkner, 2021) to run Bayesian models, with diagnostics generated used *bayesplot* v.1.11.1 (Gabry & Mahr, 2024), *tidybayes* v.3.0.7 (Kay, 2023b), and *performance* v.0.12.4 (Lüdtke et al., 2021).

We used the *dplyr* v.1.1.4 (Wickham et al., 2023), *tidyr* v.1.3.1 (Wickham, Vaughan & Girlich, 2023), *tibble* v.3.2.1 (Müller & Wickham, 2023), and *stringr* v.1.5.1 (Wickham, 2022) packages for data manipulation.

We used *sp* v.2.1.4 (Pebesma & Bivand, 2005b), *sf* v.1.0.19 (Pebesma, 2018; Cooley, 2023; Pebesma & Bivand, 2023), *raster* v.3.6.30 (Hijmans, 2023a), *terra* v.1.8.29 (Hijmans, 2023b),

tidyterra v.0.6.1 (Hernangómez, 2023), and *move* v.4.2.6 (Kranstauber, Smolla & Scharf, 2023) for manipulation of spatial data and estimation of space use not otherwise mentioned in the methods. We used *ctmm* v.1.2.1 (Fleming & Calabrese, 2023), *amt* v.0.2.2.0 (Signer, Fieberg & Avgar, 2019), *adehabitatHS* v.0.3.18 (Calenge & Mathieu Basille, 2023), *adehabitatHR* v.0.4.22 (Calenge & Scott Fortmann-Roe, 2023), *lme4* v.1.1.35.5 (Bates et al., 2015), *INLA* v.24.6.27 (Martins et al., 2013b; Lindgren & Rue, 2015), and *TwoStepCLogit* v.1.2.5 (Craiu et al., 2016) for the habitat analyses.

We used *rmarkdown* v.2.29 (Xie, Allaire & Golemund, 2018; Xie, Dervieux & Riederer, 2020; Allaire et al., 2023), *bookdown* v.0.42 (Xie, 2016, 2022a), *tinytex* v.0.54 (Xie, 2019, 2023a), and *knitr* v.1.49 (Xie, 2014, 2015, 2023b) packages to generate type-set outputs.

We generated R package citations with the aid of *grateful* v.0.2.10 (Francisco Rodríguez-Sánchez, Connor P. Jackson & Shaurita D. Hutchins, 2023).

4.6.3 Data availability

The majority of the data is already publicly available. Movement data can be accessed via Movebank (Kays et al., 2022), with the following study IDs: 1649411628, 556564170, 1093796277, 1204811528. Movement data and land use data was also available via the following data repositories: Marshall et al. (2020b); Smith et al. (2020); Hodges et al. (2021b).

For the ease of recreating the results presented here, we have uploaded all movement and land use data, alongside the code used to generate this manuscript. It is available at <https://github.com/BenMMarshall/multiverseSnakes>, and archived at <https://doi.org/10.5281/zenodo.14034844>.

If using this compiled data, we highly suggest also citing the original data paper and repository as listed above to ensure the data maintains a close connection to descriptions of its collection.

4.6.4 Author Contributions

Conceptualization: BMM, ABD, MC, IS, CTS, SNS, MDJ, TKK, and CWH; Data curation: BMM; Formal analysis: BMM; Investigation: BMM, MC, IS, CTS, SNS, MDJ, TKK, and CWH; Methodology: BMM and ABD; Software: BMM; Supervision: ABD; Visualization: BMM; Writing - original draft: BMM; Writing - review & editing: BMM, ABD, MC, IS,

CTS, SNS, MDJ, TKK, and CWH.

4.7 Supplementary Material

Table 4.1: Summary of tracking effort, tracking time lags (hours), and step lengths (metres) for each species.

Species	Individual ID	Number of Datapoints	Total Tracking Duration (days)	Mean Time Lag $\pm$ Standard Error (hours)	Number of Moves	Mean Step Length $\pm$ Standard Error (metres)
Bungarus candidus	BUCA001	116	107.50069	22.43 $\pm$ 2.26	34	31.01 $\pm$ 8.24
Bungarus candidus	BUCA002	230	222.77083	23.35 $\pm$ 1.04	64	22.05 $\pm$ 3.89
Bungarus candidus	BUCA007	130	144.47917	26.88 $\pm$ 2.56	25	23.74 $\pm$ 6.25
Bungarus candidus	BUCA012	144	161.97569	27.18 $\pm$ 2.8	25	26.08 $\pm$ 8.4
Bungarus candidus	BUCA014	77	80.65764	25.47 $\pm$ 1.49	19	22.06 $\pm$ 6.25
Bungarus candidus	BUCA016	112	116.19444	25.12 $\pm$ 1.59	17	6.14 $\pm$ 1.94
Bungarus candidus	BUCA022	51	56.67708	27.2 $\pm$ 3.25	37	37.7 $\pm$ 9.43
Bungarus candidus	BUCA027	32	36.78819	28.48 $\pm$ 4.52	9	20.23 $\pm$ 9.37
Bungarus candidus	BUCA028	117	106.53472	22.04 $\pm$ 1.21	43	17.87 $\pm$ 3.82
Bungarus candidus	BUCA032	99	107.65764	26.37 $\pm$ 1.77	27	61.47 $\pm$ 17.3
Bungarus candidus	BUCA033	179	187.49167	25.28 $\pm$ 1	34	15.86 $\pm$ 3.14
Bungarus candidus	BUCA035	141	144.82639	24.83 $\pm$ 0.82	19	20.22 $\pm$ 5.08
Bungarus candidus	BUCA036	54	68.72083	31.12 $\pm$ 5	18	55.67 $\pm$ 17.75
Bungarus fasciatus	BUFA001	475	829.00000	41.97 $\pm$ 3.43	163	102.75 $\pm$ 9.88
Bungarus fasciatus	BUFA003	217	260.00000	28.89 $\pm$ 1.12	50	26.74 $\pm$ 4.95
Bungarus fasciatus	BUFA004	182	212.00000	28.11 $\pm$ 1.2	47	33.71 $\pm$ 5.58
Ophiophagus hannah	OPHA002	667	223.75116	8.06 $\pm$ 0.7	359	476.65 $\pm$ 44.41
Ophiophagus hannah	OPHA004	122	46.74653	9.27 $\pm$ 0.81	94	665.03 $\pm$ 103.34
Ophiophagus hannah	OPHA005	664	180.68454	6.54 $\pm$ 0.29	305	406.57 $\pm$ 40.48
Ophiophagus hannah	OPHA006	2173	723.05397	7.99 $\pm$ 0.17	553	71.88 $\pm$ 4.23
Ophiophagus hannah	OPHA007	969	320.66087	7.95 $\pm$ 0.29	225	72.24 $\pm$ 6.71
Ophiophagus hannah	OPHA010	432	430.01319	23.95 $\pm$ 15.56	83	52.16 $\pm$ 10.75
Ophiophagus hannah	OPHA013	1497	561.19097	9 $\pm$ 0.17	390	48.35 $\pm$ 3.31
Ophiophagus hannah	OPHA014	247	67.83958	6.62 $\pm$ 0.24	48	20.77 $\pm$ 3.68
Ophiophagus hannah	OPHA015	1944	680.12986	8.4 $\pm$ 0.18	595	71.05 $\pm$ 4.74
Ophiophagus hannah	OPHA017	2245	774.96944	8.29 $\pm$ 0.11	747	31.84 $\pm$ 1.96
Ophiophagus hannah	OPHA018	3122	1176.09583	9.04 $\pm$ 0.28	1018	75.72 $\pm$ 4.06
Ophiophagus hannah	OPHA019	890	311.79028	8.42 $\pm$ 0.21	231	43.22 $\pm$ 4.97
Ophiophagus hannah	OPHA024	466	173.17292	8.94 $\pm$ 0.27	217	100.56 $\pm$ 9.13
Ophiophagus hannah	OPHA025	435	164.98403	9.12 $\pm$ 0.36	203	69.62 $\pm$ 7.79
Ophiophagus hannah	OPHA026	134	68.72847	12.4 $\pm$ 0.72	68	107.46 $\pm$ 20.53
Ophiophagus hannah	OPHA027	542	198.62500	8.81 $\pm$ 0.23	126	22.96 $\pm$ 2.98
Ophiophagus hannah	OPHA034	317	124.02222	9.42 $\pm$ 0.42	52	37.75 $\pm$ 7.35
Ophiophagus hannah	OPHA054	1381	520.62361	9.05 $\pm$ 0.16	383	57.41 $\pm$ 5.34
Ophiophagus hannah	OPHA055	722	353.61944	11.77 $\pm$ 0.4	247	73.12 $\pm$ 6.93
Ophiophagus hannah	OPHA056	428	162.07222	9.11 $\pm$ 0.28	140	48.96 $\pm$ 5.6
Ophiophagus hannah	OPHA058	672	288.89514	9.62 $\pm$ 0.25	254	53.71 $\pm$ 5.58
Ophiophagus hannah	OPHA059	995	345.03819	8.33 $\pm$ 0.17	213	39.42 $\pm$ 4.28
Ophiophagus hannah	OPHA086	623	241.01528	9.3 $\pm$ 0.28	196	46.65 $\pm$ 4.59
Ophiophagus hannah	OPHA096	291	132.91667	11 $\pm$ 0.41	55	50.99 $\pm$ 8.88
Ophiophagus hannah	OPHA099	276	122.01528	10.65 $\pm$ 0.37	30	46.22 $\pm$ 13.52
Python bivittatus	PYBI021	631	662.03125	25.22 $\pm$ 0.33	216	31.79 $\pm$ 3.05
Python bivittatus	PYBI022	438	485.89722	26.69 $\pm$ 1.11	154	29.44 $\pm$ 2.98
Python bivittatus	PYBI028	176	187.80417	25.76 $\pm$ 0.78	69	70.05 $\pm$ 12.29
Python bivittatus	PYBI029	486	514.91667	25.48 $\pm$ 0.83	237	56.23 $\pm$ 4.89
Python bivittatus	PYBI033	41	41.07639	24.65 $\pm$ 0.6	22	36.38 $\pm$ 7.88
Python bivittatus	PYBI055	205	207.12639	24.37 $\pm$ 0.33	26	15.67 $\pm$ 5.23
Python bivittatus	PYBI060	191	190.93819	24.12 $\pm$ 0.25	63	65.09 $\pm$ 9.91

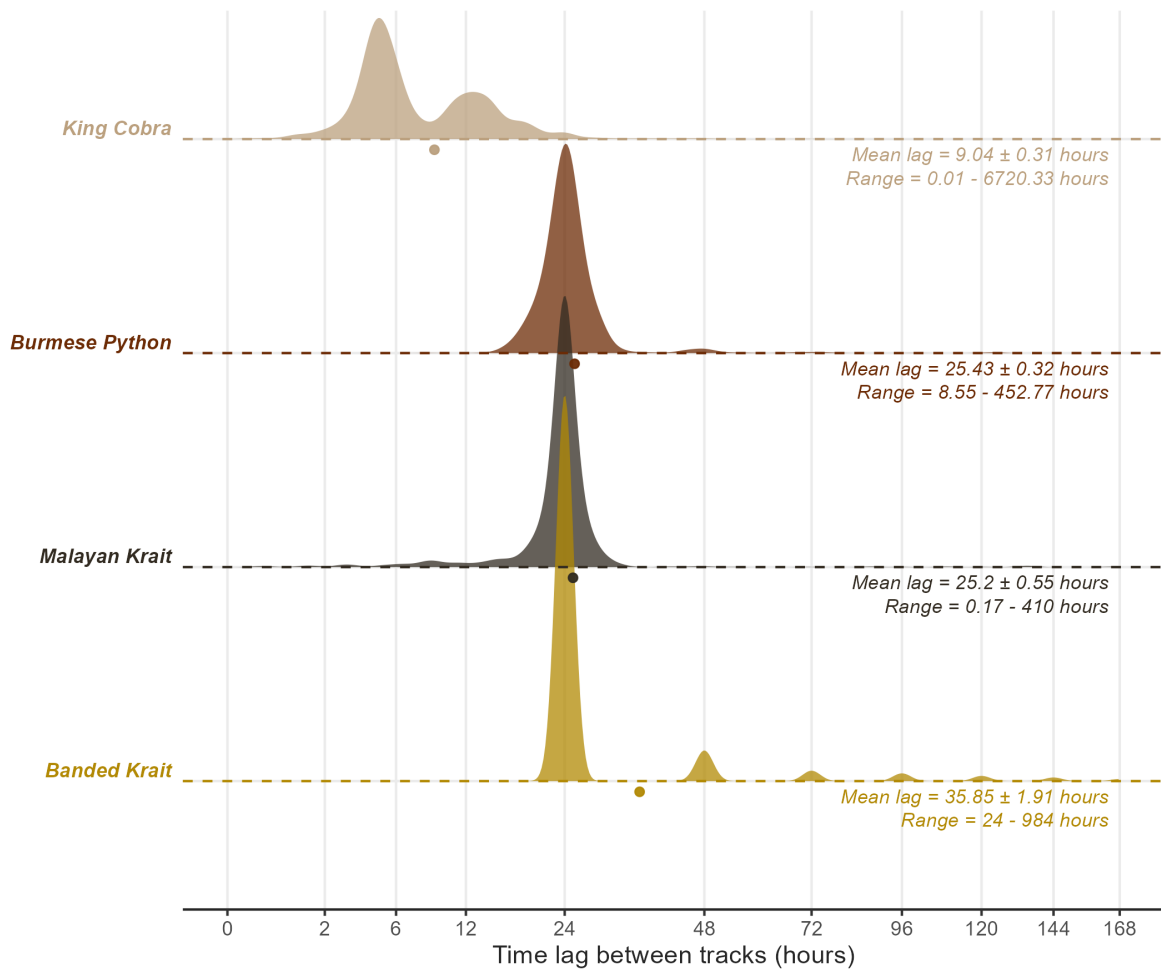


Figure 4.11: The distribution between subsequent data points for each species. Circles below the distributions show the median time lag. Reported alongside is the mean, standard error ($\pm$), and range of time lags. Large time lags are usually the result of prolonged missing periods or transmitter failures. N.b., the x axis is square-rooted to aid with visualisation.

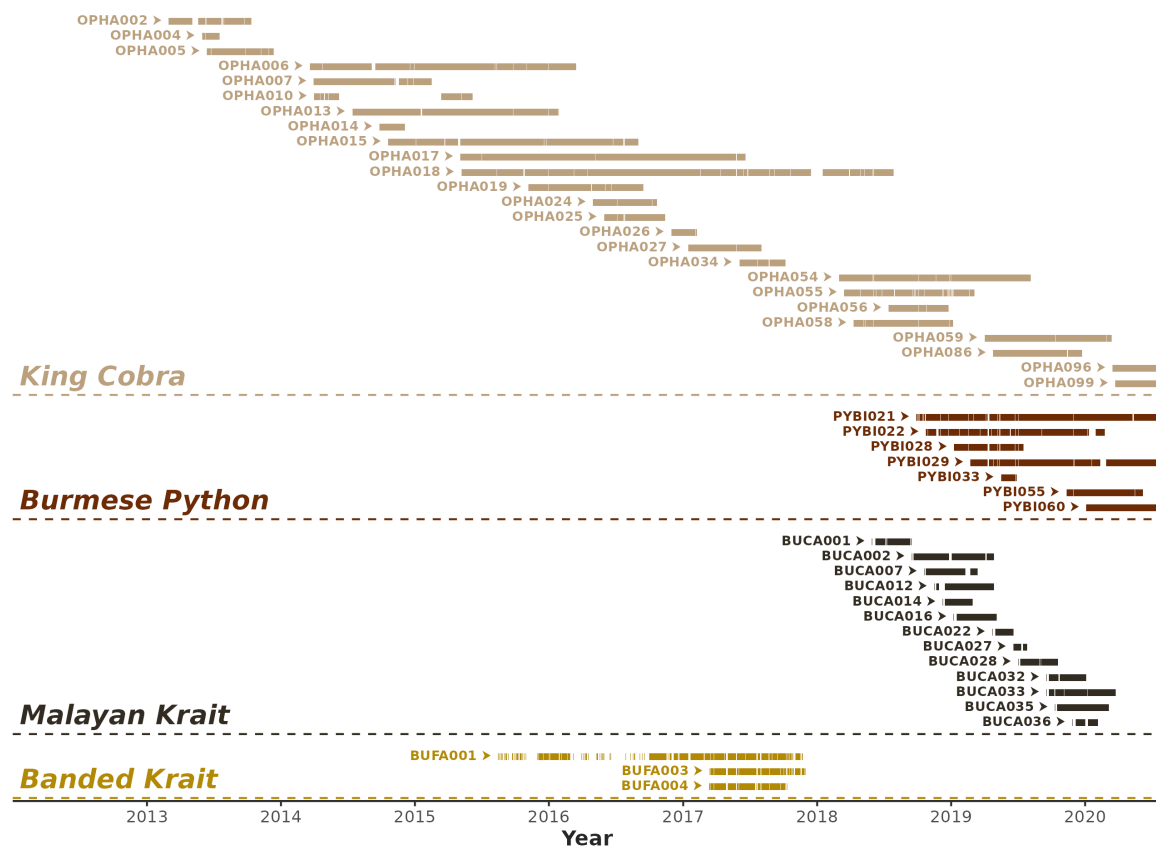


Figure 4.12: The timeline of all tracked individuals, split by species. Each recorded data point is depicted by a |.

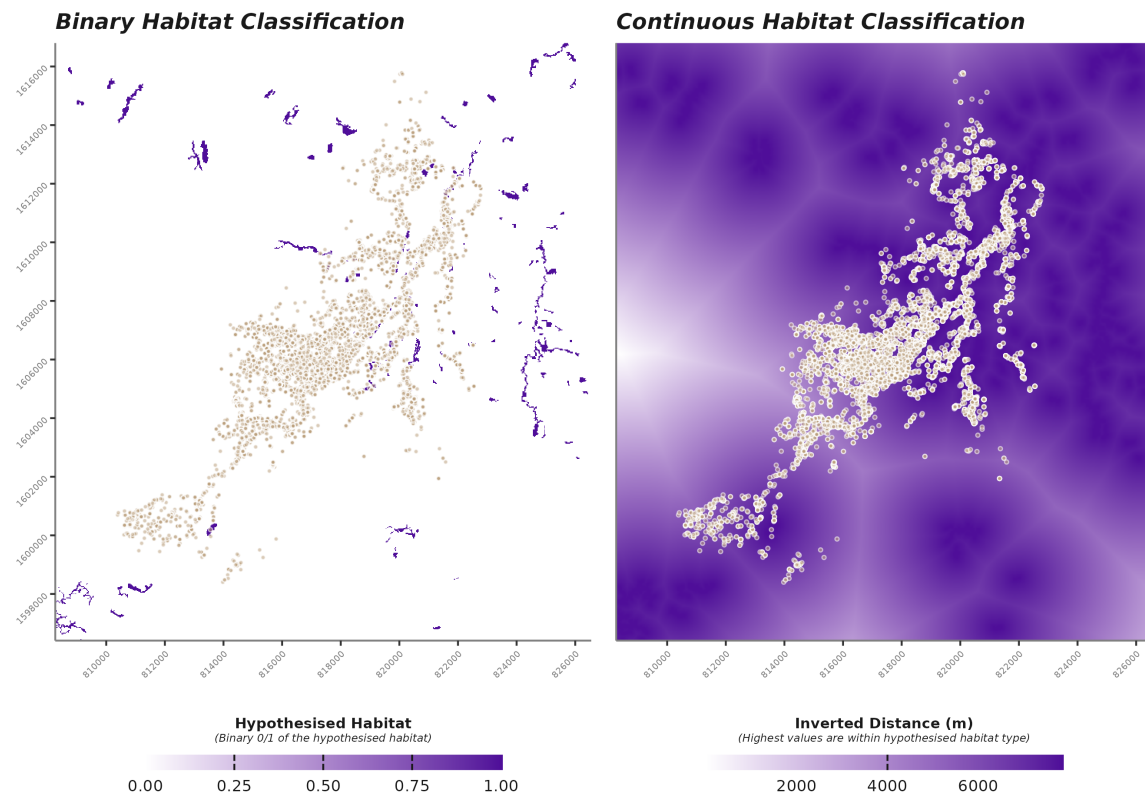
King Cobra (*Ophiophagus hannah*)

Figure 4.13: The landscape covariates used in the King Cobra models for hypothesis 1, overlaid with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat.

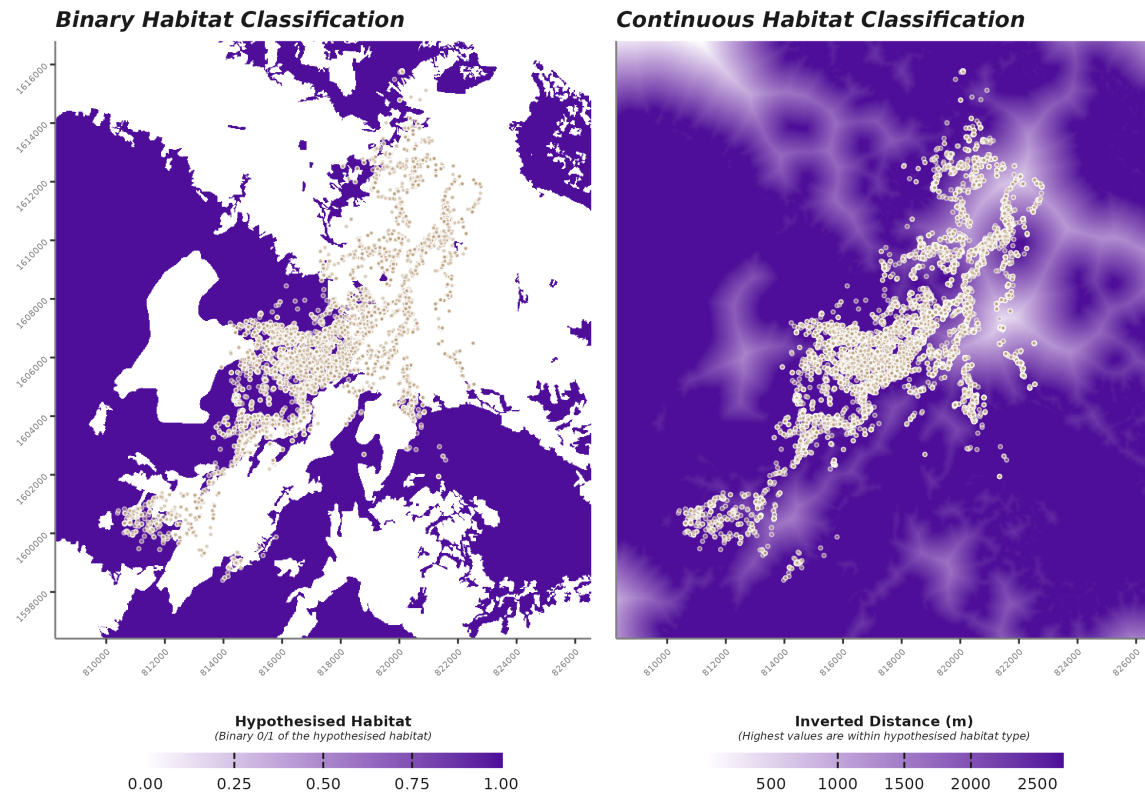
King Cobra (*Ophiophagus hannah*)

Figure 4.14: The landscape covariates used in the King Cobra models for hypothesis 2, overlaid with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat.

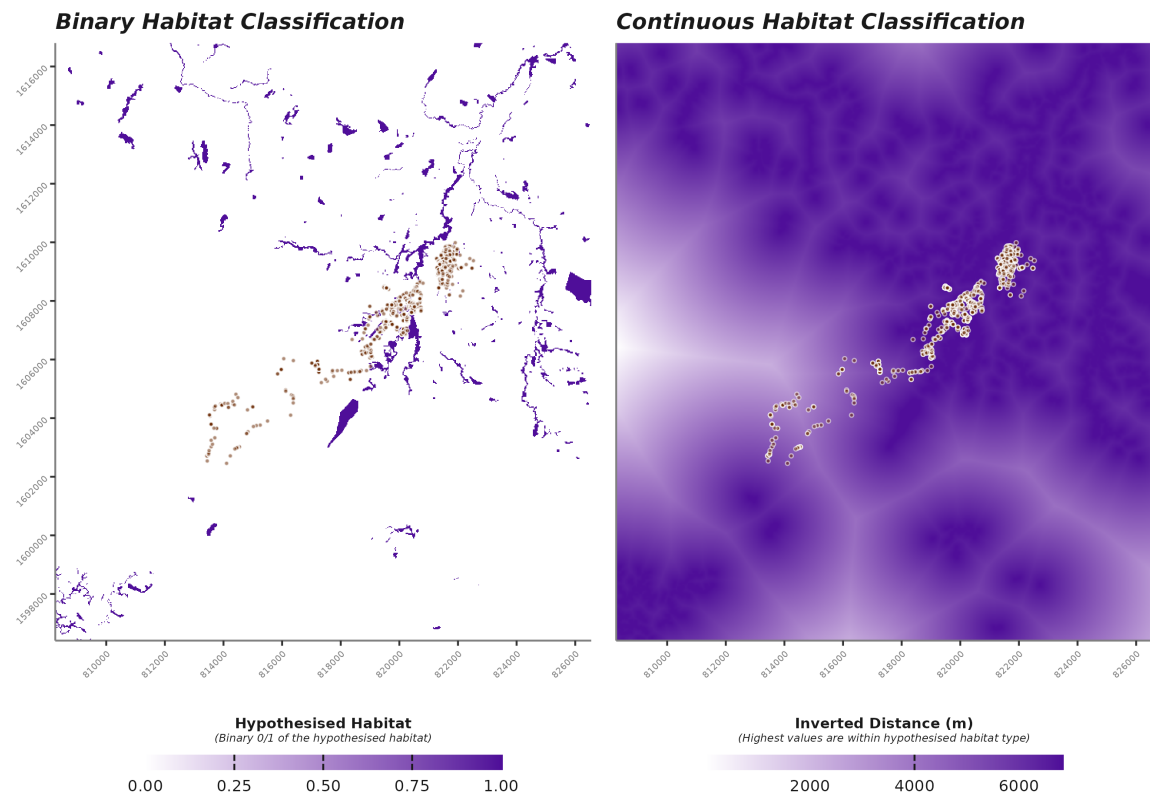
Burmese Python (*Python bivittatus*)

Figure 4.15: The landscape covariates used in the Burmese Python models for hypothesis 1, overlayed with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat.

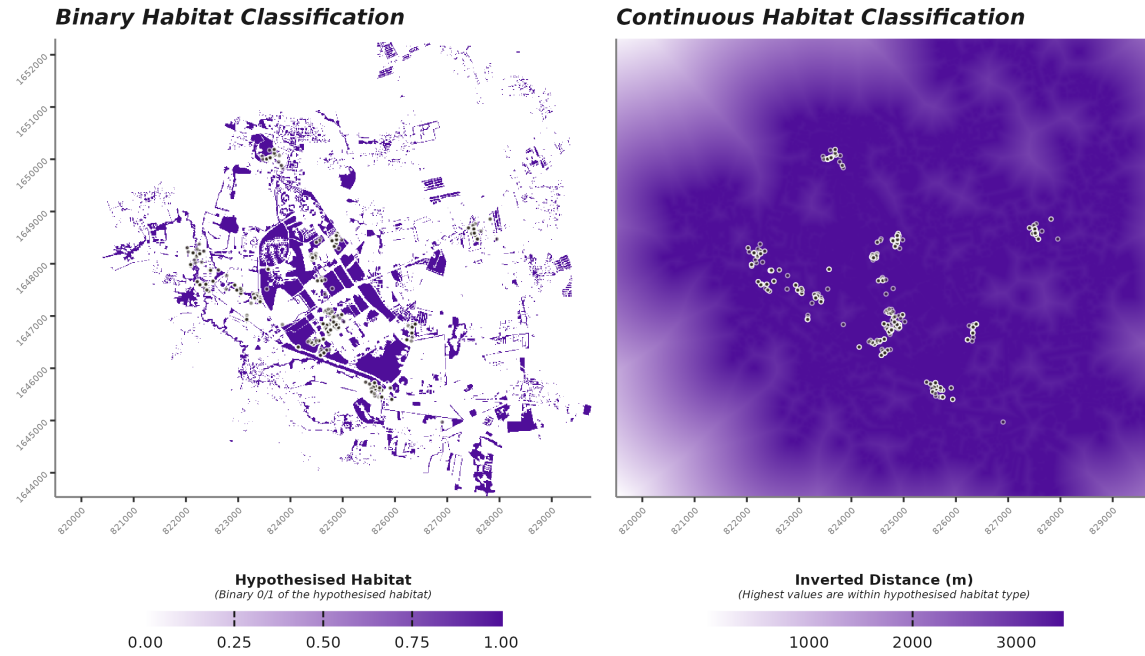
Malayan Krait (*Bungarus candidus*)

Figure 4.16: The landscape covariates used in the Malayan Krait models for hypothesis 1, overlaid with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat.

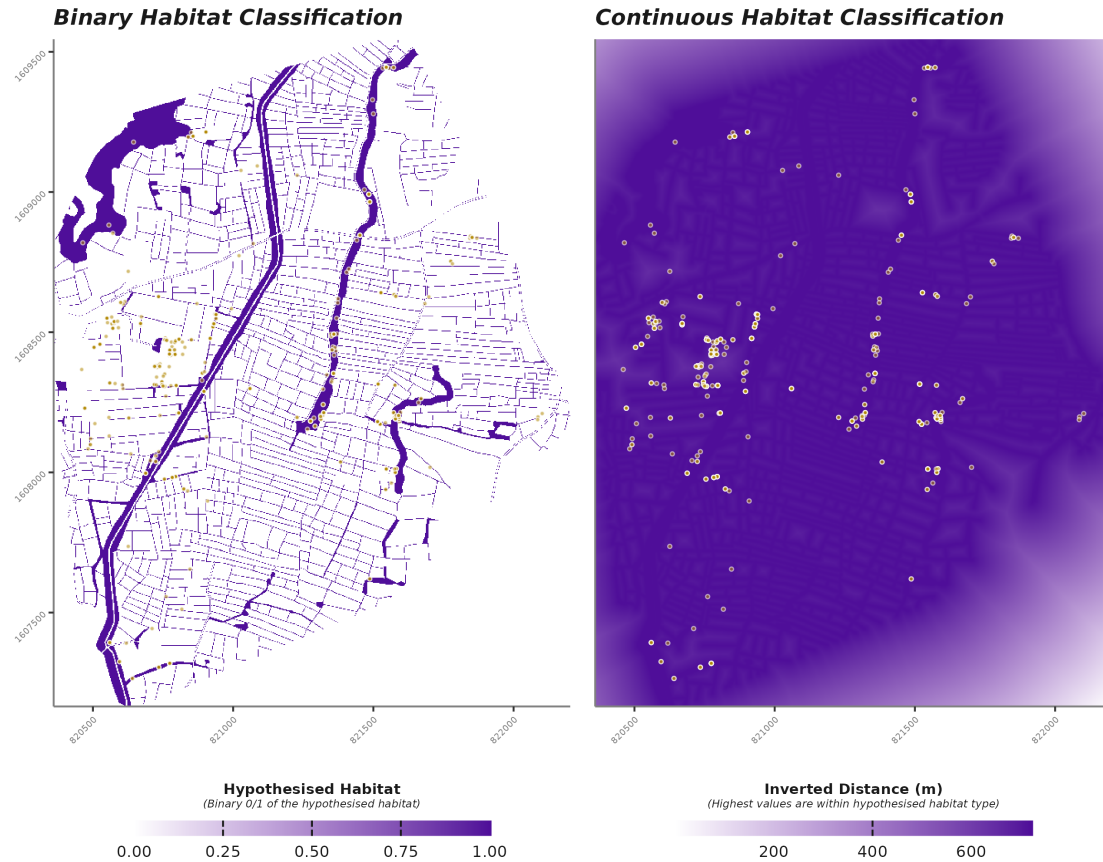
Banded Krait (*Bungarus fasciatus*)

Figure 4.17: The landscape covariates used in the Banded Krait models for hypothesis 1, overlayed with the recorded locations of the animals. Left hand: the binary classification where hypothesised habitats are classed as 1, whereas all other areas are classed as 0. Right hand: the inverted continuous landscape classification where higher values indicated proximity to a given land use or habitat.

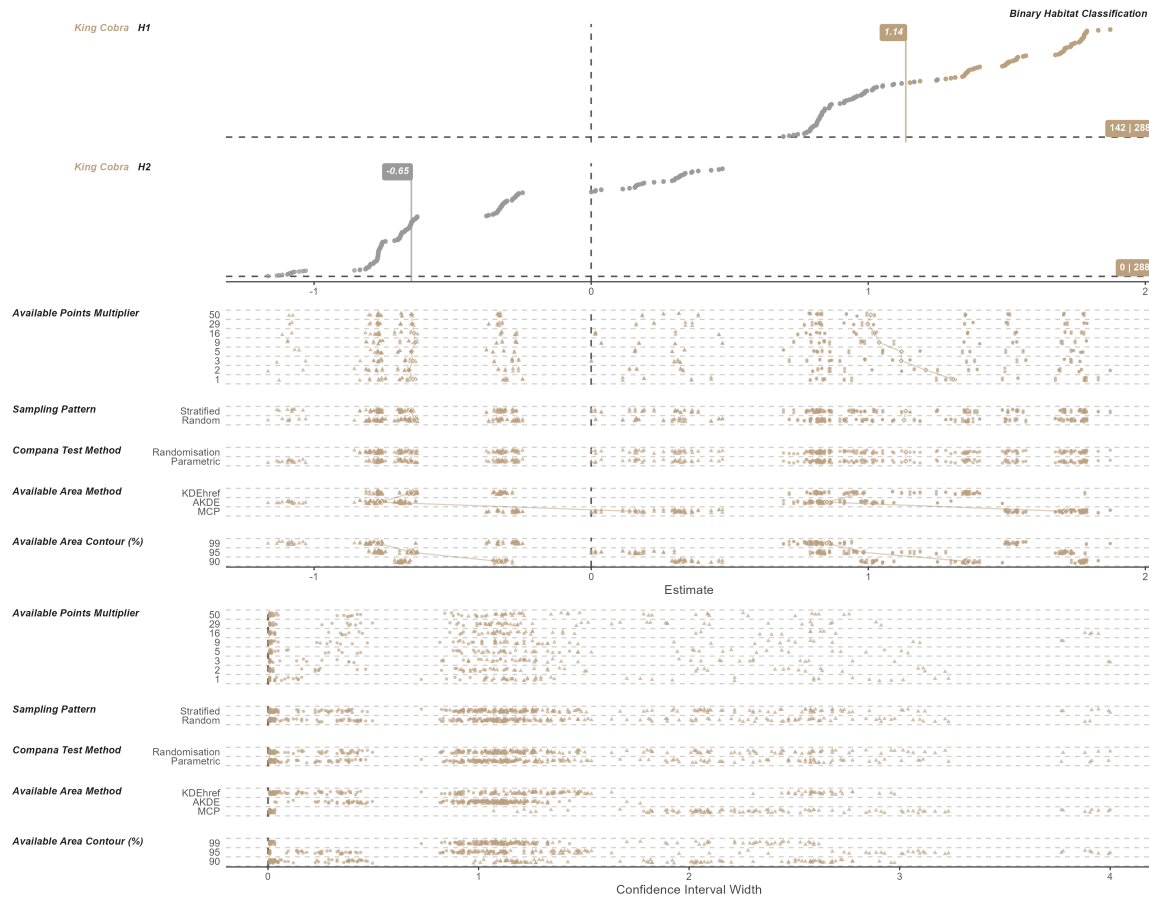


Figure 4.18: All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. Shape separates hypothesis 1 and 2 (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and hypothesis medians are connected with appropriated coloured lines.

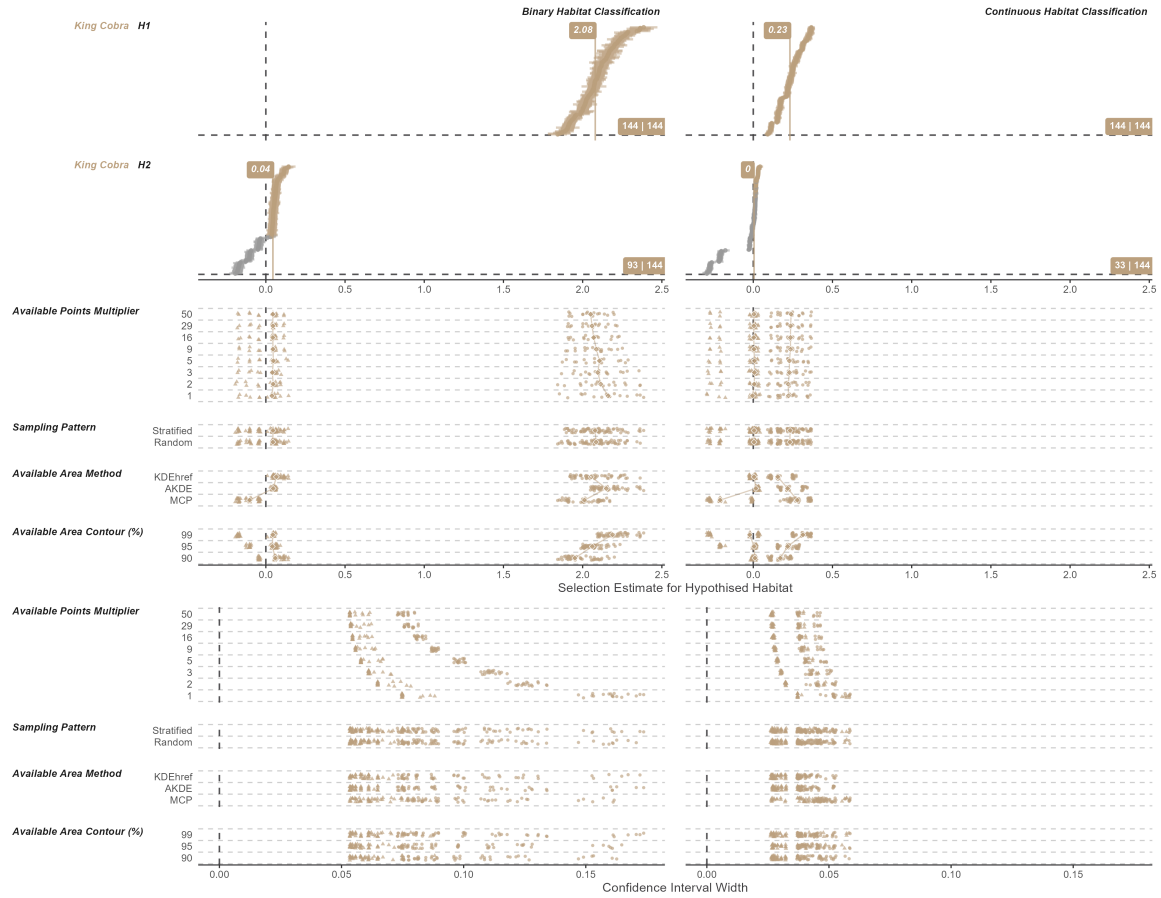


Figure 4.19: All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

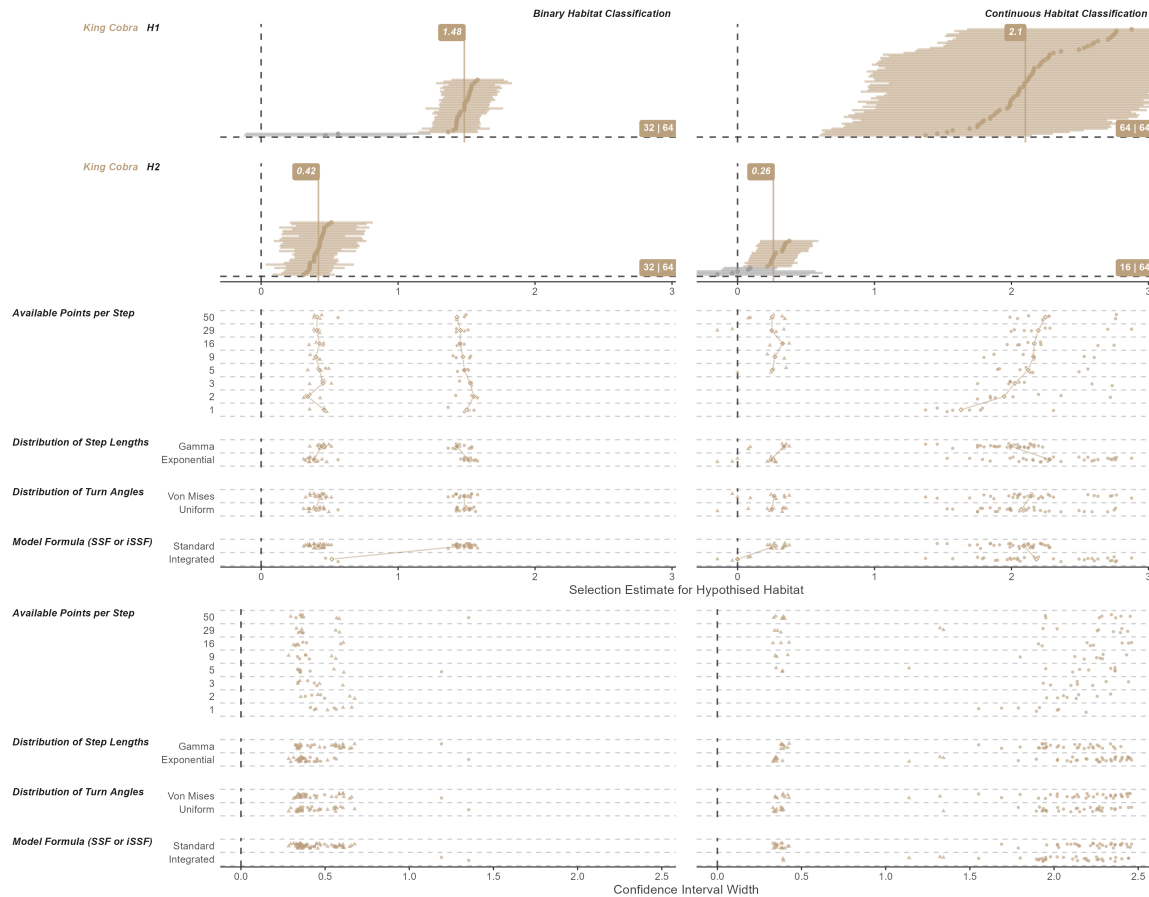


Figure 4.20: All estimates of habitat selection derived from the step-based Two-Step analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates, split by hypothesis, with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate for each hypothesis. Lower plot shows the estimates relative to each analysis choice. Shape separates hypothesis 1 and 2 (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

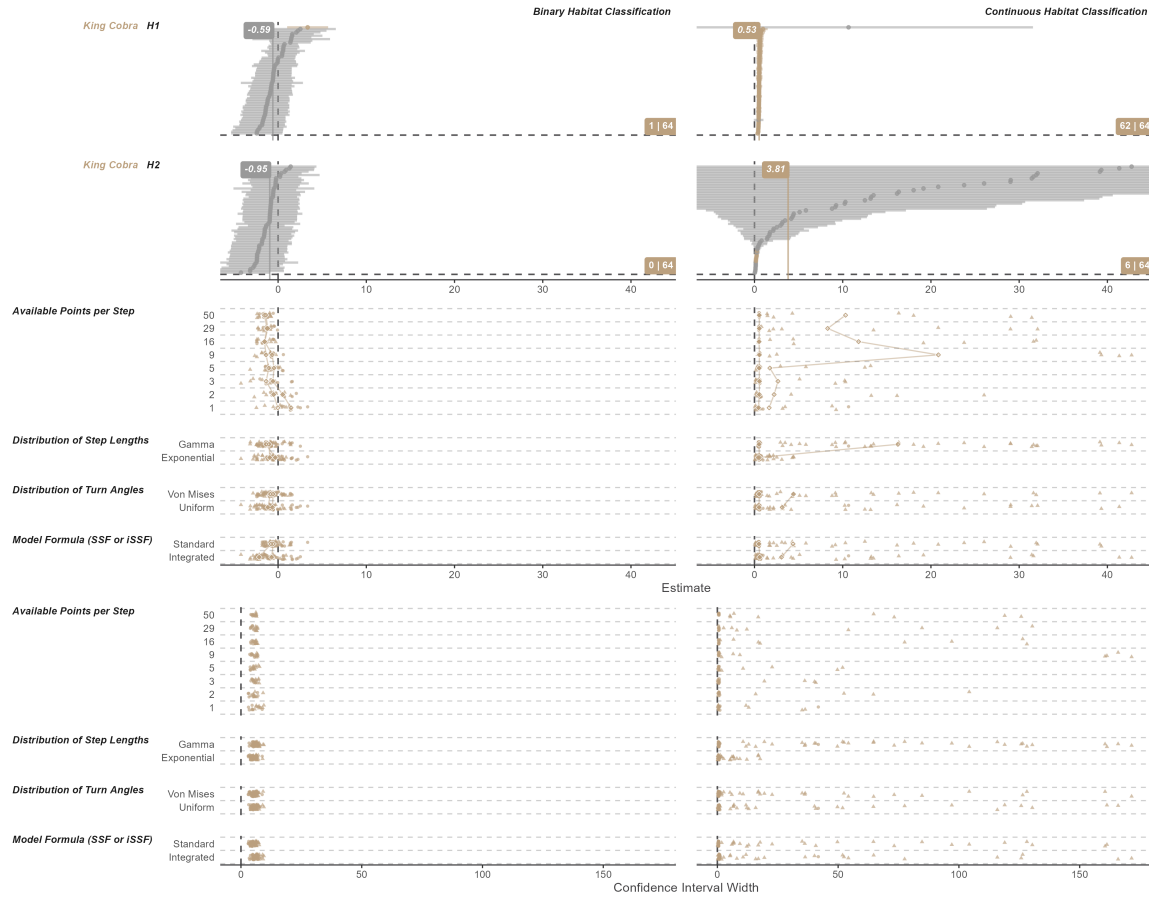


Figure 4.21: All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates, split by hypothesis, with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate for each hypothesis. Lower plot shows the estimates relative to each analysis choice. Shape separates hypothesis 1 and 2 (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

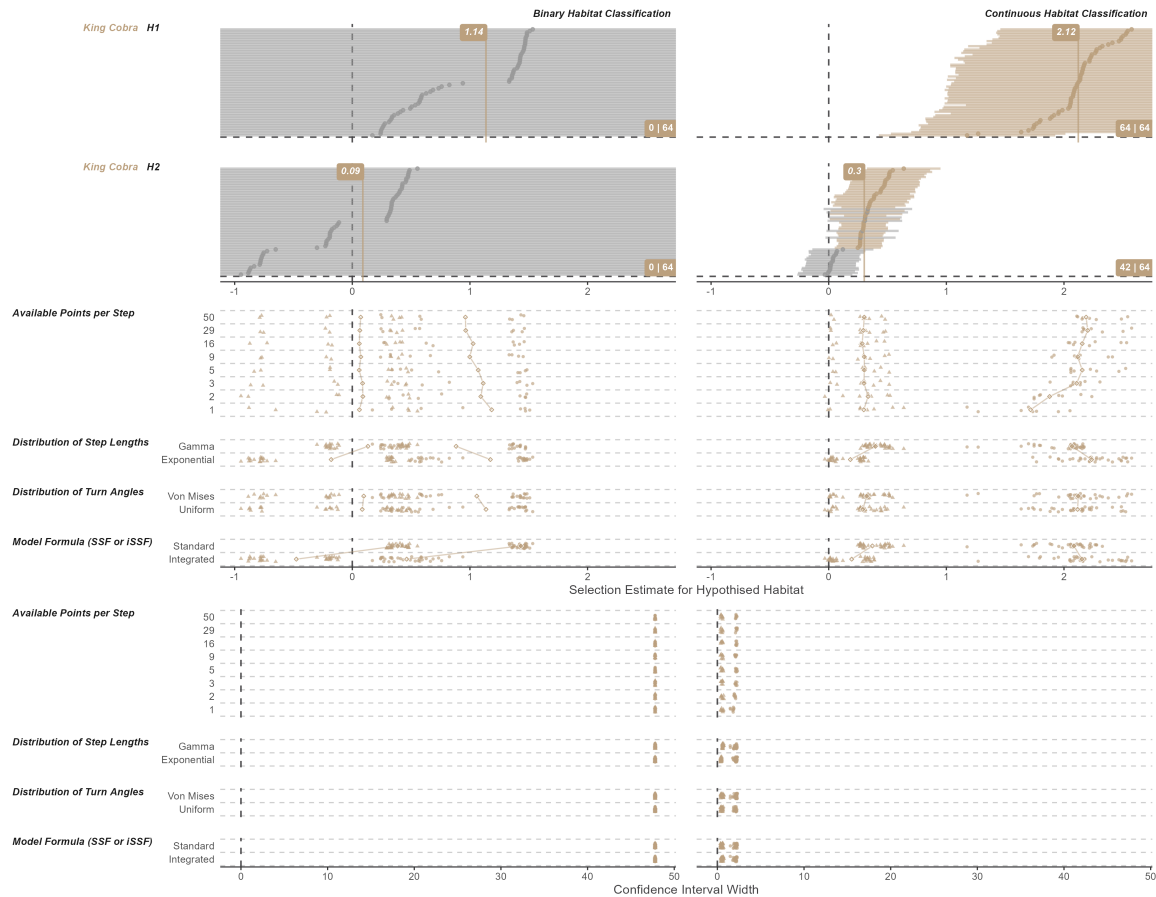


Figure 4.22: All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis for King Cobras (*Ophiophagus hannah*). Top curves show all estimates, split by hypothesis, with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate for each hypothesis. Lower plot shows the estimates relative to each analysis choice. Shape separates hypothesis 1 and 2 (circles = hypothesis 1, triangle = hypothesis 2). Median estimates are shown with hollow diamonds, and hypothesis medians are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

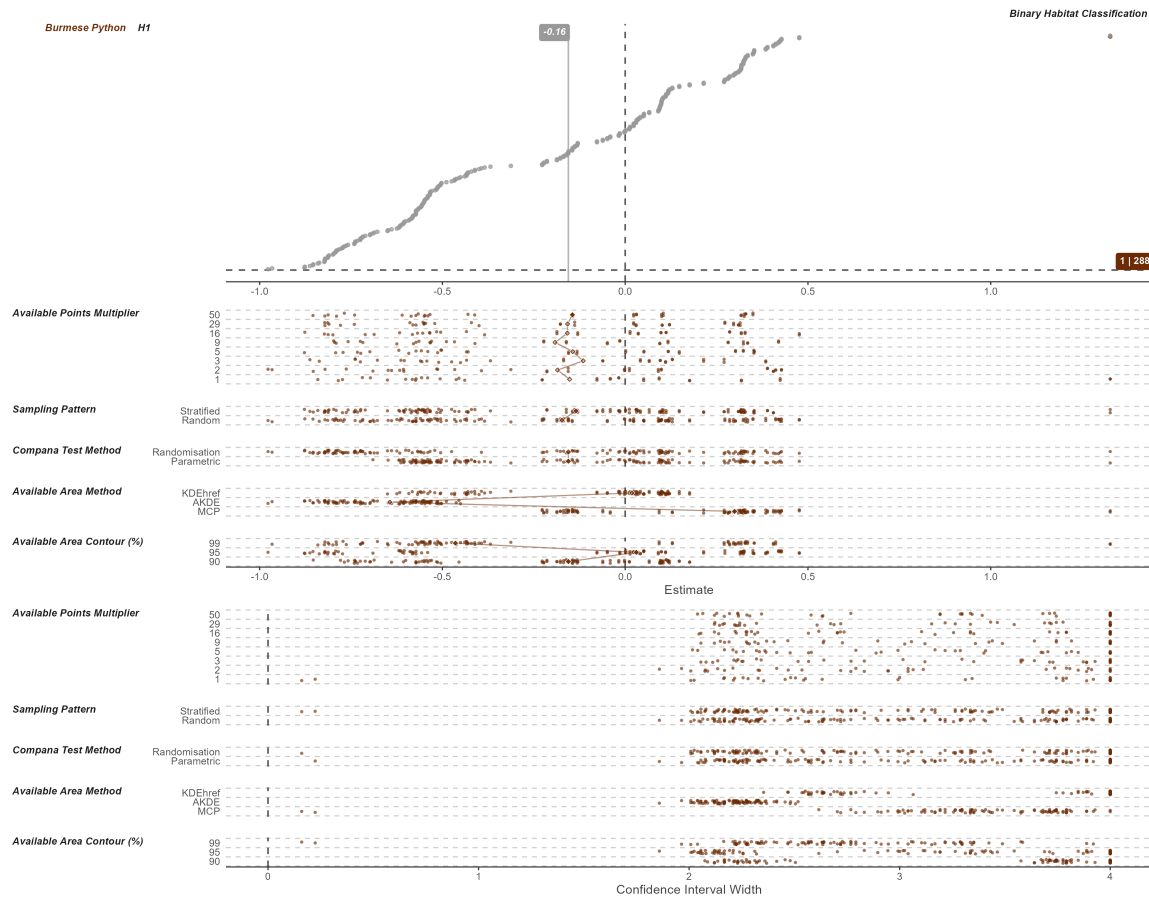


Figure 4.23: All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines.

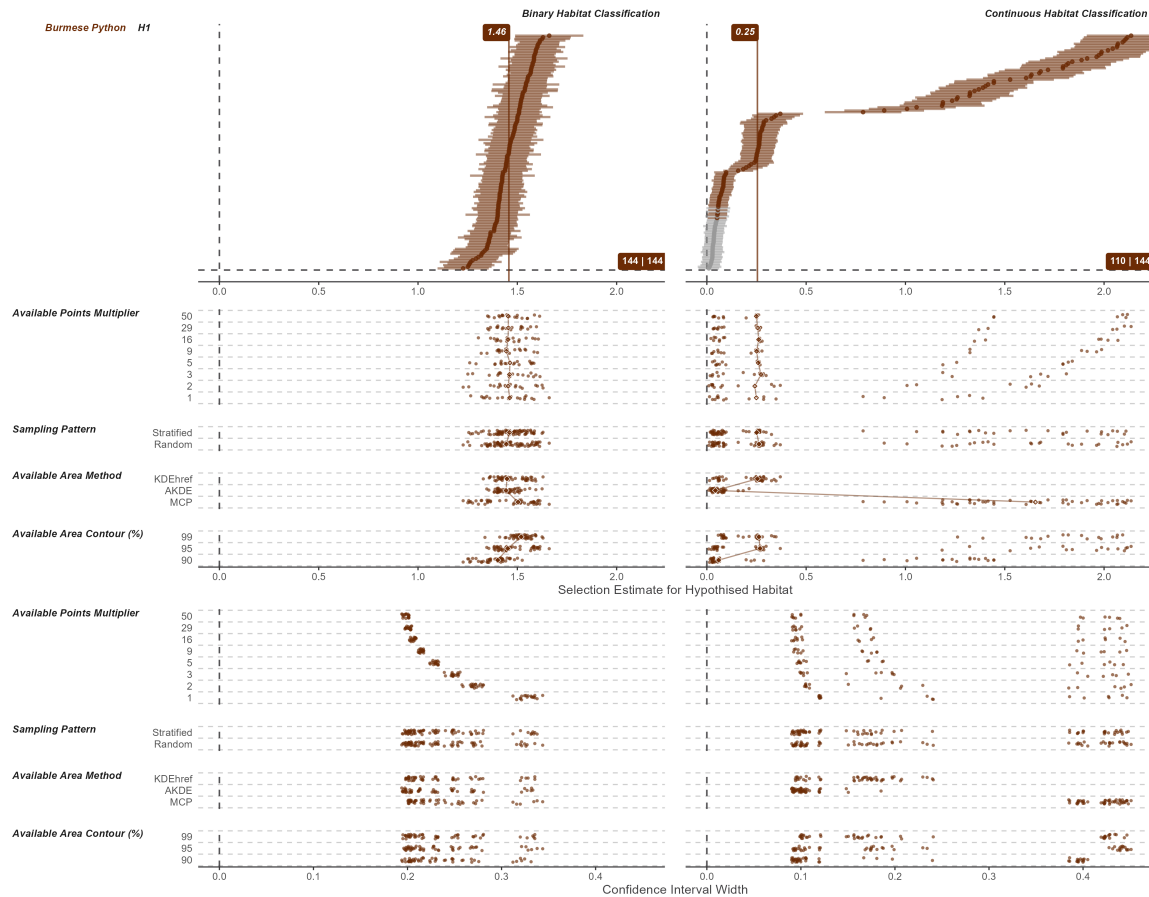


Figure 4.24: All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

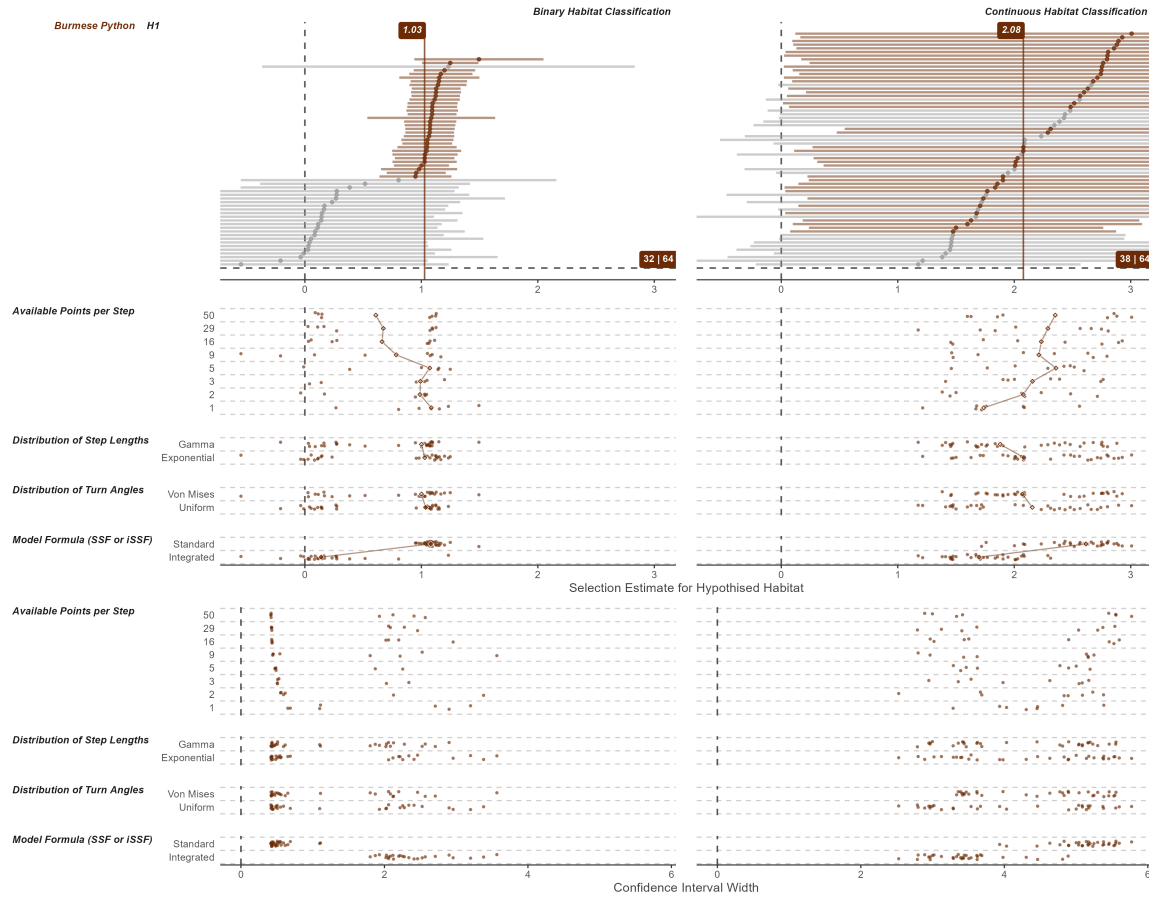


Figure 4.25: All estimates of habitat selection derived from the step-based Two-Step analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

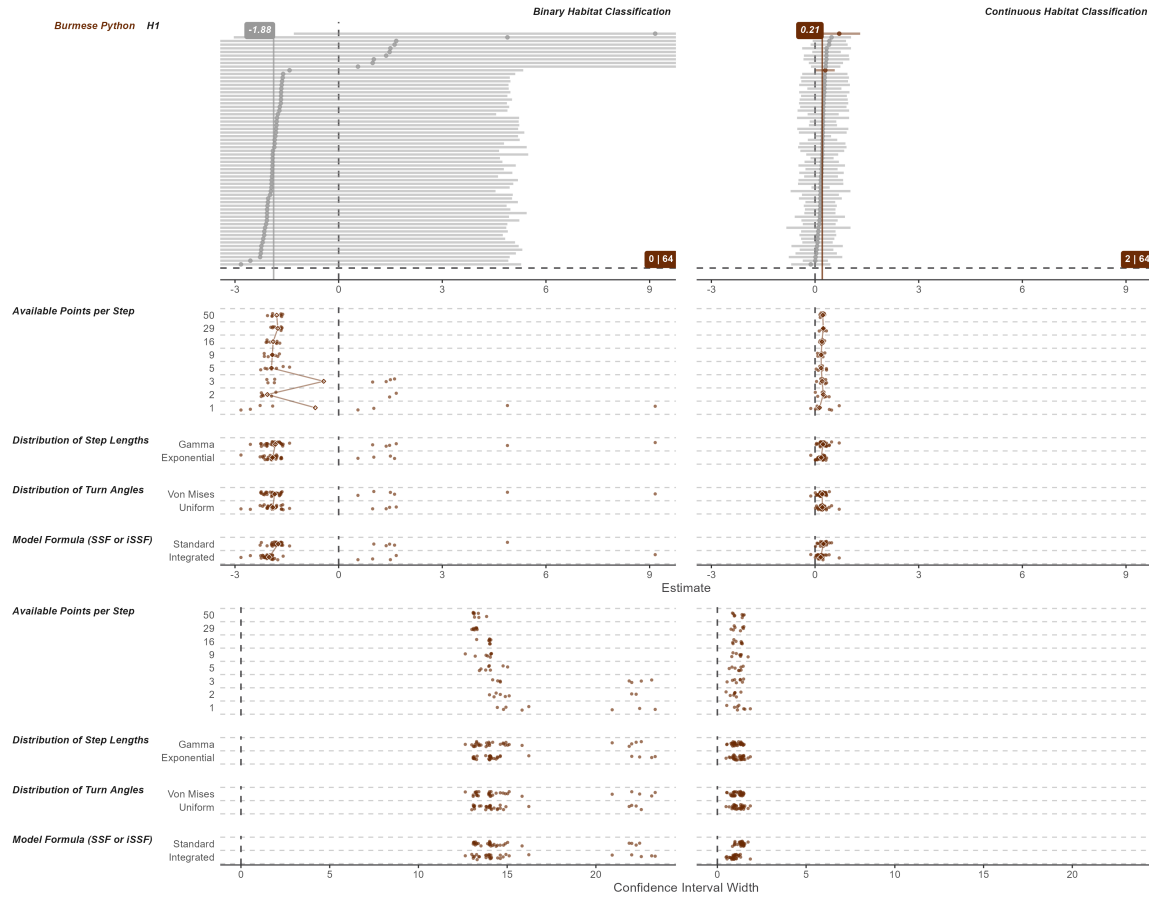


Figure 4.26: All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

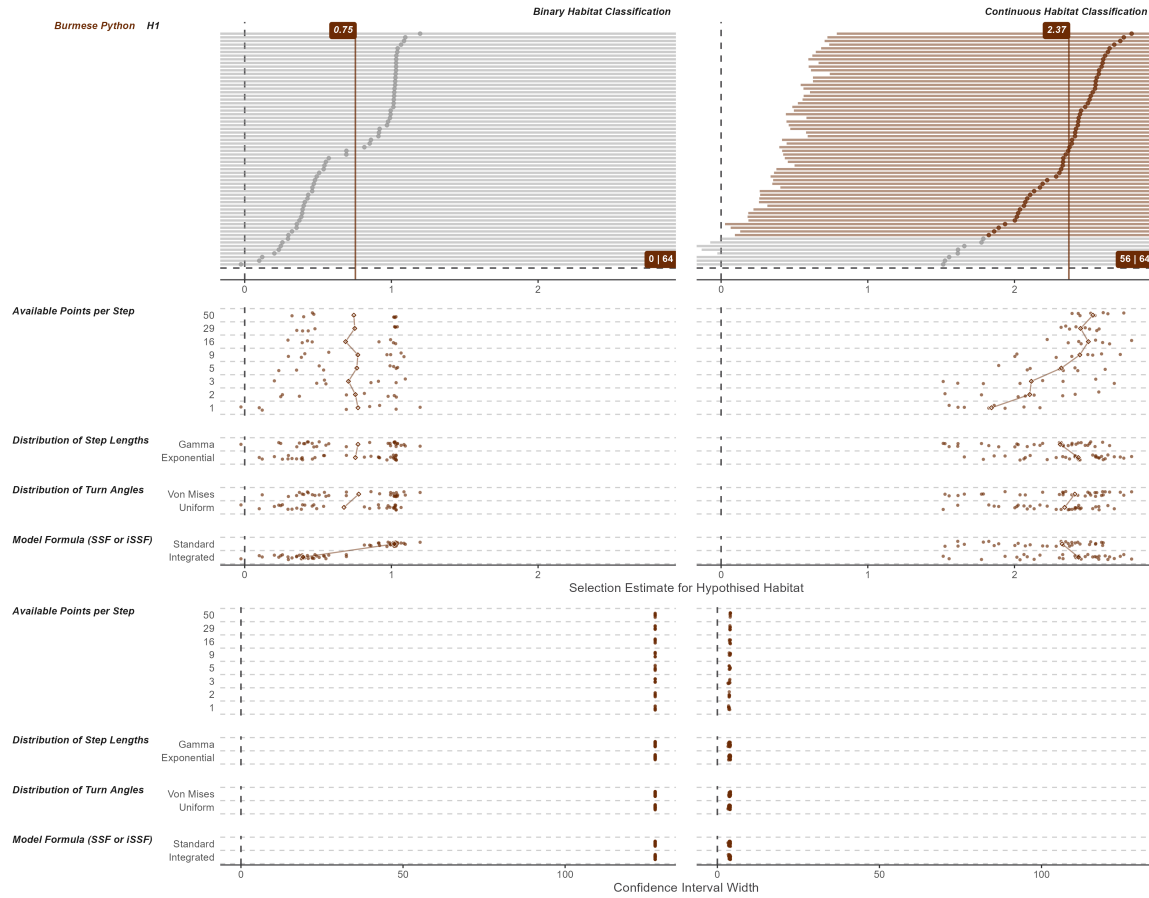


Figure 4.27: All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis for Burmese Pythons (*Python bivittatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

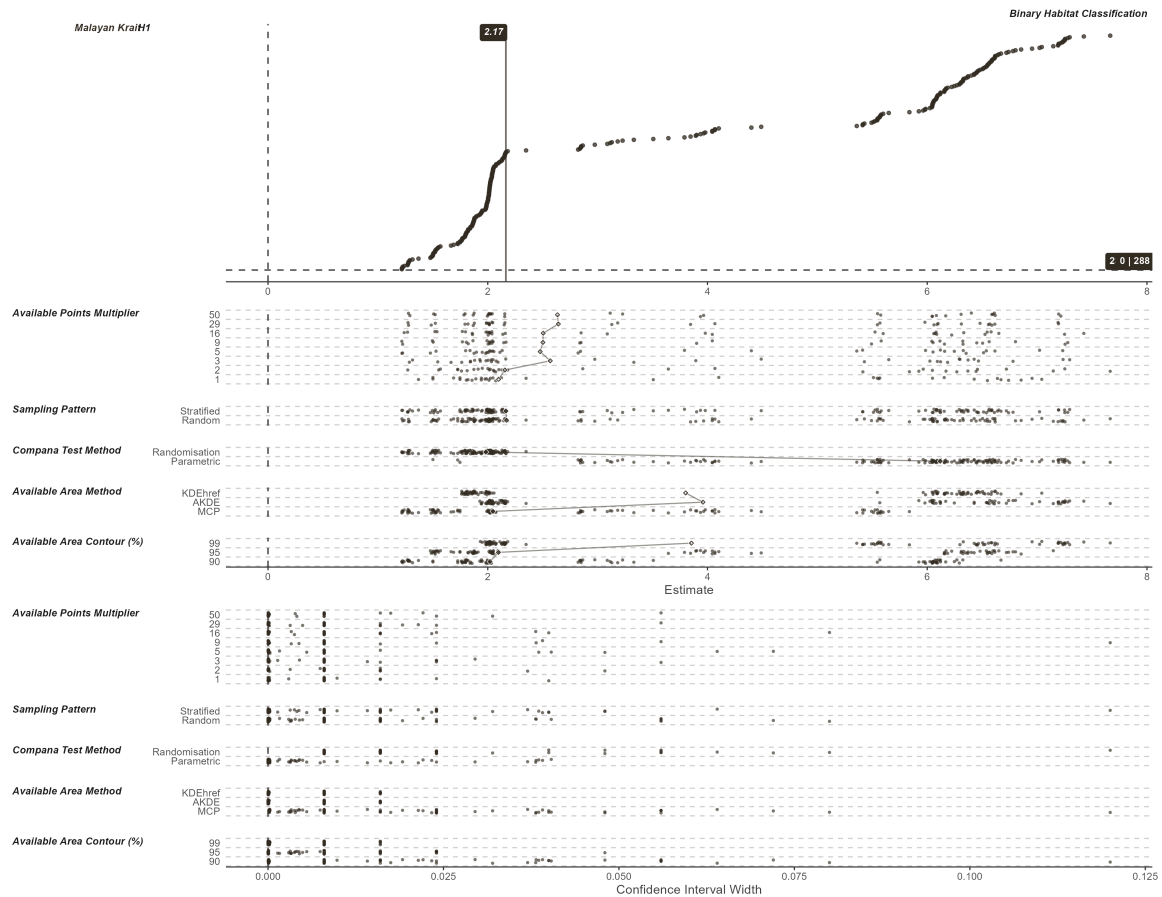


Figure 4.28: All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines.

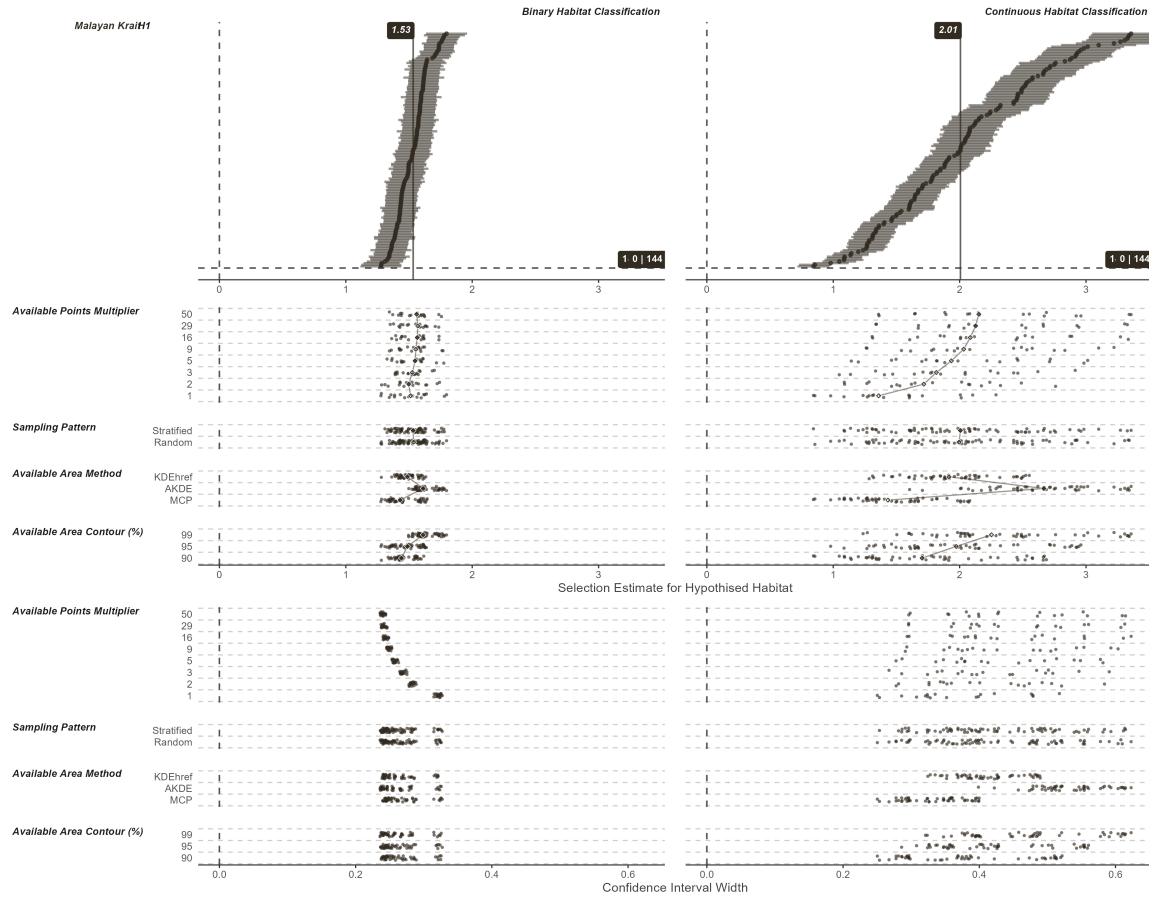


Figure 4.29: All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

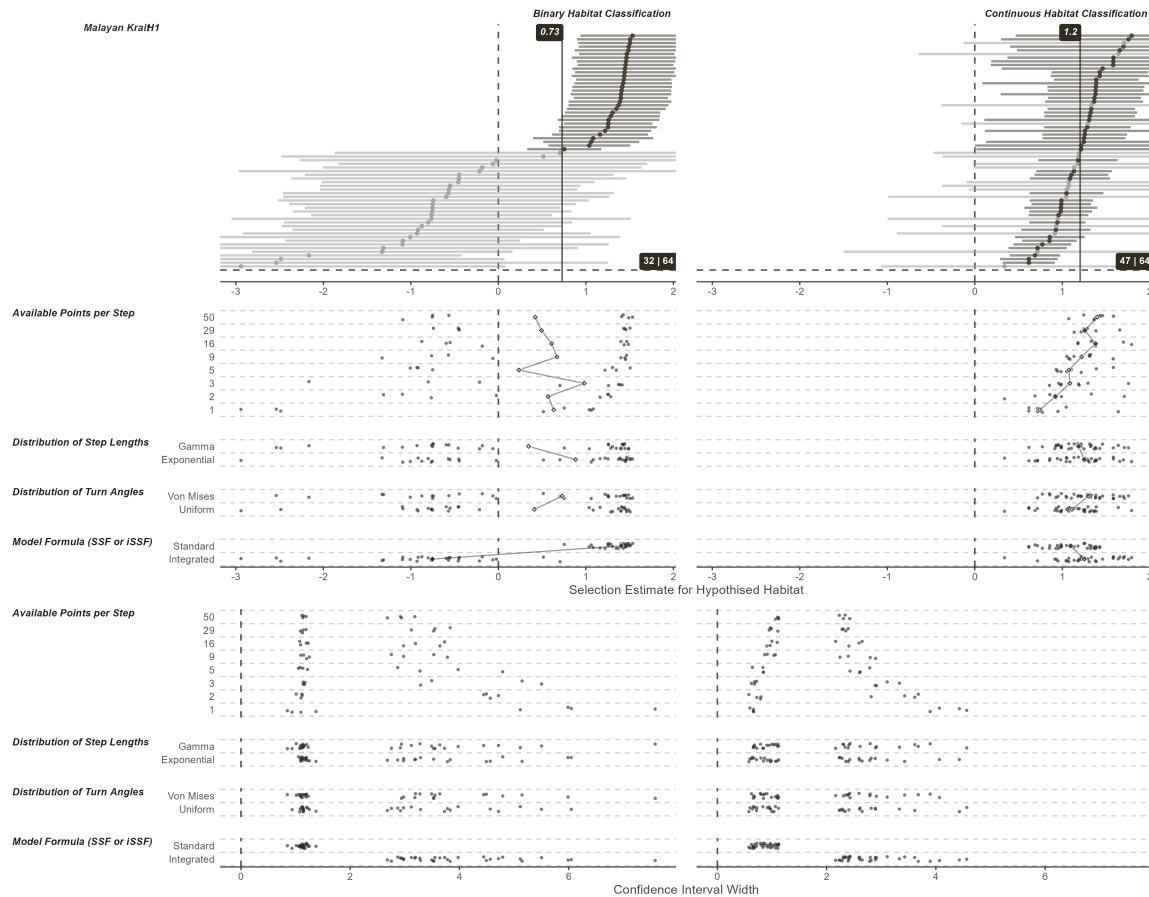


Figure 4.30: All estimates of habitat selection derived from the step-based Two-Step analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

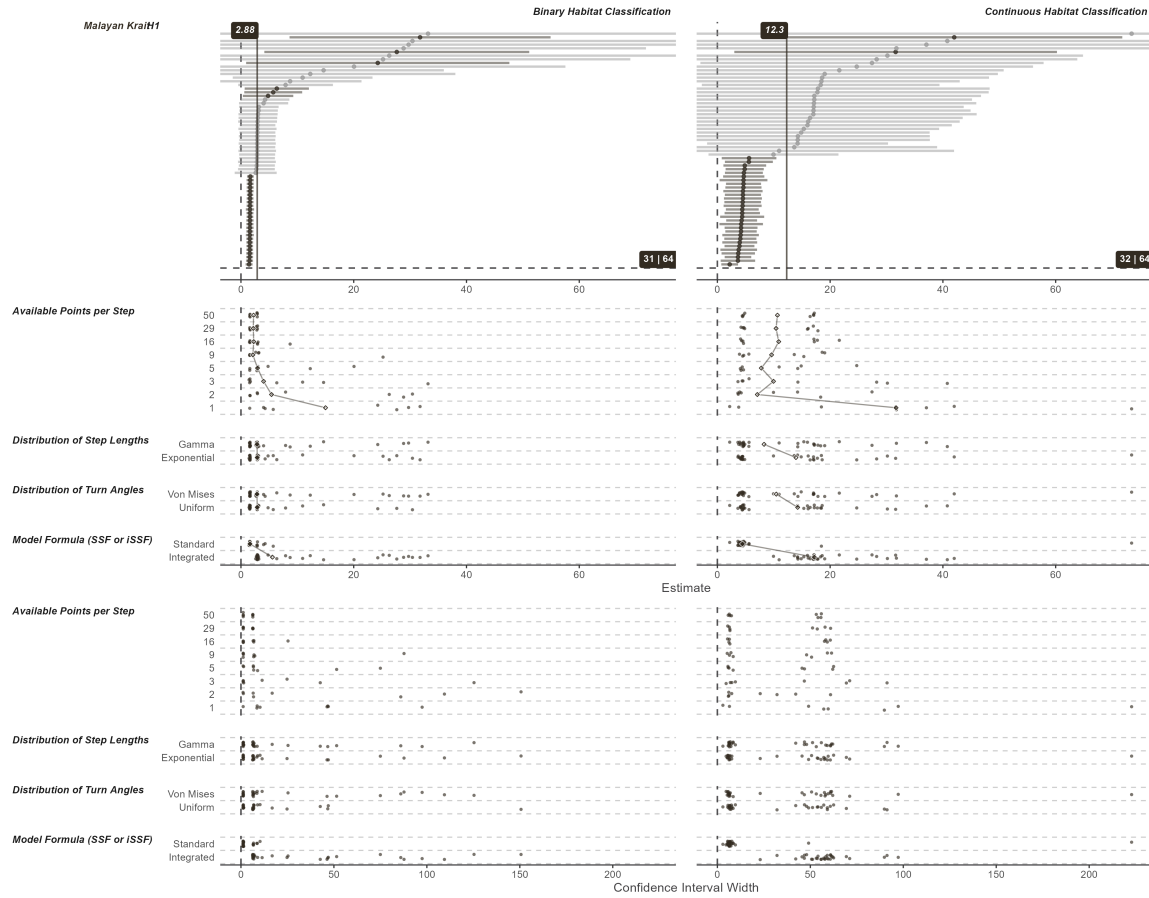


Figure 4.31: All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis for Malaysian Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

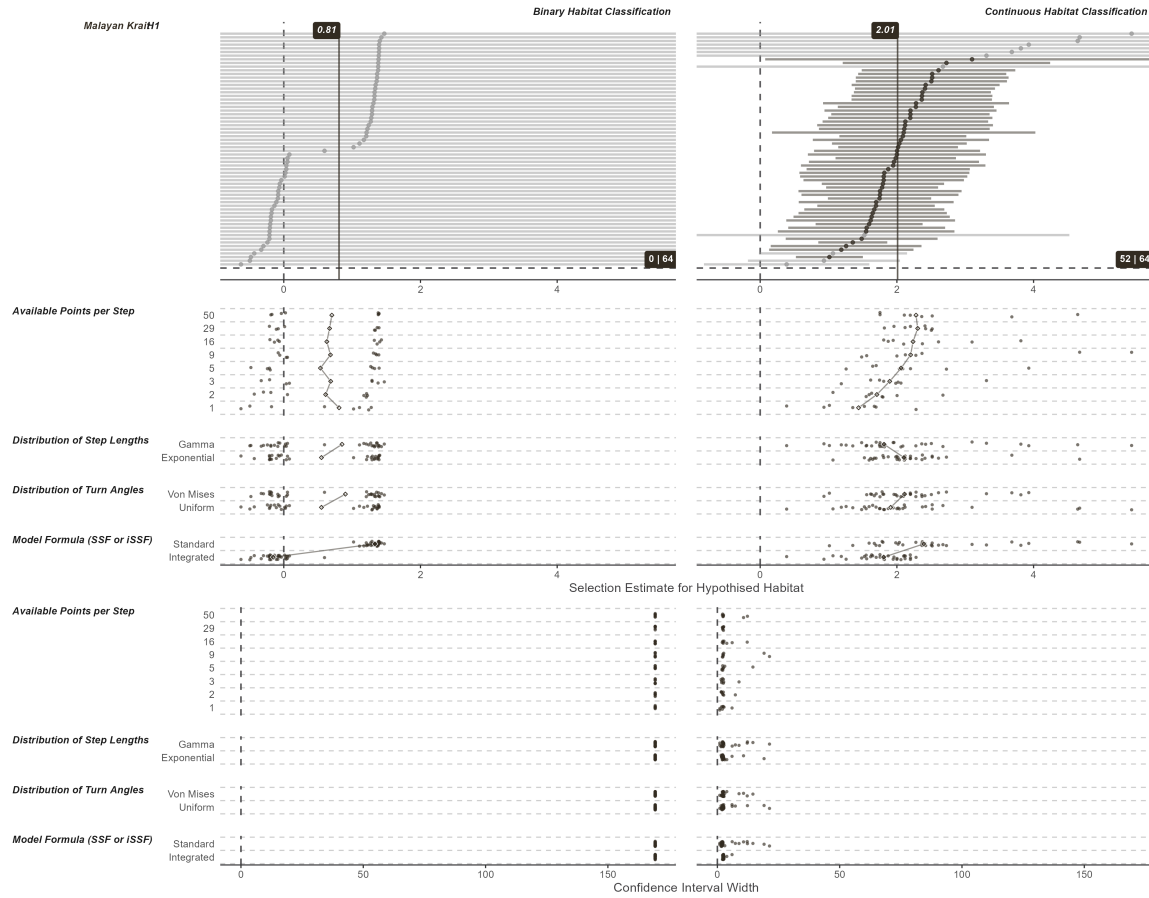


Figure 4.32: All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis for Malayan Kraits (*Bungarus candidus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

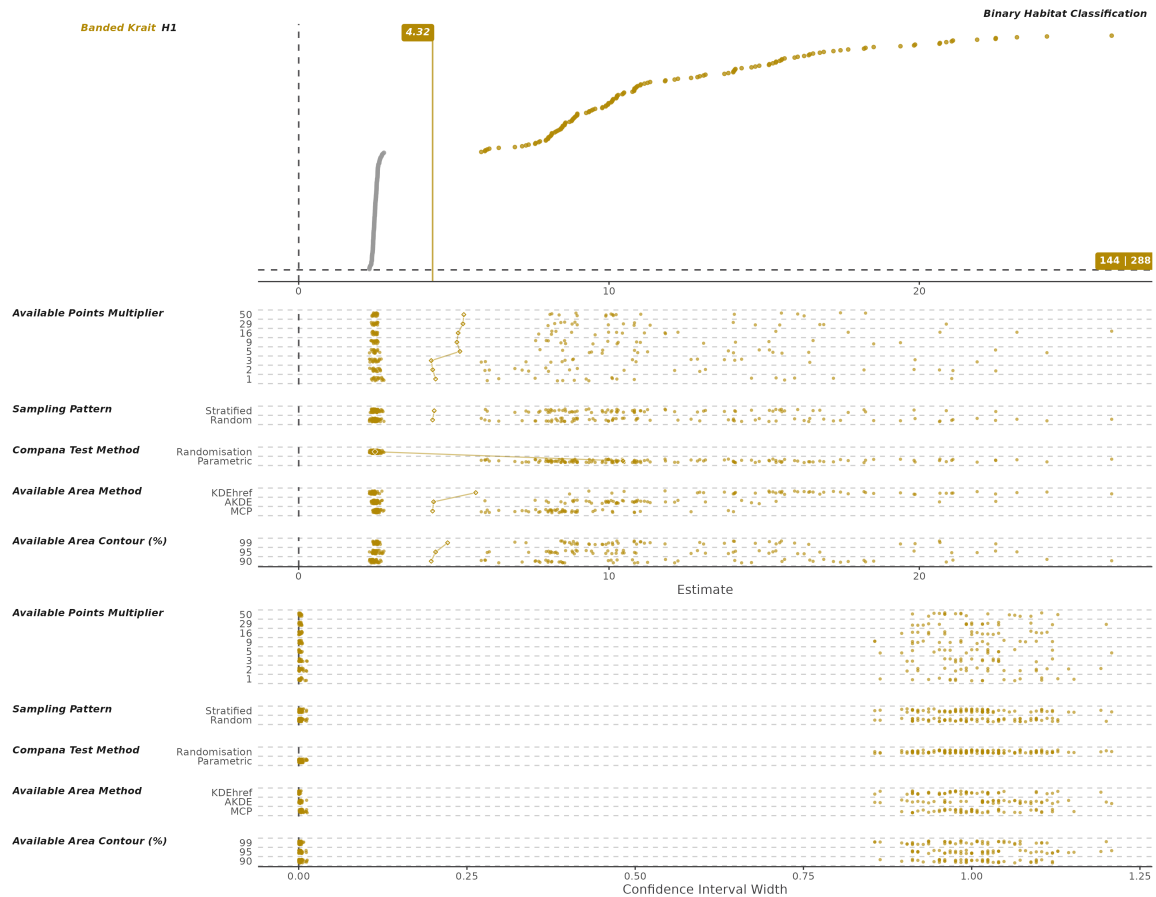


Figure 4.33: All estimates of habitat selection derived from the areas-based Compositional (Compana) analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines.

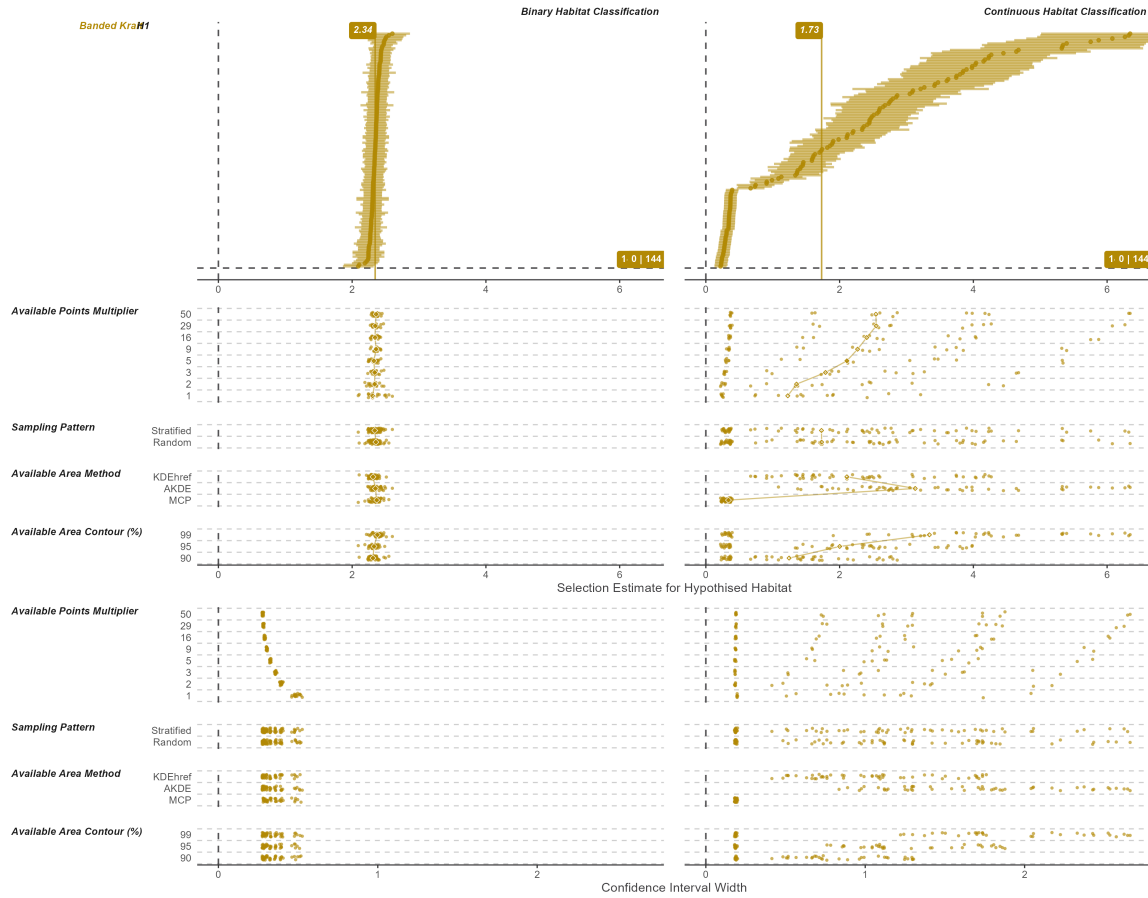


Figure 4.34: All estimates of habitat selection derived from the areas-based Resource Selection Function (RSF) analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

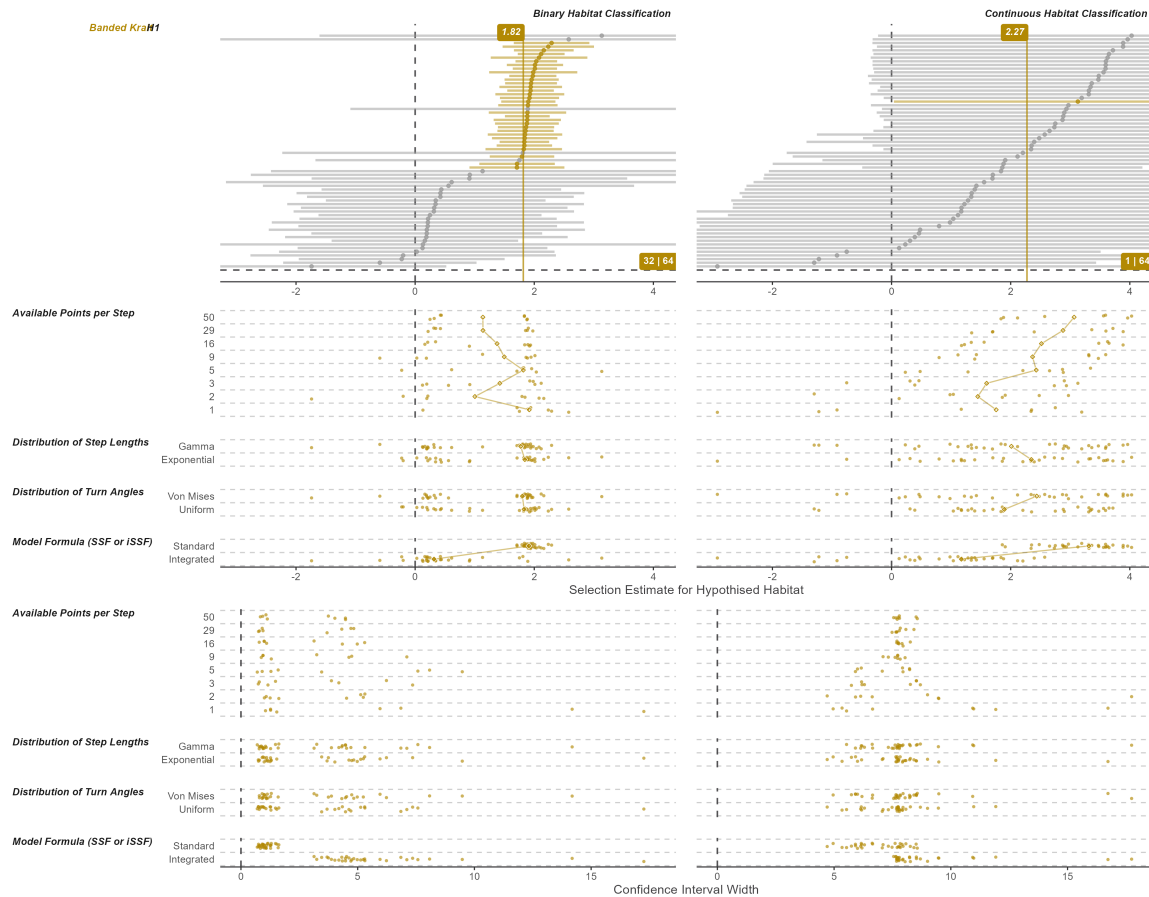


Figure 4.35: All estimates of habitat selection derived from the step-based Two-Step analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

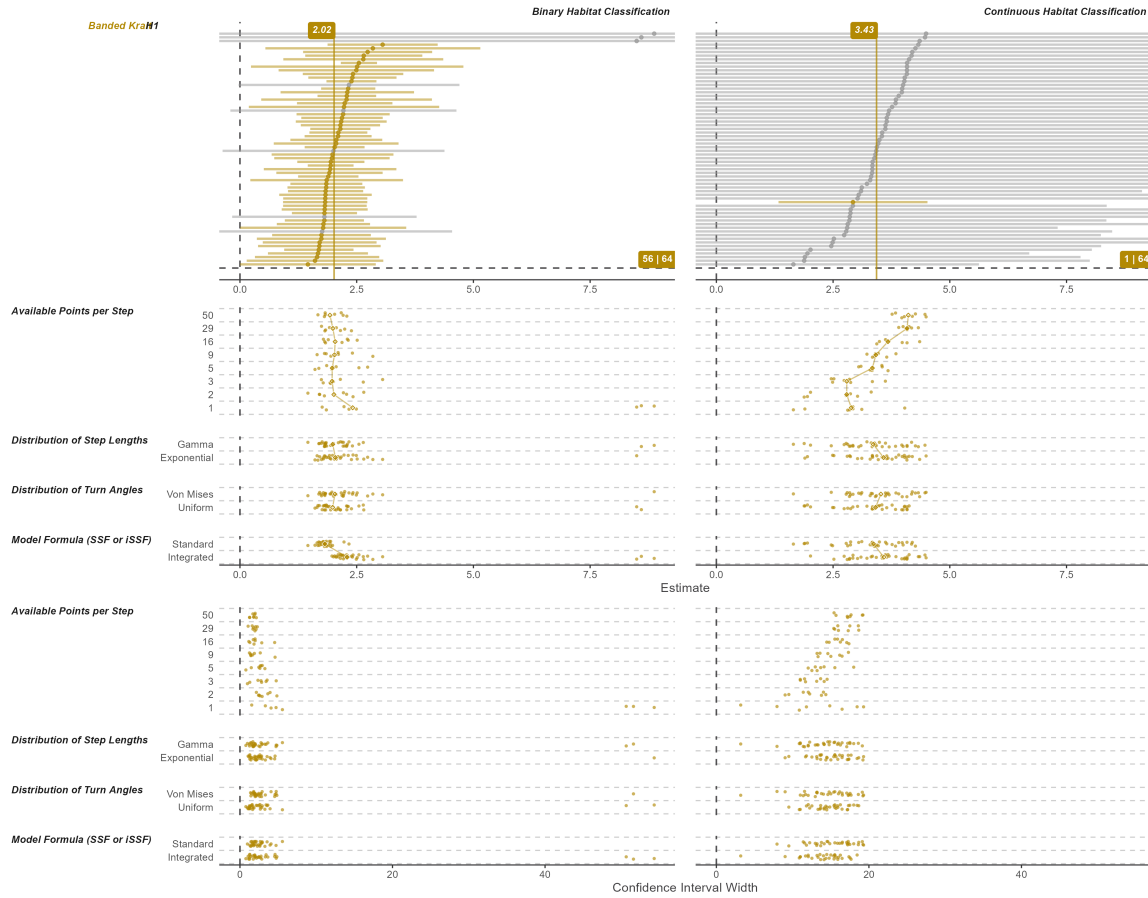


Figure 4.36: All estimates of habitat selection derived from the step-based Step Selection Function (SSF) analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

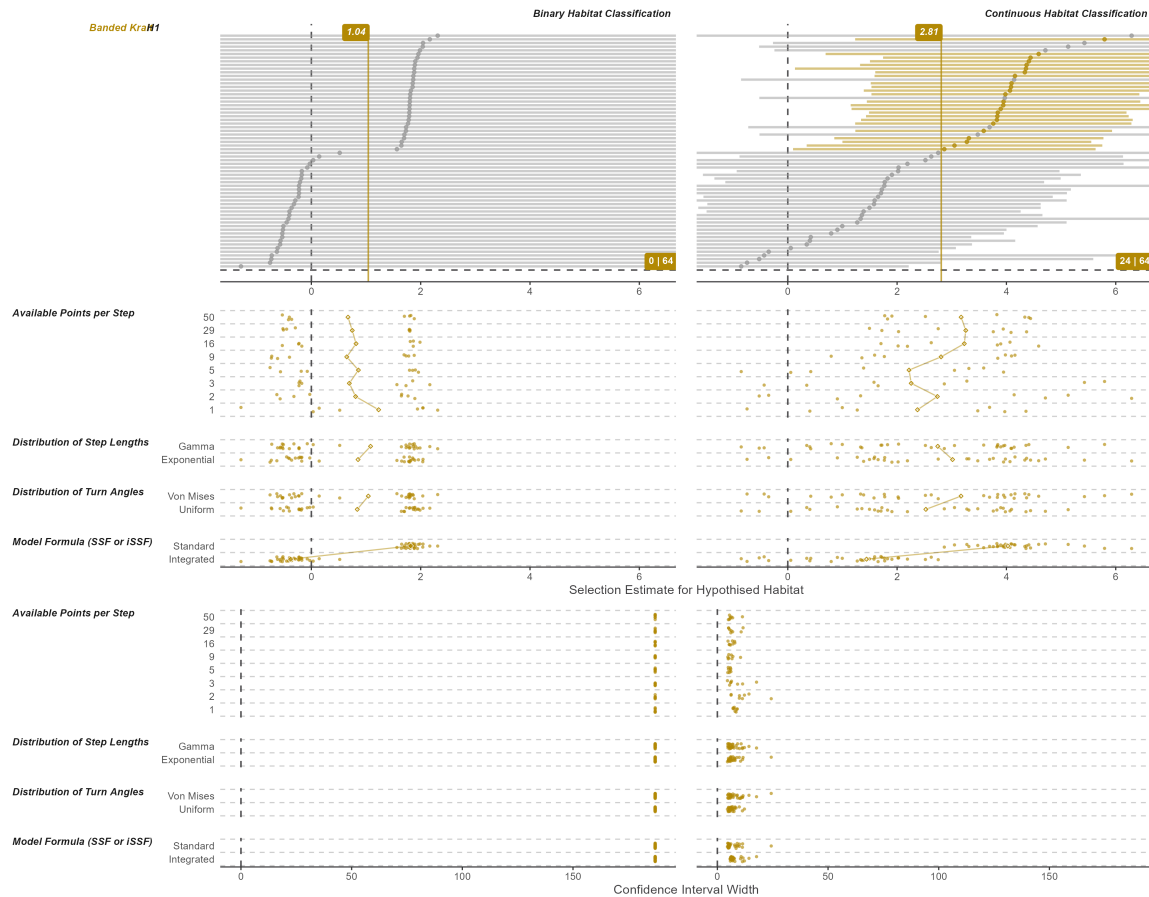
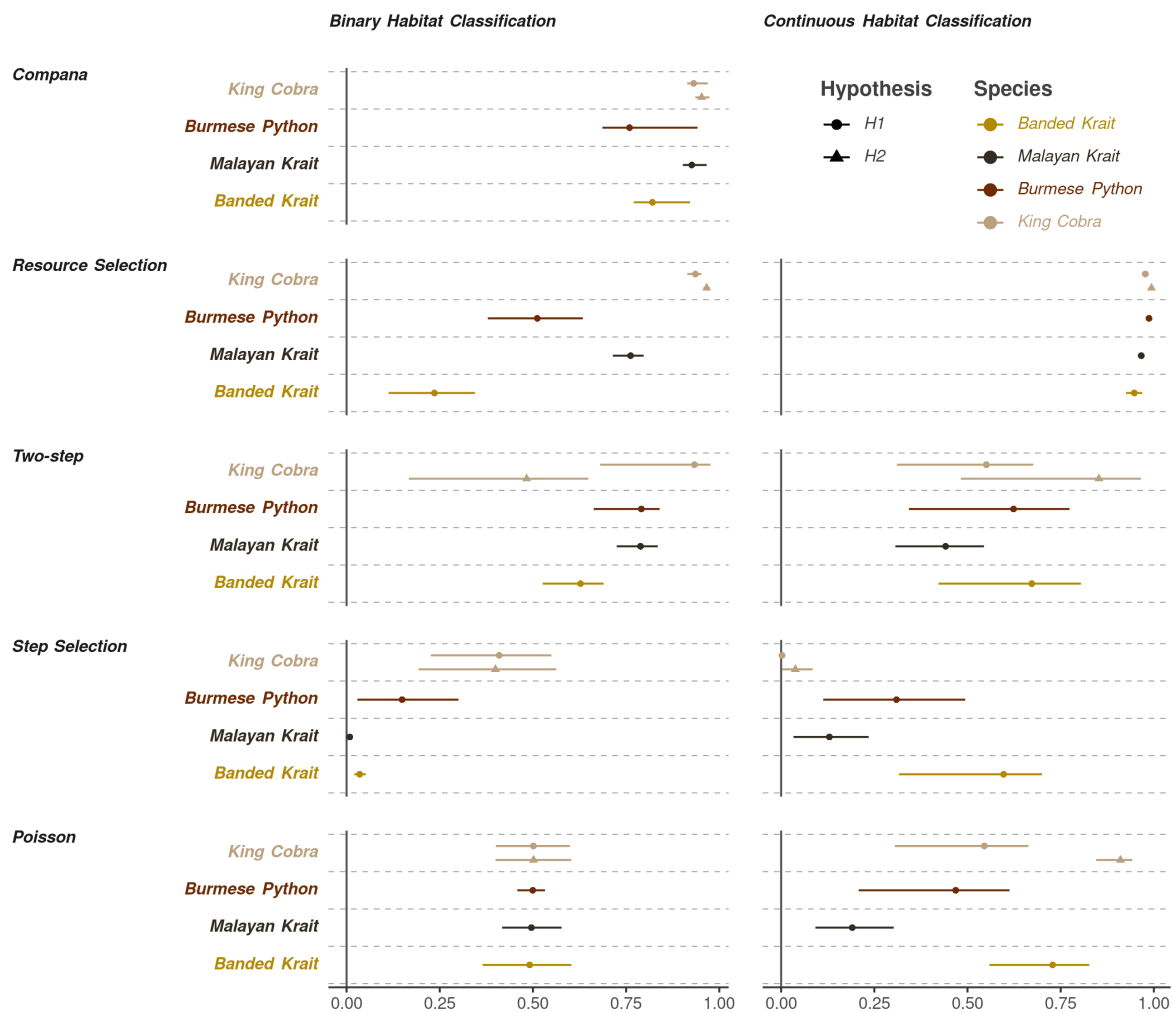


Figure 4.37: All estimates of habitat selection derived from the step-based Poisson model (Poisson) analysis for Banded Kraits (*Bungarus fasciatus*). Top curves show all estimates with coloured points and corresponding confidence intervals indicating whether those estimates significantly support the hypothesis. Bottom right labels provide a count of estimates that significantly support the hypothesis out of the total estimates. Labelled vertical lines show the median point estimate. Lower plot shows the estimates relative to each analysis choice. Median estimates are shown with hollow diamonds, and are connected with appropriately coloured lines. The plot is split left and right for the analysis using a binary classification (left), and continuous inverted distance (right).

Figure 4.38: Conditional R^2 values for all meta-analysis models, with 95% HDICs.

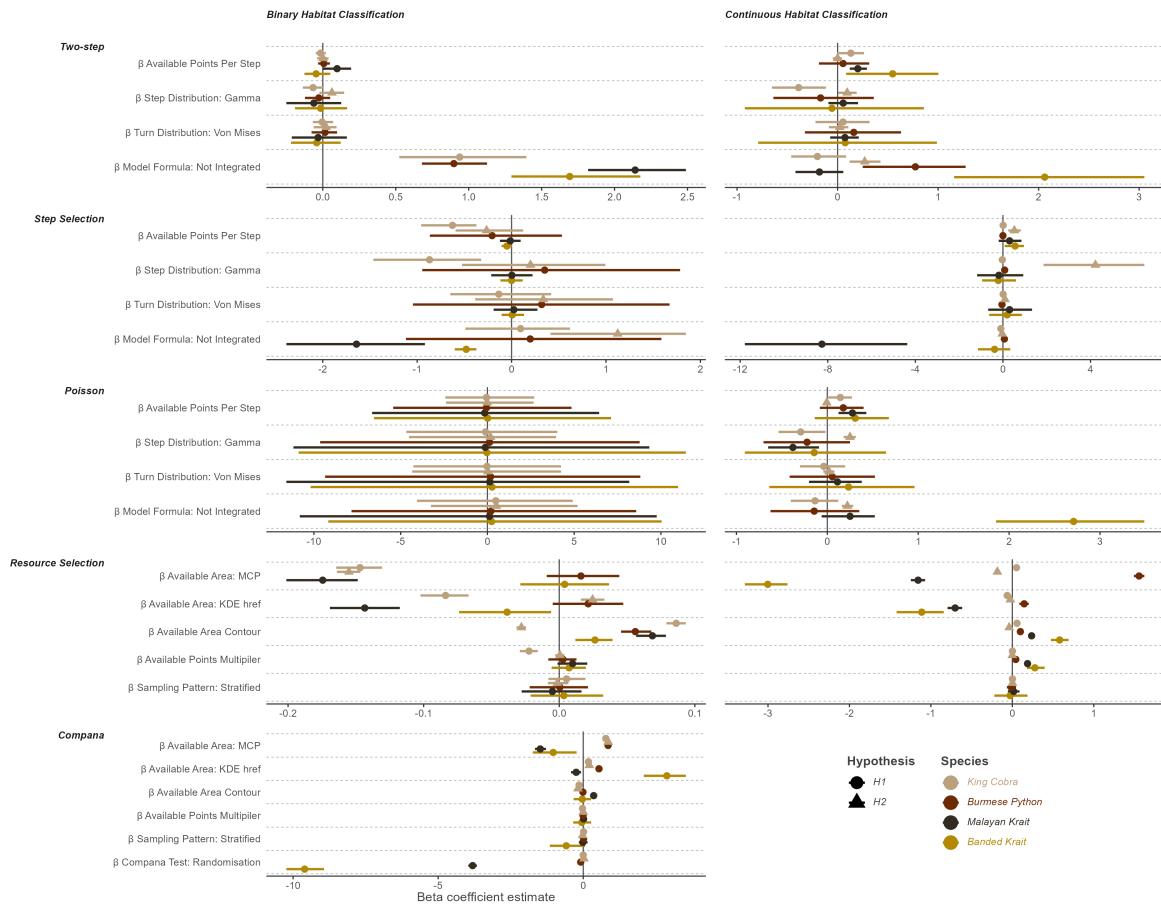


Figure 4.39: Posterior estimates of the population effects for all meta-analysis models, with 95% HDICs.

Conclusions

The motivation for this project was to determine whether the results of past habitat selection studies are reliable, given the wide breadth of analysis options available to researchers, and directly apply that to a set of snake movement case studies. Previous examination of herpetofauna spatial ecology revealed that methods are seldom reported to a sufficient level to account for differences in analysis when comparing even simpler metrics of animal movement (Crane et al., 2021). Therefore, comparisons or meta-analyses of habitat preference may be similarly impacted, with methodological differences obscuring true differences. For older studies, we also have more limited scope to re-analyse data due to very low rates of data archiving or access (Roche et al., 2015; Crane et al., 2021). This is being steadily addressed in ecology, with the use of dedicated repositories such as Movebank for movement ecology (Kays et al., 2022), and greater adoption of generic repositories (Roche et al., 2022; Jenkins et al., 2023). Shifts in publishing policies are reinforcing these requirements, enhancing the possibility of re-analyses in the future. Herpetology specifically, has scope to further improve, as many journals are yet to formally adopt positions on transparent and open science practices (Marshall & Strine, 2021).

Our assessments of simulated habitat selection and re-examination of real snake movement data suggest that the conclusions from past habitat selection studies are unlikely to be incorrect with regards to effect direction. This would indicate that the gaps in reporting and lack of data accessibility do not prohibit the results from these older studies being used to inform future study design and conservation decisions, so long as selection/no-selection is all the information required. However, in situations where the strength of selection is required, for example meta-analyses or prediction, the results of habitat selection should be carefully interpreted as analysis choices can dramatically alter the strength of effects. In these situations reanalysis of primary data is likely preferable, as that will allow confident direct comparisons and tailoring of analysis to ensure the research question is tackled directly.

Going forward, the data quantity improvements in movement ecology have the potential to lessen the impact of analysis choice for determining broad general assessments of habitat selection. We have seen in almost all presented analysis here that agreement between estimates improves with more data. However, for some species the improvements in data quantity may be slower to materialise. Snakes are a prime example, whose body layout prohibits the deployment of larger telemetry devices (Christensen & Fantuzzi, 2024). While there are ongoing efforts to miniaturise GPS devices, there remains difficulties regarding attachment and battery life (Smith et al., 2018; Gerke, Hinton & Beasley, 2021). For now, VHF and manual tracking seems likely to continue for snakes in the immediate future. Therefore, the need to compare results from studies with differing tracking regimes and downstream analysis will remain.

Much of the work here was tailored to VHF and snake tracking, where tracking frequency and duration may be more driven by external factors (e.g., equipment failure, weather, available work hours) than study design. Arguably, as with the case of more usable/affordable GPS devices compared to VHF, the findings here will diminish in relevance as new analytical methods are developed that use the underlying data structure to parametrise the optimal way to estimate habitat selection (e.g., wRSFs). These methods are better at connecting small scale movements to large scale patterns of species occurrence, either by integrating more mechanistic information or using mathematical approaches that are scale independent (Gomez et al., 2024 In Review; Michelot, Blackwell & Matthiopoulos, 2019; Michelot et al., 2020). However, these methods will not enable easy comparison with previous results if that data is not available for reanalysis. Given the monetary and welfare costs of tracking animals, replication of field studies for many species are unlikely to be justifiable (Hawkins, 2004; Bodey et al., 2018). We are left in a position where we need to maximise the existing data (Grainger et al., 2020), or where that has been lost to data entropy, the reported results of studies. Hopefully the multiverse explorations here present something to calibrate the comparisons of older results by, and better highlight how much variation can stem from a single examination of species habitat selection (even when using the same dataset).

Abolaffio M, Focardi S, Santini G. 2019. Avoiding misleading messages: Population assessment using camera trapping is not a simple task. *Journal of Animal Ecology* 88:2011–2016. DOI: 10.1111/1365-2656.13085.

Acebes P, Muñoz-Gálvez FJ, Iglesias-González Z, García-Llorente M, López-García D. 2024. Unveiling human–wildlife interactions in the context of livestock grazing abandonment and

- the return of large carnivores, ungulates and vultures: A stakeholder perspective. *People and Nature*:pan3.10769. DOI: 10.1002/pan3.10769.
- Aebischer NJ, Robertson PA, Kenward RE. 1993. Compositional Analysis of Habitat Use From Animal Radio-Tracking Data. *Ecology* 74:1313–1325. DOI: 10.2307/1940062.
- Ahearn SC, Smith JLD, Joshi AR, Ding J. 2001. TIGMOD: An individual-based spatially explicit model for simulating tiger/human interaction in multiple use forests. *Ecological Modelling* 140:81–97. DOI: 10.1016/S0304-3800(01)00258-7.
- Albers C. 2019. The problem with unadjusted multiple and sequential statistical testing. *Nature Communications* 10:1921. DOI: 10.1038/s41467-019-09941-0.
- Alberts B, Cicerone RJ, Fienberg SE, Kamb A, McNutt M, Nerem RM, Schekman R, Shiffrin R, Stodden V, Suresh S, Zuber MT, Pope BK, Jamieson KH. 2015. Self-correction in science at work. *Science* 348:1420–1422. DOI: 10.1126/science.aab3847.
- Allaire J, Xie Y, Dervieux C, McPherson J, Luraschi J, Ushey K, Atkins A, Wickham H, Cheng J, Chang W, Iannone R. 2023. *rmarkdown: Dynamic documents for r*.
- Allaire J, Xie Y, McPherson J, Luraschi J, Ushey K, Atkins A, Wickham H, Cheng J, Chang W, Iannone R. 2022. *Rmarkdown: Dynamic documents for r*.
- Alston JM, Fleming CH, Kays R, Streicher JP, Downs CT, Ramesh T, Reineking B, Calabrese JM. 2023. Mitigating pseudoreplication and bias in resource selection functions with autocorrelation-informed weighting. *Methods in Ecology and Evolution* 14:643–654. DOI: 10.1111/2041-210X.14025.
- Alston JM, Fleming CH, Noonan MJ, Tucker MA, Silva I, Folta C, Akre TSB, Ali AH, Belant JL, Beyer D, Blaum N, Böhning-Gaese K, Paula RC de, Dekker J, Drescher-Lehman J, Farwig N, Fichtel C, Fischer C, Ford AT, Janssen R, Jeltsch F, Kappeler PM, LaPoint SD, Markham AC, Medici EP, Morato RG, Nathan R, Olson KA, Patterson BD, Petroelje TR, Ramalho EE, Rösner S, Oliveira Santos LG, Schabo DG, Selva N, Sergiel A, Spiegel O, Ullmann W, Zieba F, Zwijacz-Kozica T, Wittemyer G, Fagan WF, Müller T, Calabrese JM. 2022. *Clarifying space use concepts in ecology: Range vs. Occurrence distributions*. bioRxiv. DOI: 10.1101/2022.09.29.509951.
- Altobelli JT, Dickinson KJM, Godfrey SS, Bishop PJ. 2022. Methods in amphibian biotelemetry: Two decades in review. *Austral Ecology* 47:1382–1395. DOI: 10.1111/aec.13227.
- Anderson MS, Ronning EA, De Vries R, Martinson BC. 2007. The Perverse Effects of Competition on Scientists' Work and Relationships. *Science and Engineering Ethics* 13:437–461. DOI: 10.1007/s11948-007-9042-5.

- Anvari F, Lakens D. 2018. The replicability crisis and public trust in psychological science. *Comprehensive Results in Social Psychology* 3:266–286. DOI: 10.1080/23743603.2019.1684822.
- Arrondo E, Moleón M, Cortés-Avizanda A, Jiménez J, Beja P, Sánchez-Zapata JA, Donazar JA. 2018. Invisible barriers: Differential sanitary regulations constrain vulture movements across country borders. *Biological Conservation* 219:46–52. DOI: 10.1016/j.biocon.2017.12.039.
- Auspurg K, Brüderl J. 2021. Has the Credibility of the Social Sciences Been Credibly Destroyed? Reanalyzing the “Many Analysts, One Data Set” Project. *Socius: Sociological Research for a Dynamic World* 7:237802312110244. DOI: 10.1177/23780231211024421.
- Auspurg K, Brüderl J. 2024. Toward a more credible assessment of the credibility of science by many-analyst studies. *Proceedings of the National Academy of Sciences* 121:e2404035121. DOI: 10.1073/pnas.2404035121.
- Avgar T, Betini GS, Fryxell JM. 2020. Habitat selection patterns are density dependent under the ideal free distribution. *Journal of Animal Ecology* 89:2777–2787. DOI: 10.1111/1365-2656.13352.
- Avgar T, Potts JR, Lewis MA, Boyce MS. 2016. Integrated step selection analysis: Bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution* 7:619–630. DOI: 10.1111/2041-210X.12528.
- Baker M. 2016. 1,500 scientists lift the lid on reproducibility. *Nature* 533:452–454. DOI: 10.1038/533452a.
- Barnes CH, Abdulaziz UZ, Kaenphet A, Kanlayanapaphon C. 2024. Snake diversity, occupancy, and detection on Thailand’s largest university campus. *Ecology and Evolution* 14:e70317. DOI: 10.1002/ece3.70317.
- Barnes CH, Strine CT, Suwanwaree P, Hill III JG. 2017. Movement and home range of green pit vipers (*Trimeresurus* spp.) in a rural landscape in north-east Thailand. *Herpetological Bulletin* 142:19–28.
- Bartoń K. 2023. *MuMIn: Multi-model inference*.
- Barto EK, Rillig MC. 2012. Dissemination biases in ecology: Effect sizes matter more than quality. *Oikos* 121:228–235. DOI: 10.1111/j.1600-0706.2011.19401.x.
- Bartoszek I, Andreadis PT, Prokop-Ervin C, Patel M, Reed RN. 2018. Natural History Note: Python bivittatus (Burmese Python). Diet and Prey Size. *Herpetological Review* 49:139–140.

- Bastiaansen JA, Kunkels YK, Blaauw FJ, Boker SM, Ceulemans E, Chen M, Chow S-M, Jonge P de, Emerencia AC, Epskamp S, Fisher AJ, Hamaker EL, Kuppens P, Lutz W, Meyer MJ, Moulder R, Oravecz Z, Riese H, Rubel J, Ryan O, Servaas MN, Sjobeck G, Snippe E, Trull TJ, Tschacher W, Veen DC van der, Wichers M, Wood PK, Woods WC, Wright AGC, Albers CJ, Bringmann LF. 2020. Time to get personal? The impact of researchers choices on the selection of treatment targets using the experience sampling methodology. *Journal of Psychosomatic Research* 137:110211. DOI: 10.1016/j.jpsychores.2020.110211.
- Bastille-Rousseau G, Murray DL, Schaefer JA, Lewis MA, Mahoney SP, Potts JR. 2017. Spatial scales of habitat selection decisions: Implications for telemetry-based movement modelling. *Ecography* 40:1–7. DOI: 10.1111/ecog.02655.
- Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48. DOI: 10.18637/jss.v067.i01.
- Bauder JM, Breininger DR, Bolt MR, Legare ML, Jenkins CL, Rothermel BB, McGarigal K. 2016. Seasonal Variation in Eastern Indigo Snake (*Drymarchon couperi*) Movement Patterns and Space Use in Peninsular Florida at Multiple Temporal Scales. *Herpetologica* 72:214–226. DOI: 10.1655/Herpetologica-D-15-00039.1.
- Bennett DA, Tang W. 2006. Modelling adaptive, spatially aware, and mobile agents: Elk migration in Yellowstone. *International Journal of Geographical Information Science* 20:1039–1066. DOI: 10.1080/13658810600830806.
- Beyer HL, Haydon DT, Morales JM, Frair JL, Hebblewhite M, Mitchell M, Matthiopoulos J. 2010. The interpretation of habitat preference metrics under use–availability designs. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:2245–2254. DOI: 10.1098/rstb.2010.0083.
- Bishop D. 2019. Rein in the four horsemen of irreproducibility. *Nature* 568:435–435. DOI: 10.1038/d41586-019-01307-2.
- Bivand RS, Pebesma E, Gomez-Rubio V. 2013. *Applied spatial data analysis with R, second edition*. Springer, NY.
- Boback SM, Nafus MG, Yackel Adams AA, Reed RN. 2020. Use of visual surveys and radiotelemetry reveals sources of detection bias for a cryptic snake at low densities. *Ecosphere* 11. DOI: 10.1002/ecs2.3000.
- Bodey TW, Cleasby IR, Bell F, Parr N, Schultz A, Votier SC, Bearhop S. 2018. A phylogenetically controlled meta-analysis of biologging device effects on birds: Deleterious effects and a call for more standardized reporting of study data. *Methods in Ecology and Evolution*

- 9:946–955. DOI: 10.1111/2041-210X.12934.
- Bracis C, Bildstein KL, Mueller T. 2018. Revisitation analysis uncovers spatio-temporal patterns in animal movement data. *Ecography*. DOI: 10.1111/ecog.03618.
- Brembs B. 2018. Prestigious Science Journals Struggle to Reach Even Average Reliability. *Frontiers in Human Neuroscience* 12:1–7. DOI: 10.3389/fnhum.2018.00037.
- Brembs B. 2019. Reliable novelty: New should not trump true. *PLOS Biology* 17:e3000117. DOI: 10.1371/journal.pbio.3000117.
- Breznau N, Rinke EM, Wuttke A, Adem M, Adriaans J, Alvarez-Benjumea A, Andersen HK, Auer D, Azevedo F, Bahnsen O, Balzer D, Bauer G, Bauer P, Baumann M, Baute S, Benoit V, Bernauer J, Berning C, Berthold A, Bethke FS, Biegert T, Blinzler K, Blumenberg J, Bobzien L, Bohman A, Bol T, Bostic A, Brzozowska Z, Burgdorf K, Burger K, Busch K, Castillo JC, Chan N, Christmann P, Connelly R, Czymara CS, Damian E, Ecker A, Edelmann A, Eger MA, Ellerbrock S, Forke A, Forster AG, Gaasendam C, Gavras K, Gayle V, Gessler T, Gnambs T, Godefroidt A, Grömping M, Groß M, Gruber S, Gummer T, Hadjar A, Heisig JP, Hellmeier S, Heyne S, Hirsch M, Hjerm M, Hochman O, Hövermann A, Hunger S, Hunkler C, Huth N, Ignacz Z, Jacobs L, Jacobsen J, Jaeger B, Jungkunz S, Jungmann N, Kauff M, Kleinert M, Klinger J, Kolb J-P, Kołczyńska M, Kuk JS, Kunißen K, Sinatra DK, Greinert A, Lersch PM, Löbel L-M, Lutscher P, Mader M, Madia JE, Malancu N, Maldonado L, Marahrens H, Martin N, Martinez P, Mayerl J, Mayorga OJ, McManus P, Wagner K, Meeusen C, Meierrieks D, Mellon J, Merhout F, Merk S, Meyer D, Micheli L, Mijs JJB, Moya C, Neunhoeffler M, Nüst D, Nygård O, Ochsenfeld F, Otte G, Pechenkina A, Prosser C, Raes L, Ralston K, Ramos M, Roets A, Rogers J, Ropers G, Samuel R, Sand G, Schachter A, Schaeffer M, Schieferdecker D, Schlueter E, Schmidt K, Schmidt R, Schmidt-Catran A, Schmiedeberg C, Schneider J, Schoonvelde M, Schulte-Cloos J, Schumann S, Schunck R, Schupp J, Seuring J, Silber H, Slegers WWA, Sonntag N, Staudt A, Steiber N, Steiner N, Sternberg S, Stiers D, Stojmenovska D, Storz N, Striessnig E, Stroppe A-K, Teltemann J, Tibajev A, Tung BB, Vagni G, Van Assche J, Linden M van der, Noll J van der, Van Hoogtem A, Vogtenhuber S, Voicu B, Wagemans F, Wehl N, Werner H, Wiernik BM, Winter F, Wolf C, Yamada Y, Zhang N, Ziller C, Zins S, Zóltak T, Nguyen HHV. 2021. Observing Many Researchers Using the Same Data and Hypothesis Reveals a Hidden Universe of Uncertainty. *MetaArXiv*. DOI: 10.31222/osf.io/cd5j9.
- Brown PR, Singleton GR. 2001. Habitat use and movements of the rice-field rat, *Rattus argen-*

- tiventer*, in West Java, Indonesia. *mamm* 65:151–166. DOI: 10.1515/mamm.2001.65.2.151.
- Bryan CJ, Yeager DS, O’Brien JM. 2019. Replicator degrees of freedom allow publication of misleading failures to replicate. *Proceedings of the National Academy of Sciences* 116:25535–25545. DOI: 10.1073/pnas.1910951116.
- Bürkner P-C. 2021. Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software* 100:1–54. DOI: 10.18637/jss.v100.i05.
- Calabrese JM, Fleming CH, Gurarie E. 2016. Ctm: An R Package for Analyzing Animal Relocation Data As a Continuous-Time Stochastic Process. *Methods in Ecology and Evolution* 7:1124–1132. DOI: 10.1111/2041-210X.12559.
- Calenge C, Mathieu Basille contributions from. 2023. *adehabitatHS: Analysis of habitat selection by animals*.
- Calenge C, Scott Fortmann-Roe contributions from. 2023. *adehabitatHR: Home range estimation*.
- Campitelli E. 2024. *Ggnewscale: Multiple fill and colour scales in 'ggplot2'*.
- Cappelle N, Howe EJ, Boesch C, Kühl HS. 2021. Estimating animal abundance and effort–precision relationship with camera trap distance sampling. *Ecosphere* 12:e03299. DOI: 10.1002/ecs2.3299.
- Cassey P, Ewen JG, Blackburn TM, Møller AP. 2004. A survey of publication bias within evolutionary ecology. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271. DOI: 10.1098/rsbl.2004.0218.
- Ching T. 2023. *qs: Quick serialization of r objects*.
- Christensen T, Fantuzzi J. 2024. External transmitter attachment in snakes: A systematic review of methods, efficacy, and impacts. *Animal Biotelemetry* 12:14. DOI: 10.1186/s40317-024-00371-4.
- Christie AP, Amano T, Martin PA, Shackelford GE, Simmons BI, Sutherland WJ. 2019. Simple study designs in ecology produce inaccurate estimates of biodiversity responses. *Journal of Applied Ecology* 56:2742–2754. DOI: 10.1111/1365-2664.13499.
- Christy M, Savidge J, Yackel Adams A, Gragg J, Rodda G. 2017. Experimental landscape reduction of wild rodents increases movements in the invasive brown treesnake (*Boiga irregularis*). *Management of Biological Invasions* 8:455–467. DOI: 10.3391/mbi.2017.8.4.01.
- Clark TD, Raby GD, Roche DG, Binning SA, Speers-Roesch B, Jutfelt F, Sundin J. 2020. Ocean acidification does not impair the behaviour of coral reef fishes. *Nature* 577:370–375. DOI: 10.1038/s41586-019-1903-y.

- Clements JC, Sundin J, Clark TD, Jutfelt F. 2022. Meta-analysis reveals an extreme “decline effect” in the impacts of ocean acidification on fish behavior. *PLOS Biology* 20:e3001511. DOI: 10.1371/journal.pbio.3001511.
- Community TTW. 2025. The Turing Way: A handbook for reproducible, ethical and collaborative research. DOI: 10.5281/ZENODO.15213042.
- Cooley D. 2023. *sfheaders: Converts between r objects and simple feature objects*.
- Craiu RV, Duchesne T, Fortin D, Baillargeon S. 2011. Conditional Logistic Regression With Longitudinal Follow-up and Individual-Level Random Coefficients: A Stable and Efficient Two-Step Estimation Method. *Journal of Computational and Graphical Statistics* 20:767–784. DOI: 10.1198/jcgs.2011.09189.
- Craiu RV, Duchesne T, Fortin D, Baillargeon S. 2016. *TwoStepCLogit: Conditional logistic regression: A two-step estimation method*.
- Crane MS, Silva IMS, Marshall BM, Strine CT. 2020. Supplementary files for Crane et al., “Lots of movement, little progress”: R Code, data and figures. DOI: 10.5281/ZENODO.4303887.
- Crane M, Silva I, Marshall BM, Strine CT. 2021. Lots of movement, little progress: A review of reptile home range literature. *PeerJ* 9:e11742. DOI: 10.7717/peerj.11742.
- D’souza A, Gale GA, Marshall BM, Khamcha D, Waengsothorn S, Strine CT. 2021. Space use and activity of *Boiga cyanea* – A major songbird nest predator in a seasonal tropical forest in Thailand. *Global Ecology and Conservation* 32:e01875. DOI: 10.1016/j.gecco.2021.e01875.
- Das I, Gowri Shankar P, Swamy P, Williams RC, Lalremsanga HT, Prashanth P, Sahoo G, Vijayakumar SP, Höglund J, Shanker K, Dutta SK, Ganesh SR, Wüster W. 2024. Taxonomic revision of the king cobra *Ophiophagus hannah* (Cantor, 1836) species complex (Reptilia: Serpentes: Elapidae), with the description of two new species. *European Journal of Taxonomy* 961. DOI: 10.5852/ejt.2024.961.2681.
- DeBruine LM, Barr DJ. 2021. Understanding Mixed-Effects Models Through Data Simulation. *Advances in Methods and Practices in Psychological Science* 4:1–15.
- Del Giudice M, Gangestad SW. 2021. A Traveler’s Guide to the Multiverse: Promises, Pitfalls, and a Framework for the Evaluation of Analytic Decisions. *Advances in Methods and Practices in Psychological Science* 4:251524592095492. DOI: 10.1177/2515245920954925.
- Desbureaux S. 2021. Subjective modeling choices and the robustness of impact evaluations in conservation science. *Conservation Biology* 35:1615–1626. DOI: 10.1111/cobi.13728.

- Devezer B, Navarro DJ, Vandekerckhove J, Ozge Buzbas E. 2021. The case for formal methodology in scientific reform. *Royal Society Open Science* 8:rsos.200805, 200805. DOI: 10.1098/rsos.200805.
- Doherty TS, Driscoll DA. 2018. Coupling movement and landscape ecology for animal conservation in production landscapes. *Proceedings of the Royal Society B: Biological Sciences* 285:20172272. DOI: 10.1098/rspb.2017.2272.
- Doherty TS, Fist CN, Driscoll DA. 2019. Animal movement varies with resource availability, landscape configuration and body size: A conceptual model and empirical example. *Landscape Ecology* 34:603–614. DOI: 10.1007/s10980-019-00795-x.
- Dolia J, Das A, Kelkar N. 2023. House-warming: Wild king cobra nests have thermal regimes that positively affect hatching success and hatchling size. *Journal of Thermal Biology* 112:103468. DOI: 10.1016/j.jtherbio.2023.103468.
- Dorcas ME, Willson JD, Reed RN, Snow RW, Rochford MR, Miller MA, Meshaka WE, Andreadis PT, Mazzotti FJ, Romagosa CM, Hart KM. 2012. Severe mammal declines coincide with proliferation of invasive Burmese pythons in Everglades National Park. *Proceedings of the National Academy of Sciences* 109:2418–2422. DOI: 10.1073/pnas.1115226109.
- Dragicevic P, Jansen Y, Sarma A, Kay M, Chevalier F. 2019. Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland Uk: ACM, 1–15. DOI: 10.1145/3290605.3300295.
- Durso AM, Seigel RA. 2015. A Snake in the Hand is Worth 10,000 in the Bush. *Journal of Herpetology* 49:503–506. DOI: 10.1670/15-49-04.1.
- Duvall D, Schuett GW. 1997. Straight-line movement and competitive mate searching in prairie rattlesnakes, *Crotalus viridis viridis*. *Animal Behaviour* 54:329–334. DOI: 10.1006/anbe.1996.0418.
- Eddelbuettel D. 2013. *Seamless R and C++ integration with Rcpp*. New York: Springer. DOI: 10.1007/978-1-4614-6868-4.
- Eddelbuettel D, Balamuta JJ. 2018. Extending extitR with extitC++: A Brief Introduction to extitRcpp. *The American Statistician* 72:28–36. DOI: 10.1080/00031305.2017.1375990.
- Eddelbuettel D, Emerson JW, Kane MJ. 2021. *BH: Boost c++ header files*.
- Eddelbuettel D, François R. 2011. Rcpp: Seamless R and C++ integration. *Journal of Statistical Software* 40:1–18. DOI: 10.18637/jss.v040.i08.
- Fanelli D. 2010b. “Positive” Results Increase Down the Hierarchy of the Sciences. *PLoS ONE*

- 5:e10068. DOI: 10.1371/journal.pone.0010068.
- Fanelli D. 2010a. Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. *PLoS ONE* 5:e10271. DOI: 10.1371/journal.pone.0010271.
- Fanelli D. 2012. Negative results are disappearing from most disciplines and countries. *Scientometrics* 90:891–904. DOI: 10.1007/s11192-011-0494-7.
- Farrar BG, Ostojic L. 2021. It's not just the animals that are STRANGE. *Learning & Behavior* 49:169–170. DOI: 10.3758/s13420-020-00442-5.
- Feore S, Montgomery WI. 1999. Habitat effects on the spatial ecology of the European badger (*Meles meles*). *Journal of Zoology* 247:537–549. DOI: 10.1111/j.1469-7998.1999.tb01015.x.
- Fieberg J, Signer J, Smith B, Avgar T. 2021. A “How to” guide for interpreting parameters in habitat-selection analyses. *Journal of Animal Ecology* 90:1027–1043. DOI: 10.1111/1365-2656.13441.
- Fleming CH, Calabrese JM. 2017. A new kernel density estimator for accurate home-range and species-range area estimation. *Methods in Ecology and Evolution* 8:571–579. DOI: 10.1111/2041-210X.12673.
- Fleming CH, Calabrese JM. 2023. *ctmm: Continuous-time movement modeling*.
- Fleming CH, Calabrese JM, Mueller T, Olson KA, Leimgruber P, Fagan WF. 2014. From Fine-Scale Foraging to Home Ranges: A Semivariance Approach to Identifying Movement Modes across Spatiotemporal Scales. *The American Naturalist* 183:E154–E167. DOI: 10.1086/675504.
- Fleming CH, Fagan WF, Mueller T, Olson KA, Leimgruber P, Calabrese JM. 2015. Rigorous home range estimation with movement data: A new autocorrelated kernel density estimator. *Ecology* 96:1182–1188.
- Fleming CH, Sheldon D, Fagan WF, Leimgruber P, Mueller T, Nandintsetseg D, Noonan MJ, Olson KA, Setyawan E, Sianipar A, Calabrese JM. 2018. Correcting for missing and irregular data in home-range estimation. *Ecological Applications* 28:1003–1010. DOI: 10.1002/eap.1704.
- Forester JD, Im HK, Rathouz PJ. 2009. Accounting for animal movement in estimation of resource selection functions: Sampling and data analysis. *Ecology* 90:3554–3565. DOI: 10.1890/08-0874.1.
- Forstmeier W, Schielzeth H. 2011. Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology* 65:47–55. DOI: 10.1007/s00265-010-1038-5.

- Forstmeier W, Wagenmakers E-J, Parker TH. 2017. Detecting and avoiding likely false-positive findings – a practical guide: Avoiding false-positive findings. *Biological Reviews* 92:1941–1968. DOI: 10.1111/brv.12315.
- Frair JL, Fieberg J, Hebblewhite M, Cagnacci F, DeCesare NJ, Pedrotti L. 2010. Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:2187–2200. DOI: 10.1098/rstb.2010.0084.
- Francisco Rodríguez-Sánchez, Connor P. Jackson, Shaurita D. Hutchins. 2023. *grateful: Facilitate citation of r packages*.
- Fraser H, Barnett A, Parker TH, Fidler F. 2020. The role of replication studies in ecology. *Ecology and Evolution* 10:5197–5207. DOI: 10.1002/ece3.6330.
- Fraser KC, Davies KT, Davy CM, Ford AT, Flockhart DTT, Martins EG. 2018a. Tracking the conservation promise of movement ecology. *Frontiers in Ecology and Evolution* 6:150. DOI: 10.3389/FEVO.2018.00150.
- Fraser H, Parker T, Nakagawa S, Barnett A, Fidler F. 2018b. Questionable research practices in ecology and evolution. *PLOS ONE* 13:e0200303. DOI: 10.1371/journal.pone.0200303.
- Freedman LP, Cockburn IM, Simcoe TS. 2015. The Economics of Reproducibility in Preclinical Research. *PLOS Biology* 13:e1002165. DOI: 10.1371/journal.pbio.1002165.
- Fujioka M, Don Lee S, Kurechi M. 2010. Bird use of Rice Fields in Korea and Japan. *Waterbirds* 33:8. DOI: 10.1675/063.033.s102.
- Gabry J, Mahr T. 2024. Bayesplot: Plotting for bayesian models.
- García-Jiménez R, Pérez-García JM, Margalida A. 2018. Drivers of daily movement patterns affecting an endangered vulture flight activity. *BMC Ecology* 18:39. DOI: 10.1186/s12898-018-0195-7.
- Gelman A, Loken E. 2013. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. :17.
- Gerke HC, Hinton TG, Beasley JC. 2021. Movement Behavior and Habitat Selection of Rat Snakes (*Elaphe* spp.) in the Fukushima Exclusion Zone. *Ichthyology & Herpetology* 109. DOI: 10.1643/h2019282.
- Gerstner K, Moreno-Mateos D, Gurevitch J, Beckmann M, Kambach S, Jones HP, Seppelt R. 2017. Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. *Methods in Ecology and Evolution* 8:777–784. DOI: https:

[//doi.org/10.1111/2041-210X.12758](https://doi.org/10.1111/2041-210X.12758).

- Gillies CS, Hebblewhite M, Nielsen SE, Krawchuk MA, Aldridge CL, Frair JL, Saher DJ, Stevens CE, Jerde CL. 2006. Application of random effects to the study of resource selection by animals. *Journal of Animal Ecology* 75:887–898. DOI: 10.1111/j.1365-2656.2006.01106.x.
- Goldstein E, Erinjery JJ, Martin G, Kasturiratne A, Ediriweera DS, Silva HJ de, Diggle P, Lalloo DG, Murray KA, Iwamura T. 2021. Integrating human behavior and snake ecology with agent-based models to predict snakebite in high risk landscapes. *PLOS Neglected Tropical Diseases* 15:e0009047. DOI: 10.1371/journal.pntd.0009047.
- Gomez S, English H, Alegre V, Blackwell P, Bracken A, Bray E, Evans L, Gan J, Gutmann R. C, Grecian J, Harju S, Hejzmanová P, Marshall B, Matthiopoulos J, Mnenge A, Ortega Z, Potts J, Russell C, Rutz C, Singh N, Whyte K, Börger L. 2024 In Review. Understanding and predicting animal movements and distributions in the anthropocene. *Journal of Animal Ecology*.
- Gould E, Fraser H, Parker T, Nakagawa S, Griffith S, Vesik P, Fidler F, Abbey-Lee R, Abbott J, Aguirre L, Alcaraz C, Altschul D, Arekar K, Atkins J, Atkinson J, Barrett M, Bell K, Bello S, Berauer B, Bertram M, Billman P, Blake C, Blake S, Bliard L, Bonisoli-Alquati A, Bonnet T, Bordes C, Bose A, Botterill-James T, Boyd M, Boyle S, Bradfer-Lawrence T, Brand J, Brengdahl M, Bulla M, Bussière L, Camerlenghi E, Campbell S, Campos L, Caravaggi A, Cardoso P, Carroll C, Catanach T, Chen X, Chik HYJ, Choy E, Christie A, Chuang A, Chunco A, Clark B, Cox M, Cressman K, Crouch C, D’Amelio P, De Sousa A, Döbert T, Dobler R, Dobson A, Doherty T, Drobniak S, Duffy A, Dunn R, Dunning J, Eberhart-Hertel L, Elmore J, Elsherif M, English H, Ensminger D, Ernst U, Ferguson S, Ferreira-Arruda T, Fieberg J, Finch E, Fiorenza E, Fisher D, Forstmeier W, Fourcade Y, Francesca Santostefano F, Frank G, Freund C, Gandy S, Gannon D, García-Cervigón A, Geron C, Gilles M, Girndt A, Gliksman D, Goldspiel H, Gomes D, Goslee S, Gosnell J, Gratton P, Grebe N, Greenler S, Griffith D, Griffith F, Grossman J, Güncan A, Haesen S, Hagan J, Harrison N, Hasnain S, Havird J, Heaton A, Hsu B-Y, Iranzo E, Iverson E, Jimoh S, Johnson D, Johnsson M, Jorna J, Jucker T, Jung M, Kačergytė I, Ke A, Kelly C, Keogan K, Keppeler F, Killion A, Kim D, Kochan D, Korsten P, Kothari S, Kuppler J, Kusch J, Lagisz M, Larkin D, Larson C, Lauck K, Lauterbur M, Law A, Léandri-Breton D-J, Lievens E, Lima D, Lindsay S, Macphie K, Mair M, Malm L, Mammola S, Manhart M, Mäntylä E, Marchand P, Marshall B, Martin D, Martin J, Martin C, Martinig A,

- McCallum E, McNew S, Meiners S, Michelangeli M, Moiron M, Moreira B, Mortensen J, Mos B, Muraina T, Nelli L, Nilsonne G, Nolazco S, Nooten S, Novotny J, Olin A, Organ C, Ostevik K, Palacio F, Paquet M, Pascall D, Pasquarella V, Payo-Payo A, Pedersen K, Perez G, Perry K, Pottier P, Proulx M, Proulx R, Pruett J, Ramananjato V, Randimbiarison F, Razafindratsima O, Rennison D, Riva F, Riyahi S, Roast M, Rocha F, Roche D, Román-Palacios C, Rosenberg M, Ross J, Rowland F, Rugemalila D, Russell A, Ruuskanen S, Saccone P, Sadeh A, Salazar S, Sales K, Salmón P, Sanchez-Tojar A, Santos L, Schilling H, Schmidt M, Schmoll T, Schneider A, Schrock A, Schroeder J, Schtickzelle N, Schultz N, Scott D, Shapiro J, Sharma N, Shearer C, Sitvarin M, Skupien F, Slinn H, Smith J, Smith G, Sollmann R, Stack Whitney K, Still S, Stuber E, Sutton G, Swallow B, Taff C, Takola E, Tanentzap A, Thawley C, Tortorelli C, Trlica A, Turnell B, Urban L, Van De Vondel S, Van Oordt F, Vanderwel M, Vanderwel K, Vanderwolf K, Verrelli B, Vieira M, Vollering J, Walker X, Walter J, Waryszak P, Weaver R, Weller D, Whelan S, White R, Wolfson D, Wood A, Yanco S, Yen J, Youngflesh C, Zilio G, Zimmer C, Zitomer R, Villamil N, Tompkins E. 2023. Same data, different analysts: Variation in effect sizes due to analytical decisions in ecology and evolutionary biology. *EcoEvoRxiv*. DOI: 10.32942/X2GG62.
- Grainger M, Bolam FC, Stewart G, Nilsen EB. 2019. Evidence synthesis for tackling research waste. *EcoEvoRxiv*. DOI: 10.32942/osf.io/42fkh.
- Grainger MJ, Bolam FC, Stewart GB, Nilsen EB. 2020. Evidence synthesis for tackling research waste. *Nature Ecology & Evolution* 4:495–497. DOI: 10.1038/s41559-020-1141-6.
- Griffith SC. 2019. Cooperation and Coordination in Socially Monogamous Birds: Moving Away From a Focus on Sexual Conflict. *Frontiers in Ecology and Evolution* 7:455. DOI: 10.3389/fevo.2019.00455.
- Groffen J, Borzée A, Jang Y. 2018. Preference for natural borders in rice paddies by two treefrog species. *Animal Cells and Systems* 22:205–211. DOI: 10.1080/19768354.2018.1475301.
- Guerra-Castro EJ, Cajas JC, Simões N, Cruz-Motta JJ, Mascaró M. 2021. SSP: An R package to estimate sampling effort in studies of ecological communities. *Ecography* 44:561–573. DOI: 10.1111/ecog.05284.
- Gupta A, Dilkina B, Morin DJ, Fuller AK, Royle JA, Sutherland C, Gomes CP. 2019. Reserve design to optimize functional connectivity and animal density. *Conservation Biology*:cobi.13369. DOI: 10.1111/cobi.13369.
- Gurarie E, Andrews RD, Laidre KL. 2009. A novel method for identifying behavioural changes in animal movement data. *Ecology Letters* 12:395–408. DOI: 10.1111/j.1461-

0248.2009.01293.x.

- Hall BD, Liu Y, Jansen Y, Dragicevic P, Chevalier F, Kay M. 2022. A Survey of Tasks and Visualizations in Multiverse Analysis Reports. *Computer Graphics Forum* 41:402–426. DOI: 10.1111/cgf.14443.
- Hawkins P. 2004. Bio-logging and animal welfare: Practical refinements. *Memoirs of National Institute of Polar Research. Special Issue* 58:58–68.
- Hernangómez D. 2023. Using the tidyverse with terra objects: The tidyterra package. *Journal of Open Source Software* 8:5751. DOI: 10.21105/joss.05751.
- Hijmans RJ. 2022. *Raster: Geographic data analysis and modeling*.
- Hijmans RJ. 2023a. *raster: Geographic data analysis and modeling*.
- Hijmans RJ. 2023b. *terra: Spatial data analysis*.
- Hodges CW, Barnes CH, Patungtaro P, Strine CT. 2021a. Deadly dormmate: A case study on Bungarus candidus living among a student dormitory with implications for human safety. *Ecological Solutions and Evidence* 2. DOI: 10.1002/2688-8319.12047.
- Hodges CW, D’souza A, Jintapirom S. 2020. Diurnal observation of a Malayan Krait Bungarus candidus (Reptilia: Elapidae) feeding inside a building in Thailand. *Journal of Threatened Taxa* 12:15947–15950. DOI: 10.11609/jott.5746.12.8.15947-15950.
- Hodges CW, Marshall BM, Hill III JG, Strine CT. 2021b. Malayan kraits (Bungarus candidus) show affinity to anthropogenic structures in a human dominated landscape. DOI: 10.5281/ZENODO.5495839.
- Hodges CW, Marshall BM, Hill JG, Strine CT. 2022. Malayan kraits (Bungarus candidus) show affinity to anthropogenic structures in a human dominated landscape. *Scientific Reports* 12:7139. DOI: 10.1038/s41598-022-11255-z.
- Hoffmann S, Schönbrodt F, Elsas R, Wilson R, Strasser U, Boulesteix A-L. 2021. The multiplicity of analysis strategies jeopardizes replicability: Lessons learned across disciplines. *Royal Society Open Science* 8:rsos.201925, 201925. DOI: 10.1098/rsos.201925.
- Hofman MPG, Hayward MW, Heim M, Marchand P, Rolandsen CM, Mattisson J, Urbano F, Heurich M, Mysterud A, Melzheimer J, Morellet N, Voigt U, Allen BL, Gehr B, Rouco C, Ullmann W, Holand Ø, Jørgensen NH, Steinheim G, Cagnacci F, Kroeschel M, Kaczensky P, Buuveibaatar B, Payne JC, Palmegiani I, Jerina K, Kjellander P, Johansson Ö, LaPoint S, Bayrakcismith R, Linnell JDC, Zaccaroni M, Jorge MLS, Oshima JEF, Songhurst A, Fischer C, Mc Bride RT, Thompson JJ, Streif S, Sandfort R, Bonenfant C, Drouilly M, Klapproth M, Zinner D, Yarnell R, Stronza A, Wilmott L, Meisingset E, Thaker M,

- Vanak AT, Nicoloso S, Graeber R, Said S, Boudreau MR, Devlin A, Hoogesteijn R, May-Junior JA, Nifong JC, Odden J, Quigley HB, Tortato F, Parker DM, Caso A, Perrine J, Tellaeché C, Zieba F, Zwijacz-Kozica T, Appel CL, Axsom I, Bean WT, Cristescu B, Périquet S, Teichman KJ, Karpanty S, Licoppe A, Menges V, Black K, Scheppers TL, Schai-Braun SC, Azevedo FC, Lemos FG, Payne A, Swanepoel LH, Weckworth BV, Berger A, Bertassoni A, McCulloch G, Šustr P, Athreya V, Bockmuhl D, Casaer J, Ekori A, Melovski D, Richard-Hansen C, Van De Vyver D, Reyna-Hurtado R, Robardet E, Selva N, Sergiel A, Farhadinia MS, Sunde P, Portas R, Ambarli H, Berzins R, Kappeler PM, Mann GK, Pyritz L, Bissett C, Grant T, Steinmetz R, Swedell L, Welch RJ, Armenteras D, Bidder OR, González TM, Rosenblatt A, Kachel S, Balkenhol N. 2019. Right on track? Performance of satellite telemetry in terrestrial wildlife research. *PLOS ONE* 14:e0216223. DOI: 10.1371/journal.pone.0216223.
- Hofmann DD, Cozzi G, Fieberg J. 2024. Methods for implementing integrated step-selection functions with incomplete data. *Movement Ecology* 12:37. DOI: 10.1186/s40462-024-00476-8.
- Holman L, Head ML, Lanfear R, Jennions MD. 2015. Evidence of Experimental Bias in the Life Sciences: Why We Need Blind Data Recording. *PLOS Biology* 13:e1002190. DOI: 10.1371/journal.pbio.1002190.
- Homberger B, Jenni L, Duplain J, Lanz M, Schaub M. 2021. Strong effects of radio-tags, social group and release date on survival of reintroduced grey partridges. *Animal Conservation* 24:677–688. DOI: 10.1111/acv.12673.
- Howe EJ, Buckland ST, Després-Einspenner M-L, Kühl HS. 2017. Distance sampling with camera traps. *Methods in Ecology and Evolution* 8:1558–1565. DOI: 10.1111/2041-210X.12790.
- Hribsek I, Plecas M, Skoric S, Marinkovic S. 2021. First description of movement and ranging behavior of the Griffon vulture (*Gyps fulvus*) from Serbia using GPS satellite tracking. *Archives of Biological Sciences* 73:185–195. DOI: 10.2298/ABS201210013H.
- Huntington-Klein N, Arenas A, Beam E, Bertoni M, Bloem JR, Burli P, Chen N, Grieco P, Ekpe G, Pugatch T, Saavedra M, Stopnitzky Y. 2021. The influence of hidden researcher decisions in applied microeconomics. *Economic Inquiry* 59:944–960. DOI: 10.1111/ecin.12992.
- Ihle M, Kempenaers B, Forstmeier W. 2015. Fitness Benefits of Mate Choice for Compatibility in a Socially Monogamous Species. *PLOS Biology* 13:e1002248. DOI: 10.1371/jour-

nal.pbio.1002248.

- International Union for Conservation of Nature and Natural Resources, International Union for Conservation of Nature and Natural Resources (eds.). 2001. *IUCN red list categories and criteria*. Gland, Switzerland: IUCN–The World Conservation Union.
- Jenkins GB, Beckerman AP, Bellard C, Benítez-López A, Ellison AM, Foote CG, Hufton AL, Lashley MA, Lortie CJ, Ma Z, Moore AJ, Narum SR, Nilsson J, O’Boyle B, Provete DB, Razgour O, Rieseberg L, Riginos C, Santini L, Sibbett B, Peres-Neto PR. 2023. Reproducibility in ecology and evolution: Minimum standards for data and code. *Ecology and Evolution* 13:e9961. DOI: 10.1002/ece3.9961.
- Jennions MD, Møller AP. 2002a. Relationships fade with time: A meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269:43–48. DOI: 10.1098/rspb.2001.1832.
- Jennions MD, Møller AP. 2002b. Publication bias in ecology and evolution: An empirical assessment using the ‘trim and fill’ method. *Biological Reviews of the Cambridge Philosophical Society* 77:211–222. DOI: 10.1017/S1464793101005875.
- Johnson DH. 1980. The Comparison of Usage and Availability Measurements for Evaluating Resource Preference. *Ecology* 61:65–71. DOI: 10.2307/1937156.
- Jones MD, Crane MS, Silva IMS, Artchawakom T, Waengsothorn S, Suwanwaree P, Strine CT, Goode M. 2020. Supposed snake specialist consumes monitor lizards: Diet and trophic implications of king cobra feeding ecology. *Ecology*. DOI: 10.1002/ecy.3085.
- Jones MD, Marshall BM, Smith SN, Crane M, Silva I, Artchawakom T, Suwanwaree P, Waengsothorn S, Wüster W, Goode M, Strine CT. 2022. How do King Cobras move across a major highway? Unintentional wildlife crossing structures may facilitate movement. *Ecology and Evolution* 12. DOI: 10.1002/ece3.8691.
- Joo R, Boone ME, Clay TA, Patrick SC, Clusella-Trullas S, Basille M. 2020. Navigating through the `move` packages for movement. *Journal of Animal Ecology* 89:248–267. DOI: 10.1111/1365-2656.13116.
- Joo R, Picardi S, Boone ME, Clay TA, Patrick SC, Romero-Romero VS, Basille M. 2022. Recent trends in movement ecology of animals and human mobility. *Movement Ecology* 10:26. DOI: 10.1186/s40462-022-00322-9.
- Kamath A, Losos J. 2017. The erratic and contingent progression of research on territoriality: A case study. *Behavioral Ecology and Sociobiology* 71:1–13. DOI: 10.1007/s00265-017-2319-z.

- Kamath A, Losos JB. 2018a. Estimating encounter rates as the first step of sexual selection in the lizard *Anolis sagrei*. *Proceedings of the Royal Society B: Biological Sciences* 285:20172244. DOI: 10.1098/rspb.2017.2244.
- Kamath A, Losos J. 2018b. Reconsidering territoriality is necessary for understanding *Anolis* mating systems. *Behavioral Ecology and Sociobiology* 72:106. DOI: 10.1007/s00265-018-2524-4.
- Kaplan RM, Irvin VL. 2015. Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time. *PLOS ONE* 10:e0132382. DOI: 10.1371/journal.pone.0132382.
- Kay M. 2023a. *ggdist: Visualizations of distributions and uncertainty*. DOI: 10.5281/zenodo.3879620.
- Kay M. 2023b. *tidybayes: Tidy data and geoms for Bayesian models*. DOI: 10.5281/zenodo.1308151.
- Kay GM, Mortelliti A, Tulloch A, Barton P, Florance D, Cunningham SA, Lindenmayer DB. 2016. Effects of past and present livestock grazing on herpetofauna in a landscape-scale experiment. *Conservation Biology* 31:446–458. DOI: 10.1111/cobi.12779.
- Kays R, Davidson SC, Berger M, Bohrer G, Fiedler W, Flack A, Hirt J, Hahn C, Gauggel D, Russell B, Kölzsch A, Lohr A, Partecke J, Quetting M, Safi K, Scharf A, Schneider G, Lang I, Schaeuffelhut F, Landwehr M, Storhas M, Schalkwyk L, Vinciguerra C, Weinzierl R, Wikelski M. 2022. The Movebank system for studying global animal movement and demography. *Methods in Ecology and Evolution* 13:419–431. DOI: 10.1111/2041-210X.13767.
- Kellner KF, Parsons AW, Kays R, Millspaugh JJ, Rota CT. 2022. A Two-Species Occupancy Model with a Continuous-Time Detection Process Reveals Spatial and Temporal Interactions. *Journal of Agricultural, Biological and Environmental Statistics* 27:321–338. DOI: 10.1007/s13253-021-00482-y.
- Kelly CD. 2019. Rate and success of study replication in ecology and evolution. *PeerJ* 7:e7654. DOI: 10.7717/peerj.7654.
- Kelly DJ, Gaughran A, Mullen E, MacWhite T, Maher P, Good M, Marples NM. 2020. Extra Territorial Excursions by European badgers are not limited by age, sex or season. *Scientific Reports* 10:9665. DOI: 10.1038/s41598-020-66809-w.
- Klappstein NJ, Michelot T, Fieberg J, Pedersen EJ, Mills Flemming J. 2024. Step selection functions with non-linear and random effects. *Methods in Ecology and Evolution* 15:1332–1346. DOI: 10.1111/2041-210X.14367.
- Knierim T, Barnes CH, Hodges C. 2017. BUNGARUS FASCIATUS banded krait. DIET

- SCAVENGING. *Herpetological Review* 48:215–218.
- Knierim TK, Marshall BM, Hayes L, Suwanwaree P, Strine CT. 2018. The movements and habitat preferences of a Malayan krait (*Bungarus candidus*) in an agrarian landscape. *The Herpetological Bulletin* 143:30–33.
- Knierim T, Strine CT, Suwanwaree P, Hill III JG. 2019. Spatial ecology study reveals nest attendance and habitat preference of banded kraits (*Bungarus fasciatus*). *Herpetological Bulletin*:6–13. DOI: 10.33256/hb150.613.
- Koirala BK, Dawa Tshering. 2021. Distribution, habitat use, and nesting behavior of the King Cobra (*Ophiophagus hannah*) in the Trashigang Forest Division, Eastern Bhutan. *Reptiles & Amphibians* 28:397–403. DOI: 10.17161/randa.v28i3.15786.
- Kourounis D, Fuchs A, Schenk O. 2018. Towards the next generation of multiperiod optimal power flow solvers. *IEEE Transactions on Power Systems* PP:1–10.
- Kowalczyk R, Zalewski A, Bogumiła J. 2006. Daily movement and territory use by badgers *Meles meles* in Białowieża Primeval Forest, Poland. *Wildlife Biology* 12:385–391. DOI: 10.2981/0909-6396(2006)12[385:DMATUB]2.0.CO;2.
- Kranstauber B, Kays R, Lapoint SD, Wikelski M, Safi K. 2012. A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *Journal of Animal Ecology* 81:738–746. DOI: 10.1111/j.1365-2656.2012.01955.x.
- Kranstauber B, Smolla M, Scharf AK. 2023. *move: Visualizing and analyzing animal track data*.
- Kuch U. 2001. NOTES ON THE DIET OF THE MALAYAN KRAIT, *BUNGARUS CANDIDUS* (LINNAEUS, 1758). *Herpetological Bulletin* 75.
- Lakens D, Adolphi FG, Albers CJ, Anvari F, Apps MAJ, Argamon SE, Baguley T, Becker RB, Benning SD, Bradford DE, Buchanan EM, Caldwell AR, Van Calster B, Carlsson R, Chen S-C, Chung B, Colling LJ, Collins GS, Crook Z, Cross ES, Daniels S, Danielsson H, DeBruine L, Dunleavy DJ, Earp BD, Feist MI, Ferrell JD, Field JG, Fox NW, Friesen A, Gomes C, Gonzalez-Marquez M, Grange JA, Grieve AP, Guggenberger R, Grist J, Harmelen A-L van, Hasselman F, Hochard KD, Hoffarth MR, Holmes NP, Ingre M, Isager PM, Isotalus HK, Johansson C, Juszczak K, Kenny DA, Khalil AA, Konat B, Lao J, Larsen EG, Lodder GMA, Lukavský J, Madan CR, Manheim D, Martin SR, Martin AE, Mayo DG, McCarthy RJ, McConway K, McFarland C, Nio AQX, Nilsson G, Oliveira CL de, Xivry J-JO de, Parsons S, Pfuhl G, Quinn KA, Sakon JJ, Saribay SA, Schneider IK, Selvaraju M, Sjoerds Z, Smith SG, Smits T, Spies JR, Sreekumar V, Steltenpohl CN,

- Stenhouse N, Świątkowski W, Vadillo MA, Van Assen MALM, Williams MN, Williams SE, Williams DR, Yarkoni T, Ziano I, Zwaan RA. 2018. Justify your alpha. *Nature Human Behaviour* 2:168–171. DOI: 10.1038/s41562-018-0311-x.
- Lambert M, Farrar B, Garcia-Pelegriñ E, Reber S, Miller R. 2022. ManyBirds: A multi-site collaborative Open Science approach to avian cognition and behavior research. *Animal Behavior and Cognition* 9:133–152. DOI: 10.26451/abc.09.01.11.2022.
- Land Development Department T. 2017. Land-use data of Thailand.
- Landau WM. 2021b. *Tarchetypes: Archetypes for targets*.
- Landau WM. 2021a. The targets r package: A dynamic make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. *Journal of Open Source Software* 6:2959.
- Lemoine NP, Hoffman A, Felton AJ, Baur L, Chaves F, Gray J, Yu Q, Smith MD. 2016. Underappreciated problems of low replication in ecological field studies. *Ecology* 97:2554–2561. DOI: 10.1002/ecy.1506.
- Levins R, Lewontin R. 1985. *The dialectical biologist*. Harvard University Press.
- Lindgren F, Rue H. 2015. Bayesian spatial modelling with R-INLA. *Journal of Statistical Software* 63:1–25.
- Lindgren F, Rue H, Lindström J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach (with discussion). *Journal of the Royal Statistical Society B* 73:423–498.
- Liu Y, Althoff T, Heer J. 2020. Paths Explored, Paths Omitted, Paths Obscured: Decision Points & Selective Reporting in End-to-End Data Analysis. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, 1–14. DOI: 10.1145/3313831.3376533.
- Looareesuwan S, Viravan C, Warrell DA. 1988. Factors contributing to fatal snake bite in the rural tropics: Analysis of 46 cases in Thailand. *Transactions of The Royal Society of Tropical Medicine and Hygiene* 82:930–934. DOI: 10.1016/0035-9203(88)90046-6.
- Loureiro F, Rosalino LM, Macdonald DW, Santos-Reis M. 2007. Path tortuosity of Eurasian badgers (*Meles meles*) in a heterogeneous Mediterranean landscape. *Ecological Research* 22:837–844. DOI: 10.1007/s11284-006-0325-0.
- Loveridge AJ, Valeix M, Elliot NB, Macdonald DW. 2017. The landscape of anthropogenic mortality: How African lions respond to spatial variation in risk. *Journal of Applied Ecology* 54:815–825. DOI: 10.1111/1365-2664.12794.

- Lüdecke D, Ben-Shachar MS, Patil I, Waggoner P, Makowski D. 2021. performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software* 6:3139. DOI: 10.21105/joss.03139.
- Magowan EA, Maguire IE, Smith S, Redpath S, Marks NJ, Wilson RP, Menzies F, O'Hagan M, Scantlebury DM. 2022. Dead-reckoning elucidates fine-scale habitat use by European badgers *Meles meles*. *Animal Biotelemetry* 10:10. DOI: 10.1186/s40317-022-00282-2.
- Manly B, McDonald L, Thomas DL, McDonald TL, Erickson WP. 2007. *Resource selection by animals: Statistical design and analysis for field studies*. Springer Science & Business Media.
- Many Primates, Altschul DM, Beran MJ, Bohn M, Call J, DeTroy S, Duguid SJ, Egelkamp CL, Fichtel C, Fischer J, Flessert M, Hanus D, Haun DBM, Haux LM, Hernandez-Aguilar RA, Herrmann E, Hopper LM, Joly M, Kano F, Keupp S, Melis AP, Motes Rodrigo A, Ross SR, Sánchez-Amaro A, Sato Y, Schmitt V, Schweinfurth MK, Seed AM, Taylor D, Völter CJ, Warren E, Watzek J. 2019. Establishing an infrastructure for collaboration in primate cognition research. *PLOS ONE* 14:e0223675. DOI: 10.1371/journal.pone.0223675.
- ManyDogs Project MP, Alberghina D, Bray EE, Buchsbaum D, Byosiére S-E, Espinosa J, Gnanadesikan GE, Guran C-NA, Hare E, Horschler DJ, Huber L, Kuhlmeier VA, MacLean EL, Pelgrim MH, Perez B, Ravid-Schurr D, Rothkoff L, Sexton CL, Silver ZA, Stevens JR. 2023. ManyDogs Project: A Big Team Science Approach to Investigating Canine Behavior and Cognition. *Comparative Cognition & Behavior Reviews* 18:59–77. DOI: 10.3819/CCBR.2023.180004.
- Margalida A, Pérez-García JM, Afonso I, Moreno-Opo R. 2016. Spatial and temporal movements in Pyrenean bearded vultures (*Gypaetus barbatus*): Integrating movement ecology into conservation practice. *Scientific Reports* 6:35746. DOI: 10.1038/srep35746.
- Marshall BM, Crane M, Silva I, Strine CT, Jones MD, Hodges CW, Suwanwaree P, Artchawakom T, Waengsothorn S, Goode M. 2020a. No room to roam: King Cobras reduce movement in agriculture. *Movement Ecology* 8:33. DOI: 10.1186/s40462-020-00219-5.
- Marshall BM, Crane M, Silva I, Strine CT, Jones MD, Hodges CW, Suwanwaree P, Taksin Artchawakom, Surachit Waengsothorn, Goode M. 2020b. Data set and code supporting Marshall et al. 2020. No room to roam: King Cobras reduce movement in agriculture. DOI: 10.5281/ZENODO.3666028.
- Marshall BM, Duthie AB. 2022. abmAnimalMovement: An r package for simulating animal

- movement using an agent-based model. *F1000* 0:0.
- Marshall BM, Duthie AB. 2024. A Habitat Selection Multiverse Reveals Largely Consistent Results Despite a Multitude of Analysis Options. *bioRxiv*. DOI: 10.1101/2024.06.19.599733.
- Marshall BM, Strine CT. 2021. Make like a glass frog: In support of increased transparency in herpetology. *The Herpetological Journal* 31:35–45. DOI: 10.31219/osf.io/74frd.
- Marshall BM, Strine CT, Jones MD, Artchawakom T, Silva I, Suwanwaree P, Goode M. 2019. Space fit for a king: Spatial ecology of king cobras (*Ophiophagus hannah*) in Sakaerat Biosphere Reserve, Northeastern Thailand. *Amphibia-Reptilia* 40:163–178. DOI: 10.1163/15685381-18000008.
- Marshall BM, Strine CT, Jones MD, Theodorou A, Amber E, Waengsothorn S, Suwanwaree P, Goode M. 2018. Hits Close to Home: Repeated Persecution of King Cobras (*Ophiophagus hannah*) in Northeastern Thailand. *Tropical Conservation Science* 11:194008291881840. DOI: 10.1177/1940082918818401.
- Martins TG, Simpson D, Lindgren F, Rue H. 2013b. Bayesian computing with INLA: New features. *Computational Statistics and Data Analysis* 67:68–83.
- Martins TG, Simpson D, Lindgren F, Rue H. 2013a. Bayesian computing with INLA: New features. *Computational Statistics and Data Analysis* 67:68–83.
- Michelot T, Blackwell PG, Chamaillé-Jammes S, Matthiopoulos J. 2020. Inference in MCMC step selection models. *Biometrics* 76:438–447. DOI: 10.1111/biom.13170.
- Michelot T, Blackwell PG, Matthiopoulos J. 2019. Linking resource selection and step selection models for habitat preferences in animals. *Ecology* 100:e02452. DOI: 10.1002/ecy.2452.
- Michelot T, Gloaguen P, Blackwell PG, Étienne M. 2019. The Langevin diffusion as a continuous-time model of animal movement and habitat selection. *Methods in Ecology and Evolution* 10:1894–1907. DOI: 10.1111/2041-210X.13275.
- Michelot T, Langrock R, Patterson TA. 2016. moveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. *Methods in Ecology and Evolution* 7:1308–1315. DOI: 10.1111/2041-210X.12578.
- Milleret C, Dey S, Dupont P, Turek D, Valpine P de, Bischof R. 2022. Estimating spatially variable and density-dependent survival using open-population spatial capture-recapture models. *bioRxiv*. DOI: 10.1101/2022.03.04.482982.
- Minchin PR. 1987. Simulation of multidimensional community patterns: Towards a compre-

- hensive model. *Vegetatio* 71:145–156.
- Miranda EBP, Ribeiro RP, Strüßmann C. 2016. The ecology of human-anaconda conflict: A study using internet videos. *Tropical Conservation Science* 9:43–77. DOI: 10.1177/194008291600900105.
- Miyakawa T. 2020. No raw data, no science: Another possible source of the reproducibility crisis. *Molecular Brain* 13:24, s13041-020-0552-2. DOI: 10.1186/s13041-020-0552-2.
- Mueller T, Olson KA, Dressler G, Leimgruber P, Fuller TK, Nicolson C, Novaro AJ, Bolgeri MJ, Wattles D, DeStefano S, Calabrese JM, Fagan WF. 2011. How landscape dynamics link individual- to population-level movement patterns: A multispecies comparison of ungulate relocation data: Population-level movement patterns. *Global Ecology and Biogeography* 20:683–694. DOI: 10.1111/j.1466-8238.2010.00638.x.
- Muff S, Signer J, 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* 89:80–92. DOI: 10.1111/1365-2656.13087.
- Müller K. 2020a. *Here: A simpler way to find your files.*
- Müller K. 2020b. *Here: A simpler way to find your files.*
- Müller K, Wickham H. 2023. *tibble: Simple data frames.*
- Mutascio HE, Pittman SE, Zollner PA, D’Acunto LE. 2018. Modeling relative habitat suitability of southern Florida for invasive Burmese pythons (*Python molurus bivittatus*). *Landscape Ecology* 33:257–274. DOI: 10.1007/s10980-017-0597-5.
- Nakagawa S, Parker TH. 2015. Replicating research in ecology and evolution: Feasibility, incentives, and the cost-benefit conundrum. *BMC Biology* 13:88. DOI: 10.1186/s12915-015-0196-3.
- Nathan R, Getz WM, Revilla E, Holyoak M, Kadmon R, Saltz D, Smouse PE. 2008. A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences* 105:19052–19059. DOI: 10.1073/pnas.0800375105.
- Nathan R, Spiegel O, Fortmann-Roe S, Harel R, Wikelski M, Getz WM. 2012. Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: General concepts and tools illustrated for griffon vultures. *Journal of Experimental Biology* 215:986–996. DOI: 10.1242/jeb.058602.
- Noonan MJ, Fleming CH, Tucker MA, Kays R, Harrison A, Crofoot MC, Abrahms B, Alberts SC, Ali AH, Altmann J, Antunes PC, Attias N, Belant JL, Beyer DE, Bidner LR, Blaum

- N, Boone RB, Caillaud D, Paula RC, Torre JA la, Dekker J, DePerno CS, Farhadinia M, Fennessy J, Fichtel C, Fischer C, Ford A, Goheen JR, Havmøller RW, Hirsch BT, Hurtado C, Isbell LA, Janssen R, Jeltsch F, Kaczensky P, Kaneko Y, Kappeler P, Katna A, Kauffman M, Koch F, Kulkarni A, LaPoint S, Leimgruber P, Macdonald DW, Markham AC, McMahon L, Mertes K, Moorman CE, Morato RG, Moßbrucker AM, Mourão G, O'Connor D, Oliveira-Santos LGR, Pastorini J, Patterson BD, Rachlow J, Ranglack DH, Reid N, Scantlebury DM, Scott DM, Selva N, Sergiel A, Songer M, Songsasen N, Stabach JA, Stacy-Dawes J, Swingen MB, Thompson JJ, Ullmann W, Vanak AT, Thaker M, Wilson JW, Yamazaki K, Yarnell RW, Zieba F, Zwijacz-Kozica T, Fagan WF, Mueller T, Calabrese JM. 2020. Effects of body size on estimation of mammalian area requirements. *Conservation Biology*:cobi.13495. DOI: 10.1111/cobi.13495.
- Noonan MJ, Tucker MA, Fleming CH, Akre T, Alberts SC, Ali AH, Altmann J, Antunes PC, Belant JL, Beyer D, Blaum N, Böhning-Gaese K, Cullen L, Paula RC de, Dekker J, Drescher-Lehman J, Farwig N, Fichtel C, Fischer C, Ford A, Goheen JR, Janssen R, Jeltsch F, Kauffman M, Kappeler PM, Koch F, LaPoint S, Catherine Markham A, Medici EP, Morato RG, Nathan R, Oliveira-Santos LGR, Olson KA, Patterson BD, Paviolo A, Ramalho EE, Roesner S, Schabo D, Selva N, Sergiel A, Xavier da Silva M, Spiegel O, Thompson P, Ullmann W, Zieba F, Zwijacz-Kozica T, Fagan WF, Mueller T, Calabrese JM. 2019. A comprehensive analysis of autocorrelation and bias in home range estimation. *Ecological Monographs*. DOI: 10.1002/ecm.1344.
- Northrup JM, Hooten MB, Anderson CR, Wittemyer G. 2013. Practical guidance on characterizing availability in resource selection functions under a use–availability design. *Ecology* 94:1456–1463. DOI: 10.1890/12-1688.1.
- Northrup JM, Vander Wal E, Bonar M, Fieberg J, Laforge MP, Leclerc M, Prokopenko CM, Gerber BD. 2022. Conceptual and methodological advances in habitat-selection modeling: Guidelines for ecology and evolution. *Ecological Applications* 32. DOI: 10.1002/eap.2470.
- O'Boyle EH, Banks GC, Gonzalez-Mulé E. 2014. The Chrysalis Effect: How Ugly Initial Results Metamorphosize Into Beautiful Articles. *Journal of Management* 43:376–399. DOI: 10.1177/0149206314527133.
- Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science* 349:aac4716–aac4716. DOI: 10.1126/science.aac4716.
- Palmer AR. 2000. Quasi-Replication and the Contract of Error: Lessons from Sex Ratios, Heritabilities and Fluctuating Asymmetry. *Annual Review of Ecology and Systematics*

- 31:441–480. DOI: 10.1146/annurev.ecolsys.31.1.441.
- Parker TH. 2013. What do we really know about the signalling role of plumage colour in blue tits? A case study of impediments to progress in evolutionary biology: Case study of impediments to progress. *Biological Reviews* 88:511–536. DOI: 10.1111/brv.12013.
- Parsons S. 2020. Exploring reliability heterogeneity with multiverse analyses: Data processing decisions unpredictably influence measurement reliability. *PsyArXiv*. DOI: 10.31234/osf.io/y6tcz.
- Patel CJ, Burford B, Ioannidis JPA. 2015. Assessment of vibration of effects due to model specification can demonstrate the instability of observational associations. *Journal of Clinical Epidemiology* 68:1046–1058. DOI: 10.1016/j.jclinepi.2015.05.029.
- Paton RS, Matthiopoulos J. 2015. Defining the scale of habitat availability for models of habitat selection. *Ecology*:14–2241.1. DOI: 10.1890/14-2241.1.
- Pebesma E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* 10:439–446. DOI: 10.32614/RJ-2018-009.
- Pebesma EJ, Bivand RS. 2005a. Classes and methods for spatial data in R. *R News* 5:9–13.
- Pebesma EJ, Bivand R. 2005b. Classes and methods for spatial data in R. *R News* 5:9–13.
- Pebesma E, Bivand R. 2022. *Sp: Classes and methods for spatial data*.
- Pebesma E, Bivand R. 2023. *Spatial Data Science: With applications in R*. Chapman and Hall/CRC.
- Pedersen TL. 2020. *Patchwork: The composer of plots*.
- Pedersen TL. 2021. *Ggforce: Accelerating ggplot2*.
- Pedersen TL. 2024. *Patchwork: The composer of plots*.
- Peshev H, Grozdanov A, Kmetova–Biro E, Ivanov I, Stoyanov G, Tsiakiris R, Marin S, Marinković S, Sušić G, Lisichanets E, Hribšek I, Karić Z, Kapelj S, Bonchev L, Stoynev E. 2021. New insight into spatial ecology of Griffon Vulture (*Gyps fulvus*) on the Balkans provides opportunity for focusing conservation actions for a threatened social scavenger. *Biodiversity Data Journal* 9:e71100. DOI: 10.3897/BDJ.9.e71100.
- Peterson D, Panofsky A. 2021. Self-correction in science: The diagnostic and integrative motives for replication. *Social Studies of Science*.
- Petr M, Haller BC, Ralph PL, Racimo F. 2022. Slendr: A framework for spatio-temporal population genomic simulations on geographic landscapes. DOI: 10.1101/2022.03.20.485041.
- Pochanugool C, Wildde H, Bhanganada K, Chanhom L, Cox MJ, Chaiyabutr N, Sitprija V. 1998. Venomous snakebite in Thailand. II: Clinical experience. *Mil Med* 163:318–323.

- Portugal SJ, White CR. 2022. Externally attached biologgers cause compensatory body mass loss in birds. *Methods in Ecology and Evolution* 13:294–302. DOI: 10.1111/2041-210X.13754.
- Prange S, Gehrt SD, Wiggers EP. 2004. Influences of anthropogenic resources on raccoon (*Procyon lotor*) movements and spatial distribution. *Journal of Mammalogy* 85:8.
- Quaglietta L, Porto M. 2019. SiMRiv: An R package for mechanistic simulation of individual, spatially-explicit multistate movements in rivers, heterogeneous and homogeneous spaces incorporating landscape bias. *Movement Ecology* 7:11. DOI: 10.1186/s40462-019-0154-8.
- Quintana DS. 2020. A synthetic dataset primer for the biobehavioural sciences to promote reproducibility and hypothesis generation. *eLife* 9:e53275. DOI: 10.7554/eLife.53275.
- R Core Team. 2021. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- R Core Team. 2022. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- R Studio Team. 2021. RStudio: Integrated Development Environment for R.
- Rao C, Talukdar G, Choudhury BC, Shankar PG, Whitaker R, Goode M. 2013. Habitat use of King Cobra (*Ophiophagus hannah*) in a heterogeneous landscape matrix in the tropical forests of the Western Ghats , India. *Hamadryad* 36:69–79.
- Receveur A, Bonfanti J, D’Agata S, Helmstetter AJ, Moore NA, Oliveira BF, Petit-Cailleux C, Rievers Borges E, Schultz M, Sexton AN, Veytia D. 2024. David versus Goliath: Early career researchers in an unethical publishing system. *Ecology Letters* 27:e14395. DOI: 10.1111/ele.14395.
- Reinert HK, Cundall D. 1982. An Improved Surgical Implantation Method for Radio-Tracking Snakes. *Copeia* 1982:702. DOI: 10.2307/1444674.
- Rijnhart JJM, Twisk JWR, Deeg DJH, Heymans MW. 2021. Assessing the Robustness of Mediation Analysis Results Using Multiverse Analysis. *Prevention Science*. DOI: 10.1007/s11121-021-01280-1.
- Rivera K, Fidino M, Farris ZJ, Magle SB, Murphy A, Gerber BD. 2022. Rethinking habitat occupancy modeling and the role of diel activity in an anthropogenic world. *The American Naturalist*. DOI: 10.1086/720714.
- Robstad CA, Lodberg-Holm HK, Mayer M, Rosell F. 2021. The impact of bio-logging on body weight change of the Eurasian beaver. *PLOS ONE* 16:e0261453. DOI: 10.1371/journal.pone.0261453.

- Roche DG, Amcoff M, Morgan R, Sundin J, Finnøen MH, Lawrence MJ, Henderson E, Speers-Roesch B, Brown C, Clark TD, Bshary R, Jutfelt F, Binning SA. 2020. Behavioural lateralisation in a detour test is not repeatable in fishes. *EcoEvoRxiv*:62. DOI: 10.32942/osf.io/6kcwa.
- Roche DG, Berberi I, Dhane F, Lauzon F, Soeharjono S, Dakin R, Binning SA. 2022. Slow improvement to the archiving quality of open datasets shared by researchers in ecology and evolution. *Proceedings of the Royal Society B: Biological Sciences* 289:20212780. DOI: 10.1098/rspb.2021.2780.
- Roche DG, Kruuk LEB, Lanfear R, Binning SA. 2015. Public Data Archiving in Ecology and Evolution: How Well Are We Doing? *PLOS Biology* 13:e1002295. DOI: 10.1371/journal.pbio.1002295.
- Rosalino LM, Macdonald DW, Santos-Reis M. 2005. Activity rhythms, movements and patterns of sett use by badgers, *Meles meles*, in a Mediterranean woodland. *Mammalia* 69. DOI: 10.1515/mamm.2005.031.
- RStudio Team. 2022. *RStudio: Integrated development environment for r*. Boston, MA: RStudio, PBC.
- Rue H, Martino S, Chopin N. 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society B* 71:319–392.
- Rue H, Riebler AI, Sørbye SH, Illian JB, Simpson DP, Lindgren FK. 2017. Bayesian computing with INLA: A review. *Annual Reviews of Statistics and Its Applications* 4:395–421.
- Salis A, Lena J-P, Lengagne T. 2021. How Subtle Protocol Choices Can Affect Biological Conclusions: Great Tits’ Response to Allopatric Mobbing Calls. *Animal Behavior and Cognition* 8:152–165. DOI: 10.26451/abc.08.02.05.2021.
- Samuel SP, Chinnaraju S, Williams HF, Pichamuthu E, Subharao M, Vaiyapuri M, Arumugam S, Vaiyapuri R, Baksh MF, Patel K, Trim SA, Duncombe TE, Vaiyapuri S. 2020. Venomous snakebites: Rapid action saves lives—A multifaceted community education programme increases awareness about snakes and snakebites among the rural population of Tamil Nadu, India. *PLOS Neglected Tropical Diseases* 14:e0008911. DOI: 10.1371/journal.pntd.0008911.
- Sánchez-Tójar A, Nakagawa S, Sánchez-Fortún M, Martin DA, Ramani S, Girndt A, Bókony V, Kempenaers B, Liker A, Westneat DF, Burke T, Schroeder J. 2018. Meta-analysis challenges a textbook example of status signalling and demonstrates publication bias.

- eLife* 7:e37385. DOI: 10.7554/eLife.37385.
- Saunders D, Nguyen H, Cowen S, Magrath M, Marsh K, Bell S, Bobruk J. 2022. Radio-tracking wildlife with drones: A viewshed analysis quantifying survey coverage across diverse landscapes. *Wildlife Research* 49:1–10. DOI: 10.1071/WR21033.
- Scharf AK, Belant JL, Beyer DE, Wikelski M, Safi K. 2018. Habitat suitability does not capture the essence of animal-defined corridors. *Movement Ecology* 6:18. DOI: 10.1186/s40462-018-0136-2.
- Scheel AM, Schijen MRMJ, Lakens D. 2021. An Excess of Positive Results: Comparing the Standard Psychology Literature With Registered Reports. *Advances in Methods and Practices in Psychological Science* 4:251524592110074. DOI: 10.1177/25152459211007467.
- Schlather M, Malinowski A, Menck PJ, Oesting M, Stokorb K. 2015. Analysis, simulation and prediction of multivariate random fields with package RandomFields. *Journal of Statistical Software* 63:1–25.
- Schnitzer SA, Carson WP. 2016. Would Ecology Fail the Repeatability Test? *BioScience* 66:98–99. DOI: 10.1093/biosci/biv176.
- Sciaini M, Fritsch M, Hesselbarth M, Simpkins C, Scherer C, Hanß S. 2021. *NLMR: Simulating neutral landscape models*.
- Sciaini M, Fritsch M, Scherer C, Simpkins CE. 2018a. NLMR and landscapetools: An integrated environment for simulating and modifying neutral landscape models in r. *Methods in Ecology and Evolution* 00:1–9.
- Sciaini M, Fritsch M, Scherer C, Simpkins CE. 2018b. NLMR and landscapetools : An integrated environment for simulating and modifying neutral landscape models in R. *Methods in Ecology and Evolution* 9:2240–2248. DOI: 10.1111/2041-210X.13076.
- Sciaini M, Fritsch M, Scherer C, Simpkins CE. 2018c. NLMR and landscapetools : An integrated environment for simulating and modifying neutral landscape models in R. *Methods in Ecology and Evolution* 9:2240–2248. DOI: 10.1111/2041-210X.13076.
- SEARO. 2016. *Guidelines for the management of snakebites, 2nd edition*. SEARO Regional Office for the South East Asia (RGO); WHO South-East Asia.
- Seguin A, Forstmeier W. 2012. No Band Color Effects on Male Courtship Rate or Body Mass in the Zebra Finch: Four Experiments and a Meta-Analysis. *PLoS ONE* 7:e37785. DOI: 10.1371/journal.pone.0037785.
- Shankar PG, Ganesh SR, Whitaker R, Prashanth P. 2013a. King Cobra *Ophiophagus hannah* (Cantor, 1836) encounters in human-modified rainforests of the Western Ghats, India.

Hamadryad 36:62–68.

- Shankar PG, Singh A, Ganesh SR, Whitaker R. 2013b. Factors influencing human hostility to King Cobras (*Ophiophagus hannah*) in the Western Ghats of India. *Hamadryad* 36:91–100.
- Shankar PG, Swamy P, Williams RC, Ganesh SR, Moss M, Höglund J, Das I, Sahoo G, Vijayakumar SP, Shanker K, Wüster W, Dutta SK. 2021. King or royal family? Testing for species boundaries in the King Cobra, *Ophiophagus hannah* (Cantor, 1836), using morphology and multilocus DNA analyses. *Molecular Phylogenetics and Evolution*:107300. DOI: 10.1016/j.ympev.2021.107300.
- Siers SR, Yackel Adams AA, Reed RN. 2018. Behavioral differences following ingestion of large meals and consequences for management of a harmful invasive snake: A field experiment. *Ecology and Evolution* 8:10075–10093. DOI: 10.1002/ece3.4480.
- Signer J, Fieberg J, Avgar T. 2019. Animal movement tools (amt): R package for managing tracking data and conducting habitat selection analyses. *Ecology and Evolution* 9:880–890.
- Silberzahn R, Uhlmann EL, Martin DP, Anselmi P, Aust F, Awtrey E, Bahník Š, Bai F, Bannard C, Bonnier E, Carlsson R, Cheung F, Christensen G, Clay R, Craig MA, Dalla Rosa A, Dam L, Evans MH, Flores Cervantes I, Fong N, Gamez-Djokic M, Glenz A, Gordon-McKeon S, Heaton TJ, Hederos K, Heene M, Hofelich Mohr AJ, Högden F, Hui K, Johannesson M, Kalodimos J, Kaszubowski E, Kennedy DM, Lei R, Lindsay TA, Liverani S, Madan CR, Molden D, Molleman E, Morey RD, Mulder LB, Nijstad BR, Pope NG, Pope B, Prenoveau JM, Rink F, Robusto E, Roderique H, Sandberg A, Schlüter E, Schönbrodt FD, Sherman MF, Sommer SA, Sotak K, Spain S, Spörlein C, Stafford T, Stefanutti L, Tauber S, Ullrich J, Vianello M, Wagenmakers E-J, Witkowiak M, Yoon S, Nosek BA. 2018. Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results. *Advances in Methods and Practices in Psychological Science* 1:337–356. DOI: 10.1177/2515245917747646.
- Silva R, Afán I, Gil JA, Bustamante J. 2017. Seasonal and circadian biases in bird tracking with solar GPS-tags. *PLOS ONE* 12:e0185344. DOI: 10.1371/journal.pone.0185344.
- Silva I, Crane M, Marshall BM, Strine CT. 2020. Reptiles on the wrong track? Moving beyond traditional estimators with dynamic Brownian Bridge Movement Models. *Movement Ecology* 8:43. DOI: 10.1186/s40462-020-00229-3.
- Silva I, Crane M, Suwanwaree P, Strine C, Goode M. 2018. Using dynamic Brownian Bridge Movement Models to identify home range size and movement patterns in king cobras.

- PLOS ONE* 13:e0203449. DOI: 10.1371/journal.pone.0203449.
- Silva I, Fleming CH, Noonan MJ, Alston J, Foltá C, Fagan WF, Calabrese JM. 2022. Autocorrelation-informed home range estimation: A review and practical guide. *Methods in Ecology and Evolution* 13:534–544. DOI: 10.1111/2041-210X.13786.
- Silva I, Fleming CH, Noonan MJ, Fagan WF, Calabrese JM. 2023. Movedesign: Shiny R app to evaluate sampling design for animal movement studies. *Methods in Ecology and Evolution* 14:2216–2225. DOI: 10.1111/2041-210X.14153.
- Simmons JP, Nelson LD, Simonsohn U. 2011. False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science* 22:1359–1366. DOI: 10.1177/0956797611417632.
- Simonsohn U. 2015. Small Telescopes: Detectability and the Evaluation of Replication Results. *Psychological Science* 26:559–569. DOI: 10.1177/0956797614567341.
- Simonsohn U, Simmons JP, Nelson LD. 2020. Specification curve analysis. *Nature Human Behaviour* 4:1208–1214. DOI: 10.1038/s41562-020-0912-z.
- Smaldino PE, McElreath R. 2016. The natural selection of bad science. *Royal Society Open Science* 3:160384. DOI: 10.1098/rsos.160384.
- Smith BJ, Hart KM, Mazzotti FJ, Basille M, Romagosa CM. 2018. Evaluating GPS biologging technology for studying spatial ecology of large constricting snakes. *Animal Biotelemetry* 6:1. DOI: 10.1186/s40317-018-0145-3.
- Smith SN, Jones MD, Marshall BM, Waengsothorn S, Gale GA, Strine CT. 2020. Dataset and code for supporting Smith et al. 2020. Restricted constrictors: Space use and habitat selection of native Burmese pythons in Northeast Thailand. DOI: 10.5281/ZENODO.4026927.
- Smith SN, Jones MD, Marshall BM, Waengsothorn S, Gale GA, Strine CT. 2021. Native Burmese pythons exhibit site fidelity and preference for aquatic habitats in an agricultural mosaic. *Scientific Reports* 11:7014. DOI: 10.1038/s41598-021-86640-1.
- Smith SN, Stavish J, Glosenger-Thrasher S, Gale G, Waengsothorn S. 2024. Risk versus reward: Burmese python mothers select precarious oviposition sites. *Ecology*:e4411. DOI: 10.1002/ecy.4411.
- Sperry JH, Butler LK, Romero LM, Weatherhead PJ. 2009. Effects of parasitic infection and radio-transmitters on condition, hematological characteristics and corticosterone concentrations in Texas ratsnakes. *Journal of Zoology* 278:100–107. DOI: 10.1111/j.1469-7998.2009.00549.x.
- Sridhar VH, Li L, Gorbonos D, Nagy M, Schell BR, Sorochkin T, Gov NS, Couzin ID. 2021.

- The geometry of decision-making in individuals and collectives. *Proceedings of the National Academy of Sciences* 118:e2102157118. DOI: 10.1073/pnas.2102157118.
- Steege S, Tuerlinckx F, Gelman A, Vanpaemel W. 2016. Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science* 11:702–712. DOI: 10.1177/1745691616658637.
- Street GM, Potts JR, Börger L, Beasley JC, Demarais S, Fryxell JM, McLoughlin PD, Monteith KL, Prokopenko CM, Ribeiro MC, Rodgers AR, Strickland BK, Beest FM, Bernasconi DA, Beumer LT, Dharmarajan G, Dwinnell SP, Keiter DA, Keuroghlian A, Newediuk LJ, Oshima JEF, Rhodes O, Schlichting PE, Schmidt NM, Vander Wal E. 2021. Solving the sample size problem for resource selection functions. *Methods in Ecology and Evolution* 12:2421–2431. DOI: 10.1111/2041-210X.13701.
- Strine C, Silva I, Barnes CH, Marshall BM, Artchawakom T, Hill J, Suwanwaree P. 2018. Spatial ecology of a small arboreal ambush predator, *Trimeresurus macrops* Kramer, 1977, in Northeast Thailand. *Amphibia-Reptilia* 39:335–345. DOI: 10.1163/15685381-17000207.
- Stuart B, Wogan G, Grismer L, Auliya M, Inger RF, Lilley R, Chan-Ard T, Thy N, Nguyen TQ, Srinivasulu C, Jelić D. 2012. *Ophiophagus hannah*, King Cobra. *The IUCN Red List of Threatened Species 2012*.
- Stuber EF, Carlson BS, Jesmer BR. 2022. Spatial personalities: A meta-analysis of consistent individual differences in spatial behavior. *Behavioral Ecology* 33:477–486. DOI: 10.1093/beheco/arab147.
- Studd EK, Derbyshire RE, Menzies AK, Simms JF, Humphries MM, Murray DL, Boutin S. 2021. The Purr-fect Catch: Using accelerometers and audio recorders to document kill rates and hunting behaviour of a small prey specialist. *Methods in Ecology and Evolution* 12:1277–1287. DOI: 10.1111/2041-210X.13605.
- Subedi TR, Pérez-García JM, Sah SAM, Gurung S, Baral HS, Poudyal LP, Lee H, Thomsett S, Virani MZ, Anadón JD. 2020. Spatial and temporal movement of the Bearded Vulture using GPS telemetry in the Himalayas of Nepal. *Ibis* 162:563–571. DOI: 10.1111/ibi.12799.
- Tam J, Lagisz M, Cornwell WK, Nakagawa S. 2021. Quantifying research interests in 7,521 mammalian species with h-index: A case study. *EcoEvoRxiv*. DOI: 10.32942/osf.io/gd7cv.
- Tang W, Bennett DA. 2010. Agent-based Modeling of Animal Movement: A Review. *Geography Compass* 4:682–700. DOI: 10.1111/j.1749-8198.2010.00337.x.
- Tang-Martínez Z. 2020. The history and impact of women in animal behaviour and the ABS: A North American perspective. *Animal Behaviour* 164:251–260. DOI:

- 10.1016/j.anbehav.2019.12.011.
- Tedersoo L, Küngas R, Oras E, Köster K, Eenmaa H, Leijen Ä, Pedaste M, Raju M, Astapova A, Lukner H, Kogermann K, Sepp T. 2021. Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data* 8:192. DOI: 10.1038/s41597-021-00981-0.
- Theng M, Milleret C, Bracis C, Cassey P, Delean S. 2022. Confronting spatial capture–recapture models with realistic animal movement simulations. *Ecology* n/a:e3676. DOI: 10.1002/ecy.3676.
- Thurfjell H, Ciuti S, Boyce MS. 2014. Applications of step-selection functions in ecology and conservation. *Movement Ecology* 2. DOI: 10.1186/2051-3933-2-4.
- Tomotani BM, Muijres FT, Johnston B, Jeugd HP, Naguib M. 2021. Great tits do not compensate over time for a radio-tag-induced reduction in escape-flight performance. *Ecology and Evolution* 11:16600–16617. DOI: 10.1002/ece3.8240.
- Tucker MA, Alexandrou O, Bierregaard RO, Bildstein KL, Böhning-Gaese K, Bracis C, Brzorad JN, Buechley ER, Cabot D, Calabrese JM, Carrapato C, Chiaradia A, Davenport LC, Davidson SC, Desholm M, DeSorbo CR, Domenech R, Enggist P, Fagan WF, Farwig N, Fiedler W, Fleming CH, Franke A, Fryxell JM, García-Ripollés C, Grémillet D, Griffin LR, Harel R, Kane A, Kays R, Kleyheeg E, Lacy AE, LaPoint S, Limiñana R, López-López P, Maccarone AD, Mellone U, Mojica EK, Nathan R, Newman SH, Noonan MJ, Oppel S, Prostor M, Rees EC, Ropert-Coudert Y, Rösner S, Sapir N, Schabo D, Schmidt M, Schulz H, Shariati M, Shreading A, Paulo Silva J, Skov H, Spiegel O, Takekawa JY, Teitelbaum CS, Toor ML van, Urios V, Vidal-Mateo J, Wang Q, Watts BD, Wikelski M, Wolter K, Żydelis R, Mueller T. 2019. Large birds travel farther in homogeneous environments. *Global Ecology and Biogeography* 28:576–587. DOI: 10.1111/geb.12875.
- Tucker MA, Böhning-Gaese K, Fagan WF, Fryxell JM, Van Moorter B, Alberts SC, Ali AH, Allen AM, Attias N, Avgar T, Bartlam-Brooks H, Bayarbaatar B, Belant JL, Bertassoni A, Beyer D, Bidner L, Beest FM van, Blake S, Blaum N, Bracis C, Brown D, Bruyn PJN de, Cagnacci F, Calabrese JM, Camilo-Alves C, Chamaillé-Jammes S, Chiaradia A, Davidson SC, Dennis T, DeStefano S, Diefenbach D, Douglas-Hamilton I, Fennessy J, Fichtel C, Fiedler W, Fischer C, Fischhoff I, Fleming CH, Ford AT, Fritz SA, Gehr B, Goheen JR, Gurarie E, Hebblewhite M, Heurich M, Hewison AJM, Hof C, Hurme E, Isbell LA, Janssen R, Jeltsch F, Kaczensky P, Kane A, Kappeler PM, Kauffman M, Kays R, Kimuyu D, Koch F, Kranstauber B, LaPoint S, Leimgruber P, Linnell JDC, López-López

- P, Markham AC, Mattisson J, Medici EP, Mellone U, Merrill E, Miranda Mourão G de, Morato RG, Morellet N, Morrison TA, Díaz-Muñoz SL, Mysterud A, Nandintsetseg D, Nathan R, Niamir A, Odden J, O'Hara RB, Oliveira-Santos LGR, Olson KA, Patterson BD, Cunha de Paula R, Pedrotti L, Reineking B, Rimmner M, Rogers TL, Rolandsen CM, Rosenberry CS, Rubenstein DI, Safi K, Saïd S, Sapir N, Sawyer H, Schmidt NM, Selva N, Sergiel A, Shiilegdamba E, Silva JP, Singh N, Solberg EJ, Spiegel O, Strand O, Sundaresan S, Ullmann W, Voigt U, Wall J, Wattles D, Wikelski M, Wilmers CC, Wilson JW, Wittemyer G, Zięba F, Zwijacz-Kozica T, Mueller T. 2018. Moving in the Anthropocene: Global reductions in terrestrial mammalian movements. *Science* 359:466–469. DOI: 10.1126/science.aam9712.
- Vales DJ, Nielson RM, Middleton MP. 2022. Black-tailed deer seasonal habitat selection: Accounting for missing global positioning system fixes. *The Journal of Wildlife Management* 86:e22305. DOI: 10.1002/jwmg.22305.
- Van Calster H, Oldoni D, Van Hoey S. 2023. *Inborutils: Collection of useful r utilities. Version 0.3.0.*
- Van Moorter B, Visscher D, Benhamou S, Börger L, Boyce MS, Gaillard JM. 2009. Memory keeps you at home: A mechanistic model for home range emergence. *Oikos* 118:641–652. DOI: 10.1111/j.1600-0706.2008.17003.x.
- Vines TH, Albert AYK, Andrew RL, Débarre F, Bock DG, Franklin MT, Gilbert KJ, Moore J-S, Renaut S, Rennison DJ. 2014. The Availability of Research Data Declines Rapidly with Article Age. *Current Biology* 24:94–97. DOI: 10.1016/j.cub.2013.11.014.
- Vinkers CH, Tjeldink JK, Otte WM. 2015. Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: Retrospective analysis. *BMJ*:h6467. DOI: 10.1136/bmj.h6467.
- Viravan C, Looareesuwan S, Kosakam W, Wuthiekanun V, McCarthy CJ, Stimson AF, Bun-nag D, Harinasuta T, Warrell DA. 1992. A national hospital-based survey of snakes responsible for bites in Thailand. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 86:100–106. DOI: 10.1016/0035-9203(92)90463-M.
- Vogt K, Vimercati E, Ryser A, Hofer E, Signer S, Signer C, Breitenmoser U. 2018. Suitability of GPS telemetry for studying the predation of Eurasian lynx on small- and medium-sized prey animals in the Northwestern Swiss Alps. *European Journal of Wildlife Research* 64:73. DOI: 10.1007/s10344-018-1225-7.
- Wang D, Forstmeier W, Ihle M, Khadraoui M, Jerónimo S, Martin K, Kempenaers B. 2018.

- Irreproducible text-book “knowledge”: The effects of color bands on zebra finch fitness: COLOR BANDS HAVE NO EFFECT ON FITNESS IN ZEBRA FINCHES. *Evolution* 72:961–976. DOI: 10.1111/evo.13459.
- Wang D, Forstmeier W, Kempenaers B. 2017. No mutual mate choice for quality in zebra finches: Time to question a widely held assumption: NO MUTUAL MATE CHOICE FOR QUALITY IN ZEBRA FINCHES. *Evolution* 71:2661–2676. DOI: 10.1111/evo.13341.
- Wang D, Kempenaers N, Kempenaers B, Forstmeier W. 2017. Male zebra finches have limited ability to identify high-fecundity females. *Behavioral Ecology* 28:784–792. DOI: 10.1093/beheco/axx037.
- Ware JJ, Munafò MR. 2015. Significance chasing in research practice: Causes, consequences and possible solutions: Significance chasing. *Addiction* 110:4–8. DOI: 10.1111/add.12673.
- Weatherhead PJ, Blouin-Demers G. 2004. Long-term effects of radiotelemetry on black rat-snakes. *Wildlife Society Bulletin* 32:900–906. DOI: 10.2193/0091-7648(2004)032[0900:LEOROB]2.0.CO;2.
- Weaver SJ, Westphal MF, Taylor EN. 2021. Technology wish lists and the significance of temperature-sensing wildlife telemetry. *Animal Biotelemetry* 9:29. DOI: 10.1186/s40317-021-00252-0.
- Webb SS, Demeyere N. 2021. Multiverse Analysis: A Method to Determine Researcher Degrees of Freedom in Test Validation. *PsyArXiv*. DOI: 10.31234/osf.io/nhrwq.
- Webster MM, Rutz C. 2020. How STRANGE are your study animals? *Nature* 582:337–340. DOI: 10.1038/d41586-020-01751-5.
- Whitaker R, Captain A. 2004. King Cobra Ophiophagous hannah. In: *Snakes of India - A Field Guide*. Chennai, India: Draco Books, 384–385.
- Wickham H. 2007a. Reshaping data with the reshape package. *Journal of Statistical Software* 21:1–20.
- Wickham H. 2007b. Reshaping data with the reshape package. *Journal of Statistical Software* 21:1–20.
- Wickham H. 2016a. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.
- Wickham H. 2016b. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.
- Wickham H. 2022. *stringr: Simple, consistent wrappers for common string operations*.
- Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, Woo K, Yutani H, Dunnington D. 2022a. *ggplot2: Create elegant data visualisations using the grammar of graphics*.
- Wickham H, François R, Henry L, Müller K. 2022b. *Dplyr: A grammar of data manipulation*.

- Wickham H, François R, Henry L, Müller K, Vaughan D. 2023. *dplyr: A grammar of data manipulation*.
- Wickham H, Hester J, Bryan J. 2024. *Readr: Read rectangular text data*.
- Wickham H, Hester J, Chang W, Bryan J. 2021. *Devtools: Tools to make developing r packages easier*.
- Wickham H, Vaughan D, Girlich M. 2023. *tidyr: Tidy messy data*.
- Wild TA, Wikelski M, Tyndel S, Alarcón-Nieto G, Klump BC, Aplin LM, Meboldt M, Williams HJ. 2022. Internet on animals: Wi-Fi-enabled devices provide a solution for big data transmission in biologging. *Methods in Ecology and Evolution*:2041–210X.13798. DOI: 10.1111/2041-210X.13798.
- Wilke CO. 2020. *Ggtext: Improved text rendering support for ggplot2*.
- Wilke CO. 2021. *Ggridges: Ridgeline plots in ggplot2*.
- Wilke CO. 2024. *Ggridges: Ridgeline plots in 'ggplot2'*.
- Wilke CO, Wiernik BM. 2022. *Ggtext: Improved text rendering support for 'ggplot2'*.
- Williams HJ, Taylor LA, Benhamou S, Bijleveld AI, Clay TA, Grissac S, Demšar U, English HM, Franconi N, Gómez-Laich A, Griffiths RC, Kay WP, Morales JM, Potts JR, Rogerson KF, Rutz C, Spelt A, Trevail AM, Wilson RP, Börger L. 2020. Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology* 89:186–206. DOI: 10.1111/1365-2656.13094.
- Wilmers CC, Nickel B, Bryce CM, Smith JA, Wheat RE, Yovovich V. 2015. The golden age of bio-logging: How animal-borne sensors are advancing the frontiers of ecology. *Ecology* 96:1741–1753. DOI: 10.1890/14-1401.1.
- Winne CT, Willson JD, Andrews KM, Reed RN. 2006. Efficacy of marking snakes with disposable medical cautery units. *Herpetological Review* 37:52–54.
- Xie Y. 2014. Knitr: A comprehensive tool for reproducible research in R. In: Stodden V, Leisch F, Peng RD eds. *Implementing reproducible computational research*. Chapman; Hall/CRC,.
- Xie Y. 2015. *Dynamic documents with R and knitr*. Boca Raton, Florida: Chapman; Hall/CRC.
- Xie Y. 2016. *Bookdown: Authoring books and technical documents with R markdown*. Boca Raton, Florida: Chapman; Hall/CRC.
- Xie Y. 2019. TinyTeX: A lightweight, cross-platform, and easy-to-maintain LaTeX distribution based on TeX live. *TUGboat* 40:30–32.

- Xie Y. 2021. *Knitr: A general-purpose package for dynamic report generation in r*.
- Xie Y. 2022b. *Tinytex: Helper functions to install and maintain TeX live, and compile LaTeX documents*.
- Xie Y. 2022a. *Bookdown: Authoring books and technical documents with r markdown*.
- Xie Y. 2023b. *knitr: A general-purpose package for dynamic report generation in r*.
- Xie Y. 2023a. *tinytex: Helper functions to install and maintain TeX live, and compile LaTeX documents*.
- Xie Y, Allaire JJ, Grolemond G. 2018. *R markdown: The definitive guide*. Boca Raton, Florida: Chapman; Hall/CRC.
- Xie Y, Dervieux C, Riederer E. 2020. *R markdown cookbook*. Boca Raton, Florida: Chapman; Hall/CRC.
- Young C, Holsteen K. 2017. Model Uncertainty and Robustness: A Computational Framework for Multimodel Analysis. *Sociological Methods & Research* 46:3–40. DOI: 10.1177/0049124115610347.
- Zhu H. 2021. *kableExtra: Construct complex table with 'kable' and pipe syntax*.