UCC Library and UCC researchers have made this item openly available. Please let us know how this has helped you. Thanks!

Title	Uncertainty analysis of step-selection functions: The effect of model parameters on inferences about the relationship between animal movement and the environment
Author(s)	Holloway, Paul; Miller, Jennifer A.
Publication date	2014-09
Original citation	Holloway, P. and Miller, J. A. (2014) 'Uncertainty Analysis of Step-Selection Functions: The Effect of Model Parameters on Inferences about the Relationship between Animal Movement and the Environment', GioScience 2014, Lecture Notes in Computer Science, LCNS 8728, pp. 48-63. doi: 10.1007/978-3-319-11593-1_4
Type of publication	Article (peer-reviewed) Conference item
Link to publisher's version	http://dx.doi.org/10.1007/978-3-319-11593-1_4 Access to the full text of the published version may require a subscription.
Rights	**© Springer International Publishing Switzerland 2014. This is a post-peer-review, pre-copyedit version of an article published in Lecture Notes in Computer Science. The final authenticated version is available online at: http://dx.doi.org/10.1007/978-3-319-11593-1_4**
Item downloaded from	http://hdl.handle.net/10468/9176

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Uncertainty analysis of step-selection functions – the effect of model parameters on inferences about the relationship between animal movement and the environment

Paul Holloway and Jennifer A. Miller

Department of Geography and the Environment, University of Texas at Austin

305 E. 23rd Street, CLA 3.306, A3100, Austin, TX 78712, USA

paul_holloway@utexas.edu; jennifer.miller@austin.utexas.edu

Abstract. As spatio-temporal movement data is becoming more widely available for analysis in GIS and related areas, new methods to analyze them have been developed. A step-selection function (SSF) is a recently developed method used to quantify the effect of environmental factors on animal movement. This method is gaining traction as an important conservation tool; however there have been no studies that have investigated the uncertainty associated with subjective model decisions. In this research we used two types of animals – oilbirds and hyenas – to examine how systematically altering user decisions of model parameters influences the main outcome of an SSF, the coefficients that quantify the movement-environment relationship. We found that user decisions strongly influence the results of step-selection functions and any subsequent inferences about animal movement and environmental interactions. Differences were found between categories for every variable used in the analysis and the results presented here can help to clarify the sources of uncertainty in SSF model decisions.

Keywords: Step-Selection Functions · Uncertainty Analysis · Animal Movement · Conditional Logistic Regression

Introduction

Movement data is now abundant in GIScience, resulting in applied research to extract patterns and processes from the underlying phenomena. While early applications have involved using data rich travel diaries to explore the relationship between human movement and space [1], there have also been questions related to movement ecology and animal movement [2]. The need to develop techniques to analyze the vast amount of spatio-temporal data was the driving force behind the GIScience sub-field of movement pattern analysis (MPA) [3]. The issue of uncertainty has been a long-standing research focus in GIScience, and uncertainty analysis is now considered a prerequisite for model building [4]. However, the field of MPA has been slower to explore such uncertainties (but see [5]). Uncertainty research within the MPA domain has focused on changes in the temporal scale of movement data and how this affects the calculation of movement parameters [5]. However, there is also considerable uncertainty associated with the statistical methods used to analyze movement data and their outputs that has not been investigated.

Step-selection function (SSF) is a powerful new spatial modeling approach that has been developed as an extension of resource-selection function (RSF – a function that is proportional to the probability of the use of a resource unit by an organism –[6]), and is beginning to gain traction as an important tool for studying conservation issues associated with animal movement. SSF was developed by Fortin et al. [7] by combining several methods in order to improve the ability to model resource availability in a home range. These concepts improved upon previous models that did not limit resource availability to an accessible distance of current animal location [8]. Whereas RSF considers the location of an observation, SSF considers the step between two locations. The

observed step between two successive locations is compared to a number of alternatively generated steps that the animal could have taken (Fig. 1), and the coefficients from the case-control regression identify which environmental variables characterize the movement steps actually taken. The majority of SSF studies have estimated SSF in the form of:

$$\hat{w}(x) = \exp\llbracket (\beta\_1 x\_1 + \beta\_2 x\_2 + ... + \beta\_n x\_n)\rrbracket \qquad (1)$$

where βn is the coefficient estimated by the conditional logistic regression for the variable x_n. Steps with higher SSF scores $\hat{w}(x)$ have a higher likelihood of being chosen by the animal, meaning that SSF can help to identify the influence of the environment on animal movements by revealing where they are most likely to be at the end of a movement step.
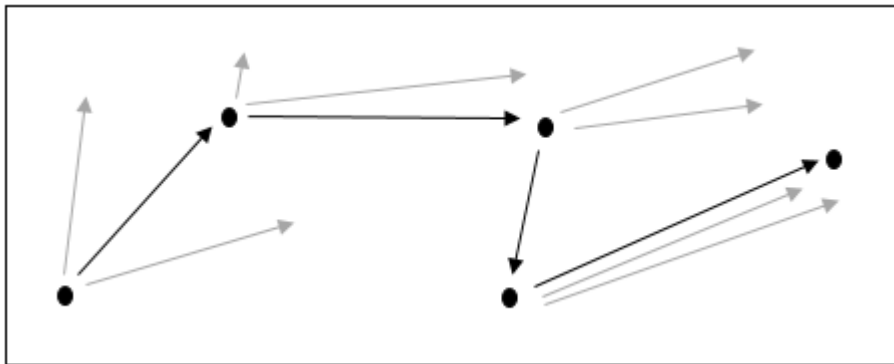

Fig. 1. Conceptual depiction of step-selection functions. Black dots represent successive telemetry locations of an individual, with black arrows representing observed steps. The grey arrows represent the available steps which the individual could have taken.

Since its development, SSF has been used to study a wide range of species, from wrens [9] to wolves [10], and is increasing in its popularity among ecologists due to the power of SSF to identify the influence of environmental variables on the movement of animals. Fitting an SSF involves a number of phases, and we addressed what we considered to be some of the most important and subjective decisions, most of which were described as potentially important by a SSF review paper that was published recently [11]. With SSF currently being used to inform researchers and conservationists about the relationship between animal movement and the environment, the uncertainty associated with the model building phases needs to be explored so that any applications and interpretations can be made with as much confidence as possible.

Method of generating available steps

As SSFs compare used versus available steps, the method used to generate the available steps could potentially have the most influence on the results [11]. The method introduced by Fortin et al. [7] is still the most commonly applied and involves generating available steps from an empirical distribution of movement parameters (turn angle and step length) from other monitored individuals. Empirical distributions are classified in frequency tables of varying interval size, and draws are then made based on a rejection algorithm, whereby an interval is randomly selected from the table alongside a random number from a uniform distribution. If the random number is less than or equal to the specified probability value, a value from that interval is returned [12]. Fortin et al. [7] used intervals of 20o and 50m for the frequency tables for turn angle and step length respectively which equated to 18 bins (360o / 20o) for turn angle and 60 bins (maximum length of 3000m / 50m). In theory, the smaller the intervals of the bins, the closer the empirical distribution will match a more continuous real distribution. Other SSF studies have followed, using evenly spaced intervals for turn

angle and step length [10], [13]. However, the intervals do not need to be even, and in portions of the distribution where the density is constant, a wider interval may be beneficial compared to when density changes rapidly and narrower bins may be favorable [12]; although, this adaptive method has yet to be employed.

Other studies have generated the available steps using a random distribution [14], a quantile distribution [15], and a parametric distribution [8]. When a random distribution has been used, turn angle was randomly drawn from between -180o to +180o, and step length was randomly drawn from any value between 0 and the maximum step length. The same method has been used for quantile distributions, but the values were capped at a specific percentage of the distribution. Bjørneraas et al. [15] used the 99% quantile for step lengths, although this method was not used to select turn angle as the distribution is in degrees and so cannot be capped in the same manner.

A recent study by Forester et al. [8] examined the use of SSFs using three methods of generating available samples (in terms of step length); random (termed uniform), empirical and parametric and compared the results on simulated data and elk data in Yellowstone National Park. They found that empirical and parametric distributions performed better than the uniform distribution which was the method furthest from resembling a continuous real distribution. The level of bias in the estimates differed between empirical and parametric distributions and was dependent on the size of the coefficient. With empirical distributions currently more widely used in SSF studies, we will choose to focus on sampling from this distribution as uncertainty still exists between using even and uneven intervals.

Number of steps

Once the available steps have been generated, the next user decision is the number of available steps used in the comparison. In their seminal paper, Fortin et al. [7] used 200 available steps, mainly because their research question was directed at rare habitat selection, although they noted that future SSF studies would not need such a high number. 200 remains the largest number of steps used in SSF studies so far, with the next highest value only 25 steps [10], while the minimum is two [15]. Subsequent studies have not specified the reasoning behind their selection of fewer steps beyond citing the statement of Fortin et al. [7] that they would not need 200. RSFs have been used more extensively than SSFs, and while some studies have suggested using 10,000 available locations (in total) in the comparison [16, 17], using an ad hoc approach still appears to be the norm [18]. Subsequently, it is unknown whether a high number of available locations/steps over-samples the environmental choices available, or whether a low number of available locations/steps under-samples.

Modelling Approach

The majority of SSF studies have used conditional logistic regression as the statistical method, but there has been variation with respect to how the analysis is conducted. Conditional logistic regression in which regression parameters are fit for all of the individuals together has been the most widely used approach. An alternative to this has recently been employed whereby model parameters are fit for each individual separately and then averaged to attain aggregate information, and has been termed individual modelling [14], [19], [20]. Individual modelling can potentially capture more individual movement traits while still being applicable to the aggregation of individuals and is beginning to be used more regularly.

Individual Variation

The effect of individual variation on results has only recently begun to be incorporated in SSF studies, but in the wider field of telemetry studies, this is slightly more developed. Lindberg and Walker [21] suggest that more than 20 animals are needed to make reliable statistical inferences about simple population comparisons, and at least 75 animals for realistically complex studies. However, due to the high costs of GPS units, sample sizes are often far from these numbers, with several of the SSF studies containing less than 20 individuals [7], [19], [22]. In studies where sample sizes are lower, individual traits may have more influence on the results, although in a recent study using seven Eleonora's falcons, Gschweng et al. [23] found that removing one from the study did not result in a significant change in habitat use. Few studies have taken into account the idiosyncratic differences among individual animals in SSF models and this is an area that needs further study.

Uncertainty Analysis

Uncertainty analysis allows researchers to assess the range of outputs associated with the model responses as a result of variations in parameter values used in the model input [4]. This uncertainty analysis will provide SSF practitioners with insight into how to choose appropriate model parameters as well as information on the level of uncertainty associated with the results. Here, we analyze the effect of 1) generating the available steps based on a) turn angle and b) step length, 2) the number of available steps used, 3) the modelling approach used, and 4) the number of individuals used.

Methods

Data Collection

Oilbird (Steatornis caripensis) data was obtained from Holland et al. [24] via Movebank [25]. GPS with remote UHF readout was used to collect locations of four individuals with ten minute intervals, resulting in approximately 800 fixes for use, with the number of observations ranging from 133 to 264. Brown hyena (Hyaena brunnea) data was collected by Maude [26] in the Makgadikgadi Pans region of northern Botswana between June 2004 and December 2007 with the support of the Makgadikgadi Brown Hyena Project. GPS locations were recorded for each hyena at 1 hour intervals. Ten hyenas were used in this analysis across a time period of a month, which resulted in 4000 fixes for use, ranging from 278-432. The environmental variables used in both of these analyses are briefly described in table 1.

Uncertainty Analysis: Generating Available Step

Steps were generated from drawing from four distributions for both turn angle and step length; empirical distribution (even bins), empirical distribution (uneven bins), random distribution, and quantile distribution (99%) (n.b. quantile distribution was not used for turn angle). Empirical distributions of even bins formed intervals of 20o for turn angles, and 187.3m and 362m for step length for oilbirds and hyenas respectively. The values of step lengths equates to 2% of the maximum step length. Empirical distributions of uneven bins formed intervals of 5.56% quantiles for turn angle and 2% quantiles for step lengths. These values ensured that the same number of bins would be used for even and uneven empirical distributions.

Uncertainty Analysis: Number of Available Steps

The number of available steps were chosen to fall between the minimum (2) and the maximum value used (200) in previous SSF studies: 2, 10, 20, 50, 100, and 200.

Table 1. Information about the environmental variables used in the regression models. The IGBP classification scheme was used for the MODIS land cover product [20], which delineates into 16 classes.

Animal	Environmental Variable	Method	Hypothesis	Source
Oilbird	Land Cover – Evergreen Broadleaf Forest	Categorical Variable - Value of land cover at the end of the step	Oilbirds eat the fruit of tropical laurels. The majority of these trees are evergreen broadleaf species. The oilbirds are more likely to move through land covers which contain food sources.	[27]
Oilbird	Cropland – Percentage of cropland	Continuous Variable - Value of cropland at the end of the step	Oilbirds also eat the fruit of oil palms, a commercial crop in South America. The higher the percentage of cropland, the higher chance of oil palms in the area.	[28]
Oilbird	Distance to Roads	Continuous Variable - If step crosses line – distance = 0. Else, distance equals the average distance from start, mid and end of the step	Birds have been found to frequent edge habitats, with roads one of the most common features splitting habitats.	[29]
Hyena	Land Cover – Savanna	Categorical Variable - Value of land cover at the end of the step	Hyenas may visit savanna habitats more often due to a higher potential of prey species.	[27]
Hyena	Land Cover – Open Shrublands	Categorical Variable - Value of land cover at the end of the step	Hyenas may visit shrubland habitats more often due to a higher potential of prey species.	[27]
Hyena	Land Cover – Fragmentation of habitat	Continuous Variable - Number of surrounding habitats of the cover at the end of the step	The more surrounding habitats results in the potential of more prey species. More habitats may mean more species	[27]
Hyena	Distance to Roads	Continuous Variable - If step crosses line – distance = 0. Else, distance equals the average distance from start, mid and end of the step	Hyenas could use roads to travel along or obtain road kill. Alternatively they may avoid them due to the dangers.	[29]

Uncertainty Analysis: Modelling Approach

Conditional logistic regression (equation 1.0) was done in R using the survival package [30]. This was either conducted with all individuals in one table (herein referred to as aggregate modelling), or for each individual separately with the coefficients averaged to generate an aggregate level model (herein referred to as individual modelling).

Uncertainty Analysis: Individual Variation

Conditional logistic regression was fit for all the individuals, while systematically dropping one individual from the analysis (n.b. this analysis was undertaken using available steps generated from the empirical distribution of all the individuals). All combinations of turn angle, step length, number of steps and modelling approach was also conducted for each model without certain individuals.

Data Analysis

In total, 720 oilbird and 1584 hyena regression models were fit. The coefficient values were compared for each variable with Wilcoxon matched pairs signed rank test using a Bonferroni corrected α of 0.05 according to the number of comparisons made. This test converts scores to ranks and compares them across the two conditions. The effect size of the test was calculated by dividing the z value by the square root on N and using the Cohen [31] criteria of 0.1 = small effect, 0.3 = medium effect and 0.5 = large effect.

Results

For both hyenas (table 2) and oilbirds (table 3), some combination of coefficients differed for each variable (n.b. tables only contain comparisons where one significant difference exists). The distribution used to generate the turn angle differed between random and empirical (even or uneven), although this effect was very small. Step length distribution caused a variety of significant differences in the coefficients, with Fig. 3 showing the range of coefficient values for savanna selection for hyenas and evergreen broadleaf forest for oilbirds. The number of available samples used in the model resulted in some small significant differences, although these were not consistent between species (tables 2 and 3). Individual modelling resulted in inflated coefficient values when compared to the aggregate model (Fig. 2), with only two hyenas and one oilbird appearing to cause the increased values (Fig. 4). Finally, systematically removing individuals caused significant differences, with the largest differences appearing to match the individuals causing the inflated coefficients in Fig. 4. Removing individuals from the analysis does result in the coefficient of most importance changing in the final regression model in some instances (Fig. 5).
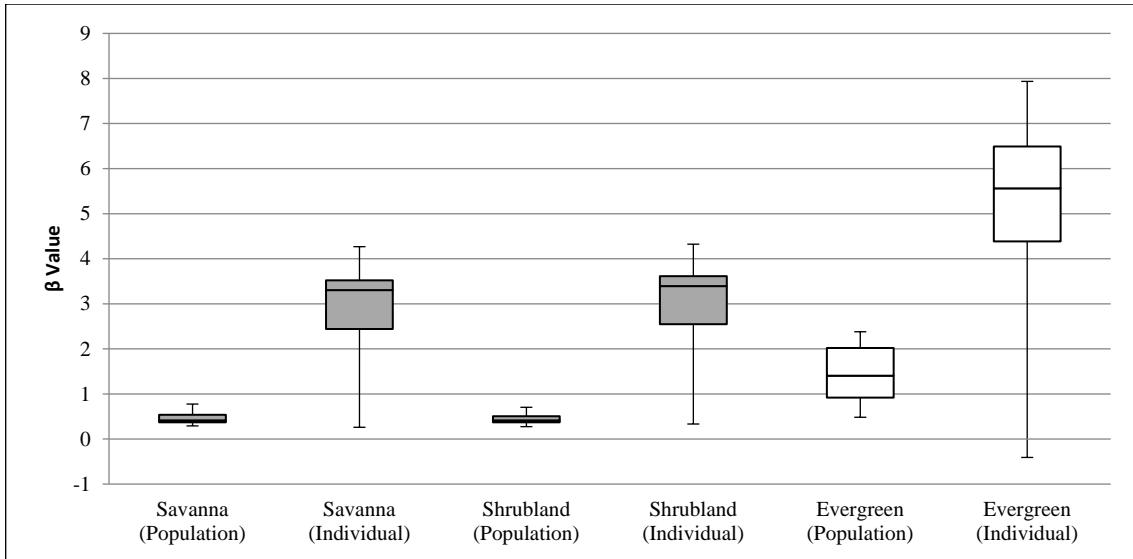
Fig. 2. β values for savanna and shrublands for individual and population modelling of hyena SSF (grey) and evergreen broadleaf forest for oilbirds (white)
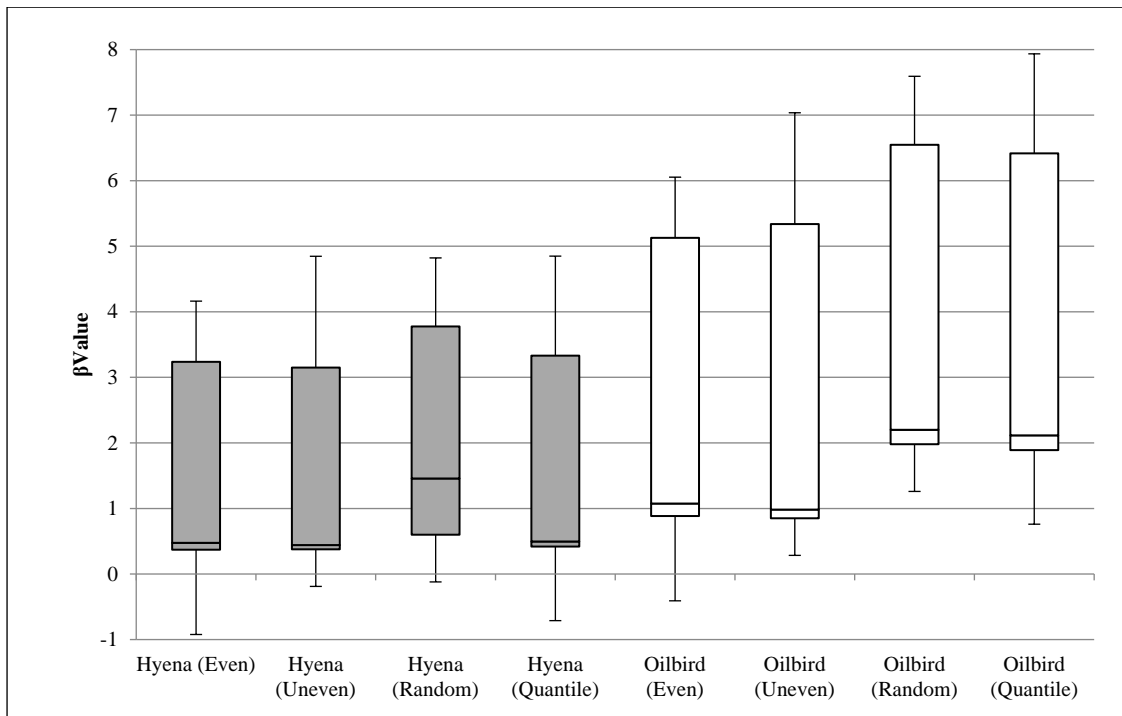

Fig. 3. β values for savanna habitat for hyenas (grey) and evergreen broadleaf forests for oilbirds (white) based on the four methods of generating step lengths of oilbird SSF

Table 2. Wilcoxon Matched Pairs Signed Rank for hyenas, value refers to second group, *medium effect, **large effect.

Variable		β Roads	β Savanna	β Shrubland	β Edge
Method	Aggregate - Individ	Lower	Higher**	Higher**	Lower**
Turn Angle	Even - Random	n.s.	Lower	Lower	Higher
	Even - Uneven	n.s.	n.s.	n.s.	Lower
	Uneven - Random	n.s.	Lower	Lower *	Lower *
Step Length	Even - Uneven	n.s.	n.s.	Lower	Higher
	Even - Random	Lower*	Higher**	Higher**	Higher**
	Even - Quantile	n.s.	Higher*	Higher	n.s.
	Uneven - Random	Lower*	Higher**	Higher*	Higher*
	Uneven - Quantile	n.s.	Higher*	Higher	Lower
	Quantile - Random	Lower*	Higher**	Higher**	Higher**
Number of Steps	2 - 50	Lower	n.s.	n.s.	n.s.
	10 - 20	n.s.	Lower*	Lower	n.s.
	10 - 50	n.s.	Lower*	Lower	n.s.
	10 - 100	n.s.	Lower	Lower	n.s.
	10 - 200	n.s.	Lower*	Lower	n.s.
	20 - 50	n.s.	n.s.	Higher	Lower
	20 - 200	n.s.	Lower	n.s.	n.s.
	50 - 200	n.s.	Lower	Lower	n.s.
	100 - 200	n.s.	Lower*	Lower	n.s.
Individuals	All - H1	Higher	Higher**	Higher**	Lower**
	All - H2	n.s.	Lower	Lower**	Lower**
	All - H3	Higher	Lower*	Higher**	Higher**
	All - H4	n.s.	Higher**	Higher**	Higher**
	All - H5	n.s.	Lower	Lower**	Lower**
	All - H6	n.s.	Lower	Lower**	Lower**
	All - H7	Higher	Lower	Lower*	Higher*
	All - H8	n.s.	Higher*	Higher	n.s.
	All - H9	n.s.	Higher**	Higher**	Lower**
	All - H10	Higher	Lower	Higher	n.s.

Table 3. Wilcoxon Matched Pairs Signed Rank for oilbirds, value refers to second group, *medium effect, **large effect.

Variable		β Roads	β Evergreen	β Croplands
Method	Aggregate - Individ	Lower	Higher*	Lower**
Turn Angle	Even - Random	n.s.	n.s.	Lower
Step Length	Even - Uneven	n.s.	n.s.	Higher
	Even - Random	n.s.	Higher**	n.s.
	Even - Quantile	n.s.	Higher**	n.s.
	Uneven - Random	n.s.	Higher**	n.s.
	Uneven - Quantile	n.s.	Higher*	Lower
	Quantile - Random	n.s.	Higher*	Higher
Number of Steps	2 - 10	Higher	n.s.	Higher
	2 - 20	Higher	n.s.	Higher
	2 - 50	Higher	n.s.	Higher
	2 - 100	Higher	n.s.	Higher
	2 - 200	Higher	n.s.	Higher
	10 - 20	n.s.	Lower	n.s.
	10 - 200	n.s.	Lower*	n.s.
	20 - 200	n.s.	Lower	n.s.
	50 - 100	n.s.	Lower	n.s.
	50 - 200	n.s.	Lower*	n.s.
	100 - 200	n.s.	Lower*	n.s.
Individuals	All - B1	n.s.	Higher**	Lower**
	All - B2	Lower	Higher**	Higher**
	All - B3	n.s.	n.s.	Lower
	All - B4	Lower	Higher*	Lower**

Discussion

The objective of this study was to investigate the uncertainty associated with user decisions and how these decisions influence the results obtained from step-selection functions for two different types of animals (oilbirds and brown hyenas). Oilbirds congregate in caves at night, and forage for food by day, with most of their food source coming from evergreen laurels. Brown hyenas forage solitarily, across large home ranges encompassing a variety of land covers, returning to a clan den at night. Both datasets had a considerable number of observations, as well as differing temporal resolutions (10 minutes for oilbirds, 1 hour for hyenas). Subsequently, any similarities or differences in the results could be attributed to either a user-decision or a species specific trait and should be used by other researchers to help inform their implementation of SSF.

Our results indicate that there is considerable uncertainty associated with the decisions of selecting model parameters and their effects on coefficient values. SSFs associate parameters of movement rules with landscape features, as well as modeling the choices actually presented to the animal as it traverses through the landscape [32]. However, if the user decisions do alter the results observed, then it is difficult to disentangle actual step/habitat preferences from model parameterization decisions. The number of significant differences found in tables 2 and 3, and the size of these differences suggests that researchers are currently unable to distinguish between whether results

are representative of step/habitat preferences or whether they are a function of researchers selecting certain methods to generate their SSF model.
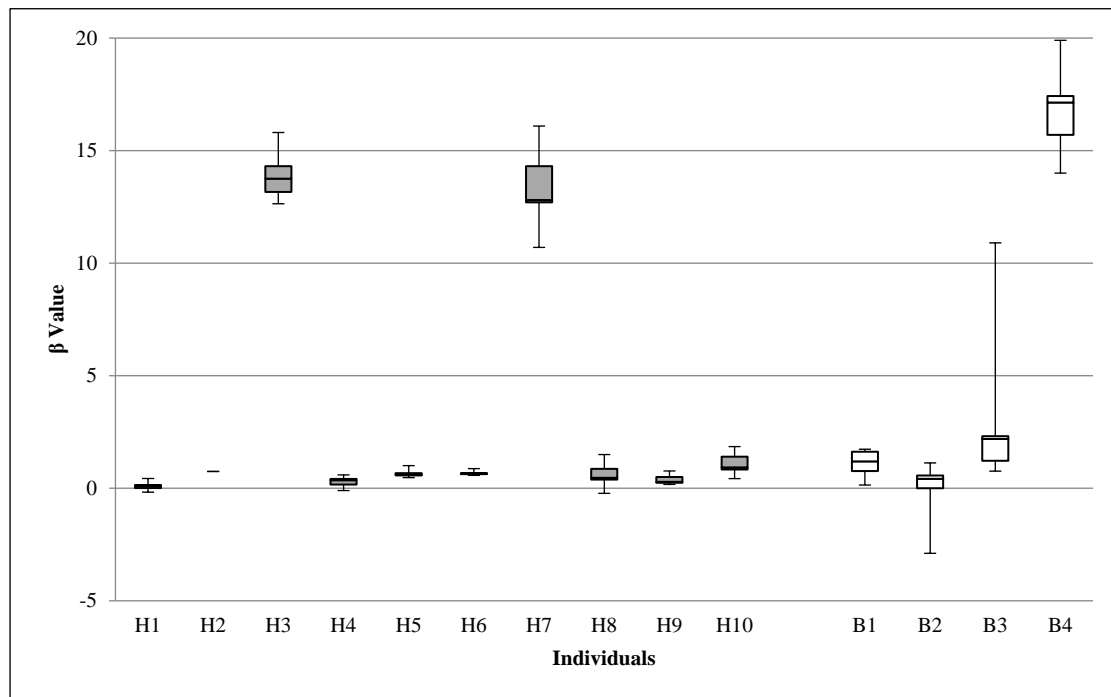

 Fig. 4. β values of shrublands for individual hyenas (grey) and the β values of evergreen broadleaf forest for individual oilbirds (white)

The largest differences appeared to be between the modelling approach used and systematically removing one individual from the model. The idea behind an individual modelling approach is that the final model will contain results that highlight environmental interactions that better represent individual to population preferences for habitat use, something that is not possible with aggregate modelling (see Figs. 2 and 4). Fig. 2 shows the range of coefficient values for savanna and shrublands (for hyenas) and evergreen forest (for oilbirds), when calculated using the aggregated and individual modelling approaches. Fig. 4 shows the range of coefficients for each individual hyenas and oilbirds that were averaged in the individual modelling approach. It can be observed that hyena 3 and 7, and oilbird 4 have much higher coefficients than the others (indicating their increased preference for moving into that specific habitat). The aggregate method appears to suppress the idiosyncratic preferences of these individuals, while the individual method incorporates it with a higher average value. Individual modelling therefore incorporates individual information into the coefficient and researchers need to be aware of these differences, as the method used in analysis could over- or under-estimate the importance of an environmental variable based on the idiosyncratic preferences of just one or two individuals (as shown in our research).

The influence of individuals is furthered highlighted by the differences in results when one was dropped from the regression (tables 2 and 3). Removing either hyena 1, 4, 8 and 9 from analysis results in a higher preference for savanna habitats and removing either hyena 1, 3, 4, 8, 9 and 10 from analysis result in a higher preference for shrublands. Similar results are observed for oilbirds (table 3). With only four oilbirds, these differences in results are less surprising, but ten hyenas are not a small sample size, and is larger than the sample size used by Gschweng et al. [23] when they concluded that removing one individual did not result in statistically different results (albeit the results were not based on conditional logistic regression). Fig. 5 identifies the number of times

removing an individual results in the coefficient of most importance shifting from the value in the complete model. Overall, these values are relatively small, but it shows that in some instances coefficient importance can change. Therefore, datasets of similar sizes as those used in this research are relatively sensitive to individual preferences. Interestingly, the hyena which had the most impact on a change in coefficient importance is not one of the hyenas already identified as having a substantially higher coefficient value when conditional logistic regression was run for each individual (Figs. 4 and 5). Individuals can therefore influence the value of the coefficients obtained as well as their subsequent importance compared to other environmental variables, although the individual that does so is not necessarily the same for both differences.

A recent review outlining the need for an uncertainty analysis of SSF suggested that the method of generating the available steps could be the most important decision when developing this model [11]. While the method of generating the step length caused significant changes, this was not the case for generating the turn angle (tables 2 and 3). While significant differences did exist between turn angle distributions, they were either of a small or medium effect for hyenas, and only one difference between even and random distributions existed for the coefficient describing the effect of croplands on oilbirds. Step length was a more important variable than turn angle, although it appeared more important for variables where the value was measured at the end of the step (savanna, shrubland, edge habitat for hyenas and broadleaf forest for oilbirds). Distributions which selected randomly up to the maximum (or 99% quantile) step length resulted in higher coefficients than even or uneven distributions (Fig. 3). These results could be compounded by infrequent movement steps that occur over great dista-
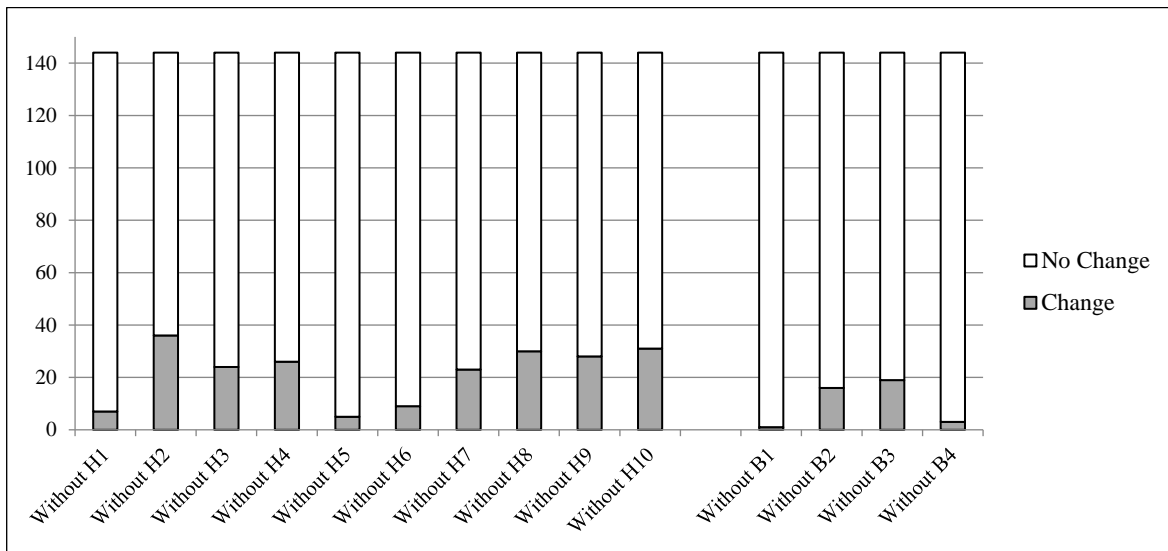

Fig. 5. The number of times removing an individual from the regression (note 144 combinations of variables) results in the coefficient of most importance shifting

nces, but are not controlled for through their low probability of occurrence as they would be with an empirical distribution. The higher coefficients obtained using a random distribution are a result of longer steps taking the available steps beyond the evergreen forest for oilbirds and the savanna and shrubland for hyenas. This increases the number of alternative habitats the animal could select, and thus increases the coefficient value of their preferred habitat.

Generating steps using a quantile distribution (99%) resulted in higher coefficients for oilbirds, but not hyenas (Fig. 3). The distribution of step lengths for hyenas is relatively normal (note a slight bimodal distribution), while for oilbirds it is L shaped (data not shown). Therefore, using a 99%

quantile distribution removes the most extreme values for the hyenas, but fails to do so with the oilbirds, meaning that the quantile distribution continues to produce results similar to a random distribution. We suggest that researcher investigate the type of distribution associated with step lengths first, before making a decision on the type of distribution they use to generate available steps.

The number of available steps was incorporated in the uncertainty analysis as it is unknown whether a high or low number of available steps over- or under-samples the environmental choices available. While results were not consistent for oilbirds and hyenas, differences in the number of available steps used were found for both hyenas and oilbirds (table 2 and 3). Coefficients describing the effects of roads and croplands on oilbird movement differed between all combinations with either 2 steps or 200 steps. This suggests that while values between 10 and 100 do not statistically differ, the smallest and largest values do. The opposite was observed for hyenas, with many of the medium to higher values of steps being significantly different and while 2 steps was not significantly different to a higher number. However, the majority of these differences are small, and it is the category which has the most not significant differences. While it may not be the variable that has the strongest effect on the results, it does have a slight impact (possibly due to rare habitats or homogenous environments) and subsequently the researcher needs to select a number of steps accordingly.

This study used a variety of environmental variables, including discrete, continuous and distance to, with values measured at various points along the step (see table 1). Significant differences in the effect of step length could be attributed to the possibility that this value was measured at the end of the step, while distance to roads (which was often less significant) was averaged across the whole length. More research into how the environmental variables are measured has been suggested by Thurfjell et al. [11] as an area of research that needs further exploration, and certain patterns in the results of this study certainly indicate that this warrants further research.

Conclusion

User decisions strongly influence the results of step-selection functions and any subsequent inferences about animal movement and environmental interactions. By assuming that results would be consistent between methods, any conservation management strategies based upon SSF research could be a function of specifying unrealistic movement options. This study found that differences in individual behaviors have the strongest influence on the results observed. Averaging coefficients across individuals results in higher values when studied at an aggregate level, indicating that individual preferences are lost when studied solely at an aggregate level. The influence of removing one individual from the study was surprisingly significant and contradicts recent research [23]. Researchers conducting analysis on such medium sized datasets need to be aware that idiosyncratic preferences could potentially influence the results in terms of coefficient values and importance of variables used in the model, and should check for such occurrences. The method of generating available steps was important, but in this study not as important as some have suggested [11], although variations within the distributions used in this study (i.e. smaller intervals for empirical distributions) could be investigated further. Finally, while the number of steps used in comparison was the least important variable in determining coefficient values, it was still significantly different for a number of combinations. Differences between the extremes (2 and 200) suggest that a medium value may be preferable, but that only 2 steps could potentially mask actual movement preferences. User decisions of SSF practitioners should subsequently be justified based on research objectives where possible and further research into other user decisions of SSF should continue.

References

1. Miller, H.J.: A measurement theory for time geography. Geogr. Anal. 37, 17-45 (2005)
2. Cagnacci, F., Boitani, L., Powell, R.A., Boyce, M.S.: Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. Philos. T. Roy. Soc. B. 365, 2157-2162 (2010)
3. Gudmundsson, J., Laube, P. Wolle, T.: Computational movement analysis. In Springer handbook of geographic information. 423-438. Springer Berlin Heidelberg (2012)
4. Crosetto, M., Tarantola, S.: Uncertainty and sensitivity analysis: tools for GIS-based model implementation. Int. J Geogr. Inf. Sci. 15, 415-437 (2001)
5. Laube, P., Purves, R.S.: How fast is a cow? Cross-Scale Analysis of Movement Data. Trans. GIS. 15, 401–418 (2011)
6. Manly, B.F., McDonald, L.L., Thomas, D.L., McDonald, T.L., Erickson, W.P.: Resource selection by animals: statistical design and analysis for field studies. Springer (2002)
7. Fortin, D., Beyer, H.L., Boyce, M.S., Smith, D.W., Duchesne, T., Mao, J.S.: Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park. Ecology. 86, 1320–1330 (2005)
8. Forester, J.D., Im, H.K., Rathouz, P.J.: Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. Ecology. 90, 3554–3565 (2009)
9. Gillies, C.S., Beyer, H.L., St. Clair, C.C.: Fine-scale movement decisions of tropical forest birds in a fragmented landscape. Ecol. Appl. 21, 944–954 (2011)
10. Latham, A.D.M., Latham, M.C., Boyce, M.S., Boutin, S.: Movement responses by wolves to industrial linear features and their effect on woodland caribou in northeastern Alberta. Ecol. Appl. 21, 2854–2865 (2011)
11. Thurfjell, H., Ciuti, S., Boyce, M.S.: Applications of step-selection functions in ecology and conservation. Mov. Ecol. 2, 4 (2014)
12. Beyer, H.L.: Geospatial Modelling Enviornment. Available at: http://www.spatialecology.com/gme/gmehelp.htm, [Accessed 12/12/13]
13. Coulon, A., Morellet, N., Goulard, M., Cargnelutti, B., Angibault, J.-M., Hewison, A.M.: Inferring the effects of landscape structure on roe deer (Capreolus capreolus) movements using a step selection function. Landsc. Ecol. 23, 603–614 (2008)
14. Squires, J.R., DeCesare, N.J., Olson, L.E., Kolbe, J.A., Hebblewhite, M., Parks, S.A.: Combining resource selection and movement behavior to predict corridors for Canada lynx at their southern range periphery. Biol. Conserv. 157, 187–195 (2013)
15. Bjørneraas, K., Solberg, E.J., Herfindal, I., Moorter, B.V., Rolandsen, C.M., Tremblay, J.-P., Skarpe, C., Sæther, B.-E., Eriksen, R., Astrup, R.: Moose Alces alces habitat use at multiple temporal scales in a human-altered landscape. Wildl. Biol. 17, 44–54 (2011)
16. Lele, S.R., Keim, J.L.: Weighted distributions and estimation of resource selection probability functions. Ecology 87, 3021-3028 (2006)
17. Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W.: Selecting pseudo-absences for species distribution models: how, where and how many? Methods in Ecology and Evolution 3, 327-338 (2012)
18. Northrup, J.M., Hooten, M.B., Anderson Jr, C.R., Wittemyer, G.: Practical guidance on characterizing availability in resource selection functions under a use-availabiltiy design. Ecology 94, 1456-1463 (2013)

19. Northrup, J.M., Pitt, J., Muhly, T.B., Stenhouse, G.B., Musiani, M., Boyce, M.S.: Vehicle traffic shapes grizzly bear behaviour on a multiple-use landscape. J. Appl. Ecol. 49, 1159–1167 (2012)

20. Sawyer, H., Nielson, R.M., Lindzey, F., McDonald, L.L.: Winter habitat selection of mule deer before and during development of a natural gas field. J Wildlife Manage. 70, 396-403 (2006)

21. Lindberg, M.S., Walker, J.: Satellite telemetry in avian research and management: sample size considerations. J. Wildl. Manag. 71, 1002–1009 (2007)

22. Pereboom, V., Mergey, M., Villerette, N., Helder, R., Gerard, J.-F., Lode, T.: Movement patterns, habitat selection, and corridor use of a typical woodland-dweller species, the European pine marten (Martes martes), in fragmented landscape. Can. J. Zool. 86, 983–991 (2008)

23. Gschweng, M., Kalko, E.K., Berthold, P., Fiedler, W., Fahr, J.: Multi-temporal distribution modelling with satellite tracking data: predicting responses of a long-distance migrant to changing environmental conditions. J. Appl. Ecol. 49, 803–813 (2012)

24. Holland, R.A., Wikelski, M., Kümmeth, F., Bosque, C.: The secret life of oilbirds: new insights into the movement ecology of a unique avian frugivore. PLoS One. 4, e8264 (2009)

25. Data from: The secret life of oilbirds: new insights into the movement ecology of a unique avian frugivore, http://hdl.handle.net/10255/move.269

26. Maude, G.: The spatial ecology and foraging behaviour of the brown hyaena (Hyaena brunnea) (Doctoral dissertation, Universtiy of Bristol) (2010)

27. Land Processes Distributed Active Archive Center (LP DAAC) at the USGS/Earth Resources Observation and Science (EROS) Center, Siox Falls, South Dakota. (2013)

28. Ramankutty, N., Evan, A.T., Mondreda, C., Foley, J.A.: Global Agricultural Lands: Croplands, 2000. Data distributed by the Socioeconomic Data and Applications Center (SEDAC): http://seda.ciesin.columbia.edu/es/algands.html. [Accessed 03/01/14]. (2010)

29. Center for International Earth Science Information Network - CIESIN - Columbia University and Information Technology Outreach Services - ITOS - University of Georgia. Global Roads Open Access Data Set, Version 1 (gROADSv1). Palisades, NY: NASA Socioecomonic Data and Applications Ceneter (SEDAC). http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1. [Accessed on 03/01/14]. (2013)

30. Therneau, T.: A package for survival analysis in S. R package version 2.37.4. (2013)

31. Cohen, J.: Statistical power analysis for the behavioral sciences. Psychology Press (1988)

32. Chekiewicz, C.L.B., St Clair, C.C., Boyce, M.S.: Corridors for conservation: integrating pattern and process. Annu. Rev. Ecol. Evo. Syst. 37, 317-342.