

## Integrated step selection analysis: bridging the gap between resource selection and animal movement

Tal Avgar<sup>1\*</sup>, Jonathan R. Potts<sup>2</sup>, Mark A. Lewis<sup>1,3</sup> and Mark S. Boyce<sup>1</sup>

<sup>1</sup>*Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada;* <sup>2</sup>*School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, UK;* and <sup>3</sup>*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada*

### Summary

1. A resource selection function is a model of the likelihood that an available spatial unit will be used by an animal, given its resource value. But how do we appropriately define availability? Step selection analysis deals with this problem at the scale of the observed positional data, by matching each ‘used step’ (connecting two consecutive observed positions of the animal) with a set of ‘available steps’ randomly sampled from a distribution of observed steps or their characteristics.

2. Here we present a simple extension to this approach, termed integrated step selection analysis (iSSA), which relaxes the implicit assumption that observed movement attributes (i.e. velocities and their temporal autocorrelations) are independent of resource selection. Instead, iSSA relies on simultaneously estimating movement and resource selection parameters, thus allowing simple likelihood-based inference of resource selection within a mechanistic movement model.

3. We provide theoretical underpinning of iSSA, as well as practical guidelines to its implementation. Using computer simulations, we evaluate the inferential and predictive capacity of iSSA compared to currently used methods.

4. Our work demonstrates the utility of iSSA as a general, flexible and user-friendly approach for both evaluating a variety of ecological hypotheses, and predicting future ecological patterns.

**Key-words:** conditional logistic regression, dispersal, habitat selection, movement ecology, random walk, redistribution kernel, resource selection, step selection, telemetry, utilisation distribution

### Introduction

Ecology is the scientific study of processes that determine the distribution and abundance of organisms in space and time (Elton 1927). Hence, asking how and why living beings change their spatial position through time is fundamental to ecological research (Nathan *et al.* 2008). Animal movement links the behavioural ecology of individuals with population and community level processes (Lima & Zollner 1996). Its study is consequently vital for understanding basic ecological processes, as well as for applications in wildlife management and conservation.

Whether basic or applied, common to many empirical studies of animal movement is the aspiration to reliably predict population density through space and time by modelling the spatiotemporal probability of animal occurrence, also known as the utilisation distribution (Keating & Cherry 2009). Despite much progress in statistical characterisation of animal movement and habitat associations, our ability to predict utilisation distributions is limited by our understanding of the underlying behavioural processes. Indeed, includ-

ing explicit movement behaviours into spatial models of animal density has led to improved predictive performance (Moorcroft, Lewis & Crabtree 2006; Fordham *et al.* 2014). Deriving predictive space-use models based on the behavioural process underlying animal movement patterns is of particular importance when dealing with altered or novel landscapes that might differ substantially from the landscape used to inform the models.

Over the past three decades, a great deal of research has been dedicated to explaining and predicting spatial population distribution patterns based on underlying habitat attributes (often termed resources). In that regard, much focus has been given to estimating resource selection functions (Manly *et al.* 2002)–phenomenological models of the relative probability that an available discrete spatial unit (e.g. an encountered patch or landscape pixel) will be used given its resource type/value (Lele *et al.* 2013). Indeed, its intuitive nature and ease of application has made resource selection analysis (RSA) the tool of choice for many wildlife scientists and managers seeking to use environmental information in conjunction with animal positional data (Boyce & McDonald 1999; McDonald *et al.* 2013; Boyce *et al.* 2015).

Whereas much progress has been gained in the application of RSAs to animal positional data, the problem of defining

\*Correspondence author. E-mail: avgar@ualberta.ca

the appropriate spatial domain available to the animal remains as a major concern (Matthiopoulos 2003; Lele *et al.* 2013; McDonald *et al.* 2013; Northrup *et al.* 2013). Weighted distribution approaches deal with this problem by modelling space-use as a function of a movement model and a selection function, but most weighted distribution models are challenging to implement (but see Johnson, Hooten & Kuhn 2013). Matched case-control logistic regressions (CLRs; also known as discrete-choice models) may be considered a simplified alternative to the weighted distribution approach where each observed location is matched with a conditional availability set, limited to some predefined spatial and/or temporal range (Arthur *et al.* 1996; McCracken, Manly & Heyden 1998; Compton, Rhymer & McCollough 2002; Boyce *et al.* 2003; Baasch *et al.* 2010). A major strength of this approach is that maximum-likelihood estimates (MLEs) of the parameters can be efficiently obtained through commonly used statistical software (often relying on a Cox Proportional Hazard routine; e.g. function *clogit* in R). One particular type of such conditional RSA is step selection analysis (SSA), where each ‘used step’ (connecting two consecutive observed positions of the animal) is coupled with a set of ‘available steps’ randomly sampled from the empirical distribution of observed steps or their characteristics (e.g. their length and direction; Fortin *et al.* 2005; Duchesne, Fortin & Courbin 2010; Thurfjell, Ciuti & Boyce 2014).

The definition of availability is challenging, however, even when using the SSA approach. The problem arises due to the sequential, rather than simultaneous, estimation of the movement and habitat-selection components of the process. Owing to this stepwise procedure, the resulting habitat-selection inference is conditional (on movement), whereas movement is assumed independent of habitat selection. In reality, the two are tightly linked, with habitat selection and availability affecting the animal’s movement patterns (Avgar *et al.* 2013b), and the animal’s movement capacity affecting its habitat-use patterns (Rhodes *et al.* 2005; Avgar *et al.* 2015). Failure to adequately account for the movement process may consequently lead to biased habitat-selection estimates (Forester, Im & Rathouz 2009).

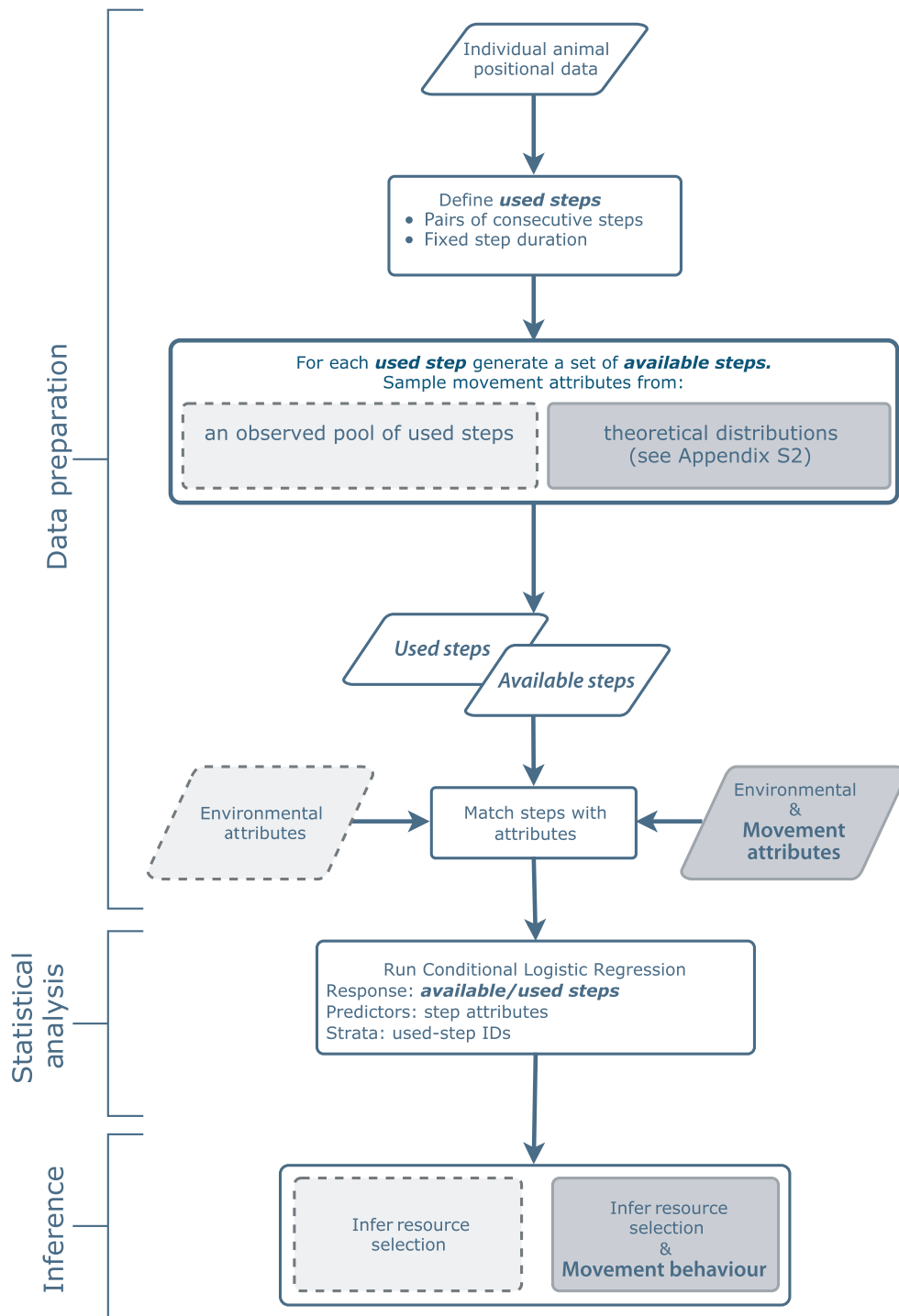
As we will show here, the benefits of adequately accounting for the movement process may extend beyond obtaining unbiased habitat-selection estimates. SSAs rely on a simple depiction of animal movement as a series of stochastic discrete steps, characterised by specific velocity and autocorrelation distributions. This same depiction underlies the mathematical modelling of animal movement as a discrete-time random walk (RW), including correlated and/or biased RW (Kareiva & Shigesada 1983; Turchin 1998; Codling, Plank & Benhamou 2008). Indeed, many SSA formulations correspond to a correlated RW process with local bias produced by resource selection (BCRW; Duchesne, Fortin & Rivest 2015). Apart from their compatibility with the way we often observe animal movement (i.e. in continuous space and at discrete times), many RW can be well approximated by diffusion equations, allowing

a much sought shift from an individual-based Lagrangian perspective to population-level Eulerian models (Turchin 1991, 1998). SSAs are thus at an interface between statistical (phenomenological) RSAs and mathematical (mechanistic) RW models (Potts, Mokross & Lewis 2014b; Potts *et al.* 2014a), models that form the backbone of much of the existing body of theory in the field of animal movement (Codling, Plank & Benhamou 2008; Benhamou 2014; Fagan & Calabrese 2014).

In this paper, we outline a CLR-based approach for simultaneous estimation of the movement and habitat-selection components, an approach we name *integrated step selection analysis* (iSSA; Fig. 1). The iSSA allows the effects of environmental variables on the movement and selection processes to be distinguished, thus providing a valuable tool for testing hypotheses (e.g. to test whether animals travel faster in certain times or through certain habitats), while resulting in an empirically parameterised mechanistic movement model (i.e. a mechanistic step selection model; Potts *et al.* 2014a), that can be used to translate individual-level observations to population-level utilisation distributions across space and time (Potts *et al.* 2014a; Potts, Mokross & Lewis 2014b; Appendix S1).

The iSSA is related to several recently published works integrating animal movement and resource selection. Both Forester, Im & Rathouz (2009) and Warton & Aarts (2013) demonstrated the inclusion of movement variables in an RSA and its marked effect on the resulting inference. Johnson, Hooten & Kuhn (2013) have shown that animal telemetry data can be viewed as a realisation of a non-homogenous space-time point process, and MLEs of this process can be obtained using a generalised linear model. These contributions focused on gaining unbiased resource selection inference while treating the movement process as nuisance that must be ‘controlled for’. Here, we seek explicit inference of this process. State-space models of animal movement (reviewed by Jonsen, Myers & Flemming 2003; and Patterson *et al.* 2008) predict the future state (e.g. spatial position) of the animal given its current state (where an ‘observation model’ provides the probability of observing these states), environmental covariates, and an explicit ‘process model’. Once parametrised, the process model can be used to generate space-use prediction, but parametrisation is often technically demanding and computationally intensive (Patterson *et al.* 2008). More recently, Potts *et al.* (2014a) demonstrated the use of a ‘mechanistic step selection model’ to infer both the drivers and the steady-state distribution of animal space-use, but the model was framed around one specific functional form of the movement kernel, and parameter estimates were obtained using a custom-made likelihood maximisation procedure. Lastly, Duchesne, Fortin & Rivest (2015) demonstrated that an SSA can be used to obtain unbiased estimates of the directional persistence and bias of a BCRW, but did not address parametrisation of the step-length distribution.

The iSSA builds and expands on these contributions. We will demonstrate that, by statistically accounting for an explicit



**Fig. 1.** Step selection analysis workflow. Light grey shading indicates conventional SSA whereas dark grey shading indicates the integrated approach advocated here (iSSA). See Appendix S4 for detailed iSSA guidelines and tips.

movement process within an SSA, a complete habitat-dependent mechanistic movement model can be parametrised from telemetry data using a standard CLR routine. In the following, we provide a detailed description of the approach and evaluate its performance (compared with standard RSA and SSA) in correctly inferring the movement and habitat-selection processes underlying observed space-use patterns, and in predicting the resulting UD.

## Materials and methods

### INTEGRATED STEP SELECTION ANALYSIS

In their work on the subject of accounting for movement in resource-selection analysis, Forester, Im & Rathouz (2009) demonstrated that including a distance function in SSA substantially reduces the bias in habitat-selection estimates. Mathematically, their argument is based

on the habitat-independent movement kernel (the function governing movement in the absence of resource selection, or across a homogeneous landscape; Hjernmann 2000; Rhodes *et al.* 2005) belonging to the exponential family, so that it can be accounted for with the logistic formulation of the SSA likelihood function. Here we shall make the assumption that, in the absence of habitat selection, step lengths follow either an exponential, half-normal, gamma or log-normal distribution. Under this assumption, the statistical coefficients associated with step length, its square, its natural logarithm and/or the square of its natural logarithm (depending on the assumed distribution), when incorporated as covariates in a standard SSA, serve as statistical estimators of the parameters of the assumed step-length distribution (see Appendices S2 and S3 for details, and below for an example). Standard model selection (e.g. likelihood ratio or AIC) then can be used to select the best-fit theoretical distribution (out of the four listed above).

The iSSA approach moreover can be applied to infer directional persistence and external bias. Assuming that the angular deviations from preferred directions (either the previous heading, the target heading or both) are von Mises distributed (an analogue of the normal distribution on the circle), the cosine of these angular deviations can be included as covariates in an SSA to obtain MLEs of the corresponding von Mises concentration parameters (Duchesne, Fortin & Rivest 2015). Hence, MLEs of iSSA coefficients affiliated with directional deviations and step lengths are directly interpretable as the parameters of distributions governing the underlying BCRW.

We shall make the assumption here that animal space-use behaviour is adequately captured by a separable model, involving the product of two kernels, a movement kernel and a habitat-selection kernel. Formally, we define a discrete-time movement kernel,  $\Phi$ , which is proportional to the probability density of occurrence in any spatial position,  $x$ , at time  $t$ , in the absence of habitat selection. The determinants of  $\Phi$  are as follows: the Euclidian distances between  $x$  and the preceding position,  $x_{t-1}$  (the step length;  $l_t = \|x - x_{t-1}\|$ ), the distances between  $x_{t-1}$  and  $x_{t-2}$  (the previous step length;  $l_{t-1} = \|x_{t-1} - x_{t-2}\|$ ), the associated step headings,  $\alpha_t$  and  $\alpha_{t-1}$  (the directions of movement from  $x_{t-1}$  to  $x$  and from  $x_{t-2}$  to  $x_{t-1}$ , respectively), and a vector of spatial and/or temporal movement predictors at time  $t$  and/or at the vicinity of  $x$  and/or  $x_{t-1}$ ,  $Y(x, x_{t-1}, t)$  (e.g. terrain ruggedness, migratory phase, snow depth, etc.). The effects of these step attributes on  $\Phi$  are controlled by the associated coefficients vector,  $\theta$ . Note that the effects of spatial attributes here are assumed to operate through local biomechanical interactions between the animal and its immediate environment, interactions that determine the rate of displacement (i.e. kinesis), not where the animal 'wants' to be (i.e. taxis). Also note that the kernel  $\Phi$  can be non-Markovian and accommodate various types of velocity autocorrelations (lack of independence between directions and/or lengths of consecutive steps), including correlated and biased random walks (if directional biases are known *a priori*).

We further define the habitat-selection function,  $\Psi$ , which is proportional to the probability density of observing the animal in any spatial position,  $x$ , at time  $t$ , in the absence of movement constraints. The determinants of  $\Psi$  are the habitat attributes in  $x$  at  $t$ ,  $H(x, t)$ , and their corresponding selection coefficients,  $\omega$ . The normalised product of  $\Phi$  and  $\Psi$  yields the probability density of occurrence in  $x$  at  $t$ , which is:

$$f(x_t | x_{t-1}, x_{t-2}) = \frac{\int_{\Omega} \Phi[l_t, l_{t-1}, \alpha_t, \alpha_{t-1}, Y(x, x_{t-1}, t), \theta] \cdot \Psi[H(x, t), \omega]}{\int_{\Omega} \Phi[l_t, l_{t-1}, \alpha_t, \alpha_{t-1}, Y(x, x_{t-1}, t), \theta] \cdot \Psi[H(x, t), \omega]} dx \quad \text{eqn 1}$$

Note that the same environmental variable (e.g. snow depth or terrain ruggedness) might be included in both  $Y$  and  $H$  and hence affect both  $\Phi$  [e.g. decreased speed in deep snow or across rugged terrain] and

$\Psi$  (e.g. selection for snow-free or flat localities). Eqn 1 is equivalent to the formulation used (for example) by Rhodes *et al.* (2005, Eqn 1), Forester, Im & Rathouz (2009) and Johnson, Hooten & Kuhn (2013, Eqn 1) and is a generalised form of a redistribution kernel – a widely used mechanistic model of animal movement and habitat selection (see Discussion for recent examples).

The denominator in Eqn 1 is an integral over the entire spatial domain,  $\Omega$ , serving as a normalisation factor to ensure the resulting probability density integrates to one. Whereas in most cases it would be impossible to solve this integral analytically, various forms of numerical (discrete-space) approximations can be used to fit redistribution kernel functions, such as Eqn 1, to data (see Avgar, Deardon & Fryxell 2013a and the Discussion). Here we focus on a simple likelihood-based alternative to such numerical methods, one that can be implemented using common statistical software and is hence accessible to most ecologists. Assuming an exponential form for both  $\Phi$  and  $\Psi$ , MLEs for the parameter vectors  $\theta$  and  $\omega$  can be obtained using conditional logistic regression, where observed positions (cases) are matched with a sample of available positions (controls; Fig. 1 and Appendices S2–S4).

#### A HYPOTHETICAL EXAMPLE

Let us assume we have obtained a set of  $T$  spatial positions, sampled at a unit temporal interval along an animal's path, and that we also have maps of two (temporally stationary) spatial covariates,  $h(x)$  and  $y(x)$ . We shall now assess the statistical support for the following propositions (examples of the sort of hypotheses that could be tested):

- A** The animal is selecting high values of  $h(x)$ .
- B** At the observed temporal scale, and in the absence of variability in  $h(x)$ , the animal's movement is directionally persistent (i.e. consecutive headings are positively correlated), and the degree of this persistence varies with  $y(x)$  (e.g. the animal moves more directionally where  $y(x)$  is lower). The resulting turn angles are von Mises distributed with mean 0 (i.e. left and right turns are equally likely) and a  $y$ -dependent concentration parameter.
- C** At the observed temporal scale, and in the absence of variability in  $h(x)$  and  $y(x)$ , the animal's movement is characterised by gamma distributed step lengths, and the shape of this step-length distribution depends on the time of day (e.g. the animal moves faster during daytime).

Note that these propositions are contingent on the temporal gap between observed relocations (i.e. step duration), as well as on the spatial resolution of our covariate maps,  $h(x)$  and  $y(x)$ . We thus explicitly acknowledge that our inference is scale dependent.

We start by sampling, for each (but the first two) of the observed points along a path ( $x_t$ ,  $t = 3, 4, \dots, T$ ), a set of  $s$  control points (available spatial positions at time  $t$ ;  $x'_{t,i}$ ,  $i = 1, 2, \dots, s$ ), where the probability of obtaining a sample at some distance,  $l'_{t,i}$ , from the previous observed point ( $l'_{t,i} = \|x'_{t,i} - x_{t-1}\|$ ) is given by the gamma PDF:

$$g(l'_{t,i} | b_1, b_2) = \frac{1}{\Gamma(b_1) \cdot b_2^{b_1}} \cdot l'_{t,i}{}^{b_1-1} \cdot e^{-\frac{l'_{t,i}}{b_2}} \quad \text{eqn 2}$$

Here,  $b_1$  and  $b_2$  are initial estimates of the gamma shape and scale parameters (respectively) obtained based on the observed step-length distribution (using either the method of moments or maximum likelihood). As noted earlier, this estimation is confounded by the process of habitat selection, and hence, a method to unravel movement inference from habitat selection is needed. The iSSA will provide estimates of the deviations of these initial values from the unobserved habitat-indepen-



dent shape and scale (Appendices S2–S3). Note that these control sets also could be sampled randomly within some finite spatial domain (e.g. within the maximal observed displacement distance; Appendices S2 and S4). Distance weighted sampling is expected to increase inferential efficiency, resulting in a smaller standard error for a given  $s$  value, but is not a necessity (Forester, Im & Rathouz 2009). In general, any increase in  $T$  and/or  $s$  will result in better approximation of the used and/or availability distributions (respectively), and hence better inference (together with larger computational costs).

Once sampled, control (available) points,  $x'_{t,i}$ , are assigned a value of 0, whereas the observed (used or case) points,  $x_t$ , are assigned a value of 1. The resulting binomial response variable can now be statistically modelled using conditional (case–control) logistic regression, as the likelihood of the observed data is exactly proportional to (Gail, Lubin & Rubinstein 1981; Forester, Im & Rathouz 2009; Duchesne, Fortin & Rivest 2015):

$$\prod_{t=3}^T \frac{\exp[b_3 \cdot h(x_t) + [b_4 + b_5 \cdot y(x_{t-1})] \cdot \cos(\alpha_{t-1} - \alpha_t) + b_6 \cdot l_t + (b_7 + b_8 \cdot D_t) \cdot \ln(l_t)]}{\sum_{i=0}^s \exp[b_3 \cdot h(x'_{t,i}) + [b_4 + b_5 \cdot y(x_{t-1})] \cdot \cos(\alpha_{t-1} - \alpha'_{t,i}) + b_6 \cdot l_{t,i} + (b_7 + b_8 \cdot D_t) \cdot \ln(l'_{t,i})]}, \quad \text{eqn 3}$$

where  $\alpha'_{t,i}$  is the direction of movement from  $x_{t-1}$  to  $x'_{t,i}$ , and  $D_t$  is an indicator variable having the value 1 when  $t$  is daytime and 0 otherwise. Note that the summation in the denominator starts at  $s = 0$  (rather than 1) to indicate that the used step is included in the availability set ( $x'_{t,i=0} = x_t$ ). Also note that it is the inclusion of turn angles that necessitates the exclusion of the first two positions ( $x_{t=1}$  and  $x_{t=2}$ ); if no velocity autocorrelation is modelled, only the first position is excluded. Lastly, note that this formulation implies that the degree of directional persistence is affected by the value of  $y$  at the onset of the step only; in the next section, we provide an example of modelling habitat effects on movement along the step.

Equation 3 is a discrete-choice approximation of Eqn 1 (specifically tailored according to propositions A–C), and we provide its full derivation in Appendix S3. In summary,  $b_3$  is the habitat-selection coefficient (corresponding to proposition A and estimating the only component of the parameter vector  $\omega$  in Eqn 1),  $b_4$  and  $b_5$  are the basal (habitat-independent) and  $y$ -dependent directional persistence coefficients (corresponding to proposition B and estimating two components of the parameter vector  $\theta$  in Eqn 1), and  $b_6$ ,  $b_7$  and  $b_8$  are the modifiers of the step-length shape and scale coefficients (corresponding to proposition C and estimating the remaining components of the parameter vector  $\theta$ ). Once maximum-likelihood estimates are obtained, the shape and scale parameters of the basal step-length distribution can be calculated (Appendix S3), where the shape is given by:  $[(b_1 + b_7) + b_8 \cdot D_t]$ , and the scale is given by:  $[1/(b_2^{-1} - b_6)]$ . Similarly,  $b_4$  can be shown to be an unbiased estimator of the concentration parameter of the (habitat-independent) von Mises turn angles distribution (Duchesne, Fortin & Rivest 2015).

Including movement attributes as covariates in SSA, which we termed here iSSA, thus allows simple likelihood-based estimation of explicit ecological hypotheses within a framework of a mechanistic habitat-mediated movement model. Such hypotheses might include, in addition to those mentioned thus far, long- and short-term target prioritisation (Duchesne, Fortin & Rivest 2015), barrier crossing and avoidance behaviour (Beyer *et al.* 2015), and interactions with conspecifics and intraspecifics (Latombe, Fortin & Parrott 2014; Potts, Mokross & Lewis 2014b; Potts *et al.* 2014a). In fact, many of the facets of the generic approach developed by Langrock *et al.*

(2012) can be included in an iSSA with the MLEs obtained using standard statistical packages. An iSSA thus holds promise as a user-friendly yet versatile approach in the movement ecologist's toolbox. In Appendix S4, we provide practical guidelines for the application of iSSA. In the next sections, we explore the utility of this approach using computer simulations.

## SIMULATIONS

Testing the inferential and predictive capacities of any statistical space-use model is challenging because we are often ignorant of the true process giving rise to the observed patterns, as well as of the true distribution of space-use from which these patterns are sampled (Avgar, Deardon & Fryxell 2013a; Van Moorter *et al.* 2013). To deal with this challenge, we employ here a simple process-based movement

simulation framework. We provide full details of the simulation procedure and its statistical analysis in Appendix S5.

Fine-scale space-use dynamics were simulated using stochastic 'stepping-stone' movement across a hexagonal grid of cells. Each cell,  $x$ , is characterised by habitat quality,  $h(x)$  with spatial autocorrelation set by an autocorrelation range parameter,  $\rho$  ( $=0, 1, 5, 10$  and  $50$ ). For each  $\rho$  value, 1000 trajectories were simulated and then rarefied (by retaining every 100th position). Each of these rarefied trajectories was then separately analysed using RSA and 10 different (i)SSA formulations, including one or more of the following covariates (Table 1): habitat values at the end of each step,  $h(x_t)$ , the average habitat value along each step,  $h(x_{t-1}, x_t)$ , the step length,  $l_t$  ( $=|x_{t-1} - x_t|$ ), its natural-log transformation,  $\ln(l_t)$ , and the statistical interactions between  $l_t$ ,  $\ln(l_t)$  and  $h(x_{t-1}, x_t)$ . Models that included only  $h(x_t)$  and/or  $h(x_{t-1}, x_t)$  correspond to traditionally used SSA (relying on empirical movement distributions with no movement attribute included as covariates; models  $a$ ,  $b$  and  $c$  in Table 1), whereas models that additionally included  $l_t$  and  $\ln(l_t)$  correspond to iSSA. The predictive capacity of the models was estimated based on the agreement between their predicted utilisation distributions (UD) and the 'true' UD, generated by the true underlying movement process. We refer the reader to Appendix S5 for further details.

A separate simulation study was conducted to evaluate the identifiability and estimability of the iSSA parameters as function of sample size and habitat-selection strength (Appendix S6).

## Results

### PARAMETERISATION

All models converged in a timely manner and the convergence time for the most complex model (model  $j$  in Table 1) was approximately 1 CPU sec. Of the 10 (i)SSA formulations specified in Table 1, AIC ranking indicated support for only four ( $d, f, h$  and  $j$ ), all of which include the habitat value at the step's endpoint (with coefficient  $\beta_3$ ) and the step length and its natural logarithm (with coefficients  $\beta_5$  and  $\beta_6$ ) as covariates. Hence,

**Table 1.** The 11 different models fitted here and their relative performance ranking at five different levels of habitat spatial autocorrelation (with 1000 realisations at each level). To enable AIC comparison, RSA's were run with only those positions included in the SSA (i.e. excluding the first position)

Model	Covariates						% Score as best (based on AIC)				
	$R(x_t)$	$R(x_t, x_{t-1})$	$l_t$	$\ln(l_t)$	$R(x_t, x_{t-1})$ $l_t$	$R(x_t, x_{t-1})$ $\ln(l_t)$	$\rho = 0$	$\rho = 1$	$\rho = 5$	$\rho = 10$	$\rho = 50$
RSA	$\beta_{\text{RSA}}$	0	0	0	0	0	0	0	20.3	<b>45.7</b>	<b>44.7</b>
SSA											
<i>a</i>	$\beta_3$	0	0	0	0	0	0	0	0	0	0
<i>b</i>	0	$\beta_4$	0	0	0	0	0	0	0	0	0
<i>c</i>	$\beta_3$	$\beta_4$	0	0	0	0	0	0	0	0	0
iSSA											
<i>d</i>	$\beta_3$	0	$\beta_5$	$\beta_6$	0	0	0	0	0	1.1	1.1
<i>e</i>	0	$\beta_4$	$\beta_5$	$\beta_6$	0	0	0	0	0	0	0
<i>f</i>	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	0	0	10.4	4.4	13.9	24.0	28.4
<i>g</i>	0	0	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	0	0	0	0	0
<i>h</i>	$\beta_3$	0	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	<b>66.2</b>	28.3	10.7	5	2.3
<i>i</i>	0	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	0	0	0	0	0
<i>j</i>	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	23.4	<b>67.3</b>	<b>55.1</b>	23.3	23.5

Bolded numbers mark the best performing model at each level of spatial autocorrelation.

iSSAs better explain our simulated data than traditionally used SSAs (excluding step length as a covariate), but only as long as an endpoint effect (i.e. selection for/against the habitat value at the end of the step) is included. In fact, models that excluded the habitat value at the step's endpoint (models *b*, *e*, *g* and *i*) had AIC scores that were typically two orders of magnitude larger than those including it. In comparison to RSA, iSSA formulations had unequivocal AIC support at low habitat spatial autocorrelation levels, but only partial support at high autocorrelation levels (Table 1).

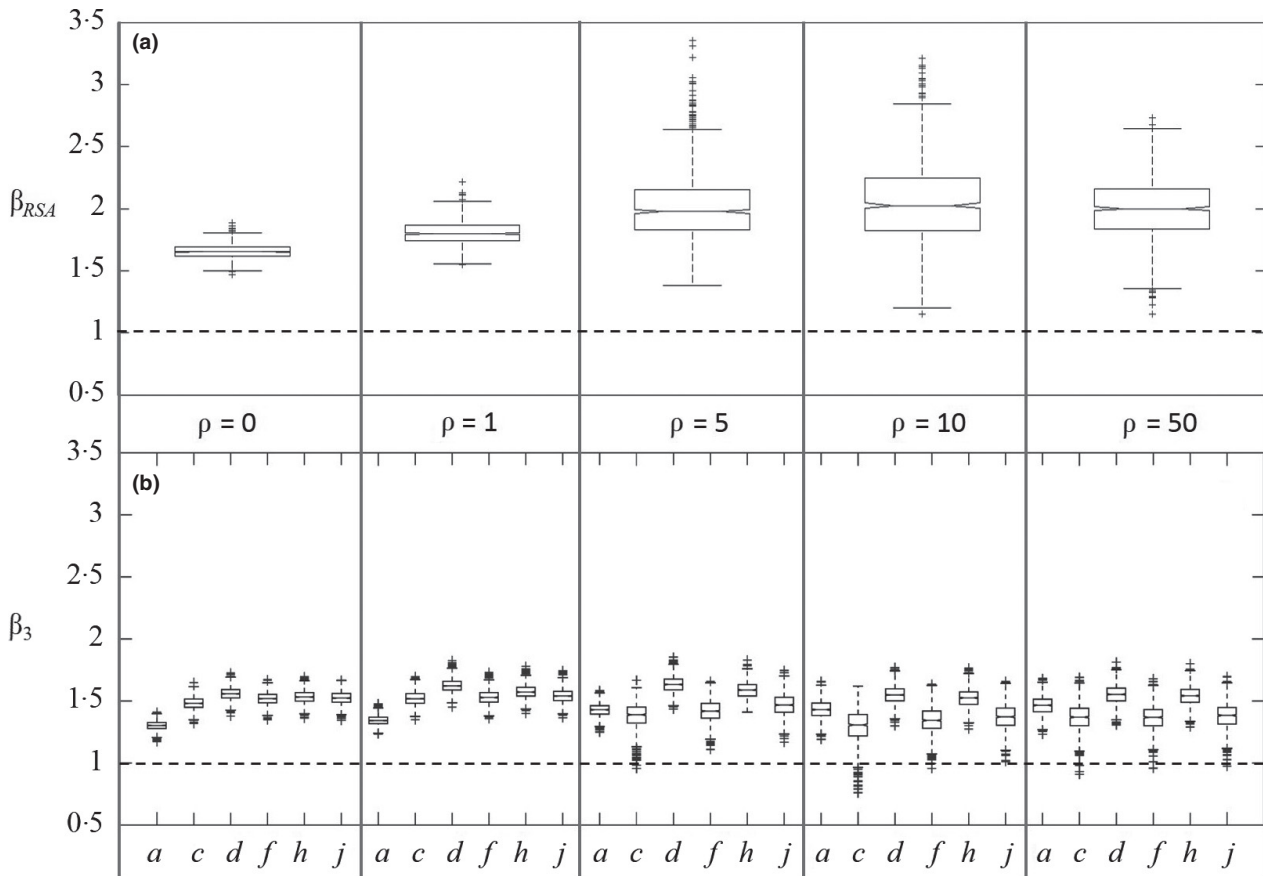
Estimated habitat-selection strengths, as indicated by our RSA and SSA coefficient estimates ( $\beta_{\text{RSA}}$  and  $\beta_3$ , respectively), were appreciably larger than the true habitat-selection strength ( $\omega = 1$ ), and more so for RSA estimates than for SSA (Fig. 2). Note that this in itself does not mean these estimates are 'biased' but rather reflects the inherent difference between the intensity of the true process and that of the emerging pattern, at the scale of observation (see further discussion below). These estimates showed little sensitivity to the level of habitat spatial autocorrelation, although a substantial increase in variance is observed in the RSA case (Fig. 2a). As found before by Forester, Im & Rathouz (2009), the strength of SSA-inferred habitat selection is larger when step lengths are included as a covariate in the analysis (iSSA), but this effect is fairly weak and diminishes as the habitat's spatial autocorrelation increases (Fig. 2b).

Overall, SSA-inferred habitat selections were substantially less variable than RSA-based estimates and showed little sensitivity to the inclusion or exclusion of other covariates in the model fit (Fig. 2). This is not the case, however, for the effect of the mean habitat value along the step ( $\beta_4$ ), which varied substantially with both the level of habitat spatial autocorrelation and the inclusion of an endpoint effect ( $\beta_3$ ). Where  $\beta_3$  was not included in the model fit (models *b*, *e* and *i* in Table 1),  $\beta_4$  increased with  $\rho$ , whereas where  $\beta_3$  was included (models *c*, *f*

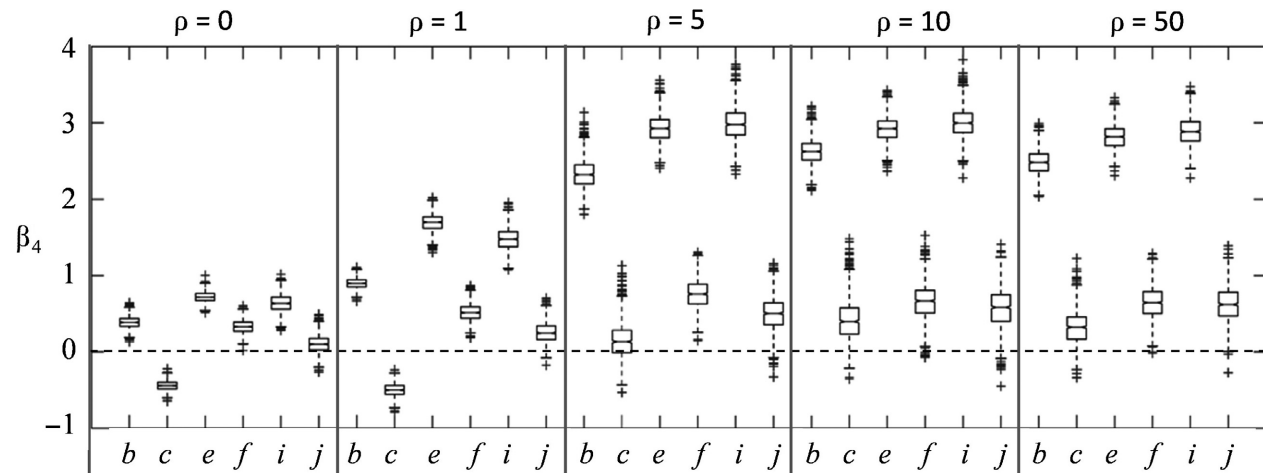
and *j*),  $\beta_4$  was closer to zero (Fig. 3). Interestingly, when only the habitat at the end of the step and the habitat along the step were included in the model (i.e. model *c*; a commonly used SSA formulation), and at low  $\rho$  values ( $=0, 1$ ),  $\beta_4$  was negative, indicating 'selection against' high-quality steps. In fact, this reflects the low probability of observing a 'used' step that traverses high-quality habitat but does not end there.

As explained above (and in Appendices S2 and S3), iSSA coefficients affiliated with the step length ( $\beta_5$ ) and its natural logarithm ( $\beta_6$ ), when combined with the estimated shape and scale values of the observed step-length distribution ( $\beta_1$  and  $\beta_2$ ; on which sampling was conditioned; Appendices S3 and S5), could be used to infer the shape and scale of the 'habitat-independent' step-length distribution [i.e. assuming  $h(x_t, x_{t+1}) = 0$ ]. Under most imaginable scenarios, we would expect this basal movement kernel to be wider (i.e. with larger mean) than the observed one, as animals tend to linger in preferred habitats and hence display more restricted movements compared to the basal expectation. Indeed, the mean of these inferred distributions (the product of their shape and scale:  $\frac{(\beta_1 + \beta_6)}{(\beta_2 - \beta_5)}$ ) corresponds exactly to the observed mean, as long as no other covariates are included in the analysis (model *x* in Fig. 4). Once other covariates are included in the model (and hence habitat selection is at least partially accounted for), inferred mean step-length values were significantly higher from the observed values, showed little sensitivity to model structure, but increased with  $\rho$  (as do the observed mean step lengths). One exception is model *g*, which strongly underestimated the mean step length at moderate-high  $\rho$  values as it does not include any main habitat effects.

Even at high  $\rho$  values, inferred mean step length slightly but consistently underestimates the true habitat-independent step-length distribution (calculated by simulating the process based on Eqn S5-1 with  $\omega = 0$ ; Fig. 4). This bias is a result of an iSSA's limited capacity to account for the full movement



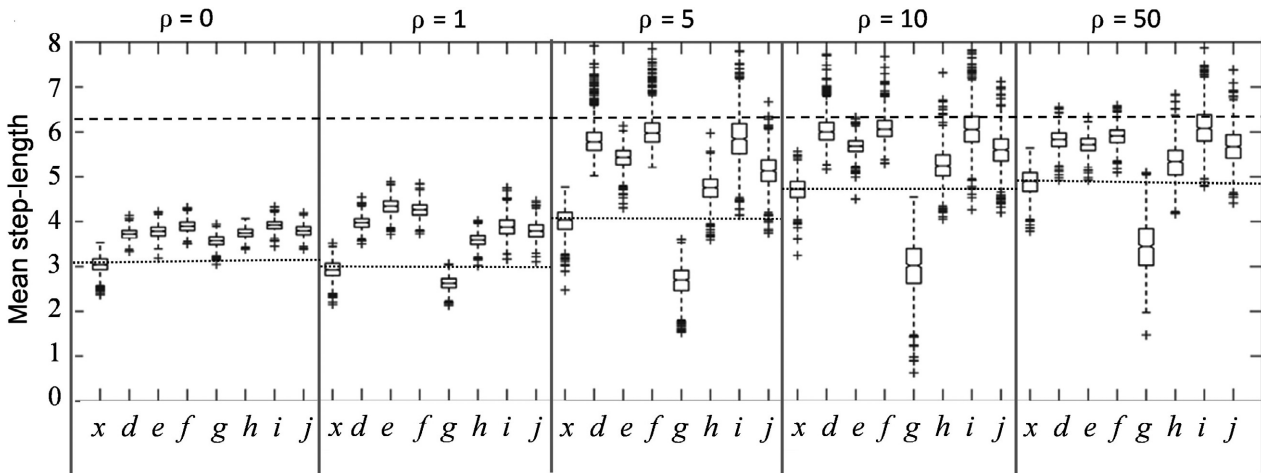
**Fig. 2.** Statistically inferred habitat-selection coefficient estimates for RSA (a) and SSA (b; letters along the *x*-axis refer to the SSA formulations listed in Table 1), for five levels of habitat spatial autocorrelation,  $\rho$ . Each box-and-whiskers is based on 1000 independent estimates, where the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers (i.e. within approximately 99% coverage if the data are normally distributed), and outliers are plotted individually. Horizontal dashed lines represent the true habitat-selection intensity,  $\omega = 1$ . See Appendix S5 for further details.



**Fig. 3.** Statistically inferred effects of the mean habitat along the step. The dashed line represents no effect. Other details are as in Fig 2.

process as it unfolds in between observations. The animal does not actually travel along the straight lines that we term ‘steps’ and, even if it would, the mean habitat value along the step does not exactly correspond to its probability to travel farther. As long as the scale of the observation is coarser than the scale

of the underlying movement process, the animal’s true movement capacity is never fully manifested in the observed relocation pattern and is thus always underestimated. Note, however, that this bias is negligibly small where the spatial autocorrelation of habitats is high ( $\rho > 1$ ; Fig. 4).



**Fig. 4.** Mean of gamma step-length distributions (displacement in spatial units per  $\Delta t$ ; Appendix S5) inferred based on the different iSSA formulations (see Table 1). Model  $x$  is a null model, including only the step length and its natural logarithm (with no habitat effects), added here to demonstrate that the conditional logistic regression produces unbiased MLEs. The dotted lines correspond to the observed mean step length across all 1000 realisations at each of the five levels of habitat spatial autocorrelation. The dashed line corresponds to the ‘true’ habitat-free mean step length, calculated by simulating the process using Eqn S5.1 but with  $\omega = 0$ . Other details are as in Fig. 2.

Finally, despite apparent support for iSSA formulations including interaction between the step length and habitat quality along the step (Table 1; models  $h$  and  $j$ ), the estimated values of the interaction coefficients,  $\beta_7$  and  $\beta_8$ , mostly overlapped zero (Appendix S7). Generally speaking, the mean habitat value along the step has a weak negative effect on both the shape (through  $\beta_8$ ) and the scale (through its inverse relationship with  $\beta_7$ ) of the step-length distribution – long steps are less likely through high-quality habitats.

#### PREDICTIVE CAPACITIES

At approximate steady state, RSA-based UD predictions are slightly more accurate and precise than SSA-based predictions (Fig. 5 and Appendix Table S8). The RSA’s predictive capacity increases with  $\rho$  (while its precision dramatically decreases), whereas the opposite is true for SSA predictions, where the minimum  $KLD$  value (Kullback-Leibler Divergence; see Appendix S5) is reached when  $\rho = 0$  (Appendix Table S8).  $KLD$  values coarsely mirror the AIC ranking of the different SSA formulations in distinguishing those that include an endpoint effect ( $\beta_3$ ), but the best performing formulations based on  $KLD$  are simpler than the ones selected based on AIC (Tables 1 and S8). That said, all iSSA formulations including an endpoint effect performed well overall, with  $G_{KLD}$  scores (a measure of goodness of fit; Appendix S5) ranging from  $\sim 0.84$  (model  $f$  when  $\rho = 50$ ) to  $\sim 0.98$  (model  $d$  when  $\rho = 0$ ). For reference, the  $G_{KLD}$  scores for RSA-based predictions ranged from  $\sim 0.98$  ( $\rho = 0$ ) to  $\sim 0.99$  ( $\rho = 50$ ).

To test the sensitivity of the models’ predictions to the sampling scale (see Appendix S9 for relating  $\rho$  to the sampling scale), we generated predicted UD, both SSA-based and RSA-based, across a highly autocorrelated landscape ( $\rho = 50$ ) using parameter estimates obtained from samples of a random landscape ( $\rho = 0$ ), and vice versa. RSA-based

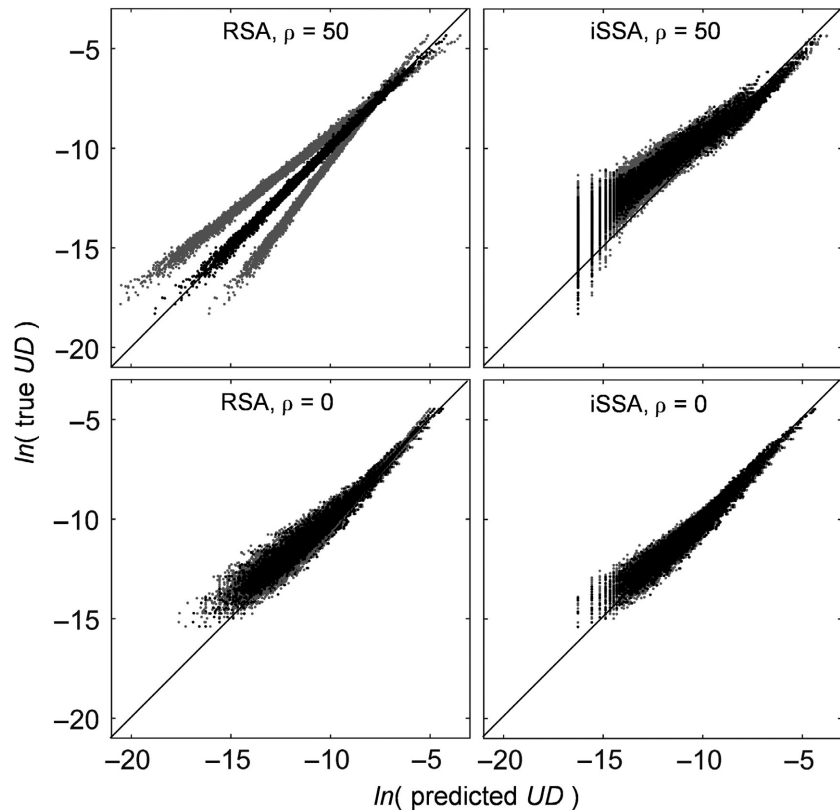
predictions were robust to these scale mismatches, with  $G_{KLD}$  scores of  $\sim 0.98$  and  $\sim 0.96$ , for the  $\rho = 50$  landscape (with parameter estimates based on  $\rho = 0$  data) and the  $\rho = 0$  landscape (with parameter estimates based on  $\rho = 50$  data), respectively. Similarly, all step selection models including an endpoint effect performed well, with  $G_{KLD}$  scores ranging from  $\sim 0.94$  (model  $f$ ) to  $\sim 0.98$  (model  $h$ ) for the  $\rho = 0$  landscape (with parameter estimates based on  $\rho = 50$  data), and  $G_{KLD}$  scores ranging from  $\sim 0.83$  (model  $j$ ) to  $\sim 0.97$  (model  $a$ ) for the  $\rho = 50$  landscape (with parameter estimates based on  $\rho = 0$  data). Overall, iSSA-based predictive capacity remained mostly unaltered by mismatches between the data’s landscape structure and the structure of the landscape on which projections are made.

As can be expected, step selection models predict transient UD, better than the inherently stationary RSA (except when  $\rho = 50$ ; Appendix Table S8). In comparison with steady-state predictions, complex iSSA-based predictions perform better than simpler ones (Appendix Table S8). As for the steady-state predictions, all iSSA formulations including an endpoint effect performed well in predicting transient UD, with  $G_{KLD}$  scores ranging from  $\sim 89\%$  (model  $h$  when  $\rho = 0$ ) to  $\sim 0.98$  (model  $d$  when  $\rho = 1$ ).  $G_{KLD}$  scores for RSA-based predictions showed substantial sensitivity to the level of spatial autocorrelation, ranging from  $\sim 0.69$  ( $\rho = 0$ ) to  $\sim 0.97$  ( $\rho = 50$ ).

#### ISSA IDENTIFIABILITY AND ESTIMABILITY

Results are presented in detail in Appendix S6. In short, our analysis revealed that, under the test scenario, all iSSA parameters are fully identifiable, that estimates are unbiased in relation to the true values of the kernel generating functions, and that an increase in sample size beyond  $\sim 400$  observed positions does not seem to substantially enhance precision (and hence estimability). That said, our results also





**Fig. 5.** Log–log plots of the true UD vs the predicted UD. Each dot represents the utilisation probability of a single spatial cell. Black dots correspond to the median parameter estimates, whereas grey dots correspond to the 2.5 and 97.5 percentiles of the estimated parameters distribution. Black diagonal lines represent a perfect 1:1 mapping – dots appearing above these lines are spatial cells where the true UD value exceeded the predicted UD value (under-predictions), whereas dots appearing below these lines represent over-predictions. Note that iSSA results are presented for the simplest iSSA including an endpoint effect, formulation *d* in Table 1.

indicate that inferential accuracy of movement related parameters may be highly variable, leading to compromised precision (with up to 1000% departure from the true value; Appendix S6) even at a fairly large sample size. This may be particularly true given the inherent trade-off between sampling extent and frequency (Fieberg 2007). Estimability of certain parameters, under certain scenarios, may thus be weak and must be evaluated on a case-by-case basis.

## Discussion

The ideas, simulations and results presented above are aimed at providing a comprehensive assessment of using integrated step selection analysis, iSSA, with emphasis on its predictive capacity. The iSSA allows simultaneous inference of habitat-dependent movement and habitat selection and is hence a powerful tool for both evaluating ecological hypotheses and predicting ecological patterns. We have shown that iSSA-based habitat-selection inference is relatively insensitive to model structure and landscape configuration, and that iSSA-based UD predictions perform well across different temporal and spatial scales (we discuss the connection between the temporal resolution of the data and the habitat spatial autocorrelation in Appendix S9). On the other hand, our results indicate that movement and habitat selection may not be completely separable once observations are collected at a coarser temporal resolution than the underlying behavioural process. Consequently, stationary RSA-based predictions, whereas much simpler to obtain, provide slight but consistent better fit

to the true UD when the time-scale is long (and hence approaches the steady-state limit).

Two caveats are in place here. First, in our analysis the definition of the availability set for the RSA was exact (i.e. the entire domain), a situation that seldom occurs in empirical studies where availability is unknown. This is not the case for iSSA where the availability set always can be adequately defined (but is conditional on the temporal resolution of the positional data). Secondly, the high variability characterising the RSA coefficient estimates, and its resulting predictions (Figs 2 and 5) indicate substantial risk of erroneous inference. This may be particularly true when sample size is smaller than the relatively large sample used here, resulting in data that are not adequate unbiased samples of the steady-state UD, which is likely the case in most empirical studies. The more mechanistic nature of the iSSA makes it less sensitive to stochastic differences between specific realisations of the space-use process (e.g. due to differences in landscape configuration) and thus leads to more precise inference. Hence, even if the sole objective of a given study is to predict the long-term (steady-state) utilisation distribution, the more complicated iSSA-based predictions might be more reliable than those based on RSA. Moreover, in many real-world ecological scenarios, a steady state is never reached, and consequently, the static RSA-based approach is less appropriate than the dynamical iSSA. We thus conclude that iSSA should be the method of choice whenever: (i) RSA availability cannot be properly defined, (ii) predicting across a landscape different from the landscape used for parametrisation, (iii) the data used for parametrisation are not

an adequate sample of the true steady-state UD or (iv) predicting transient space-use dynamics.

Many movement and selection processes could be considered plausible, and the particular details of the mechanistic model used to simulate space-use data might substantially alter our conclusions. Our aim here was to use the simplest, and hence most general, mechanistic process imaginable, leading us to choose a stepping-stone movement process as our pattern-generating process. Interestingly, this simple process, governed by only two parameters (Eqn S5.1), gave rise to complex patterns once rarefied. In particular, the emerging step-length distributions fit remarkably well with a gamma distribution, with shape and scale that reflect the underlying landscape structure. Note that this is a purely phenomenological description of the movement kernel, as the true underlying process had a fixed, habitat-independent movement parameter (Appendix S5). Ideally, a truly mechanistic approach will involve maximising the likelihood over all possible paths the animal might have taken between two observed locations, and hence allowing inference of the true underlying process (Mattiopoulos 2003). In most cases, however, this approach is forbiddingly computationally expensive. We showed that the approximation based on samples of straight-line movements between observed positions, which is the underlying assumption of any SSA, performs well over a range of conditions. An iSSA thus provides a reasonable compromise between computationally intensive mechanistic models and the purely phenomenological RSA.

According to Barnett & Moorcroft (2008), the steady-state UD should scale linearly with the underlying habitat-selection function  $\Psi$  (Eqn 1) when informed movement capacity exceeds the scale of spatial variation in  $\Psi$ , but should scale with the square of  $\Psi$  if informed movement capacity is much shorter than the scale of habitat variation. In the particular case of the exponential habitat-selection function used here (Eqn S5.1), we would thus expect the following loglinear relationship:  $\ln [UD(x)] = a + b \cdot \omega \cdot h(x)$ , where  $a$  is a scaling parameter [the utilisation probability where  $h(x) = 0$ ], and  $b$  ( $1 \leq b \leq 2$ ) is some increasing function of the habitat spatial autocorrelation,  $\rho$ . Our results, emerging from a very different model than the continuous-space continuous-time analytical approximation of Barnett & Moorcroft (2008), corroborate this expectation. The slope of the loglinear regression model described above increases from  $b \approx 1.4$  to  $b \approx 2$  as  $\rho$  increases from 0 to 50 (Appendix S10). RSA-based coefficient estimates,  $\beta_{RSA}$ , closely mirror this pattern, increasing from  $\sim 1.6$  to  $\sim 2$  as  $\rho$  increases (Fig. 2a). Hence, as can be expected from a phenomenological model, RSA-based inference reflects the steady-state UD rather than the underlying habitat-selection process.

Recent years have seen a proliferation of sophisticated modelling approaches aimed at mechanistically capturing animal space-use behaviours. Many of these models share the theoretical underpinning of iSSA (as formulated in Eqn 1), relying on a depiction of animal space-use as emerging from the product of a resource-selection process and a selection-independent movement kernel (e.g. Rhodes

*et al.* 2005; Getz & Saltz 2008; Avgar, Deardon & Fryxell 2013a; Potts *et al.* 2014a; Beyer *et al.* 2015). Unlike the iSSA, however, fitting these kernel-based models to empirical data relies on complex, and often specifically tailored likelihood maximisation algorithms (namely discrete-space approximations of the integral in Eqn 1). The statistical machinery used in iSSA, based on obtaining a small set of random samples from an inclusive availability domain, is accessible to most ecologists because it relies on software that is already used (Thurfjell, Ciuti & Boyce 2014). Through the addition of appropriate covariates and interaction terms, iSSA can moreover address many of the questions that were the focus of other kernel-based approaches, such as home-range behaviour (Rhodes *et al.* 2005), memory-use (Avgar, Deardon & Fryxell 2013a; Merkle, Fortin & Morales 2014; Avgar *et al.* 2015; Schlägel & Lewis 2015), habitat-dependent habitat selection (Potts *et al.* 2014a) and barrier effects (Beyer *et al.* 2015). Hence, iSSA allows ecologists to tackle complicated questions using simple tools.

To conclude, our work complements several recent contributions advocating the use of movement covariates within step selection analysis (Forester, Im & Rathouz 2009; Johnson, Hooten & Kuhn 2013; Warton & Aarts 2013; Duchesne, Fortin & Rivest 2015). We believe a convincing body of theoretical evidence now indicates the suitability of integrated step selection analysis as a general, flexible and user-friendly approach for both evaluating ecological hypotheses and predicting future ecological patterns. Our work highlights the importance of including endpoint effects in the analysis together with some caveats regarding the interpretation of SSA results, specifically when dealing with the effects of the habitat along the step. We also recommend careful consideration of parameter estimability, particularly with regard to the movement components of the model, which may be prone to strong cross-correlations (as discussed in Appendix S6). Based on our current experience in applying iSSA to empirical data (T. Avgar, work in progress) we have provided practical guidelines in Appendix S4. Additional theoretical work is needed to investigate the effects of the underlying movement process on iSSA performance, as well as to come up with computationally efficient iSSA-based simulations to enable rapid generation of predicted utilisation distribution (as discussed in Appendix S1). Most importantly, the utility of iSSA must now be evaluated by applying it to real data sets, and using it to solve real ecological problems.

## Acknowledgements

TA gratefully acknowledges supported by the Killam and Banting Postdoctoral Fellowships. MAL gratefully acknowledges NSERC Discovery and Accelerator Grants and a Canada Research Chair. MSB thanks NSERC and the Alberta Conservation Association for funding. The authors thank L. Broitman for designing Fig. 1 and C. Prokopenko for helpful editorial comments, and Dr. Geert Aarts, Dr. John Fieberg and an anonymous reviewer, for their instructive comments and suggestions.

## Data accessibility

All data used in this paper were simulated as described in the online Supporting Information.

## References

- Arthur, S.M., Manly, B.F.J., McDonald, L.L. & Garner, G.W. (1996) Assessing habitat selection when availability changes. *Ecology*, **77**, 215–227.
- Avgar, T., Deardon, R. & Fryxell, J.M. (2013a) An empirically parameterized individual based model of animal movement, perception and memory. *Ecological Modelling*, **251**, 158–172.
- Avgar, T., Mosser, A., Brown, G.S. & Fryxell, J.M. (2013b) Environmental and individual drivers of animal movement patterns across a wide geographical gradient. *Journal of Animal Ecology*, **82**, 96–106.
- Avgar, T., Baker, J.A., Brown, G.S., Hagens, J., Kittle, A.M., Mallon, E.E. *et al.* (2015) Space-use behaviour of woodland caribou based on a cognitive movement model. *Journal of Animal Ecology*, **84**, 1059–1070.
- Baasch, D.M., Tyre, A.J., Millsaugh, J.J., Hygnstrom, S.E. & Vercouteren, K.C. (2010) An evaluation of three statistical methods used to model resource selection. *Ecological Modelling*, **221**, 565–574.
- Barnett, A.H. & Moorcroft, P.R. (2008) Analytic steady-state space use patterns and rapid computations in mechanistic home range analysis. *Journal of Mathematical Biology*, **57**, 139–159.
- Benhamou, S. (2014) Of scales and stationarity in animal movements. *Ecology Letters*, **17**, 261–272.
- Beyer, H.L., Gurarie, E., Borger, L., Panzacchi, M., Basille, M., Herfindal, I., Van Moorter, B., Lele, S. & Matthiopoulos, J. (2015) 'You shall not pass!': quantifying barrier permeability and proximity avoidance by animals. *Journal of Animal Ecology*, **85**, 43–53.
- Boyce, M.S. & McDonald, L.L. (1999) Relating populations to habitats using resource selection functions. *Trends in Ecology & Evolution*, **14**, 268–272.
- Boyce, M.S., Mao, J.S., Merrill, E.H., Fortin, D., Turner, M.G., Fryxell, J.M. & Turchin, P. (2003) Scale and heterogeneity in habitat selection by elk in Yellowstone National Park. *Ecoscience*, **14**, 421–431.
- Boyce, M.S., Johnson, C.J., Merrill, E.H., Nielsen, S.E., Solberg, E.J. & Van Moorter, B. (2015) Can habitat selection predict abundance? *Journal of Animal Ecology*, **85**, 11–20.
- Codling, E.A., Plank, M.J. & Benhamou, S. (2008) Random walk models in biology. *Journal of the Royal Society Interface*, **5**, 813–834.
- Compton, B., Rhymer, J. & McCollough, M. (2002) Habitat selection by wood turtles (*Clemmys insculpta*): an application of paired logistic regression. *Ecology*, **83**, 833–843.
- Duchesne, T., Fortin, D. & Courbin, N. (2010) Mixed conditional logistic regression for habitat selection studies. *Journal of Animal Ecology*, **79**, 548–555.
- Duchesne, T., Fortin, D. & Rivest, L. (2015) Equivalence between step selection functions and biased correlated random walks for statistical inference on animal movement. *PLoS ONE*, **10**, e0122947.
- Elton, C.S. (1927, 2001) *Animal ecology*. University of Chicago Press, Chicago, IL, USA.
- Fagan, W. & Calabrese, J. (2014) The correlated random walk and the rise of movement ecology. *Bulletin of the Ecological Society of America*, **95**, 204–206.
- Fieberg, J. (2007) Kernel density estimators of home range: smoothing and the autocorrelation red herring. *Ecology*, **88**, 1059–1066.
- Fordham, D.A., Shoemaker, K.T., Schumaker, N.H., Akçakaya, H.R., Clisby, N. & Brook, B.W. (2014) How interactions between animal movement and landscape processes modify local range dynamics and extinction risk. *Biology Letters*, **10**, 20140198.
- Forester, J.D., Im, H.K. & Rathouz, P.J. (2009) Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, **90**, 3554–3565.
- Fortin, D., Beyer, H.L., Boyce, M.S., Smith, D.W., Duchesne, T. & Mao, J.S. (2005) Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park. *Ecology*, **86**, 1320–1330.
- Gail, M.H., Lubin, J.H. & Rubinstein, L.V. (1981) Likelihood calculations for matched case-control studies and survival studies with tied death times. *Biometrika*, **68**, 703–707.
- Getz, W.M. & Saltz, D. (2008) A framework for generating and analyzing movement paths on ecological landscapes. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 19066–19071.
- Hjermann, D.Ø. (2000) Analyzing habitat selection in animals without well-defined home ranges. *Ecology*, **81**, 1462–1468.
- Johnson, D.S., Hooten, M.B. & Kuhn, C.E. (2013) Estimating animal resource selection from telemetry data using point process models. *Journal of Animal Ecology*, **82**, 1155–1164.
- Jonsen, I.D., Myers, R.A. & Flemming, J.M. (2003) Meta-analysis of animal movement using state-space models. *Ecology*, **84**, 3055–3063.
- Kareiva, P.M. & Shigesada, N. (1983) Analyzing insect movement as a correlated random-walk. *Oecologia*, **56**, 234–238.
- Keating, K.A. & Cherry, S. (2009) Modeling utilization distributions in space and time. *Ecology*, **90**, 1971–1980.
- Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. & Morales, J.M. (2012) Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, **93**, 2336–2342.
- Latombe, G., Fortin, D. & Parrott, L. (2014) Spatio-temporal dynamics in the response of woodland caribou and moose to the passage of grey wolf. *Journal of Animal Ecology*, **83**, 185–198.
- Lele, S.R., Merrill, E.H., Keim, J. & Boyce, M.S. (2013) Selection, use, choice and occupancy: clarifying concepts in resource selection studies. *Journal of Animal Ecology*, **82**, 1183–1191.
- Lima, S.L. & Zollner, P.A. (1996) Towards a behavioral ecology of ecological landscapes. *Trends in Ecology & Evolution*, **11**, 131–135.
- Manly, B.F.J., McDonald, L.L., Thomas, D.L., McDonald, T.L. & Erickson, W.P. (2002) *Resource selection by animals: statistical design and analysis for field studies*, 2nd edn. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Matthiopoulos, J. (2003) The use of space by animals as a function of accessibility and preference. *Ecological Modelling*, **159**, 239–268.
- McCracken, M.L., Manly, B.F.J. & Heyden, M.V. (1998) The use of discrete-choice models for evaluating resource selection. *Journal of Agricultural, Biological, and Environmental Statistics*, **3**, 268–279.
- McDonald, L., Manly, B., Huettmann, F. & Thogmartin, W. (2013) Location-only and use-availability data: analysis methods converge (G. Hays, Ed.). *Journal of Animal Ecology*, **82**, 1120–1124.
- Merkle, J., Fortin, D. & Morales, J. (2014) A memory-based foraging tactic reveals an adaptive mechanism for restricted space use. *Ecology Letters*, **17**, 924–931.
- Moorcroft, P.R., Lewis, M.A. & Crabtree, R.L. (2006) Mechanistic home range models capture spatial patterns and dynamics of coyote territories in Yellowstone. *Proceedings of the Royal Society B*, **273**, 1651–1659.
- Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse, P.E. (2008) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 19052–19059.
- Northrup, J., Hooten, M., Anderson, C.R. Jr & Wittemyer, G. (2013) Practical guidance on characterizing availability in resource selection functions under a use-availability design. *Ecology*, **94**, 1456–1463.
- Patterson, T.A., Thomas, L., Wilcox, C., Ovaskainen, O. & Matthiopoulos, J. (2008) State-space models of individual animal movement. *Trends in Ecology & Evolution*, **23**, 87–94.
- Potts, J.R., Mokross, K. & Lewis, M.A. (2014b) A unifying framework for quantifying the nature of animal interactions. *Journal of the Royal Society Interface*, **11**, 20140333.
- Potts, J.R., Bastille-Rousseau, G., Murray, D.L., Schaefer, J.A. & Lewis, M.A. (2014a) Predicting local and non-local effects of resources on animal space use using a mechanistic step-selection model. *Methods in Ecology and Evolution*, **5**, 253–262.
- Rhodes, J., McAlpine, C., Lunney, D. & Possingham, H. (2005) A spatially explicit habitat selection model incorporating home range behavior. *Ecology*, **86**, 1199–1205.
- Schlägel, U. & Lewis, M. (2015) Detecting effects of spatial memory and dynamic information on animal movement decisions. *Methods in Ecology and Evolution*, **11**, 1236–1246.
- Thurfjell, H., Ciuti, S. & Boyce, M.S. (2014) Applications of step-selection functions in ecology and conservation. *Movement Ecology*, **2**, 4.
- Turchin, P. (1991) Translating foraging movements in heterogeneous environments into the spatial distribution of foragers. *Ecology*, **72**, 1253–1266.
- Turchin, P. (1998) *Quantitative analysis of movement*. Sinauer, Sunderland, MA.
- Van Moorter, B., Visscher, D., Herfindal, I., Basille, M. & Mysterud, A. (2013) Inferring behavioural mechanisms in habitat selection studies – getting the null-hypothesis right for functional and familiarity responses. *Ecography*, **36**, 323–330.
- Warton, D. & Aarts, G. (2013) Advancing our thinking in presence-only and used-available analysis. *Journal of Animal Ecology*, **82**, 1125–1134.

Received 18 August 2015; accepted 4 December 2015

Handling Editor: Luca Börger

## Supporting Information

Additional Supporting Information may be found in the online version of this article.

**Appendix S1.** From step selection to utilisation distribution.

**Appendix S2.** Inferring step-length distributions.

**Appendix S3.** Deriving an iSSA likelihood function.

**Appendix S4.** iSSA practical user guide.

**Appendix S5.** Simulation experiments.

**Appendix S6.** Evaluating the iSSA parameter identifiability and estimability.

**Appendix S7.**  $\beta_7$  and  $\beta_8$ .

**Appendix S8.** Predictive performance.

**Appendix S9.** Interpreting  $\rho$ .

**Appendix S10.** Habitat selection and utilisation distribution.

**Appendix S11.** Appendices reference list.