Journal of Animal Ecology

Energy-based step selection analysis: modelling the energetic drivers of animal movement and habitat use

Journal:	*Journal of Animal Ecology*
Manuscript ID	JAE-2021-00397.R2
Manuscript Type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Klappstein, Natasha; University of Alberta, Potts, Jonathan; University of Sheffield, School of Mathematics and Statistics Michelot, Théo; University of Saint Andrews, School of Mathematics and Statistics; Börger, Luca; Swansea University, Biosciences Pilfold, Nicholas; San Diego Zoo Wildlife Alliance Science, Conservation Science and Wildlife Health Lewis, Mark; University of Alberta, Mathematics and Statistical Sciences Derocher, Andrew; University of Alberta, Biology
Key-words:	animal movement, energetics, energy landscapes, habitat selection, movement ecology, optimal foraging theory, polar bear, step selection functions

Energy-based step selection analysis: modelling the energetic drivers of animal movement and habitat use

Natasha J. Klappstein[1,*], Jonathan R. Potts[2], Théo Michelot[3], Luca Börger[4,5], Nicholas W. Pilfold[6], Mark A. Lewis[1,7], Andrew E. Derocher[1]

[1] Department of Biological Sciences, University of Alberta, Edmonton, Canada

[2] School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield S3 7RH, UK

[3] Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, UK

[4] Department of Biosciences, Swansea University, Swansea, UK

[5] Centre for Biomathematics, College of Science, Swansea University, Swansea, UK

[6] Conservation Science and Wildlife Health, San Diego Zoo Wildlife Alliance, San Diego, USA

[7] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Canada

[*] *Corresponding Author:* nklappst@ualberta.ca

Abstract

1. The energetic gains from foraging and costs of movement are expected to be key drivers of animal decision-making, as their balance is a large determinant of body condition and survival. This fundamental perspective is often missing from habitat selection studies, which mainly describe correlations between space use and environmental features, rather than the mechanisms behind these correlations.

2. To address this gap, we present a novel parameterisation of step selection functions (SSFs), that we term the energy selection function (ESF). In this model, the likelihood of an animal selecting a movement step depends directly on the corresponding energetic gains and costs, and we can therefore assess how moving animals choose habitat based on energetic considerations.

3. The ESF retains the mathematical convenience and practicality of other SSFs and can be quickly fitted using standard software. In this paper, we outline a workflow, from data-gathering to statistical analysis, and use a case study of polar bears (*Ursus maritimus*) to demonstrate application of the model.

4. We explain how defining gains and costs at the scale of the movement step allows us to include information about resource distribution, landscape resistance, and movement patterns. We further

demonstrate this process with a case study of polar bears, and show how the parameters can be interpreted in terms of selection for energetic gains and against energetic costs.

5. The ESF is a flexible framework that combines the energetic consequences of both movement and resource selection, thus incorporating a key mechanism into habitat selection analysis. Further, because it is based on familiar habitat selection models, the ESF is widely applicable to any study system where energetic gains and costs can be derived, and has immense potential for methodological extensions.

Key words: animal movement, energetics, energy landscapes, habitat selection, movement ecology, optimal foraging theory, polar bear, step selection functions

1 Introduction

Understanding the mechanisms behind the spatial distributions of animals is a core consideration for animal ecology (Kays et al., 2015). In addition to other considerations (e.g., predator-avoidance), space use patterns are largely driven by the movement decisions made by animals whilst foraging. To increase their chances of survival, animals should distribute themselves in space so as to maximize their access to energetically rich resources, while minimizing the costs of travel (Pyke et al., 1977; Pyke, 2019). Consequently, models of animal movement often attempt to link space use to optimal foraging decisions, and to uncover the eco-evolutionary mechanisms behind observed spatial patterns.

Energy-based models are a tool to assess how animals make decisions based on movement costs and associated nutritional benefits. Despite long-standing interest in these fundamental mechanisms of space use, it remains a challenge to consider both energy expenditure and acquisition in a unified framework (Owen-Smith et al., 2010, but see Hooten et al. 2019). Optimal foraging research often focuses on foraging benefits, using approximate measures of forage quality or resource availability (e.g. Bastille-Rousseau et al., 2020), which may not be proportional to energetic gains. Even when energy intake can be evaluated more realistically (e.g. the energetic profitability of resources combined with biomass; Fortin et al., 2003), movement costs are often ignored or assumed to increase linearly with time and/or distance. In reality, the costs vary depending on factors such as movement speed (Taylor et al., 1970), mode of transport (Griffen, 2018), and environmental conditions (e.g. topography, weather, substrate; Crête & Larivière, 2003; Wilson et al., 2012). Environmentally-varying movement costs can be quantified in a framework known as an energy landscape, which can incorporate factors such as air velocity for birds (Shepard et al., 2013), water depth for diving animals (Wilson et al., 2012), or habitat type (Pagano et al., 2020). Energy landscapes are useful to quantify the energetic costs of moving through heterogeneous or dynamic environments, and there are clear benefits to integrating them with foraging theory.

Habitat selection models assess how animals distribute themselves in space relative to environmental features, and therefore may be modified to assess the energetic contributions of movement decisions. Resource selection functions (RSFs) have been used to jointly estimate the effects of foraging resources and energetically-costly environmental features on large-scale space-use (Long et al., 2014). However, that approach does not explicitly model movement, and it therefore ignores costs incurred at the scale of the movement step (i.e. movement between successive recorded locations). At this scale, animals may make decisions primarily based on the need to maximize energy intake, minimize energy use, or balance the two

(Schoener, 1971; Shepard et al., 2009; Cornioley et al., 2016). Although the idea of energy-based movement models is not entirely new, methods remain analytically complex. For example, Hooten et al. (2019) described a movement model with a binary response "recharge" function (e.g. decision to move to food patch), based on a latent process for the physiological state of an animal (e.g. its satiation or cumulative energy balance). A simpler, yet effective, approach could be to extend step selection functions (SSFs; i.e. movement-based habitat selection models) to answer similar questions in a widely used, flexible framework.

In this paper, we propose an energy-based SSF, that we refer to in short-hand as the energy selection function (ESF). The ESF models the effect of energetic gains and costs on an animal's movement decisions, therefore integrating movement and habitat selection in an energetic context. Mathematically, the ESF is a special case of SSFs, and can be implemented using standard logistic regression software, facilitating its uptake in applied studies. However, the ESF is conceptually different in defining movement and habitat availability through covariates that can be explicitly linked to energetic trade-offs in decision-making. To demonstrate our approach, we provide practical guidance to implement the ESF and define covariates, and further outline the workflow and interpretation through an illustrative case study of polar bears (*Ursus maritimus*) in the Beaufort Sea, Canada.

2 Materials and Methods

2.1 The ESF

We present the energy selection function (ESF) as an energy-based step selection model, quantifying responses to both energy gain and expenditure. The ESF defines the likelihood of a step ending at location $\boldsymbol{y}$ given that it started at location $\boldsymbol{x}$ as

$$f(\boldsymbol{y}|\boldsymbol{x}) = \frac{w(\boldsymbol{x},\boldsymbol{y})}{\int_{\boldsymbol{z}\in\Omega} w(\boldsymbol{x},\boldsymbol{z})d\boldsymbol{z}}, \tag{1}$$

where Ω is the study region, the denominator is a normalization constant that ensures the ESF likelihood is a probability density function with respect to $\boldsymbol{y}$ (Forester et al., 2009; Potts et al., 2014), and with energy preference modelled as,

$$w(\boldsymbol{x},\boldsymbol{y}) = \exp\{\beta_1 G(\boldsymbol{x},\boldsymbol{y}) - \beta_2 C(\boldsymbol{x},\boldsymbol{y})\}. \tag{2}$$

In Equation 2, we refer to $w(\boldsymbol{x},\boldsymbol{y})$ as the ESF, and $G(\boldsymbol{x},\boldsymbol{y})$ and $C(\boldsymbol{x},\boldsymbol{y})$ are the energetic gain and energetic cost of the step, respectively (illustrated in Figure 1). In this form, β_1 quantifies the selection for energetic gains $G(\boldsymbol{x},\boldsymbol{y})$, formulated in terms of energetically beneficial resources, and β_2 is the strength of selection

against energetic costs $C(\boldsymbol{x}, \boldsymbol{y})$, formulated as habitat-specific movement costs (see Section 2.1.3 for guidance on how to define covariates). The likelihood can be optimised with respect to β_1 and β_2, over all steps, to estimate the set of parameters that maximise the likelihood of an animal selecting the observed locations relative to the rest of the available habitat. The ESF can be viewed as a joint model of habitat selection (captured as selection for energetic gain) and movement (with selection against energetic costs).

When estimated together, the ESF parameters can provide support for different energy maximization strategies in optimal foraging theory: an increase in β_1 represents stronger selection for gains (i.e. stronger intake-maximization; Figure 1, scenario 1) and an increase in β_2 represents stronger selection against energetic costs (i.e. stronger cost-minimization; Figure 1, scenario 2). Either of these outcomes represent strategies to maximize net energetic balance, and the ESF can assess whether animal movement is driven by gains, costs, or both. If $G(\boldsymbol{x}, \boldsymbol{y})$ and $C(\boldsymbol{x}, \boldsymbol{y})$ are in the same units, then β_1 and β_2 can be directly compared within the same model. However, even if the units are unknown, the ESF may still be used to compare how strategies vary between individuals, time periods (e.g. seasons), demographic groups (e.g. juveniles vs. adults), or any other ecologically-important grouping.

2.1.1 Comparison to other habitat selection models

The ESF uses the same basic modelling approaches as SSFs. Both models are used to analyse autocorrelated animal tracking data and describe habitat selection at the scale of the observed movement step. SSFs consider that movement constraints limit the habitat availability, and give the likelihood of a movement step ending at location $\boldsymbol{y}$ given that it started at location $\boldsymbol{x}$ in the study region Ω as

$$f(\boldsymbol{y}|\boldsymbol{x}) = \frac{\Phi(\boldsymbol{y}|\boldsymbol{x})w(\boldsymbol{x}, \boldsymbol{y})}{\int_{\boldsymbol{z}\in\Omega} \Phi(\boldsymbol{z}|\boldsymbol{x})w(\boldsymbol{x}, \boldsymbol{z})d\boldsymbol{z}}. \quad (3)$$

The first term, $\Phi(\boldsymbol{y}|\boldsymbol{x})$, is the resource-independent movement kernel, which describes how an animal would move in a homogeneous landscape or in the absence of resource preference (Forester et al., 2009). The second term, $w(\boldsymbol{x}, \boldsymbol{y})$, is a weighting function and represents resource selection without movement constraints. The weighting function is typically defined as: $w(\boldsymbol{x}, \boldsymbol{y}) = \exp\{\boldsymbol{\beta} \cdot \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{y})\}$, where $\boldsymbol{\beta}$ is a vector of parameters representing the strength of selection for $\boldsymbol{h}(\boldsymbol{x}, \boldsymbol{y})$, a vector of habitat covariates. Therefore, the step density of an SSF model is the product of resource selection $w(\boldsymbol{x}, \boldsymbol{y})$ and movement $\Phi(\boldsymbol{y}|\boldsymbol{x})$.

The ESF can be viewed as a special case of an SSF, where $w(\boldsymbol{x}, \boldsymbol{y})$ is based on two energetic covariates, and where $\Phi(\boldsymbol{y}|\boldsymbol{x})$ can be viewed as an "energy-independent" movement kernel, which is uniform over the whole study region (i.e. it cancels out in Equation 3). We consider $\Phi(\boldsymbol{y}|\boldsymbol{x})$ to be uniform based on the assumption

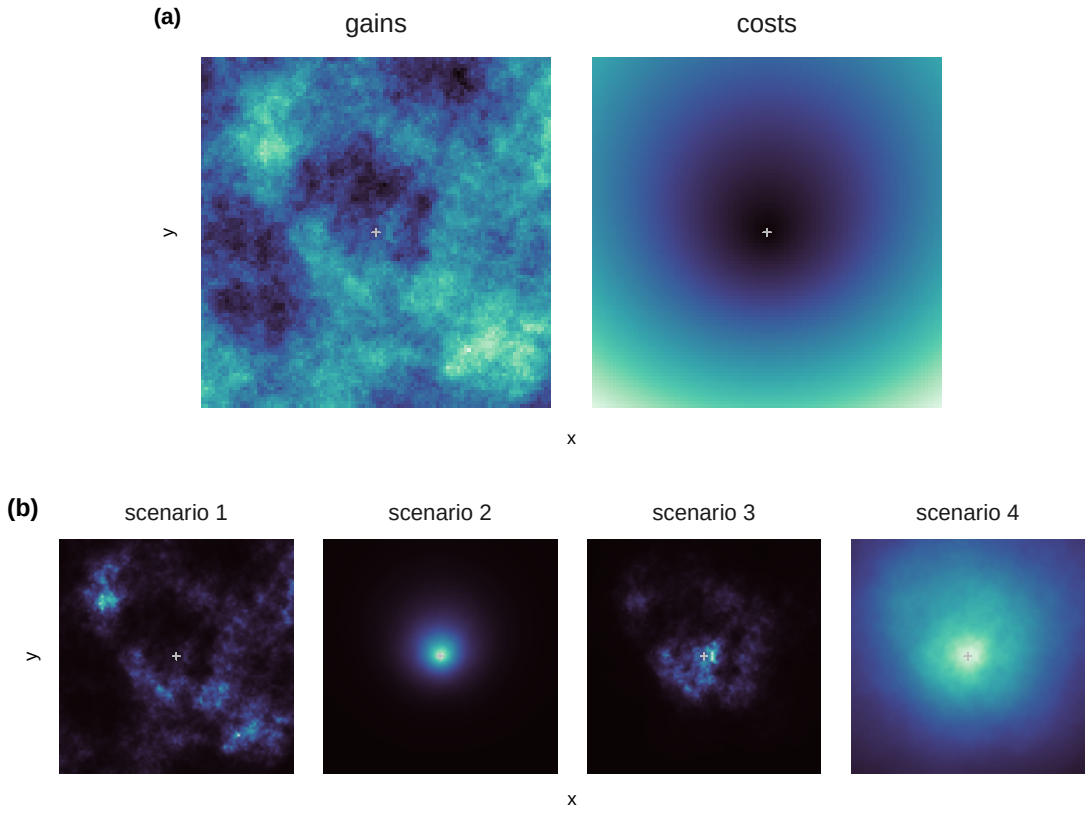


Figure 1: (a) Simulated rasters of energetic gains and energetic costs, and (b) corresponding ESFs. In (a), energetic gains were simulated as a random covariate field and energetic costs were defined as a product of the step length and turning angle from the central location (+), assuming that the animal was facing up the y-axis before turning. In (b), the ESF was calculated as $ESF = \exp(\beta_1 \times gains - \beta_2 \times costs)$ with different values of β_1 and β_2 to represent different movement scenarios: scenario 1 represents movement driven primarily by gain-maximisation (high β_1, low β_2), scenario 2 represents movement driven primarily by cost-minimisation (high β_2, low β_2), scenario 3 represents movement driven strongly by both gains and costs (high β_1, high β_2), and scenario 4 represents a low effect of both (low β_1, low β_2). In all panels, lighter colours represent higher values.

that, in the absence of energetic constraints (i.e. a hypothetical animal that can move arbitrarily fast without expending energy), animals could travel anywhere in the study region. Here, selection against costs defines a movement model, as $C(\boldsymbol{x}, \boldsymbol{y})$ can account for aspects of animal movement that affect energy expenditure (e.g. speed and tortuosity). For a simple example, in the absence of selection for gains ($\beta_1 = 0$) and if the costs are proportional to the step length squared, the ESF reduces to a Gaussian random walk with variance inversely proportional to β_2 (Appendix A). Regardless of the exact cost formulation, the ESF eliminates the need to make the assumption that availability is known, as this is captured by the selection against energetic costs. Therefore, similar to integrated step selection analysis (iSSA; Avgar et al., 2016), the ESF circumvents the difficulty of defining availability in habitat selection studies (Beyer et al., 2010), while simultaneously assessing energetic processes. For this reason, it stands in contrast with the approach of Latombe et al. (2014), where the ratio of gains over costs was included as a covariate in an SSF with a non-uniform function Φ (based on empirical distributions of step lengths and turning angles). That model measures the effects of gains and costs once movement constraints have been accounted for, rather than integrate all aspects of movement and habitat selection into energetic variables.

2.1.2 Implementation

Consider a set of locations of an animal $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$. The ESF defines the likelihood of the entire track as $L(\beta_1, \beta_2 \mid \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \prod_{i=1}^{n-1} f(\boldsymbol{x}_{i+1}|\boldsymbol{x}_i)$, where $f(\boldsymbol{x}_{i+1}|\boldsymbol{x}_i)$ is the likelihood of a single step (Equation 1). In many cases, the exact likelihood is computationally intractable, as it includes the integral of w over the study region Ω. In practice, we can approximate this likelihood using a case-control (i.e. use-availability) design (Forester et al., 2009; Thurfjell et al., 2014). Case-control designs contrast habitat characteristics at locations where the animal went (i.e. what it used) with where it could have gone (i.e. available habitat), and they are common in habitat selection analyses. For each observed location $\boldsymbol{x}_i$ (hereafter, a case), we generate a set of random locations (hereafter, controls) which represent a sample of the available habitat. Using Monte Carlo integration over the control locations $\{\boldsymbol{z}_{i1}, \boldsymbol{z}_{i2}, \ldots, \boldsymbol{z}_{iK}\}$, we calculate the approximate likelihood as

$$\tilde{L}(\beta_1, \beta_2|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \prod_{i=1}^{n-1} \frac{w(\boldsymbol{x}_i, \boldsymbol{x}_{i+1})}{\sum_{k=0}^{K} w(\boldsymbol{x}_i, \boldsymbol{z}_{ik})}, \tag{4}$$

where we define $\boldsymbol{z}_{i0} = \boldsymbol{x}_{i+1}$.

This numerical integration requires sampling control locations $\{\boldsymbol{z}_{i1}, \boldsymbol{z}_{i2}, \ldots, \boldsymbol{z}_{iK}\}$ from a uniform distribution across the whole habitat, Ω, which can be computationally intensive if Ω is large. To reduce the

computational cost, we notice that the ESF (Equation 2) will typically be negligible when $\boldsymbol{z}_{ij}$ is sufficiently far from the starting point of the step, $\boldsymbol{x}_i$ (i.e. as costs get very large). We therefore generate control locations over a sufficiently large neighbourhood of the start point (i.e. a disc of radius R) to obtain a good approximation of the likelihood (Appendix B). In most cases, it may be sufficient to choose R to be slightly larger than the animal's longest observed step, based on the assumption that the likelihood of a step longer than this is negligible (see Appendix B for evaluation of R size). This sampling is not a model of movement or availability (e.g. as in Arthur et al., 1996). Rather, we use control locations over a disc for computational convenience, and the availability can be viewed as arising from the effect of energetic costs on movement, $C(\boldsymbol{x}, \boldsymbol{y})$. Replacing Ω by a disc of radius R can be viewed as a special case of importance sampling, to increase the precision of the Monte Carlo integration in Equation 4.

Given that the ESF uses the same general formulation and case-control design as SSFs, model fitting can be done using the same statistical techniques and software. We can estimate β_1, β_2 with maximum likelihood estimation (MLE), with regards to Equation 4, using numerical optimizers (e.g. *optim* in R) or software for conditional logistic regression (e.g. the R function *clogit*, package *survival*). Therefore, implementation is fast and based on techniques that many practitioners may be familiar with. In Appendix D, we verify these implementation methods with a technical simulation, which showed accurate inferences under different levels of spatial autocorrelation and number of control locations used (Figure S5).

2.1.3 Defining the energetic covariates

The ESF is applicable to any system where there is adequate energetic data, and the covariates must be formulated specifically to each study. Gains, G, and costs, C, need to be defined in units of energy (which could be standard units such as kJ or any other convenient unit), based on ecological and physical principles. Here, we focus on the energetic gains of foraging and costs of movement, but the approach can be extended if other environmental factors (e.g. temperature, weather) are important to energy gain or expenditure. In fact, we may not always be able to or even want to include all aspects of energetics, particularly when they are not important for inferences (e.g. when they are consistent between individuals and/or are not dependent on habitat). These covariates should be formulated carefully, as the inferences from the ESF necessarily depend on the quality of the estimated energetic variables. Definitions of G and C rely on carefully evaluating complex interactions between movement and habitat, and mischaracterization may lead to parameter estimates that do not truly represent energy selection.

Energetic gains mainly arise from the consumption of energetically beneficial resources, whose distribution

can be derived from environmental data (e.g. NDVI, prey or vegetation biomass; Fortin et al., 2003; Pilfold et al., 2014). Many recent studies have presented methods to derive energy intake of herbivores at the scale of a movement step based on vegetation biomass (Latombe et al., 2014; Merkle et al., 2014; Brooke et al., 2020; Duparc et al., 2020). Metrics will contain relevant information to represent consumable and digestible biomass, and more than one resource can be incorporated into G if they are weighted based on their energetic contribution. Furthermore, resources may be combined with movement data, in cases with strong empirical or hypothesized relationships between foraging potential and movement speed (Figure S4). Under this formulation, we can evaluate selection for foraging resources under a common energetic currency, rather than preference for individual resources.

Energetic costs will typically be formulated primarily in terms of movement (tortuosity and speed; Taylor et al., 1970; Wilson et al., 2013), which can be quantified from geographical positioning system (GPS) location data. To inform the costs of movement steps, GPS metrics can be combined with captive studies (e.g. treadmills; Bidder et al., 2017) and field measures that more directly measure energy expenditure and behaviour (e.g. doubly-labelled water, heart rate, dynamic body acceleration; Pagano & Williams, 2019; Wilson et al., 2020). The cost of movement is also affected by habitat factors, such as substrate penetrability (soft vs. hard; Crête & Larivière, 2003), slope (Halsey, 2016), and resistivity (e.g. wind and water currents; Shepard et al., 2013). Therefore, habitat features and movement data can be combined into a synthetic model of energy expenditure (see Section 2.2, Figure S4) or correlated to field measures to estimate the energetic costs of control steps.

2.2 Case study

In this section, we present a case study of polar bears as an example of the ESF framework, in which we detail the process to define the energetic covariates and draw inferences. Polar bears are apex predators that forage on fat-rich prey (primarily seals) and inhabit highly dynamic sea ice environments. They have pronounced seasonal patterns of energy acquisition (i.e. hyperphagia in the spring followed by a hypophagic period; Pilfold et al., 2012), as well as high energy expenditure associated with locomotion (Hurst et al., 1982a). Additionally, movement costs can be highly affected by the local sea ice conditions, such as ice concentration and drift speed (Durner et al., 2017; Griffen, 2018). Although these spatiotemporal interactions indicate that energetics may largely influence polar bear movement and space use, these mechanisms have yet to be analysed in a framework that considers selection of gains and costs. In our case study, we applied the ESF to solitary adult ($>$ 5 years old) GPS-collared female polar bears in the Canadian Beaufort Sea.

We used tracking data at a 4-hour resolution from the spring (March - June) of 2007-2011 (Figure 2; see Appendix E.1 for details of the study area, data acquisition, and data processing). All field research was conducted under Government of Northwest Territories Department of Environment and Natural Resources permits (WL003322, WL005372, WL005596, WL007376) and animal handling procedures were approved by the University of Alberta BioSciences Animal Care and Use Committee. We estimated energetic gains from an energetically-weighted RSF of seal kills, and develop a cost model for use with GPS telemetry data. Therefore, we consider polar bear energetics at the scale of movement-based habitat selection.

2.2.1 Energetic gains G

Polar bears forage primarily on ringed seals (*Pusa hispida*) and bearded seals (*Erignathus barbatus*), particularly during a hyperphagic period from mid-April to June (Pilfold et al., 2012). Therefore, we derived energetic gains from an RSF model of forage quality that covered our study (2007 – 2011) from Pilfold et al. (2014). This RSF modelled locations of seals killed by polar bears, weighted by biomass, relative to habitat characteristics. The most important habitat covariates in the weighted RSF model were distance from land, bathymetry, distance from shorefast ice boundary, regional sea ice concentration, floe edge, and ice type (for details, see Pilfold et al., 2014). Because the RSF incorporated both seal kill biomass and abundance, we assumed the RSF value to be proportional to energetic return. We extended the temporal and spatial extent of the original rasters, but not beyond the original range of habitat characteristics and season (Figure 2), and we created daily rasters which encompassed approximately 100km off-shore along the coast of Alaska and Canada (from approximately 160°W to 115°W), including the Amundsen Gulf and regions adjacent to Banks Island. The resolution of the rasters was 6.25km and RSF values were zero in locations where sea ice was absent.

2.2.2 Energetic costs C

We formulated costs based on the movement costs of captive polar bears, combined with environmental covariates to represent field conditions (Figure 3). Telemetry locations arise from a combination of active bear movement and passive displacement caused by ice drift. Therefore, we defined a step as the active bear movement between telemetry locations, corrected for ice drift following Klappstein et al. (2020), using drift data from the National Snow and Ice Data Center (Polar Pathfinder Daily 25km EASE-Grid Sea Ice Motion Vectors; Tschudi et al., 2019). In the following, we used the GPS locations to evaluate environmental variables, whereas we used the tracks corrected for ice drift to measure movement speed. At each step, a

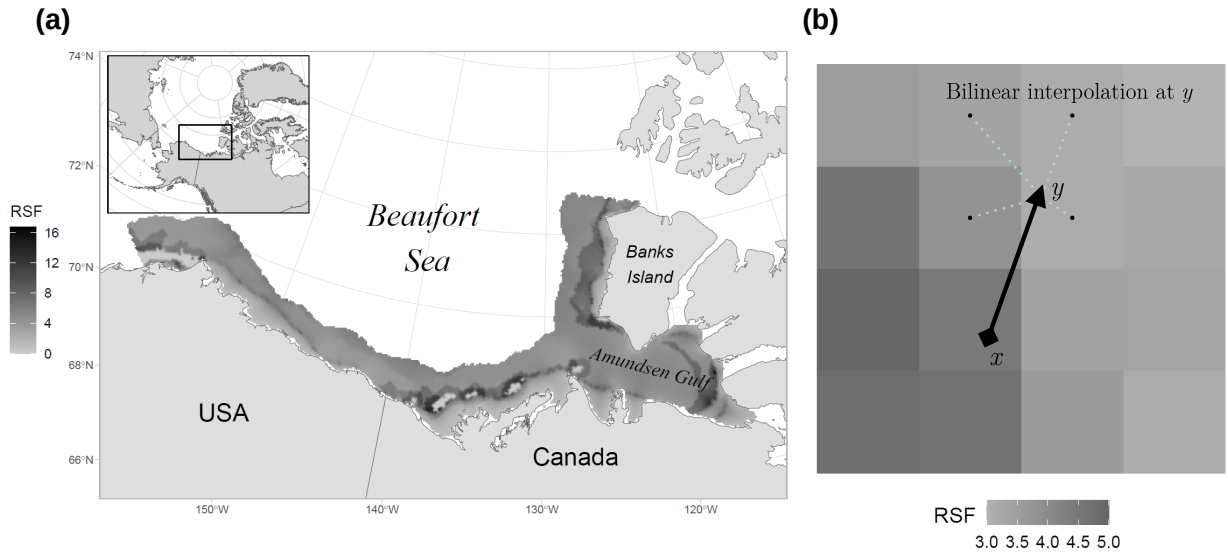


Figure 2: Illustration of energetic gains in polar bear case study. (a) Map of study area overlaid with an example of the seal biomass RSF. (b) Schematic representation of energetic gain evaluation for a step from $\boldsymbol{x}$ to $\boldsymbol{y}$, using bilinear interpolation at $\boldsymbol{y}$ based on the four adjacent cells (black dots).

bear can either be swimming or walking on sea ice, which have different energetic costs (e.g. Griffen, 2018; Pagano et al., 2018). Using spring/summer (March - August) aquatic sensor data from Lone et al. (2018), we modelled the relationship between the proportion of time in water and sea ice concentration as a generalized additive model (GAM) in the *mgcv* R package (Wood, 2017). This model does not have an intercept at 1, as expected, suggesting that time spent in water may be underestimated for low ice concentration (Figure 3b). This can be partially be explained by the tendency of satellites to underestimate ice concentration (e.g., Castro de la Guardia et al. 2017 found that bears were still on ice when satellite imagery indicated ice concentration of 0%). Although underestimation of time spent in water could result in conservative cost estimates, it is the best data available. Using this GAM, we estimated proportion of time spent in water for each polar bear step, based on its interpolated ice concentration value (AMSR-E 3.125 km grid Version 5.4; Spreen et al., 2008), and assumed the proportion of time in water to be the same as the proportion of the distance travelled. Lastly, we modelled the relationship between travel speed and energy expenditure, using combined estimates from five treadmill studies (Øritsland & Jonker, 1976; Hurst et al., 1982a,b; Watts et al., 1991; Pagano et al., 2018). These data included resting metabolic rates and accounted for the weight of the bear. We modelled energy expenditure as a function of walking speed as a GAM with a gamma response distribution and a positive monotonic constraint in the R package *scam* (Pya & Wood, 2015). When the bear

was assumed to be walking, the cost was derived directly from this curve, and when the bear was swimming, we multiplied this cost by five to represent the higher energy expenditure (Griffen, 2018). Importantly, our modelling approach estimated similar daily costs as those obtained from doubly-labelled water (Figure S9; Pagano & Williams, 2019).

As suggested in Section 2.1.3, movement tortuosity (i.e. turning angle) may also affect the costs of movement, and this effect could be included in analyses of high-resolution movement data (Wilson et al., 2013). Although small-scale tortuosity might not appear at a 4-hour resolution, directional persistence may be in place even at this coarser scale, due to large-scale targets of attraction (e.g. foraging patches, mating sites), environmental features (e.g., the floe edge), or behaviours (e.g. migration, site fidelity), on which we do not have sufficient information. In principle, if the energetic return of large-scale processes was known, it could be captured using a multi-scale ESF similar to the approach of Bastille-Rousseau et al. (2018). Instead, in Appendix F, we show how turning angle can be incorporated as a separate term in the ESF (similar to Avgar et al., 2016; Brooke et al., 2020) that encapsulates in a single energy-agnostic term all the drivers of directional persistence about which we do not have information, and compare this to the ESF presented below.

2.2.3 Fitting the ESF

We eliminated locations that were outside the spatiotemporal extent of the seal raster (Figure 2), which removed 34.1% of all locations (Figure S6). We generated 20 control locations on a disc around each observed location, with radius $R = 1.1 \times l_m$, where l_m is the maximum step length of all observed locations (see Appendix B for justification of R and Appendix D for number of controls). We calculated energetic gains and costs of each step as described above, using environmental covariate values at each end location. We omitted steps from analysis when there were ≥ 10 control locations without an energetic gain estimate (i.e. outside the raster extent) and accounted for this in the ESF likelihood. We fitted the ESF with the *clogit* function (from the R package *survival*) separately for each individual bear. Following Forester et al. (2009), we calculated robust standard errors to account for residual autocorrelation, which can lead to the underestimation of variance (details in Appendix G). We used Monte Carlo cross-validation to investigate the predictive performance of the model (as described by Fortin et al., 2009, full details in Appendix H). The general idea of cross-validation is to repeatedly fit the model to a random subset of the data (training data set), and check how frequently it correctly distinguishes between the observed steps and control steps in the remaining data (validation data set). The output of each iteration of this procedure is a Spearman rank

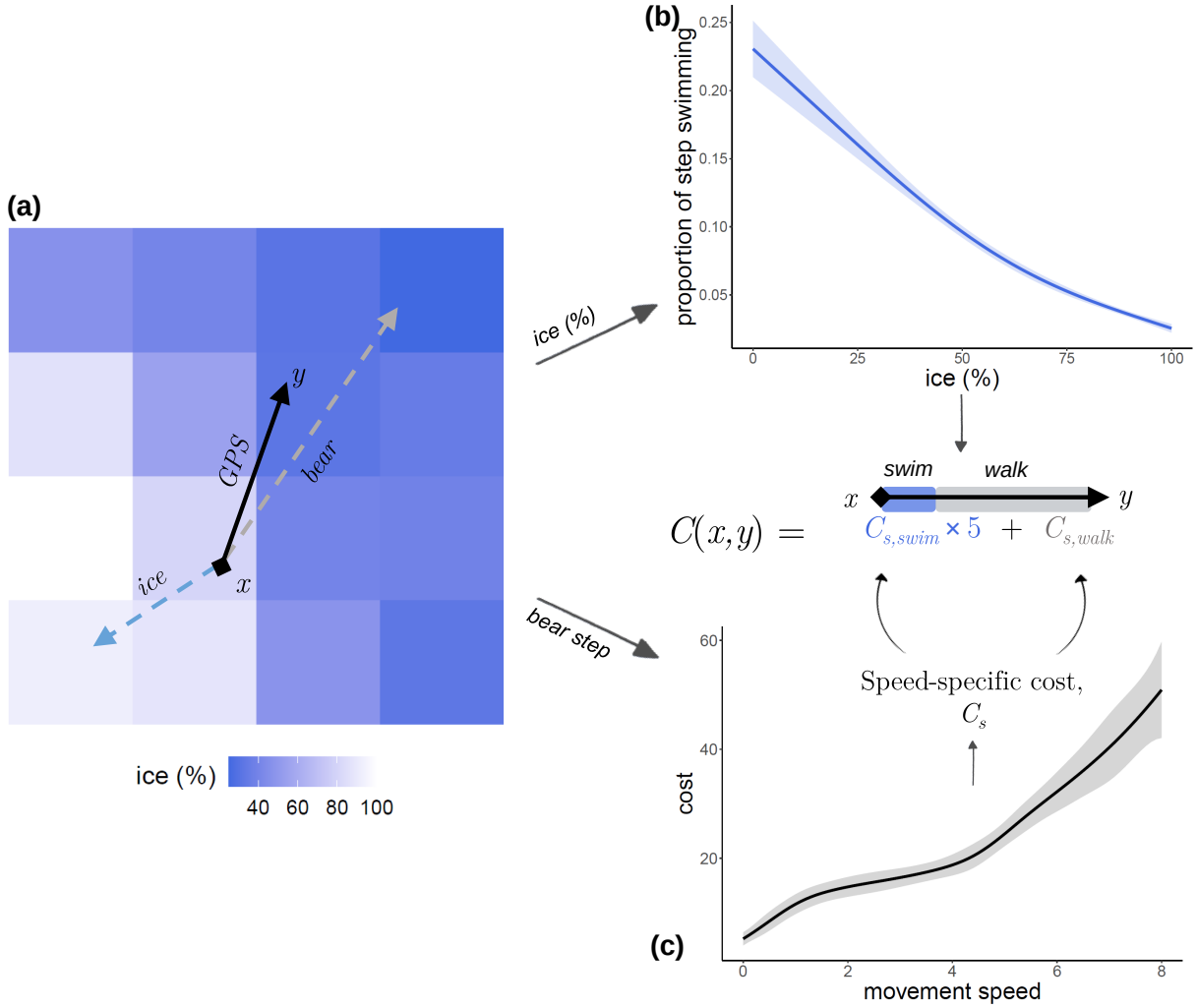


Figure 3: Schematic illustration of energetic cost evaluation in the polar bear case study, for an observed step starting at $\boldsymbol{x}$ and ending at $\boldsymbol{y}$. (a) Arrows represent the observed movement step (black; GPS), ice drift (blue dashed; ice), and actual bear movement (grey dashed; calculated as GPS - ice). Ice concentration (%) is estimated at $\boldsymbol{y}$ with bilinear interpolation (illustrated in Figure 2b). (b) Modelled relationship between ice concentration (ice %) and the proportion of the step spent swimming based on Lone et al. (2018). We use the estimated ice % from (a) to estimate the proportion of the bear step spent swimming and walking. (c) Modelled relationship between polar bear movement speed (km/h) and energetic cost (kJ/kg/h) from treadmill studies. Using the bear speed and weight, we calculate C_s for the step. $C_{s,swim}$ and $C_{s,walk}$ are C_s multiplied by the proportion of time in each behaviour. The total energy expenditure of the step $C(\boldsymbol{x}, \boldsymbol{y})$ is the sum of $C_{s,swim}$ multiplied by 5 (to represent the higher costs of swimming) and $C_{s,walk}$.

correlation coefficient, $-1 \leq r_s \leq 1$, which measures how consistently the case locations of the validation data set are ranked higher than the corresponding control locations. If the model has high predictive power, then the rank of the case locations is high and this is reflected in a high Spearman rank correlation coefficient, i.e., close to 1 . For comparison, we also repeated this procedure on the control locations only (i.e., without the case) to assess how often a random location would be chosen by chance alone.

2.2.4 Comparison to an SSF

To illustrate how an SSF based on environmental features differs from the ESF, we fitted an SSF model (given by Equation 3), with non-energetic covariates and a resource-independent movement kernel $\Phi(\boldsymbol{y}|\boldsymbol{x})$ derived from the polar bear tracking data. We included environmental covariates that contribute to G and C, but in a non-energetic form that are more typical in polar bear habitat selection analyses (e.g., McCall et al., 2016; Johnson & Derocher, 2020). Ice concentration is an important covariate for two reasons: i) it is a central component of C (i.e. it determines the proportion of time swimming; Lone et al., 2018), and ii) it is an important predictor in the seal RSF of G (i.e. there is a quadratic effect of ice concentration, which predicts highest biomass at approximately 85%; Pilfold et al., 2014). Therefore, we included both a linear and quadratic effect of ice concentration (*ice* and ice^2). We also included ocean depth (*depth*), interpolated from the International Bathymetric Chart of the Arctic Ocean (Version 4; Jakobsson et al., 2020), which contributes to energetic gains (i.e. greater seal kill biomass at greater depths; Pilfold et al., 2014). Lastly, we included sea ice drift speed (*drift*) as a covariate, as it affects the costs of moving any given geographic distance (Durner et al., 2017; Klappstein et al., 2020). We fitted the SSF using the same implementation techniques as the ESF, but generated controls based on the observed movement of the polar bears: we fitted a gamma distribution to step lengths and a wrapped Cauchy distribution to turning angles, and used these to randomly sample 20 control locations for each observed (case) location (as in Forester et al., 2009). We fitted and tested the predictive power of an SSF with cross-validation for each individual, following the same procedures described in Section 2.2.3.

3 Results

3.1 ESF results

We analysed 7,526 GPS steps from 23 GPS-collared adult female polar bears (steps per individual: $78-946$). Among case steps, the median energetic gain was 4.4 (range: $0-20.3$; arbitrary units, assumed to be

proportional to energy), while controls had a median energetic gain of 4.3 (range: $0 - 28.1$). The median energetic cost of case steps was 8.1 MJ (range: $3.3 - 44.1$ MJ) and control steps had a median cost of 19.7 MJ (range: $3.6 - 124$ MJ). The median β_1 estimate was 0.10 (range $-0.32, 0.96$), with three estimates with 95% CIs that did not overlap zero (Figure 4). Of these, two bears appeared to select for energetic gains ($\beta_1 \pm \text{SE} = 0.42 \pm 0.14$; 0.40 ± 0.17) and one bear appeared to select against energetic gains ($\beta_1 \pm \text{SE} = -0.32 \pm 0.16$). Conversely, all β_2 estimates showed a selection against costs, with a median of 0.57 (range $0.27, 0.96$), and no CIs overlapped zero. For every individual, cross-validation indicated that the ESF better predicted polar bear movement than what would be expected by chance (i.e. mean individual r_s was always larger for the observed than random). Across all individuals, mean r_s (range) was 0.56 (0.17, 0.88) for observed and -0.009 (-0.69, 0.64) for random. Further, we performed cross-validation on both a gain-only and a cost-only model, which showed generally low predictive power of energetic gains and higher predictive power of energetic costs (Appendix H.1).

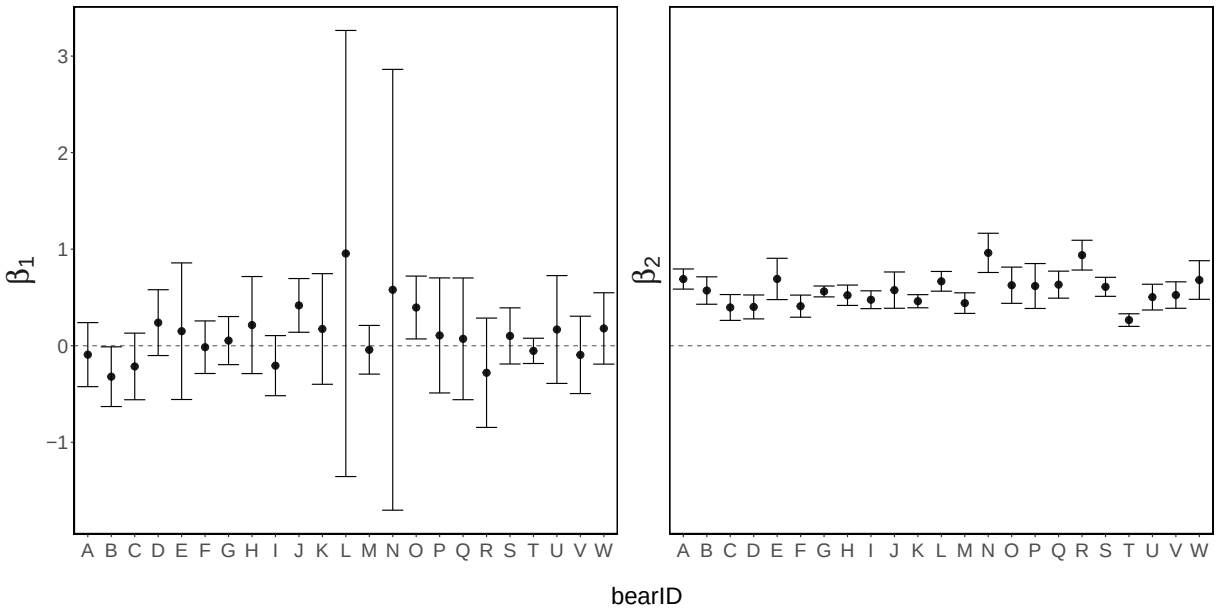


Figure 4: Estimated β_1 (selection for gains) and β_2 (selection against costs) coefficients of lone adult female polar bears ($n = 23$). Error bars are robust 95% CIs.

3.2 SSF results

Of the 23 individual bears, 8 had a statistically clear linear effect of ice concentration (i.e., robust 95% CIs for β_{ice} did not overlap 0), 7 of which also showed evidence of a quadratic effect (i.e., ice^2). For these 7

bears, estimates of β_{ice} were all positive (median: 8.05, range: 4.80 to 15.0), and estimates of β_{ice^2} were all negative (median: -5.59, range: -8.83 to -1.39). These results indicate that some bears select for an optimal range of ice concentration, which we estimated to range from 61% to 100%. Ocean depth was only found to be an important predictor for 2 bears, 1 with a clear positive effect (i.e. selection for shallower regions; $\beta_{depth} \pm SE = 3.08 \pm 0.99$) and 1 with a clear negative effect ($\beta_{depth} = -20.3 \pm 9.70$). Ice drift only had a clear negative effect for 1 bear ($\beta_{drift} = -3.95 \pm 1.62$), and all other estimates had robust CIs that overlapped 0. Here, we only presented the ranges of estimates (across individuals) with statistically clear effects, but full results including uncertainty are in Figure S11 of the supplementary material. Cross-validation scores indicated that the SSF did not better predict polar bear movement better than would be explained by random chance alone: individual mean Spearman rank correlation coefficients (r_s) of case locations had a mean of 0.02 (range: -0.75 to 0.74), whereas the controls (proxy for random) had a mean of -0.07 (range: -0.29 to 0.64). Note, these cross-validation results cannot directly be compared to those of the ESF, due to differences in null models (i.e. distribution of control locations), but see Appendix I for an AIC-based model comparison.

4 Discussion

Our proposed method to estimate the energy preferences of animals is similar to recent approaches to combine movement and habitat (Avgar et al., 2016; Michelot et al., 2019), but the ESF uniquely integrates both factors into energetic covariates. Therefore, the ESF adopts a purely energetic perspective, moving away from habitat selection analyses that focus on environmental attributes. In this discussion, we start with an interpretation and contextualisation the results of the polar bear case study, followed by a more general discussion of the ESF method.

4.1 Discussion of the polar bear case study

In the case study, we found a strong pattern of selection against energetic costs in all individuals, but only three showed evidence of selection for ($n = 2$) or against ($n = 1$) energetic gains. There was large uncertainty in β_1 for most bears. One possible factor is that, if the values of gains for case and control locations did not cover a wide enough range, it would provide little information about the effect. This can happen when spatial autocorrelation in the covariate is high, or if the raster grid cells are large compared to observed steps, such that most steps do not overlap multiple cells (Boyce et al., 2003; Boyce, 2006; Northrup et al., 2013).

As both gains and costs incorporate ice concentration, we assessed the collinearity between the energetic covariates. Although we only found a low to moderate correlation (Pearson correlation = −0.3), this might have inflated the uncertainty of the parameter estimates. Further, the energetic covariates were imperfect representations of intake and expenditure; we defined gains from an RSF model, which is only an estimation of the true prey availability (Pilfold et al., 2014), and cost estimates were also model-based. The error in these covariates could lead to either an underestimation of effects (e.g. if bears do select for seals, but the gains do not accurately represent this), or overestimation of effects, as we did not propagate this covariate error when estimating uncertainty. In studies where the energetic gains can be more directly measured (e.g. ungulates and vegetation biomass; Brooke et al., 2020), we expect that the corresponding selection parameter may be estimated more precisely, leading to more detailed insights into movement strategies of animals.

Our results suggest that movement of the solitary female polar bears in our study was driven by reducing energy expenditure (i.e. cost-minimisation), but we found little evidence of an effect of energetic gains (i.e. gain-maximisation). Cost-minimisation could arise due to high predictability of energetic costs, based on internal factors and mechanical movement constraints, while intake-maximisation would require polar bears to have knowledge of the unpredictable seal distribution (Ramsay & Stirling, 1986). At the landscape level, polar bear distribution correlates positively to seal kill biomass (Pilfold et al., 2014), but we did not observe this at the scale of the movement step, which may be reflective of local variability in the probability of killing a seal. This is similar to the general observation of Fryxell (1997) that "sloppy foragers" (i.e., animals whose movement is primarily determined by cost-minimisation rather than gain-maximisation) tend to do well in situations where the distribution of energetic gains is unpredictable. Although all estimates of β_2 were positive, there was some inter-individual variability: for an increase in costs of 1MJ (with fixed gains), the likelihood of taking a step decreased by 24% (i.e., lowest β_2) to 62% (i.e., highest β_2). This variability in selection against costs reveals differences in movement behaviour, suggesting that some bears are more averse to the energetic risk of costly steps. The absence of a statistically clear effect of energetic gains could also be due to temporal variation in foraging behaviour. Polar bears do not enter the main foraging period until mid-April (Pilfold et al., 2012), solitary females forage less when sequestered into less favourable habitats by courting males (for up to 18 days at a time; Stirling et al., 2016), and larger-scale drivers such as site fidelity (Mauritzen et al., 2001) could also dominate movement patterns. However, we did not directly test these hypotheses, which could be investigated in future research.

The energetic inferences are fundamentally different from those of the SSF, the latter of which measures selection for environmental features separated from their energetic interpretation. Of the covariates included

in the SSF, ice concentration had the most common effect and indicated that bears select for an optimal level of ice cover. However, because ice concentration is detached from energetics, it is unclear to what extent this effect is explained by avoidance of open water (i.e., costs) or selection of seal-rich areas (i.e., gains). Sea ice movement poses an additional challenge for applying a standard SSF, as it is unclear whether the movement kernel Φ should be based on the observed GPS steps or the drift-corrected steps. Both definitions of Φ are flawed: the observed steps (our chosen approach) are poor descriptors of the bear's movement because they include ice drift, and while the drift-corrected steps account for this problem, they cannot be tied to spatial locations where the bear was. This standard SSF formulation separates drift from its contribution to energetic costs, which depends on both speed and directionality, and we therefore lose information about how bears move relative to ice.

4.2 General discussion and future directions

In our case study, we demonstrated how the ESF can be used in a system where movement and habitat are related to energetics through complex interactions, and derived energetic costs based partially on the effects of a moving environmental substrate. Animal movement in moving habitats is an under-explored area of habitat selection modelling, for which the ESF provides an integrated solution. Without this integrated approach, energetic contributions of movement and habitat are typically modelled separately. Therefore, the ESF provides a parameterisation, with convenient implementation procedures, to obtain energy-based inferences that are conceptually different from the environment-based inferences of SSFs. Since the ESF is a movement-based model, its inferences differ from previous work linking energetics to RSFs that cannot directly incorporate the costs of locomotion (e.g. Long et al., 2014). Eisaguirre et al. (2020) used a two-stage energetic approach to incorporate costs into the SSF availability kernel to estimate the selection for resources once such costs had been accounted for. Therefore, Eisaguirre et al. (2020) can account for bias in habitat selection parameters when animals make energetic trade-offs; for example, in a case where the acquisition of an energy-rich resource requires moving through a highly resistive habitat, ignoring the increased energetic costs may lead to underestimating the resource selection parameter. The ESF can be used to achieve a similar goal, but offers a one-stage integrated solution where gains and costs are combined directly into the selection function. Therefore, in contrast to Eisaguirre et al. (2020), the ESF treats energetic mechanisms as the goal of inference. The approach of Hooten et al. (2019) has a similar motivation of describing the effects of physiological processes on animal movement. In their framework, the effects of environmental features on energetic balance are estimated during model fitting, whereas we assume that they are known from data.

Their approach therefore allows for more detailed inferences into underlying physiological mechanisms, and may be more appropriate when the gains and costs cannot be evaluated *a priori*. The ESF offers an alternative to understand the contribution of energy to observed space use patterns in the standard framework of SSFs.

Inherently, step selection analyses are sensitive to the spatial and temporal scale of the telemetry and covariate data (Munden et al., 2021). In the ESF, the spatial scale of the energetic covariates needs to be fine enough that it is possible to observe preference at the scale of the observed movement steps. When covariate data is coarse or spatially autocorrelated, contrast between case and control locations may be low and provide little evidence of selection (Northrup et al., 2013). Another feature of step selection models is that the temporal resolution of the tracking data determines the scale at which the animal's behaviour is examined (Bastille-Rousseau et al., 2018). For example, selection for foraging resources at the scale of the landscape (e.g. this case study) or energy accumulation over a long period (e.g. in the case of kill sites) may not be apparent at the scale of finer resolution movement steps. The interpretation of the ESF parameters is therefore tied to the time interval of observation, and lack of evidence for selection at the movement step may not translate to a biologically relevant time-scale. Thoughtful selection of intervals between telemetry locations, continuous-time analogues of SSFs, and high-frequency data may prove useful to overcome these scale dependencies. We hope this study will motivate the collection of more precise data, suited to understand the energetic mechanisms behind animal space use (see Williams et al., 2020, for a review of available technologies).

The ESF has great potential for methodological extensions, due to its close theoretical and practical links to existing methods (SSFs and iSSA). For example, we could consider a state-switching ESF model, where an unobserved behavioural state determines the selection parameters for energetic gains and costs. A state-switching ESF model would allow us to assess behaviour-specific energy selection, and could be written as a hidden Markov model, similar to the state-switching SSF model of Nicosia et al. (2017). It may also be possible to incorporate temporal dynamics and energetic state (e.g. satiation or cumulative energy balance; Latombe et al., 2014) to better reflect the physiological demands and constraints of gaining energy (i.e. recharge dynamics; Hooten et al., 2019). Further, it should be straightforward to include non-energetic terms (e.g., associated with predator risk) as interactions with gains and costs to assess how animals make energetic trade-offs with other aspects of survival. Here, we presented a re-imagining of SSFs, in which we use energy to summarise both habitat selection and movement. This approach is widely applicable, flexible enough to include a range of extensions, and can ultimately be used to build on foraging and movement theories from an energetic perspective.

Acknowledgements

We acknowledge funding from Mitacs Canada, the Canadian Association of Zoos and Aquariums, Canadian Wildlife Federation, Environment and Climate Change Canada, Hauser Bears, Natural Sciences and Engineering Research Council of Canada, Polar Bears International, Polar Continental Shelf Project, Quark Expeditions, United States Department of the Interior (Bureau of Ocean Energy Management), and World Wildlife Fund Canada. We also thank those who assisted with data collection in the field: Marie Auger-Méthé, Oliver Barker, Seth Cherry, Stephen Hamilton, Alysa McCall, Jodie Pongracz, Vicki Sahanatien, and Mike Woodcock.

Conflicts of interests

We declare no conflict of interest.

Author contributions

JR Potts conceived the original idea for the model, with input from L Börger. NJ Klappstein and AE Derocher conceived and designed the polar bear application and energetics modelling. NJ Klappstein conducted the statistical analysis with assistance from T Michelot and input from all co-authors. NW Pilfold produced the seal RSF rasters and AE Derocher provided the polar bear telemetry data. NJ Klappstein wrote the manuscript with contributions from all co-authors.

Data accessibility

If the manuscript is accepted for publication, we will publish our data and accompanying code to Zenodo, where it will be archived with a permanent doi.

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., Potts, J.R., Lewis, M.A. & Boyce, M.S. (2016). Integrated step selection analysis: bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution*, *7*, 619–630. https://doi.org/10.1111/2041-210X.12528

Bastille-Rousseau, G., Murray, D.L., Schaefer, J.A., Lewis, M.A., Mahoney, S.P. & Potts, J.R. (2018). Spatial scales of habitat selection decisions: implications for telemetry-based movement modelling. *Ecography*, *41*, 437–443. https://doi.org/10.1111/ecog.02655

Bastille-Rousseau, G., Wall, J., Douglas-Hamilton, I., Lesowapir, B., Loloju, B., Mwangi, N. & Wittemyer, G. (2020). Landscape-scale habitat response of African elephants shows strong selection for foraging opportunities in a human dominated ecosystem. *Ecography*, *43*, 149–160. https://doi.org/10.1111/ecog.04240

Beyer, H.L., Haydon, D.T., Morales, J.M., Frair, J.L., Hebblewhite, M., Mitchell, M. & Matthiopoulos, J. (2010). The interpretation of habitat preference metrics under use-availability designs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*, 2245–2254. https://doi.org/10.1098/rstb.2010.0083

Bidder, O.R., Goulding, C., Toledo, A., Van Walsum, T.A., Siebert, U. & Halsey, L.G. (2017). Does the treadmill support valid energetics estimates of field locomotion? *Integrative and Comparative Biology*, *57*, 301–319. https://doi.org/10.1093/icb/icx038

Boyce, M.S. (2006). Scale for resource selection functions. *Diversity and Distributions*, *12*, 269–276. https://doi.org/10.1111/j.1366-9516.2006.00243.x

Boyce, M.S., Mao, J.S., Merrill, E.H., Fortin, D., Turner, M.G., Fryxell, J. & Turchin, P. (2003). Scale and heterogeneity in habitat selection by elk in Yellowstone National Park. *Ecoscience*, *10*, 421–431. https://doi.org/10.1080/11956860.2003.11682790

Brooke, C.F., Fortin, D., Kraaij, T., Fritz, H., Kalule-Sabiti, M.J. & Venter, J.A. (2020). Poaching impedes the selection of optimal post-fire forage in three large grazing herbivores. *Biological Conservation*, *241*, 108393. https://doi.org/10.1016/j.biocon.2019.108393

Castro de la Guardia, L., Myers, P.G., Derocher, A.E., Lunn, N.J. & Terwisscha van Scheltinga, A.D. (2017). Sea ice cycle in western Hudson Bay, Canada, from a polar bear perspective. *Marine Ecology Progress Series*, *564*, 225–233. https://doi.org/10.3354/meps11964

Cornioley, T., Börger, L., Ozgul, A. & Weimerskirch, H. (2016). Impact of changing wind conditions on foraging and incubation success in male and female wandering albatrosses. *Journal of Animal Ecology*, *85*, 1318–1327. https://doi.org/10.1111/1365-2656.12552

Crête, M. & Larivière, S. (2003). Estimating the costs of locomotion in snow for coyotes. *Canadian Journal of Zoology*, *81*, 1808–1814. https://doi.org/10.1139/Z03-182

Duparc, A., Garel, M., Marchand, P., Dubray, D., Maillard, D. & Loison, A. (2020). Through the taste buds of a large herbivore: foodscape modeling contributes to an understanding of forage selection processes. *Oikos*, *129*, 170–183. https://doi.org/10.1111/oik.06386

Durner, G.M., Douglas, D.C., Albeke, S.E., Whiteman, J.P., Amstrup, S.C., Richardson, E.S., Wilson, R.R. & Ben-David, M. (2017). Increased Arctic sea ice drift alters adult female polar bear movements and energetics. *Global Change Biology*, *23*, 3460–3473. https://doi.org/10.1111/gcb.13746

Eisaguirre, J.M., Booms, T.L., Barger, C.P., Lewis, S.B. & Breed, G.A. (2020). Novel step selection analyses on energy landscapes reveal how linear features alter migrations of soaring birds. *Journal of Animal Ecology*, *89*, 2567–2583. https://doi.org/10.1111/1365-2656.13335

Forester, J., Kyung Im, H. & Rathouz, P. (2009). Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, *90*, 3554–3565. https://doi.org/10.1890/08-0874.1

Fortin, D., Fortin, M.E., Beyer, H.L., Duchesne, T., Courant, S. & Dancose, K. (2009). Group-size-mediated habitat selection and group fusion – fission dynamics of bison under predation risk. *Ecology*, *90*, 2480–2490.

Fortin, D., Fryxell, J.M., O'brodovich, L. & Frandsen, D. (2003). Foraging ecology of bison at the landscape and plant community levels: the applicability of energy maximization principles. *Oecologia*, *134*, 219–227. https://doi.org/10.1007/s00442-002-1112-4

Fryxell, J.M. (1997). Evolutionary dynamics of habitat use. *Evolutionary Ecology*, *11*, 687–701. https://doi.org/10.1023/A:1018434302138

Griffen, B.D. (2018). Modeling the metabolic costs of swimming in polar bears (*Ursus maritimus*). *Polar Biology*, *41*, 491–503. https://doi.org/10.1007/s00300-017-2209-x

Halsey, L.G. (2016). Terrestrial movement energetics: current knowledge and its application to the optimising animal. *Journal of Experimental Biology*, *219*, 1424–1431. https://doi.org/10.1242/jeb.133256

Hooten, M.B., Scharf, H.R. & Morales, J.M. (2019). Running on empty: recharge dynamics from animal movement data. *Ecology Letters*, *22*, 377–389. https://doi.org/10.1111/ele.13198

Hurst, R.J., Leonard, M.L., Watts, P.D., Beckerton, P. & Øritsland, N.a. (1982a). Polar bear locomotion: body temperature and energetic cost. *Canadian Journal of Zoology*, *60*, 40–44. https://doi.org/10.1139/z82-005

Hurst, R.J., Øritsland, N.A. & Watts, P. (1982b). Body mass, temperature and cost of walking in polar bears. *Acta Physiologica Scandinavica*, *115*, 391–395. https://doi.org/10.1111/j.1748-1716.1982.tb07096.x

Jakobsson, M., Jakobsson, M., Mayer, L.A., Bringensparr, C., Castro, C.F., Mohammad, R., Johnson, P., Ketter, T., Accettella, D., Amblas, D., An, L., Arndt, J.E., Canals, M., Casamor, J.L., Chauche, N., Coakley, B., Danielson, S., Demarte, M., Dickson, M.L., Dorschel, B., Dowdeswell, J.A. & Dreutter, S. (2020). The International Bathymetric Chart of the Arctic Ocean Version. *Scientific Data*, p. 176. https://doi.org/10.1038/s41597-020-0520-9

Johnson, A.C. & Derocher, A.E. (2020). Variation in habitat use of Beaufort Sea polar bears. *Polar Biology*, *43*, 1247–1260. https://doi.org/10.1007/s00300-020-02705-3

Kays, R., Crofoot, M.C., Jetz, W. & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*, aaa2478. https://doi.org/10.1126/science.aaa2478

Klappstein, N.J., Togunov, R.R., Reimer, J.R., Lunn, N.J. & Derocher, A.E. (2020). Patterns of sea ice drift and polar bear (*Ursus maritimus*) movement in Hudson Bay. *Marine Ecology Progress Series*, *641*, 227–240. https://doi.org/10.3354/meps13293

Latombe, G., Parrott, L., Basille, M. & Fortin, D. (2014). Uniting statistical and individual-based approaches for animal movement modelling. *PLOS ONE*, *9*, e99938. https://doi.org/10.1371/journal.pone.0099938

Lone, K., Kovacs, K.M., Lydersen, C., Fedak, M., Andersen, M., Lovell, P. & Aars, J. (2018). Aquatic behaviour of polar bears (*Ursus maritimus*) in an increasingly ice-free Arctic. *Scientific Reports*, *8*, 9677. https://doi.org/10.1038/s41598-018-27947-4

Long, R.A., Terry Bowyer, R., Porter, W.P., Mathewson, P., Monteith, K.L. & Kie, J.G. (2014). Behavior and nutritional condition buffer a large-bodied endotherm against direct and indirect effects of climate. *Ecological Monographs*, *84*, 513–532. https://doi.org/10.1890/13-1273.1

Mauritzen, M., Derocher, A.E. & Wiig, Ø. (2001). Space-use strategies of female polar bears in a dynamic sea ice habitat. *Canadian Journal of Zoology*, *79*, 1704–1713. https://doi.org/10.1139/z01-126

McCall, A.G., Pilfold, N.W., Derocher, A.E. & Lunn, N.J. (2016). Seasonal habitat selection by adult female polar bears in western Hudson Bay. *Population Ecology*, *58*, 407–419. https://doi.org/10.1007/s10144-016-0549-y

Merkle, J.A., Fortin, D. & Morales, J.M. (2014). A memory-based foraging tactic reveals an adaptive mechanism for restricted space use. *Ecology Letters*, *17*, 924–931. https://doi.org/10.1111/ele.12294

Michelot, T., Blackwell, P.G. & Matthiopoulos, J. (2019). Linking resource selection and step selection models for habitat preferences in animals. *Ecology*, *100*, e02452. https://doi.org/10.1002/ecy.2452

Munden, R., Börger, L., Wilson, R.P., Redcliffe, J., Brown, R., Garel, M. & Potts, J.R. (2021). Why did the animal turn? Time-varying step selection analysis for inference between observed turning points in high frequency data. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210x.13574

Nicosia, A., Duchesne, T., Rivest, L.P. & Fortin, D. (2017). A multi-state conditional logistic regression model for the analysis of animal movement. *The Annals of Applied Statistics*, *11*, 1537–1560. https://doi.org/10.1214/17-AOAS1045

Northrup, J.M., Hooten, M.B., Anderson, C.R.J. & Wittemyer, G. (2013). Practical guidance on characterizing availability in resource selection functions under a use-availability design. *Ecology*, *94*, 1456–1463. https://doi.org/10.1890/12-1688.1

Øritsland, N. & Jonker, W. (1976). A respiration chamber for exercising polar bears. *Norwegian Journal of Zoology*, *24*, 65–67.

Owen-Smith, N., Fryxell, J.M. & Merrill, E.H. (2010). Foraging theory upscaled: the behavioural ecology of herbivore movement. *Philosophical Transactions of the Royal Society Series B: Biological Sciences*, *365*, 2267–2278. https://doi.org/10.1098/rstb.2010.0095

Pagano, A.M., Atwood, T.C., Durner, G.M. & Williams, T.M. (2020). The seasonal energetic landscape of an apex marine carnivore, the polar bear. *Ecology*, *101*, e02959. https://doi.org/10.1002/ecy.2959

Pagano, A.M., Durner, G.M., Rode, K.D., Atwood, T.C., Atkinson, S.N., Peacock, E., Costa, D.P., Owen, M.A. & Williams, T.M. (2018). High-energy, high-fat lifestyle challenges an Arctic apex predator, the polar bear. *Science*, *359*, 568–572. https://doi.org/10.1126/science.aan8677

Pagano, A.M. & Williams, T.M. (2019). Estimating the energy expenditure of free-ranging polar bears using tri-axial accelerometers: A validation with doubly labeled water. *Ecology and Evolution*, *9*, 4210–4219. https://doi.org/10.1002/ece3.5053

Pilfold, N.W., Derocher, A.E. & Richardson, E.S. (2014). Influence of intraspecific competition on the distribution of a wide-ranging, non-territorial carnivore. *Global Ecology and Biogeography*, *23*, 425–435. https://doi.org/10.1111/geb.12112

Pilfold, N.W., Derocher, A.E., Stirling, I., Richardson, E.S. & Andriashek, D.S. (2012). Age and sex composition of seals killed by polar bears in the Eastern Beaufort Sea. *PLoS ONE*, *7*, e41429. https://doi.org/10.1371/journal.pone.0041429

Potts, J.R., Bastille-Rousseau, G., Murray, D.L., Schaefer, J.A. & Lewis, M.A. (2014). 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. https://doi.org/10.1111/2041-210X.12150

Pya, N. & Wood, S.N. (2015). Shape constrained additive models. *Statistics and Computing*, *25*, 543–559. https://doi.org/10.1007/s11222-013-9448-7

Pyke, G.H. (2019). Animal movements - an optimal foraging theory approach. *Encyclopedia of Animal Behavior*, volume 2, pp. 149–156. Elsevier, second edition. https://doi.org/10.1016/B978-0-12-809633-8.90160-2

Pyke, G.H., Pulliam, H.R. & Charnov, E.L. (1977). Optimal foraging: A selective review of theory and tests. *The Quarterly Review of Biology*, *52*, 137–154.

Ramsay, M.A. & Stirling, I. (1986). On the mating system of polar bears. *Canadian Journal of Zoology*, *64*, 2142–2151. https://doi.org/10.1139/z86-329

Schoener, T.W. (1971). Theory of feeding strategies. *Annual Review of Ecology and Systematics*, *2*, 369–404. https://doi.org/10.1146/annurev.es.02.110171.002101

Shepard, E.L.C., Wilson, R.P., Quintana, F., Gó Mez Laich, A. & Forman, D.W. (2009). Pushed for time or saving on fuel: fine-scale energy budgets shed light on currencies in a diving bird. *Proceedings of the Royal Society B*, *276*, 3149–3155. https://doi.org/10.1098/rspb.2009.0683

Shepard, E.L.C., Wilson, R.P., Rees, W.G., Grundy, E., Lambertucci, S.A. & Vosper, S.B. (2013). Energy landscapes shape animal movement ecology. *The American Naturalist*, *182*, 298–312. https://doi.org/10.1086/671257

Spreen, G., Kaleschke, L. & Heygster, G. (2008). Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal of Geophysical Research: Oceans*, *113*, C02S03. https://doi.org/10.1029/2005JC003384

Stirling, I., Spencer, C. & Andriashek, D. (2016). Behavior and activity budgets of wild breeding polar bears (*Ursus maritimus*). *Marine Mammal Science*, *32*, 13–37. https://doi.org/10.1111/mms.12291

Taylor, C.R., Schmidt-Nielsen, K. & Raab, J.L. (1970). Scaling of energetic cost of running to body size in mammals. *The American Journal of Physiology*, *219*, 1104–1107. https://doi.org/10.1152/ajplegacy.1970.219.4.1104

Thurfjell, H., Ciuti, S. & Boyce, M.S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, *4*, 1–12. https://doi.org/10.1186/2051-3933-2-4

Tschudi, M., Meier, W.N., Stewart, J.S., Fowler, C. & Maslanik, J. (2019). Polar Pathfinder Weekly 25 km EASE-Grid Sea Ice Motion Vectors, Version 4, Northern Hemisphere.

Watts, P., Ferguson, K.L. & Draper, B. (1991). Energetic output of subadult polar bears (*Ursus maritimus*): Resting, disturbance and locomotion. *Comparative Biochemistry and Physiology*, *98A*, 191–193. https://doi.org/10.1016/0300-9629(91)90518-H

Williams, T.M., Peter-Heide Jørgensen, M., Pagano, A.M. & Bryce, C.M. (2020). Hunters versus hunted: New perspectives on the energetic costs of survival at the top of the food chain. *Functional Ecology*, *34*, 2015–2029. https://doi.org/10.1111/1365-2435.13649

Wilson, R.P., Griffiths, I.W., Legg, P.A., Friswell, M.I., Bidder, O.R., Halsey, L.G., Lambertucci, S.A. & Shepard, E.L.C. (2013). Turn costs change the value of animal search paths. *Ecology Letters*, *16*, 1145–1150. https://doi.org/10.1111/ele.12149

Wilson, R.P., Börger, L., Holton, M.D., Scantlebury, D.M., Gómez-Laich, A., Quintana, F., Rosell, F., Graf, P.M., Williams, H., Gunner, R., Hopkins, L., Marks, N., Geraldi, N.R., Duarte, C.M., Scott, R., Strano, M.S., Robotka, H., Eizaguirre, C., Fahlman, A. & Shepard, E.L. (2020). Estimates for energy expenditure in free-living animals using acceleration proxies: A reappraisal. *Journal of Animal Ecology*, *89*, 161–172. https://doi.org/10.1111/1365-2656.13040

Wilson, R.P., Quintana, F. & Hobson, V.J. (2012). Construction of energy landscapes can clarify the movement and distribution of foraging animals. *Proceedings of the Royal Society B*, *279*, 975–980. https://doi.org/10.1098/rspb.2011.1544

Wood, S.N. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2nd edition.

Appendices

A A random walk as a special case of the ESF

The simple isotropic random walk (SRW) is unbiased and uncorrelated, with diffusion coefficient D, such that the density of a step starting at $\boldsymbol{x}$ and ending at $\boldsymbol{y}$ over time interval t is normally distributed around the start point: $y \sim N(x, \sigma^2)$, where $\sigma^2 = 4Dt$ (Codling et al., 2008). The likelihood of a step ending at $\boldsymbol{y}$ given that it started at $\boldsymbol{x}$ is $f(\boldsymbol{y}|\boldsymbol{x}) = \psi(\boldsymbol{y}|\boldsymbol{x}, \sigma^2)$, where ψ is the probability density function (PDF) of a normal distribution of $\boldsymbol{y}$ with mean $\boldsymbol{x}$ and variance σ^2.

We can write the SRW as a special case of the ESF. The ESF likelihood of a step ending at $\boldsymbol{y}$ given that it started at $\boldsymbol{x}$ is

$$f(\boldsymbol{y}|\boldsymbol{x}) = S^{-1} \exp[\beta_1 G(\boldsymbol{x}, \boldsymbol{y}) - \beta_2 C(\boldsymbol{x}, \boldsymbol{y})] \tag{1}$$

where S^{-1} is a normalization constant that ensures it is a PDF of $\boldsymbol{y}$. To write it as an SRW, we can set $\beta_1 = 0$ to represent no effect of energetic gains and $C(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{x})^2$. The likelihood then becomes $f(\boldsymbol{y}|\boldsymbol{x}) = S^{-1} \exp[-\beta_2 (\boldsymbol{y} - \boldsymbol{x})^2]$, which can be rewritten as

$$f(\boldsymbol{y}|\boldsymbol{x}) = S^{-1} \exp\left[-\frac{(\boldsymbol{y} - \boldsymbol{x})^2}{2\sigma^2}\right] \tag{2}$$

where $\beta_2 = \frac{1}{2\sigma^2}$. We recognize this as the PDF of a bivariate normal distribution with variance σ^2, mean $\boldsymbol{x}$, and $S = 2\pi\sigma^2$. This shows that an ESF with no gains and $C(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{x})^2$ (i.e. costs formulated as the step length squared) is equivalent to an SRW model.

B Radius size, R

To approximate Equation 2, we generate controls on a disc (Section 2.3). By using this approximation method, we therefore assume that the probability density function of a step ending at $\boldsymbol{y}$ given that it started at $\boldsymbol{x}$ over the area of the disc, is

$$f(\boldsymbol{y}|\boldsymbol{x}) = \begin{cases} \exp\{\beta_1 G(\boldsymbol{x}, \boldsymbol{y}) - \beta_2 C(\boldsymbol{x}, \boldsymbol{y})\} & \text{if } l_{xy} \leq R \\ 0 & \text{if } l_{xy} > R \end{cases} \tag{3}$$

Therefore, for this approximation to be accurate, the disc needs to be large enough so that the probability of a step longer than R is very small (Figure S1). If we define the radius as $R = l_m \times \gamma$, where l_m is the maximum

observed step length, the approximation will improve as γ increases. However, as the size of the disc becomes larger, so does the number of controls needed for the approximation. There is no straightforward way to assess this trade-off (i.e. the optimal size of R), but we can use importance sampling, based on where we expect the ESF to take large values. We explore the effect of the size of R on the approximation using simulated data, as well as comparing individual bear estimates approximated with different values of γ.

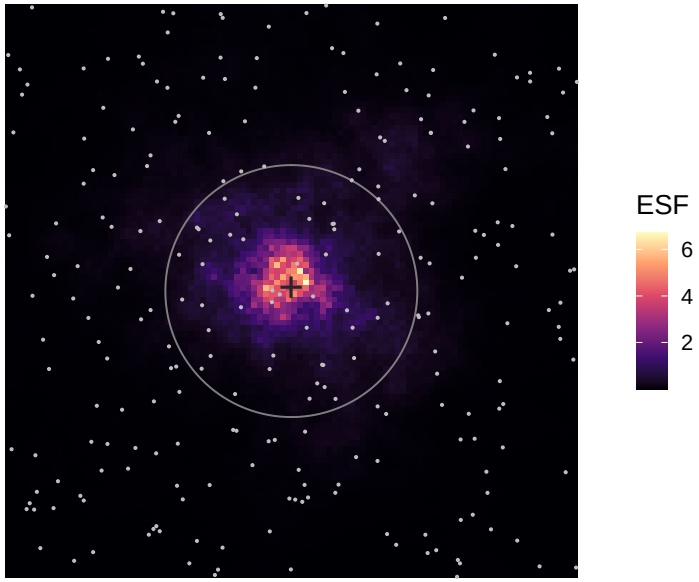


Figure S1: Plot illustrating importance sampling in the ESF. Ideally, we wish to sample uniformly over the entire study area (white dots). However, the ESF will decay with distance from the start point (+), due to the effect of step length on costs, and controls generated outside the disc will contribute very little to the approximation (i.e. their ESF is nearly zero). Therefore, for computational convenience, we can just sample within the disc, as long as the radius is large enough.

Simulated data We simulated 250 movement tracks $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ of length $n = 250$, as described in Section D. For each step, we generated 50 controls on a disc with a radius of the size $R = l_m \times \gamma$, where $\gamma = 0.5, 1.1, 2$. We fit the ESF for each movement track. As expected, β_2 was estimated with the lowest precision with the smallest radius ($\gamma = 0.5$). It was estimated correctly when $\gamma \geq 1.1$ (Figure S2). However, this represents a simplistic example, where the costs are *only* dependent on step length and $\beta_2 = 15$ is fairly strong selection against costs (i.e. the step length distribution should quickly decay to 0).

Real data We checked the effect of radius size on our polar bear telemetry data. We generated controls on a disc with $R_1 = 1.1 \times l_m$ and $R_2 = 2 \times l_m$ for each individual, fit the models separately, and then

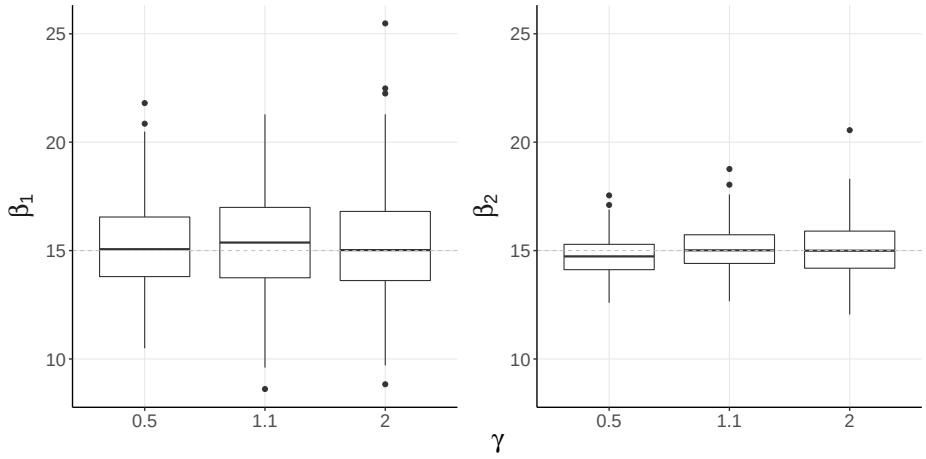


Figure S2: Estimates of β_1 and β_2 with $R = \gamma \times l_m$, where l_m is the maximum observed step length and $\gamma = 0.5, 1.1, 2$. Dashed line represents the true parameter value.

compared parameter estimates. Estimates varied up to ± 0.15 for β_2 (Figure S3), but followed the same general pattern. There was no evidence of systematic bias (i.e. underestimation or overestimation), and variation may be explained by the random generation of controls (which varied between the two trials). β_1 also varied between the two radius sizes, but this is likely attributable to high uncertainty in the estimates (i.e. no clear selection for gains).

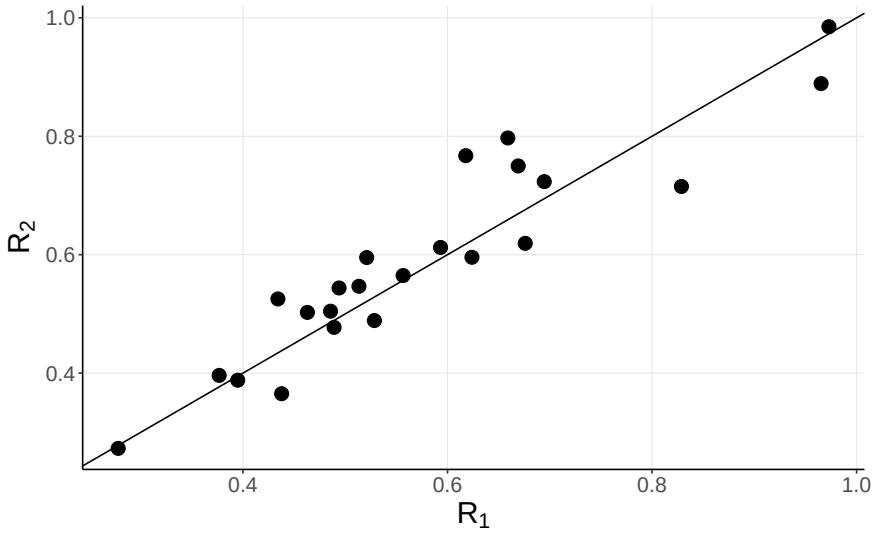


Figure S3: Individual estimates of β_2 with $R_1 = 1.1 \times l_m$ and $R_2 = 2.0 \times l_m$. Each point is an individual polar bear and the straight line represents a 1:1 relationship.

C Examples of gains and costs formulations

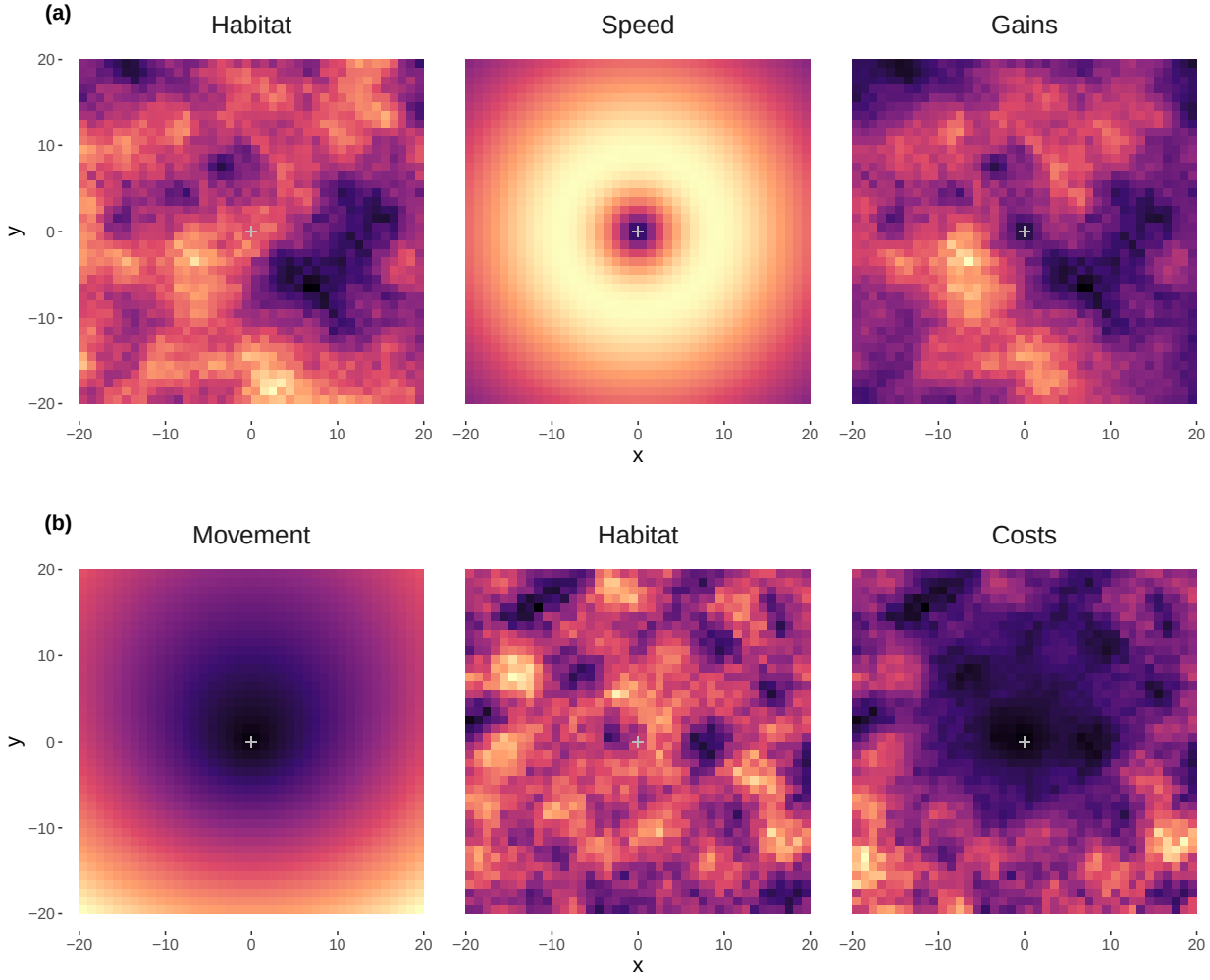


Figure S4: Example energetic gain G (a) and energetic cost C (b) formulations. In all panels, higher values are lighter in colour. In both (a) and (b), the third panel is a product of the first two panels, which represent movement and habitat components. In (a), energetic gains are composed of an energetically beneficial habitat covariate (e.g. forage biomass) scaled to the speed travelled. In this case, the effect of movement speed is gamma distributed ($k = 2$, $\theta = 2.2$) about the start point to represent decreased foraging potential at low and high speeds. In (b) the energetic costs are defined by the distance and turning angle from the start point (+; assuming movement up the y-axis), combined with a habitat covariate in which higher values increase energy expenditure.

D Simulation study

We ran simulations to assess the performance of the ESF inference method under different scenarios (i.e. to test the accuracy of the approximation; Section 2.3). The main objective was to recover model parameters from movement tracks simulated directly from the ESF, with known parameter values. For all simulations, G was defined as a random covariate field and C was calculated as the step length, both from $[0, 1]$ and assumed to be in the same units.

Algorithm We generated n locations $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$, with $\boldsymbol{x}_1$ selected randomly from the study area Ω. For each iteration $i = 1, \ldots, n-1$, we followed these steps to generate $\boldsymbol{x}_{i+1}$:

1. Simulate possible endpoints $\{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_K\}$ uniformly on a disc centred on $\boldsymbol{x}_i$, with a radius $R = 1$ and $K = 10,000$. We chose a very large value of K to ensure any bias was the result of the approximation, rather than the simulation itself.

2. Evaluate G and C at each endpoint.

3. For $k \in 1, 2, ...K$, sample x_{i+1} from $\{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_K\}$, with probabilities defined by

$$p_k = \frac{w(\boldsymbol{x}_i, \boldsymbol{z}_k)}{\sum_{j=1}^{K} w(\boldsymbol{x}_i, \boldsymbol{z}_j)}, \tag{4}$$

   where w is the ESF (Equation 2 of the main text).

Scenarios First, we assessed whether the selection strength affected our ability to estimate the parameters, and subsequently, whether certain foraging strategies may be harder to identify. For both β_1 and β_2, we considered 15 as a high parameter value. We considered low parameter values to be 0 for β_1 (no selection for gains), and 5 for β_2 (very weak selection against costs). We could not use 0 for β_2, as the ESF simulation algorithm would artificiality constrain the step length through the radius model. We tested different values of β_2 and found that 5 was the lowest parameter value where the size of the radius no longer affected the simulated step lengths. We combined these parameter values to represent the following movement patterns: i) optimal movement (high values of both parameters), ii) intake maximization (high β_1, low β_2), iii) cost minimization (low β_1, high β_2), and iv) movement nearly free of energetic considerations (low values of both parameters). Next, we altered the level of spatial autocorrelation in G. We simulated the study area Ω as a 1000×1000 raster with a resolution of 0.25, and assigned each grid cell a random value [$\sim U(0, 1)$]. We

calculated the covariate field for G by using a circular moving average window with diameter ρ (measured in grid units) to control the degree of spatial autocorrelation (Avgar et al., 2016; Michelot, 2019). We created random rasters of G with $\rho = 1, 5, 10, 25$ to reflect four levels of spatially autocorrelated habitat. For each of the 16 scenarios (parameter sets and spatial autocorrelation), we generated 250 movement tracks $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ of length $n = 250$. For each track, we tested the inference method using 20 and 200 control locations in the Monte Carlo integration procedure (Section 2.3). All parameters were estimated using MLE.

Results In most cases, the parameters were estimated accurately, although β_2 was generally estimated more precisely than β_1 (Figure S5). The median (min, max) difference between estimated and known parameter values was -0.04 (-65, 27) for β_1 and 0.04 (-4.8, 4.7) for β_2. Spatial autocorrelation in G had a noticeable effect on the precision of β_1 estimates, but not β_2. When β_1 was high, there was a pattern of decreased precision with increased autocorrelation. When β_1 was low, precision was lowest when spatial autocorrelation was very low ($\rho = 1$) and very high ($\rho = 50$). Spatial autocorrelation is a documented issue in resource selection analyses, which can lead to biased parameter estimates (Northrup et al., 2013). In SSFs and ESFs, high spatial autocorrelation decreases the range of the covariate space that may be explored for each movement step, which may decrease the ability to infer selection, particularly when the number of control locations is low (Northrup et al., 2013). However, in our simulations, the number of control locations used in Monte Carlo integration had negligible effects on the precision or accuracy of parameter estimations. Therefore, in most cases, 20 control locations should be adequate to approximate the likelihood, although we still recommend caution when working with highly spatially autocorrelated environmental covariates. As noted in Fortin et al. (2005), more controls may also be necessary when covariates are rare.

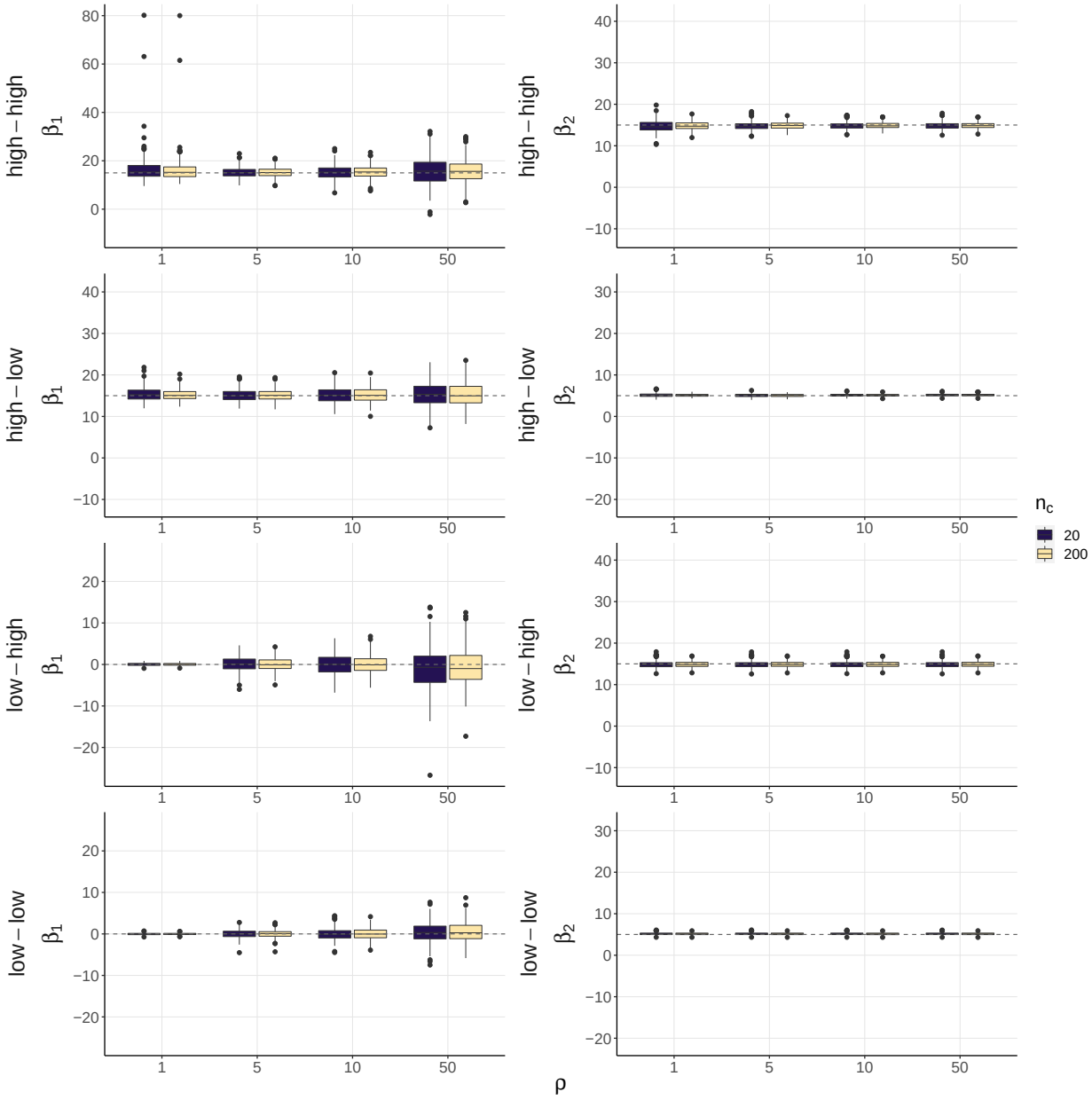


Figure S5: Parameter estimates from the simulations, under 32 different scenarios. Tracks were simulated either of four sets of parameters: "high-high" ($\beta_1 = 15$, $\beta_2 = 15$), "high-low" ($\beta_1 = 15$, $\beta_2 = 5$), "low-high" ($\beta_1 = 0$, $\beta_2 = 15$), and "low-low" ($\beta_1 = 0$, $\beta_2 = 5$). ρ refers to the level of spatial autocorrelation in the energetic gain covariate G, and n_c is the number of controls used in Monte Carlo integration. Dashed line is the true parameter value.

E Case study

E.1 Study area, field sampling, and data processing

Field sampling was done in Beaufort Sea, Canada (Figure S6). Sea ice in the area is mostly annual, with a flaw lead that separates near-shore areas of stable landfast ice and off-shore drifting pack ice (Carmack & Macdonald, 2002). The lead widens in spring and forms an active sea ice zone with high productivity (Pilfold et al., 2014), before most ice disappears by mid-summer (Stern & Laidre, 2016). Sea ice drift is characterized by the clockwise Beaufort Gyre, which is strengthening with climate change (Hutchings & Rigor, 2012; Petty et al., 2016), and increasing the energetic expenditure of polar bears in the area (Durner et al., 2017). Following standard capture procedures (Stirling et al., 1989), polar bears were sighted and immobilized in April-May of 2007-2011. Bears were fitted with GPS collars (Telonics, Mesa, AZ) set to a collect locations at a 4-hour resolution (relayed via the Argos satellite system; CLS America, Lanham, MD), and programmed to release after 1-2 years. The age of each bear was determined by analysing cementum growth layers of an extracted vestigial premolar (Calvert & Ramsay, 1998), and sex was determined in the field. The weight of each bear was calculated as

$$Weight(kg) = 0.00010309 \times AXG^{1.729} \times SLEN^{1.179} \tag{5}$$

where AXG is axillary girth (cm) and SLEN is straight body length (cm) taken at the time of capture (Thiemann et al., 2011). Capture and handling was approved by the University of Alberta BioSciences Animal Care and Use Committee following guidelines from the Canadian Council on Animal Care.

We omitted GPS locations from dropped collars or deceased bears, following Togunov et al. (2020). We defined a movement burst as a sequence of locations with no gaps > 24 hours and only kept bursts with ≥ 10 locations. We calculated step length as the Euclidean distance between projected GPS locations (NAD83 UTM Zone 9N, EPSG:3156), and removed unrealistic locations where the step speed was > 5.4 km/h (Whiteman et al., 2015). Then, we imputed missing locations of each burst with a continuous-time correlated random walk model (Johnson et al., 2008), implemented in *momentuHMM* (McClintock & Michelot, 2018).

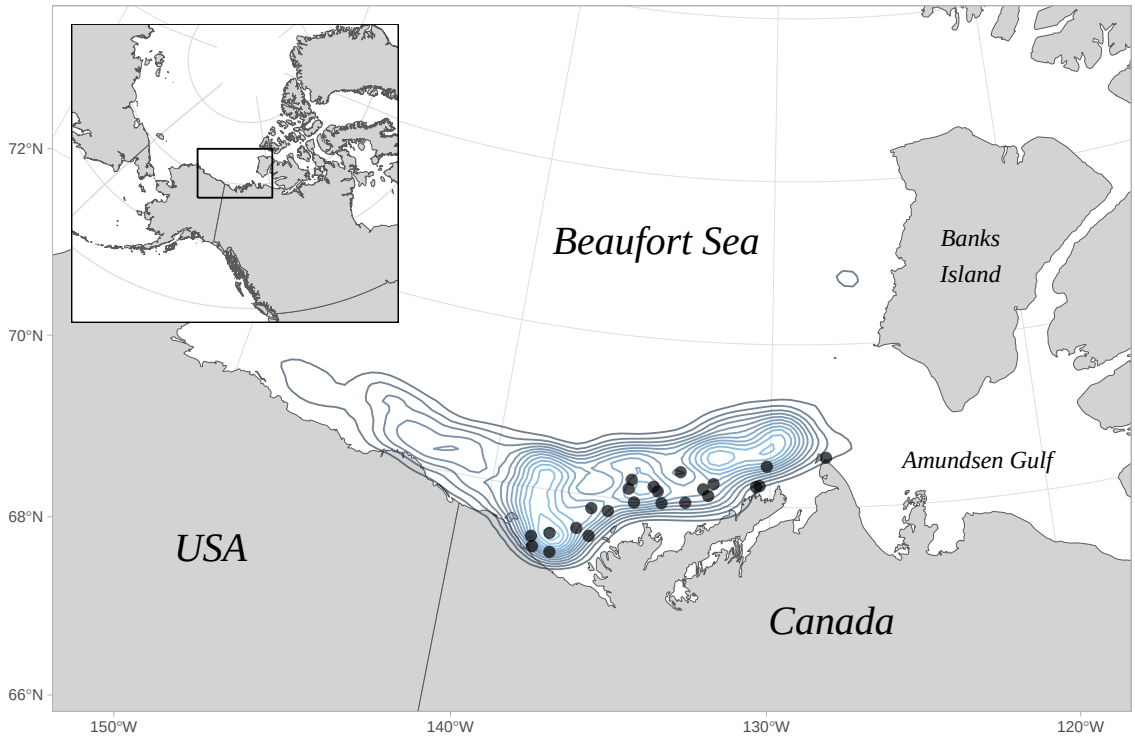


Figure S6: Study area in the Beaufort Sea, Canada. Circle points are polar bear collar deployment locations, and contour lines show the density of satellite telemetry data for all individuals (once regularised and limited to the spatiotemporal extent of the energetic gains raster).

E.2 Polar bear cost modelling

In this appendix, we present high-resolution complements to Figure 3 of the main text (Figures S7, S8), as well as a comparison of our cost model to doubly labelled water (Figure S9). In figure S9, we compared the estimated movement costs from our cost model (described in Section 3.2) with estimates of energy expenditure from doubly labelled water (Pagano & Williams, 2019). We estimated mean daily movement speeds and energy expenditure for adult female polar bears with > 6 locations per day for ≥ 25 days. We used linear regression to estimate the relationship between mean bear speed and daily energetic costs and compared this to the same relationship modelled in Pagano & Williams (2019). Our modelled costs were closely related to those from field DLW measurements, although slightly underestimated (Figure S9).

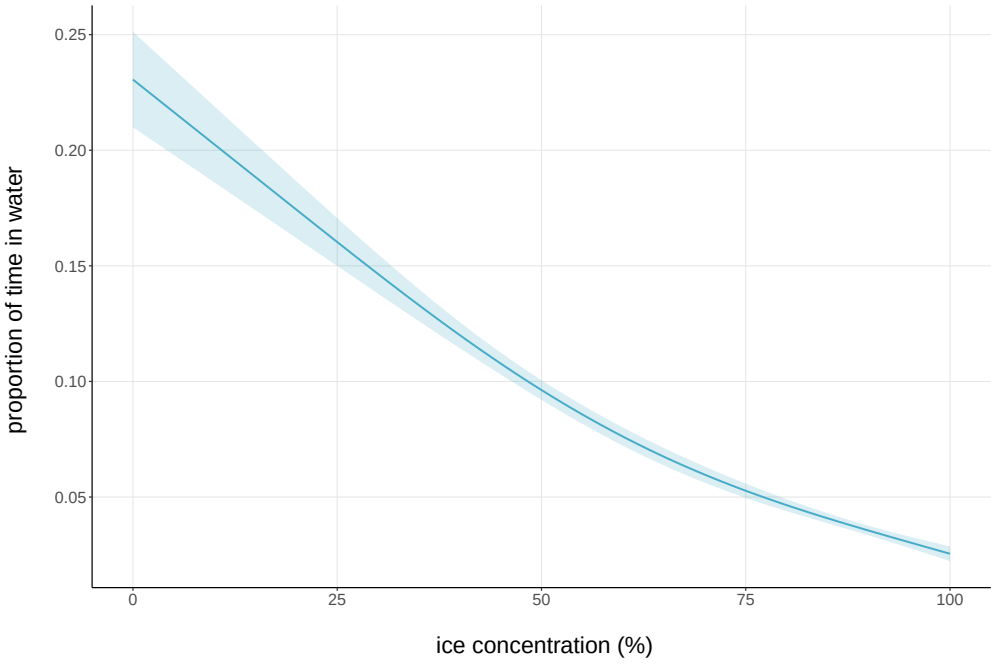


Figure S7: Estimated proportion of time spent in water relative to ice concentration, modelled with a generalized additive model (GAM). Data from Lone et al. (2018). Shaded area represents the standard error of the model fit.

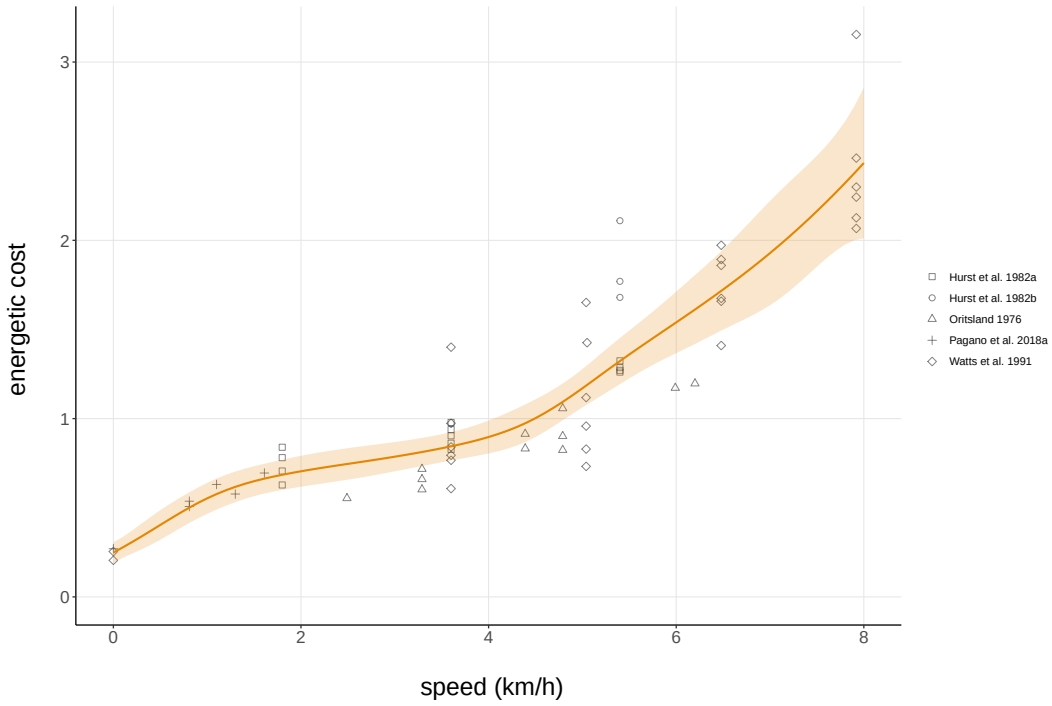


Figure S8: Relationship between polar bear walking speed (km/h) and energy expenditure (oxygen consumption; mL O2/g/hr) from six reference studies. Solid line is the predicted relationship from a monotonically constrained generalized additive model (gamma distribution, logit link function). Shaded area represents the standard error of the model fit.

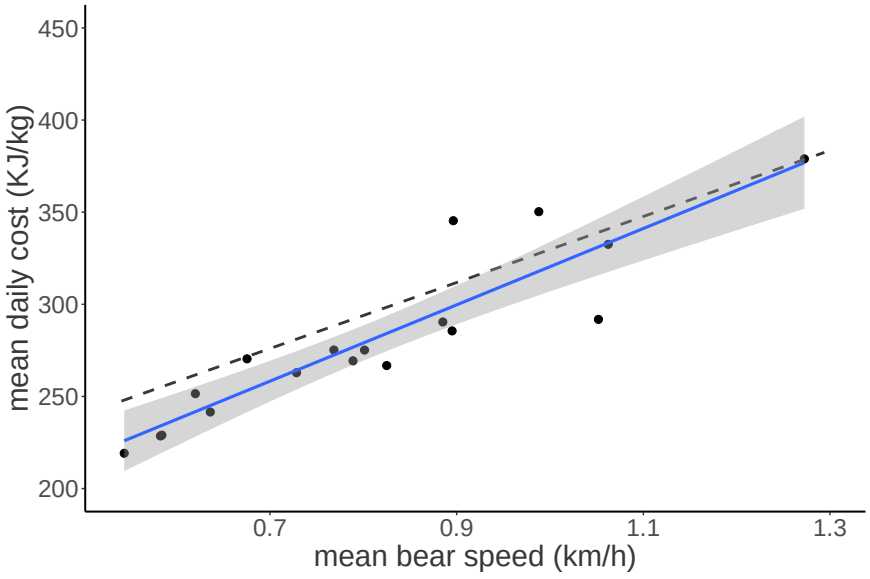


Figure S9: Relationship between mean daily movement speed (km/h) and daily energy expenditure (kJ/kg) for individuals with more than 25 days of locations with 6 locations (blue line; this study) compared to the estimated relationship from doubly-labelled water (dashed line; Pagano & Williams, 2019). Data from this study was from March - June of 2007-2011; data from Pagano & Williams (2019) was from April 2014 - 2016.

F ESF with turning angle, ESF_θ

Many animals display directional persistence at the time resolution of modern telemetry data, and this results in a non-uniform distribution of turning angles. Turning angles are therefore important predictors of an animal's movement, and there is incentive to model their contribution in movement and habitat selection models. For this reason, some recent SSF studies have proposed including turning angles as a covariate (Avgar et al., 2016; Brooke et al., 2020). In an ESF analysis, the turning angles can in principle be included through their contribution to the energetic costs directly (e.g., large turning angle requires high energy expenditure). However, this might only be possible when high-resolution data are available, as low-resolution turning angles cannot capture an animal's fine-scale movement tortuosity. In coarser data, one can adopt a hybrid approach where the turning angle (θ) is included as an additional covariate in ESF, to assess the contribution of directional persistence. Following (Avgar et al., 2016), we propose including the cosine of θ, so the ESF formula becomes $w(\boldsymbol{x}, \boldsymbol{y}) = \exp\{\beta_1 G(\boldsymbol{x}, \boldsymbol{y}) - \beta_2 C(\boldsymbol{x}, \boldsymbol{y}) + \beta_3 \cos(\theta)\}$. The cosine of θ ranges from -1 (for $\theta = \pm\pi$) to 1 (for $\theta = 0$), where a higher value corresponds to stronger directional persistence (smaller angle). A large coefficient β_3 therefore suggests strong selection for small angles, i.e., strong directional persistence. We fit this model to the polar bear tracking data, following the same procedures described in Section 3.3. We found that all bears showed selection for $\cos(\theta)$ (i.e. positive β_3), indicating that all individuals show directional persistence in their movement (Figure S10).

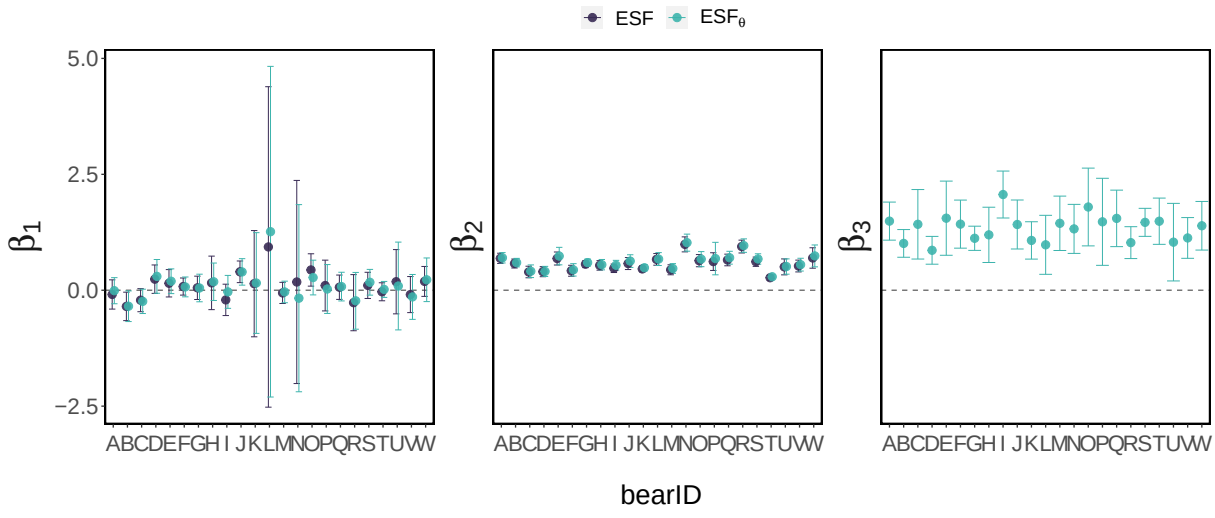


Figure S10: Estimated β coefficients of lone adult female polar bears ($N = 23$) for the ESF and ESF_θ. β_1 represents selection for gains, β_2 selection against costs, and β_3 selection for directional persistence.

G Robust uncertainty estimation

In step selection analyses, the standard error of selection parameters can be underestimated due to autocorrelation in the data (Fortin et al., 2005; Forester et al., 2009). This problem may also arise with the ESF, and we used the method described by Forester et al. (2009) to obtain robust standard errors.

1. Fit ESF model using either `clogit` or (equivalently) `coxph` in the survival R package.
2. Evaluate deviance residuals of the fitted model using the `residuals` function.
3. Sum the residual of each observed location with the residuals of the corresponding control locations, to get one residual for each stratum.
4. Use `acf` to plot the autocorrelation function of the residuals.
5. Visually determine the time lag over which the autocorrelation function decays to zero.
6. Split the data set into chunks of length that time lag, i.e., add a column 'cluster' in the data set to index the chunks.
7. Create two data sets: one containing chunks with even indices, and one containing chunks with odd indices. The chunks within each of the two data sets are approximately uncorrelated. If we denote as n the number of data points, and as p the time lag of autocorrelation, then this defines n/p clusters (rounded up), half of which go into the first sub-data set, and half into the second one.
8. Fit the ESF to each of the two data sets with the function `coxph`, using the option 'cluster' to specify that robust standard errors are required. (This automatically sets the 'robust' option to `TRUE`.)
9. Extract variance estimates from the two fitted models (i.e., diagonal elements of the robust covariance matrix returned by `coxph`), and take their mean to obtain the overall robust variance estimates.
10. Calculate Wald-type confidence intervals based on the robust variance estimates, i.e.,

$$\mathrm{CI}(\hat{\beta}_j) = \left[\hat{\beta}_j \pm 1.96 \times \mathrm{SE}(\hat{\beta}_j)\right],$$

for a 95% confidence interval, where $\hat{\beta}_j$ is the estimate for the parameter β_j, and $\mathrm{SE}(\hat{\beta}_j)$ is the corresponding robust standard error.

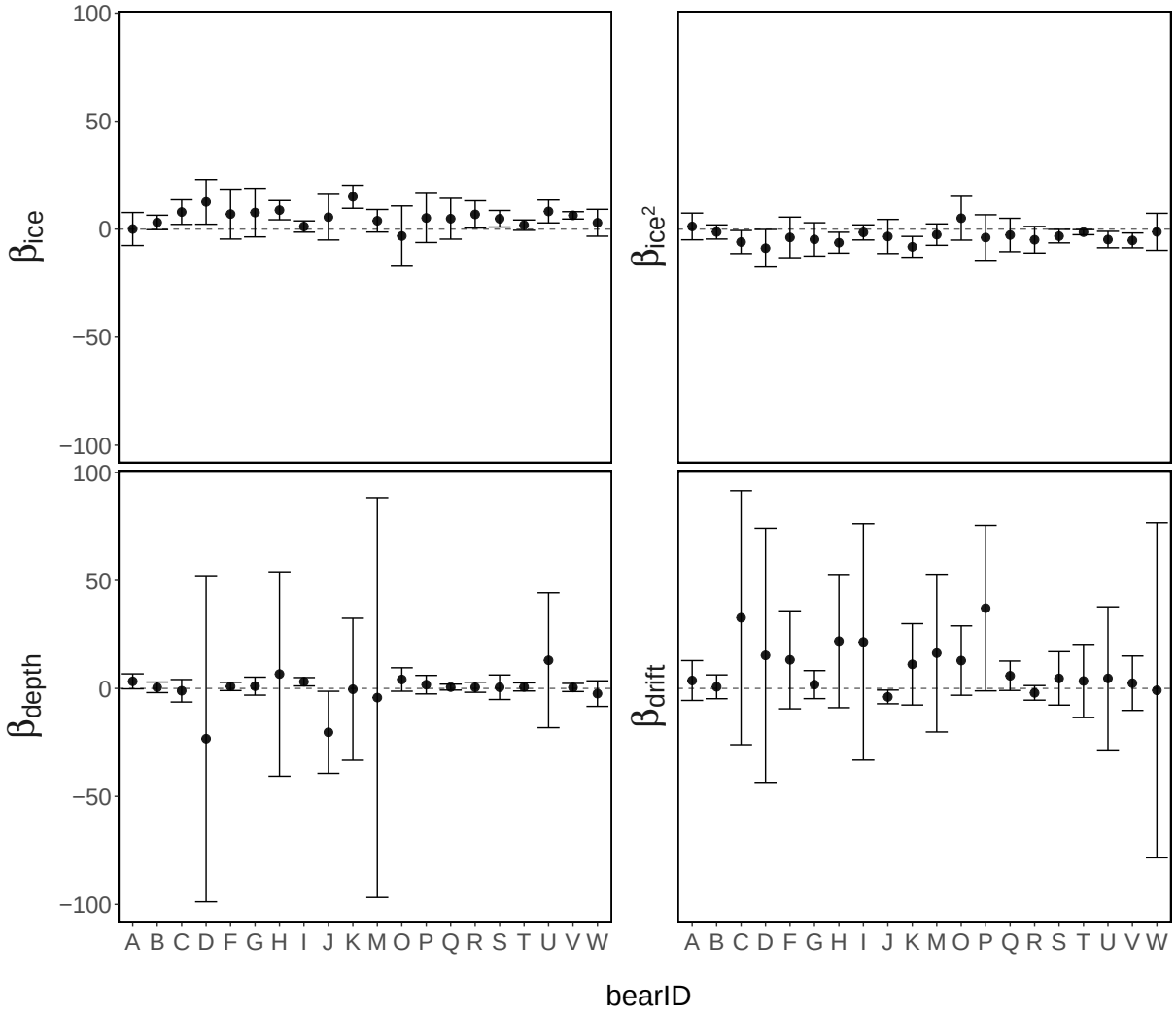


Figure S11: Estimated β coefficients of lone adult female polar bears ($N = 23$) for ice concentration (*ice*) and its square (ice^2), ocean depth (*depth*), and ice drift speed (*drift*). Error bars are robust 95% CIs.Three bears (E, L, N) were omitted for data visualization purposes. Coefficients ± SE for bears E, L, N (in order) are: $\beta_{ice} = 20.6 \pm 79.2, 875 \pm 1382, 44.8 \pm 119$, $\beta_{ice^2} = -15.0 \pm 45.6, -454 \pm 707, -22.9 \pm 66.1$, $\beta_{depth} = -37.8 \pm 66.2, 0.08 \pm 1.23, 0.69 \pm 1.90$, $\beta_{drift} = 7.0 \pm 19.9, -6.51 \pm 8.07, 9.13 \pm 8.63$

H Model checking using cross-validation

Monte Carlo cross-validation is the most common method to investigate goodness-of-fit of habitat selection models (Boyce et al., 2003). The idea is to assess how good the model is at correctly predicting which step, from the set of the observed and control steps, was selected at each time point. We followed the procedure suggested by Fortin et al. (2009), summarised below.

1. Randomly select a given proportion of strata in the data (e.g., 80%) to create a training data set. All other strata go into the validation data set.

2. Fit the ESF model to the training data set to obtain estimates of the selection parameters β_1 and β_2.

3. Evaluate the ESF w based on the estimated parameters, for each observed and control location in the validation data set.

4. For each stratum, compute the rank of the ESF score of the observed location (compared to ESF scores of control locations for that stratum). For example, the rank would be 21 if the observed location had a higher score than the 20 control locations, or it would be 1 if the ESF score of the observed location was the lowest. This returns n ranks (where n is the number of observations, or of strata in the data).

5. Evaluate Spearman's rank correlation between the ranks and their frequencies. This correlation measures how monotonic the relationship is, i.e., whether the frequency consistently increases with the rank.

6. Repeat steps 1-5 a large number (K) of times, with different random partitions for the training and validation data.

The Spearman's rank correlation is defined between -1 and 1. A correlation close to 1 indicates that the model consistently predicts that the observed location is the most likely, whereas a correlation close to -1 suggests that the model often ranks the observed location as unlikely (compared to control locations). Step 6 returns K correlation values, and their distribution must be inspected to assess the predictive power of the model. The approach proposed by Fortin et al. (2009) is to compare summary statistics (e.g., quantiles) of that distribution to those of the distribution of correlations obtained from chance alone (i.e., by running steps 1-6 on the control steps only, excluding the observed steps).

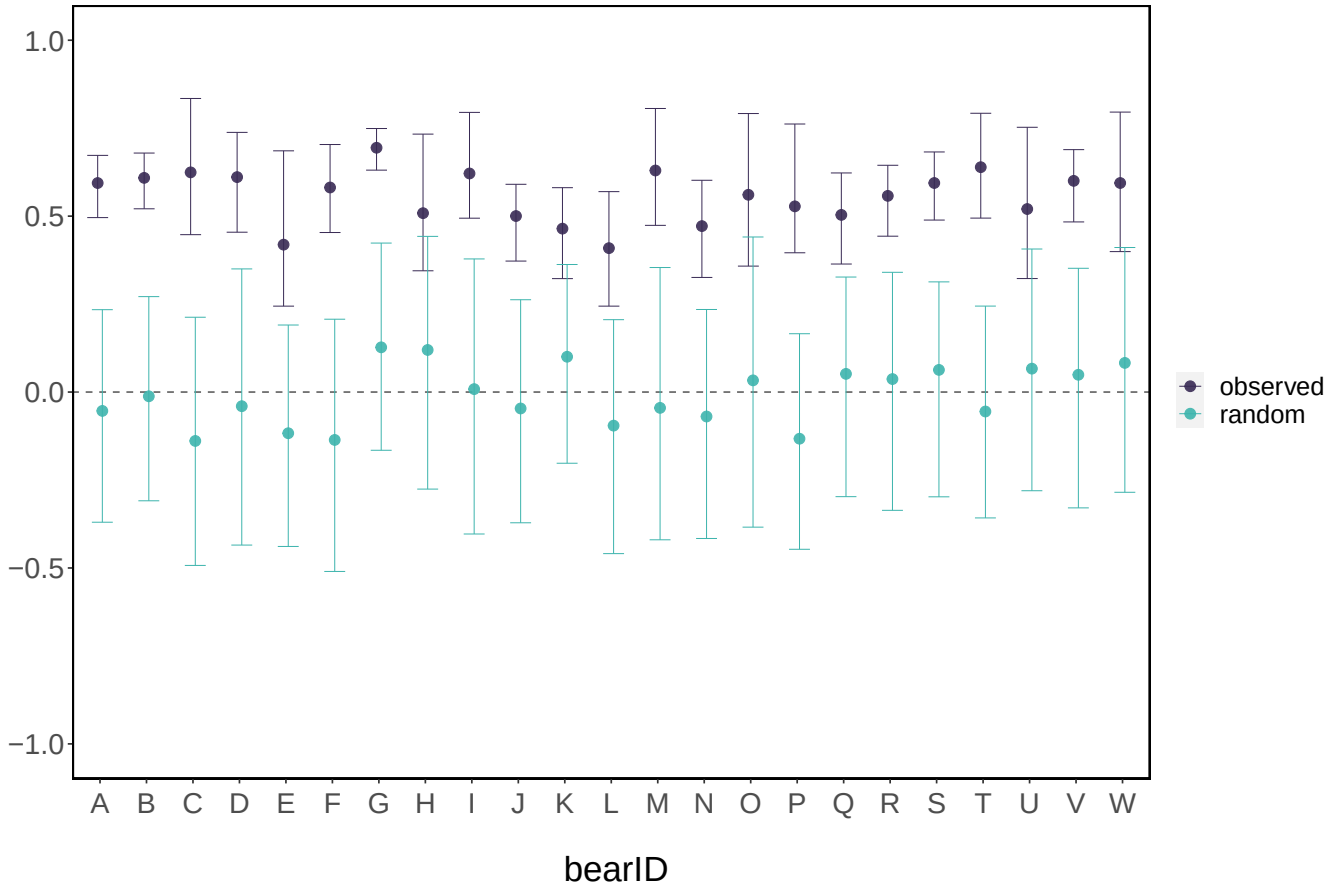


Figure S12: Monte Carlo cross-validation results for the ESF, performed with $K = 250$ data partitions for each individual (bearID). Each point is the mean Spearman rank correlation coefficient r_s for that individual and the error bars represent the range of 95% of the estimates. Observed refers to the cross-validation of the case locations, and random refers to cross-validation assessed on only control locations (i.e. by chance alone).

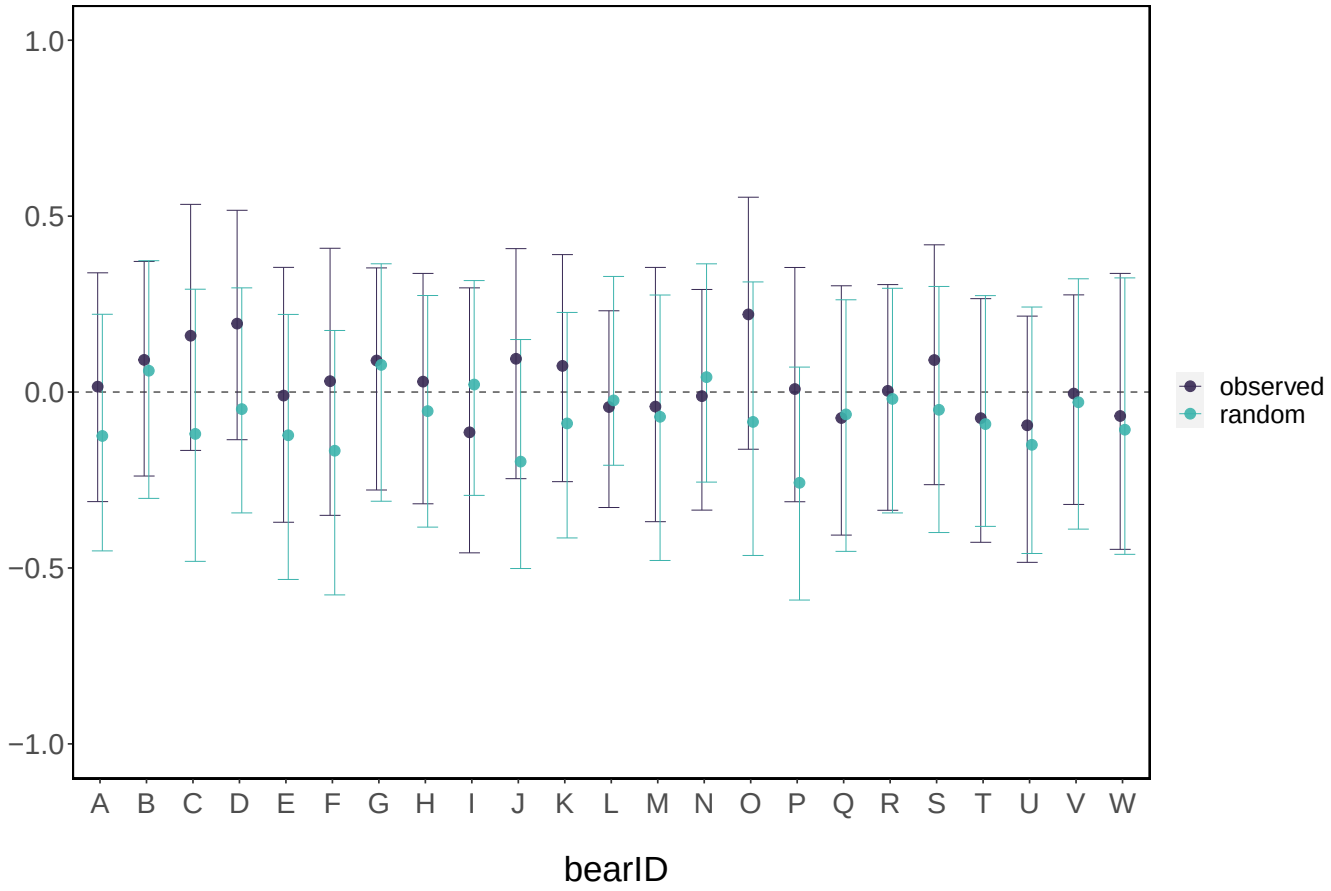


Figure S13: Monte Carlo cross-validation results for the SSF of Section 3.5, performed with $K = 250$ data partitions for each individual (bearID). Each point is the mean Spearman rank correlation coefficient r_s for that individual and the error bars represent the range of 95% of the estimates. Observed refers to the cross-validation of the case locations, and random refers to cross-validation assessed on only control locations (i.e. by chance alone).

H.1 Predictive power of gains and costs

Using the same Monte Carlo cross-validation procedure (described above), we also assessed the predictive power of a gains-only model $w(x, y) = \exp\{\beta G(x, y)\}$, and a cost-only model $w(x, y) = \exp\{\beta C(x, y)\}$. Figure S14 shows the range of Spearman rank correlation coefficients r_s obtained from cross-validation for each bear, which measures predictive performance. The figure contrasts the coefficients obtained based on the fitted model to those that would be obtained from chance alone. In the gain-only model, there is great overlap between the range of coefficients for the fitted model and the random experiment, suggesting that the model predictions are no better than chance alone (for most bears). Conversely, the cost-only model leads to consistently higher coefficients r_s, indicating better predictive performance. This is consistent with the observation made in the main text that, for this data set, there is strong evidence that movement is driven by avoidance of costly steps (whereas there is no evidence of selection of energetic gains).

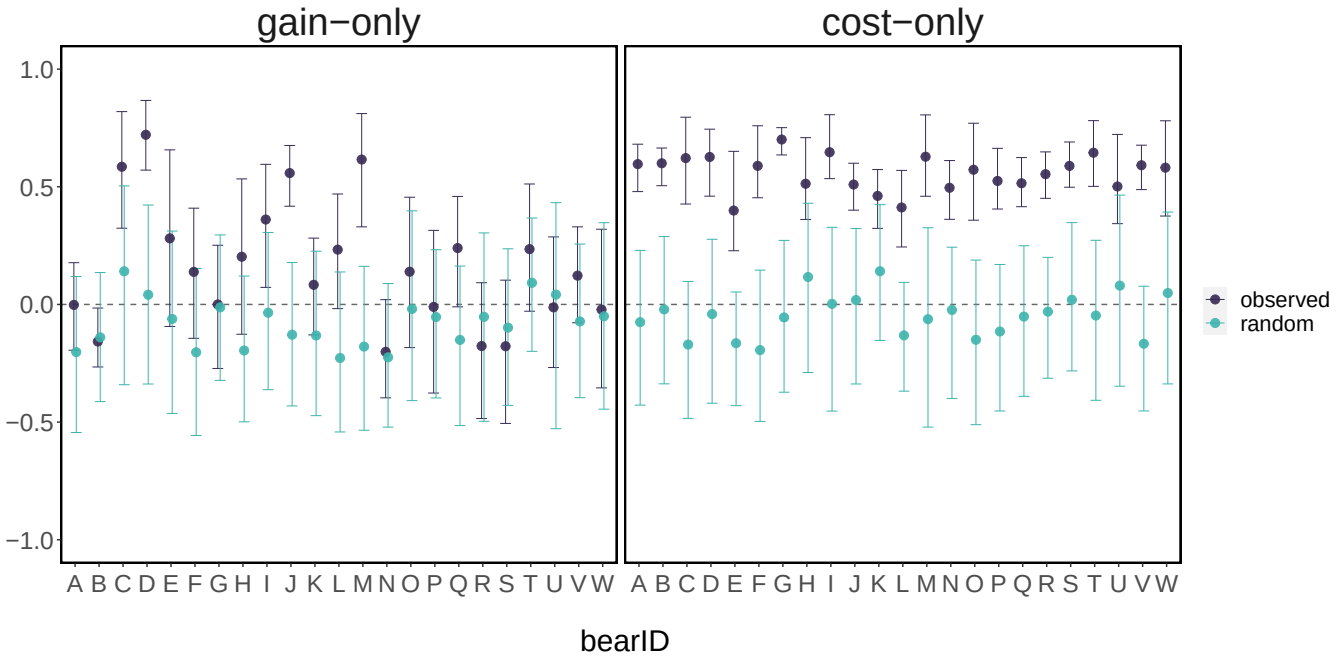


Figure S14: Monte Carlo cross-validation results for a gains-only and a cost-only model, performed with $K = 250$ data partitions for each individual (bearID). Each point is the mean Spearman rank correlation coefficient r_s for that individual and the error bars represent the range of 95% of the estimates. Observed refers to the cross-validation of the case locations, and random refers to cross-validation assessed on only control locations (i.e. by chance alone).

I Comparison to other models

In this appendix, we compare the model fit of the ESF (Section 2.1) to several other models: i) a null model given as a simple random walk (SRW), ii) the ESF with turning angle presented in Appendix F, and iii) an SSF with non-energetic covariates. We compared the models using the Akaike Information Criterion (AIC), defined as $AIC = 2 \times nllk + 2k$ where $nllk$ is the negative log-likelihood and k is the number of parameters in each model. Comparison using AIC requires the same data, and so we used the same case-control data set for each model. This required excluding locations for which any of the covariates were missing, to ensure that those locations were not included in some models are excluded from others. Following Sections 2.2 and 4.1, we fitted each model separately for each individual using the *survival* R package. We note that AIC might be unreliable in the presence of residual autocorrelation, and future work could focus on developing an alternative, such as the QIC used for generalised estimating equations.

1. **Simple random walk (null model):** We compared the ESF to a null model, given by the simple random walk (SRW) described in Appendix A. Costs were defined as l^2, where l is the ice drift-corrected bear step length (km), and ranged from $0 - 509$ km^2.

2. **ESF with turning angle, ESF$_\theta$:** See Appendix F.

3. **SSF with movement covariates:** Comparing ESF and SSF models using AIC requires not only the case locations to be identical, but also the control locations. Lastly, we therefore compared the ESF to an SSF with uniform control locations over a disc, for which step length and turning angle were included as covariates (similarly to Brooke et al., 2020), in addition to the environmental covariates described in Section 3.5.

Results The ESF$_\theta$ or the SSF were indicated to be the best models in all cases: ESF$_\theta$ was the best model for 7 bears, the SSF was the best in 9 cases, and 7 cases had no clear best model between the two (i.e. the ΔAIC of the closest competing model was between 0 and 5). The SRW was most commonly the worst fitting model, and was better than the ESF in 5 cases. The best models always have turning angle θ included, which indicates that the ESF$_\theta$ may be useful in cases where θ cannot easily be integrated into C.

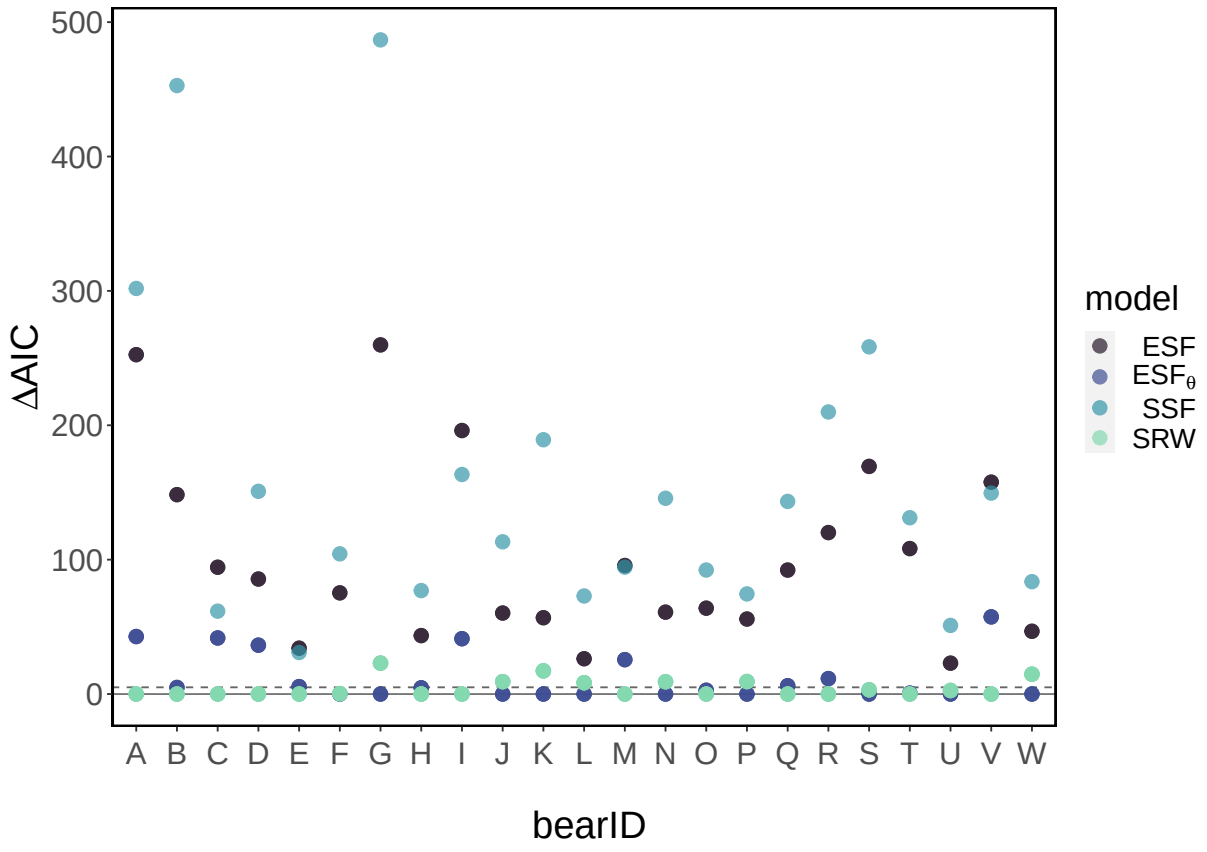


Figure S15: ΔAIC scores of each model for each individual (bearID). Solid line at 0 indicates the best model, dashed line at 5 is the threshold for strong support (i.e. any competing models within this threshold are also plausible).

References

Avgar, T., Potts, J.R., Lewis, M.A. & Boyce, M.S. (2016). Integrated step selection analysis: bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution*, *7*, 619–630. https://doi.org/10.1111/2041-210X.12528

Boyce, M.S., Mao, J.S., Merrill, E.H., Fortin, D., Turner, M.G., Fryxell, J. & Turchin, P. (2003). Scale and heterogeneity in habitat selection by elk in Yellowstone National Park. *Ecoscience*, *10*, 421–431. https://doi.org/10.1080/11956860.2003.11682790

Brooke, C.F., Fortin, D., Kraaij, T., Fritz, H., Kalule-Sabiti, M.J. & Venter, J.A. (2020). Poaching impedes the selection of optimal post-fire forage in three large grazing herbivores. *Biological Conservation*, *241*, 108393. https://doi.org/10.1016/j.biocon.2019.108393

Calvert, W. & Ramsay, M.A. (1998). Evaluation of age determination of polar bears by counts of cementum growth layer groups. *Ursus*, *10*, 449–453.

Carmack, E.C. & Macdonald, R.W. (2002). Oceanography of the Canadian shelf of the Beaufort Sea: A setting for marine life. *Arctic*, *55*, 29–45. https://doi.org/10.14430/arctic733

Codling, E.A., Plank, M.J. & Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal Society Interface*, *5*, 813–834. https://doi.org/10.1098/rsif.2008.0014

Durner, G.M., Douglas, D.C., Albeke, S.E., Whiteman, J.P., Amstrup, S.C., Richardson, E.S., Wilson, R.R. & Ben-David, M. (2017). Increased Arctic sea ice drift alters adult female polar bear movements and energetics. *Global Change Biology*, *23*, 3460–3473. https://doi.org/10.1111/gcb.13746

Forester, J., Kyung Im, H. & Rathouz, P. (2009). Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, *90*, 3554–3565. https://doi.org/10.1890/08-0874.1

Fortin, D., Beyer, H.L., Boyce, M.S., Smith, D.W., Duchesne, T. & Mao, J.S. (2005). Wolves influence elk movements: behaviour shapes a trophic cascade in Yellowstone National Park. *Ecology*, *86*, 1320–1330. https://doi.org/10.1890/04-0953

Fortin, D., Fortin, M.E., Beyer, H.L., Duchesne, T., Courant, S. & Dancose, K. (2009). Group-size-mediated habitat selection and group fusion – fission dynamics of bison under predation risk. *Ecology*, *90*, 2480–2490.

Hutchings, J.K. & Rigor, I.G. (2012). Role of ice dynamics in anomalous ice conditions in the Beaufort Sea during. *Journal of Geophysical Research*, *117*, C00E04. https://doi.org/10.1029/2011JC007182

Johnson, D.S., London, J.M., Lea, M.A. & Durban, J.W. (2008). Continous-time correlated random walk model for animal telemetry data. *Ecology*, *89*, 1208–1215.

Lone, K., Kovacs, K.M., Lydersen, C., Fedak, M., Andersen, M., Lovell, P. & Aars, J. (2018). Aquatic behaviour of polar bears (*Ursus maritimus*) in an increasingly ice-free Arctic. *Scientific Reports*, *8*, 9677. https://doi.org/10.1038/s41598-018-27947-4

McClintock, B.T. & Michelot, T. (2018). momentuHMM: R package for generalized hidden Markov models of animal movement. *Methods in Ecology and Evolution*, *9*, 1518–1530. https://doi.org/10.1111/2041-210X.12995

Michelot, T. (2019). *Stochastic models of animal movement and habitat selection*. Ph.D. thesis, University of Sheffield.

Northrup, J.M., Hooten, M.B., Anderson, C.R.J. & Wittemyer, G. (2013). Practical guidance on characterizing availability in resource selection functions under a use-availability design. *Ecology*, *94*, 1456–1463. https://doi.org/10.1890/12-1688.1

Pagano, A.M. & Williams, T.M. (2019). Estimating the energy expenditure of free-ranging polar bears using tri-axial accelerometers: A validation with doubly labeled water. *Ecology and Evolution*, *9*, 4210–4219. https://doi.org/10.1002/ece3.5053

Petty, A.A., Hutchings, J.K., Richter-Menge, J.A. & Tschudi, M.A. (2016). Sea ice circulation around the Beaufort Gyre: the changing role of wind forcing and the sea ice state. *Journal of Geophysical Research: Oceans*, *121*, 3278–3296. https://doi.org/10.1002/2015JC010903

Pilfold, N.W., Derocher, A.E. & Richardson, E.S. (2014). Influence of intraspecific competition on the distribution of a wide-ranging, non-territorial carnivore. *Global Ecology and Biogeography*, *23*, 425–435. https://doi.org/10.1111/geb.12112

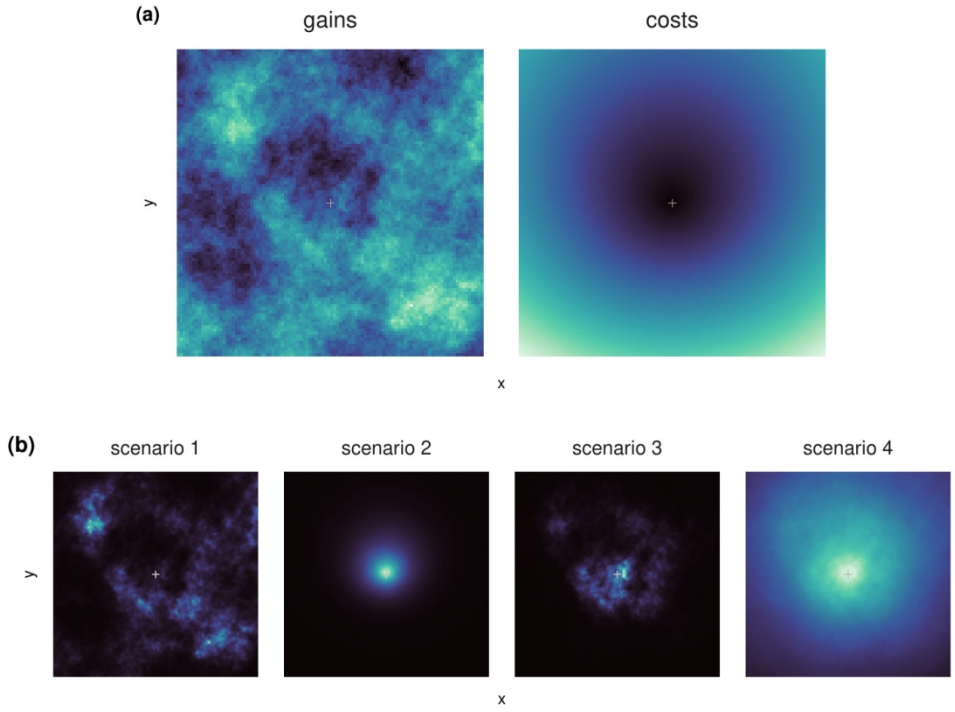
Stern, H.L. & Laidre, K.L. (2016). Sea-ice indicators of polar bear habitat. *The Cryosphere*, *10*, 2027–2041. https://doi.org/10.5194/tc-10-2027-2016

Stirling, I., Spencer, C. & Andriashek, D.S. (1989). Immobilization of polar bears (*Ursus maritimus*) with Telazol in the Canadian Arctic. *Journal of Wildlife Diseases*, *25*, 159–168.

Thiemann, G.W., Sciences, B. & Tg, C.A.B. (2011). Temporal Change in the Morphometry – Body Mass Relationship of Polar Bears. *Journal of Wildlife Management*, *75*, 580–587. https://doi.org/10.1002/jwmg.112

Togunov, R.R., Klappstein, N.J., Lunn, N.J., Derocher, A.E. & Auger-Méthé, M. (2020). Opportunistic evaluation of modelled sea ice drift using passively drifting telemetry collars in Hudson Bay, Canada. *The Cryosphere*, *14*, 1937–1950. https://doi.org/10.5194/tc-14-1937-2020

Whiteman, J.P., Harlow, H.J., Durner, G.M., Anderson-Sprecher, R., Albeke, S.E., Regehr, E.V., Amstrup, S.C. & Ben-David, M. (2015). Summer declines in activity and body temperature offer polar bears limited energy savings. *Science*, *349*, 295–298. https://doi.org/10.1126/science.aaa8623



533x394mm (72 x 72 DPI)