Statistical Methodology 17 (2014) 82–98

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

Temporal variation and scale in movement-based resource selection functions

M.B. Hooten [a,b,c,*], E.M. Hanks [c], D.S. Johnson [d], M.W. Alldredge [e]

[a] U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Fort Collins, CO, USA
[b] Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO, USA
[c] Department of Statistics, Colorado State University, Fort Collins, CO, USA
[d] National Marine Mammal Laboratory, National Oceanic and Atmospheric Administration, Seattle, WA, USA
[e] Colorado Parks and Wildlife, Fort Collins, CO, USA

ARTICLE INFO

Article history:
Received 9 June 2012
Received in revised form
8 December 2012
Accepted 9 December 2012

Keywords:
Animal movement
Kullback–Leibler
Telemetry data

ABSTRACT

A common population characteristic of interest in animal ecology studies pertains to the selection of resources. That is, given the resources available to animals, what do they ultimately choose to use? A variety of statistical approaches have been employed to examine this question and each has advantages and disadvantages with respect to the form of available data and the properties of estimators given model assumptions. A wealth of high resolution telemetry data are now being collected to study animal population movement and space use and these data present both challenges and opportunities for statistical inference. We summarize traditional methods for resource selection and then describe several extensions to deal with measurement uncertainty and an explicit movement process that exists in studies involving high-resolution telemetry data. Our approach uses a correlated random walk movement model to obtain temporally varying use and availability distributions that are employed in a weighted distribution context to estimate selection coefficients. The temporally varying coefficients are then weighted by their contribution to selection and combined to provide inference at the population level. The result is an intuitive and accessible statistical procedure that uses readily available software and is computationally feasible for large datasets. These methods are

* Corresponding author at: Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523-1484, USA.
E-mail address: Mevin.Hooten@colostate.edu (M.B. Hooten).

1572-3127/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.stamet.2012.12.001

demonstrated using data collected as part of a large-scale mountain lion monitoring study in Colorado, USA.

Published by Elsevier B.V.

1. Introduction

An explosion of recent papers on the statistical analysis of animal movement indicates rapid growth in this emerging area of animal ecology. The formal mathematical description of animal movement is quite old, dating back hundreds of years, and even though the seminal text on the topic written by Turchin [47] is quite relevant, it is currently out of print and lacks a contemporary statistical perspective. Despite the existence of highly technical literature describing ways to model animal movement (e.g., [7,15]), most applied studies that have actual management or conservation objectives have sought to focus more on what is termed "space use", in which they seek to characterize the geographical and/or environmental space used by either individual animals or populations or both.

Types of space use analyses vary widely and include: (1) describing an individual's home range or core area (e.g., [52]) (2) describing the spatial distribution (probability density function or "utilization distribution", e.g. [34]) from which individual's locations ($\mathbf{s}_t$, for time $t \in \mathcal{T}$) might arise and (3) the estimation of resource selection functions (or RSFs, e.g., [32]). In the latter, statistical inference is focused on identifying the probability of resource use *given* resource availability (i.e., selection).

1.1. Traditional resource selection

The basic statistical approach put forth by Manly et al. [32], and since extended in several different directions (e.g., [23,30,24,29]), specifies that the distribution of use $[\mathbf{x}]_u$ is equal to a weighted distribution of availability $[\mathbf{x}]_a$:

$$\begin{aligned}[\mathbf{x}]_u &= \frac{g(\mathbf{x}, \boldsymbol{\beta})[\mathbf{x}]_a}{\int g(\boldsymbol{\nu}, \boldsymbol{\beta})[\boldsymbol{\nu}]_a d\boldsymbol{\nu}}, \\ &= c\, g(\mathbf{x}, \boldsymbol{\beta})\, [\mathbf{x}]_a, \\ &= [\mathbf{x}|\boldsymbol{\beta}]_u, \end{aligned} \tag{1}$$

where, the square bracket notation '[. . .]' denotes a probability density function, $\mathbf{x}$ corresponds to a vector of resource covariates, $\boldsymbol{\beta}$ is a set of selection parameters (often regression coefficients), c is a normalizing constant, and $g(\mathbf{x}, \boldsymbol{\beta})$ is the resource selection function (often the inverse logit or exponential function, depending on the desired inference). The last line in (1) is not commonly used elsewhere, but is a notation that we will make use of in what follows.

Importantly, as others point out (e.g., [40,30,49,12]), the equation in (1) is referred to as a "weighted distribution" and can be arrived at by an application of Bayes rule. Making a slight modification to the typical weighted distribution specification, we index the resource observations by the location $\mathbf{s}_t$ at which they were observed at time t. Thus, without loss of generality we may write:

$$\begin{aligned}[\mathbf{x}(\mathbf{s}_t)]_u &= \frac{g(\mathbf{x}(\mathbf{s}_t), \boldsymbol{\beta})[\mathbf{x}(\mathbf{s}_t)]_a}{\int g(\mathbf{x}(\mathbf{s}), \boldsymbol{\beta})[\mathbf{x}(\mathbf{s})]_a d\mathbf{s}}, \\ &= [\mathbf{x}(\mathbf{s}_t)|\boldsymbol{\beta}]_u. \end{aligned} \tag{2}$$

Now, assuming independent observations $\mathbf{x}(\mathbf{s}_t)$ for $t = 1, \ldots, T$ and a known availability distribution $[\mathbf{x}(\mathbf{s}_t)]_a$, the likelihood can be written as the product of the right-hand-side of (2):

$$\prod_{t=1}^{T} [\mathbf{x}(\mathbf{s}_t)|\boldsymbol{\beta}]_u, \tag{3}$$

and can be maximized with respect to the selection coefficients $\boldsymbol{\beta}$, given the data, as long as the integral in the denominator of (2) can be computed. Conveniently, it has been shown that the likelihood in