Statistical inference for home range overlap

Kevin Winner [a1], Michael J. Noonan [b2,3], Chris H. Fleming [c2,3], Kirk A. Olson [d1,4], Thomas Mueller [e2,5,6], Dan Sheldon [f1,7], and Justin M. Calabrese [g2,3]

[1]College of Information and Computer Sciences, University of Massachusetts Amherst, 140 Governors Drive, Amherst, MA 01003, USA
[2]Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Rd., Front Royal, VA 22630, USA
[3]Department of Biology, University of Maryland, College Park, MD 20742, USA
[4]Wildlife Conservation Society, Mongolia Program. Ulaanbaatar, Mongolia
[5]Senckenberg Biodiversity and Climate Research Centre, Senckenberg Gesellschaft für Naturforschung, Senckenberganlage 25, 60325 Frankfurt (Main), Germany
[6]Department of Biological Sciences, Goethe University, Max-von-Laue-Straße 9, 60438, Frankfurt (Main), Germany
[7]Department of Computer Science, Mount Holyoke College, 50 College Street, South Hadley, MA 01075, USA

Running title: Inferential overlap

Article Type: Research article

Words in the abstract: 319

Words in the main text: 7831

Number of references: 46

Number of figures: 5

Number of tables: 0

Number of text boxes: 0

Supplementary material: 2

Author contributions: KW and MJN contributed equally to this work. CHF and JMC conceived the study. KW and CHF developed the methods. KAO and TM collected the data. MJN conducted the analyses. MJN and JMC drafted the manuscript. All authors contributed to study concepts and writing.

[a]kwinner@cs.umass.edu [b]NoonanM@si.edu [c]FlemingC@si.edu [d]kirkaolson@hotmail.com [e]muellert@gmail.com [f]sheldon@cs.umass.edu [g]**Corresponding Author**: Tel: +1 540-635-5682; Email: CalabreseJ@si.edu

Abstract

1. Despite the routine nature of estimating overlapping space use in ecological research, to date no formal inferential framework for home range overlap has been available to ecologists. Part of this issue is due to the inherent difficulty of comparing the estimated home ranges that underpin overlap across individuals, studies, sites, species, and times. Because overlap is calculated conditionally on a pair of home range estimates, biases in these estimates will propagate into biases in overlap estimates. Further compounding the issue of comparability in home range estimators is the historical lack of confidence intervals on overlap estimates. This means that it is not currently possible to determine if a set of overlap values are statistically different from one another.

2. As a solution, we develop the first rigorous inferential framework for home range overlap. Our framework is based on the AKDE family of home range estimators, which correct for biases due to autocorrelation, small effective sample size, and irregular sampling in time. Collectively, these advances allow AKDE estimates to validly be compared even when sampling strategies differ. We then couple the AKDE estimates with a novel bias-corrected Bhattacharyya Coefficient (BC) to quantify overlap. Finally, we propagate uncertainty in the AKDE estimates through to overlap, and thus are able to put confidence intervals on the BC point estimate.

3. Using simulated data, we demonstrate how our inferential framework provides accurate overlap estimates, and reasonable coverage of the true overlap, even at small sample sizes. When applied to empirical data, we found that building an interaction network for Mongolian gazelles (*Procapra gutturosa*) based on all possible ties, versus only those ties with statistical support, substantially influenced the network's properties and any potential biological inferences derived from it.

4. Our inferential framework permits researchers to calculate overlap estimates that

can validly be compared across studies, sites, species, and times, and test whether observed differences are statistically meaningful. This method is available via the `R` package `ctmm`.

Keywords: Animal movement, Bhattacharyya Coefficient, AKDE, KDE, Kernel Density Estimate, Autocorrelation, ctmm

Introduction

Ecologists have long been interested in patterns and drivers of animal space use (Burt, 1943; Brown & Orians, 1970; Jetz, 2004). Decisions on what areas to occupy can influence fitness through a wide range of pathways such as foraging efficiency (Mitchell & Powell, 2012) or predator-prey dynamics (Mitchell & Lima, 2002), and even drive evolutionary trajectories (Lukas & Clutton-Brock, 2013). Related to this is the question of overlapping space use between individuals and/or populations. Quantifying overlap can provide an informative metric for testing hypotheses on inter-specific competition (Berger & Gese, 2007), territoriality (Grant *et al.*, 1992), and mating systems (Powell, 1979). Furthermore, overlap can be used to underpin analyses of social network structure (Frère *et al.*, 2010), and contact rates, with implications for disease transmission (Sanchez & Hudgens, 2015; Dougherty *et al.*, 2018). Trends in overlapping space use are also routinely used in determining allometric scaling laws (Grant *et al.*, 1992; Jetz, 2004). The rapid increase in both the availability and quality of tracking data in recent years (Kays *et al.*, 2015) has made the concept of home range (HR) overlap increasingly relevant. Ecologists are now in a position to address overlap-related questions for a larger number of species and individuals, in more ecosystems, and with more accurate data than ever before.

Despite these advances, a formal inferential framework for HR overlap is still lacking. Overlap is typically quantified by first estimating HRs from tracking data, and then applying an overlap metric to the range estimates (Millspaugh *et al.*, 2004; Fieberg & Kochanny, 2005). A wide range of overlap metrics have been proposed in the literature, spanning the gamut from *ad hoc* indices to more formal measures. These different metrics have contrasting properties and can produce highly different overlap estimates on the same data (see Millspaugh *et al.*, 2004; Fieberg & Kochanny, 2005). Further compounding this problem is the inherent difficulty of comparing the estimated HRs that underpin overlap across studies, sites, species, and times (Fleming & Calabrese, 2017). There is broad agreement in the literature that HR estimates based on different sampling strategies are difficult to compare, as they may be exposed to different degrees of bias (Frair *et al.*, 2010; Fieberg & Börger, 2012; Fleming *et al.*, 2018). More subtly, even identical sampling strategies can still produce differentially biased HR estimates if the underlying parameters of movement differ among individuals in the comparison (Fleming & Calabrese, 2017). Because overlap is calculated conditionally on a pair of HR estimates, biases in the HR estimates will propagate into biases in overlap estimates (Fieberg & Kochanny, 2005). It follows then that *differential* biases in HR estimates among different groups of interest will tend to propagate into differential biases in overlap estimates, rendering comparisons difficult to interpret and potentially unreliable.

Additionally, none of the overlap metrics of which we are aware come equipped with confidence intervals to quantify the uncertainty in the estimates. This means that it is currently not possible to determine if a set of overlap values are statistically different from one another, or from a reference value of interest. To see this, consider a case where one wishes to compare two overlap estimates from two pairs of individuals: 0.35 and 0.55. If the 95% confidence intervals for each estimate are disjoint, then we may conclude that the two pairs have significantly different measures of overlap. If, on the other hand, the 95% confidence intervals are not disjoint, then the point estimates may not be significantly different. In other words, without confidence intervals, one cannot properly interpret differences between estimates (Pawitan, 2001).

Here, we develop the first inferential framework for HR overlap by building on previous work in quantifying overlap (Fieberg & Kochanny, 2005) and by leveraging recent advances in HR estimation (Fleming *et al.*, 2015a; Fleming & Calabrese, 2017; Fleming *et al.*, 2018). We base our approach on the Bhattacharyya Coefficient (BC; Bhattacharyya, 1943, also called the Bhattacharyya Affinity), which has a formal basis as a measure of similarity between two probability distributions, and is straightforward to calculate, and interpret (Fieberg & Kochanny, 2005). We couple the BC with autocorrelated-Kernel Density Estimation (AKDE) as a general HR estimator (Fleming & Calabrese, 2017). Basing overlap estimation on AKDE has two primary advantages. First, AKDE corrects for bias due to autocorrelation (Fleming *et al.*, 2015a), ordinary small-sample-size bias

(Fleming & Calabrese, 2017), and temporal sampling bias (Fleming *et al.*, 2018). The net result is that AKDE HR estimates can validly be compared across studies, sites, species, and times, even when sampling strategies and underlying movement parameters differ (Fleming & Calabrese, 2017; Fleming *et al.*, 2018, Noonan et al. *under review*). Second, the error propagation techniques used to develop confidence intervals on AKDE area estimates (Fleming & Calabrese, 2017) can be extended to overlap estimation, allowing us to develop confidence intervals for overlap estimates. In addition, overlap estimates can exhibit negative bias (Fieberg & Kochanny, 2005), where part of this problem is the result of small-sample-size bias in the BC (Djouadi & Snorrason, 1990). As a solution, we derive an approximate, first order bias correction to the BC.

We use a combination of simulated and empirical data to demonstrate the power of our inferential framework. First, based on simulations, we study the bias in BC estimates as a function of the amount of autocorrelation in the data and of the effective sample size, both in cases where the underlying HR estimators account for these biases (AKDE), and where they do not (conventional KDE; Worton, 1989). We use a similar approach to quantify the realized coverage of our confidence intervals. We then show how our framework can be used to accurately estimate overlap, even when individuals exhibited different movement strategies and/or were subject to completely different sampling designs, whereas conventional methods fail. Finally, we show how our approach can be used in 'downstream' applications that depend on overlap. Specifically, we build an interaction network (Wey *et al.*, 2008) for Mongolian gazelles (*Procapra gutturosa*) where edges are established only between individuals whose overlap estimates received statistical support.

Methods

Our inferential framework consists of bias-corrected HR estimates, a bias-corrected BC estimator, and confidence intervals on the BC point estimate. We describe each of these elements in turn. We then describe how our framework can be used in practice via the ctmm R package by extending the workflow for HR analysis described in Calabrese *et al.* (2016), or through the web based graphical user interface at ctmm.shinyapps.io/ctmmweb/ (Dong *et al.*, 2017).

Home range estimation

At a minimum, calculating overlap requires a pair of HR estimates (Millspaugh *et al.*, 2004; Fieberg & Kochanny, 2005). More generally, comparisons of overlap among different groups, species, places or times may also be of interest. Nonetheless, as overlap estimates are conditional on estimated HRs, those underlying HR estimates must be directly comparable across the different groups the researcher wishes to evaluate. Unfortunately, HR estimates are subject to a number of biases, and differences in either sampling schedule, underlying movement parameters, or both can expose different datasets to different degrees of bias (Fieberg & Börger, 2012; Fleming & Calabrese, 2017). Datasets characterized by one of more of these forms of bias, which are the norm in practice, can thus render comparison of HR estimates across groups of interest highly misleading. The propagation of differentially biased HR estimates into differentially biased overlap estimates has been a key impediment to the development of a reliable inferential framework for HR overlap.

In decreasing order of importance, the three main sources of bias in HR estimation are unmodeled autocorrelation (Fleming *et al.*, 2015a), small effective sample sizes (Fleming & Calabrese, 2017), and temporally biased sampling (Fleming *et al.*, 2018). The magnitude of the negative bias in HR estimates that results from assuming the data are Independent and Identically Distributed (IID) when, in fact, they are autocorrelated can be arbitrarily large (Fleming & Calabrese, 2017). All else being equal, the bias will increase with the strength of autocorrelation in the data. In contrast, small sample size bias will be estimator-specific, and will tend to be of smaller magnitude than autocorrelation-related bias for modern GPS data. For example, KDEs based on the conventional Gaussian Reference Function (GRF) approximation tend to overestimate HR areas at small sample size (Fleming & Calabrese, 2017). Temporally biased sampling occurs when some times are oversampled while others are under-sampled (Frair *et al.*, 2010), which can produce data that are not representative of the individual's space use (Fleming *et al.*, 2018). Bias due to non-representative sampling in time will tend to increase with the degree of unevenness in the sampling schedule.

These three sources of bias must be mitigated to validly compare HR estimates, and, by extension, to validly compare overlap estimates. We now describe HR estimation methods that, when used in combination, largely corrects these biases. Autocorrelated-KDE is a generalization of the GRF-KDE (Fleming *et al.*, 2015a). The core advance in AKDE is that the optimization of the smoothing bandwidth, $\boldsymbol{\sigma}_B$, explicitly accounts for autocorrelation in the data. Specifically, an autocorrelated movement model is used to represent the autocorrelation structure of the data in the bandwidth optimization (Fleming *et al.*, 2014c, 2015b). Model selection (detailed below) can be used to arrive at an appropriate model for the data (Calabrese *et al.*, 2016). When the data exhibit no autocorrelation, the IID model would be selected, and AKDE conditional on the IID model is exactly equivalent to the well known GRF-KDE. Recently, Fleming & Calabrese (2017) derived a small-sample-size, area-based correction that mitigates the tendency of KDEs based on the GRF approximation, including AKDE, to over-smooth the data. Finally, (Fleming *et al.*, 2018) developed an optimal weighting scheme, termed 'wAKDE', that leverages the autocorrelation structure of the data to appropriately up-weight under-sampled times and down-weight over-sampled times. When used in concert, these innovations result in more accurate HR estimates that are directly comparable across groups of interest. A technical introduction to these estimators is provided in Appendix A.1.

The Bhattacharyya Coefficient (BC)

There are many different measures which quantify the relative similarity (overlap) or dissimilarity (distance) of two probability distributions. While both types of metrics can be used to describe the degree of shared space use between individuals, measures of overlap are used more commonly in biological contexts than measures of distance (but see Kranstauber *et al.*, 2016). In their comparative analysis of overlap metrics, Fieberg

& Kochanny (2005) concluded that the BC, and Volume of Intersection statistic (VI; also known as the overlap coefficient; Inman & Bradley Jr, 1989) were the most robust overlap estimators. While these two valid choices exist, we suggest that, for inferential purposes, an overlap estimator should satisfy the following criteria:

i) **Statistical validity** An appropriate overlap estimator should be based on an established measure of statistical distance or divergence that satisfies related mathematical properties.

ii) **Geometric interpretability** For uniform distributions, overlap should be proportional to the area of intersection.

iii) **Objectivity** Overlap should not depend on *ad hoc* parameters such as particular isopleths (e.g., 95% or 50%), or discretized distributions.

iv) **Computational efficiency** Computing the overlap of two distributions should scale efficiently with the sample size and extent of both distributions.

v) **Asymptotic consistency** An overlap estimator should converge to the true overlap in the large sample size limit.

vi) **Minimal bias** An overlap estimator should have good small sample size behavior.

vii) **Quantifiable uncertainty** Overlap is an estimate derived from data and should be accompanied by a measure of the confidence in that estimate (Pawitan, 2001).

The BC (Bhattacharyya, 1943) is a solid basis for inference on HR overlap because it satisfies criteria i-v, and has the additional benefit of being well known to the ecological community (Fieberg & Kochanny, 2005). Although the VI also meets these criteria (Fieberg & Kochanny, 2005), approximating confidence intervals on the VI for the case of unequal variances presents severe difficulties (Reiser & Faraggi, 1999). Consequently, we base our approach on the BC. The BC between two continuous distributions p_1 and p_2 is given by

$$\mathrm{BC}(p_1, p_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \sqrt{p_1(x,y)\, p_2(x,y)} \quad dx\, dy. \tag{1}$$

The BC is thus a function of the product of the two distributions, ranging from $0 \leq \mathrm{BC} \leq 1$, with $\mathrm{BC} = 0$ only when p_1 and p_2 have no shared support and $\mathrm{BC} = 1$ only when $p_1 = p_2$. We now turn our attention to criteria vi and vii and derive a confidence interval approximation, and bias correction that allow the BC to satisfy these additional criteria.

Confidence intervals for the BC

When measuring the overlap of two HRs, the BC, as given above, is a point estimate of the overlap between the two distributions, but does not capture any of our uncertainty in the HR estimation procedure. To address this limitation, we derive confidence intervals for the BC, in the Gaussian reference function (GRF) approximation. AKDE's first step involves fitting stochastic movement models (Fleming *et al.*, 2015a) to estimate the mean and covariance parameters

$$\boldsymbol{\mu} = \langle \mathbf{r}(t) \rangle\,, \qquad\qquad \boldsymbol{\sigma} = \mathrm{COV}[\mathbf{r}(t), \mathbf{r}(t)]\,, \tag{2}$$

where $\mathbf{r}(t) = (x(t), y(t))$ denotes the individual's location. In the GRF approximation, the individual spatial density estimates are given by

$$p(\mathbf{r}) = \frac{e^{-\frac{1}{2}(\mathbf{r}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\sigma}^{-1}(\mathbf{r}-\boldsymbol{\mu})}}{\sqrt{\det(2\pi\boldsymbol{\sigma})}}\,, \tag{3}$$

and so the BC between Gaussian density estimates resolves to

$$\mathrm{BC} = \sqrt{\det\left(\frac{\mathrm{GM}}{\mathrm{AM}}\right) e^{-\frac{1}{4}\mathrm{MD}^2}}\,, \tag{4}$$

in terms of the arithmetic and geometric means of the covariance matrices

$$\mathrm{AM} = \frac{\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2}{2}\,, \qquad \mathrm{GM} = \sqrt{\boldsymbol{\sigma}_1\,\boldsymbol{\sigma}_2}\,, \tag{5}$$

and the Mahalanobis distance (Mahalanobis, 1936) between the two distributions

$$\mathrm{MD} = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\mathrm{T}}\mathrm{AM}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}\,. \tag{6}$$

The closely related Bhattacharyya distance ($\mathrm{BD} = -\log \mathrm{BC}$; Bhattacharyya, 1946) is defined

$$\mathrm{BD} = -\log \mathrm{BC}\,, \qquad 0 \leq \mathrm{BD} < \infty\,, \tag{7}$$

which here resolves to

$$\mathrm{BD} = \frac{1}{8}\mathrm{MD}^2 + \frac{1}{2}\mathrm{tr}\log\left(\frac{\mathrm{AM}}{\mathrm{GM}}\right). \tag{8}$$

Term-by-term all components of the BD are non-negative, with the first set of terms involving the Mahalanobis distance being zero only for identical mean locations, and the second set of terms invoking the AM-GM inequality being zero only for identical covariance matrices.

First we propagate uncertainty in the mean and covariance parameters into uncertainty in $\widehat{\mathrm{BD}}$ via the delta method (Cox, 2005) to obtain $\mathrm{VAR}[\widehat{\mathrm{BD}}]$. Second, as an improvement over asymptotically normal CIs, and as the BD roughly takes the form of a square distance, we approximate the BD statistic as being chi-squared with degrees of freedom equal to

$$\mathrm{DOF} = \frac{2\,\mathrm{BD}^2}{\mathrm{VAR}[\widehat{\mathrm{BD}}]}\,, \tag{9}$$

in accord with the chi-square variance formula. We then transform the BD CIs back into BC CIs via $\mathrm{BC} = \exp(-\mathrm{BD})$. Finally, for the kernel density BC CIs, we apply the same χ^2 approximation (9), but with the AKDE point estimate for the BD and the GRF estimate for $\mathrm{VAR}[\widehat{\mathrm{BD}}]$.

Bias correction for the BC

As noted by Fieberg & Kochanny (2005), overlap is likely to be negatively biased at small sample sizes. In addition to negative biases in HR estimation driven by unmodeled autocorrelation, part of this problem is the result of small sample size bias in the BC (Djouadi & Snorrason, 1990), which is a common property of asymptotically consistent estimators (Basu, 1956). As a solution, here we derive an approximate bias correction for the BD

$$\widehat{\mathrm{BD}} = \frac{1}{8}\left(\hat{\boldsymbol{\mu}}_1-\hat{\boldsymbol{\mu}}_2\right)^{\mathrm{T}} \hat{\boldsymbol{\sigma}}^{-1} \left(\hat{\boldsymbol{\mu}}_1-\hat{\boldsymbol{\mu}}_2\right) + \frac{1}{2}\log\det\hat{\boldsymbol{\sigma}} - \frac{1}{4}\log\det\hat{\boldsymbol{\sigma}}_1 - \frac{1}{4}\log\det\hat{\boldsymbol{\sigma}}_2\,, \tag{10}$$

$$\hat{\boldsymbol{\sigma}} \equiv \frac{1}{2}\left(\hat{\boldsymbol{\sigma}}_1 + \hat{\boldsymbol{\sigma}}_2\right), \tag{11}$$

which we will also apply to the AKDE BD point estimate. Even if the two distributions are Gaussian, the BD plug-in estimator — which calculates the BD directly by assuming that the density estimates are true — is severely biased. This bias correction will be exact in the case of IID processes of equal variance, which is known to be solvable (Djouadi & Snorrason, 1990), but approximately generalized for the movement processes we consider and verified with simulation (Appendix A.2). Most of the bias is due to the fact that uncertainty in the centroids translates strictly into positive BD, even if the two distributions are identical. First we address this largest source of bias, by decomposing the mean estimates into their expectation values and (mean-zero) error

$$\hat{\boldsymbol{\mu}} = \boldsymbol{\mu} + \boldsymbol{\xi}\,, \qquad \langle\boldsymbol{\xi}\rangle = \mathbf{0}\,, \qquad \mathrm{COV}[\boldsymbol{\xi}] = \mathrm{COV}[\hat{\boldsymbol{\mu}}]\,, \tag{12}$$

whereupon we can express the first expected BD term

$$\left\langle (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^{\mathrm{T}}\, \hat{\boldsymbol{\sigma}}^{-1}\, (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) \right\rangle = \mathrm{tr}\left\langle (\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\, (\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)^{\mathrm{T}}\, \hat{\boldsymbol{\sigma}}^{-1} \right\rangle + (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\mathrm{T}}\, \langle \hat{\boldsymbol{\sigma}}^{-1} \rangle\, (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) + \cdots\,, \tag{13}$$

plus terms like $\hat{\boldsymbol{\sigma}}^{-1}\boldsymbol{\xi}$ that we ignore because $\boldsymbol{\xi}$ is mean zero and asymptotically uncorrelated with $\hat{\boldsymbol{\sigma}}$. Next we note the approximation

$$\widehat{\mathrm{COV}}[\hat{\boldsymbol{\mu}}] \propto \hat{\boldsymbol{\sigma}}\,, \tag{14}$$

which is exact for many stationary processes (e.g., Fleming *et al.*, 2014c), with a proportionality constant equal to the effective sample size of the mean. Therefore we have

$$\mathrm{tr}\left\langle (\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\, (\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)^{\mathrm{T}}\, \hat{\boldsymbol{\sigma}}^{-1} \right\rangle \approx \mathrm{tr}\big[\mathrm{COV}[\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2]\, \boldsymbol{\sigma}^{-1}\big]\,, \tag{15}$$

when the two covariances are similar, allowing us to here ignore the biases in $\hat{\boldsymbol{\sigma}}^{-1}$. We note that, in general, this term related to home-range centroid uncertainty is by far the largest source of bias in BD estimation. Furthermore, if the two movement process are independent of each other, then we have

$$\mathrm{COV}[\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2] = \mathrm{COV}[\hat{\boldsymbol{\mu}}_1] + \mathrm{COV}[\hat{\boldsymbol{\mu}}_2]\,. \tag{16}$$

For the remaining terms of the plug-in BD estimator, we require some distributional assumptions on the covariance estimates $\hat{\boldsymbol{\sigma}}_1$, $\hat{\boldsymbol{\sigma}}_2$, and $\hat{\boldsymbol{\sigma}}$. We take $\hat{\boldsymbol{\sigma}}_1$ and $\hat{\boldsymbol{\sigma}}_2$ to be Wishart distributed (Wishart, 1928) where effective sample sizes N_1 and N_2 are estimated with the parameters (Fleming & Calabrese, 2017). For the average covariance $\hat{\boldsymbol{\sigma}}$, we construct a Welch-Satterthwaite (Satterthwaite, 1946) like approximation that is exact for equal covariances. If $\hat{\boldsymbol{\sigma}}$ were χ^2 distributed, the ordinary Welch-Satterthwaite approximation would fix its degrees of freedom via the relationship between its variance and that of its constituents. However, $\hat{\boldsymbol{\sigma}}$ is matrix valued and has many variances. We choose to conserve the trace variance, which is both additive and rotationally invariant:

$$\operatorname{tr}\mathrm{VAR}[\hat{\boldsymbol{\sigma}}] = \frac{1}{4}\operatorname{tr}\mathrm{VAR}[\hat{\boldsymbol{\sigma}}_1] + \frac{1}{4}\operatorname{tr}\mathrm{VAR}[\hat{\boldsymbol{\sigma}}_2]\,, \tag{17}$$

$$\frac{\operatorname{tr}\operatorname{diag}(\hat{\boldsymbol{\sigma}})^2}{N} = \frac{\operatorname{tr}\operatorname{diag}(\hat{\boldsymbol{\sigma}}_1)^2}{4\,N_1} + \frac{\operatorname{tr}\operatorname{diag}(\hat{\boldsymbol{\sigma}}_2)^2}{4N_2}\,, \tag{18}$$

$$N = \frac{4\operatorname{tr}\operatorname{diag}(\hat{\boldsymbol{\sigma}})^2}{\frac{\operatorname{tr}\operatorname{diag}(\hat{\boldsymbol{\sigma}}_1)^2}{N_1} + \frac{\operatorname{tr}\operatorname{diag}(\hat{\boldsymbol{\sigma}}_2)^2}{N_2}}\,. \tag{19}$$

Next the expected inverse estimate matrix resolves to

$$\langle \hat{\boldsymbol{\sigma}}^{-1} \rangle = \frac{N}{N - \dim(\boldsymbol{\sigma}) - 1}\boldsymbol{\sigma}\,, \tag{20}$$

and so we clamp our effective sample size estimates to $N \geq \dim(\boldsymbol{\sigma}) + 2$, which is the smallest discrete number of IID locations with which one can estimate properly. Below this value the estimate is likely not approximately Wishart distributed and N is likely not well estimated. So by clamping N we effectively clamp our bias correction. Next, the expected log determinant terms resolve to

$$\langle \log\det\hat{\boldsymbol{\sigma}} \rangle = \log\det\boldsymbol{\sigma} + \psi_{\dim(\boldsymbol{\sigma})}(N/2) - \log(N/2)^{\dim(\boldsymbol{\sigma})}\,, \tag{21}$$

in terms of the multivariate digamma function ψ_d.

Finally, as $\mathrm{BD} \geq 0$, we debias the plug-in estimator by dividing by a large number rather than by subtracting a large number:

$$\hat{\theta} \to \left(\frac{\hat{\theta}}{\hat{\theta} + \widehat{\mathrm{BIAS}}[\hat{\theta}]} \right) \hat{\theta} = \hat{\theta} - \widehat{\mathrm{BIAS}}[\hat{\theta}] + \mathcal{O}\big(N^{-2}\big)\,, \tag{22}$$

which is the same to first order. This serves as a first order bias correction to both the BD and the BC.

Workflow

The resulting centerpiece of our inferential framework is a bias corrected BC estimate, with confidence intervals, that is comparable across studies. To get to that point, the user must first proceed through a workflow designed to produce the best possible estimates from their data, but warn when such an analysis is inappropriate. This workflow builds on that described in Calabrese *et al.* (2016) for HR analysis.

The first step is ensuring that the data at hand are appropriate for HR analysis, which means that there must be clear evidence of range-residency. Data from non-range-resident individuals, or from range-resident intervals that were only briefly tracked may not satisfy this criterion. When the data do not show evidence of range-residency, HR estimation is not appropriate (Calabrese *et al.*, 2016; Fleming & Calabrese, 2017), which implies that HR overlap analysis is also not appropriate. We therefore strongly recommend starting with visual verification of range-residency via variogram analysis (Fleming *et al.*, 2014b). Specifically, the variogram of a range-resident individual should show a clear asymptote.

Once range-residency has been verified, the next step is to fit a series of range-resident movement models to the data, such as the IID, Ornstein-Uhlenbeck (OU; Uhlenbeck & Ornstein, 1930), and OU-Foraging (OUF; Fleming *et al.*, 2014b,c) processes. Model selection should then be employed to identify the best model for the data (Fleming *et al.*, 2014c, 2015b). The selected model should then be visually compared to the variogram to ensure that the model is capturing the key features in the data. Models that fail to converge, or that do not provide a reasonable fit to the data are another indication that HR analysis may be inappropriate (Calabrese *et al.*, 2016).

With a fitted, selected movement model in hand, AKDE HR estimates can then be calculated, and these can be used to obtain BC estimates and CIs. These overlap estimates may either be the final product of the analysis, or be used in subsequent analyses. Importantly, the confidence intervals attached to each BC estimate can be straightforwardly propagated into derived quantities, such as the mean overlap within a group, which can facilitate testing hypotheses on similarity or differences among groups of interest. While the workflow we describe involves several steps, the `ctmm` package, and graphical user interface (Dong *et al.*, 2017) streamline this procedure. A full example of the workflow is shown in Appendix B.

Simulation study

To examine the statistical properties of the BC, and the coverage of our CIs, we simulated tracking data with variable sampling durations and frequencies. Data were simulated based on pairs of both IID processes, and OUF processes (Fleming *et al.*, 2014b,c), parameterized such that the true overlap between these pairs was fixed at 0.5. Simulating from an OUF process generates relocations that feature autocorrelated positions and velocities, as well as restricted space use, and are representative of modern GPS tracking data commonly used in HR analyses (Fleming & Calabrese, 2017).

Importantly, the timescale over which autocorrelation in position decays, τ_p (also termed the HR crossing time; Calabrese *et al.*, 2016), is a key parameter for HR estimation (Noonan et al. *under review*). Formally, τ_p can be quantified from the data as the timescale over which an individual's positional autocorrelation decays by a factor of $\frac{1}{e}$, and its movement process reverts to the mean location (Fleming *et al.*, 2015a; Fleming & Calabrese, 2017). The duration of the observation period (T), in relation to τ_p, will thus dictate the effective sample size (n_e) of a dataset via

$$n_e \approx \frac{T}{\tau_p}, \tag{23}$$

which may be interpreted as the approximate number of range crossings that occurred during the sampling period. We tailored our simulations according to their relative effects on n_e. These were:

i) **Sampling duration**. Observations were recorded eight times/day, and we manipulated sampling duration (ranging from 1 to 4096 days in a doubling series). For OUF simulations, the HR crossing time was set to one day, and the velocity autocorrelation timescale to 1/5 of a day. Notably, this parameterization was such that in these simulations the sampling duration in days exhibited a 1:1 relationship with n_{e}.

ii) **Sampling frequency**. Here, the sampling duration was fixed at 32 days, and we manipulated the sampling frequency (ranging from 1 to 1024 fixes/day in a doubling series). Again, for the OUF process HR crossing time was set to one day, and the velocity autocorrelation timescale to 1/5 of a day. The fixed sampling duration in these simulations resulted in n_{e} being fixed at 32, irrespective of variation in the sampling frequency.

We then compared the accuracy of the underlying HR estimates, the accuracy of the estimated overlap, and the realized coverage of the confidence intervals. Results were averaged over 1000 simulations per manipulation. The computations were conducted on the Smithsonian Institution High Performance Cluster (SI/HPC).

Empirical study

We demonstrate the functionality of this method using GPS data from Mongolian gazelles. Mongolian gazelles are medium sized herbivores that cross their ranges on seasonal timescales (Fleming *et al.*, 2014c,b). Positional data for 36 Mongolian gazelle were collected in Mongolia's Eastern Steppe between 2007 and 2011 (Fleming *et al.*, 2014a). Both variogram analysis (Fleming *et al.*, 2014c) and model selection (Calabrese *et al.*, 2016) were used to confirm that there was evidence of range-residency in the data. From these diagnostic checks, 13 individuals showed no signs of range-resident behavior, and we restricted our analyses to the 23 range-resident individuals. HR estimation was then carried out using KDE and AKDE as described above. We then computed all pairwise BCs $\pm$ 95% CIs on the KDE and AKDE estimates. Notably, the long HR crossing timescales ($\bar{x} = 111.5$ days; range = 8.0 – 443.2), and comparatively short tracking durations ($\bar{x} = 381.0$ days; range = 67.2 – 755.0), here produced a mean n_{e} of 6.1 (ranging from 0.7 – 24.6). This is a regime where the negative bias of conventional KDE is known to have serious implications for HR estimates on autocorrelated data

(Fleming & Calabrese, 2017).

Downstream analyses

To further highlight the utility of these confidence intervals, we used the estimated overlap to quantify the edges of a spatial interaction network (Wey *et al.*, 2008). Because point estimates were accompanied by CIs, we were able to subset edges into two categories:

i) **Supported**. Well supported edges were identified as cases where two individuals exhibited overlapping space use, with a minimum CI that was greater than 0.01 – i.e., there was a 95% certainty that the overlap was ≥ 0.01

ii) **Unsupported**. Unsupported edges were identified as cases where the point estimate suggested overlapping space use, but with a minimum CI that was less than 0.01 – i.e., there was insufficient evidence to be certain that the overlap differed significantly from 0.

We then quantified a number of commonly used diagnostics (i.e., network density, mean path length, and closeness centrality; Wey *et al.*, 2008), to investigate how these might differ when the network was based only on statistically supported edges, versus the inclusion of unsupported edges.

All analyses were conducted in the `R` environment (R Core Team, 2016), using the methods implemented in the package `ctmm` (Calabrese *et al.*, 2016).

Results

Simulation results

Asymptotic properties of the BC

Simulations revealed that for IID data, both AKDE and KDE HR estimates provided identical results, and were relatively unbiased except at very small sample sizes (Fig. 1a). The resulting overlap was also identical between estimators, and increasing the number of fixes, by either increasing the sampling duration (Fig. 1b) or frequency (Fig.

1e), had the expected effect of increasing the accuracy of the overlap estimate and decreasing the uncertainty. Notably, the CIs on the BC offered reasonable coverage of the true overlap across all sampling regimes, albeit with some persistent negative bias at large sample sizes (Fig. 1c,f). This was the result of bias in the BC decaying too slowly relative to the variance (see Appendix A.3).

For autocorrelated data in contrast, AKDE 95% HR estimates were generally accurate across the range of sample durations (Fig. 2a), and frequencies (Fig. 2d) we simulated, whereas KDE HR estimates were severely biased for all but the largest datasets. As a result, while the estimated overlap between AKDE and KDE estimates both converged to the truth as sampling duration increased (Fig. 2b), asymptotic consistency for KDE estimates was severely delayed. Furthermore, increasing the sampling frequency increased the negative bias in overlap estimates derived from KDE, but, appropriately, did not influence overlap estimates based on AKDE (Fig. 2e).

The coverage of 95% CIs for the KDE derived overlap estimates was severely biased under all of the scenarios we tested (Fig. 2c, f). In contrast, the coverage of CIs on the AKDE estimates consistently provided close to nominal coverage of the true overlap.

Comparability of estimates

Our baseline simulation study controlled the effect of the movement parameters by assuming the individuals exhibited identical movement strategies, and were sampled at the exact same times. Under these conditions, the improved accuracy of AKDE HRs estimates resulted in more accurate overlap estimates, with 95% CIs that provided close to nominal coverage (Fig. 3a). There are realistic complications to our basic simulation strategy, however, including cases where individuals are subject to the same sampling design, but exhibit different movement strategies, and cases where both movement strategies and sampling designs differ. Importantly we found that AKDE based overlap still provided reasonable coverage for both of these cases (Fig. 3c,e). In contrast, because of the differential bias in KDE HR estimates, the estimated overlap differed substantially between each of these scenarios, and in every case failed to provide coverage of the true value (Fig. 3b,d,f).

Empirical case study

Consistent with our simulated findings of negative bias in KDE HR and BC estimates at mid to low n_e on autocorrelated data, empirical AKDE HR estimates were larger than KDE estimates for all pairs (Fig. 4a). Median pairwise overlap between the 276 pairs of individuals was 0.66 (95% CI 0.58 – 0.76) when the overlap was estimated from AKDE HR estimates, but five-fold lower when estimated from KDE estimates (median = 0.13; 95% CI 0.06 – 0.22).

The severe negative bias of KDE derived overlap was persistent across all individuals. This can be illustrated in a specific example, where the KDE HR estimates resulted in an estimated overlap of 0.02 (95% CI 0.01 – 0.03), whereas the AKDE HRs resulted in an overlap of 0.80 (95% CI 0.22 – 0.99). Visual inspection of the range estimates for these individuals revealed substantial negative bias in the KDE HR, whereas the AKDE HR was larger, with appropriately wide CIs considering the small n_e of ~ 4 for each HR estimate (Fig. 4 b–c).

Downstream analyses

Because these overlap estimates were accompanied by confidence intervals, the uncertainty can be used to inform downstream analyses. For instance, a spatial network analysis based on the estimated overlap revealed 461 edges of variable strength (Fig. 5). Of these, 275 were well supported, whereas 186 had no statistical support. We found that basing the network off of all possible edges, versus only those edges with statistical support, influenced its properties and any potential biological inferences that would be derived from it. For instance, network density was reduced from 0.86 to 0.63 when the analysis was restricted to only the well supported edges. Furthermore, only utilizing statistically supported edges increased the mean path length from 1.13 to 1.39. Interestingly, despite decreasing density and increasing the mean path length, constructing the network based on only well supported edges resulted in a two-fold increase in the closeness centrality compared to the network constructed with both supported and unsupported edges (0.45 vs. 0.23 respectively).

Discussion

Despite the routine nature of estimating overlapping space use (e.g., Berger & Gese, 2007; Frère *et al.*, 2010; Sanchez & Hudgens, 2015; Dougherty *et al.*, 2018), there exists no formal inferential framework for this analysis. This is largely due to the inherent difficulties associated with HR estimation (Fieberg & Börger, 2012) and exacerbated by the historical lack of CIs on both HR, and overlap estimates. As a solution, we have demonstrated how AKDE HR estimates (Fleming *et al.*, 2015a; Fleming & Calabrese, 2017) can serve as a reliable foundation on which to base statistical inference. In addition, we have implemented a small-sample-size bias correction for the BC and derived well-behaved, approximate CIs on the point estimate. Collectively, these advances permit researchers to accurately quantify HR overlap, even when sampling strategies and underlying movement parameters differ among groups being compared, and test whether any observed differences are statistically meaningful.

Home range and overlap estimation: an intrinsic relationship

A crucial component of any statistical inference is having comparable measures on which to base analyses. Overlap is typically conditional on HR estimates (Millspaugh *et al.*, 2004; Fieberg & Kochanny, 2005), which are themselves estimated from animal tracking data. Because overlap estimation relies on at least three separate estimates (two HR estimates, and their overlap), it follows that this analysis is particularly vulnerable to issues of estimator bias. Accurate HR estimation is a deceptively challenging problem however, as autocorrelation (Fleming *et al.*, 2015a), small-sample-size bias (Fleming & Calabrese, 2017), and sampling irregularities (Frair *et al.*, 2010; Fleming *et al.*, 2018) will significantly influence any statistical analyses applied to animal tracking data. More subtly, even identical sampling strategies can still produce differentially biased HR estimates if the underlying parameters of movement differ markedly between individuals (Fleming & Calabrese, 2017, Noonan et al. *under review*). As these are nearly ubiquitous aspects of animal tracking data, accurate overlap estimation requires statistical methods that can handle these complications, without introducing artifactual differences due purely to estimator bias.

In this respect, our simulation study revealed that, for autocorrelated data, KDE regularly underestimated HR sizes (Fleming & Calabrese, 2017, Noonan et al. *under review*), and this negative bias was directly propagated to overlap estimates. For KDE, the amount of data required to achieve an accurate measure of overlap was very large, and most empirical cases are likely to underestimate the true overlap (Fieberg & Kochanny, 2005). In contrast, AKDE HRs were larger, but significantly more accurate, which translated to more accurate overlap estimates. Crucially, when we varied the sampling design and movement strategies between the individuals we were comparing, AKDE based estimates provided reliable coverage of the true overlap, whereas this was not the case for KDE. Consistent with the results of our simulation study, empirical AKDE HRs from autocorrelated Mongolian gazelle GPS data were ca. twice as large as KDE estimates. This resulted in the median pairwise overlap being five-fold larger when based on AKDE versus KDE. Had an analysis been based on the biased KDE estimates, one would have erroneously concluded that there was little spatial overlap in this system, whereas, results based on AKDE's more rigorous estimates revealed these individuals actually exhibited extensive overlap. Although these empirical estimates could not be compared to a truth, as per our simulations, this finding is also consistent with a recent analysis by Noonan et al. (*under review*). In a large scale comparative study encompassing 369 individuals across 30 species, they found that AKDE 95% HR estimates consistently included ∼95% of holdout observations, whereas KDE estimates included ∼92% at high n_e (> 256), but only ∼75% at low n_e. This means AKDE's larger estimates are accurate, while those produced by conventional KDE on the same data are consistently, and often grossly, too small. The net result is that AKDE provides a solid foundation for estimating overlap under realistic sampling regimes, resulting in accurate overlap estimates that can validly be compared across studies.

As described above, a fundamental component of estimating HR overlap is having comparable measures on which to base analyses. Notably, in this study, we consider *range* estimators in the sense of Burt (1943), which estimate long-run space use, assuming the focal individual does not change its movement process (Fleming *et al.*, 2015a). This includes KDEs, Minimum Convex Polygons (MCP; Mohr, 1947), and time-naive Local Convex Hulls (LoCoH) (Getz *et al.*, 2007). Also of interest are *occurrence* distribution estimators such as the Brownian bridge (Horne *et al.*, 2007), or t-LoCoH (Lyons *et al.*, 2013) which quantify uncertainty in the animal's location during the sampling period, including times not sampled. Crucially, this uncertainty vanishes in the limit where both the sampling interval and telemetry error approach zero. Although these two mathematically distinct classes of distributions have been historically conflated under the umbrella term of "utilization distributions", they have very different interpretations and use cases (Fleming *et al.*, 2015a). Consequently, overlap based on *occurrence* estimates have very different meanings from overlap based on *range* estimates, and are beyond the scope of the present work.

We also note that extending our bias-correction and CIs to other HR estimators, such as MCP, LoCoH, or non-GRF KDE bandwidth optimizers, is not a tractable problem. First, our methods are explicitly based on the GRF approximation, so they are not consistent with non-GRF estimators. Second, the GRF-based methods implemented in `ctmm` are, to our knowledge, the only HR estimators that quantify uncertainty. As an uncertainty estimate is a prerequisite for our error propagation techniques, it would not currently be possible to adapt our approach to other estimators. Finally, the target distributions and expectation values of geometric methods such as MCP and LoCoH are usually unknown, which makes these estimators incompatible with the methods developed here.

Properties of the overlap estimator

In addition to utilizing reliable HR estimates, the overlap estimator itself should have desirable properties (Fieberg & Kochanny, 2005). While several valid estimators exist, the BC (Bhattacharyya, 1943) stands out because of its statistical validity, geometric interpretability, computational efficiency, and asymptotic consistency. As noted by Fieberg & Kochanny (2005) however, the BC is prone to exhibiting negative, small-sample-size bias (Djouadi & Snorrason, 1990). To correct for this, we derived a small-sample-size bias correction, which improved the accuracy of BC estimates (see also Djouadi & Snorrason, 1990).

Also problematic is the historical lack of CIs on overlap estimates. Overlap is an estimate derived from data and should be accompanied by a measure of the uncertainty (Pawitan, 2001). Without this, one cannot properly infer the importance of a given estimate. As a solution, we have derived CIs on the BC based on a GRF approximation. Using simulated data, we demonstrated how this implementation will provide reasonable coverage of the true overlap. We note, however, that, while generally well behaved, there was some persistent negative bias in the coverage of these CIs. The biased coverage is likely the result of the bias in the BC point estimate decaying too slowly relative to the variance as $\mathrm{n_e}$ increased (Fig. A.2). With asymptotically efficient estimators, this ratio would decay at a rate of $1/\sqrt{N}$ or better, whereas here it increases at a rate of $\sim \sqrt{N}$. As such, their coverage should be treated with caution, particularly at large $\mathrm{n_e}$. Furthermore, because we approximate the HRs as Gaussian when estimating uncertainty, the CIs may exhibit unintended behavior when the overlap is dependent on non-Gaussian features.

Despite these limitations, well-behaved CIs for HR overlap is a novel feature, and permits true statistical inference on overlap estimates. For instance, these CIs can be applied to a reference value of interest (e.g., the mean overlap between individuals of the same species studied elsewhere) to test for significant differences between these, as opposed to relying on *ad hoc* comparisons. Additionally, if overlap is being used to inform subsequent analyses, CIs can be used to improve these. For example, we found that differentiating between the 275 overlap estimates that were well supported by the data and the 186 that may have been artifactual significantly influenced the properties of an interaction network of Mongolian gazelle. When based on all possible edges, the network suggested a larger number of edges, but with a low closeness centrality. Conversely, when based only on edges with statistical support, the network density decreased, but closeness increased. The supported, and unsupported, networks would each lead to a unique set of biological interpretations, with only the former being supported by the data.

Conclusion

In conclusion, we have developed the first inferential framework for HR overlap tailored for the specific needs of ecologists that is both statistically valid and computationally efficient. Collectively, the more accurate and comparable HR estimates provided by AKDE (Fleming *et al.*, 2015a; Fleming & Calabrese, 2017, Noonan et al. *under review*) and our novel bias correction and CIs on the BC permit rigorous overlap estimation. This method is now available via command line interface through the `ctmm` package (Calabrese *et al.*, 2016), or through the web based graphical user interface at ctmm.shinyapps.io/ctmmweb/ (Dong *et al.*, 2017).

Acknowledgments

This work was supported by the US NSF Advances in Biological Informatics program (ABI-1458748 to JMC). MJN was supported by a Smithsonian Institution CGPS grant. TM was funded by the Robert Bosch Foundation.

Data Accessibility

The Mongolian gazelle data used in this manuscript are available from the Dryad online repository (Fleming *et al.*, 2014a, DOI: 10.5061/dryad.45157).

References

Basu, D. (1956) The Concept of Asymptotic Efficiency. *Sankhyā: the indian journal of statistics*, **17**, 193–196.

Berger, K.M. & Gese, E.M. (2007) Does interference competition with wolves limit the distribution and abundance of coyotes? *The Journal of Animal Ecology*, **76**, 1075–1085.

Bhattacharyya, A. (1943) On a measure of divergence between two statistical populations defined by their probability distribution. *Bulletin of the Calcutta Mathematical Society*, **35**, 99–109.

Bhattacharyya, A. (1946) On a measure of divergence between two multinomial populations. *Sankhyā: the indian journal of statistics*, **7**, 401–406.

Brown, J.L. & Orians, G.H. (1970) Spacing Patterns in Mobile Animals. *Annual review of Ecology and Systematics*, **1**, 239–262.

Burt, W.H. (1943) Territoriality and Home Range Concepts as Applied to Mammals. *Journal of Mammalogy*, **24**, 346–352.

Calabrese, J.M., Fleming, C.H. & Gurarie, E. (2016) ctmm: an r package for analyzing animal relocation data as a continuous-time stochastic process. *Methods in Ecology and Evolution*, **7**, 1124–1132.

Cox, C. (2005) Delta Method. *Encyclopedia of Biostatistics*, p. 41. John Wiley & Sons, Ltd, Chichester, UK.

Djouadi, A. & Snorrason, O. (1990) The quality of training sample estimates of the bhattacharyya coefficient. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**.

Dong, X., Fleming, C. & Calabrese, J. (2017) *`ctmm` `webapp`: A graphical user interface for the `ctmm` R package.*

Dougherty, E.R., Seidel, D.P., Carlson, C.J., Spiegel, O. & Getz, W.M. (2018) Going through the motions: incorporating movement analyses into disease research. *Ecology letters*, **21**, 588–604.

Fieberg, J. & Börger, L. (2012) Could you please phrase "home range" as a question? *Journal of Mammalogy*, **93**, 890–902.

Fieberg, J. & Kochanny, C.O. (2005) Quantifying home-range overlap: the importance of the utilization distribution. *Journal of Wildlife Management*, **69**, 1346–1359.

Fleming, C.H., Fagan, W.F., Mueller, T., Olson, K.A., Leimgruber, P. & Calabrese, J.M. (2015a) Rigorous home range estimation with movement data: a new autocorrelated kernel density estimator. *Ecology*, **96**, 1182–1188.

Fleming, C.H., Sheldon, D., Fagan, W.F., Leimgruber, P., Mueller, T., Nandintsetseg, D., Noonan, M.J., Olson, K.A., Setyawan, E., Sianipar, A. & Calabrese, J.M. (2018) Correcting for missing and irregular data in home-range estimation. *Ecological Applications*.

Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2014a) Data from: From fine-scale foraging to home ranges: a semi-variance approach to identifying movement modes across spatiotemporal scales. *Dryad Digital Repository*. http://dx.doi.org/10.5061/dryad.45157.

Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2014b) From Fine-Scale Foraging to Home Ranges: A Semivariance Approach to Identifying Movement Modes across Spatiotemporal Scales. *The American Naturalist*, **183**, E154–E167.

Fleming, C.H., Subaşı, Y. & Calabrese, J.M. (2015b) Maximum-entropy description of animal movement. *Physical review E, Statistical, nonlinear, and soft matter physics*, **91**, 032107.

Fleming, C.H. & Calabrese, J.M. (2017) A new kernel density estimator for accurate home-range and species-range area estimation. *Methods in Ecology and Evolution*, **8**, 571–579.

Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2014c) Non-Markovian maximum likelihood estimation of autocorrelated movement processes. *Methods in Ecology and Evolution*, **5**, 462–472.

Frair, J.L., Fieberg, J., Hebblewhite, M., Cagnacci, F., DeCesare, N.J. & Pedrotti, L. (2010) Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **365**, 2187–2200.

Frère, C.H., Krützen, M., Mann, J., Watson-Capps, J.J., Tsai, Y.J., Patterson, E.M., Connor, R., Bejder, L. & Sherwin, W.B. (2010) Home range overlap, matrilineal and biparental kinship drive female associations in bottlenose dolphins. *Animal Behaviour*, **80**, 481–486.

Getz, W.M., Fortmann-Roe, S., Cross, P.C., Lyons, A.J., Ryan, S.J. & Wilmers, C.C. (2007) LoCoH: Nonparameteric Kernel Methods for Constructing Home Ranges and Utilization Distributions. *PLoS ONE*, **2**, e207.

Grant, J.W.A., Chapman, C.A. & Richardson, K.S. (1992) Defended versus undefended home range size of carnivores, ungulates and primates. *Behavioral Ecology and Sociobiology*, **31**, 149–161.

Horne, J.S., Garton, E.O., Krone, S.M. & Lewis, J.S. (2007) Analyzing animal movements using Brownian bridges. *Ecology*, **88**, 2354–2363.

Inman, H.F. & Bradley Jr, E.L. (1989) The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Communications in Statistics-Theory and Methods*, **18**, 3851–3874.

Jetz, W. (2004) The Scaling of Animal Space Use. *Science*, **306**, 266–268.

Kays, R., Crofoot, M.C., Jetz, W. & Wikelski, M. (2015) Terrestrial animal tracking as an eye on life and planet. *Science*, **348**, aaa2478–aaa2478.

Kranstauber, B., Smolla, M. & Safi, K. (2016) Similarity in spatial utilization distributions measured by the earth mover's distance. *Methods in Ecology and Evolution*, **8**, 155–160.

Lukas, D. & Clutton-Brock, T.H. (2013) The Evolution of Social Monogamy in Mammals. *Science*, **341**, 526–530.

Lyons, A.J., Turner, W.C. & Getz, W.M. (2013) Home range plus: a space-time characterization of movement over real landscapes. *Movement Ecology*, **1**, 2.

Mahalanobis, P.C. (1936) On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, **2**, 49–55.

Millspaugh, J.J., Gitzen, R.A., Kernohan, B.J., Larson, M.A. & Clay, C.L. (2004) Comparability of three analytical techniques to assess joint space use. *Wildlife Society Bulletin*, **32**, 148–157.

Mitchell, M.S. & Powell, R.A. (2012) Foraging optimally for home ranges. *Journal of Mammalogy*, **93**, 917–928.

Mitchell, W.A. & Lima, S.L. (2002) Predator-prey shell games: large-scale movement and its implications for decision-making by prey. *Oikos*, **99**, 249–259.

Mohr, C.O. (1947) Table of Equivalent Populations of North American Small Mammals. *American Midland Naturalist*, **37**, 223.

Pawitan, Y. (2001) *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Clarendon Press, Oxford.

Powell, R.A. (1979) Mustelid Spacing Patterns: Variations on a Theme by Mustela. *Ethology*, **50**, 153–165.

R Core Team (2016) R: A language and environment for statistical computing. *Vienna, Austria: R foundation for statistical computing.*

Reiser, B. & Faraggi, D. (1999) Confidence Intervals for the Overlapping Coefficient: the Normal Equal Variance Case. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **48**, 413–418.

Sanchez, J.N. & Hudgens, B.R. (2015) Interactions between density, home range behaviors, and contact rates in the Channel Island fox (*Urocyon littoralis*). *Ecology and evolution*, **5**, 2466–2477.

Satterthwaite, F.E. (1946) An approximate distribution of estimates of variance components. *Biometrics Bulletin*, **2**, 110–114.

Uhlenbeck, G.E. & Ornstein, L.S. (1930) On the Theory of the Brownian Motion. *Physical review*, **36**, 823–841.

Wey, T., Blumstein, D.T., Shen, W. & Jordán, F. (2008) Social network analysis of animal behaviour: a promising tool for the study of sociality. *Animal Behaviour*, **75**, 333–344.

Wishart, J. (1928) The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, **20A**, 32–52.

Worton, B.J. (1989) Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. *Ecology*, **70**, 164–168.

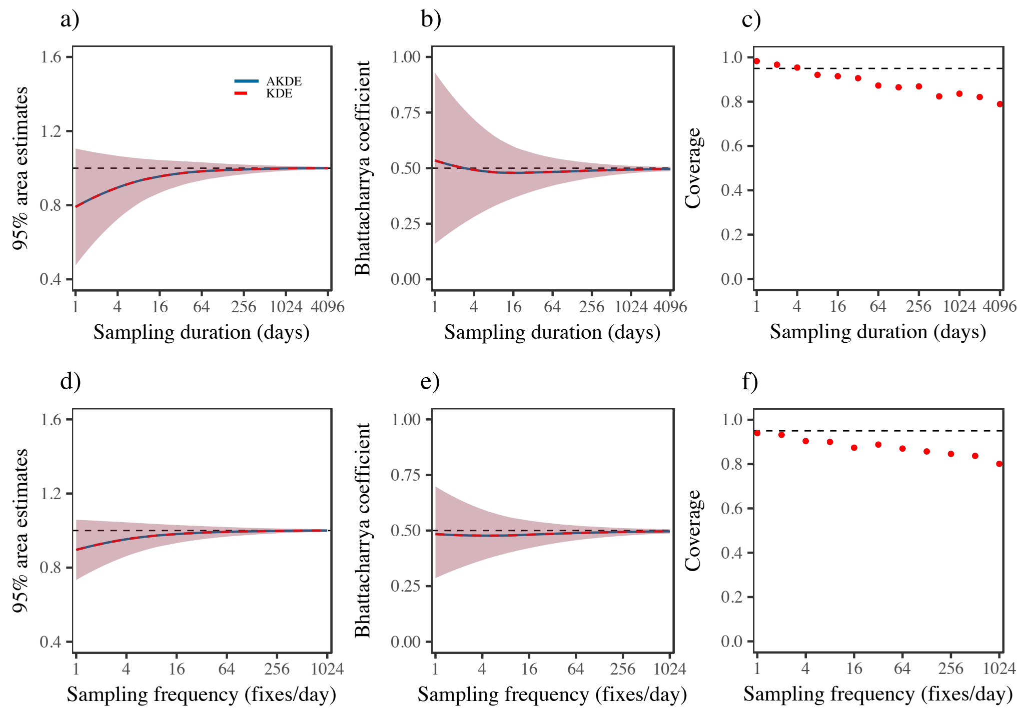

Figure 1: The asymptotic properties of KDE and AKDE HR estimators (panels a and d), and the BC (panels b and e) for simulated, IID data, as well as the coverage of the CIs (panels c and f), as a function of sampling duration (top row), and frequency (bottom row). In all panels the dashed horizontal lines depict the truth, the solid line the mean point estimate, and the shaded regions the 95% CIs.

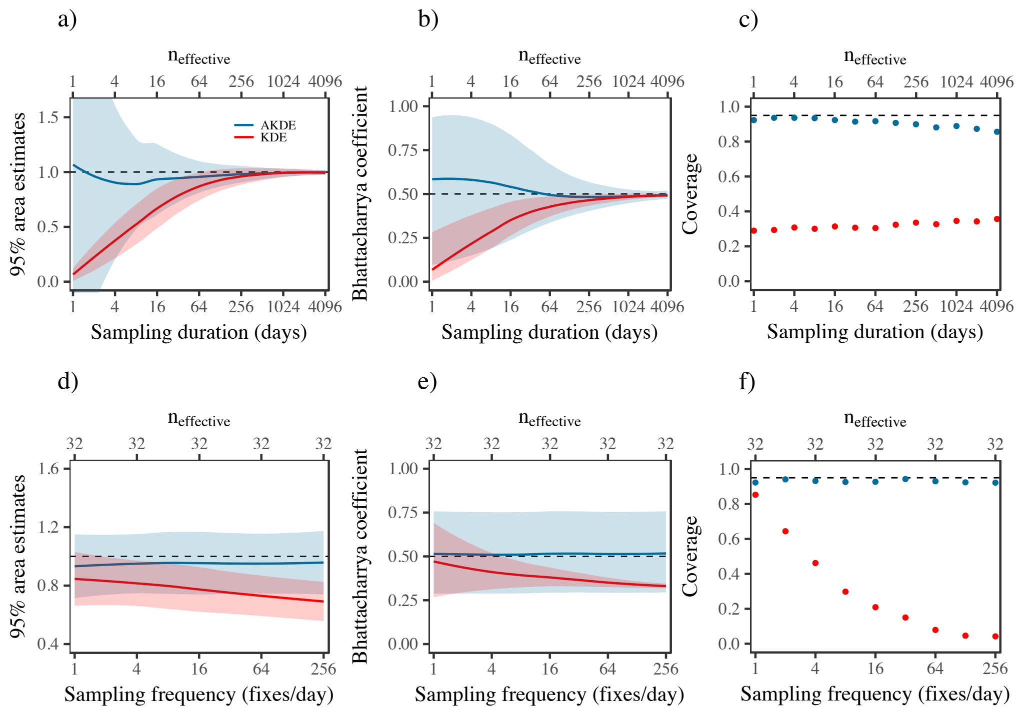

Figure 2: The asymptotic properties of KDE and AKDE HR estimators (panels a and d), and the BC (panels b and e) for simulated, autocorrelated tracking data, and the coverage of the CIs (panels c and f), as a function of sampling duration (top row), and frequency (bottom row). In all panels the dashed horizontal lines depict the truth, the solid line the mean point estimate, and the shaded regions the 95% CIs. Notably, convergence to the truth was much slower for KDE, and the coverage of KDE's CIs was far from appropriate in all cases.

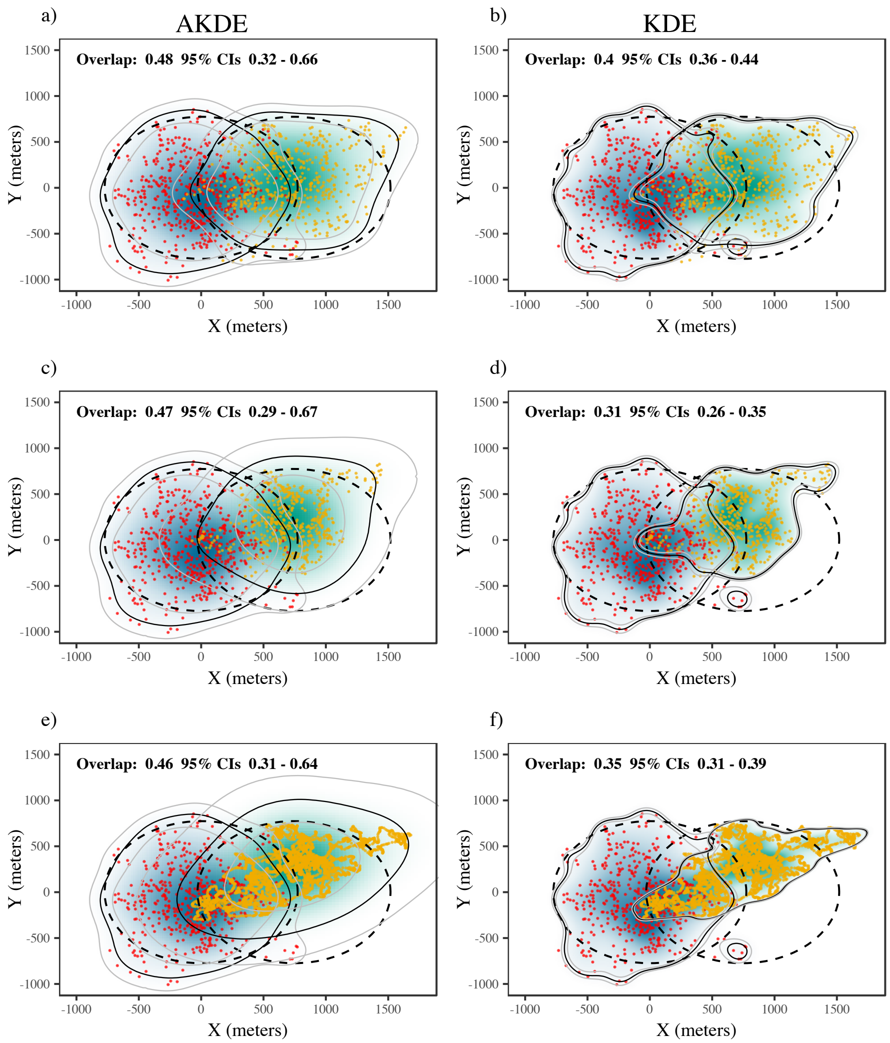

Figure 3: HR and overlap estimates for two simulated individuals with a true overlap of 0.50. In all panels, the dashed circles depict the true 95% areas, the solid black lines the estimated 95% areas, and the grey lines the 95% CIs on the area estimates. In the first row, relocations were simulated from OUF models with identical movement parameters and sampling times. In the second row, sampling was held consistent, but the individual plotted in yellow had a HR crossing time of 1 week versus 1 day for the individual in red. In the third row, movement again differed between individuals, but here the individual in yellow was sampled once every 30min, versus once every 3hrs for the individual in red. Note how in all cases AKDE based overlap estimates were relatively consistent, and provided coverage of the true overlap, whereas KDE based overlap estimates varied substantially, and consistently failed to provide coverage of the truth.

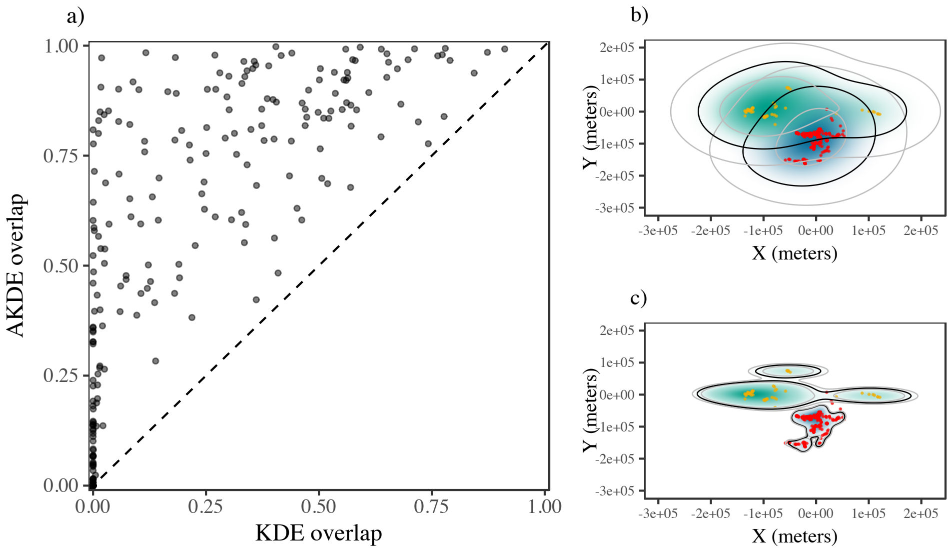

Figure 4: Panel a) depicts the relationship between pairwise estimates of the BC for Mongolian gazelle, computed from KDE and AKDE HR estimates. The dashed 1:1 line depicts parity between these. Note how all cases fall above this line, highlighting how AKDE derived BC suggests more overlap than KDE derived BC. An example of this discrepancy is depicted in panel b), with AKDE BC suggesting extensive overlap 0.80 (0.22 – 0.99), whereas in c) the negative bias in KDE propagates to produce a biased estimate of the overlap 0.02 (0.01 – 0.03). Crucially, with effective sample sizes of ca. 4 for each HR estimate, the CIs approximated from the AKDE estimates were appropriately wide, versus KDE's deceivingly narrow CIs.

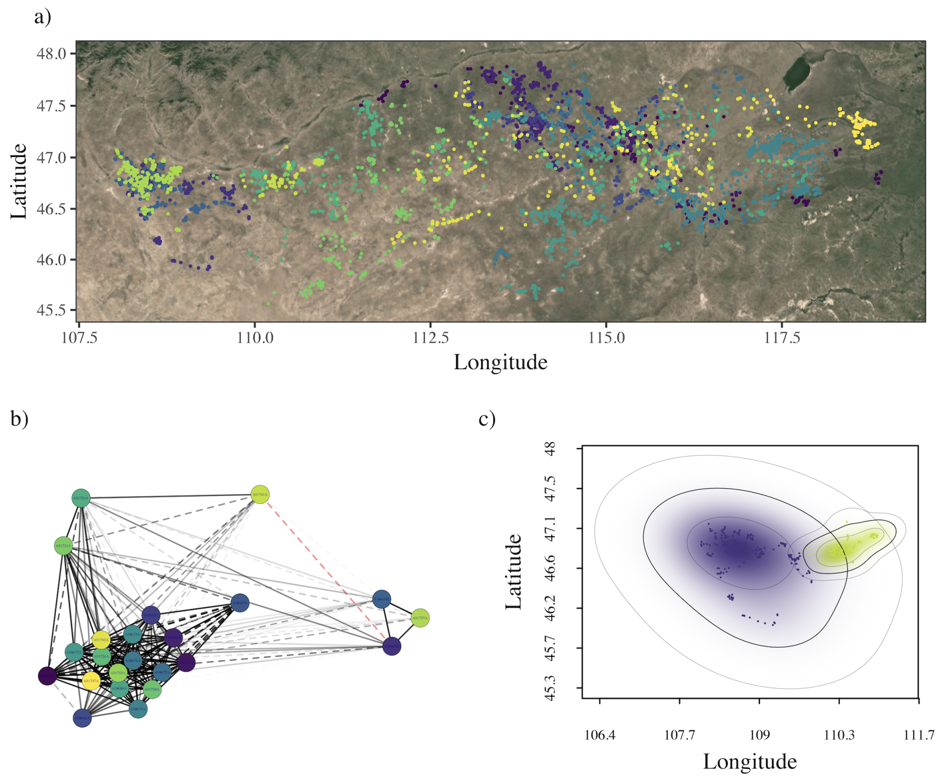

Figure 5: Figure depicting a) the GPS locations for 23 Mongolian gazelle tracked in Mongolia's Eastern Steppe; b) a network diagram with edge weights based on overlap values; and c) an example case of two HR estimates where the point estimate of the overlap suggests a connection, but the CIs on the estimates suggest that connection might not be statistically significant. The dashed lines in b) depict pairs where the point estimate suggests a connection, but with CIs that include 0.01 and thus may not be statistically significant. The transparency of the lines is proportional to the point estimate of the BC. The connection depicted in red on the right-hand side of panel b) corresponds to the pair in panel c).