

# Non-Markovian maximum likelihood estimation of autocorrelated movement processes

Christen H. Fleming<sup>1,2\*</sup>, Justin M. Calabrese<sup>1</sup>, Thomas Mueller<sup>1,2</sup>, Kirk A. Olson<sup>1</sup>, Peter Leimgruber<sup>1</sup> and William F. Fagan<sup>2</sup>

<sup>1</sup>Conservation Ecology Center, Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Road, Front Royal, VA 22630, USA; and <sup>2</sup>Department of Biology, University of Maryland, College Park, MD 20742, USA

## Summary

**1** By viewing animal movement paths as realizations of a continuous stochastic process, we introduce a rigorous likelihood method for estimating the statistical parameters of movement processes. This method makes no assumption of a hidden Markov property, places no special emphasis on the sampling rate, is insensitive to irregular sampling and data gaps, can produce reasonable estimates with limited sample sizes and can be used to assign AIC values to a vast array of qualitatively different models of animal movement at the individual and population levels.

**2** To develop our approach, we consider the likelihood of the first two cumulants of stochastic processes, the mean and autocorrelation functions. Together, these measures provide a considerable degree of information regarding searching, foraging, migration and other aspects of animal movement. As a specific example, we develop the likelihood analyses necessary to contrast performance of animal movement models based on Brownian motion, the Ornstein–Uhlenbeck process and a generalization of the Ornstein–Uhlenbeck process that includes ballistic bouts.

**3** We then show how our framework also provides a new and more accurate approach to home-range estimation when compared to estimators that neglect autocorrelation in the movement path.

**4** We apply our methods to a data set on Mongolian gazelles (*Procapra gutturosa*) to identify the movement behaviours and their associated time and length scales that characterize the movement of each individual. Additionally, we show that gazelle annual ranges are vastly larger than those of other non-migratory ungulates.

**Key-words:** autocorrelation function, characteristic scale, home range, irregular sampling, Mongolian gazelle, movement ecology, *Procapra gutturosa*, satellite collar movement data, time series, tracking data gaps

## Introduction

Movement ecology is currently undergoing a period of rapid expansion, fuelled by the increasing availability of high-quality animal relocation data sets. While the methods used to analyse movement data are also evolving, most are still based on discrete-time correlated random walks (CRWs) and their continuous-time diffusion approximations. For example, recently developed composite random walk methods represent elaborations of the basic CRW framework (Kareiva & Shigesada 1983; Turchin 1998; Codling, Plank & Benhamou 2008) While these models enjoy widespread use in ecology, they have a range of known limitations, (Bovet & Benhamou 1988; Turchin 1998; Codling & Hill 2005; Nouvellet, Bacon & Waxman 2009; Gautestad 2012; Fleming *et al.* 2014) whose severity is not fully acknowledged or appreciated, despite repeated demonstrations. Key among these drawbacks is the unrealistic assumption that estimated step lengths and turn angles

actually correspond to start and end points of discrete movement behaviours, whereas location fixes usually follow a programmed schedule that may have little relationship with important behavioural events (Bovet & Benhamou 1988). When discretizing a continuous movement path, step-length and turn-angle distributions are heavily influenced by the choice of sampling rate (Nouvellet, Bacon & Waxman 2009; Fleming *et al.* 2014), which is often based on logistical considerations such as managing GPS-collar battery life rather than *a priori* considerations of the important time-scales of the movement process. As a result, these quantities often reveal little about the movement process. The step-length and turn-angle distributions only play a fundamental role for truly discrete movement processes, where the sampling locations and times correspond to actual behavioural events (i.e. steps and turns).

The more realistic continuous-time stochastic processes (CTSP) models have seen relatively limited use in movement ecology (but see Dunn & Gipson 1977; Brillinger & Stewart 1998; and Blackwell 2003; Horne *et al.* 2007; Johnson *et al.* 2008; Gurarie, Andrews & Laidre 2009), and they have

\*Correspondence author. E-mail: chris.h.fleming@gmail.com

always suffered from the key drawbacks of assuming a particular CTSP model and then fitting it to the data in a fashion analogous to discrete CRW analysis. Prior to Fleming *et al.* (2014), we have not seen any examples of exploratory signal-processing analyses being used to motivate model selection, whereby movement behaviours are first identified in a general way, without assuming any particular model, and then, models incorporating these behaviours are tested. Instead, very particular CTSPs, such as the Orstein–Uhlenbeck process, have always been assumed *a priori*, rather than selected from a set of plausible candidate models. Finally, the likelihood functions used in all studies that fit CTSPs to movement of which we are aware condition only on movements that occur at the sampling rate and effectively difference the data to yield a Markovian representation of the movement process that is only valid if the model is true. This results in a Markovian likelihood function that, while straightforward to implement, is heavily influenced by the choice of sampling rate. As we will show, this does not allow CTSPs to be used to their full potential.

Animal movement data often feature rich autocorrelation structure, including a broad range of spatial and temporal scales and, in some cases, persistence over very long periods of time (Polansky *et al.* 2010; Fleming *et al.* 2014). Although autocorrelation has long been considered a nuisance factor in spatiotemporal ecological data that must be ‘dealt with’ or ‘accounted for’ (Swihart & Slade 1985; Worton 1989), it can, when properly harnessed, reveal a wealth of information about the underlying movement behaviour (Boyce *et al.* 2010; Polansky *et al.* 2010; Fleming *et al.* 2014) and provide new paradigms for analysis (Legendre 1993). While there is growing awareness of the importance of autocorrelation in understanding animal movement, most current analytical methods can only deal with autocorrelation in a very limited way. For example, the CRW framework de-emphasizes the importance of autocorrelation by its very design. These models are constructed as Markov chains, where the current movement step depends only on the previous step. This limits the autocorrelation structure to simple autoregressive models, whether or not this type of model is the most relevant. And while CTSPs are inherently capable of incorporating a much broader array of autocorrelations, currently used Markovian likelihood functions effectively limit them to the same kind of short-range, fast-decaying autocorrelations as CRWs. For more general autocorrelated processes, movement at the current point in time may depend upon a long continuation of past movements. The strength of that dependence decays with time-lag according to characteristic movement time-scales that are generally unrelated to the sampling schedule. Therefore, the inclusion of dependence only on a single time step in the past is an artificial, sampling-dependent construction. While the perspective we advocate is much more difficult to treat mathematically (see, for instance, Adelman 1976; Calzetta, Roura & Verdaguer 2003), only non-Markovian statistical methods that use all possible time-lags in the data can account for and take advantage of the full sweep of autocorrelated behaviour in movement data.

Here, we introduce a non-Markovian likelihood function that can unlock the full power of CTSPs for analysing and understanding animal movement. This likelihood function can fit autocorrelated movement models that can be constructed without any assumption of stationarity or hidden Markov property, whereas previously considered likelihood functions assume that autocorrelated telemetry data can be differenced to obtain ‘hidden’ Markov processes. The sampling schedule is given no special importance in our treatment, and instead, all unobserved times are properly marginalized out of the full-time distribution. In other words, we assume an underlying continuous-time path and correctly treat discrete relocation data as an incomplete sample of this path. Autocorrelation in the data is not avoided by differencing or thinning, but, instead, it is modelled and fit to using all lags in the data. In this way, all noted drawbacks of the previous analyses can be avoided, and arbitrary autocorrelation structures can be tested and compared via likelihood ratio test or AIC. As a case study, we consider a hierarchy of increasingly complex models, from Brownian motion to a multiscale movement model derived in (Fleming *et al.* 2014), with and without assumptions of isotropy, at the population and individual level. The range of included behaviours is beyond the scope of previous analyses and is only possible due to the efficiency and generality of the non-Markovian likelihood function. This family of models is applied to sample of 36 Mongolian gazelles – a species previously described as nomadic (Olson *et al.* 2010; Mueller *et al.* 2011) and later identified as exhibiting movement behaviours at multiple scales (Fleming *et al.* 2014).

#### THE TIME-SERIES ANALYSIS FRAMEWORK

In the CRW framework, there is a straightforward procedure for inspecting the step-length and turn-angle distributions, proposing candidate models and then selecting between them. There is an analogous procedure for continuous-time models, although with a slight difference. In general, one cannot estimate the distribution of a continuous-time process, but one can estimate the moments and cumulants of the process, such as the mean (first cumulant) and autocorrelation function (second cumulant). The first two cumulants alone are sufficient to describe drift and diffusion and completely define any Gaussian stochastic process that has a finite range. Furthermore, the cumulants are not limited to any assumption of stationarity or Markov property. Therefore, the cumulants, which may be considered the fundamental statistics of the stochastic process (Fleming *et al.* 2014), are given primacy in our analysis, rather than the distribution.

Our framework begins by estimating of the movement processes’ autocorrelation structure nonparametrically. The variogram (Cressie 1993; Fleming *et al.* 2014) is an excellent estimator of the mean squared displacement (MSD) for stationary processes. It uses the data much more thoroughly than the conventional MSD statistic (Fleming *et al.* 2014, appendix A.1), allowing investigators to better test the MSD patterns’ fit to a particular movement mechanism and model. The periodogram (Lomb 1976; Scargle 1982), which represents the autocor-

**Table 1.** Comparison of parametric autocorrelation estimators: weighted variogram regression,  $\chi^2$  periodogram regression and maximum likelihood

	Variogram	Periodogram	ML
Estimation error	Worse	$\mathcal{O}(1/n)$	$\mathcal{O}(1/n^2)$
Confidence intervals	Unreliable	Asymptotically reliable	Asymptotically reliable*
Computational complexity	$\mathcal{O}(n^2)^\dagger$	$\mathcal{O}(n \log n)$	$\mathcal{O}(n^3)$
Data quality sensitivity	Intermediate	Fragile	Robust
Non-stationarity	Problematic	$\mu(t)$ only	$\mu(t)$ and $\sigma(t, t')$

\*The asymptotic properties of ML are typically achieved fairly easily for most time-series data.

<sup>†</sup>A fast  $\mathcal{O}(n \log n)$  algorithm for the method-of-moments estimator is given by Marcotte (1996).

relation structure in the frequency domain, can be used to identify cycles and periodicities in the movement. One might be inclined to skip nonparametric estimation and perform a model selection based upon likelihood and AIC. However, this assumes that all relevant models will be included in the AIC comparison (and in their relevant parameter regimes). Our previous analysis of Mongolian gazelles offers a good example against this practice, as our best movement model was invented only after noting significant discrepancy between the variogram and Ornstein–Uhlenbeck motion (Fleming *et al.* 2014).

The second step, which this work focuses on, is to select the best model and estimate model parameters. There are three classes of approaches to parametric autocorrelation estimation: weighted variogram regression (Cressie 1993), Chi-squared ( $\chi^2$ ) periodogram regression (Whittle 1953) and non-Markovian maximum likelihood (Table 1). Weighted variogram regression, which we used in Fleming *et al.* (2014), is the most accessible of the three methods, but it does not produce reliable confidence intervals (Fleming & Calabrese 2013).  $\chi^2$  periodogram regression is the most computationally efficient of the three methods, but it only applies to evenly sampled data. Here, we derive the likelihood function for data of a specified mean and autocorrelation function, as given by a particular model of movement. Non-Markovian maximum likelihood accurately and efficiently uses limited data and is completely insensitive to the irregular sampling and data gapping often found in animal relocation data sets. We emphasize that in our approach there are no assumptions pertaining to the data quality (in terms of gaps), amount of data,<sup>1</sup> or hidden Markov property (i.e. that one can difference the data enough times and rid it of autocorrelations).

Once a suitable modelling approach is selected, the autocorrelation estimate can then be applied to better answer movement-related questions, as any movement-related quantity of

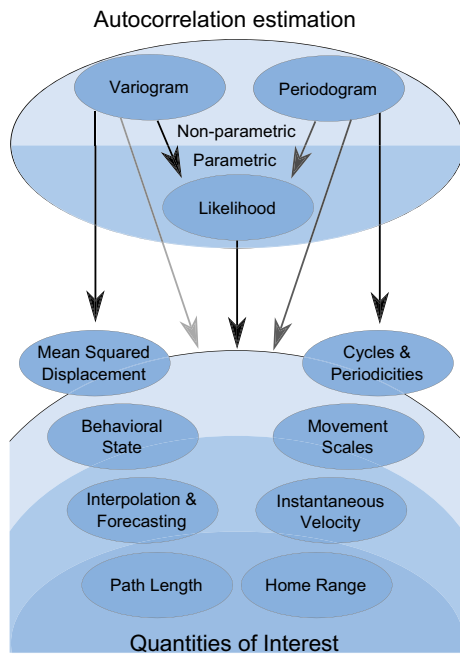
<sup>1</sup>Although likelihood-based estimates, confidence intervals and AIC comparisons are only asymptotically valid, this is typically not an issue of primary concern for time-series data where ‘small’ data sets consist of numerous observations.

interest necessarily requires knowledge of the autocorrelation function (ACF) to make optimal estimates. We can compare characteristic time and length scales based on different behavioural movement modes identified by the best-fit model. Using the likelihood function, we can also estimate probabilities for different movement behaviours in individuals and groups, and even at different times, as with composite random walk models (Morales *et al.* 2004; Jonsen, Flemming & Myers 2005; McClintock *et al.* 2012). Finally, current strategies to estimate home ranges such as kernel-density estimates (KDE) (Worton 1989) and minimal convex polygons (MCP) (Southwood 2000) are incomplete because they do not adequately take autocorrelation into account, which results in the underestimation of home-range sizes (Fleming *et al.* 2014, appendix E). To be valid, these estimators require a coarsening of the data, which is then discarding information that a better estimator could use to improve the estimate (for an example of how small-scale details can improve a large scale estimate, see Appendix S4). Similarly, the problem of predicting an animal’s location when it is not sampled (interpolation and forecasting) is nearly impossible without a probabilistic movement model that fully accounts for autocorrelation (see, for the spatial-field analogue, Krige 1951; Cressie 1993). All of these quantities require knowledge of the movement processes’ autocorrelation structure to be optimally estimated. Estimators that ignore the underlying autocorrelation structure can be heavily biased, resulting in over- or underestimation of the desired quantities. In Fig. 1, we summarize the principal analysis flow from the nonparametric estimation of autocorrelation structure, to the parametric estimation of model parameters (and model selection) and to the estimation of quantities of interest related to movement.

## THE CASE STUDY

To demonstrate the power of a full non-Markovian likelihood analysis, we analyse 5 years of relocation data for Mongolian gazelles collected in the eastern steppe of Mongolia using GPS-Argos collars (Fleming *et al.* 2013). Mongolian gazelles are intermediate herbivores (Jiang *et al.* 2002) that live in one of the world’s largest remaining grasslands and have previously been described as nomadic (Olson *et al.* 2010; Mueller *et al.* 2011). They exhibit long distance movements, and although their individual ranges are smaller than the available grassland (Fleming *et al.* 2014), they do not reside in small home ranges that are typical for range resident species. It takes months for the gazelles to cross their range, yet they do not undertake regular migrations, nor do gazelles exhibit any obvious cycles or periodicities (Fleming *et al.* 2014). While the likelihood method can easily handle complexities such as migrations and repeating (periodic) behaviours, with the gazelle we only have to address the autocorrelation function proper and not a time-dependent mean. We consider a family of models including Brownian motion with the possibility of home-range fidelity at long time-scales and a ballistic phase at short time-scales, which might be related to commuting or foraging behaviours.

Model identification for this species was previously performed in (Fleming *et al.* 2014) using variogram analysis. We



**Fig. 1.** A visual overview of the analysis flow from the estimation of autocorrelation to the estimation of important movement quantities, such as characteristic scales, path length, velocity, area use (e.g. home range), cyclicity and behaviour. Darker arrows indicate a preferred method for approaching the endpoint. That is, the periodogram is more suitable for addressing the existence of periodicities than the variogram, whereas maximum likelihood is generally preferable when preceded by a nonparametric analysis. Certain quantities of interest, such as mean squared displacement, can be estimated quickly with little effort, while quantities near the bottom of the figure require more analysis beyond autocorrelation estimation.

repeat the same model selection here for demonstration purposes. However, we emphasize that, although maximum likelihood is the superior estimator, it does not yield any visual representation of the ACF, and so a variogram or periodogram analysis should always be the first step of analysis to ensure that all important movement behaviours are included in the modelling. Here in step two of our framework (Fig. 1), we employ maximum likelihood to obtain better parameter estimates and confidence intervals. The mean and ACF estimates can then be applied to calculating quantities of interest, which is what we do to estimate the individual- and population-level utilization distributions and 95% confidence region ranges.

To ensure a variety of sampling rates, individual collars were programmed for a pair of alternating collection intervals, with periods of 1, 5 and 25 h. Additionally, to preserve battery life, collars were programmed to have 10 day gaps after every 5 days of data collection. There was also data loss from malfunction; however, estimated telemetry errors were very small (details of the error estimates can be found in Fleming *et al.* 2014). It is important to note that this sampling schedule, which represents an attempt to balance fine-scale detail against collar longevity, is far from optimal and cannot be recommended. Specifically, the mix of different sampling rates and data gaps it introduced caused conventional variogram, periodogram and composite random walk methods to fail on

this data set. After modifying conventional variogram analysis to handle such heterogeneity and data gaps (Fleming *et al.* 2014), we were limited to the analysis of the time-averaged and population-averaged isotropic ( $x$ - $y$  averaged) autocorrelations. As a result, previously we could only make inferences about how an average gazelle moves in an average direction. However, the irregular sampling schedule and random gaps do not pose any problem for the non-Markovian likelihood function, and here, we can analyse individual movement paths and relax the isotropy assumption.

## Statistical framework and movement models

The first two cumulants of the stochastic process provide the backbone for our calculations: the mean  $\mu(t)$ :

$$\mu(t) = \langle \mathbf{r}(t) \rangle, \quad \mathbf{r}(t) \equiv (x(t), y(t)), \quad \text{eqn 1}$$

and the autocorrelation function  $\sigma(t, t')$ :

$$\sigma(t, t') = \left\langle [\mathbf{r}(t) - \mu(t)][\mathbf{r}(t') - \mu(t')]^T \right\rangle, \quad \text{eqn 2}$$

where  $t$  denotes the time index,  $\mathbf{r} = (x, y)$  denotes the location, and  $\langle \dots \rangle$  denotes the average over realizations of the process. Different movement models predict different functional forms for the first two cumulants. Our approach to fitting movement models to relocation data is to first calculate the model mean and ACF in terms of the movement parameters, such as the characteristic time and length scales, and then to maximize the likelihood of the location data with respect to these parameters. A good method of fitting the first two cumulants to the data is to assume a Gaussian distribution for the relocation time series. This produces standard formulas for method-of-moments estimates (Pawitan 2001), standard formulas for the periodogram in the case of stationary cumulants (Dembo 1986), an asymptotically normal estimator for any autoregressive moving-average (ARMA) model (Fan & Yao 2003), and it is the distribution of maximum entropy given no further knowledge of the higher-order cumulants (Cover & Thomas 2006). For time-correlated movement, this distributional assumption is fairly permissive, in that it restricts neither the step-length distribution (given a discrete sampling) nor the home range to be a Gaussian function, and can incorporate any number of movement behaviours.

The Gaussian probability density functional  $P[\mathbf{r}]$  of a realized trajectory  $\mathbf{r}(t)$  is formally given by

$$P[\mathbf{r}] = \frac{e^{-\frac{1}{2} \int \int dt dt' [\mathbf{r}(t) - \mu(t)]^T \sigma^{-1}(t, t') [\mathbf{r}(t') - \mu(t')]} }{\sqrt{\det 2\pi\sigma}}. \quad \text{eqn 3}$$

This is a continuous generalization of the multivariate Gaussian probability distribution, where  $t$  acts as an index for the Gaussian random variable  $\mathbf{r}(t)$ , which is distinct from, yet correlated to, the Gaussian random variable  $\mathbf{r}(t')$  at a different time  $t'$ . The continuous trajectory is only sampled a countable number of times, and so we marginalize over all times for which data are not observed. This results in the marginal probability distribution of the observed time-series data



$$P(\mathbf{R}) = \frac{e^{-\frac{1}{2}(\mathbf{R} - \mathbf{M})^T \Sigma^{-1} (\mathbf{R} - \mathbf{M})}}{\sqrt{\det 2\pi \Sigma}}, \quad \text{eqn 4}$$

in terms of the location data  $\mathbf{R}$ , sampled mean  $\mathbf{M}$  and sampled autocorrelation  $\Sigma$ , with components

$$[\mathbf{R}]_i = \mathbf{r}(t_i), \quad [\mathbf{M}]_i = \boldsymbol{\mu}(t_i), \quad [\Sigma]_{ij} = \boldsymbol{\sigma}(t_i, t_j), \quad \text{eqn 5}$$

where the observed times  $t_i$  are indexed by  $i \in \{1, 2, 3, \dots\}$ . Arbitrary mean and autocorrelation functions can be considered for the stochastic process, which can incorporate various movement behaviours and parameters. The likelihood function does not care how we sampled the data nor if it contains gaps. The only mathematical requirement is that the autocorrelation function is positive definite:<sup>2</sup>  $\Sigma$  must have all positive eigenvalues. Note that for  $q$  spatial dimensions and  $n$  sampled times,  $\mathbf{R}$  and  $\mathbf{M}$  are  $n \times q$  dimensional, while  $\Sigma$  is  $n \times n \times q \times q$  dimensional. These arrays must be properly flattened into  $nq$ -dimensional vectors and  $(nq) \times (nq)$ -dimensional matrices to use conventional matrix operations. We will explicitly do this in the case of two dimensions.

Following probability distribution (4), the log-likelihood function is given by

$$\ell = -\frac{1}{2} \log \det \Sigma - \frac{1}{2} (\mathbf{R} - \mathbf{M})^T \Sigma^{-1} (\mathbf{R} - \mathbf{M}), \quad \text{eqn 6}$$

to within a constant. Similar likelihood functions have been derived in other contexts (e.g. see Mardia & Marshall 1984, for spatial fields). An AIC comparison or likelihood ratio test will then allow us to determine what behaviours, as expressed by parameters in the mean and autocorrelation function, are suggested by the data. In the family of models we consider, there are two movement processes that have been historically considered for movement processes: Brownian motion (BM) and Ornstein–Uhlenbeck (OU) motion. The OU process (Gardiner 2009) describes a random search within a defined area that grows more-and-more slowly in time and asymptotes to a finite value.

The OU autocorrelation function is given by

$$\sigma(t, t') = \sigma_H e^{-\frac{|t-t'|}{\tau_H}}, \quad \text{eqn 7}$$

where  $\sigma_H$  is the variance of the animal's utilization distribution, while  $\tau_H$  is the time-scale in which the animal crosses this range. For small time-lags,  $\tau = |t-t'| < \tau_H$ , the diffusion is regular, and in the limit  $1/\tau_H \rightarrow 0$ , this model reduces to Brownian motion with a diffusion rate given by the limit of  $\sigma_H/\tau_H$ . For larger lags,  $\tau > \tau_H$ , the variance is asymptotically given by  $\sigma_H$  and so the displacement is subdiffusive.

In Fleming *et al.* (2014), the OU model was generalized by the inclusion of random periods of autocorrelated velocity (i.e. periods of time in which an animal tended to maintain its present rate of movement or lack thereof). These ballistic periods are characterized by a short time-scale  $\tau_F$ , which were hypothe-

sized to correspond to foraging behaviour, given that a grazing ungulate tends to move in a relatively straight line as it feeds (Bailey *et al.* 1996). However, this effect also arises from CRW behaviour at short time-scales (see Gurarie & Ovaskainen 2011, for the OU velocity process). The corresponding autocorrelation function is given by

$$\sigma(t, t') = \sigma_H \frac{\tau_H e^{-\frac{|t-t'|}{\tau_H}} - \tau_F e^{-\frac{|t-t'|}{\tau_F}}}{\tau_H - \tau_F}, \quad \text{eqn 8}$$

and the quantity  $\sigma_F \equiv (\tau_F/\tau_H)\sigma_H$  then corresponds to the range associated with the autocorrelated movements, which could possibly be associated with a foraging range. We will refer to this model as the OUF process. The OUF model contains within it the Ornstein–Uhlenbeck (OU) model and within that the classical Brownian motion (BM) model. From BM to OU to OUF, the likelihood cannot decrease; however, each step adds one parameter. Like the OU process, for large lags,  $\tau > \tau_H$ , the variance is asymptotically constant, and for intermediate lags,  $\tau_F < \tau < \tau_H$ , the diffusion is regular. For small lags,  $\tau < \tau_F$ , the displacement is technically superdiffusive, with the semi-variance function and mean squared displacement proportional to  $\tau^2$ . However, the mean squared displacement in this regime is always less than that from the regular diffusion at intermediate lags, as the semi-variance and mean squared displacement strictly grow larger with time-lag. In the limit  $\tau_F \rightarrow 0$ , this model reduces to the OU model.

The BM, OU and OUF models all describe stationary processes, and yet ecological processes are assuredly non-stationary, as they are driven by diurnal and seasonal cycles. Although the mean and autocorrelation functions can be specified to incorporate non-stationary behaviours, this requires some additional effort. The easiest, though, least informative approach is to treat the non-stationarity as a nuisance parameter and average over all time dependence. This is effectively what is done when fitting a stationary model to non-stationary processes via maximum likelihood (Appendix S1). To construct specific parametric models of non-stationary behaviour, one can rely upon nonparametric estimates of the non-stationary mean and autocorrelation function, such as the population mean and Wigner–Ville function (Ville 1948; Hillery *et al.* 1984). Finally, we note that most ideas for addressing non-stationarity in movement ecology can be adapted to the formalism we propose, so that they better take into account the non-Markovian behaviours. In particular, behavioural change point analysis (BCPA, Gurarie *et al.*, 2009) requires only a minimal substitution of likelihood functions.

## The likelihood function

In the case of a two-dimensional stochastic movement process with isotropic correlations, the autocorrelation function can be represented

$$\sigma(t, t') = \begin{bmatrix} \sigma_0 & 0 \\ 0 & \sigma_0 \end{bmatrix} c(t, t'), \quad \text{eqn 9}$$

<sup>2</sup>Not all positive functions are positive definite and not all positive definite functions are everywhere positive.

where  $\sigma_0$  is the variance and  $c(t, t')$  is a dimensionless, scalar correlation function – it carries no units of space or time and is not a matrix. E.g.,  $c(t, t') = \exp(-|t-t'|/\tau_H)$  for an OU process. Under this assumption, log-likelihood function (6) simplifies to

$$\ell = -\log \det(\mathbf{C}) - n \log \sigma_0 - n \frac{s_{xx} + s_{yy}}{2\sigma_0}, \quad \text{eqn 10}$$

to within a constant, where  $n$  is the number of times that the movement path is sampled (i.e. the number of relocations per individual animal), the correlation matrix  $\mathbf{C}$  is defined by  $C_{ij} \equiv c(t_i, t_j)$ , and the sample correlations are given by

$$s_{xx} \equiv \frac{1}{n} (\mathbf{X} - \mathbf{M}_x)^T \mathbf{C}^{-1} (\mathbf{X} - \mathbf{M}_x), \quad \text{eqn 11}$$

$$s_{yy} \equiv \frac{1}{n} (\mathbf{Y} - \mathbf{M}_y)^T \mathbf{C}^{-1} (\mathbf{Y} - \mathbf{M}_y), \quad \text{eqn 12}$$

$$s_{xy} \equiv \frac{1}{n} (\mathbf{X} - \mathbf{M}_x)^T \mathbf{C}^{-1} (\mathbf{Y} - \mathbf{M}_y), \quad \text{eqn 13}$$

in terms of the data and mean vectors

$$X_i = x(t_i), \quad Y_i = y(t_i), \quad \text{eqn 14}$$

$$M_i^x = \mu_x(t_i), \quad M_i^y = \mu_y(t_i). \quad \text{eqn 15}$$

So for the OU process, the location data enter into the likelihood function via the vectors  $\mathbf{X}$  and  $\mathbf{Y}$  and the static mean  $\mu_x$ ,  $\mu_y$ , variance  $\sigma_0$ , and range crossing time-scale  $\tau_H$  would be the parameters allowed to vary in the likelihood function. Each time the parameter  $\tau_H$  is varied, the matrix  $\mathbf{C}$ , its inverse and its determinant must all be updated to evaluate the likelihood function.

The assumption of isotropy can be relaxed at the cost of additional model parameters, which can be justified by comparing AIC values or with a likelihood ratio test. We consider a two-dimensional process with anisotropic correlations that are uniform in time:

$$\sigma(t, t') = \sigma_0 c(t, t'), \quad \sigma_0 \equiv \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix}. \quad \text{eqn 16}$$

Log-likelihood function (6) can then be expressed

$$\ell = -\log \det(\mathbf{C}) - \frac{n}{2} \log \det \sigma_0 - n \frac{\sigma_{yy}s_{xx} + \sigma_{xx}s_{yy} - 2\sigma_{xy}s_{xy}}{2 \det \sigma_0}, \quad \text{eqn 17}$$

after some algebraic manipulations (using block-matrix inverse and determinant relations) to the multivariate Gaussian likelihood function. Given the likelihood function in equations (10) or (17), the mean (and its standard error) and the variance (and its standard error) can be solved (i.e. maximized or profiled) via some straightforward linear algebra (Appendix S2). This allows for model parameterization by standard maximum likelihood techniques and for model comparison via AIC or likelihood ratio test.

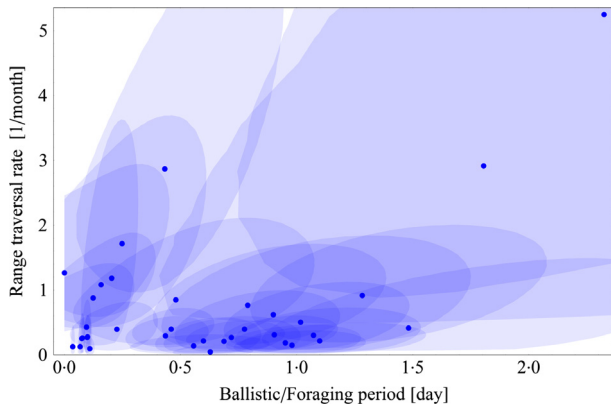
The non-Markovian likelihood functions we have introduced take into account an animal's autocorrelations over all possible time-lags, regardless of the model chosen and sampling rate employed. Our approach stands in contrast to

conditional, Markovian likelihood functions that have been used in ecology to estimate the parameters of Brownian and OU movement processes (Dunn & Gipson 1977; Brillinger & Stewart 1998; Blackwell 2003; Horne *et al.* 2007; Johnson *et al.* 2008; Gurarie, Andrews & Laidre 2009). The Markovian likelihood functions are conditioned upon data differences and are constructed to be uncorrelated under the assumption of a BM or OU model. The Markov property vastly simplifies the resulting likelihood functions, which do not contain any matrix operations if only non-overlapping segments of time are considered. There are three significant drawbacks to this construction. First, the Markov property only applies under the assumption that the particular model is correct, and this assumption only holds when the model is a very accurate representation of the process that generated the data *at the specific sampling rate* for which the differences are calculated. In contrast to this, our more general non-Markovian likelihood function accounts for all possible lags in the data via the correlation matrix  $C_{ij} = c(t_i, t_j)$ , and therefore, it is much more robust to situations where the stochastic movement process is misspecified. Second, we prove in Appendix S3 that for any realistic movement process  $x(t)$  that has continuous and bounded velocities, with fine sampling rate  $\Delta t = t_{i+1} - t_i$ , the diffusion rates and corresponding standard errors estimated by Markovian likelihood methods will be proportional to  $\Delta t$ , which vanishes in the limit of continuously sampled data. Finally, by construction, the Markovian likelihood functions are limited to a specific order of continuous-time ARMA models. For instance, the Markovian likelihood function that conditions upon pairs of locations  $\mathcal{A}(x_i, x_j)$  can be applied to compare the AIC values of the BM and OU models, but it cannot be applied to the OUF model, which requires location triplets  $\mathcal{A}(x_i, x_j, x_k)$ . In contrast, the more general likelihood function can be applied to arbitrary mean and autocorrelation functions.

## Analysis of gazelle movement

Overall the anisotropic individual-level OUF model was best supported by the data. We plot the likelihood profiles for this model in Fig. 2 and summarize the model fits in Table 2. Starting with the population-level isotropic OUF model, which was selected by variogram analysis, we fit the models to each of the gazelles. The gazelles had individual sample sizes ranging from 20 to 886 relocations. Allowing each gazelle to have independent time and length scale parameters, rather than a single set of scale parameters describing the entire population of gazelles, resulted in a further AIC decrease of 2690, indicating genuine variability in scales among the gazelles (Table 2). Compared to the population scales, the individual  $\tau_F$  range up to a day and  $\tau_H$  range down to tens of days (Fig. 2). This places the previous population-level results of variogram regression (Fleming *et al.* 2014; Table 2) within the range of individual variability. Allowing each gazelle to have an independent anisotropic covariance  $\sigma_0$  resulted in a further AIC decrease of 416, with very little effect on estimates of the time-scale parameters.

To repeat the model selection of our variogram analysis (Fleming *et al.* 2014), we also considered the OU and BM



**Fig. 2.** ML point estimates and 95% likelihood-profile confidence regions for the individual gazelle time-scales. Foraging time-scales  $\tau_F$  range from two hours to a day, while ranging time-scales  $\tau_H$  range from tens to hundreds of days. The left  $\tau_F = 0$  boundary corresponds to the OU model, while the bottom  $1/\tau_H = 0$  boundary corresponds to a model with regular diffusion at time-scales greater than  $\tau_F$  and thus no finite asymptotic ranging area. The origin corresponds to classical Brownian motion. One can see that the ML estimate from the population-level OUF model (Table 2) is located in the cluster of highly peaked individual profiles at small  $\tau_F$  and small  $1/\tau_H$ . While these individuals have very tight confidence intervals on their  $\tau_F$  estimates, it is also clear that many gazelles have much larger foraging time-scales ranging from half a day to a day. One can see that a few gazelles were consistent with OU motion, whereas none of the gazelles were consistent with Brownian motion. What is not clearly visible in this plot is that none of the individual gazelles touch the  $1/\tau_H = 0$  boundary, and so they are all predicted to have finite ranging areas, as one would reasonably expect.

models. The isotropic OU model resulted in a marginally larger range crossing time  $\tau_H$  estimate and was roughly consistent with some aspects of the data, while the BM model was found to be a very poor model of gazelle movement (Appendix S7). For comparison purposes, we also fit the data with the Markovian OU likelihood function that conditions only upon differences in the data, which corresponds to the techniques of Gurarie, Andrews & Laidre (2009). Relative to all non-Markovian estimates (including OU), this method resulted in an

approximately 50% smaller range crossing time estimate  $\tau_H$ , with grossly underestimated confidence intervals (Table 2). Moreover, as we prove in Appendix S3, in the limit of continuously sampled data, the Markovian likelihood function produces an estimate of  $\tau_H$  that diverges. Ironically, the Markovian likelihood actually benefited in this case from the relatively infrequent sampling and excessive gapping in the data. This means that if one could improve the data quality by hand, perhaps by tracking down the collar and extracting data that was not uploaded to the satellite, then the Markovian likelihood estimate of  $\tau_H$  would increase through the correct range of values and ultimately become arbitrarily large, yet with deceptively small confidence intervals.

We also performed a data simulation using the same sampling schedule, maximum likelihood (ML) parameter estimates and estimated GPS errors, as detailed and performed in Fleming *et al.* (2014). The results of an OUF fit showed marked improvement over the variogram regression and Markovian likelihood function, with the true values of all parameters being contained inside the 95% confidence intervals of the non-Markovian ML estimates. The simulations also served to demonstrate that  $\tau_H$  was genuinely difficult to estimate with the given data schedule, even with the underlying model being correctly specified, as the confidence intervals for  $\tau_H$  were as large with the simulated data as they were with the real data. This results from the fact that the period of the data samples,  $T = t_n - t_1$ , was not significantly larger than the true value of  $\tau_H$ , as the effective sample size necessary to estimate a long time-scale  $\tau_c$  in a single time series is roughly given by  $T/\tau_c$ , and not by the total amount of data  $n$ .

### Analysis of gazelle area utilization

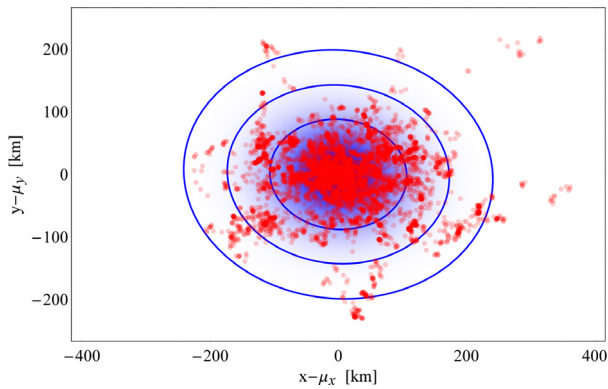
Using the probability relationships outlined in Appendix S6, the population-level anisotropic OUF model resulted in a ranging area of  $A_{95\%} = 77\,600 \pm 1200 \text{ km}^2$ , which we plot in Fig. 3, with  $A_P$  referring to the most probable area in which the 36 gazelles in the sample population will be located a proportion  $P$  of all time. The breakdown of the individual area scales

**Table 2.** Comparison of alternative movement model fits to Mongolian gazelle data, including Akaike Information Criterion and best-fit parameters

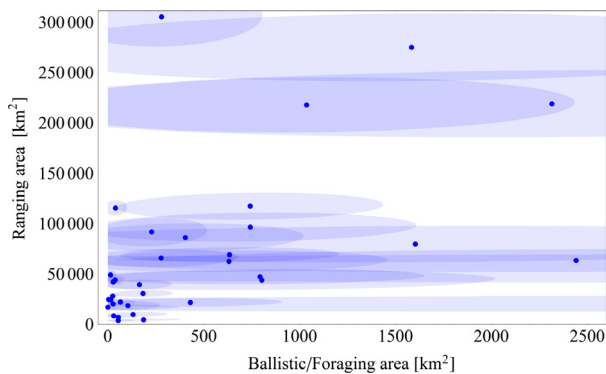
			$\Delta\text{AIC}$	$\tau_H$ [days]	$\tau_F$ [hours]	$A_H$ [km <sup>2</sup> ]	$A_F$ [km <sup>2</sup> ]
Population-level model							
(Markovian)	Isotropic	OU	NA	$107 \pm \begin{smallmatrix} 13 \\ 11 \end{smallmatrix}$	NA	$30\,120 \pm 660$	NA
(Variogram)	Isotropic	OUF	NA	$74.1 \pm 3.6$	$6.16 \pm 0.58$	$91\,000 \pm 21\,000$	$315 \pm 73$
	Isotropic	BM	>8729	NA	NA	NA	NA
	Isotropic	OU	8229	$274 \pm \begin{smallmatrix} 68 \\ 48 \end{smallmatrix}$	NA	$80\,500 \pm 1800$	NA
	Isotropic	OUF	3106	$199 \pm \begin{smallmatrix} 45 \\ 32 \end{smallmatrix}$	$2.459 \pm \begin{smallmatrix} 0.083 \\ 0.082 \end{smallmatrix}$	$80\,500 \pm 1800$	$41.3 \pm 8.0$
	Anisotropic	OUF	2797	$198 \pm \begin{smallmatrix} 45 \\ 32 \end{smallmatrix}$	$2.415 \pm \begin{smallmatrix} 0.083 \\ 0.081 \end{smallmatrix}$	$77\,600 \pm 1200$	$39.5 \pm 8.1$
Individual-level model							
	Isotropic	OUF	416	$170 \pm 380$	$1.8 \pm 3.0$	$14\,000 \pm 34\,000$	$10 \pm 80$
	Anisotropic	OUF	0	$160 \pm 370$	$1.8 \pm 2.8$	$12\,000 \pm 30\,000$	$11 \pm 72$

$A_H \equiv z_{95\%}^2 \pi \sqrt{\det \sigma_0}$  denotes the 95% confidence ranging area, while  $A_F \equiv z_{95\%}^2 \pi (\tau_F / \tau_H) \sqrt{\det \sigma_0}$  denotes the 95% confidence ballistic-motion or foraging area. For the individual-level models, the maximum likelihood estimates are given with their 95% confidence intervals, whereas for the individual-level models, the weighted mean of the parameter estimates are given with corresponding 95% confidence intervals for the population's parameter distribution (not standard errors).





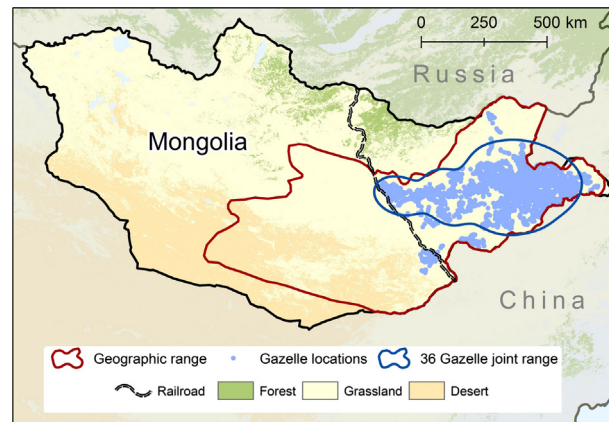
**Fig. 3.** Superimposed distributions of all mean-detrended gazelle locations in red, with the predicted density function for the population-level anisotropic OUF model in blue, and blue contours delineating the 68%, 95% and 99.7% confidence regions. Structure present in the sample distribution represents not only deviation from the model, but also the artefacts of bad gapping in the data and significant autocorrelation, both of which cause clumping.



**Fig. 4.** ML point estimates and 95% likelihood-profile confidence regions for the individual gazelle (95% CR) ranging and foraging areas. Foraging areas vary from tens to hundreds of square kilometres, while ranging areas vary from thousands to several hundred thousand kilometres.

are given in Fig. 4 and the joint utilization distribution of the 36 collared gazelles is given in Fig. 5, resulting in a joint-distribution area of  $A_{95\%}^{\text{joint}} = 200\,000 \pm 3100 \text{ km}^2$ . Obviously, the joint-distribution area of the population of  $\sim 1.1$  million gazelles will be even larger. However, the joint-distribution area of a mere 36 gazelles already almost completely fills the total available area of the Eastern Steppe, when bounded by the fenced borders to China and Russia, as well as the Trans-Mongolian railway, which is also fenced. Therefore, one can easily predict that an analysis featuring more individuals would indicate that this entire area of land is being utilized (Fig. 6), and so its encroachment would have a significant impact on gazelle behaviour.

For comparison purposes, we also fit the data with both Markovian and non-Markovian likelihood functions for the isotropic OU model. The non-Markovian likelihood function yielded a utilization range estimate similar to that of the OUF model, whereas the Markovian likelihood function resulted in a significantly smaller ranging area estimate of



**Fig. 5.** An ecosystem and land coverage map of Mongolia outlining the current geographical range of the Mongolian gazelle, which is bisected by the Trans-Mongolian railway, and the result of our 36 gazelle study, with individual locations in light blue and a contour delineating the 95% confidence region in dark blue. The 36 gazelle utilization area does not correspond to a population range, but instead to the most probable area containing the sampled gazelles 95% of the time, with 5% of locations spilling out into less probable regions. We further note that individual gazelle locations that are close together in time receive lower weight in the calculation of the joint utilization area than do observations that are farther apart. In other words, the uneven time sampling of the gazelle data makes it all but impossible to visually judge where the 95% probability contour should fall.

$30\,120 \pm 660 \text{ km}^2$ . This mismatch is another indication of significant bias in the Markovian likelihood function, as even a bivariate Gaussian estimate (for an anisotropic 2D Gaussian distribution without autocorrelation) yields an area estimate of  $A_{95\%} = 63\,480 \pm 500 \text{ km}^2$ . In other words, the Markovian likelihood function has produced an area estimate which is even worse than the simplest possible analysis. By contrast, weighted variogram regression, which is also a convenient estimator, yielded a more reasonable estimate of  $A_{95\%} = 91\,000 \pm 21\,000 \text{ km}^2$  (Fleming *et al.* 2014).

## Discussion

We have introduced a rigorous method of modelling animal movement focusing on the first two cumulants of a stochastic process: the mean and autocorrelation function, which are sufficient to describe drift and diffusion (Adelman 1976; Calzetta, Roura & Verdaguer 2003). This method can incorporate a diverse range of behaviours, including home-range fidelity, searching and foraging. Using different movement models beyond the three discussed here (BM, OU and OUF) allows the approach to be extended to additional behaviour types, such as migration. We express the models by their mean and autocorrelation functions and then fit them to the data by maximizing the (non-Markovian) likelihood function. This novel approach to movement analysis allows us to identify mixtures of movement processes at the individual level, along with their corresponding characteristic time and length scales.

The flow of analysis we advocate (Fig. 1) involves three stages, with the likelihood method being central. In the first stage, one applies nonparametric estimators, such as



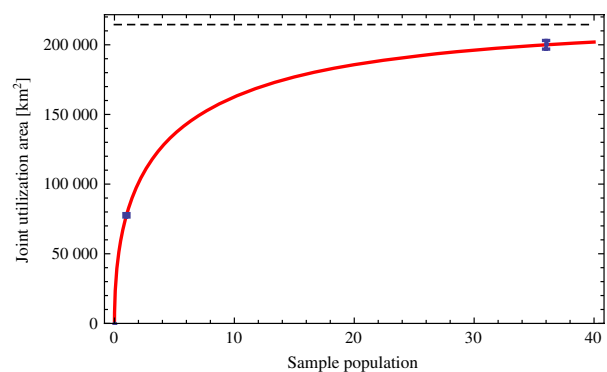
variogram analysis (Fleming *et al.* 2014), to identify reasonable movement models and obtain rough estimates of the movement parameters. The identified movement models should be capable of explaining the statistically significant and visually apparent features of the mean, variogram and periodogram. In stage two, a likelihood analysis is used to obtain more accurate parameter estimates, reliable confidence intervals and a more efficient use of the data, allowing for fine-grained analyses such as the estimation of anisotropy and individual-level variability. In stage three, the best-fit mean and autocorrelation functions are applied to obtain reliable and statistically efficient estimates of quantities of interest that are conditioned upon movement, and therefore based upon the autocorrelation structure of the data.

There are four advantages of our likelihood approach over variogram regression: heightened accuracy; reliable confidence-interval estimation (Fleming & Calabrese 2013); robustness to irregularities in the sampling schedule; and the ease with which it can more naturally incorporate non-stationary movement behaviours, such as migration. However, the likelihood function requires that researchers identify a suitable choice of movement model. For this purpose, a nonparametric estimator, such as the variogram approach, remains useful as a model identification tool. Our analysis has confirmed the existence of multiscale movement processes suggested by the population-level variogram analysis of Fleming *et al.* (2014), from small-scale foraging to an asymptotic (multiyear) range (Table 2). The superior efficiency of maximum likelihood analysis allowed us to obtain parameter estimates at the individual level and in each of the spatial dimensions independently. This analysis revealed significant variability among individuals as well as pronounced directional bias in their movement.

Like the variogram method introduced in Fleming *et al.* (2014), our non-Markovian likelihood approach conditions upon all time-lags in the data, thereby avoiding the sampling problem inherent in currently standard movement analyses. For the conventional step-length and turn-angle analysis, such limitations are well known (Bovet & Benhamou 1988; Turchin 1998; Codling & Hill 2005; Gautestad 2012; Fleming *et al.* 2014). However, we have demonstrated that similar limitations also apply to Markovian likelihood functions that condition upon differenced data (see Appendix S3). In contrast to ad hoc estimators such as minimum convex polygons and kernel densities, the home-range and area utilization estimation provided by our likelihood analysis is inherently mechanistic, as it conditions upon the appropriate autocorrelation structure present in the data. Moreover, as the underlying model is stochastic and self-consistent, there is no fine tuning of ancillary parameters, such as a degree of refinement or bandwidth size. Compared to mechanistic home-range analysis (MHRA) (Moorcroft, Lewis & Crabtree 2006; Moorcroft & Lewis 2006), the parametric home-range estimator introduced here is limited because the stationary distributions of individuals will always be Gaussian. This stationary approach is inappropriate for migratory species. However, this limitation can be resolved through the use of non-stationary movement processes and related approaches, which we leave for a future effort. Two

factors give our non-Markovian likelihood estimator an advantage over MHRA. First, MHRA requires a much more detailed understanding of the underlying movement mechanisms. Second, MHRA conditions upon advection-diffusion equations, which are inherently incapable of generating the multitime correlations that can be generated by Langevin equations (Calzetta, Roura & Verdaguer 2003), which are the stochastic (and possibly integro-differential) equations of motion for the animal trajectories. The limitation implies that the larger class of non-Markovian processes, and OUF in particular, cannot be modelled with MHRA, at present. On the other hand, the MHRA has a more detailed and mechanistic description of the movement process, and so it allows for the prediction of movement behaviours under hypothetical conditions.

The likelihood analysis is efficient and robust enough to fit to each individual gazelle, which paves the way for many future analyses that require individual-level estimates of movement path autocorrelation structure. Quantities of interest that require such estimates (Fig. 1) include optimal (nonparametric) home-range estimation, interpolation and forecasting of animal locations, path length and instantaneous velocity estimation, and transitions between behavioural states. The individual-level analysis, which was precluded in the variogram analysis by too little data (Fleming *et al.* 2014), also enables the exploration of hypotheses on the relationships between different characteristic movement scales. Although no obvious trends presented themselves in Figs 2 and 4, any relation between time-scales, length scales, diffusion or rate scales can be tested with the likelihood function. It is clear from our analysis that the population scales are well defined by a single global maximum in the likelihood profile, but the individual scales vary significantly within the population (Table 2). This naturally suggests the hypothesis that some of the variability in movement behaviour among individuals may be attributable to variability in the environment. We will explore this idea in the future using individual-level analyses that relate movement processes to environmental covariates.



**Fig. 6.** Joint utilization areas for collared gazelles as a function of (sample) population size in blue. The thick red curve outlines the asymptotic utilization behaviour, whereby a very small number of gazelle utilize the entire Eastern Steppe, and the dashed black line represents the hypothetical asymptotic joint utilization area, which is quickly filled by a relatively small sample of gazelles.

## Acknowledgments

The project was funded by the US National Science Foundation ABI 1062411. We thank Caroline Wick for editorial assistance.

## Data accessibility

All gazelle location data can be accessed through the Dryad digital repository: <http://dx.doi.org/10.5061/dryad.45157> (Fleming *et al.* 2013). Commented Mathematica scripts have been uploaded as supporting information.

## References

- Adelman, S.A. (1976) Fokker–Planck equations for simple non-Markovian systems. *The Journal of Chemical Physics* **64**, 124–130.
- Bailey, D.W., Gross, J.E., Laca, E.A., Rittenhouse, L.R., Coughenour, M.B., Swift, D.M. & Sims P.L. (1996) Mechanisms that result in large herbivore grazing distribution patterns. *Journal of Range Management* **49**, 386–400.
- Blackwell, P.G. (2003) Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika* **90**, 613–627.
- Bovet, P. & Benhamou, S. (1988) Spatial analysis of animals' movements using a correlated random walk model. *Journal of theoretical biology* **131**, 419–433.
- Boyce, M.S., Pitt, J., Northrup, J.M., Morehouse, A.T., Knopff, K.H., Cristescu, B. & Stenhouse, G.B. (2010) Temporal autocorrelation functions for movement rates from global positioning system radiotelemetry data. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 2213–2219.
- Brillinger, D.R. & Stewart, B.S. (1998) Elephant-seal movements: modelling migration. *Canadian Journal of Statistics* **26**, 431–443.
- Calzetta, E., Roura, A. & Verdaguer, E. (2003) Stochastic description for open quantum systems. *Physica A: Statistical Mechanics and its Applications* **319**, 188–212.
- Codling, E.A. & Hill, N.A. (2005) Sampling rate effects on measurements of correlated and biased random walks. *Journal of Theoretical Biology* **233**, 573–588.
- Codling, E.A., Plank, M.J. & Benhamou, S. (2008) Random walk models in biology. *Journal of the Royal Society Interface* **5**, 813–834.
- Cover, T.M. & Thomas, J.A. (2006) *Elements of Information Theory*. Wiley Interscience, Hoboken, NJ.
- Cressie, N. (1993) *Statistics for Spatial Data*, Revised edn. Wiley, New York.
- Dembo, A. (1986) The relation between maximum likelihood estimation of structured covariance matrices and periodograms. *Acoustics, Speech and Signal Processing, IEEE Transactions on* **34**, 1661–1662.
- Dunn, J.E. & Gipson, P.S. (1977) Analysis of radio telemetry data in studies of home range. *Biometrics* **33**, 85–101.
- Fan, J. & Yao, Q. (2003) *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer-Verlag, New York.
- Fleming, C.H. & Calabrese, J.M. (2013) On the estimators of autocorrelation model parameters ArXiv:1301.4968 [stat.ME].
- Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2013) Data from: From fine-scale foraging to home ranges: A semi-variance approach to identifying movement modes across spatiotemporal scales. *Dryad Digital Repository*. doi:10.5061/dryad.45157.
- Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2014) From fine-scale foraging to home ranges: A semi-variance approach to identifying movement modes across spatiotemporal scales. *The American Naturalist* **183**. doi:10.1086/675504.
- Gardiner, C.W. (2009) *Handbook of Stochastic Methods*. Springer-Verlag, Berlin, Heidelberg, New York.
- Gautesstad, A.O. (2012) Animal space use: distinguishing a two-level superposition of scale-specific walks from scale-free Lévy walk. *Oikos* **122**, 612–620.
- Gurarie, E. & Ovaskainen, O. (2011) Characteristic spatial and temporal scales unify models of animal movement. *The American Naturalist* **178**, 113–123.
- Gurarie, E., Andrews, R.D. & Laidre, K.L. (2009) A novel method for identifying behavioural changes in animal movement data. *Ecology Letters* **12**, 395–408.
- Hillery, M., O'Connell, R.F., Scully M.O. & Wigner E.P. (1984) Distribution functions in physics: Fundamentals. *Physics Reports* **106**, 121–167.
- Horne, J.S., Garton, E.O., Krone, S.M. & Lewis, J.S. (2007) Analyzing animal movements using Brownian bridges. *Ecology* **88**, 2354–2363.
- Jiang, Z., Takatsuki, S., Li, J., Wang, W., Ma, J. & Gao, Z. (2002) Feeding type and seasonal digestive strategy of Mongolian gazelles in China. *Journal of Mammalogy* **83**, 91–98.
- Johnson, D.S., London, J.M., Lea, M.-A. & Durban, J.W. (2008) Continuous-time correlated random walk model for animal telemetry data. *Ecology* **89**, 1208–1215.
- Jonsen, I.D., Flemming, J.M. & Myers, R.A. (2005) Robust state-space modeling of animal movement data. *Ecology* **86**, 2874–2880.
- Kareiva, P.M. & Shigesada, N. (1983) Analyzing insect movement as a correlated random walk. *Oecologia* **56**, 234–238.
- Krige, D.G. (1951) A statistical approach to some basis mine valuation problems on the Witwatersrand. *Journal of the Chemical Metal and Mining Society of South Africa* **52**, 119–139.
- Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* **74**, 1659–1673.
- Lomb, N.R. (1976) Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science* **39**, 447–462.
- Marcotte, D. (1996) Fast variogram computation with FFT. *Computers & Geosciences* **22**, 1175–1186.
- Mardia, K.V. & Marshall, R.J. (1984) Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* **71**, 135–146.
- McClintock, B.T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B.J. & Morales, J.M. (2012) A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs* **82**, 335–349.
- Moorcroft, P.R. & Lewis, M.A. (2006) *Mechanistic Home Range Analysis*. Princeton University Press, Princeton, NJ.
- Moorcroft, P.R., Lewis, M.A. & Crabtree, R.L. (2006) Mechanistic home range models capture spatial patterns and dynamics of coyote territories in Yellowstone. *Proceedings of the Royal Society B: Biological Sciences* **273**, 1651–1659.
- Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E. & Fryxell, J.M. (2004) Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology* **85**, 2436–2445.
- Mueller, T., Olson, K.A., Dressler, G., Leimgruber, P., Fuller, T.K., Nicolson, C. *et al.* (2011) How landscape dynamics link individual- to population-level movement patterns: A multispecies comparison of ungulate relocation data. *Global Ecology and Biogeography* **20**, 683–694.
- Nouvellet, P., Bacon, J. & Waxman, D. (2009) Fundamental insights into the random movement of animals from a single distance-related statistic. *The American Naturalist* **174**, 506–514.
- Olson, K.A., Fuller, T.K., Mueller, T., Murray, M.G., Nicolson, C., Odonkhuu, D., Bolortsetseg, S. & Schaller, G.B. (2010) Annual movements of Mongolian gazelles: Nomads in the eastern steppe. *Journal of Arid Environments* **74**, 1435–1442.
- Pawitan, Y. (2001) *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press, Oxford.
- Polansky, L., Wittemyer, G., Cross, P.C., Tambling, C.J. & Getz, W.M. (2010) From moonlight to movement and synchronized randomness: Fourier and wavelet analyses of animal location time series data. *Ecology* **91**, 1506–1518.
- Scargle, J.D. (1982) Studies in astronomical time series analysis. II – Statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal* **263**, 835–853.
- Southwood, T.R. (2000) *Ecological Methods*, 3rd edn. Wiley Blackwell, Hoboken, NJ.
- Swihart, R.K., & Slade N.A. (1985) Testing for independence of observations in animal movements. *Ecology* **66**, 1176–1184.
- Turchin, P. (1998) *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. Sinauer Associates, Sunderland, MA.
- Ville, J. (1948) Théorie et applications de la notion de signal analytique. *Câbles et Transmission* **2**, 61–74.
- Whittle, P. (1953) Estimation and information in stationary time series. *Arkiv for Matematik* **2**, 423–434.
- Worton, B.J. (1989) Kernel methods for estimating the utilization distribution in home-range studies. *Ecology* **70**, 164–168.

Received 30 December 2013; accepted 26 February 2014

Handling Editor: Justin Travis

## Supporting Information

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

**Appendix S1.** Stationary models with non-stationary processes.

**Appendix S2.** Maximizing likelihood.

**Appendix S3.** Analysis of the Markovian likelihood function.

**Appendix S4.** Demonstration of the robustness of non-Markovian ML estimation to gapped data.

**Appendix S5.** Proof of  $\hat{\sigma}_0 = 0$ .

**Appendix S6.** Stationary Gaussian home-range analysis.

**Appendix S7.** Likelihood of Brownian motion.