Ecology Letters, (2009) **12**: 395–408 doi: 10.1111/j.1461-0248.2009.01293.x

LETTER

A novel method for identifying behavioural changes in animal movement data

Eliezer Gurarie,[1]* Russel D. Andrews[2] and Kristin L. Laidre[3]
[1]Department of Biological and Environmental Sciences, University of Helsinki, PO Box 65, FI-00014 Helsinki, Finland
[2]School of Fisheries and Ocean Sciences, University of Alaska, Fairbanks, Alaska SeaLife Center, Seward, AK 99664, USA
[3]Polar Science Center, Applied Physics Laboratory, University of Washington, Seattle, WA 98105, USA
**Correspondence: E-mail: eliezer.gurarie@helsinki.fi*

Abstract

A goal of animal movement analysis is to reveal behavioural mechanisms by which organisms utilize complex and variable environments. Statistical analysis of movement data is complicated by the fact that the data are multidimensional, autocorrelated and often marked by error and irregular measurement intervals or gappiness. Furthermore, movement data reflect behaviours that are themselves heterogeneous. Here, we model movement data as a subsampling of a continuous stochastic processes, and introduce the behavioural change point analysis (BCPA), a likelihood-based method that allows for the identification of significant structural changes. The BCPA is robust to gappiness and measurement error, computationally efficient, easy to implement and reveals structure that is otherwise difficult to discern. We apply the analysis to a GPS movement track of a northern fur seal (*Callorhinus ursinus*), revealing an unexpectedly complex diurnal behavioural profile, and demonstrate its robustness to the greater errors associated with the ARGOS tracking system. By informing empirical interpretation of movement data, we suggest that the BCPA can eventually motivate the development of mechanistic behavioural models.

Keywords

ARGOS data, behavioural heterogeneity, continuous autoregressive model, gappy time series, GPS data, movement ecology, structural shifts.

Ecology Letters (2009) 12: 395–408

INTRODUCTION

In recent years, there has been a rapidly growing body of work devoted to the detailed study of free-ranging animal movements in the wild, reflecting the increase in the technological ability to accumulate data (Ferguson *et al.* 1998; Heide-Jørgensen *et al.* 2003; Block *et al.* 2005; Ropert-Coudert & Wilson 2005; Andrews *et al.* 2008). Many fundamental demographic parameters can be directly related to animal movements, including foraging success, breeding success, migration and dispersal (Guinet *et al.* 2001; Mauritzen *et al.* 2001; Laidre *et al.* 2003; Matthiopoulos 2003; Buskirk & Millspaugh 2006). Individual animal movements are a measurable behavioural response to a combination of internal states, physiological constraints and environmental factors. Ideally, analysis of movement data can yield insights into the behavioural mechanisms that allow organisms to exploit their temporally variable and spatially heterogeneous environments (Schick *et al.* 2008).

Analysis of movement data is far from straightforward primarily because the data are multidimensional and autocorrelated in space and time. A common model of movement is some variety of the correlated random walk (CRW) model (Skellam 1951; Shigesada 1980; Kareiva & Shigesada 1983; Marsh & Jones 1988; Bergman *et al.* 2000; Bartumeus *et al.* 2008), which hypothesizes some distribution of step-lengths and turning angles. Models have been constructed that successfully relate changes in CRW parameters to environmental and landscape features (Morales *et al.* 2004; Forester *et al.* 2007; Aarts *et al.* 2008; Haydon *et al.* 2008). CRWs have also been generalized into diffusion-based approximations used to model movements in complex environments, such as habitat types in patchy landscapes (Ovaskainen 2004; Ovaskainen *et al.* 2008) or within the framework of mechanistic home range models (Moorcroft *et al.* 1999; Moorcroft & Lewis 2006).

There are several important features of movement data that complicate the straightforward application of CRWs. The first is error in the measurement process. A fruitful body of research has emerged recently that addresses measurement error with the use of state-space models (Jonsen *et al.* 2003, 2005; Royer *et al.* 2005; Patterson *et al.*

2008). While these models are effective at separating mechanistic processes from observation error, their effectiveness is limited by prior specification of the movement process and their estimation typically involves computationally intensive Bayesian estimation procedures.

A second important and commonly encountered feature of movement data that confounds the application of CRWs is the irregular timing of measurements, or gappiness in the data. This issue is particularly important for marine species such as pinnipeds, cetaceans, penguins, turtles and large pelagic fish (Guinet *et al.* 2001; Laidre *et al.* 2004; Block *et al.* 2005; Jonsen *et al.* 2007), species that typically spend only a small fraction of their time close to the surface. Not only are the intervals unpredictable, but the quality of data that is obtained while the satellites pass overhead is compromised by the brief periods of emergence. Estimates of turning angle distributions and estimated velocities are necessarily affected by irregular intervals: at very small time intervals, turning angles are all close to zero and displacements are locally linear, whereas at very long time intervals the correlation is completely lost. Typically, researchers deal with irregular sampling intervals by using interpolation to estimate a most likely best location at a fixed interval (Laidre *et al.* 2004; Jonsen *et al.* 2005). This has the consequence of either adding additional complication to a movement process or of sacrificing potentially informative data points.

Independent of data collection issues, a fundamental property of animal movements is behavioural heterogeneity. Over an extended period of observation, a organism's movement is often the result of multiple behavioural modes, such as travelling, searching, feeding, resting, milling. Previous work on modelling animal movements as mixtures of different behaviours have limited the behavioural variability to a few categorical modes (usually 'travelling' and 'foraging') and rely on Bayesian estimation methods to model probabilities of being in any given mode (Jonsen *et al.* 2007; Bailey *et al.* 2008). Alternatively, complex movements are modelled as responses to environmental covariates such as habitat type (Morales *et al.* 2004; Ovaskainen 2004; Ovaskainen *et al.* 2008) and predator density (Forester *et al.* 2007). While both kinds of approaches are versatile and instructive, they both rely on some structural *a priori* assumptions about the movement.

Here we develop a novel, robust and efficient method for identifying behavioural changes in behaviourally heterogeneous and temporally gappy movement data without any prior assumptions. Step lengths and turning angles are transformed into orthogonal persistence and turning components of velocity and characterized as continuous autocorrelated time series described locally by three parameters: a mean, a variance and a continuous autocorrelation. Likelihood estimation can be used to identify moments where the parameter values change significantly, corresponding to shifts in behaviour. By sweeping an analysis window over an entire movement path, an aggregated behavioural summary of movement is obtained. The complete suite of steps is termed a *behavioural change point analysis* (BCPA). We illustrate the BCPA with an application to GPS-based data on the foraging tracks of a northern fur seal (*Callorhinus ursinus* Linneaus).

METHODS

Orthogonal decomposition of movement data

Raw movement data consist of $n + 1$ observations (numbered from $i = 0, 1, 2,...,n$) of absolute positions $\mathbf{Z} = \{\mathbf{X},\mathbf{Y}\}$ collected at times $\mathbf{T}$. Rather than analyse the absolute positions ($\mathbf{Z}$) or absolute compass orientation ($\mathbf{\Phi}$) directly, we examine two variables directly controlled by the organism from its reference point: estimated speeds ($\mathbf{V}$) and turning angles ($\mathbf{\Psi}$). These are obtained directly via: $V(T_i) = ||Z_i - Z_{i-1}||/(T_i - T_{i-1})$ and $\Psi(T_i) = \Phi_i - \Phi_{i-1}$.

We further transform the data by decomposing every speed estimate and turning angle into orthogonal components of persistence velocity $V_\mathrm{p}(t)$ and turning velocity $V_\mathrm{t}(t)$ defined as

$$V_\mathrm{p}(T_i) = V(T_i)\cos(\Psi(T_i)) \quad (1)$$

$$V_\mathrm{t}(T_i) = V(T_i)\sin(\Psi(T_i)) \quad (2)$$

V_p captures the tendency and magnitude of a movement to persist in a given direction while V_t captures the tendency of movement to head in a perpendicular direction in a given time interval. Thus, the primary descriptive features of movement, namely speed, directional persistence and variability are captured in these variables. A further fundamental advantage of these transformations is that the resulting variables are well modelled by stationary, Gaussian, autoregressive time-series models. Empirical explorations of movement data via histograms or normal Q–Q plots suggest that both V_p and V_t are well approximated by mixed normal distributions, with V_t having a mean very close to zero. This appealing statistical property allows for the application of an arsenal of more or less standard time-series techniques for characterizing autocorrelated data.

It should be noted that, though geometrically orthogonal, these two variables are not independent. The eventual technique we present is, however, ultimately descriptive, and as the interpretation of each of these variables is somewhat unique, we chose to analyse them separately. In practice, assessing a scatter plot of the two variables either visually or with some statistical technique is sufficient to determine whether the assumption of independence is supported by the data, as is the case in the data analysed further. A deeper

investigation of the correlation structure between these variables is the subject of future work, and could lead to an informed modelling of the joint process.

Autocorrelated time-series model

The basic assumption underlying the analysis of movement is that the persistence velocity V_p (eqn 1) is a sample from a continuous space–time, stationary Gaussian process $W(t)$ with the following properties:

$$W(0) = W_0, \tag{3}$$

$$E[W(t)] = \mu, \tag{4}$$

$$\mathrm{Var}[W(t)] = \sigma^2, \tag{5}$$

$$\mathrm{Corr}[W(t), W(t-\tau)] = \rho^{\tau}, \tag{6}$$

where $0 < \rho < 1$ is the first order autocorrelation at a time lag 1, i.e. at whatever units the time is measured. (Note, all subsequent discussion applies analogously to V_t, with the additional simplification that $\mu_t = 0$.)

Consider n observations from the continuous process $\mathbf{W} = \{W_0, W_1, W_2, \ldots, W_n\}$ collected at times $\mathbf{T} = \{t_0, t_1, t_2, \ldots, t_n\}$. The observation W_i can be described as

$$W_i = \mu + \rho^{\tau_i}(W_{i-1} - \mu) + \varepsilon_i, \tag{7}$$

where $i \in \{1, \ldots, n\}$, $\tau_i = t_i - t_{i-1}$ is the time interval between subsequent observations, $\rho^{\tau}{}_i$ is the autocorrelation as a function of the time gap and ε_i is a stochastic error term. Given the conditions in eqns 3–6, ε_i can be shown to have mean 0 and variance $\sigma^2(1 - \rho_i^{2\tau})$. The derivation of the variance is obtained as follows:

$$\begin{aligned}\mathrm{Var}[\varepsilon_i] &= \mathrm{Var}[X_i - \rho^{\tau_i} X_{i-1}] \\ &= \sigma^2 + \rho^{2\tau_i}\sigma^2 - 2\rho^{\tau_i}\mathrm{Cov}[X_i X_{i-1}]\end{aligned} \tag{8}$$

Substituting in the correlation structure assumption (6) into (8) yields

$$\mathrm{Var}[\varepsilon_i] = \sigma^2(1 - \rho^{2\tau_i}) \tag{9}$$

The most important statistical assumptions in this modelling of the movement data are the Gaussian error structure and the exponential decay in the autocorrelation. While these assumptions are supported by satellite data collected on large animals in the marine environment and other datasets that we have analysed, they are not fundamental properties of movement. For example, oscillatory or zig-zaggy movements have theoretically negative autocorrelations at the half-period of the oscillation. The assumptions can be readily verified empirically: the former by visual inspection of a normal Q–Q plot and the latter by discretizing the data at an appropriate interval and inspecting a discrete autocorrelation function.

Estimating irregular time-series parameters

Estimates for the mean and variance parameters $\hat{\mu}$ and $\hat{\sigma}$ are given by the mean ($\bar{X}$) and standard deviation (S) of the data respectively. In order to estimate the continuous autocorrelation parameter ρ, a likelihood for the process given the data and the other estimates is needed. Because any W_i depends only on the previous observed value W_{i-1}, the conditional likelihood of ρ given all W is

$$L(\rho|\mathbf{W}, \mathbf{T}, \hat{\mu}, \hat{\sigma}) = \prod_{i=1}^{n} f(W_i|W_{i-1}, \tau_i, \rho, \hat{\mu}, \hat{\sigma}) \tag{10}$$

where $\mathbf{W}$ and $\mathbf{T}$ represent the vector of observations and times of observation respectively, and the distribution function f is the probability density function of the conditional distribution $W_i|W_{i-1}$. For process (7) with a Gaussian error structure, f is given by

$$f(W_i|W_{i-1}) = \frac{1}{\hat{\sigma}\sqrt{2\pi(1-\rho^{2\tau_i})}} \exp\left(\frac{(W_i - \rho^{\tau_i}(W_{i-1} - \hat{\mu}))^2}{2\hat{\sigma}^2(1-\rho^{2\tau_i})}\right) \tag{11}$$

This likelihood (10) is smooth and easily maximized over the range of possible values for $0 < \rho < 1$. The estimate for ρ is denoted as

$$\hat{\rho} = \underset{\rho}{\mathrm{argmax}}\, L(\rho|\mathbf{W}, \mathbf{T}, \hat{\mu}, \hat{\sigma}) \tag{12}$$

The precision and accuracy of this estimate were confirmed with a quick simulation study in which ρ was estimated from a gappy series generated by randomly sampling 100 observations from 100 AR(1) processes with $\rho = 0.8$ and length $n = 1000$ (Fig. 1a,b). The estimate $\hat{\rho}$ had mean 0.797 with standard error 0.052. Further estimates for a wider range of values are summarized in the Simulation study section and Table 2.

The autocorrelation coefficient is a fundamental property of the movement process and readily interpreted in terms of what might be termed 'movement inertia'. A characteristic time scale of movement can be defined, for example, as the expected time duration at which the autocorrelation drops to 0.5, such that

$$t_{1/2} = \frac{\log(0.5)}{\log(\rho)} \tag{13}$$

For example, if the measurement times in the example above were collected in units of hours, the characteristic time scale of the movement is *c.* 3.1 h.

This time scale also provides a range over which the estimation method can be considered valid. As means and variances estimates of time series are affected by the autocorrelation, the method for estimating the parameters presented here is only valid if the length of the entire time

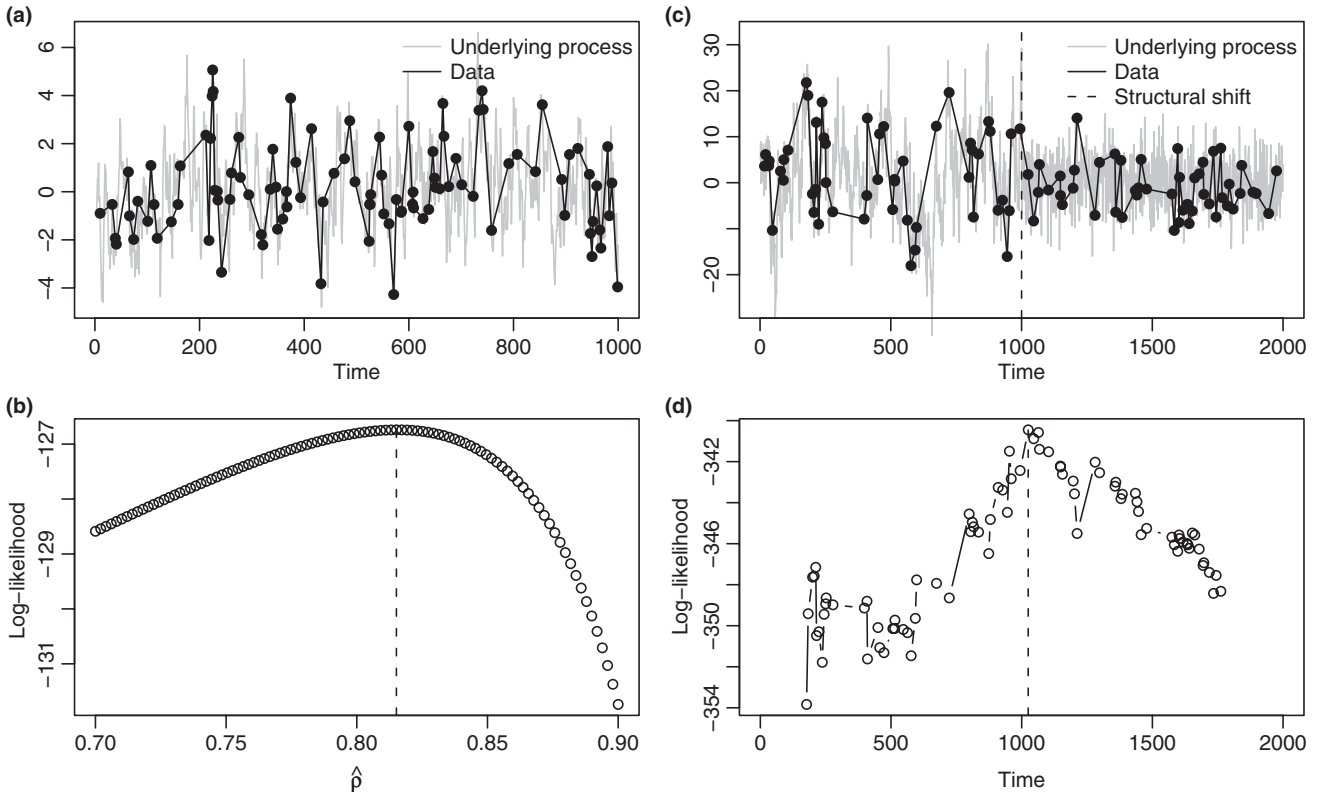

Figure 1 Illustration of the gappy change point estimation method. In the top left plot (a), a simulated AR(1) time series of length 1000 with autocorrelation coefficient $\rho = 0.8$ is plotted in grey, with the black circles indicating a subsampling of 100 points from this time series. Below that (b), the log-likelihood profile of ρ yields an estimated value for ρ of 0.815. The plots on the right (c and d) illustrate estimation of a single-structural shift. The time series in (c) was generated by sampling 50 values from a time series of length 2000 with a breakpoint at $T_{\text{br}} = 1000$. The true parameter values are $\mu_1 = 5$, $\mu_2 = 2$, $\sigma_1 = 10$, $\sigma_2 = 5$, $\rho_1 = 0.9$ and $\rho_2 = 0.2$. In this simulation, the resulting parameter estimates were: $\hat{\mu}_1 = 1.6$, $\hat{\mu}_2 = 0.24$, $\hat{\sigma}_1 = 8.75$, $\hat{\sigma}_2 = 4.73$, $\hat{\rho}_1 = 0.91$ and $\hat{\rho}_1 = 0.35$. The log-likelihood profile over possible change points is in the lower right (d), suggesting an MLCP at $t = 1025$.

series is much greater than the characteristic time scale of movement, i.e. $t_{\max} >> t_{1/2}$. For almost all actual data, this condition holds.

It should, however, be noted that it is possible and tractable to estimate all three parameters μ, σ and ρ simultaneously using the likelihood function (10), and the estimates thus obtained would more correctly be considered maximum likelihood estimates. As long as the temporal range of the data is much longer than the time scale of autocorrelation, however, the differences between the two kinds of estimates are minor, whereas the existence of directly calculated estimates for μ and σ significantly reduces computational intensity. This is an important consideration, as a typical application of the complete analysis requires estimating these parameters tens of thousands of times.

Identifying structural shifts

The likelihood method described above can be expanded to identify structural shifts, or sudden changes in the parameters describing the underlying process. The biological hypothesis underlying this analysis is that behavioural changes can occur discretely and suddenly (e.g. the transition from a goal-oriented travelling mode of movement to a search and forage mode). We refer to the time point at which the structural shift occurs as the 'change point' (CP). The mathematical challenge is to estimate the change point in a gappy time series.

Consider a continuous stochastic process $X(t)$ for $0 < t < T$ defined by parameter set $\Theta(t)$ whose values change at an unknown time point T^* such that

$$\boldsymbol{\Theta}(t) = \begin{cases} \boldsymbol{\Theta}_1 & \text{if } 0 < t \leq T^* \\ \boldsymbol{\Theta}_2 & \text{if } T^* < t \leq T \end{cases} \tag{14}$$

We now take N samples from the continuous process $X(t)$ to obtain a time series X_i at times T_i. If n is defined as the number of measurements within the first regime, such that $T_n \equiv \max(T_i < T^*)$, then the likelihood of the parametrization $(\Theta_1, \Theta_2, T^*)$ given data X_i is the product of the two likelihoods

$$L(\boldsymbol{\Theta}|\mathbf{X},\mathbf{T}) = \prod_{i=1}^{n} f(X_i|X_{i-1},\boldsymbol{\Theta}_1) \prod_{j=n+1}^{N} f(X_j|X_{j-1},\boldsymbol{\Theta}_2) \quad (15)$$

This likelihood can be maximized by sweeping over all possible values of n (from 1 to N), and obtaining the estimates for the remaining parameters at each side of the possible change point using the method described in the previous section.

The parameter estimates can be expressed as

$$\hat{n} = \underset{n}{\operatorname{argmax}}\, L(\boldsymbol{\Theta}|\mathbf{X},\mathbf{T}) \quad (16)$$

$$\hat{\mu}_j = \bar{X}_j \quad (17)$$

$$\hat{\sigma}_j = S_j \quad (18)$$

$$\hat{\rho}_j = \underset{\rho}{\operatorname{argmax}}\, L(\rho|\mathbf{X}_j,\mathbf{T}_j,\hat{\mu}_j,\hat{\sigma}_j) \quad (19)$$

where $j \in (1,2)$ indexes the two regimes, such that if $j = 1$ the data and estimates are taken in the range $i = (1,2,\ldots,\hat{n})$ and if $j = 2$, they are taken from $(i = \hat{n}+1, \hat{n}+2,\ldots,N)$. We term the estimate for the change point $(\widehat{T^*} = t_{\hat{n}})$, the 'most likely change point' (MLCP). The corresponding estimates for the turning component of velocity $V_t = V \sin \psi$ are simplified by the fact that the mean can be assumed to be zero.

Identifying models

Each of the parameters used in this approach to characterize a movement can change, and changes in each parameter correspond to a different behavioural interpretation. Thus, for $V \cos \psi$, an increase in μ corresponds to a combination of faster and more directed movement. An increase in σ indicates more variable movement, e.g. streaky bouts of movement and stopping and sudden changes in direction. A higher ρ indicates more directed and correlated movements, whether fast or slow. For $V \sin \psi$, higher values of σ indicate more turning or tortuous movements, while higher values of ρ indicate longer turning radii. It is consequently of great interest to be able to identify which, if any, of the parameters actually change at the MLCP.

There are eight possible models to consider when analysing a change point. We label these M0–M7 and define them as follows: M0 is the null hypothesis of no significant change ($\mu_1 = \mu_2$, $\sigma_1 = \sigma_2$ and $\rho_1 = \rho_2$); M1, M2 and M3 represent one parameter changing ($\mu_1 \neq \mu_2$, $\sigma_1 \neq \sigma_2$ and $\rho_1 \neq \rho_2$ respectively) while the other two parameters remain constant; M4, M5 and M6 have one equality each while two other parameters change (μ and σ, μ and ρ, σ and ρ respectively); and M7 is the most alternate hypothesis, in which all parameter values change at the MLCP.

Because the conditional likelihood is well defined, we can apply an information-based criterion to select our models. Two such criteria are the AIC (Akaike's Information Criterion) and BIC (Bayesian Information Criterion) defined as

$$I_A(\mathbf{X},\mathbf{T}) = -2n\log(L(\boldsymbol{\Theta}|\mathbf{X},\mathbf{T})) + 2d \quad (20)$$

$$I_B(\mathbf{X},\mathbf{T}) = -2n\log(L(\boldsymbol{\Theta}|\mathbf{X},\mathbf{T})) + d\log(n), \quad (21)$$

where $L(\boldsymbol{\Theta}\,|\,\mathbf{X},\mathbf{T})$ is the likelihood (15) and d is the number of parameters in each of the eight models, ranging from $d = 4$ for M0 to $d = 7$ for M7. The value of the criterion with the more negative value is selected. The BIC is generally more conservative than the AIC with respect to model complexity.

Simulation study

Figure 1c,d illustrates an example of the single change point estimation routine. The simulated data in this example come from an underlying process of length $N = 2000$ with a discrete structural shift at $t = 1000$ where the mean drops from 5 to 0, the variance decreases from 10 to 5 and the autocorrelation coefficient drops from 0.9 to 0.2. This simulation represents a shift from a highly variable, positively biased, correlated process associated with directed movement to a zero mean, lower variance, less-correlated process that mimics foraging. We randomly sampled 100 points from the complete time series and obtained the MLCP according to the method described above. The resulting estimated parameters are listed in the figure caption.

The plots of the log-likelihood (Fig. 1d) give an idea of how precise the T^* estimate is. It should be noted that the time series in this example is rather gappy, only 100 points were sampled out of 2000, and the location of the change point is difficult to discern by eye.

We explored the properties of this experiment for a variety of parameter value changes by simulating 100 gappy processes (true data $N = 400$, subsampled data $n = 50$, $T^* = 200$) for each of eight different models (see Table 1) ranging from the null model of no change in parameter values to the most alternate model of change in all parameter values. Resulting parameter estimates appear unbiased and precise (Table 2), all falling well within one standard deviation of the true value. Perhaps most importantly, the MLCP estimates are highly accurate, with means between 197 and 203. We also applied AIC and BIC to assess the eight models and report their results (Table 2). The BIC performs far better than the AIC at identifying the true model, with, notably, a 78% rate of correctly identifying the null hypothesis and 72–94% probabilities of identifying parameters with a single-parameter changing.

Table 1 Parameter values in a simulation study used to explore the accuracy and power of the single structural break estimates and the model selection criteria

	μ_1	μ_2	σ_1	σ_2	ρ_1	ρ_2
S0	0	0	1	1	0.5	0.5
S1	**−1**	**1**	1	1	0.5	0.5
S2	0	0	**0.5**	**2**	0.5	0.5
S3	0	0	1	1	**0.2**	**0.9**
S4	**−1**	**1**	**0.5**	**2**	0.5	0.5
S5	**−1**	**1**	1	1	**0.2**	**0.9**
S6	0	0	**0.5**	**2**	**0.2**	**0.9**
S7	**−1**	**1**	**0.5**	**2**	**0.2**	**0.9**

Time series of length 400 were generated with a breakpoint at $T_{\text{br}} = 200$ for eight models, ranging from a null model of no change in μ, σ and ρ, to a model where all three parameters change value. Boldface indicates the parameters that undergo a shift. From these series, 50 points were randomly selected, generating gappiness, and estimates were obtained using the method described in the text. The process was repeated 100 times for each of the eight models. The results are presented in Table 2.

For those models where two parameters shift, BIC tends to falsely favour the most complex model. In contrast, AIC performs poorly, never selecting the null model correctly and in other cases almost always choosing more complex models than necessary. The power of the BIC-based model selection is weakened somewhat for very short time series ($n = 30$) and improves for longer ($n = 80$) time series (Fig. 2a,c).

When analysing actual data, a researcher may have greater interest in identifying the location and direction of significant, detectable structural shifts than in estimating the parameters themselves. For this reason, we prefer the more conservative criterion. A complete power analysis is very difficult to perform for the selection mechanism, as the ability of the method to correctly identify the model depends in complicated ways on the magnitude of the difference between the parameters, the extent of the gappiness in the dataset and the length of the series. However, an effort was made in the simulation study to look at differences in parameter values that are comparable to or smaller than those observed in real movement data (see application below). The ability of the BIC to pick the appropriate level of complexity inspires confidence in this method of model selection.

Multiple change points

Heretofore, we have only discussed the identification of a single most likely breakpoint. Within a single track, an organism likely exhibits multiple behaviours. Rigorously, a continuous process $W(t)$ can be defined such that any interval between $0 < t < \tau_m$ is defined by parameter set $\Theta(t)$ whose values change at unknown time points $\mathbf{T} = \{T_1,..., T_m\}$. Thus,

$$\Theta(t) = \begin{cases} \Theta_1 & \text{if } 0<t \leq T_1 \\ \Theta_2 & \text{if } T_1<t \leq T_1 \\ \vdots & \vdots \\ \Theta_m & \text{if } T_{m-1}<t \leq T_m \end{cases} \quad (22)$$

Estimating the parametrization Θ for the multiple change point model is a computationally non-trivial problem that has been discussed in the literature, though only for regularly sampled time series and with applications primarily to financial markets (Chib 1998; Hawkins 2001). An important complication is determining the number of change points to select in a long, structurally complex process. Our solution to this problem is to pass a window over the complete time series and apply the single change point analysis described above at each window. The complete procedure can be summarized as follows:

Table 2 Parameter estimates obtained from the simulation experiment described in text

	T_{br}	$\hat{\mu_1}$	$\hat{\mu_2}$	$\hat{\sigma_1}$	$\hat{\sigma_2}$	$\hat{\rho_1}$	$\hat{\rho_2}$	N_{A}	N_{B}
S0	206 (113)	0.0 (0.21)	0.0 (0.17)	1.0 (0.11)	1.0 (0.11)	0.48 (0.16)	0.47 (0.15)	0	78
S1	197 (13.2)	−1.0 (0.18)	1.0 (0.19)	1.2 (0.21)	1.2 (0.22)	0.59 (0.25)	0.63 (0.19)	1	84
S2	203 (10.3)	−0.03 (0.12)	−0.04 (0.27)	0.50 (0.04)	2.0 (0.19)	0.46 (0.16)	0.47 (0.16)	28	72
S3	201 (26.1)	−0.02 (0.22)	0.04 (0.43)	1.0 (0.08)	1.0 (0.12)	0.23 (0.21)	0.92 (0.056)	29	92
S4	199 (5.49)	−1.0 (0.08)	1.1 (0.33)	0.5 (0.04)	2.0 (0.17)	0.47 (0.2)	0.53 (0.17)	22	40
S5	199 (12.0)	−0.96 (0.17)	1.0 (0.51)	1.0 (0.12)	1.0 (0.14)	0.19 (0.18)	0.92 (0.033)	7	15
S6	200 (8.51)	0.0 (0.09)	0.02 (0.93)	0.49 (0.05)	2.0 (0.17)	0.20 (0.19)	0.92 (0.036)	22	40
S7	201 (15.6)	−0.99 (0.08)	0.94 (0.87)	0.50 (0.05)	2.0 (0.18)	0.19 (0.18)	0.93 (0.033)	97	97

True values are tabulated in Table 1. Nearly all estimates are within one standard error of the true value. The last two columns (N_{A} and N_{B}) are the number of times that AIC and BIC respectively select the correct model out of 100 attempts. The experiment was repeated for the BIC with 30 and 80 points, as illustrated in Fig. 2.

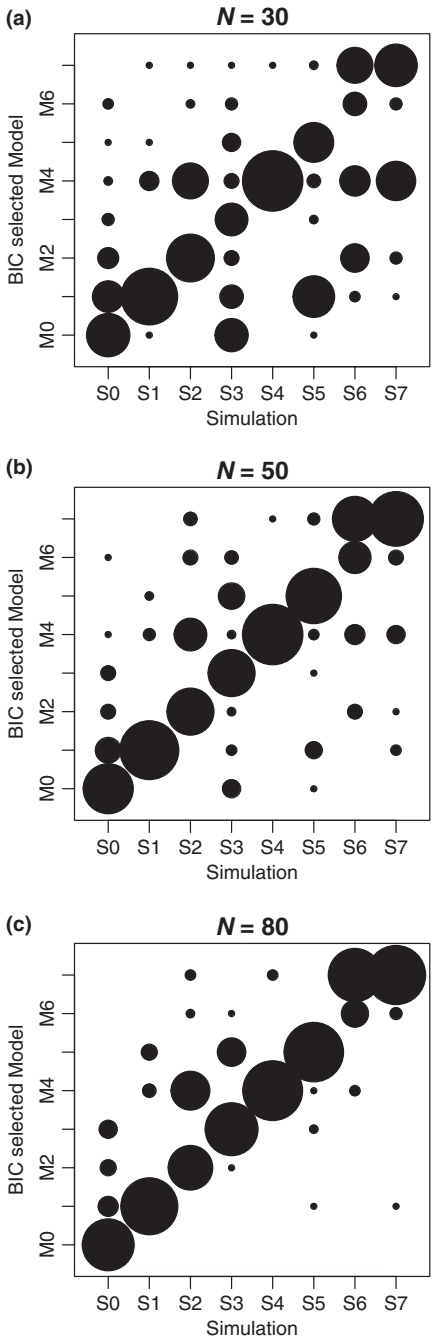

Figure 2 Results of model selection simulation using BIC for gappy time series of length $N = 30$, 50 and 80, (Fig 2a, b and c respectively) subsampled from a true process of length $T = 200$. The simulated time series are parameterized as in Table 1, ranging from the null model (S0) of no difference before and after the gap to the extreme model (S7) with differences in all three parameters. Gappy time series were simulated for each parameter set and the BIC used to select a most parsimonious model (M0–M7). The area of the circles is proportional to the number of times a model is selected for a given simulation out of 100 repetitions. Larger circles along the *x*–*y* diagonal indicate more correct model selections.

(1) Select a window of length $30 \leq l < N$.
(2) Find the MLCP in a subsample of the data $X_1,...,X_l$.
(3) Use the BIC criterion to accept or reject the 'significance' of the change point for each of the parameters μ, σ and ρ.
(4) Based on the result of the test, record the location of the behavioural change point and the resulting estimated parameter values: $\hat{\mu}_1$, $\hat{\sigma}_1$, $\hat{\rho}_1$, $\hat{\mu}_2$, $\hat{\sigma}_2$ and $\hat{\rho}_2$.
(5) Shift the window forward by one data point, and repeat.

Throughout the running of this analysis, the estimates according to the model chosen by the BIC are stored. Thus, if M0 is chosen, no significant shifts are detectable and a single value is reported for each of the parameters for the entire window. If M7 is chosen, we separately estimate parameters at each side of the MLCP. All of the estimates are logged and averaged after the sweep is performed.

This analysis has the important innovation of being able to capture any range of behaviours determined by location in the three-dimensional parameter space of μ, σ and ρ. By returning and averaging the model estimated parameters at every time step multiple times, both gradual shifts in behavioural parameters and potentially dramatic jumps can be identified. The aggregated information including both the parameter estimates, the significant change points and their nature and direction (i.e. increasing μ, decreasing ρ, etc.) can be considered a distilled behavioural model of movement. We term the aggregated output, with estimated change points, model specifications and recorded averaged parameter estimates, a *behavioural change point analysis of animal movement data*, or BCPA.

It should be noted that the appropriate window size is the sole parameter specified by the user. The greater the size of the window, the greater the power of the model selection (Fig. 2) at a cost of identifying smaller-scale behavioural shifts. A smaller window size will reveal finer-scale structure in the data, but the risk of spurious change points increases. Based on the results of the BIC simulation and experimentation with datasets, a window size of 30 is the minimum recommended window size for obtaining reasonable statistical power for model selection.

Essential sections of code in the R programming language (R Development Core Team 2008) are provided in Appendix S1.

APPLICATION TO DATA: NORTHERN FUR SEAL

We applied the BCPA to global positioning system (GPS) data collected from a female northern fur seal tagged in summer 2007 in the central Kuril Islands (Fig. 3). The female was instrumented with a Fastloc GPS data-logging device (MK10-F; Wildlife Computers Inc., Redmond, WA, USA). The tag allows for quick satellite fixes and high

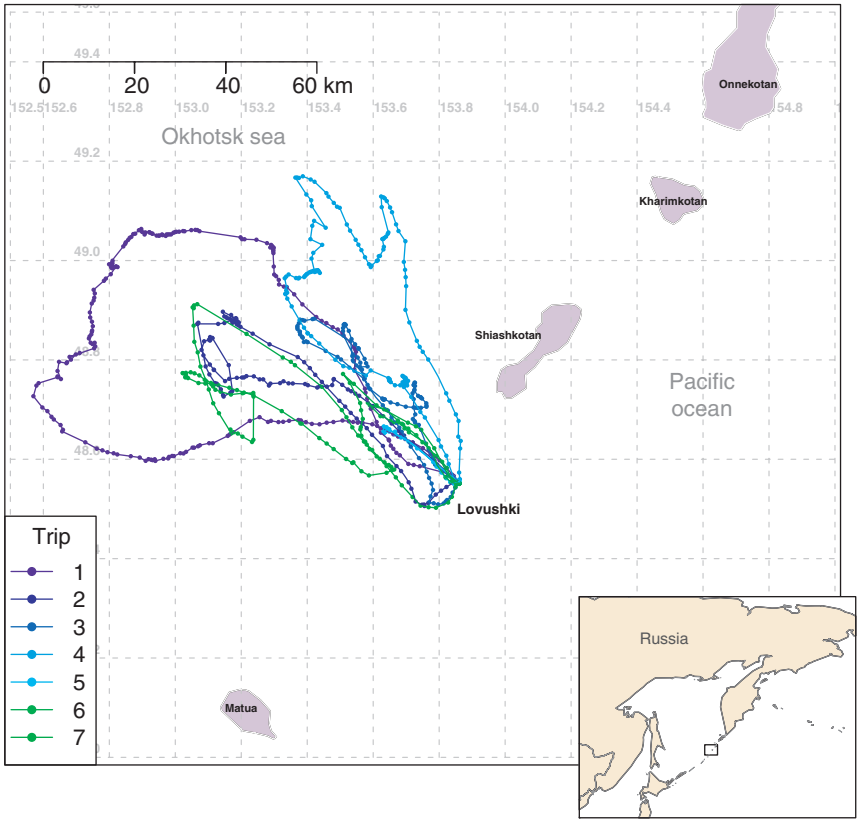

Figure 3 Trips taken by Northern fur seal female (NFS07-03) in Summer 2007 in the Kuril Islands, Russia. The square in the inset map indicates the study area.

accuracy, with at least 60% of the locations within 100 m (Bryant 2007).

The animal was monitored for 38 days, during which seven foraging trips occurred. A total of 763 position fixes were obtained over all seven trips, with individual trips ranging from 30 to 205 position fixes each. The time intervals between the fixes ranged widely from less than a minute to 700 min, with the majority (over 80%) between 15 and 45 min. The data were filtered to exclude implausible swim speeds (> 11 km h^{-1}). Geographic positions were converted to displacements in kilometres, velocities were estimated by dividing displacements by time intervals and turning angles were calculated directly from the positions.

The BCPA was applied to five trips (those with more than 70 location fixes) with window sizes of 30 and 50. The results presented here (Figs 4–6) are for the 30 datapoint window, which covered an average period of 18 h. The output of analyses with window size 50, corresponding to about a 24-h period on average, provided fewer and more robust change points, but masked changes that occurred more frequently than once a day.

In order to more clearly describe the complex output of the BCPA, it is worthwhile to look closely at a few trips. Trip 1 is the longest (5.1 days long) and furthest (maximum distance from rookery 96 km, Fig. 4), as is typical for female fur seals taking their first feeding trip after a fasting period immediately post-parturition. The initial departure from the rookery was marked by high values of V_p ($\hat{\mu}$, *c.* 5 km h^{-1}) and a high estimated per-hour autocorrelation $\rho = 0.36$. Note, because the time data are in hours, $\hat{\rho}$ estimates what the first order autocorrelation coefficient would be if the movement data were collected at exactly 1-h intervals without gaps.

The first significant change point was selected as the MLCP by all the windows that passed over it. The models chosen by the BIC were split about evenly between Model 4 (M4: μ and σ change) and Model 7 (M7: all three parameters change), with all three parameters values decreasing. Similarly strong support was given for a change point that occurred at 19:00 h before the last night of the trip (Fig. 4b,c), where values for all three parameters increased significantly. Referencing this change point against the track (Fig. 4a) indicates that this final change point is associated with a sudden turn south, and a fairly correlated, moderately fast (*c.* 3 km h^{-1}) return journey to the rookery.

The interim period between these two travelling bouts was marked by lower autocorrelation (bluer colours between CP1 and CP2 in Fig. 4a,b), < 0.14 during the entire period. The mean persistence velocity and standard deviations,

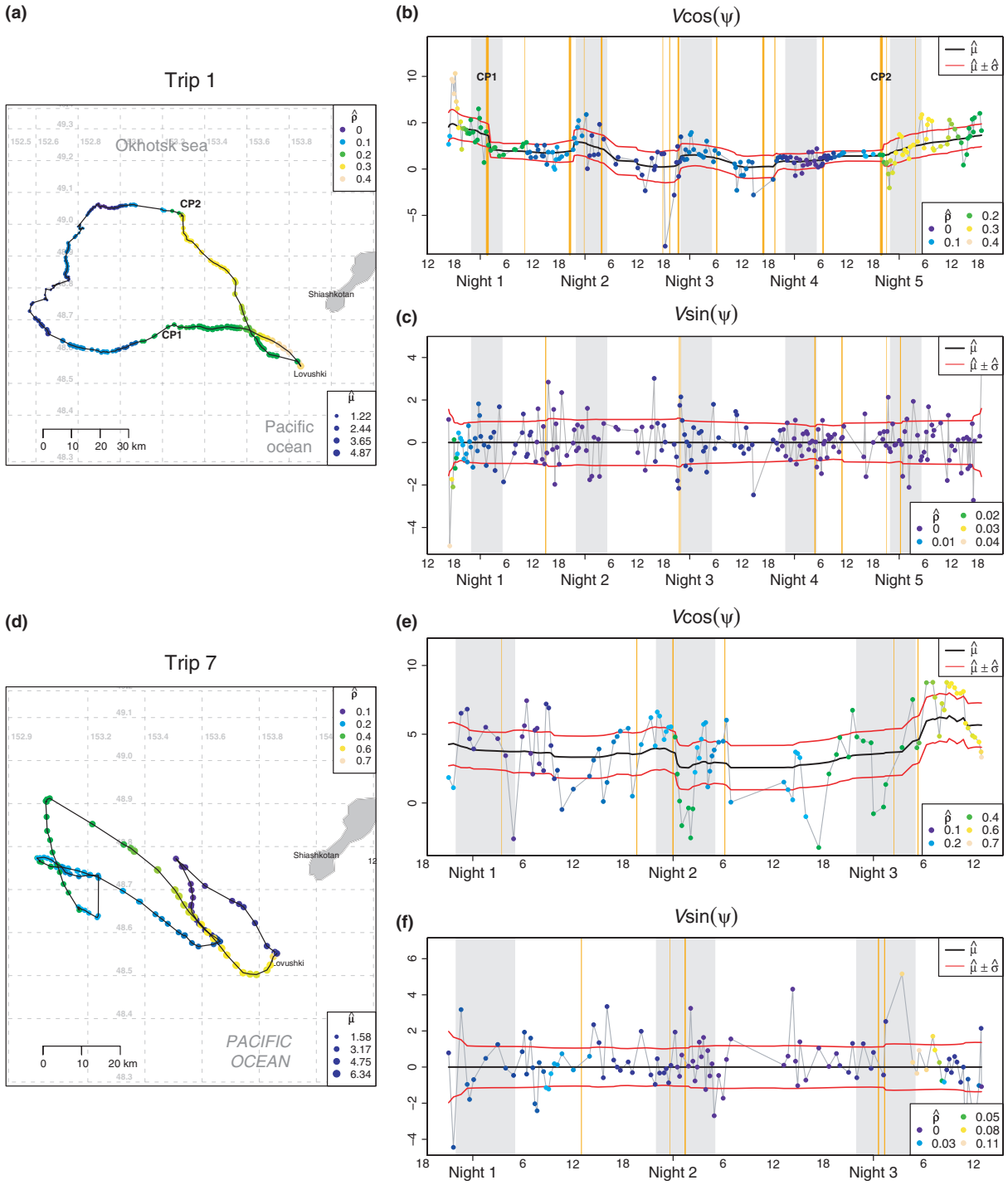

Figure 4 Plots of northern fur seal (NFS07-03) foraging trips 1 and 7 and BCPA output. The tracks are in (a) and (d) respectively. Plots (b, c) and (e,f) show the time series of velocity and turning angles decomposed into persistence ($V_p = V\cos\psi$) and turning ($V_t = V\sin\psi$) components. The black line is the estimate for the mean $\hat{\mu}$, the red line represents the estimate for the standard deviation $\hat{\sigma}$ around the mean, and the colour reflects the autocorrelation $\hat{\rho}$: bluer colours indicate less autocorrelated movement, while yellow colours indicate more autocorrelated movement. Vertical orange lines indicate change points. Thicker lines correspond to a higher number of selected shifts. Wide grey bands indicate nighttime. In the map of the tracks (a and d), colours again indicate the autocorrelation estimate for the persistence component (V_p), while the size of the dots is proportional to the estimated local mean value V_p. The two change points discussed in the text for trip 1 (CP1 and CP2) are pointed out on the respective plots. The window size used in this analysis was 30.

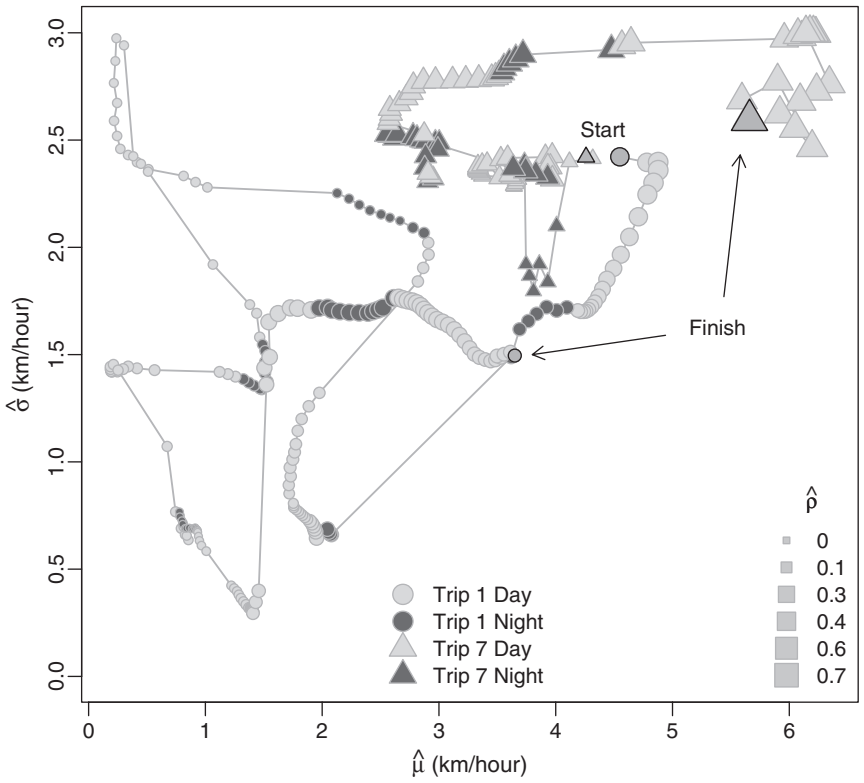

Figure 5 A mapping of all three parameters for $V_p = V \cos \psi$ for trips 1 (circles) and 7 (triangles). Each point represents the three-parameter estimates associated with each observation: $\hat{\mu}$ on the x-axis, $\hat{\sigma}$ on the y-axis and $\hat{\rho}$ proportional to the area of the points. Darker and lighter shapes represent night-time and daytime observations respectively. The connecting lines illustrate the progress of the walk through the parameter space, beginning and ending at the labelled points. The longer connecting segments are associated with the most significant structural shifts in the behaviours.

however, varied considerably and significant change points were identified at dusk and dawn on all three interim nights. Mean velocities tended to increase at night, while standard deviations and autocorrelations were generally constant.

The analysis of turning velocity V_t for Track 1 proved less informative than the analysis of V_p with relatively few significant change points identified (Fig. 4c). The general patterns of trip 1, i.e. higher velocities and autocorrelations during the outbound and inbound portions of the foraging trip and higher general activity at night, were also reflected in trips 2, 3 and 4.

Trip 7 (Fig. 4d–f) displayed a markedly different pattern. Notably, the fur seal's persistence velocity was much higher (an aggregate V_p mean of 3.8 km h^{-1} compared to 2.0 km h^{-1} for trip 1, see Fig. 6) and there were fewer highly significant change points. The fur seal left the rookery in the evening, at high velocities but low autocorrelation and made several hairpin-like changes in direction identified by the MLCPs. High autocorrelations (up to 0.11) in the turning velocity (Fig. 4f) on the return portion of the trip reflects the large turning radius as the fur seal adjusted its trajectory.

The greater values for all three parameters in trip 7 are visible in the behavioural phaseplot (Fig. 5). The phaseplot further summarizes the complex range of behaviours demonstrated by the fur seal within any given single track. Kernelized distribution estimates of aggregated BCPA parameter outputs (Fig. 6) indicate bimodality in the persistence velocity means and, to a lesser extent, in the persistence velocity autocorrelation. Furthermore, aggregated differences between night- and daytime movement parameters suggest that movement during the day is generally faster and more autocorrelated.

We used the time to half-correlation $t_{1/2}$ (eqn 13) to characterize the time scale of the movement process. For trip 1, this value ranged between 30 and 40 min for the travelling modes, and between 4 and 15 min during the remainder of the trip. Similar patterns held broadly for the remaining trips.

DISCUSSION

Top marine predators are extremely well adapted to exploiting the heterogeneous environment of the open ocean to fulfil their survival needs. Their movements are complicated, measurable manifestations of behavioural responses to environmental cues and physiological constraints. In practice the data collected on movement is an irregular and often imprecise subsampling of a structurally complex, continuously autocorrelated, multidimensional process. Analysis of such data requires a method that is robust, flexible, holds few assumptions and outputs biologically interpretable and useful synthesis, requirements that are largely satisfied by the BCPA.

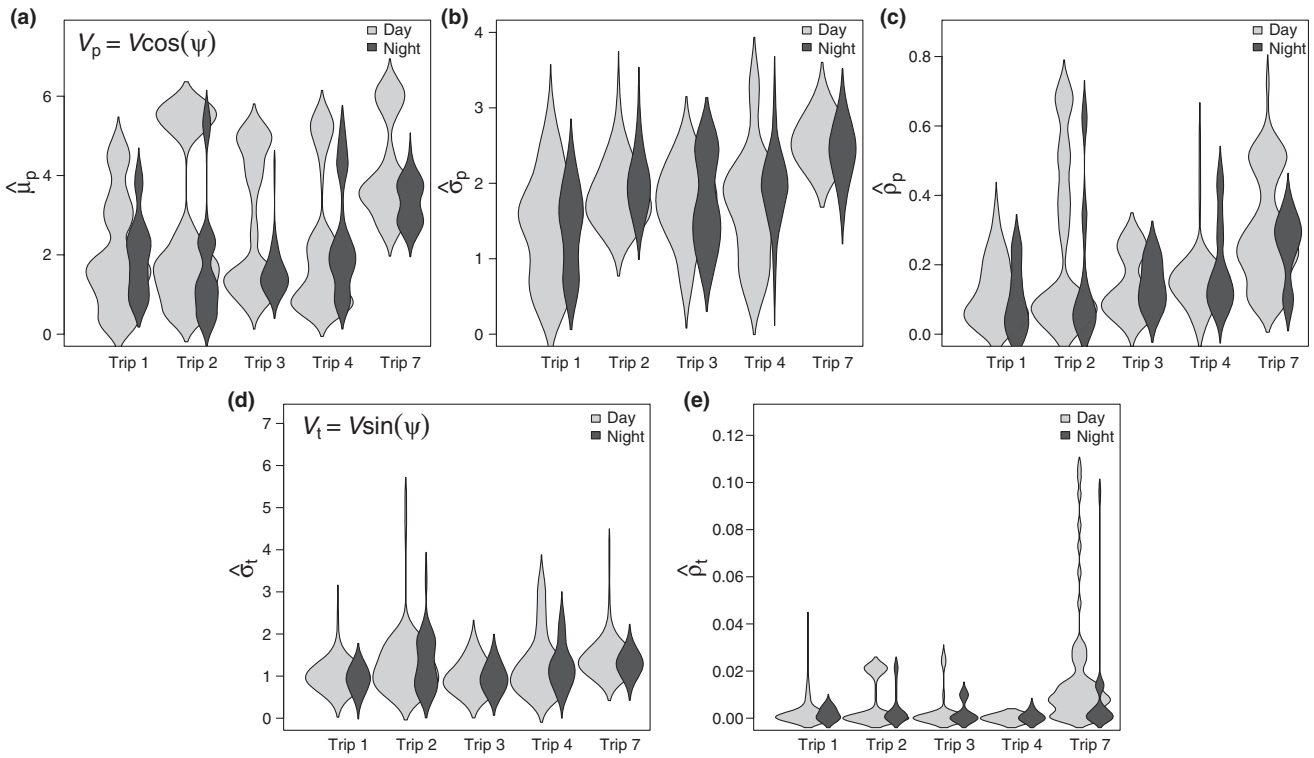

Figure 6 Violinplots showing kernelized distribution estimates of all parameter values for each of the five analysed trips. Darker violins represent nighttime values, while lighter violins represent daytime values. The top row represents the three paramaters used to model the persistence velocity $V\cos\psi$ respectively: (a) μ_p, (b) σ_p and (c) ρ_p, while the lower row represents estimates of the parameters describing the turning velocity $V\sin\psi$: (d) σ_t and (e) ρ_t. Violinplots allow for a visualization of the multimodal distributions, evident particularly in the (a) μ_p, and (c) ρ_p estimates, for which higher values are more frequent during the daytime than at night.

The fundamental contribution of the BCPA is the ability to detect and characterize significant behavioural shifts without *a priori* assumptions. The framework permits not only the identification of discrete shifts, but also the detection of gradual changes in the parameter values. While similar movement analyses have presupposed a few distinct behaviours and classified behaviour into categories (Morales *et al.* 2004; Forester *et al.* 2007; Jonsen *et al.* 2007; Bailey *et al.* 2008), the BCPA reveals striking complexity in behavioural modes, both within a single movement track and in comparisons between tracks of a single individual (Fig. 5).

Autocorrelation is an important and often overlooked intrinsic feature of movement data, and its rigorous estimation and accounting in our method is an important innovation. The autocorrelation is readily interpreted in terms of time scales at which an organism's basic movement patterns (persistence and turning) change. For example, the most significant distinction between the travelling modes and non-travelling behavioural modes in the northern fur seal tracks was associated with a greater increase in the autocorrelation than in the persistence means (Fig. 4). The associated characteristic time scales of movement are 30–40 min compared to 4–15 min for travelling and searching respectively. The existence of modes of movement with high mean persistence but low autocorrelation indicates a fast and active searching mode likely associated with rapid but erratic movements, punctuated by dives, turns and loops, but covering a large average distance. This is contrasted both with the slower, uncorrelated movements associated with drifting or milling and the more correlated, but not necessarily faster, oriented movement associated with directed travel.

In principle, the BCPA should be able to pick out behavioural shifts in more error-ridden data as long as their changes are greater than the noise. For example, much remotely tracked animal movement data are obtained via the ARGOS system (Ferguson *et al.* 1998; Heide-Jørgensen *et al.* 2003; Block *et al.* 2005; Ropert-Coudert & Wilson 2005). ARGOS data is reported with location classes (3, 2, 1, 0, A and B) associated with spatial errors varying from 100 m to several kilometres (Hays *et al.* 2001; Vincent *et al.* 2002; CLS Argos 2008). In a typical dataset on marine organism movements, over 80% of the locations have errors greater than a kilometre.

In order to test the robustness of the procedure against ARGOS-like error, we simulated a behaviourally

heterogeneous movement process that emulated the behaviour of narwhals (*Monodon monoceros*) based on results reported and quantified using fractal dimension by Laidre *et al.* (2004). Movement data were generated by simulating a high resolution (15-min intervals) dataset with four different behaviours ranging from highly tortuous, slower feeding like movement to high-speed migratory-type movement over a 100-day period. The entire time series of observations was randomly subsampled and ARGOS-like errors were attributed to the simulated data according to proportions characteristic of real narwhal movement data. The BCPA was capable of identifying all three breakpoints using the higher location classes 3, 2 and 1 (accounting for only 140 of 1200 datapoints) and estimated most parameter values within one standard deviation. The addition of LC 0 locations, while increasing the number of data points to 422, made the data too noisy to detect any changes. However, a very coarse filtering of the data that eliminated velocities > 250 km day^{-1} allowed for an identification of the largest shift from migration to tortuous movements. The actual parameters values, in particular ρ and σ, were less precisely estimated with the error; however, the direction of the behavioural shifts was well captured.

This experiment suggests that the BCPA can identify large-scale shifts in movement processes that occur on the order of months and over distances of hundreds of kilometres and are measured with ARGOS-like error. While the method works best with the top three locations classes, it is possible to include LC0 if some filtering is applied. Thankfully, many sophisticated filtering methods exist and are routinely applied (Coyne & Godley 2005). The BCPA, however, does not require any further interpolation or discretization of the data.

The appropriate application of the BCPA depends on the temporal scale, resolution and error structure of the data. The assumptions that underlie the continuous time-series model need to be tested and a window size selected that balances the minimum temporal scope within which changes are expected against the desired power. Perhaps most importantly, meaningfully interpreting the output of the analysis depends very much on the goals of the application. In the fur seal example, we present several ways to visualize complex analysis output in order to highlight the high level of behavioural heterogeneity within a single animal's movement profile. In other applications, single-parameter values can be simplified by estimating the parameters between the most significant behavioural change points, or narrowing the question. For example, dates at which feeding or breeding behaviour of whales turns into migratory behaviour can be almost unambiguously identified and compared between individuals or years.

It should be reemphasized that the BCPA provides an empirical, descriptive distillation of movement data rather than an explanatory model of the factors that influence movement behaviour. Its power lies in revealing underlying structures in movement data that are otherwise difficult to discern. Deeper analysis of behavioural processes would relate movement to environmental covariates and associated behavioural measurements. For marine mammals, for example, diving behaviour and foraging success can occasionally be measured, and information can be obtained about bathymetry, currents, ice cover or other potentially relevant environmental factors. Once a movement profile is distilled and parallel time series of observed behaviours and immediate environmental cues are obtained, sophisticated mechanistic models can be developed incorporating such processes as foraging strategies, bioenergetic constraints and intrinsic forcings such as homing and orientation or, as in the case of the northern fur seal, the need to return to a colony and nurse.

While the development of the BCPA was motivated primarily by addressing some of the practical limitations of data on marine organisms, the underlying autocorrelated change point model can be productively implemented on terrestrial movements where missing data is less of an issue. Indeed, because the terrestrial environment is much better mapped and observable than the marine environment, obtaining useful parallel time series is generally easier. Thus, the behavioural phases of a large herbivore or carnivore can be compared to various kinds of habitat types and seasonal variables. In this context, the BCPA can be used to empirically explore the validity of some of the structural hypotheses that underly mechanistic movement models such as those of Moorcroft *et al.* (1999), Ovaskainen (2004), Morales *et al.* (2004) and Forester *et al.* (2007). Given an appropriate hierarchical framework, models that hypothesize responses in movement behaviour to certain stimuli can be estimated, compared and assessed against the output of the BCPA. These are all important avenues for future work.

On a final note, the BCPA can be adapted to a variety of ecological data not directly related to movement. Many time series in ecology are opportunistically and irregularly sampled, yet practical recommendations in the ecological literature on analysing gappy time series are scant. The methods described here provide a versatile, robust and efficient framework with which to investigate a wide range of temporally heterogeneous autocorrelated processes, regardless of the regularity of sampling intervals.

ACKNOWLEDGEMENTS

Discussions with H. Nesse, T. Gneiting and V. Minin provided valuable insights into the analysis. The data from the Russian Far East were collected with the support of

V. Burkanov and a host of field assistants and with funding provided by a grant from NOAA to the Alaska SeaLife Center. The manuscript benefited greatly from the suggestions of two anonymous referees.

REFERENCES

Aarts, G., MacKenzie, M., McConnell, B., Fedak, M. & Matthiopoulos, J. (2008). Estimating space-use and habitat preference from wildlife telemetry data. *Ecography*, 31, 140–160.

Andrews, R.D., Pitman, R.L. & Ballance, L.T. (2008). Satellite tracking reveals distinct movement patterns for Type B and Type C killer whales in the southern Ross Sea, Antarctica. *Polar Biol.*, 31, 1461–1468.

Bailey, H., Shillinger, G., Palacios, D., Bograd, S., Spotila, J., Paladino, F. *et al.* (2008). Identifying and comparing phases of movement by leatherback turtles using state-space models. *J. Exp. Mar. Biol. Ecol.*, 356, 128–135.

Bartumeus, F., Catalan, J., Viswanathan, G.M., Raposod, E.P. & da Luze, M.G.E. (2008). The influence of turning angles on the success of non-oriented animal searches. *J. Theor. Biol.*, 252, 43–55.

Bergman, C.M., Schaefer, J.A. & Luttich, S.N. (2000). Caribou movement as a correlated random walk. *Oecol*, 123, 364–374.

Block, B.A., Teo, S.L.H., Walli, A., Boustany, A., Stokesbury, M.J.W., Farwell, C.J. *et al.* (2005). Electronic tagging and population structure of Atlantic bluefin tuna. *Nature*, 434, 1121–1127.

Bryant, E. (2007). *2D Location Accuracy Statistics for Fastloc RCores Running Firmware Versions 2.2 & 2.3, Technical Report TR01.* Technical Report, Wildtrack Telemetry Systems Ltd. http://www.wildtracker.com/results_files/TechnicalReportTR01.pdf.

Buskirk, S.W. & Millspaugh, J.J. (2006). Metrics for studies of resource selection. *J. Wildl. Manage.*, 70, 358–366.

Chib, S. (1998). Estimation and comparison of multiple change-point models. *J. Econom.*, 86, 221–241.

CLS Argos (2008). *Argos User*'s Manual. Available at: http://www.argos-system.org/manual/ (accessed 14 October 2008).

Coyne, M.S. & Godley, B.J. (2005). Satellite Tracking and Analysis Tool (STAT): an integrated system for archiving, analyzing and mapping animal tracking data. *Mar. Ecol. Prog. Ser.*, 301, 1–7.

Ferguson, S.H., Taylor, M.K., Born, E.W. & Messier, F. (1998). Fractals, sea ice landscape and spatial patterns of polar bears. *J. Biogeogr.*, 25, 1081–1092.

Forester, J.D., Ives, A.R., Anderson, D.P., T, M.G., Fortin, D., Beyer, H.L. *et al.* (2007). State-space models link elk movement patterns to landscape characteristics in yellowstone national park. *Ecol. Monogr.*, 77, 285–299.

Guinet, C., Dubroca, L., Lea, M., Goldsworthy, S., Cherel, Y., Duhamel, G. *et al.* (2001). Spatial distribution of foraging in female Antarctic fur seals Arctocephalus gazella in relation to oceanographic variables: a scale-dependent approach using geographic information systems. *Mar. Ecol. Prog. Ser.*, 219, 251–264.

Hawkins, D.M. (2001). Fitting multiple change-point models to data. *Comput. Stat. Data Anal.*, 37, 323–341.

Haydon, D.T., Morales, J.M., Yott, A., Jenkins, D.A., Rosatte, R. & Fryxell, J.M. (2008). Socially informed random walks: incorporating group dynamics into models of population spread and growth. *Proc. R. Soc. Lond., B, Biol. Sci.*, 275, 1101–1109.

Hays, G.C., Arkesson, S., Godley, B.J., Luschi, P. & Santidrian, P. (2001). The implications of location accuracy for the interpretation of satellite-tracking data. *Anim. Behav.*, 61, 1035–1040.

Heide-Jørgensen, M.P., Dietz, R., Laidre, K.L., Richard, P., Orr, J. & Schmidt, H.C. (2003). The migratory habits of narwhals. *Can. J. Zool.*, 81, 1298–1305.

Jonsen, I.D., Myers, R.A. & Flemming, J.M. (2003). Meta-analysis of animal movement using state-space models. *Wildl. Res.*, 21, 149–161.

Jonsen, I.D., Myers, R.A. & Flemming, J.M. (2005). Robust state-space modeling of animal movement data. *Ecology*, 86, 2874–2880.

Jonsen, I.D., Myers, R.A. & James, M.C. (2007). Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. *Mar. Ecol. Prog. Ser.*, 337, 255–264.

Kareiva, P.M. & Shigesada, N. (1983). Analyzing insect movement as a correlated random walk. *Oecol*, 56, 234–238.

Laidre, K.L., Heide-Jørgensen, M.P., Dietz, R., Hobbs, R.C. & Heide-Jørgensen, O.A. (2003). Deep-diving by narwhals, *Monodon monoceros*: differences in foraging behavior between wintering areas? *Mar. Ecol. Prog. Ser.*, 261, 269–281.

Laidre, K.L., Heide- Jørgensen, M.P., Logsdon, M.L., Hobbs, R.C., Dietz, R. & VanBlaricom, G.R. (2004). Fractal analysis of narwhal space use patterns. *Zoology*, 107, 3–11.

Marsh, L. & Jones, R. (1988). The form and consequence of random walk movement models. *J. Theor. Biol.*, 133, 113–131.

Matthiopoulos, J. (2003). The use of space by animals as a function of accessibility and preference. *Ecol. Modell.*, 159, 239–268.

Mauritzen, M., Derocher, A.E. & Wiig, Ø. (2001). Space-use strategies of female polar bears in a dynamic sea ice habitat. *Can. J. Zool.*, 79, 1704–1713.

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. (1999). Home range analysis using a mechanistic home range model. *Ecology*, 80, 1656–1665.

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.

Ovaskainen, O. (2004). Habitat-specific movement parameters estimated using mark-recapture data and a diffusion model. *Ecology*, 85, 242–257.

Ovaskainen, O., Luoto, M., Ikonen, I., Rekola, H., Meyke, E. & Kuussaari, M. (2008). An empirical test of a diffusion model: predicting clouded apollo movements in a novel environment. *Am. Nat.*, 171, 610–619.

Patterson, T.A., Thomas, L., Wilcox, C., Ovaskainen, O. & Matthiopoulos, J. (2008). State-space models of individual animal movement. *Trends Ecol. Evol.*, 23, 87–94.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.r-project.org.

Ropert-Coudert, Y. & Wilson, R.P. (2005). Trends and perspectives in animal-attached remote sensing. *Front. Ecol. Environ.*, 3, 437–444.

Royer, F., Fromentin, J.M. & Gaspar, P. (2005). A state-space model to derive bluefin tuna movement and habitat from archival tags. *Oikos*, 109, 473–484.

Schick, R.S., Loarie, S.R., Colchero, F., Best, B.D., Boustany, A., Conde, D.A. *et al.* (2008). Understanding movement data and movement processes: current and emerging directions. *Ecol. Lett.*, 11, 1338–1350.

Shigesada, N. (1980). Spatial distribution of dispersing animals. *J. Math. Biol.*, 9, 85–96.

Skellam, J.G. (1951). Random dispersal in theoretical populations. *Biometrika*, 38, 196–218.

Vincent, C., McConnell, B.J., Ridoux, V. & Fedak, M.A. (2002). Assessment of ARGOS location accuracy from satellite tags deployed on captive gray seals. *Mar. Mamm. Sci.*, 18, 156–166.

SUPPORTING INFORMATION

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

Appendix S1 R-Code for implementing behavioural change point analysis.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Editor, Jean Clobert
Manuscript received 11 December 2008
First decision made 14 January 2009
Manuscript accepted 30 January 2009