Stochastic differential equation based on a multimodal potential to model movement data in ecology

Pierre Gloaguen,

AgroParisTech, Paris, Institut Français de Recherche pour l'Exploitation de la Mer, Nantes, and Institut de Recherche en Informatique et Systèmes Aléatoires, Vannes, France

Marie-Pierre Etienne

AgroParisTech, Paris, France

and Sylvain Le Corff

Université Paris-Saclay, Orsay, France

Summary. The paper proposes a new model for individuals' movement in ecology. The movement process is defined as a solution to a stochastic differential equation whose drift is the gradient of a multimodal potential surface. This offers a new flexible approach among the popular potential-based movement models in ecology. To perform parameter inference, the widely used Euler method is compared with two other pseudolikelihood procedures and with a Monte Carlo expectation–maximization approach based on exact simulation of diffusions. Performances of all methods are assessed with simulated data and with a data set of fishing vessel trajectories. We show that the usual Euler method performs worse than the other procedures for all sampling schemes.

Keywords: Exact simulation; Global Positioning System data; Local linearization; Movement model; Pseudolikelihood methods; Stochastic differential equation

1. Introduction

Statistical inference of movement models provides many insights into the ecological features that explain population level dynamics. These analyses are crucial to wildlife managers to understand complex animal behaviours (Chavez, 2006). In fisheries science, understanding the underlying patterns that are responsible for spatial use of the ocean is a key aspect of sustainable management (Chang, 2011). Both fields promote now large programmes to deploy Global Positioning System (GPS) devices. For instance, we mention, among others, the 'Tagging of pelagic predators' programme (`www.gtopp.org`), the 'Tortues marine du Sud Ouest de l'Océan Indien' (`www.torsooi.com`) programme for marine animals, the 'Elephant without borders' (`www.elephantswithoutborders.org/tracking.php`) or Wolf GPS (`www.wolfgps.com`) programmes for terrestrial wildlife. In the European Union, since

Address for correspondence: Pierre Gloaguen, Institut de Recherche en Informatique et Systèmes Aléatoires, Campus de Tohannic, Rue André Lwoff, 56000 Vannes, France.
E-mail: pierre.gloaguen@irisa.fr

January 1st, 2012, fishing vessels above 12 m long are mandatorily equipped with a vessel monitoring system which has become a standard tool of fisheries monitoring world wide. As a result, such programmes produce huge data sets which have been largely used to understand, explain and predict animals' or vessels' movements. The GPS-type loggers can be set with several acquisition frequencies which should be adjusted to the experimental setting.

In the last few years, growing attention has been dedicated to continuous time and continuous space Markovian models as a realistic and flexible framework to model such data (Blackwell, 1997; Brillinger *et al.*, 2002; Preisler *et al.*, 2004). These papers focus on stochastic differential equations (SDEs) to describe and analyse animal trajectories. Since the model based on a pure diffusion process that was introduced in Skellam (1951) a large variety of SDEs has been proposed to model wildlife behaviour; see Preisler *et al.* (2004) for references about elephant seal or bird migrations. Brillinger *et al.* (2001a) introduced SDEs based on a potential surface to capture the directional bias of animal patterns. This potential function is assumed to reflect the attractiveness of the environment and regions where species are likely to travel to. In the framework that was proposed by Brillinger (2010), the drift of the SDE is given by the gradient of a potential map P_η which depends on an unknown parameter η whereas the diffusion coefficient is a smooth function. This flexible framework has been widely used in movement ecology for the last 20 years (Blackwell, 1997; Brillinger *et al.*, 2001a, 2002, 2011; Harris and Blackwell, 2013; Preisler *et al.*, 2004, 2013). Blackwell (1997), Blackwell *et al.* (2015) and Harris and Blackwell (2013) introduced a quadratic potential function which corresponds to a bivariate Ornstein–Uhlenbeck process to model the position of an animal. The Ornstein–Ulhenbeck process offers a convenient framework to represent attractiveness of a given area but it remains restrictive as animals are not prone to revert to a single attractive zone. This process owes its popularity to the simple form of the maximum likelihood estimator of η and of the diffusion coefficient as its transition density is Gaussian. Preisler *et al.* (2013) proposed different potential functions, with more complex features such as multiple attractive regions, as a baseline of a flexible framework to describe animal movement.

In this paper, we propose a new model where the potential function is a mixture of Gaussian-shaped functions. This model was suggested in Preisler *et al.* (2013), but, to the best of our knowledge, it has never been used in movement ecology or applied to practical cases. Each attractive region is modelled as a Gaussian-shaped function with its own mean location and information matrix characterizing its dispersion. The drift of the SDE ruling animal movement is a mixture of these functions, with unknown weights, centres and information matrices that should be inferred from movement data.

Designing efficient procedures to infer this model is a very challenging task as some elementary properties, such as the transition density of the solution to the SDE, are not available in closed form. Various statistical methods have been proposed to perform maximum likelihood estimation with unknown transition densities (see Iacus (2009) and Kessler *et al.* (2012)). The most convenient procedures to estimate discretely observed diffusion processes rely on discretization schemes to approximate the SDE between two observations and use a surrogate of the likelihood function to compute an approximate maximum likelihood estimator of η. The simplest likelihood approximation is the Euler–Maruyama method, which uses constant drift and diffusion coefficients on each interval. To the best of our knowledge, maximum likelihood inference procedures in movement ecology in the case of unknown transition densities are based only on the Euler–Maruyama method (see Brillinger (2010) and Preisler *et al.* (2013)). The maximum likelihood estimator that is computed with this method may be proved to be consistent under some assumptions on the total number of observations n and on the time step Δ_n between observations (which is assumed to vanish as $n \to \infty$). However, there is no guarantee that using

GPS tagging in ecology provides observations at a sufficiently high sampling rate and it might produce a highly biased estimator. Indeed, the chosen frequency is a crucial parameter to monitor animal behaviours over a given time horizon since GPS batteries depend on the number of recorded relocations. Therefore, depending on the needs of the given study, the recording frequency can be small. Moreover, real life experiments in complex environments (water for marine mammals for instance) might result in strongly irregular time step acquisitions. Therefore, alternatives to the Euler–Maruyama method should be considered to estimate parameters in potential-based movement models. For instance, the bias of the Euler–Maruyama method may be reduced by using higher order discretization schemes.

A particularly interesting higher order scheme is given by the Ozaki discretization which proposes a linear approximation of the drift coefficient combined with a constant approximation of the diffusion term (Ozaki, 1992; Shoji and Ozaki, 1998a). More recently, Kessler (1997), Kessler *et al.* (2012) and Uchida and Yoshida (2012) introduced another Gaussian-based approximation of the transition density between two consecutive observations by using higher order expansions of the conditional mean and variance of an observation given the observation at the previous time step. Another approach to approximate the likelihood of the observations based on Hermite polynomials expansion was introduced by Ait-Sahalia (1999, 2002, 2008) and extended in several directions recently; see Li (2013) and all the references therein. However, these methods also induce, theoretically, a bias due to the likelihood approximation.

To avoid this systematic bias associated with every discretization procedure, Beskos and Roberts (2005) and Beskos, Papaspiliopoulos and Roberts (2006) proposed an exact algorithm to draw samples from finite dimensional distributions of potential-based diffusion processes. This sampling method is based on a rejection sampling step and may be used to sample from any finite dimensional distribution of the target SDE. This algorithm is a pivotal tool to obtain an unbiased estimate of the intermediate quantity of the expectation–maximization (EM) algorithm (Dempster *et al.*, 1977). As a by-product, this also enables sampling trajectories exactly distributed according to the estimated model which is not possible with the three other methods. However, the computational complexity of the acceptance–rejection step of these exact algorithms, which requires sampling skeletons of trajectories between each pair of observations, grows with the time step between observations.

There are no theoretical nor practical studies to choose the method which achieves the best trade-off between fast convergence and accuracy in movement ecology. We investigate the performance of four different estimation methods for the new model proposed in this paper:

(a) the Euler–Maruyama method;
(b) the Ozaki local linearization method, proposed in Ozaki (1992) and Shoji and Ozaki (1998a);
(c) the Kessler high order Gaussian approximation method, proposed in Kessler (1997) and refined using an adaptive procedure in Uchida and Yoshida (2012);
(d) the exact algorithm Monte Carlo EM (EAMCEM) method, proposed in Beskos, Papaspiliopoulos, Roberts and Fearnhead (2006).

The paper is organized as follows. The potential-based model is introduced in Section 2 and the inference methods that are used in the paper are presented in Section 3. A simulation study to evaluate the robustness of the statistical methods with the specific potential function that is proposed in the paper is given in Section 4.1. To evaluate the effect of the sampling rate, three sampling scenarios are considered: high frequency, intermediate frequency and low frequency. Finally, an application of our new model to real data, together with an evaluation of

the performance of each inference method, is done in Section 4.2 for the estimation of attractive zones of two French fishing vessels.

The data that are analysed in the paper and the programs that were used to analyse them can be obtained from

```
http://wileyonlinelibrary.com/journal/rss-datasets
```

2. Potential-based movement model

In this section, we propose a new model to describe the position process of an individual. The latter is modelled as a stochastic process $(\mathbf{X}_t)_{0 \leqslant t \leqslant T}$ observed between times 0 and T. $(\mathbf{X}_t)_{0 \leqslant t \leqslant T}$ is assumed to be the solution to the following time homogeneous SDE:

$$\mathbf{X}_0 = \mathbf{x}_0,$$
$$\mathrm{d}\mathbf{X}_t = b_\eta(\mathbf{X}_t)\mathrm{d}t + \gamma\,\mathrm{d}\mathbf{W}_t, \tag{1}$$

where $(\mathbf{W}_t)_{0 \leqslant t \leqslant T}$ is a standard Brownian motion on $\mathbb{R}^2$, and $\gamma \in \mathbb{R}_+^\star$ is an unknown diffusion parameter. Moreover, we assume that the drift follows the gradient of mixtures function, i.e.

$$b_\eta(\mathbf{x}) := \nabla P_\eta(\mathbf{x}), \tag{2}$$

$$P_\eta(\mathbf{x}) = \sum_{i=1}^{K} \pi_k \varphi_k^\eta(\mathbf{x}) \qquad \varphi_k^\eta(\mathbf{x}) := \exp\{-\tfrac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_k)^\mathrm{T} C_k(\mathbf{x} - \boldsymbol{\mu}_k)\}, \tag{3}$$

where K is the number of components of the mixture, $\pi_k \in \mathbb{R}^+$ is the relative weight of the kth component, with $\Sigma_{k=1}^{K} \pi_k = 1$, $\boldsymbol{\mu}_k \in \mathbb{R}^2$ is the centre of the kth component and $C_k \in \mathcal{S}_2^+$ is the information matrix of the kth component, where $\mathcal{S}_2^+$ is the set of 2×2 symmetric positive definite matrices.

Equations (1) and (2) define a potential-based continuous time movement model, and follow a popular idea in movement ecology (Blackwell, 1997; Brillinger *et al.*, 2001a,b, 2002; Preisler *et al.*, 2004, 2013; Brillinger, 2010; Harris and Blackwell, 2013). In this framework, the movement is supposed to reflect the attractiveness of the environment, which is modelled by using a real-valued potential function defined on $\mathbb{R}^2$. Equation (3) describes the new potential function that is proposed in this paper. This function is a smooth multimodal surface, having K different components. In this paper, K is supposed to be known and we need to infer η from movement data where

$$\eta = \{(\pi_k)_{k=1,\dots,K-1}, (\boldsymbol{\mu}_k)_{k=1,\dots,K}, (C_k)_{k=1,\dots,K}\} \in \mathbb{R}^{6K-1}.$$

Potential-based continuous time movement models have been widely studied in ecology for the last 20 years. The two most popular SDE-based models in movement ecology are

(a) Brownian motion, with or without drift, corresponding to a constant b_η (corresponding to a linear potential function), as in Skellam (1951), and
(b) the Ornstein–Uhlenbeck process, corresponding to a linear drift function b_η (corresponding to a linear potential function), as in Blackwell (1997).

These two models are appealing as their transition density function is available explicitly, and the inference in these cases is straightforward. However, it contains only a very restricted class of potential functions P_η (linear, or quadratic) that might be unrealistic in ecological applications (unimodal attractiveness surfaces, for instance). The model that is introduced in this paper is an extension of one of those suggested in Preisler *et al.* (2013); it has multiple modes and is

sufficiently flexible to accommodate a wide range of situations. However, the process solution to the SDE (1) with the potential function (3) has no explicit transition density so that approximate maximum likelihood procedures such as the methods that are presented in Section 3 may be considered.

3. Maximum likelihood estimation procedures

The aim of this section is to present four maximum likelihood inference procedures to estimate the parameter $\theta = (\eta, \gamma)$ by using a set of discrete time observations. Among them, the Euler–Maruyama method is the only one that has been used so far in movement ecology, regardless of its potential weaknesses. The performances of these methods are evaluated in Section 4.

The observations are given by G independent realizations of the continuous process given by equations (1)–(3). The set of observations is then the collection of trajectories $(x^g)_{g=1,\ldots,G}$ where each trajectory x^g is made of $n_g + 1$ exact observations at times $t_0^g = 0 < t_1^g < \ldots < t_{n^g}^g = T^g$ and starting at $\mathbf{x}_0^g$. The probability density of the distribution of $\mathbf{X}_\Delta$ given $\mathbf{X}_0$ (i.e. the transition density) is denoted by $q_{\boldsymbol{\theta}}^\Delta$ when the model is parameterized by $\boldsymbol{\theta}$: for all bounded measurable functions h on $\mathbb{R}^2$,

$$\mathbb{E}_{\boldsymbol{\theta}}[h(\mathbf{X}_\Delta)|\mathbf{X}_0] = \int h(\mathbf{y}) q_{\boldsymbol{\theta}}^\Delta(\mathbf{X}_0, \mathbf{y}) \, \mathrm{d}\mathbf{y},$$

where $\mathbb{E}_{\boldsymbol{\theta}}$ is the expectation when the model is driven by $\boldsymbol{\theta}$.

As the G trajectories $(x^g)_{g=1,\ldots,G}$ are independent, and by the Markov property of the solution to equation (1), the likelihood function may be written as

$$L(\boldsymbol{\theta}; x) = \prod_{g=1}^{G} \prod_{i=0}^{n^g - 1} q_{\boldsymbol{\theta}}^{\Delta_i^g}(\mathbf{x}_i^g, \mathbf{x}_{i+1}^g), \tag{4}$$

where $\Delta_i^g := t_{i+1}^g - t_i^g$. The maximum likelihood estimator of $\boldsymbol{\theta}$ is defined as

$$\hat{\boldsymbol{\theta}} := \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}; x).$$

In the cases of the Brownian motion and Ornstein–Uhlenbeck processes, the function $q_{\boldsymbol{\theta}}^\Delta$ is the probability density function of a Gaussian random variable with known mean and covariance so the direct computation of $\hat{\boldsymbol{\theta}}$ is an easy task. However, in more general settings such as in the model that was considered in Section 2, $q_{\boldsymbol{\theta}}^\Delta$ is unknown. Whereas many statistical methods have been proposed to compute approximations of $\hat{\boldsymbol{\theta}}$, the only statistical method that is used in movement ecology is the Euler–Maruyama discretization procedure. It is known that the quality of this approximation depends highly on the sampling rate, which means that this method might therefore not be well suited to real life animal GPS tags. Therefore, several alternatives are investigated in the remainder of this paper to compute approximations of $\hat{\boldsymbol{\theta}}$. Since all g trajectories are exchangeable, for clarity, the aforementioned approximation methods are detailed for a given trajectory and references to g are omitted.

3.1. Euler–Maruyama method

The Euler–Maruyama discretization is an order 0 Taylor series expansion of the drift function (as the diffusion coefficient is constant no expansion is required). The drift is therefore assumed to be constant between two observations. For any $0 \leqslant i \leqslant n - 1$ the target process is approximated by the process $(\mathbf{X}_t^E)_{t_i \leqslant t < t_{i+1}}$, the solution to the SDE $\mathbf{X}_{t_i}^E = \mathbf{x}_i$ and, for $t_i \leqslant t \leqslant t_{i+1}$,

$$\mathrm{d}\mathbf{X}_t^E = b_\eta(\mathbf{x}_i) \, \mathrm{d}t + \gamma \, \mathrm{d}\mathbf{W}_t.$$

For all $0 \leqslant i \leqslant n-1$, the transition density $q_{\boldsymbol{\theta}}^{\Delta_i}(\mathbf{x}_i, \mathbf{x}_{i+1})$ is approximated by the transition density of the process $(\mathbf{X}_t^{\mathrm{E}})_{t_i \leqslant t \leqslant t_{i+1}}$, denoted by $q_{\boldsymbol{\theta}}^{\Delta_i, \mathrm{E}}$, evaluated for the observations $(\mathbf{x}_i, \mathbf{x}_{i+1})$. This is the probability density function of a Gaussian random variable with mean $\boldsymbol{\mu}_i^{\mathrm{E}}$ and variance Σ_i^{E}:

$$\boldsymbol{\mu}_i^{\mathrm{E}} = \mathbf{x}_i + \Delta_i b_\eta(\mathbf{x}_i),$$
$$\Sigma_i^{\mathrm{E}} = \gamma^2 \Delta_i,$$

where I_k denotes the $k \times k$ identity matrix. Hence, the Euler–Maruyama estimate is given by maximizing equation (4) when $q_{\boldsymbol{\theta}}^{\Delta_i^g}$ is replaced by $q_{\boldsymbol{\theta}}^{\Delta_i^g, \mathrm{E}}$.

3.2.   Ozaki method

The Ozaki method, proposed in Ozaki (1992) and Shoji and Ozaki (1998a), provides a local linearization of the drift term to improve the Euler scheme. For any $0 \leqslant i \leqslant n-1$ the target process is approximated by the process $(\mathbf{X}_t^{\mathrm{O}})_{t_i \leqslant t < t_{i+1}}$, the solution to the SDE $\mathbf{X}_{t_i}^{\mathrm{O}} = \mathbf{x}_i$ and, for $t_i \leqslant t \leqslant t_{i+1}$,

$$\mathrm{d}\mathbf{X}_t^{\mathrm{O}} = \{J_{i,\eta}(\mathbf{X}_t^{\mathrm{O}} - \mathbf{x}_i) + b_\eta(\mathbf{x}_i)\}\mathrm{d}t + \gamma \, \mathrm{d}\mathbf{W}_t,$$

where $J_{i,\eta}$ is the 2×2 Jacobian matrix of the drift function b_η evaluated at $\mathbf{x}_i$. Therefore, the target process is now approximated between each pair of observations by a two-dimensional Ornstein–Uhlenbeck process. For all $0 \leqslant i \leqslant n-1$, the transition density $q_{\boldsymbol{\theta}}^{\Delta_i}(\mathbf{x}_i, \mathbf{x}_{i+1})$ is approximated by the transition density of the process $(\mathbf{X}_t^{\mathrm{O}})_{t_i \leqslant t \leqslant t_{i+1}}$, denoted by $q_{\boldsymbol{\theta}}^{\Delta_i, \mathrm{O}}$, evaluated for the observations $(\mathbf{x}_i, \mathbf{x}_{i+1})$. If the potential function is such that, for all $0 \leqslant i \leqslant n-1$ and all $\boldsymbol{\theta}$, $J_{i,\boldsymbol{\theta}}$ is non-singular and symmetric (as is the case of the potential that was introduced in Section 2); following Shoji and Ozaki (1998b), this transition density is the probability density function of a Gaussian random variable with mean $\boldsymbol{\mu}_i^{\mathrm{O}}$ and variance Σ_i^{O}:

$$\boldsymbol{\mu}_i^{\mathrm{O}} = \mathbf{x}_i + \{\exp(J_{i,\eta}\Delta_i) - I_2\}(J_{i,\eta})^{-1} b_\eta(\mathbf{x}_i),$$
$$\mathrm{vec}(\Sigma_i^{\mathrm{O}}) = (J_{i,\eta} \oplus J_{i,\eta})^{-1} \exp[\{(J_{i,\eta} \oplus J_{i,\eta})\Delta_i\} - I_4] \, \mathrm{vec}(\gamma^2 I_2),$$

where 'vec' is the stack operator and '$\oplus$' is the Kronecker sum. The local linearization estimate is then given by maximizing equation (4) when $q_{\boldsymbol{\theta}}^{\Delta_i^g}$ is replaced by $q_{\boldsymbol{\theta}}^{\Delta_i^g, \mathrm{O}}$.

3.3.   Kessler method

The procedure that was proposed in Kessler (1997) generalizes the Euler–Maruyama method by using a higher order approximation of the transition density. Following the same steps as Euler or Ozaki discretizations, Kessler (1997) proposed considering a Gaussian approximation of the transition density between consecutive observations. The procedure aims at replacing $q_{\boldsymbol{\theta}}^{\Delta_i}(x_i, x_{i+1})$ by a Gaussian probability density function with mean and variance given by the exact mean $\boldsymbol{\mu}_i^{\mathrm{target}}$ and variance $\Sigma_i^{\mathrm{target}}$ of the target process at time t_{i+1} conditionally on the value of the target process at time t_i. This leads to an approximation with the same first two moments as the target transition density. This approximation can be written similarly to the previous discretization schemes by defining, for any $0 \leqslant i \leqslant n-1$, the process $(\mathbf{X}_t^{\mathrm{target}})_{t_i \leqslant t < t_{i+1}}$ as the solution to the SDE $\mathbf{X}_{t_i}^{\mathrm{target}} = \mathbf{x}_i$ and, for $t_i \leqslant t \leqslant t_{i+1}$,

$$\mathrm{d}\mathbf{X}_t^{\mathrm{target}} = \frac{\boldsymbol{\mu}_i^{\mathrm{target}} - \mathbf{x}_i}{\Delta_i} \mathrm{d}t + \left(\frac{\Sigma_i^{\mathrm{target}}}{\Delta_i}\right)^{1/2} \mathrm{d}\mathbf{W}_t.$$

However, as the transition density of the target process is unknown, $\boldsymbol{\mu}_i^{\mathrm{target}}$ and $\Sigma_i^{\mathrm{target}}$ are not available in closed form. Nonetheless, Taylor series expansions of these two conditional

moments are available (see Florens-zmirou (1989), or Uchida and Yoshida (2012) for the multi-dimensional case). Kessler (1997) proposed replacing μ_i^{target} and Σ_i^{target} by these approximations. For all $0 \leqslant i \leqslant n-1$, the transition density $q_{\boldsymbol{\theta}}^{\Delta_i}(\mathbf{x}_i, \mathbf{x}_{i+1})$ is approximated by the transition density of the process $(\mathbf{X}_t^K)_{t_i \leqslant t \leqslant t_{i+1}}$ (approximated with these expansions). These expansions can be computed directly from the drift and diffusion functions and their partial derivatives. The order of the expansion of the true conditional moments is left to the user and the performance of the estimator highly depends on this parameter, order 1 being the Euler–Maruyama method. In this paper, it was performed up to the second order. In this case, and for the process defined by equations (1) and (2), the function $q_{\boldsymbol{\theta}}^{\Delta_i, K}(\mathbf{x}_i, \cdot)$ is the probability density function of a Gaussian random variable with mean $\boldsymbol{\mu}_i^K$ and variance Σ_i^K:

$$
\begin{aligned}
\boldsymbol{\mu}_i^K &= \mathbf{x}_i + \Delta_i b_\eta(\mathbf{x}_i), \\
\Sigma_i^K &= \gamma^2 \Delta_i (I_2 + \Delta_i J_{i,\eta}).
\end{aligned}
\tag{5}
$$

Note that $\boldsymbol{\mu}_i^K = \boldsymbol{\mu}_i^E$. For numerical stability, $(\Sigma_i^K)^{-1}$ and its determinant can be replaced by a Taylor series expansion; the associated contrast function to optimize in multi-dimensional cases is given in Uchida and Yoshida (2012). A drawback of this method is the fact that equation (5) does not necessarily define a positive semidefinite matrix. For instance, if $\text{tr}(J_{i,\eta}) < -2/\Delta_i$, then $\text{tr}(\Sigma_i^K) < 0$, which is likely to occur when Δ_i is large. In the following applications, whenever Σ_i^K is not positive definite, the associated observation is thrown out from the computation of the likelihood, as proposed in Iacus (2009). The Kessler estimate is then given by maximizing equation (4) when $q_{\boldsymbol{\theta}}^{\Delta_i^g}$ is replaced by $q_{\boldsymbol{\theta}}^{\Delta_i^g, K}$.

3.4. Exact-algorithm-based Monte Carlo expectation–maximization method

The EAMCEM method proposed by Beskos, Papaspiliopoulos, Roberts and Fearnhead (2006) does not use Gaussian approximations of the transition densities to compute the maximum likelihood estimator. Details of the method are given in Beskos, Papaspiliopoulos, Roberts and Fearnhead (2006) and the important results for the experiments are given on line as supplementary material.

Applying the exact algorithm to estimate $\boldsymbol{\theta}$ requires the target SDE to be reducible to a unit diffusion by using the Lamperti transform which is obtained in our case by setting $(\mathbf{Y}_t := \gamma^{-1}\mathbf{X}_t)_{0 \leqslant t \leqslant T}$. Then,

$$
d\mathbf{Y}_t = \alpha_{\boldsymbol{\theta}}(\mathbf{Y}_t)dt + d\mathbf{W}_t \qquad \alpha_{\boldsymbol{\theta}}(\cdot) := \gamma^{-1} b_\eta(\gamma \cdot) = \gamma^{-1} \nabla P_\eta(\gamma \cdot).
\tag{6}
$$

The EAMCEM approach that is used in this work relies on the following assumptions.

(a) *Conservative assumption*: for all $\boldsymbol{\theta} = (\eta, \gamma) \in \mathbb{R}^{d+1}$, there exists $H_{\boldsymbol{\theta}} : \mathbb{R}^2 \mapsto \mathbb{R}$ such that, for all $\mathbf{x} \in \mathbb{R}^2$,

$$
\alpha_{\boldsymbol{\theta}}(\mathbf{x}) = \nabla H_{\boldsymbol{\theta}}(\mathbf{x}).
\tag{7}
$$

(b) *Boundedness condition*: for all $\boldsymbol{\theta} \in \mathbb{R}^{d+1}$, there exist $m_{\boldsymbol{\theta}}$ and $M_{\boldsymbol{\theta}}$, such that, for all $\mathbf{x} \in \mathbb{R}^2$,

$$
m_{\boldsymbol{\theta}} \leqslant \|\alpha_{\boldsymbol{\theta}}(\mathbf{x})\|^2 + \Delta H_{\boldsymbol{\theta}}(\mathbf{x}) \leqslant M_{\boldsymbol{\theta}},
\tag{8}
$$

where Δ is the Laplace operator,

$$
\Delta H_{\boldsymbol{\theta}} : \mathbf{x} \mapsto \frac{\partial \alpha_{\boldsymbol{\theta},1}}{\partial x_1}(\mathbf{x}) + \frac{\partial \alpha_{\boldsymbol{\theta},2}}{\partial x_2}(\mathbf{x}).
$$

Both conditions are somewhat restrictive in general; however, the first turns out to be satisfied
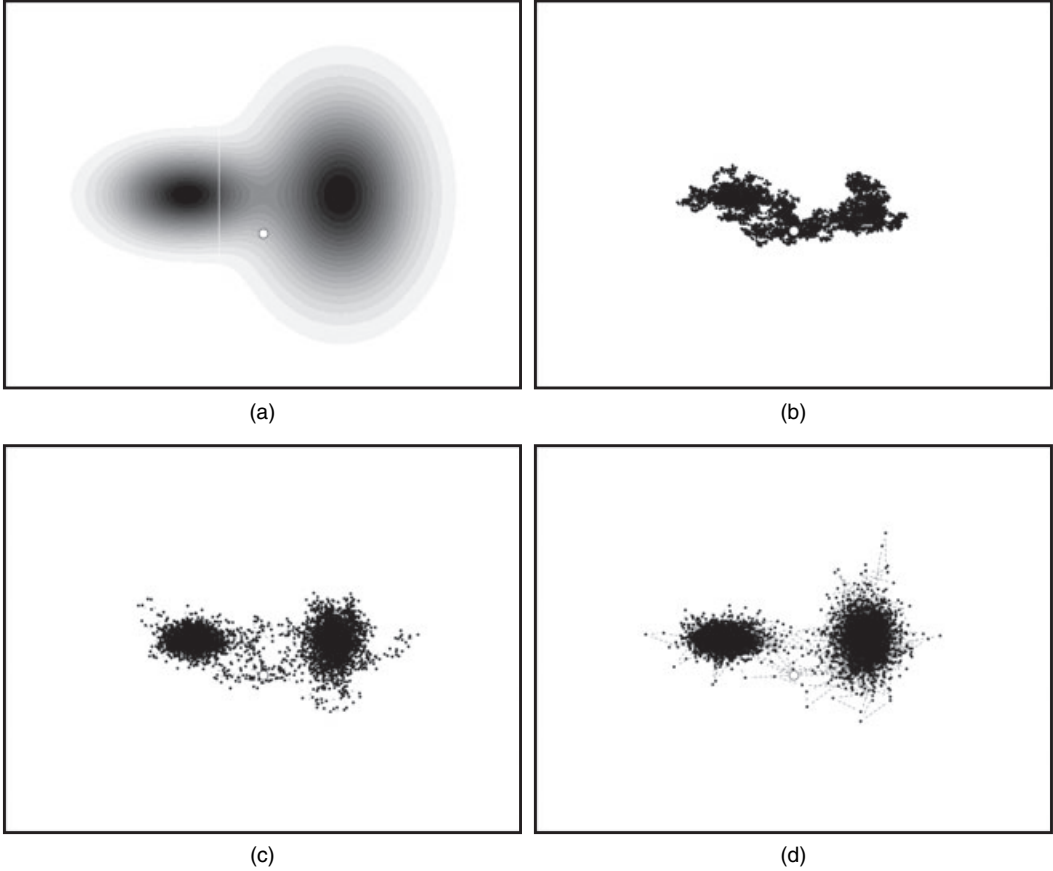
Fig. 1. Simulation of the process solution to equation (1) with a potential as in equation (3) (in (b)–(d) 10 simulated trajectories with three different sampling schemes are given; $\bigcirc$, starting point x_0 of each trajectory): (a) hidden potential map driving the movement (dark zones present high potential whereas white zones have low potential); (b) $\Delta = 0.1$; (c) $\Delta = 1$; (d) $\Delta = 10$

automatically for potential-based models that are studied in ecology, and described in Section 2. Given the model that is described by equations (1)–(3), H_θ is given as

$$H_\theta : \mathbf{x} \mapsto \sum_{k=1}^{K} \pi_k \varphi_k^\eta (\gamma \mathbf{x}) / \gamma,$$

with $\theta = (\eta, \gamma)$. The second condition can be relaxed at the cost of additional computations (Beskos, Papaspiliopoulos, Roberts and Fearnhead, 2006) but is satisfied by the potential function of equation (3).

This method offers the advantage of avoiding approximation errors due to time discretization of the SDE. The only error comes from the Monte Carlo simulations that are used to approximate the expectation in the E-step (see the details in the on-line supplementary material). In theory, if conditions (7) and (8) are satisfied, it is always possible to perform these simulations. However, the simulation algorithm is based on rejection sampling and the acceptance probability is equal to

$$\exp\left\{ -\frac{1}{2}(M_\theta - m_\theta) \int_{t_i}^{t_{i+1}} \phi_\theta(\mathbf{Y}_s) \mathrm{d}s \right\}, \tag{9}$$
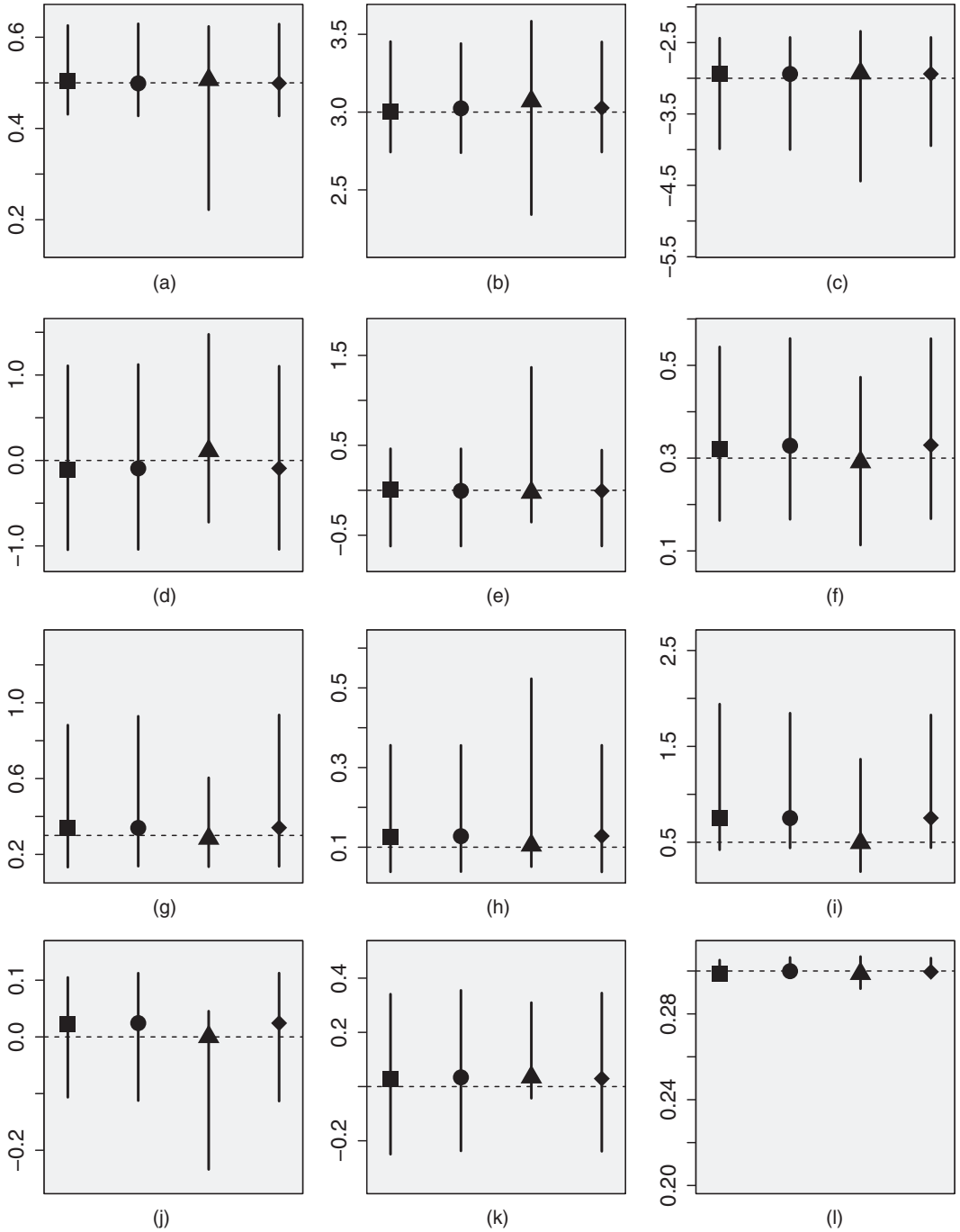
Fig. 2. Estimation of the parameters in the case where $\Delta = 0.1$ (■ (Euler), ● (Ozaki), ▲ (Kessler), ♦ (EAMCEM), medians; |, 95% range of estimations; ------, true value): (a) π_1; (b) $\mu_1^{(1)}$; (c) $\mu_1^{(2)}$; (d) $\mu_2^{(1)}$; (e) $\mu_2^{(2)}$; (f) $C_{11}^{(1)}$; (g) $C_{11}^{(2)}$; (h) $C_{22}^{(1)}$; (i) $C_{22}^{(2)}$; (j) $C_{12}^{(1)}$; (k) $C_{12}^{(2)}$; (l) γ
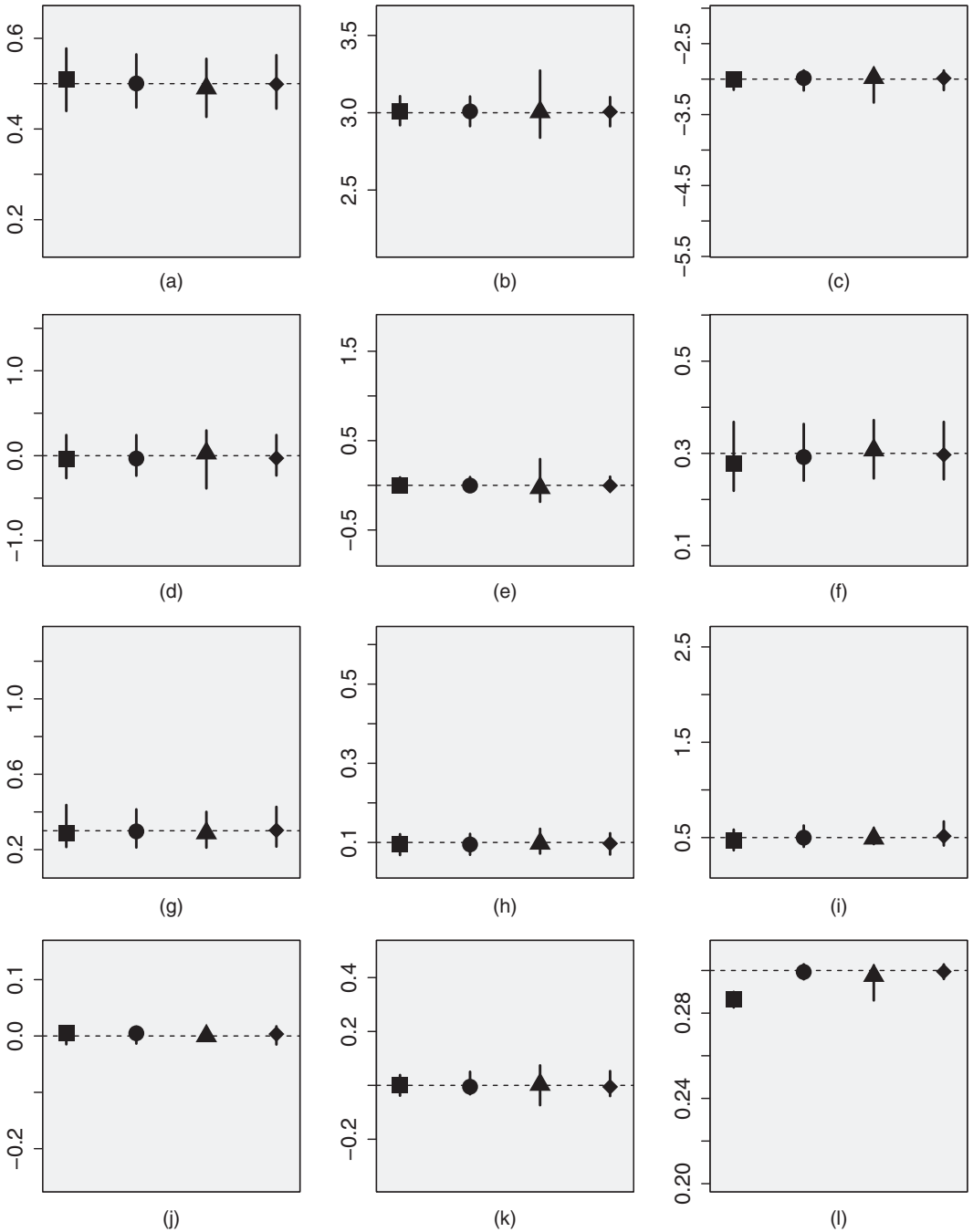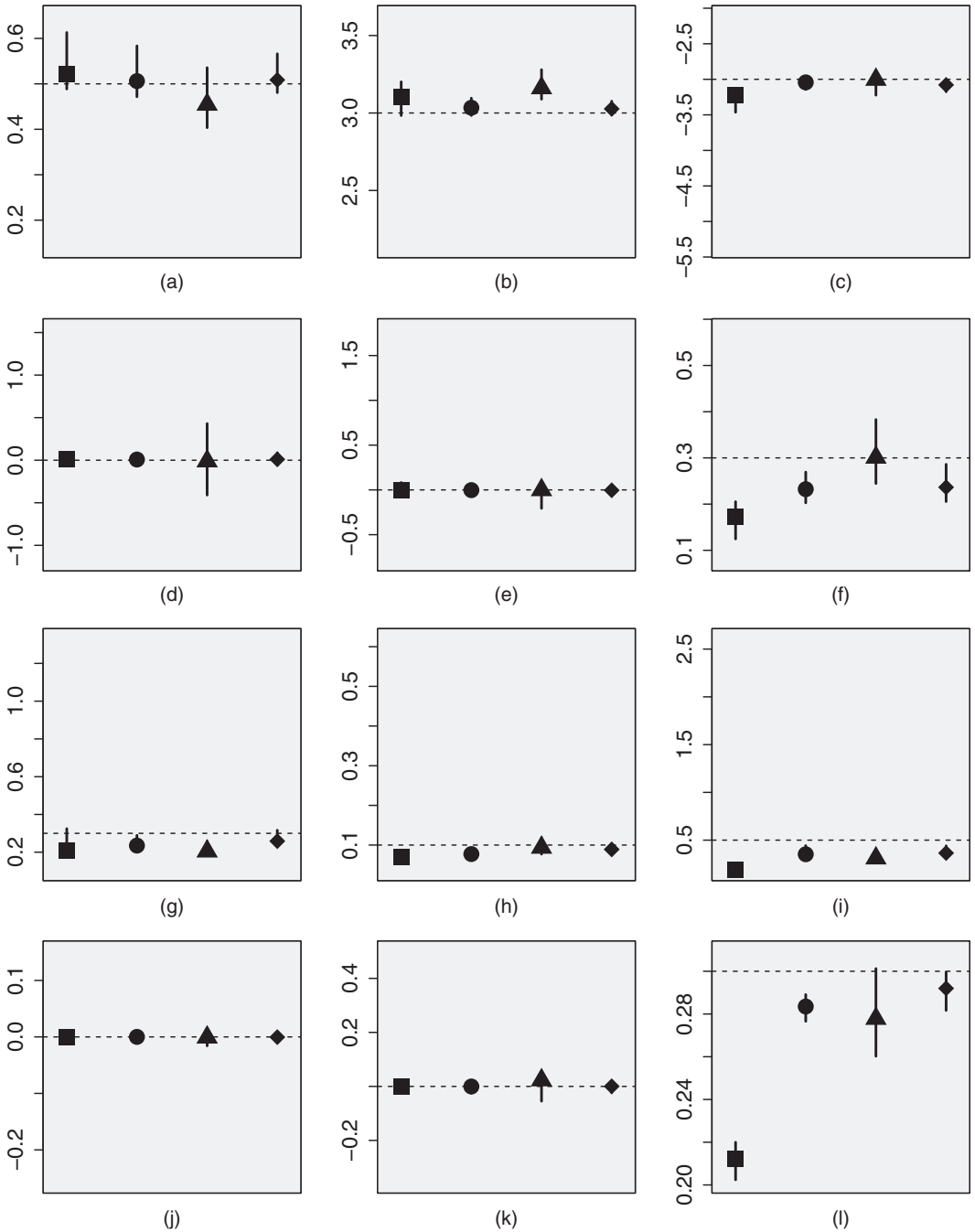
Fig. 3. Estimation of the parameters in the case where $\Delta = 1$ ($\blacksquare$ (Euler), $\bullet$ (Ozaki), $\blacktriangle$ (Kessler), $\blacklozenge$ (EA-MCEM), medians; $|$, 95% range of estimations; ------, true value): (a) π_1; (b) $\mu_1^{(1)}$; (c) $\mu_1^{(2)}$; (d) $\mu_2^{(1)}$; (e) $\mu_2^{(2)}$; (f) $C_{11}^{(1)}$; (g) $C_{11}^{(2)}$; (h) $C_{22}^{(1)}$; (i) $C_{22}^{(2)}$; (j) $C_{12}^{(1)}$; (k) $C_{12}^{(2)}$; (l) γ

Fig. 4. Estimation of the parameters in the case where $\Delta = 10$ (■ (Euler), ● (Ozaki), ▲ (Kessler), ♦ (EAMCEM), medians; |, 95% range of estimations; ------, true value): (a) π_1; (b) $\mu_1^{(1)}$; (c) $\mu_1^{(2)}$; (d) $\mu_2^{(1)}$; (e) $\mu_2^{(2)}$; (f) $C_{11}^{(1)}$; (g) $C_{11}^{(2)}$; (h) $C_{22}^{(1)}$; (i) $C_{22}^{(2)}$; (j) $C_{12}^{(1)}$; (k) $C_{12}^{(2)}$; (l) γ

where $m_{\boldsymbol{\theta}}$ and $M_{\boldsymbol{\theta}}$ are defined as in condition (8), and $\phi_{\boldsymbol{\theta}}(\mathbf{x}) := \{\| \alpha_{\boldsymbol{\theta}}(\mathbf{x}) \|^2 + \Delta H_{\boldsymbol{\theta}}(\mathbf{x}) - m_{\boldsymbol{\theta}}\}/2 \geqslant 0$. Then, the integral (9) increases with the time step, and the acceptance probability decreases to 0. Therefore, in practice, this method also depends on the sampling step to be used in a reasonable time.

To compare the performance of the three methods, an approximated likelihood criterion is used. The log-likelihood of two given estimates $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ cannot be computed exactly. However, in the case where conditions (7) and (8) hold, and following Beskos *et al.* (2009), an unbiased estimator l_{EA} of the log-likelihood based on Monte Carlo simulations may be computed. This estimator has no intrinsic bias as the error comes only from the Monte Carlo procedure. In what follows, this criterion allowed us to compare the quality of the various estimations resulting from the four methods.

4. Application

4.1. Simulation study

The performance of the four estimation methods that are presented in this paper is evaluated by using several sampling schemes. The hidden potential map has $K = 2$ components and is represented in Fig. 1(a). All simulated trajectories start from the same $\mathbf{x}_0$ and are sampled from the true distribution of the process $(X_t)_{t \geqslant 0}$ with parameters η^* and γ^* by using the acceptance–rejection procedure that was proposed in Beskos and Roberts (2005). $G = 10$ independent trajectories are simulated with $n_g = 500$ recorded points. Three different sampling time steps are considered, with high frequency ($\Delta = 0.1$), intermediate frequency ($\Delta = 1$) and low frequency ($\Delta = 10$). In an actual sampling design, this would correspond to three different GPS tags settings with a limited number of emissions. For each method and in each scenario, this procedure is repeated independently to produce 30 different data sets.

The optimization problem that is associated with the M-step of the EM algorithm is performed by using the covariance matrix adaptation evolution strategy algorithm that was proposed by Hansen and Ostermeier (2001). For each estimation method, 30 initial parameters $\boldsymbol{\theta}_0$ are used and the parameter estimate with the maximum objective function is considered as the estimator $\hat{\boldsymbol{\theta}}$. Figs 2, 3 and 4 display the parameter estimate for the three sampling cases. The best estimate for all parameters and techniques is obtained with the intermediate frequency sampling rate. As illustrated by Fig. 1, this frequency enables a good exploration of the map by the process with 500 observations and all methods provide reliable estimates. When $\Delta = 0.1$ with only 500 observations per trajectory, the time horizon 50 might be insufficient for good parameter estimation using this model. In contrast, when $\Delta = 10$, there is a strong bias in the estimation of the shape parameters $C^k, k = 1, 2$, and the diffusion coefficient for all methods.

The Euler method performs much worse than the other methods when $\Delta = 1$ and $\Delta = 10$ whereas the three other methods are more robust to the sampling frequency. The Ozaki method provides similar results to the EAMCEM method. The Kessler estimation method provides results which may differ for some parameters (for instance $\boldsymbol{\mu}_1^{(1)}$ or $C_{11}^{(1)}$ in Fig. 4) as about 50% of the observations lead to a non-positive semidefinite matrix in equation (5) and have therefore been dropped. This is a major drawback of this method as it induces a supplementary bias. Finally, note that the EAMCEM method has a much larger computation cost than all other procedures but it is less sensitive to the starting point of the algorithm. The mean acceptance rate decreases from 99.6% ($\Delta = 0.1$) to 97.6 ($\Delta = 1$) and 67.5% ($\Delta = 10$).

Fig. 5 displays the absolute estimation errors $|P_{\hat{\eta}}(\mathbf{x}) - P_{\eta^*}(\mathbf{x})|$ for each method. The map is
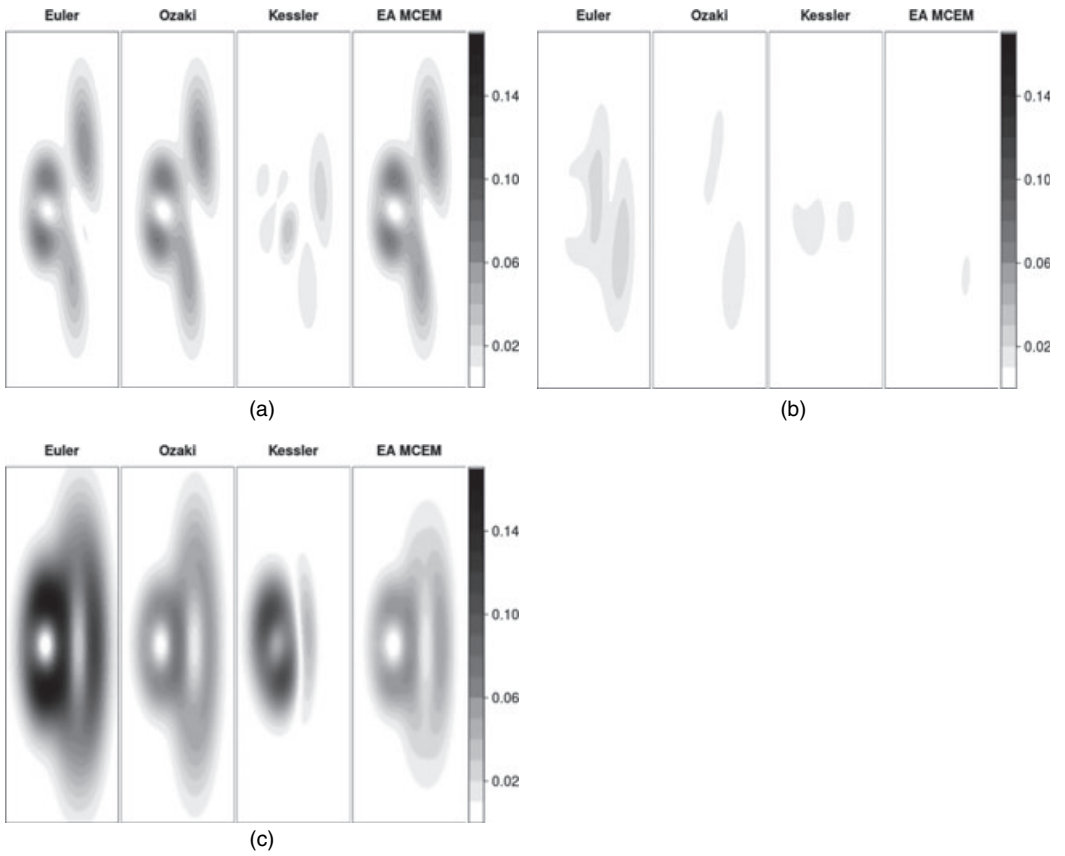
(a)

(b)

(c)

Fig. 5. Error between the median map estimate and the true map: results are presented for each method with (a) $\Delta = 0.1$, (b) $\Delta = 1$ and (c) $\Delta = 10$

produced by using the median of all estimated values as a set of parameters. The Euler method performs much worse than the other methods when $\Delta = 1$ and $\Delta = 10$.

Fig. 6 shows the distribution of the relative integrated error

$$\frac{\int |P_{\hat{\eta}}(\mathbf{x}) - P_{\eta*}(\mathbf{x})| d\mathbf{x}}{\int |P_{\eta*}(\mathbf{x})| d\mathbf{x}}$$

for each method. According to this simulation study, it seems preferable not to use the Euler method in comparison with any of the three other procedures.

4.2.   *Inferring fishing zones*

Delimiting fishing zones with high potential represents a key step in fisheries management, for instance to set up some marine protected areas. To illustrate the performance of each method to define such zones by using actual GPS data, the model that was presented in Section 2 is estimated by using two sets of trajectories of two French fishing vessels named V1 and V2. The two sets of trajectories are represented in Fig. 7.

(a)   (b)

(c)
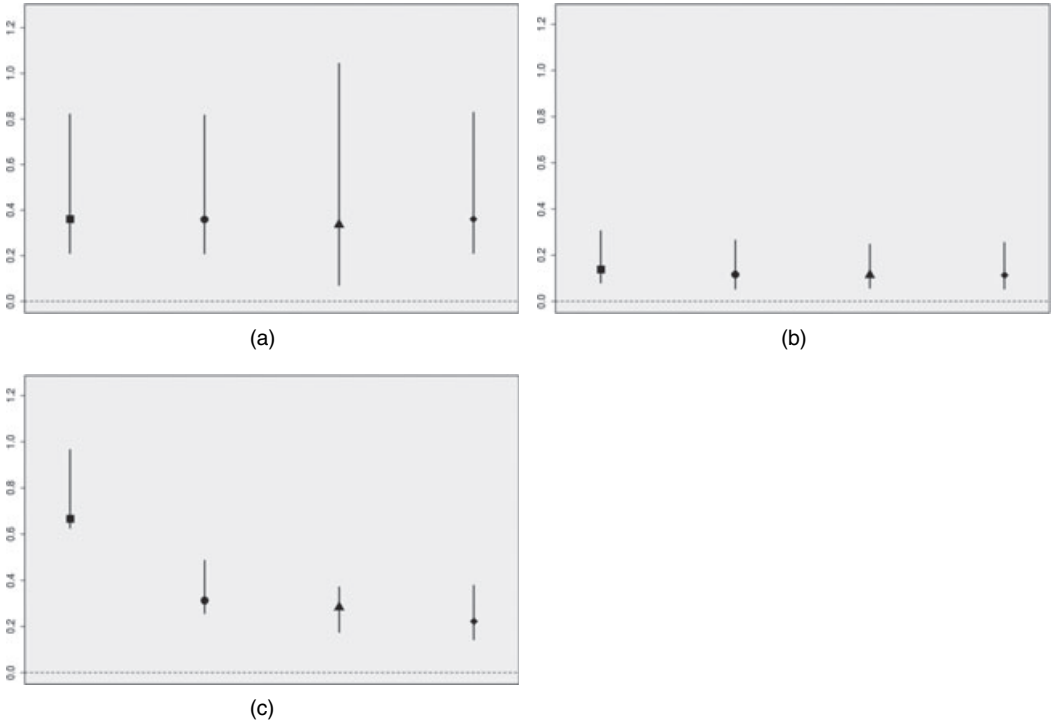
Fig. 6. Distribution of the relative integrated error between the estimated map and the true map (■, Euler; ●, Ozaki; ▲, Kessler; ♦, EAMCEM): results are presented for each method with (a) $\Delta = 0.1$, (b) $\Delta = 1$ and (c) $\Delta = 10$

Table 1. Value of the contrast function to maximize for each estimated parameter, for each data set†

Estimate	Results for the following contrasts and vessel V1:				Results for the following contrasts and vessel V2:			
	l_E	l_O	l_K	l_{EA}	l_E	l_O	l_K	l_{EA}
θ_E	-1.32	-1.29	-0.37	-1.29	-2.17	-2.16	-1.07	-2.16
θ_O	-1.50	-0.91	0.39	-0.93	-2.21	-2.09	-0.90	-2.09
θ_K	-1.50	-0.96	0.70	-0.99	-2.26	-2.18	-0.67	-2.10
θ_{EA}	-1.49	-0.92	0.06	-0.92	-2.21	-2.11	-0.98	-2.07

†Each entry is the value of the contrast function (given in the columns) evaluated for an estimation of θ (given in the rows). The highest value for each column is on the diagonal. Values are normalized contrasts (divided by the number of segments of trajectories).

(a) For vessel V1 (black in Fig. 7) the data set is composed of 15 short trajectories with a total of 724 GPS locations and a sampling time step of around 12 min, with some irregularities up to 1 h.

(b) For vessel V2 (grey in Fig. 7) the data set is composed of 25 trajectories with a total of 3111 GPS locations and an irregular sampling time step mostly between 15 and 50 min, with some irregularities up to 4 h.

For each data set, the four methods that were presented above were used to estimate θ. The
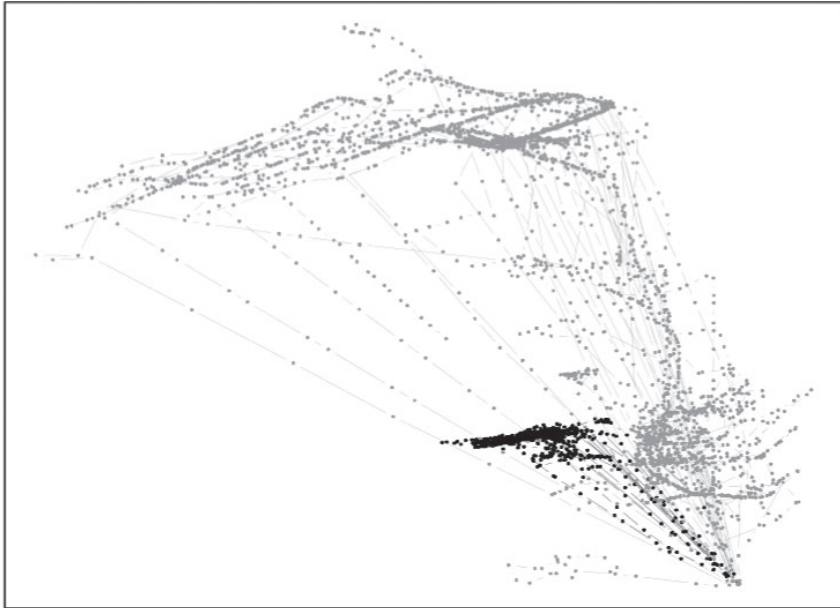
(a)
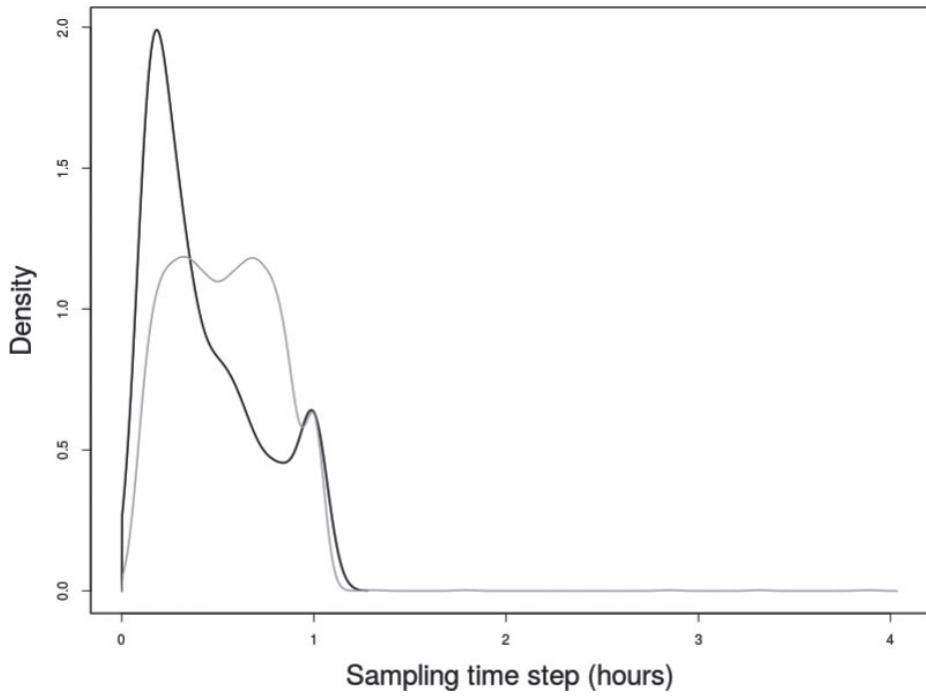
(b)

Fig. 7.   (a) Trajectories for vessel V1 (●) and vessel V2 (●) and (b) distribution of sampling time step for GPS tags for both data sets: for confidentiality, the actual recorded positions may not be shown and therefore no labels are presented on the axes
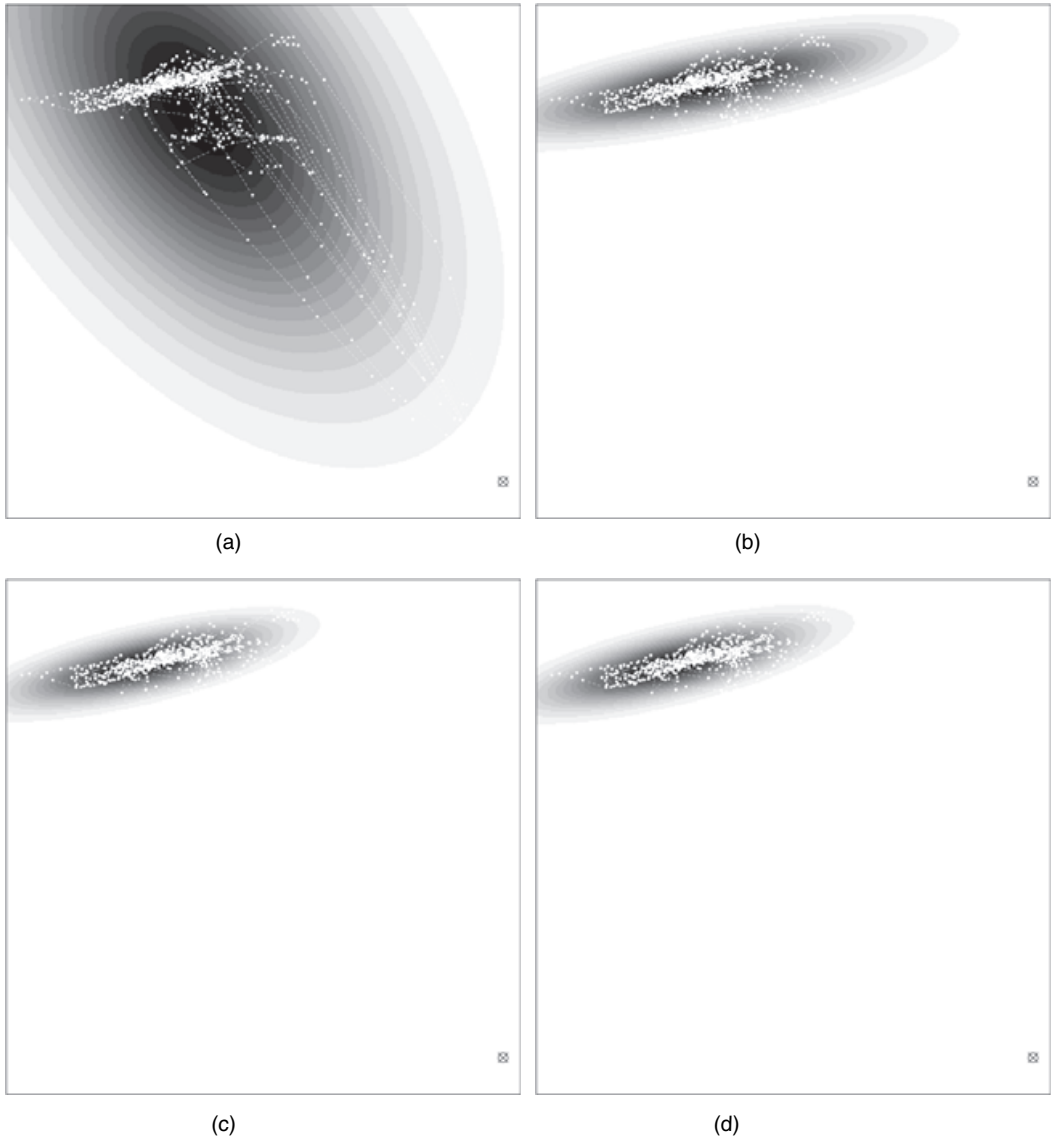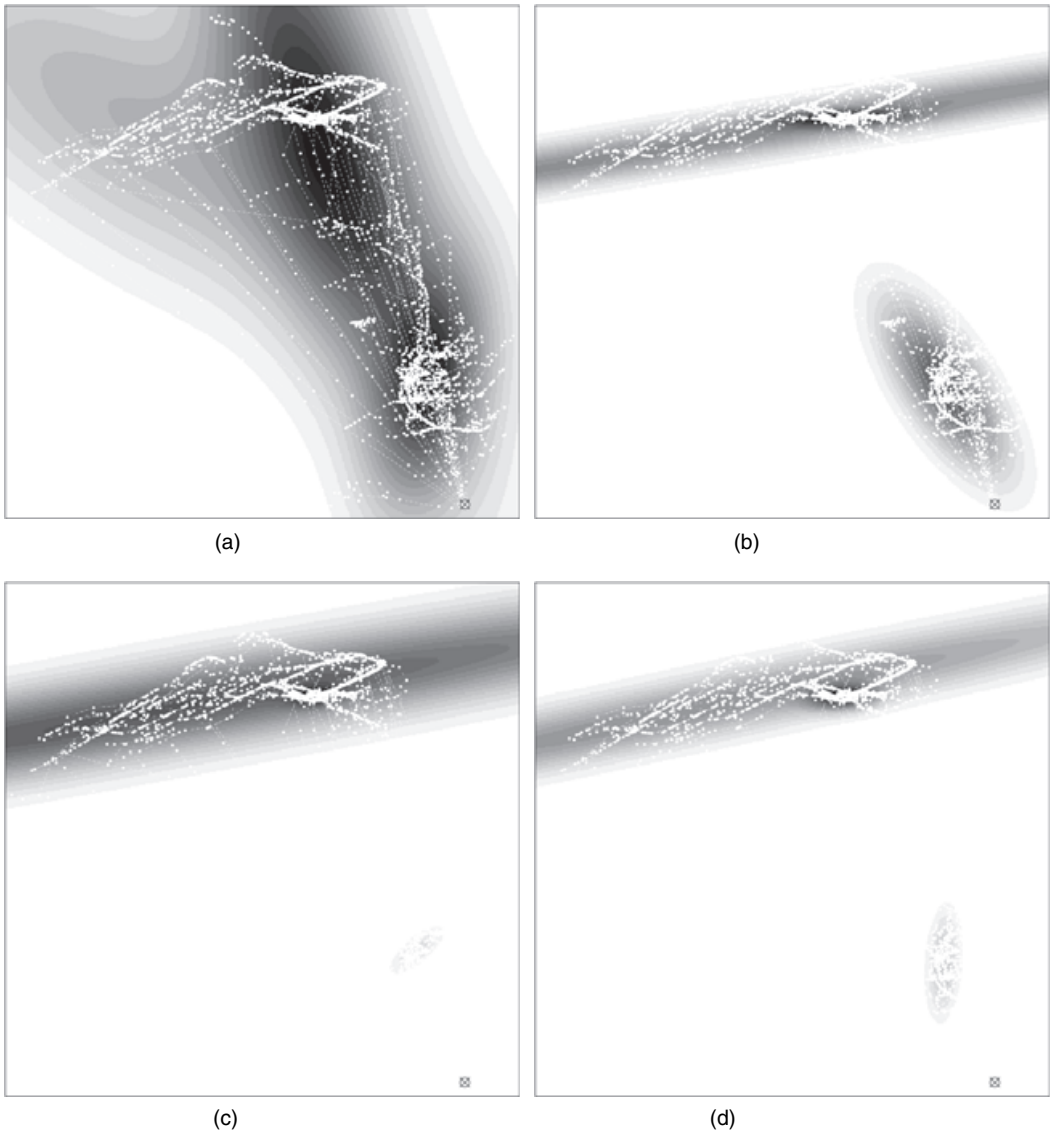
Fig. 8.   Estimated map for the data set of vessel V1 (see Fig. 7), using four different estimation methods (×, departure harbour) (the darker a zone is, the more attractive it is for the given vessel; observed points are plotted in white to see the superposition between maps and trajectories): (a) Euler; (b) Kessler; (c) Ozaki; (d) EAMCEM

number of components K is not assumed to be known anymore and must be estimated. We do not focus on this problem, which has already been addressed in the case of discretely observed SDEs (see Uchida and Yoshida (2005), for instance) and assume that $K_1 = 1$ and $K_2 = 3$ respectively for vessels V1 and V2. In what follows, this approximation of the likelihood proposed by Beskos *et al.* (2009) is used to choose which estimation method provides the best estimated map.

4.2.1.   *Results for vessel V1*
Results for vessel V1 are presented in Fig. 8. The best estimate in the sense of the greatest

(a)

(b)

(c)

(d)

Fig. 9.   Estimated map for the data set of vessel V2 (see Fig. 7), using four different estimation methods (×, departure harbour) (the darker a zone is, the more attractive it is for the given vessel; observed points are plotted in white to see the superposition between maps and trajectories): (a) Euler; (b) Kessler; (c) Ozaki; (d) EAMCEM

approximated log-likelihood is given by the map estimated with the EAMCEM method, followed by the Ozaki, Kessler and Euler methods (see Table 1 for the values of each log-likelihood, for each estimated parameter). Again $\hat{\theta}_K$ is unstable as the correction term for the variance (given in equation (5)) leads to a non-positive semidefinite matrix for 28% of the observations. However, the estimated map is similar to that given by the Ozaki and EAMCEM methods although less concentrated around the data. The acceptance rate for the conditional simulation of the EAMCEM method is 1.4%. The Euler discretization method estimates a potential map with a much more spread attractive zone than those of the three other methods. Moreover, the

orientation of this zone does not follow what seems to be the main axis of trawling (mainly east–west). Both the Ozaki and the EAMCEM methods provide a similar estimated map (in terms of parameters and of likelihood): a Gaussian form wrapping what appears to be the main trawling zone. The axis of the resulting Gaussian form conforms to the trawling directions of the vessels.

4.2.2.   *Results for vessel V2*

Results for vessel V2 are presented in Fig. 9. As in the case of vessel V1, the best estimate in the sense of the greatest approximated log-likelihood is given by the map estimated with the EAMCEM method, followed by the Ozaki, Kessler and Euler methods (see Table 1 for the values of each log-likelihood, for each estimated parameter). Again, $\hat{\theta}_K$ is unstable, as the correction term for the variance leads to a non-positive semidefinite matrix for 48% of observations. The acceptance rate for the conditional simulation of the EAMCEM method is 6.2%. As for vessel V1, the Euler method produces a smooth estimated map, where attractive zones are connected and wrap almost all observed points. The Kessler, Ozaki and EAMCEM methods lead to maps where zones are disconnected: one close to the harbour and an offshore zone. In the case of the Kessler method, the extension and the relative weight of this first zone is larger than for the two other methods. For the three methods, the second offshore zone is a mixture of a general attractive zone with an east–west orientation, and a smaller hotspot that gathers a large amount of observed points. Here, a natural extension would be to use a larger K. The Gaussian model that was proposed by Preisler *et al*. (2013) would have imposed a potential function whose orientation is given by the *x*- and *y*-axis, with circular contour lines as a consequence of the independence between directions. The generalization that is proposed in this paper models more complex attractive zones. The contour lines of the potential might be used to define more realistic high potential fishing zones.

5.   Conclusions

This paper proposes a new model based on an SDE to describe movement in ecology. The model belongs to the popular class of continuous time models described by SDEs, whose drift is supposed to be the gradient of a potential function. This potential function is assumed to reflect the attractiveness of the environment in which the moving individual evolves. We propose a model where this potential is a mixture of smooth Gaussian-shaped functions, giving a flexible framework to describe multimodal potential surfaces. As a counterpart, this model leads to a difficult inference task, as some key quantities to express the likelihood are unknown. This led us to explore several possible approaches for parameters estimation.

   Several extensions of the proposed potential-based model could be considered to capture sharply the individuals' dynamics. For instance, in this paper, the first part of the fishing vessels' trajectories, which consists in steaming towards a fishing zone, seems to be more deterministic than the second part, when the vessel is in the fishing zone. We could use a Markov switching SDE where the state might follow several dynamics depending on a regime indicator given by a finite state space hidden Markov chain. In the case that is presented here, this deterministic phase contains a small proportion of the trajectory and therefore does not affect the inference result, but a better model for this first phase might be preferable.

   We also make the assumptions that observed trajectories are independent. In many cases, this is likely to be false. Therefore, the G observed trajectories could also be modelled as the solution

to a $2G$-dimensional SDE with interactions in the drift function and the diffusion matrix. As the individuals are not observed at the same time, a partially observed SDE can be defined which implies additional challenges to estimate all parameters by using algorithms that were designed for hidden Markov models.

Finally, an undesirable property of the potential that is introduced in the paper is that, far from the attractive zones, the process behaves like Brownian motion. This can be easily overcome for instance by adding another component in the mixture proportional to $-\|\mathbf{x} - G\|^{1/2}$, where G is a fixed attractive point. This ensures that, even if an individual is far from the attractive zones, he is likely to revert around G. We used this potential and obtained very similar numerical results. We decided not to add this component to the model as it seemed artificial to ensure stationarity for theoretical reasons.

From the inference point of view, this work explores the performance of four inference methods to estimate the parameters in such potential-based movement models. The first three methods are pseudolikelihood approaches, including the widely used Euler–Maruyama procedure, with similar implementations and requiring little assumptions on the model. The fourth method, based on a Monte Carlo EM using the exact algorithm, requires additional mathematical assumptions on the model and is more technically and computationally intensive but has the nice property of not using Gaussian approximations, inducing a structural bias.

Numerical experiments with simulated data illustrate that the Euler discretization method is not robust to low sampling rates, which is a common situation in movement ecology, and that the three other approximate maximum likelihood methods perform better, suggesting that the Euler approach might not be well suited to movement ecology.

These four methods were also compared to estimate potential maps (from a specific model) by using actual movement data of two fishing vessels. It is clear here that the different methods can lead to different maps with various interpretations. Specifically, in the examples that were presented here, we would like to point out that the Euler–Maruyama results lead to a totally different map, implying different interpretations, from the three other methods. Therefore, users should be careful when using this method and ensure that their conclusions are not biased by the choice of the Euler discretization. Moreover, using the criterion of the unbiased approximated likelihood of Beskos *et al.* (2009), the Euler method shows the worst results.

Among the alternative likelihood approaches the Kessler approximate method was more unstable than the Ozaki discretization as the proposed approximation of the covariance matrix can be non-positive semidefinite for some observed points. The Kessler method is only used in this work with an expansion of conditional moments up to order 2 and higher order expansions could lead to more robust estimates. The Ozaki method showed results similar to those of the unbiased EAMCEM approach, which were the best according to the approximated likelihood criterion.

Overall, the EAMCEM-based method seems to be the most appealing as it does not introduce any bias and, in our study, was much less sensitive to the starting point than were the pseudolikelihood methods. However, the difficulty of its implementation and the computation cost might reduce its interest. Moreover, its application is restricted to a specific kind of potential functions.

On the basis of simulations and the real data set that were presented in this paper, the Ozaki method seems to provide a close estimate to that given by the EAMCEM method and is as easy as the Euler approach to implement. In addition, it requires only weak assumptions (the drift must have an invertible Jacobian matrix at observed points) and, unlike the Kessler approach, this method is numerically stable. Therefore, we suggest to users in movement ecology the use of the Ozaki method to estimate parameters from movement data.

Acknowledgements

References

Ait-Sahalia, Y. (1999) Transition densities for interest rate and other nonlinear diffusions. *J. Finan.*, **54**, 1361–1395.

Ait-Sahalia, Y. (2002) Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, **70**, 223–262.

Ait-Sahalia, Y. (2008) Closed-form likelihood expansions for multivariate diffusions. *Ann. Statist.*, **36**, 906–937.

Beskos, A., Papaspiliopoulos, O. and Roberts, G. (2006) Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli*, **12**, 1077–1098.

Beskos, A., Papaspiliopoulos, O. and Roberts, G. (2009) Monte Carlo maximum likelihood estimation for discretely observed diffusions processes. *Ann. Statist.*, **37**, 223–245.

Beskos, A., Papaspiliopoulos, O., Roberts, G. O. and Fearnhead, P. (2006) Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion). *J. R. Statist. Soc.* B, **68**, 333–382.

Beskos, A. and Roberts, G. (2005) Exact simulation of diffusions. *Ann. Appl. Probab.*, **15**, 2422–2444.

Blackwell, P. (1997) Random diffusion models for animal movement. *Ecol. Modllng*, **100**, 87–102.

Blackwell, P. G., Niu, M., Lambert, M. S. and LaPoint, S. D. (2015) Exact Bayesian inference for animal movement in continuous time. *Meth. Ecol. Evoln*, **7**, 184–195.

Brillinger, D. (2010) *Handbook of Spatial Statistics*, ch. 26. Boca Raton: Chapman and Hall–CRC.

Brillinger, D., Haiganoush, K., Ager, A., Kie, J. and Stewart, B. (2002) Employing stochastic differential equations to model wildlife motion. *Bull. Braz. Math. Soc.*, **33**, 385–408.

Brillinger, D. R., Preisler, H. K., Ager, A. A. and Kie, J. G. (2001a) The use of potential functions in modelling animal movement. In *Data Analysis from Statistical Foundations* (ed. A. K. M. E. Saleh), pp. 369–386. Huntington: Nova.

Brillinger, D., Preisler, H., Ager, A., Kie, J. and Stewart, B. (2001b) Modeling movements of free-ranging animals. *Statistics Technical Report 610*. University of California at Berkeley, Berkeley.

Brillinger, D., Preisler, H. and Wisdom, M. (2011) Modeling particles moving in a potential field with pairwise interactions and an application. *Braz. J. Probab. Statist.*, **25**, 421–436.

Chang, S.-K. (2011) Application of a vessel monitoring system to advance sustainable fisheries management—benefits received in Taiwan. *Mar. Poly*, **35**, 116–121.

Chavez, A. S. G. E. M. (2006) Landscape use and movements of wolves in relation to livestock in a wildland–agriculture matrix. *J. Wildlif. Mangmnt*, **70**, 1079–1086.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. R. Statist. Soc.* B, **39**, 1–38.

Florens-zmirou, D. (1989) Approximate discrete-time schemes for statistics of diffusion processes. *Statistics*, **20**, 547–559.

Hansen, N. and Ostermeier, A. (2001) Completely derandomized self-adaptation in evolution strategies. *Evoln. Computn*, **9**, 159–195.

Harris, K. J. and Blackwell, P. G. (2013) Flexible continuous-time modeling for heterogeneous animal movement. *Ecol. Modllng*, **255**, 29–37.

Iacus, S. M. (2009) *Simulation and Inference for Stochastic Differential Equations: with R Examples*, vol. 1. New York: Springer Science and Business Media.

Kessler, M. (1997) Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.*, **24**, 211–229.

Kessler, M., Lindner, A. and Sorensen, M. (2012) *Statistical Methods for Stochastic Differential Equations*. Boca Raton: CRC Press.

Li, C. (2013) Maximum-likelihood estimation for diffusion processes via closed-form density expansions. *Ann. Statist.*, **41**, 1350–1380.

Ozaki, T. (1992) A bridge between nonlinear time series models and nonlinear stochastic dynamical systems: a local linearization approach. *Statist. Sin.*, **2**, 113–135.

Preisler, H., Ager, A., Johnson, B. and Kie, J. (2004) Modeling animal movements using stochastic differential equations. *Environmetrics*, **15**, 643–657.

Preisler, H., Ager, A. and Wisdom, M. (2013) Analyzing animal movement patterns using potential functions. *Ecosphere*, **4**, 1–13.

Shoji, I. and Ozaki, T. (1998a) Estimation for nonlinear stochastic differential equations by a local linearization method 1. *Stoch. Anal. Appl.*, **16**, 733–752.

Shoji, I. and Ozaki, T. (1998b) A statistical method of estimation and simulation for systems of stochastic differential equations. *Biometrika*, **85**, 240–243.

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

Uchida, M. and Yoshida, N. (2005) AIC for ergodic diffusion processes from discrete observations. *Preprint MHF 12*. Kyushu University, Kyushu.

Uchida, M. and Yoshida, N. (2012) Adaptive estimation of an ergodic diffusion process based on sampled data. *Stoch. Processes Appl.*, **122**, 2885–2924.