



# Varying-Coefficient Stochastic Differential Equations with Applications in Ecology

Théo MICHELOT , Richard GLENNIE, Catriona HARRIS, and Len THOMAS

Stochastic differential equations (SDEs) are popular tools to analyse time series data in many areas, such as mathematical finance, physics, and biology. They provide a mechanistic description of the phenomenon of interest, and their parameters often have a clear interpretation. These advantages come at the cost of requiring a relatively simple model specification. We propose a flexible model for SDEs with time-varying dynamics where the parameters of the process are nonparametric functions of covariates, similar to generalized additive models. Combining the SDE and nonparametric approaches allows for the SDE to capture more detailed, non-stationary, features of the data-generating process. We present a computationally efficient method of approximate inference, where the SDE parameters can vary according to fixed covariate effects, random effects, or basis-penalty smoothing splines. We demonstrate the versatility and utility of this approach with three applications in ecology, where there is often a modelling trade-off between interpretability and flexibility.

Supplementary materials accompanying this paper appear online.

**Key Words:** Continuous time; Diffusion process; Nonparametric; Smoothing splines; Generalized additive models; Animal movement.

## 1. INTRODUCTION

Stochastic differential equations (SDEs) describe the evolution of a system that involves stochastic noise (Allen 2007). We present a general approach to improving the flexibility of such models. To introduce it, we focus on the most popular form of SDE,

$$dZ_t = \mu(Z_t, t) dt + \sigma(Z_t, t) dW_t, \quad Z_0 = z_0, \quad (1)$$

where  $(W_t)$  is a Wiener process and  $z_0$  is a known initial condition. The terms of the equation describe the evolution of the process  $(Z_t)$ : the drift  $\mu$  measures the expected change in the process over an infinitesimal time interval, and the diffusion  $\sigma$  captures variability. In most

---

Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews KY169LZ, UK  
(E-mail: [tm75@st-andrews.ac.uk](mailto:tm75@st-andrews.ac.uk)).

© 2021 The Author(s)  
*Journal of Agricultural, Biological, and Environmental Statistics*  
<https://doi.org/10.1007/s13253-021-00450-6>

applications, the drift and diffusion are chosen as simple parametric functions; the objective is to estimate their parameters and obtain a mechanistic description of the system. SDEs have, for example, been applied in finance to study asset pricing (Aït-Sahalia and Kimmel 2007), in biology to describe population dynamics (Dennis et al. 1991), and in epidemiology to predict disease spread (Allen and Van den Driessche 2006). Equation 1 includes Brownian motion, geometric Brownian motion, and the Ornstein–Uhlenbeck process as special cases.

SDEs are used to formulate a (simplified) description of a stochastic system, and the challenge is to build models flexible enough to reflect features of the system within the assumed structure of Eq. 1. For this purpose, there has been interest in specifying SDEs with time-varying dynamics. Regime-switching models have been developed, where a process switches between a finite number of SDEs, often based on an underlying continuous-time Markov chain (Mao and Yuan 2006). These models combine the convenience of simple parametric models with the flexibility provided by multiple regimes and have been used to describe, for example, the movement of animals switching between behavioural states (Blackwell 1997) or the time-varying dynamics of oil prices (Liechty and Roberts 2001). An alternative is to specify the parameters of a SDE as continuous-valued random processes (e.g. Duan et al. 2009); e.g. in stochastic volatility models, the variance parameter of the diffusion function is itself specified as a diffusion process to account for changes in the variability (Aït-Sahalia and Kimmel 2007). Other approaches have been developed to model the drift and diffusion of SDEs as nonparametric functions of time or of the value of the process  $Z_t$ , e.g. using Gaussian processes (Archambeau et al. 2007) or orthogonal Legendre polynomials (Rajabzadeh et al. 2016). For a particular class of SDEs applied to animal movement studies, Preisler et al. (2004) and Russell et al. (2018) suggested using splines to model the dependence of SDE parameters on spatial covariates.

We propose a general approach where the parameters of a SDE are specified as basis-penalty smoothing splines, similar to generalized additive models [GAMs; (Wood 2017)]. This allows for a rich class of models including linear covariate effects, factor variables, independent random effects, and smooth (nonparametric) covariate effects. It stands in contrast to the regime-switching models where parameters are piecewise constant rather than smooth. It generalizes the models where parameters are specified as Gaussian processes, given the equivalent interpretation of Gaussian processes and smoothing splines (Wood 2017, Sect. 5.8.2), and it also extends ideas from Preisler et al. (2004) and Russell et al. (2018) to allow for flexible covariate dependence in a more general class of SDEs. We develop a method of inference, using a smoothing penalty in the SDE likelihood to control the roughness of nonparametric terms. We present a computationally efficient implementation based on the R packages *mgcv* and *TMB*, for model specification and model fitting, respectively.

We illustrate the potential of this new framework using three case studies from ecology. SDEs have great theoretical and practical appeal for the analysis of ecological data, because their continuous-time formulation does not depend on the sampling resolution of the data. Inferences from these models can therefore be compared across studies with different sampling schemes, and they can be fitted to data collected at irregular time intervals (e.g. Michélot and Blackwell 2019). Despite their advantages, continuous-time models have been underutilized in this field, in part because they have lacked flexibility to specify time-

varying dynamics and covariate effects, or have required computationally costly model fitting procedures. The three case studies illustrate the utility of the new model over existing parametric approaches and highlight its flexibility and computational convenience. In the supplementary materials, we give implementation details, a simulation study, another case study from finance, and the source code to reproduce all analyses presented in the paper. The method that we describe is implemented in an R package, available at <https://github.com/TheoMichelot/smoothSDE>.

## 2. VARYING-COEFFICIENT STOCHASTIC DIFFERENTIAL EQUATIONS

### 2.1. MODEL FORMULATION

We consider a stochastic process  $(Z_t)$  defined by

$$dZ_t = \mu(Z_t, \theta_t) dt + \sigma(Z_t, \theta_t) dW_t, \quad (2)$$

where the drift  $\mu$  and diffusion  $\sigma$  depend on a time-varying parameter vector  $\theta_t$ . We assume that  $\mu$  and  $\sigma$  are known functions of  $Z_t$  and  $\theta_t$ ; they determine the type of stochastic process (e.g. Brownian motion, Ornstein–Uhlenbeck process). The parameter  $\theta_t$  depends on time through its relationship with  $J$  temporal covariates  $x_{1t}, x_{2t}, \dots, x_{Jt}$ , and we write each component  $\theta_t$  of  $\theta_t$  as

$$h(\theta_t) = \beta_0 + f_1(x_{1t}) + f_2(x_{2t}) + \dots + f_J(x_{Jt}),$$

where  $h$  is a link function,  $\beta_0$  is an intercept parameter and, for  $j = 1, \dots, J$ ,  $f_j$  could be a linear effect of a covariate, an independent random effect, or a smooth function. A simple example would be to have  $x_{1t} = t$ , to express that the dynamics of the process depend on time. For smooth functions or random effects, we employ the basis-penalty approach (Wood 2017), writing the functions as linear combinations of  $m_j$  basis functions  $\{\psi_{jk}\}$ ,

$$f_j(x) = \sum_{k=1}^{m_j} \beta_{jk} \psi_{jk}(x), \quad (3)$$

where several standard bases could be considered, e.g. cubic splines, thin plate regression splines, or B-splines. We will refer to this model as a varying-coefficient stochastic differential equation, as an analogy with the varying-coefficient models of Hastie and Tibshirani (1993).

### 2.2. MODEL FITTING

We consider  $n$  observations  $(z_1, z_2, \dots, z_n)$  from the process  $(Z_t)$ , collected at (possibly irregular) times  $t_1 < t_2 < \dots < t_n$ . The aim is to estimate the relationship between the parameters  $\theta_t$  governing the drift and diffusion of the process and the covariates. The method

that we propose is based on (1) the likelihood of the observations under the SDE model and (2) a penalty added to the likelihood to control the roughness of nonparametric terms in  $\theta_t$ .

### 2.2.1. Likelihood

Diffusion processes are Markovian, so the likelihood of  $n$  observations can be obtained as the product of the likelihoods of the individual transitions,

$$\begin{aligned} L(\alpha, \beta | z_1, \dots, z_n) &= [Z_{t_1} = z_1, \dots, Z_{t_n} = z_n] \\ &= [Z_{t_1} = z_1] \prod_{i=1}^{n-1} [Z_{t_{i+1}} = z_{i+1} | Z_{t_i} = z_i], \end{aligned} \quad (4)$$

where  $\beta$  contains the basis coefficients from Eq. 3 and  $\alpha$  is the vector of other parameters of the model (e.g. linear covariate effects), and where  $[\cdot]$  is the pdf. The dependence on  $\alpha$  and  $\beta$  is omitted in the right-hand side of Eq. 4 for notational simplicity. We assume that the first value  $z_1$  is deterministic, such that  $[Z_{t_1} = z_1] = 1$ .

Evaluating the likelihood requires computation of the transition density  $[Z_{t_{i+1}} | Z_{t_i}]$  of the process. For many common processes, such as Brownian motion, geometric Brownian motion, and the Ornstein–Uhlenbeck process, this density has an analytical expression. In such cases, the transition density of the corresponding varying-coefficient process can be approximated by assuming that the parameter  $\theta_t$  is fixed over each time interval of observation. Then, the time-varying parameter  $\theta_{t_i}$  can be substituted into the transition density of the standard process. However, many SDEs of the form given in Eq. 2 do not have a closed-form transition density. More generally, we can then use the Euler–Maruyama discretization, and approximate the transition density  $[Z_{t_{i+1}} | Z_{t_i} = z_i]$  by the pdf of a normal distribution with mean  $z_i + \mu(z_i, \theta_{t_i})\Delta_i$  and variance  $\sigma(z_i, \theta_{t_i})^2\Delta_i$ , where  $\Delta_i = t_{i+1} - t_i$ . This approximation assumes that the drift and diffusion terms are constant over each interval  $[t_i, t_{i+1})$  between two observations. We present the varying-coefficient versions of several common processes in Appendix A and give their approximate transition densities. Substituting the approximate transition density into Eq. 4 yields the approximate likelihood for the full data set.

This method of inference is not exact, because it uses the transition density of the time-discretized diffusion process. The Euler–Maruyama discretization has the advantage of being widely applicable and easy to implement, but the accuracy of the estimation will decrease for longer time intervals between observations. To mitigate the effects of this approximation, we could include additional time points in the time series of observed data, corresponding to “missing” observations, and integrate over them, e.g. using either Markov chain Monte Carlo methods or the Laplace approximation (Elerian et al. 2001; Albertsen 2019). Adding these missing values to the grid of observations leads to a finer time resolution and improves the accuracy of the approximation, such that the error can be made arbitrarily small.

The process  $(Z_t)$  might sometimes not be observed directly, in which case the problem of inference is slightly different. This can be viewed as a state-space model, where the state equation is given by the transition density  $[Z_{t_{i+1}} | Z_{t_i}]$  (e.g. obtained using the Euler approximation), and the observation equation is the density  $[\tilde{Z}_t | Z_t]$ , where  $\tilde{Z}_t$  denotes the

observations. In this context,  $\tilde{Z}_t$  could, for example, include measurement error or be a more general function of  $Z_t$ . The diffusion process of interest is a latent process in the model, and it must be marginalized over to obtain the likelihood of the observed data. In the case of a Gaussian linear state-space model, the Kalman filter can be implemented, with time-varying parameters, and the likelihood obtained as a by-product. In this case, the Kalman filter can also be used to integrate over missing data in a data augmentation scheme to improve the discretization approximation. One example of a latent-state SDE is the velocity Ornstein–Uhlenbeck model described by [Johnson et al. \(2008\)](#), where the observed process (location) is the integral of a diffusion process (velocity). In that model, the location process is smooth, and it is therefore convenient to describe the persistent movement of an animal or particle, or other processes with strong autocorrelation. Non-Gaussian state-space models can also be accommodated using Markov chain Monte Carlo or the Laplace approximation to marginalize over the state process, as suggested, for example, by [Albertsen et al. \(2015\)](#). We present two examples of state-space SDE models in Sect. 3.

### 2.2.2. Smoothing Penalty

Within the basis-penalty approach of GAMs, the roughness of the smoothing splines can be penalized in the likelihood, to obtain smooth relationships between the parameter  $\theta_t$  and the covariates. The penalized log-likelihood is

$$l_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}|z_1, \dots, z_n) = \log\{L(\boldsymbol{\alpha}, \boldsymbol{\beta}|z_1, \dots, z_n)\} - \sum_j \lambda_j \boldsymbol{\beta}_j^T \mathbf{S}_j \boldsymbol{\beta}_j, \quad (5)$$

where  $L(\boldsymbol{\alpha}, \boldsymbol{\beta}|z_1, \dots, z_n)$  is the unpenalized likelihood given in Eq. 4,  $\boldsymbol{\beta}_j$  is the vector of basis coefficients,  $\mathbf{S}_j$  is the smoothing matrix associated with the chosen penalty, and  $\lambda_j$  is a smoothness parameter for the  $j$ -th smooth term in  $\theta_t$  ([Wahba 1990](#)).  $\mathbf{S}_j$  is a matrix of known coefficients, and it is constructed such that  $\boldsymbol{\beta}_j^T \mathbf{S}_j \boldsymbol{\beta}_j$  measures the roughness (wiggleness) of the corresponding smooth term ([Wood 2017](#)). The penalized log-likelihood can then be used to perform maximum likelihood estimation, or Bayesian inference can be performed if the penalty is viewed as an improper prior on the basis coefficients.

In Eq. 5, the penalized log-likelihood is expressed in terms of the degrees of smoothness  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots)$  of the smoothing splines. In most applications,  $\boldsymbol{\lambda}$  is unknown, and it must be estimated from the data. Here, we consider the marginal likelihood approach, i.e. we treat the basis coefficients  $\boldsymbol{\beta}$  of the splines as random effects, and integrate them out of the likelihood. This yields the marginal likelihood of the smoothness parameters  $\boldsymbol{\lambda}$  and other fixed parameters  $\boldsymbol{\alpha}$ ,

$$L(\boldsymbol{\alpha}, \boldsymbol{\lambda}|z_1, \dots, z_n) = \int L(\boldsymbol{\alpha}, \boldsymbol{\beta}|z_1, \dots, z_n) [\boldsymbol{\beta}|\boldsymbol{\lambda}] d\boldsymbol{\beta}, \quad (6)$$

where  $[\boldsymbol{\beta}|\boldsymbol{\lambda}]$  is the density of a multivariate normal distribution with mean zero and block-diagonal precision matrix. Each block of the precision matrix corresponds to the penalty for the basis coefficients of one smoothing spline, and it can be written  $\lambda_j \mathbf{S}_j$ . As with standard GAMs, various basis-penalty smooths could be used. In the applications of Sect. 3, we considered thin plate regression splines, which are optimal in the sense defined by [Wood](#)

(2003), with a shrinkage penalty to ensure that the smooth terms shrink to zero when the penalty tends to infinity (Marra and Wood 2011).

### 2.2.3. Implementation

The marginal likelihood can be implemented in the R package TMB, which uses the Laplace approximation to integrate over the random effects (Kristensen et al. 2016), and the design matrices for the basis functions and the penalty matrices can be computed with the R package mgcv (Wood 2017). A numerical optimizer (e.g. `optim` or `nlm`) can then be used to minimize the marginal likelihood and obtain estimates of the smoothness parameter  $\lambda$  and fixed effects  $\alpha$ . We can get predicted values for the random effects  $\beta$ , analogous to best linear unbiased predictors in linear mixed effect models, to infer the smooth relationships between SDE parameters and covariates. The joint precision matrix of fixed and random effects can be used for uncertainty quantification.

TMB makes it relatively simple to include other random effects in the model. It is often the case, e.g. in animal movement or financial studies, that the data arise from multiple instances of the SDE (multiple animals, stocks, etc.) and one wishes to fit a model that combines these instances while also allowing for inter-individual variation. The case of i.i.d. normal random effects is easily handled, as it is another type of basis-penalty smoother, where the penalty matrix is the identity matrix (Wood 2017, Sect. 7.7).

We describe the details of the implementation of this method, using mgcv and TMB, in Appendix B. We ran simulation experiments to investigate the performance of the proposed approach to recover the relationship between the SDE parameters and the covariates, under several model formulations. In those simulations, we thinned the simulated data to irregular time intervals, to mimic a real data set, and the method performed well in all scenarios (Appendix C). We also performed a simulation experiment to check the coverage of confidence intervals derived for  $\theta_t$  using the precision matrix given by TMB and found that they correctly represented the uncertainty in the estimates (Appendix C).

## 2.3. MODEL SELECTION AND MODEL CHECKING

In this framework, it might be useful to discriminate between competing model formulations, e.g. different forms of the drift and diffusion terms. The problem of model selection in models involving basis-penalty smooths is relatively understudied outside standard GAMs. Wood (2017) describes two versions of the Akaike information criterion (AIC), the *marginal* AIC and the *conditional* AIC, based on different forms of the likelihood and AIC penalty (i.e. number of parameters in basic AIC). The marginal AIC uses the marginal likelihood defined in Eq. 6 with a penalty for the number of fixed effects  $\alpha$  and smoothing parameters  $\lambda$  and would be straightforward to implement in the framework of Sect. 2.2.3. The conditional AIC is based on the joint penalized likelihood of fixed and random effects (Eq. 5), with an additional AIC penalty on the complexity of smooth terms [given by the number of effective degrees of freedom; see Sect. 5.4.2 of Wood (2017)]. For more detail about the respective limitations of these two criteria and possible solutions, see Sect. 6.11 of Wood (2017). An alternative approach for model selection is to include an additional penalty in the likelihood,

so that model components can be shrunk to zero (i.e. removed from the model) as part of the smoothness parameter estimation (Marra and Wood 2011).

For a chosen formulation, we propose a simple diagnostic to investigate goodness-of-fit in varying-coefficient SDE models. Based on the Euler–Maruyama discretization of the process, a natural choice for model residuals  $\epsilon_i$  is

$$\epsilon_i = \frac{z_{i+1} - (z_i + \mu(z_i, \hat{\theta}_{t_i})\Delta_i)}{\sigma(z_i, \hat{\theta}_{t_i})\sqrt{\Delta_i}},$$

for  $i = 1, \dots, n - 1$ , using the notation of Sect. 2.2.1, and where  $\hat{\theta}_{t_i}$  is the estimate of  $\theta_t$  over  $[t_i, t_{i+1})$ . Under the assumptions of the model (and of the discretization), the residuals should be independent and approximately follow a standard normal distribution. In the analysis of Sect. 3.3, we use quantile–quantile plots of the residuals to investigate lack of fit, and autocorrelation function plots to identify residual autocorrelation.

### 3. ILLUSTRATIVE EXAMPLES

In this section, we present three analyses based on ecological data, to illustrate different applications of the models presented in Sect. 2. We stress, however, that the varying-coefficient approach is general to SDE modelling, and our focus is chosen only because we are most familiar with these ecological problems. To further demonstrate the generality of the method, we also provide the analysis of a financial data set of oil prices in Appendix D.

#### 3.1. LINKING ELEPHANT MOVEMENT TO ENVIRONMENTAL CONDITIONS

We illustrate the utility of varying-coefficient SDEs to analyse animal movement data, using the trajectory of an African elephant (*Loxodonta africana*) presented by Wall et al. (2014b) and available on the Movebank data repository (Wall et al. 2014a). We restricted the analysis to a period from May to September 2009 to avoid seasonal effects. The data set consisted of a time series of 3652 Easting–Northing locations and also included the air temperature measured by the tag, at a time resolution of 1 hour (with a few missing observations).

We used a varying-coefficient version of the continuous-time correlated random walk model presented by Johnson et al. (2008); the original model has been used extensively to analyse animal location data. In this model, the (unobserved) velocity  $V_t$  of the animal is formulated as a varying-coefficient Ornstein–Uhlenbeck process,

$$dV_t = -r_t V_t dt + s_t dW_t,$$

where  $r_t$  and  $s_t$  can be linked to the speed and sinuosity of the movement. This is a special case of the varying-coefficient SDE of Eq. 2 where  $\mu(V_t, t) = -r_t V_t$  and  $\sigma(V_t, t) = s_t$ , i.e. with parameters  $\theta_t = (r_t, s_t)$ . (Although the process  $V_t$  is bivariate, it is isotropic and the two dimensions can therefore be treated as two univariate processes driven by

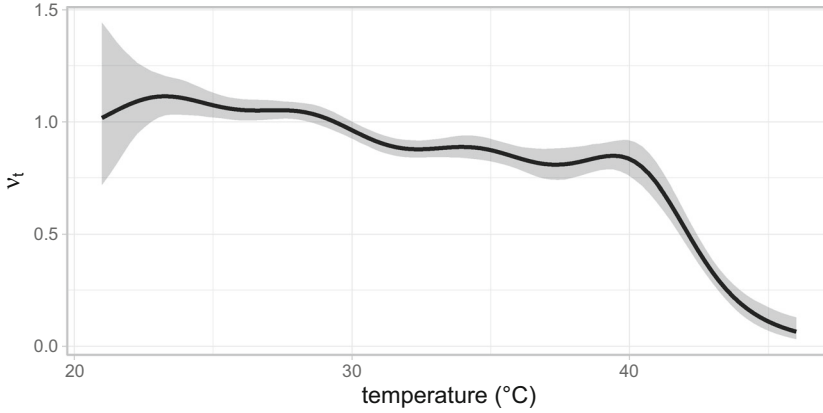


Figure 1. Results of the elephant analysis. Estimates of the speed parameter  $v_t = \sqrt{\pi s_t} / (2\sqrt{r_t})$ , as a function of temperature. The black line is the mean estimate, and the grey shaded area is a 95% confidence band .

the same parameters.) Because the velocity is unobserved, the model can be written as a state-space model where  $V_t$  is latent, and where the observed process is the location of the animal, obtained as  $Z_t = Z_0 + \int_0^t V_s ds$ . As described in Sect. 2.2.1, we implemented the likelihood using a Kalman filter with time-varying parameters. To investigate the effects of environmental conditions on the elephant’s behaviour, we estimated the parameters  $r_t$  and  $s_t$  of the velocity process as functions of the air temperature. For interpretation, we then derived the parameter  $v_t = \sqrt{\pi s_t} / (2\sqrt{r_t})$ , described by Gurarie et al. (2017) as a measure of the speed of movement of the animal. Model fitting took about 5 min on a 1.3 GHz Intel i7 CPU.

Figure 1 shows estimates of the speed parameter  $v_t$  as a function of temperature. A small value of  $v_t$  corresponds to slow movement, encompassing behaviours with little activity (e.g. resting), and a large value corresponds to more active behaviours (e.g. exploration, transit). The speed parameter was highest at low temperatures (20–30 degrees) and decreased for higher temperatures, with a very steep decline above 40 degrees. This is consistent with what is known of the species: elephants are very sensitive to heat and spend much of their time resting or waiting in the shade during periods of high temperatures (Mole et al. 2016). Our nonparametric approach illuminated the nonlinear relationship between the temperature and the speed of movement of this elephant.

Other varying-coefficient models have been proposed in movement ecology to capture the effects of time-varying covariates on animal behaviour. In particular, Hanks et al. (2015) modelled animal movement over a discrete spatial grid as a continuous-time Markov chain with time-varying transition rates and linked transition rates to covariates using basis functions (similarly to Eq. 3). This continuous-time discrete-space model and its extensions have, for example, been used to investigate the effects of time-varying environmental conditions on the movements of cougars (Hanks et al. 2015; Buderman et al. 2018), fur seals, and ants (Hanks and Hughes 2016). The varying-coefficient SDEs presented in this paper offer an alternative framework to incorporate similar covariate effects in the dynamics of continuous-valued random processes (e.g. the continuous location or velocity of an animal).



SDEs have also been popular for analysing animal tracking data, e.g. to model animal behaviour (Blackwell 1997), the effects of environmental features on movement decisions (Preisler et al. 2004; Michelot et al. 2019), and the emergence of home ranges (Dunn and Gipson 1977). Regime-switching SDEs have been developed to allow for time-varying dynamics in the movement, where the latent state represents the behaviour of the animal (Blackwell 1997; Michelot and Blackwell 2019). However, discrete behavioural states may lack the flexibility to capture the wide range of behaviours that animals display. It has also been difficult to include general covariate effects in that context; inference has typically required computationally costly custom algorithms (Blackwell et al. 2016). The method we propose to include factor covariates, linear or smooth effects of continuous covariates, and random effects in SDE models is an important step forward to link animal movement behaviour to environmental and individual-specific conditions.

The output of regime-switching models (classification of data into clusters) may sometimes be more readily interpretable than the smoothly varying parameters suggested here. In such cases, we could use a clustering algorithm on the estimated  $\theta_t$  values to identify different regimes in the time series and interpret each cluster based on its centre, say. This procedure could be repeated on posterior draws of  $\theta_t$  to account for uncertainty in the clustering.

Preisler et al. (2004) and Russell et al. (2018) presented application-specific methods to define smooth relationships between movement parameters and spatial covariates in SDEs. The approach presented in this paper is a generalization of their work to a wider class of SDEs, where any parameters can be specified using basis-penalty smooths. Note that the varying-coefficient CTCRW model could be fitted with the R package *crawl* (Johnson et al. 2008; Johnson and London 2018), using the joint likelihood of all parameters rather than the marginal likelihood of Eq. 6. That package does not implement the smoothness parameter estimation, and this would need to be done in an additional model selection stage.

### 3.2. BODY CONDITION OF ELEPHANT SEALS

We considered a study of body condition of elephant seals described by Schick et al. (2013), where the authors modelled body fat content over time, and how it was affected by environmental conditions. We used the data set from Pirota et al. (2019), which includes information about drift dives of 26 Northern elephant seals (*Mirounga angustirostris*). The goal of the study was to investigate the dynamics of the body fat content of seals during migratory foraging trips, which last several months. Animals' body fat content cannot be observed directly when they are at sea. However, using telemetry tags fitted with depth sensors, we can measure the rate at which seals drift vertically in the water column during non-active dives ["drift dives"; (Biuw et al. 2003)]. This drift rate is linked to the percentage of body fat because fat content affects buoyancy. A natural modelling approach, proposed by Schick et al. (2013), is therefore to treat body fat content as a latent process in a state-space model and estimate how it changes during a foraging trip from the drift rate observations.

We formulated a continuous-time analogue of the model of Schick et al. (2013) and defined the body fat content  $L_t$  as a Brownian motion with time-varying drift, i.e.  $dL_t = r_t dt + \sigma dW_t$  (Eq. 2 with  $\mu(L_t, t) = r_t$  and  $\sigma(L_t, t) = \sigma$ ). The time-varying parameter

$r_t$  measured the daily rate of change of the lipid content, with larger values indicating faster accumulation of fat mass. We combined the above SDE with the observation equation proposed by [Schick et al. \(2013\)](#) to obtain the following state-space model formulation

$$\begin{aligned} \text{Observation process } D_i &\sim N\left(\alpha_1 + \alpha_2 \frac{L_i}{R_i}, \frac{\tau^2}{h_i}\right) \\ \text{State process } L_{i+1} &\sim N(L_i + r_i \Delta_i, \sigma^2 \Delta_i) \end{aligned}$$

where  $i$  is the day index,  $D_i$  is the mean drift rate,  $L_i$  is the lipid content,  $R_i$  is the non-lipid content,  $h_i$  is the number of drift dives, and  $\alpha_1$ ,  $\alpha_2$  and  $\tau$  are parameters of the observation process. We followed [Schick et al. \(2013\)](#) in assuming a constant diffusion  $\sigma$  and investigated the effects of two covariates on the lipid change rate  $r_t$ : (1) surface transit per day and (2) distance to the colony where the animals were tagged. We chose these covariates to link fat gains (i.e. foraging behaviour) to movement patterns and geographical location. In preliminary analyses, we included two other covariates from [Schick et al. \(2013\)](#) (body fat proportion at departure and daily number of drift dives), but found no evidence of an effect. We included a random normal intercept in  $r_t$  to account for differences between seals.

We implemented the likelihood of this model with the Kalman filter, and estimated the effects of the covariates on the drift parameter following Sect. 2. [Schick et al. \(2013\)](#) fitted their state-space model within a Bayesian framework, using informative priors for  $\tau^2$  and  $\sigma^2$  based on biological knowledge. We included the same prior distributions as multiplicative terms in the likelihood, therefore performing maximum posterior estimation for those parameters. Model fitting took about 2 min on a 1.3 GHz Intel i7 CPU.

Results are shown in Fig. 2. The lipid gain rate  $r_t$  was estimated to decrease with daily transit distance, which is consistent with the findings of [Schick et al. \(2013\)](#). This suggests that lipid gains are low when seals are travelling at high speeds, and that foraging is characterized by less horizontal movement. We also found that lipid gains increased with distance to the colony, in particular between 0 and 2000 km. This indicates that animals must travel a considerable distance from their breeding colony to find foraging grounds that are rich enough for them to start gaining fat. Figure 2 shows a map of the movement tracks of the seals, coloured by the predicted value of  $r_t$ , which highlights portions of the trips with high lipid gains. Our results provide a mechanistic justification for the assumption often made in elephant seal studies that slow horizontal movement at sea is associated with foraging behaviour (e.g. [Michélot et al. 2017](#)).

This application shows how SDEs can be built as alternatives to discrete-time models [such as the state-space model of [Schick et al. \(2013\)](#)], which do not depend on the time resolution of the data and can be applied to data collected at irregular intervals. Another difference with [Schick et al. \(2013\)](#) is that we implemented the Kalman algorithm and Laplace approximation (with TMB) to integrate over latent components of the model, whereas they used computationally costly Markov chain Monte Carlo methods.

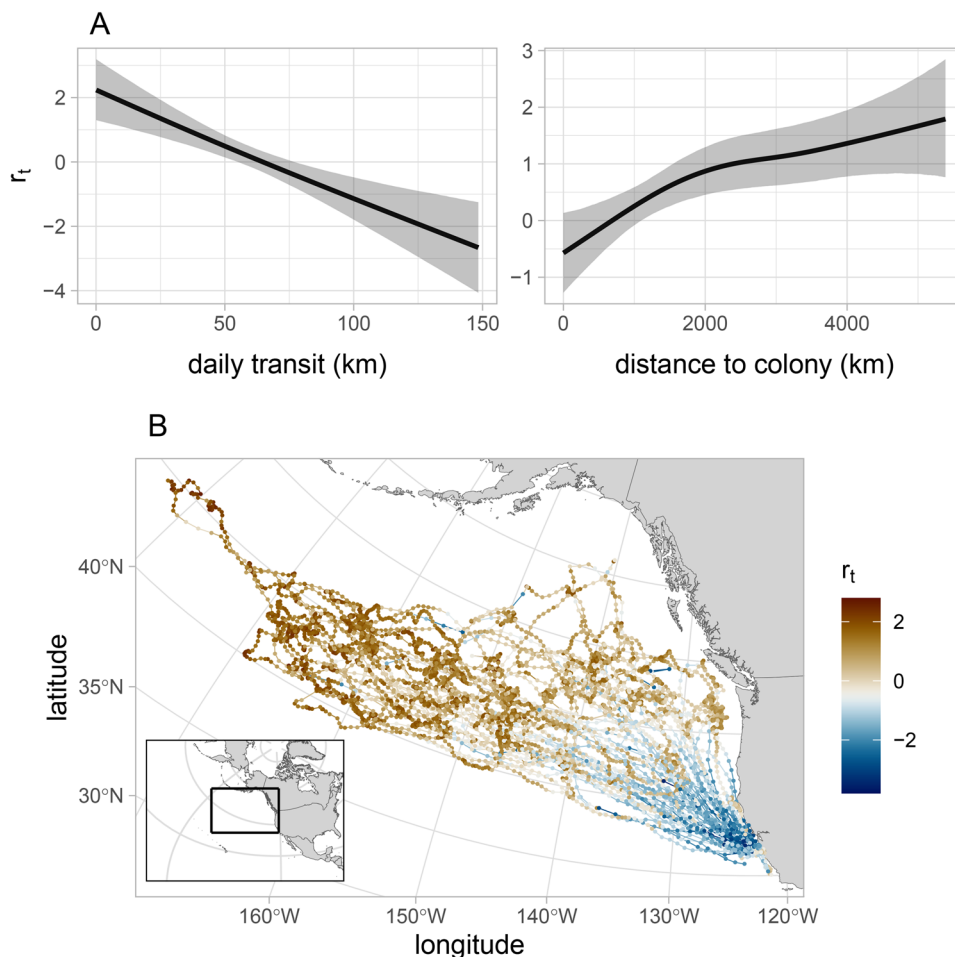


Figure 2. Results of elephant seal body condition study. **a** Estimated relationship between lipid gain parameter  $r_t$  and two covariates: daily transit distance (left) and distance to colony (right). The black lines are mean estimates, and the grey shaded areas are 95% confidence bands. Each estimate was obtained by fixing the other covariate to its mean. **b** Elephant seal movement tracks, coloured by predicted value of  $r_t$ . This figure appears in colour in the electronic version of this article .

### 3.3. DIVING BEHAVIOUR OF BEAKED WHALES

Beaked whales are marine mammals that routinely dive to depths in excess of 1 km for periods of over an hour. Animal-borne telemetry tags allow study of their diving behaviour (Johnson and Tyack 2003; DeRuiter et al. 2013). Here, we consider data collected from high-resolution tags that include accelerometer and magnetometer sensors [“DTAGs”; (Johnson and Tyack 2003)], attached to four Cuvier’s beaked whales (*Ziphius cavirostris*). In general, beaked whales display two different types of dives with different physiological functions: deep and shallow dives. The structures of deep and shallow dives are very distinct and would require separate models. Here, we focused on shallow dives and also excluded sections of the data where the animals were at the sea surface (depth < 15 m). The data set comprised  $n_d = 73$  shallow dives from the four whales, with a median duration of 23 min. Multiple

variables can be derived from DTAG data, and we computed the Euler angles (pitch, roll, and heading), which describe the posture of the animal in the water (Johnson and Tyack 2003). The pitch is the angle between the main body axis and the horizontal, the roll is the angle around the main body axis, and the bearing is the angle in the horizontal plane (Figure S4 of supplementary material). The sampling rate of the raw data varied between 5 and 25 Hz, and we downsampled by taking averages over non-overlapping 5-s windows, to reduce the computational cost while keeping a sufficiently fine resolution to detect behavioural changes over each dive. This resulted in a total of 20,041 observations for each variable.

Visual inspection of the data suggested that shallow dives all had a similar structure, with different phases of each dive displaying different levels of activity. Our aim was therefore to characterize the typical behaviour of beaked whales, as measured by their postural dynamics, during the different diving phases (e.g. descent, ascent, bottom). In preliminary analyses, we tried modelling each variable (pitch, roll, heading) with Brownian motion, but residual analysis revealed that the model did not capture heavy tails in increments of the process. We therefore replaced the Gaussian transition density with a generalized t distribution with fixed degrees of freedom ( $\nu = 3$ , based on visual data exploration) and estimated the location parameter  $r_t$  and the scale parameter  $s_t$  as time-varying. Here, we used the Euler–Maruyama discretization of the process (rather than the SDE itself) as the “model of record” to build a more complex model, as suggested by Brillinger (2010) in a similar context. Although this model is not a special case of the SDE given in Eq. 2, the method described in Sect. 2 can be applied, with the likelihood defined as the pdf of a t distribution for each observed transition. The details of the model are described in Appendix E of the supplementary material. For this example, the three processes were treated as independent.

The model had two parameters: the location  $r_t$  and scale  $s_t$  of the t-distributed increments. To investigate the time-varying behaviour of beaked whales, we specified  $r_t$  and  $s_t$  as smooth functions of the proportion of time through the dive  $x_t \in [0, 1]$ . We treated the dives as independent and included random intercepts for the dive in  $r_t$  and  $s_t$ , to account for variability between individuals and between dives. In summary, there were six time-varying parameters ( $r_t$  and  $s_t$  for each of the three data variables), modelled for each variable as

$$\begin{aligned} r_t &= \zeta_{d_t} + f_r(x_t), & \zeta_j &\sim N(\mu_\zeta, (\sigma_\zeta)^2) \text{ for } j \in \{1, 2, \dots, n_d\}, \\ \log(s_t) &= \xi_{d_t} + f_s(x_t), & \xi_j &\sim N(\mu_\xi, (\sigma_\xi)^2) \text{ for } j \in \{1, 2, \dots, n_d\}, \end{aligned}$$

where  $d_t \in \{1, 2, \dots, n_d\}$  is the dive index at time  $t$ ,  $f_r$  and  $f_s$  are basis-penalty smooths, and  $\{\mu_\zeta, \sigma_\zeta, \mu_\xi, \sigma_\xi\}$  are unknown hyper-parameters. Model fitting took 30 min on a 1.3 GHz Intel i7 CPU. The estimated relationships between the parameters and the proportion of time through the dive are shown in Fig. 3.

The estimated drift parameter  $r_t$  for pitch was positive over the whole dive, suggesting that pitch tended to increase during a typical dive. This is consistent with the observed convex shape of the dives: pitch increases between the descent and bottom phases, and again between the bottom and ascent phases. The estimated drift for roll was close to zero and did not seem to be affected by the phase of the dive, suggesting that there were no particular trend in that variable. The estimated drift in heading was negative during the final part of the dive (ascent), but the confidence bands were wide and overlapped

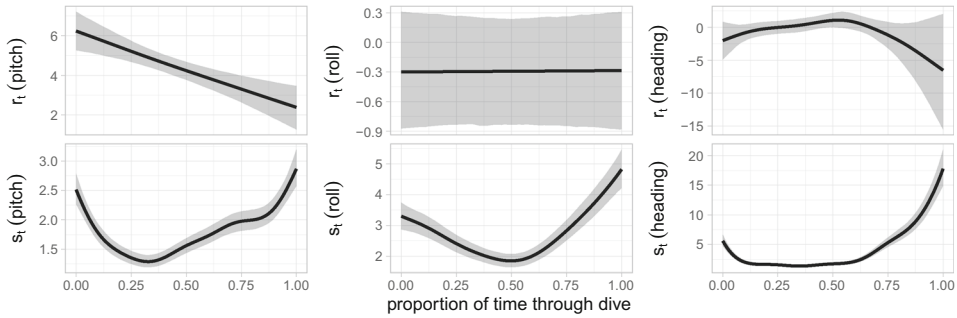


Figure 3. Estimates of the drift ( $r_t$ , top row) and diffusion ( $s_t$ , bottom row) parameters as functions of the proportion of time through the dive, in the beaked whale analysis. The three columns correspond to the three modelled variables: pitch (left), roll (middle), and heading (right). The black lines are the mean estimates, and the grey shaded areas are 95% confidence bands. All estimates were obtained with the mean random effect intercept .

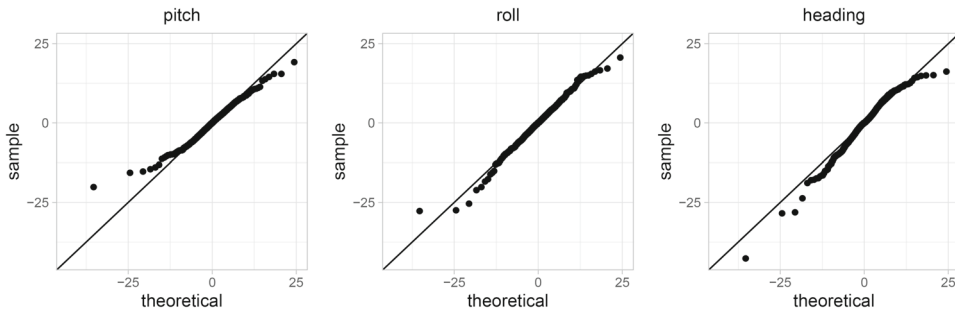


Figure 4. Quantile–quantile plots of the residuals against the Student’s  $t$  distribution with  $\nu = 3$  degrees of freedom, for the beaked whale analysis .

zero. All three diffusion parameters  $s_t$  suggested that there was more variability at the start and end of each dive, i.e. during ascent and descent, than when the whale was at the bottom. This variability can be viewed as a proxy for the level of activity: more diffusion suggests more frequent postural changes. Variability in pitch was low during the bottom phase, which may be associated with gliding motion, whereas it was high during descent and ascent, suggesting continued stroking or “stroke-and-glide” motion. These changes in the pitch diffusion parameter showed how whales alternate between different swimming styles over each dive, which has been linked to energetic efficiency in response to drag forces and buoyancy (Miller et al. 2004; Martín López et al. 2015). Roll displayed highest variability during the ascent phase, and the diffusion parameter for heading was much higher during the final phase of the dive, just before the whales surfaced again, corresponding to more directional changes in the horizontal plane. These postural changes may have several functions, such as locating predators before surfacing (when the whales are most vulnerable), socializing with conspecifics, and orienting to sea currents.

Quantile–quantile plots of the residuals against the appropriate standardized  $t$  distributions are shown in Fig. 4, and suggest appropriate fit. However, we found autocorrelation in the residuals using autocorrelation function plots, pointing to features of the whales’ movement that were not included in the model (Fig. S5 of the supplementary material). Most

notably, there was positive autocorrelation in the heading residuals over about 1–2 min, suggesting that some persistence in heading was not captured by the model. It might be more adequate to model heading with a process that induces correlation between increments, such as a one-dimensional version of the continuous-time correlated random walk used in Sect. 3.1. There was no correlation between residuals of the three processes (pitch, roll, and heading), supporting our decision to model them separately.

Most existing analyses of DTAG data have been based on dive-by-dive summary statistics of activity (e.g. dive duration, maximum depth), and have looked at broad behavioural patterns (e.g. DeRuiter et al. 2013; Quick et al. 2017). This contrasts with our approach, where the behaviour of beaked whales is modelled at a fine time resolution over each dive, and the dives are treated as realizations of an underlying random process. Using varying-coefficient SDEs, we could estimate a more detailed description of the within-dive activity of whales. As another alternative, discrete-time regime-switching models have been proposed to analyse data of this kind [i.e. hidden Markov models (Isojunno and Miller 2015; Leos-Barajas et al. 2017)], and a similar continuous-time approach could be implemented. In that setting, the states of the latent process would represent discrete regimes of activity. For high-resolution data, however, it may often be preferable to model behaviour as changing smoothly in time (rather than switching between discrete states). The method that we use also makes it straightforward to investigate the effects of covariates (if available) and to include random effects to capture differences between individuals.

## 4. DISCUSSION

In this paper, we have focused on a univariate diffusion process ( $Z_t$ ) and suggested using independent diffusion processes for each dimension in the case of multivariate data (e.g. the three postural angles in Sect. 3.3). The proposed method could similarly be applied to the  $N$ -dimensional diffusion process ( $\mathbf{Z}_t$ ) defined by the equation  $d\mathbf{Z}_t = \boldsymbol{\mu}(\mathbf{Z}_t, \boldsymbol{\theta}_t)dt + \boldsymbol{\sigma}(\mathbf{Z}_t, \boldsymbol{\theta}_t)d\mathbf{W}_t$ , where  $\boldsymbol{\mu}(\mathbf{Z}_t, \boldsymbol{\theta}_t) \in \mathbb{R}^N$ ,  $\mathbf{W}_t \in \mathbb{R}^N$ , and  $\boldsymbol{\sigma}(\mathbf{Z}_t, \boldsymbol{\theta}_t) \in \mathbb{R}^{N \times N}$ . In this case, the drift  $\boldsymbol{\mu}$  and diffusion  $\boldsymbol{\sigma}$  might be functions of several time-varying parameters. We can use the Euler–Maruyama discretization to obtain the (approximate) transition density as a multivariate normal distribution and estimate the model parameters as in Sect. 2. A simulation study may be required to investigate identifiability in such models, when a large number of time-varying parameters must be estimated jointly.

The method of inference that we presented in Sect. 2 is approximate, as it relies on the time discretization of the drift and diffusion functions. As suggested in Sect. 2.2.1, data augmentation can be used to improve the accuracy of the method, but there are no general guidelines as to when this might be necessary. Recent studies have evaluated the approximation error of the Euler–Maruyama method for special cases of SDEs (Albertsen 2019; Michélot et al. 2019), but the performance of the approach will be application-specific. The discretization is based on the assumption that the SDE parameters  $\boldsymbol{\theta}_t$  are fixed over each time interval, i.e. the error depends on how fast  $\boldsymbol{\theta}_t$  varies over time. Future work could focus on developing diagnostics to assess the time resolution of discretization, e.g. based on the distribution of first-order differences in estimated values of  $\boldsymbol{\theta}_t$ .

The parameters of a varying-coefficient SDE are specified using basis-penalty smooths, and we therefore assume a smooth relationship between parameters and covariates. There is some flexibility in this formulation, because the smoothness parameter is estimated from the data, but it may not always be appropriate. In particular, if the relationship involves abrupt changes (e.g. discontinuities), then it may not be well captured by a smooth function, and regime-switching models may be preferable. To better understand this issue, it would be interesting to compare the results of the varying-coefficient approach and a regime-switching model on the same data set. However, this may not be straightforward in practice, as there is no generally applicable method to include covariates in continuous-time regime-switching models. We could also consider integrating regime switches into varying-coefficient SDEs, i.e. defining each SDE parameter by several smooths between which the process switches through time. This model can be viewed as a continuous-time hidden Markov model where the state-dependent observation distribution is given by the transition density of the SDE, and the approximate likelihood of this model could be obtained using existing methodology.

We suggested using the package TMB to implement the marginal likelihood and to integrate over random effects using the Laplace approximation. TMB also uses automatic differentiation to evaluate the gradient of the log-likelihood, which improves computational speed (Kristensen et al. 2016). This fast implementation has a downside: to build the gradient function, TMB needs to create the “computational graph” of the likelihood, i.e. its representation in terms of elementary functions (for which the analytical gradient is known). In our experience, the construction of this graph can be memory-intensive for large data sets or complex model formulations, and may not be feasible on standard desktop computers. In those cases, high-performance computing systems with more memory may be required.

The model presented for the time-varying parameters of SDEs relies on the general methodology of generalized additive models (GAMs), which has been greatly extended beyond the basic formulation presented herein. In particular, an interesting direction for future research will be the implementation of hierarchical GAMs (Pedersen et al. 2019) in this framework. Here, the smooth relationship between response and covariates can vary across groups, while retaining some common features (related to shape and degree of smoothness). This extension could be applied to investigate inter-individual differences in ecological analyses, with more nuance than the simple random-intercept model mentioned in Sect. 2.2.3. We could, for example, define the response of several animals to an environmental covariate with functions comprising a population-level mean component and individual-level components measuring the individual deviations from the mean. Other extensions of GAMs, such as adaptive smoothing (Wood 2017, Sect. 5.3.5) or tensor product smooth interactions (Wood 2017, Sect. 5.6), could further increase the applicability of varying-coefficient SDEs.

## ACKNOWLEDGEMENTS

We would like to thank Simon Chamaillé-Jammes, Stacy DeRuiter, Alan Gelfand, Josh Hewitt, Dave Miller, Nicola Quick, and Rob Schick for fruitful discussions about this research. We are grateful to the editors and reviewers whose suggestions improved the presentation of this work. The beaked whale data of Section 3.3 were collected as part of the SOCAL-BRS project, primarily funded by the US Navy’s Chief of Naval Operations Environmental

Readiness Division and subsequently by the US Navy's Living Marine Resources Program. Additional support for environmental sampling and logistics was also provided by the office of Naval Research, Marine Mammal Program. All research activities for that study were authorized and conducted under US National Marine Fisheries Service permit 14534; Channel Islands National Marine Sanctuary permit 2010-004; US Department of Defense Bureau of Medicine and Surgery authorization; a federal consistency determination by the California Coastal Commission; and numerous institutional animal care and use committee authorizations. We are grateful to Jake Wall, George Wittemyer, Valerie LeMay, Iain Douglas-Hamilton, and Brian Klinkenberg for making the elephant data used in Section 3.1 publicly available and to Enrico Pirota, Lisa Schwarz, Daniel Costa, Patrick Robinson, and Leslie New for making the elephant seal data used in Section 3.2 publicly available

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Funding** This work was funded by the US office of Naval Research, Grant N000141812807

#### **Declarations**

**Conflict of interest** The authors declare no conflict of interest.

*[Received November 2020. Revised January 2021. Accepted March 2021.]*

## REFERENCES

- Aït-Sahalia Y, Kimmel R (2007) Maximum likelihood estimation of stochastic volatility models. *J Financ Econ* 83(2):413–452
- Albertsen CM (2019) Generalizing the first-difference correlated random walk for marine animal movement data. *Sci Rep* 9(1):1–14
- Albertsen CM, Whoriskey K, Yurkowski D, Nielsen A, Flemming JM (2015) Fast fitting of non-Gaussian state-space models to animal movement data via Template Model Builder. *Ecology* 96(10):2598–2604
- Allen E (2007) Modeling with Itô stochastic differential equations, vol 22. Springer, Berlin
- Allen LJ, Van den Driessche P (2006) Stochastic epidemic models with a backward bifurcation. *Math Biosci Eng* 3(3):445
- Archambeau C, Cornford D, Opper M, Shawe-Taylor J (2007) Gaussian process approximations of stochastic differential equations. *J Mach Learn Res* 1:1–16
- Biuw M, McConnell B, Bradshaw CJ, Burton H, Fedak M (2003) Blubber and buoyancy: monitoring the body condition of free-ranging seals using simple dive characteristics. *J Exp Biol* 206(19):3405–3423
- Blackwell PG (1997) Random diffusion models for animal movement. *Ecol Model* 100(1–3):87–102
- Blackwell PG, Niu M, Lambert MS, LaPoint SD (2016) Exact Bayesian inference for animal movement in continuous time. *Methods Ecol Evol* 7(2):184–195
- Brillinger DR (2010) Modeling spatial trajectories. In: Gelfand A, Diggle P, Guttorp P, Fuentes M (eds) *Handbook of spatial statistics*. CRC Press, Boca Raton, pp 463–474
- Buderman FE, Hooten MB, Alldredge MW, Hanks EM, Ivan JS (2018) Time-varying predatory behavior is primary predictor of fine-scale movement of wildland-urban cougars. *Mov Ecol* 6(1):22
- Dennis B, Munholland PL, Scott JM (1991) Estimation of growth and extinction parameters for endangered species. *Ecol Monogr* 61(2):115–143



- DeRuiter SL, Southall BL, Calambokidis J, Zimmer WM, Sadykova D, Falcone EA, Friedlaender AS, Joseph JE, Moretti D, Schorr GS et al (2013) First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. *Biol Lett* 9(4):20130223
- Duan JA, Gelfand AE, Sirmans C et al (2009) Modeling space-time data using stochastic differential equations. *Bayesian Anal* 4(4):733–758
- Dunn JE, Gipson PS (1977) Analysis of radio telemetry data in studies of home range. *Biometrics* 33(1):85–101
- Elerian O, Chib S, Shephard N (2001) Likelihood inference for discretely observed nonlinear diffusions. *Econometrica* 69(4):959–993
- Gurarie E, Fleming CH, Fagan WF, Laidre KL, Hernández-Pliego J, Ovaskainen O (2017) Correlated velocity models as a fundamental unit of animal movement: synthesis and applications. *Mov Ecol* 5(1):13
- Hanks EM, Hooten MB, Alldredge MW et al (2015) Continuous-time discrete-space models for animal movement. *Ann Appl Stat* 9(1):145–165
- Hanks EM, Hughes DA (2016) Flexible discrete space models of animal movement. Preprint [arXiv:1606.07986](https://arxiv.org/abs/1606.07986)
- Hastie T, Tibshirani R (1993) Varying-coefficient models. *J Roy Stat Soc Ser B (Methodol)* 55(4):757–779
- Isojunno S, Miller PJ (2015) Sperm whale response to tag boat presence: biologically informed hidden state models quantify lost feeding opportunities. *Ecosphere* 6(1):1–46
- Johnson DS, London JM (2018) Crawl: an R package for fitting continuous-time correlated random walk models to animal movement data. Zenodo, New York
- Johnson MP, Tyack PL (2003) A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE J Oceanic Eng* 28(1):3–12
- Johnson DS, London JM, Lea M-A, Durban JW (2008) Continuous-time correlated random walk model for animal telemetry data. *Ecology* 89(5):1208–1215
- Kristensen K, Nielsen A, Berg C, Skaug H, Bell B (2016) TMB: automatic differentiation and Laplace approximation. *J Stat Softw* 70(5):1–21
- Leos-Barajas V, Photopoulou T, Langrock R, Patterson TA, Watanabe YY, Murgatroyd M, Papastamatiou YP (2017) Analysis of animal accelerometer data using hidden Markov models. *Methods Ecol Evol* 8(2):161–173
- Liechty JC, Roberts GO (2001) Markov chain Monte Carlo methods for switching diffusion models. *Biometrika* 88(2):299–315
- Mao X, Yuan C (2006) Stochastic differential equations with Markovian switching. Imperial College Press, London
- Marra G, Wood SN (2011) Practical variable selection for generalized additive models. *Comput Stat Data Anal* 55(7):2372–2387
- Martín López LM, Miller PJ, Aguilar de Soto N, Johnson M (2015) Gait switches in deep-diving beaked whales: biomechanical strategies for long-duration dives. *J Exp Biol* 218(9):1325–1338
- Michélot T, Blackwell PG (2019) State-switching continuous-time correlated random walks. *Methods Ecol Evol* 10(5):637–649
- Michélot T, Langrock R, Bestley S, Jonsen ID, Photopoulou T, Patterson TA (2017) Estimation and simulation of foraging trips in land-based marine predators. *Ecology* 98(7):1932–1944
- Michélot T, Gloaguen P, Blackwell PG, Étienne M-P (2019) The Langevin diffusion as a continuous-time model of animal movement and habitat selection. *Methods Ecol Evol* 10(11):1894–1907
- Miller PJ, Johnson MP, Tyack PL, Terray EA (2004) Swimming gaits, passive drag and buoyancy of diving sperm whales *Physeter macrocephalus*. *J Exp Biol* 207(11):1953–1967
- Mole MA, Rodrigues-D'Áraujo S, Van Aarde RJ, Mitchell D, Fuller A (2016) Coping with heat: behavioural and physiological responses of savanna elephants in their natural habitat. *Conserv Physiol* 4(1):044
- Pedersen EJ, Miller DL, Simpson GL, Ross N (2019) Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ* 7:e6876
- Pirotta E, Schwarz LK, Costa DP, Robinson PW, New L (2019) Modeling the functional link between movement, feeding activity, and condition in a marine predator. *Behav Ecol* 30(2):434–445
- Preisler HK, Ager AA, Johnson BK, Kie JG (2004) Modeling animal movements using stochastic differential equations. *Environmetrics* 15(7):643–657

- Quick NJ, Isojunno S, Sadykova D, Bowers M, Nowacek DP, Read AJ (2017) Hidden Markov models reveal complexity in the diving behaviour of short-finned pilot whales. *Sci Rep* 7(1):1–12
- Rajabzadeh Y, Rezaie AH, Amindavar H (2016) A robust nonparametric framework for reconstruction of stochastic differential equation models. *Physica A* 450:294–304
- Russell JC, Hanks EM, Haran M, Hughes D et al (2018) A spatially varying stochastic differential equation model for animal movement. *Ann Appl Stat* 12(2):1312–1331
- Schick RS, New LF, Thomas L, Costa DP, Hindell MA, McMahon CR, Robinson PW, Simmons SE, Thums M, Harwood J et al (2013) Estimating resource acquisition and at-sea body condition of a marine predator. *J Anim Ecol* 82(6):1300–1315
- Wahba G (1990) Spline models for observational data. SIAM, Philadelphia
- Wall J, Wittemyer G, LeMay V, Douglas-Hamilton I, Klinkenberg B (2014a) Data from: elliptical time-density model to estimate wildlife utilization distributions. Movebank Data Repos. <https://doi.org/10.5441/001/1.f321pf80/1>
- Wall J, Wittemyer G, LeMay V, Douglas-Hamilton I, Klinkenberg B (2014b) Elliptical time-density model to estimate wildlife utilization distributions. *Methods Ecol Evol* 5(8):780–790
- Wood SN (2003) Thin plate regression splines. *J R Stat Soc Ser B (Stat Methodol)* 65(1):95–114
- Wood SN (2017) Generalized additive models: an introduction with R, 2nd edn. CRC Press, London

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.