Spatiotemporal modelling of marine movement data using Template Model Builder

9 authors, including:

Marie Auger-Méthé
Dalhousie University

32 PUBLICATIONS 281 CITATIONS

SEE PROFILE

Ian D Jonsen
Macquarie University

126 PUBLICATIONS 2,809 CITATIONS

SEE PROFILE

Damian Lidgard
Dalhousie University

24 PUBLICATIONS 201 CITATIONS

SEE PROFILE

William Don Bowen
Bedford Institute of Oceanography

194 PUBLICATIONS 7,974 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Short-term behavioural profiles of white sturgeon released from angling View project

Project Atlantic salmon post-spawning condition: implications of individual's physiological state on migration timing and survival of kelts on their journey back to sea. View project

Spatiotemporal modelling of marine movement data using Template Model Builder (TMB)

Marie Auger-Méthé[1],*, Christoffer M. Albertsen[2], Ian D. Jonsen[3], Andrew E. Derocher[4], Damian C. Lidgard[5], Katharine R. Studholme[5], W. Don Bowen[6], Glenn T. Crossin[5], Joanna Mills Flemming[1]

[1]Department of Mathematics and Statistics, Dalhousie University, Halifax, NS B3H 4R2, Canada
[2]National Institute of Aquatic Resources, Technical University of Denmark, 2920 Charlottenlund, Denmark
[3]Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia
[4]Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada
[5]Department of Biology, Dalhousie University, Halifax, NS B3H 4R2, Canada
[6]Bedford Institute of Oceanography, Dartmouth, NS B2Y 4A2, Canada

ABSTRACT: Tracking of marine animals has increased exponentially in the past decade, and the resulting data could lead to an in-depth understanding of the causes and consequences of movement in the ocean. However, most common marine tracking systems are associated with large measurement errors. Accounting for these errors requires the use of hierarchical models, which are often difficult to fit to data. Using 3 case studies, we demonstrate that Template Model Builder (TMB), a new R package, is an accurate, efficient and flexible framework for modelling movement data. First, to demonstrate that TMB is as accurate but 30 times faster than bsam, a popular R package used to apply state-space models to Argos data, we modelled polar bear *Ursus maritimus* Argos data and compared the locations estimated by the models to GPS locations of these same bears. Second, to demonstrate how TMB's gain in efficiency and frequentist framework facilitate model comparison, we developed models with different error structures and compared them to find the most effective model for light-based geolocations of rhinoceros auklets *Cerorhinca monocerata*. Finally, to maximize efficiency through TMB's use of the Laplace approximation of the marginal likelihood, we modelled behavioural changes with continuous rather than discrete states. This new model directly accounts for the irregular sampling intervals characteristic of Fastloc-GPS data of grey seals *Halichoerus grypus*. Using real and simulated data, we show that TMB is a fast and powerful tool for modelling marine movement data. We discuss how TMB's potential reaches beyond marine movement studies.

KEY WORDS: TMB · State-space model · Telemetry · Argos · GLS · FastLoc GPS

INTRODUCTION

The tracking of marine animals has grown rapidly over the past 30 yr, with a 6-fold increase in the past decade (Hussey et al. 2015). While marine telemetry has been primarily used to describe movement patterns such as transoceanic migrations (e.g. Block et al. 2011), this phenomenal increase in data should unlock a more in-depth understanding of the causes

and consequences of movement in the ocean (Hussey et al. 2015). However, the shortcomings of many marine tracking systems limit our capacity to understand the behavioural mechanisms behind the movement of marine animals. In particular, most common tracking technologies (e.g. GPS) do not work underwater, and marine tracking systems are often associated with large measurement errors (e.g. Costa et al. 2010, Lam et al. 2010). Modelling such movement data

often requires the use of complex models that are computationally demanding to fit to data (e.g. state-space models, see Jonsen et al. 2005, 2013, Albertsen et al. 2015). This computational burden limits our capacity to fit a variety of models to the same dataset and explore complex behavioural mechanisms. Here, we demonstrate how Template Model Builder (TMB, Kristensen et al. 2016), a new R package, can facilitate the fitting of movement models to 3 of the most common marine telemetry data types and discuss how this flexible framework can be used to advance the study of animal movement.

The movement of some marine species can be tracked with the conventional GPS system common in terrestrial studies (e.g. Young et al. 2015, Auger-Méthé et al. 2016b), but this system is usually inappropriate because most marine animals do not surface long enough for the receiver to acquire the satellite data needed to estimate a location (Tomkiewicz et al. 2010, Dujon et al. 2014). As a result, 3 alternative systems are commonly used. Argos telemetry is often used to track air-breathing animals, such as marine mammals, turtles, and penguins, as well as shark species that surface regularly (e.g. Fitzpatrick et al. 2012, Pütz et al. 2014, McClintock et al. 2015). The accuracy of Argos locations partly relies on the number of successive satellite uplinks and most commonly obtained locations for marine species are of low quality (4–36 km, Vincent et al. 2002, Costa et al. 2010, Hoenner et al. 2012; but see new algorithm, McClintock et al. 2015). Fastloc-GPS can be used on animals that surface for only brief periods (10s of milliseconds) and is orders of magnitude more accurate than Argos (<200 m; Bryant 2007, Tomkiewicz et al. 2010, Dujon et al. 2014). However, the higher cost of Fastloc-GPS tags can limit the number of individuals tracked, and the low transmission rate of Fastloc-GPS compared to Argos tags can limit the description of fine-scale movement for high-latitude species that cannot be recaptured (Kuhn et al. 2009, Breed et al. 2011, Lowther et al. 2015). Light-based geolocation is popular in seabird studies because the small size of the loggers reduces some of the effects of tags on animals (Bost et al. 2009, Thiebot & Pinaud 2010), and is often the only telemetry option available to track fish species (e.g. Greene et al. 2009, Lam et al. 2010, Hammerschlag et al. 2011). However, light-based geolocations are associated with large measurement errors (~200 km; Phillips et al. 2004, Pollet et al. 2014). This telemetry system estimates latitude based on day length, and longitude based on the time difference with Greenwich Mean Time (Thiebot & Pinaud 2010). As a result, latitude estimates are particu-

larly inaccurate close to each equinox, when day length is nearly equivalent everywhere on earth (Musyl et al. 2001, Thiebot & Pinaud 2010).

Because many of these tracking systems are associated with large measurement errors, modelling the movement of marine animals often requires the use of state-space models, a class of hierarchical model that accounts for measurement errors. Many movement state-space models have been fitted using Bayesian tools (e.g. JAGS, WinBUGS), which can be computationally demanding (Jonsen et al. 2005, 2013, Albertsen et al. 2015). In particular, the discrete-time model of Jonsen et al. (2005), and its associated R package (bsam, available at http://web.science.mq.edu.au/ ~ijonsen/code.html), has been used extensively for Argos data (e.g. Fitzpatrick et al. 2012, Reisinger et al. 2015). Recently, Albertsen et al. (2015) demonstrated that TMB could estimate the true movement path of an animal much faster than bsam. To do so, they modified a continuous-time analogue of the model of Jonsen et al. (2005) that was developed by Johnson et al. (2008) and compared it to bsam. TMB, which is similar to AD Model Builder (Fournier et al. 2012), is generally used to perform maximum likelihood estimation for hierarchical models (Kristensen et al. 2016). The computational efficiency of TMB can be attributed to its use of reverse-mode automatic differentiation and the Laplace approximation for high-dimension integrals when estimating the parameters (Albertsen et al. 2015, Kristensen et al. 2016). The Laplace approximation accelerates the fitting process by calculating the marginal likelihood of the model with a deterministic equation that approximates integrals (i.e. a second-order Taylor expansion; see Skaug & Fournier 2006 for more details). In addition to computational efficiency, TMB is a rare tool that can be used to fit an extensive variety of non-linear and non-Gaussian hierarchical models in a frequentist framework. For example, it can be used to compute simple metrics for model comparison and it does not require the assignment of priors.

While Albertsen et al. (2015) showed the promise of TMB in modelling Argos data, they did not directly compare the efficiency and accuracy of TMB and bsam, nor did they demonstrate the wider applicability of TMB in movement studies. Here, our first goal was to directly compare the efficiency of TMB to bsam by fitting the same discree-time model of Jonsen et al. (2005). This state-space model was developed for Argos data and can be used to provide more accurate estimates of the true locations of the animal (Jonsen et al. 2005). In our first case study, we applied this state-space model to Argos data col-

lected from polar bear *Ursus maritimus* collars and compared the locations estimated by TMB and bsam to GPS locations collected from the same individuals. We also used simulations to explore the effect of data quality on the accuracy of the 2 packages.

Model comparison is a powerful tool to identify the most adequate model for a dataset (Burnham & Anderson 2002), but it is rarely used to choose among state-space models. In fact, many movement studies rely on a single state-space model (e.g. Fitzpatrick et al. 2012, Pollet et al. 2014, Reisinger et al. 2015). Our second goal was to demonstrate how TMB's increase in computational efficiency, as well as its frequentist framework, allows researchers to develop and compare a set of models. In our second case study, we used light-based geolocation data from rhinoceros auklets *Cerorhinca monocerata* and a simulation study to demonstrate how TMB can be used to compute information theoretic metrics, such as Akaike's information criterion (AIC, Burnham & Anderson 2002), to find the best way to handle the error structure.

Our third goal was to demonstrate the use of TMB for applications beyond assessing measurement errors and addressing questions of greater ecological interest. We focussed on using TMB to quantify temporal changes in movement behaviour. For this, we developed a model with continuous behavioural changes that allowed us to use the Laplace approximation and account for irregular sampling intervals. In our third case study, we tested this model using simulations and archival Fastloc-GPS data from grey seals *Halichoerus grypus*.

MATERIALS AND METHODS

Case study 1 – Argos tracks from polar bears

We used Argos location data from 4 adult female polar bears collared in the Beaufort Sea, Northwest Territories, Canada, in April of 2009 to 2011 (for collaring details, see Auger-Méthé et al. 2016b). These bears had Gen 4 Telonics collars, which collected both Argos and conventional GPS locations. As standard GPS data are extremely accurate (≤30 m, Tomkiewicz et al. 2010) compared to Argos data (mean accuracy for lowest and highest Argos quality classes: 0.5–36 km, Costa et al. 2010), we considered the GPS locations as the true locations of the animals and compared these to the estimated locations obtained by fitting our state-space models to the Argos data.

The collars collected multiple Argos locations daily, but 99% of them were collected between 17:00 and 21:00 h. They collected GPS locations every 4 h, but we used only the daily 17:00 h location, which was closest in time to the bulk of the Argos locations. We chose these 4 bears for 2 reasons. First, they had long time series (>1500 Argos locations), which are more computationally demanding to fit and could thus be used to adequately assess the difference in efficiency between TMB and bsam. Second, they had no data gaps longer than 2 wk, which ensured that the comparison between the location estimates and the GPS locations was based on days where we had recent Argos information. The Argos categories provided by these collars were 3, 2, 1, 0, A and B, and as in other studies (Costa et al. 2010, Albertsen et al. 2015, Lowther et al. 2015), most locations were in the worst quality categories (Table S1.1 in Supplement 1 at www.int-res.com/articles/suppl/m565p237_supp/). Note that we did not pre-process the data with a speed filter and kept all Argos locations in the analysis.

The goal was to directly compare the efficiency and accuracy of TMB to bsam when fitting the same model. To do so, we used the 1-behaviour first-difference correlated random walk (DCRW) available in bsam (v 0.43.1), which is based on the model described by Jonsen et al. (2005). Unless otherwise indicated, we used the default settings. Our version of the DCRW in TMB was identical to the version in bsam, with the exception that it did not use priors as, unlike bsam, we used a frequentist approach (see Supplement 1 for a description of the priors used in bsam). The model was written as:

Initialisation eqn $\quad \mathbf{x}_1 = \mathbf{x}_0 + \boldsymbol{\epsilon}_1$ (1)

Process eqn $\quad \mathbf{x}_t = \mathbf{x}_{t-1} + \gamma \mathbf{T}(\theta)(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}) + \boldsymbol{\epsilon}_t$ (2)

Measurement eqn $\quad \mathbf{y}_{t,i} = (1 - j_i)\mathbf{x}_{t-1} + j_i \mathbf{x}_t + \boldsymbol{\eta}_t$ (3)

where

$$\boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Sigma}_\epsilon) \qquad \boldsymbol{\Sigma}_\epsilon = \begin{bmatrix} \sigma^2_{\epsilon,lon} & \rho\sigma_{\epsilon,lon}\sigma_{\epsilon,lat} \\ \rho\sigma_{\epsilon,lon}\sigma_{\epsilon,lat} & \sigma^2_{\epsilon,lat} \end{bmatrix} \quad (4)$$

$$\mathbf{T}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad (5)$$

$$\boldsymbol{\eta}_t = \begin{bmatrix} \eta_{t,lon,q} \\ \eta_{t,lat,q} \end{bmatrix} \qquad \eta_{t,c,q} \sim tdist\left(\frac{\phi_{c,q}}{\sqrt{\psi}}, df_{c,q}\right) \quad (6)$$

and

$$\sigma_{\epsilon,c} > 0, \quad -1 < \rho < 1, \quad 0 < \gamma < 1, \quad -\pi < 0 < \pi, \quad \psi > 0 \quad (7)$$

In these equations, $\mathbf{y}_{t,i}$ was a 2-dimensional vector containing an observed location: $\mathbf{y}_{t,i} = \begin{bmatrix} y_{t,i,lon} \\ y_{t,i,lat} \end{bmatrix}$.

The 2-dimensional vector, $\mathbf{x}_t$, represented the true unobserved locations at time t: $\mathbf{x}_t = \begin{bmatrix} x_{t,lon} \\ x_{t,lat} \end{bmatrix}$. j_i represented the proportion of the regular time interval between $t-1$ and t at which the observation $\mathbf{y}_{t,i}$ was made. The model used a t-distribution for the measurement errors and most measurement parameters ($\phi_{c,q}$ and $df_{c,q}$) associated with the coordinates (c = latitude or longitude) and Argos quality classes (q = 3, 2, 1, 0, A or B) were assumed to be known. Jonsen et al. (2005) estimated these parameters based on the observed error distribution from Vincent et al. (2002). We had 6 remaining parameters to estimate $\Theta = (\theta, \gamma, \rho, \sigma_{\epsilon,lon}, \sigma_{\epsilon,lat}, \psi)$, which included a correction factor for the fixed measurement parameters, ψ. The other parameters to estimate were associated with the process equation. The mean relative turning angle (θ) represented the level of directional persistence, with $\theta = 0$ representing an animal that generally continued in the same direction. γ represented the autocorrelation in terms of both movement speed and direction, with $\gamma = 0$ representing an animal using a simple random walk. The 3 parameters associated with the process stochasticity (ρ, $\sigma_{\epsilon,lon}$, $\sigma_{\epsilon,lat}$) also controlled the level with which the displacement at time t was close to the previous displacement in terms of speed and directionality. The process variance along each coordinate axis was represented by $\sigma^2_{\epsilon,lon}$ and $\sigma^2_{\epsilon,lat}$, while ρ represented the correlation coefficient between the coordinates.

To demonstrate the difference in efficiency between TMB and bsam, we estimated the time it took to estimate the parameters of the models. For the TMB timings, we included the time it took to compile the C++ file with compile(), load the compile function with dyn.load(dynlib()), construct the objective function with MakeADFun() and minimize the negative log likelihood of the objective function using nlminb(). Minimizing the negative log likelihood allowed us to find the maximum likelihood estimate (MLE) for the model parameters. For the bsam timings, we included the time it took to fit the model using the function fitSSM(). We ran 2 chains for 55000 steps, for which the first 35000 allowed for adaptive tuning. We only kept every 50th sample of the last 20000 steps. We chose a number of steps that would result in convergence but would be short enough to limit excessive computing time (see Supplement 1). Our chain length is within the range of those from similar movement analyses (e.g. 40000–90000; Kennedy et al. 2014, Silva et al. 2014, Lowther et al. 2015).

Similar to Patterson et al. (2010), we evaluated the performance of the method in terms of estimating the value of the states, $\mathbf{x}_t$, by calculating the root mean square error (RMSE) for each coordinate, c (i.e. latitude and longitude):

$$\text{RMSE}_c = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{x}_{t,c} - x_{t,c})^2} \qquad (8)$$

where $\hat{x}_{t,c}$ is the estimated true coordinate value of the bear at time t, and $x_{t,c}$ is the associated GPS coordinate value for that bear. The RMSE_c returned a decimal degree value. We also calculated the mean great circle distance between the estimated true locations and GPS locations.

To investigate whether the quality of the data affected the relative accuracy of TMB and bsam, we performed a simulation study. For the bears, the number of Argos locations per day varied from 3.7 to 5.2 (Table S1.1). To explore the effect of data frequency on the estimation process, we created 3 simulation scenarios that varied in the number of daily Argos locations: 0.5, 1 and 5 locations per day. For each scenario, we simulated 365 d movement tracks ($n = 100$) using the model presented in Eqs. (1)–(7) and the parameter values estimated with TMB for polar bear PB1 (Table S1.4). To select the Argos class category of an observation, we sampled based on the overall occurrence of class categories in the polar bear dataset (Table S1.1). We applied the model to the simulated data using both TMB and bsam as described above, and used the same approach to calculate efficiency and accuracy.

Case study 2 – light geolocation tracks from rhinoceros auklets

We used the daily light-based geolocations obtained from loggers deployed on 4 female rhinoceros auklets breeding on Lucy Island, British Columbia, Canada. The loggers (LAT2900, Lotek Wireless) were attached to the tarsus and weighed ~0.4% of the auklets' body mass, well below the recommended 3% (Phillips et al. 2003). Birds were captured from their nest chamber during the chick-provisioning period (July 2013), and the loggers were recovered at the time of arrival and egg-laying (May 2014). The processing system onboard the loggers used the template-fit algorithm to return location estimates. This algorithm matches the shape of each observed light curve with a theoretical curve parametrized with latitude, longitude and cloudiness factors (Ekstrom 2004), and provides an error estimate for each location coordinate (Ekstrom 2007).

Because light-based geolocation is associated with large measurement errors and frequent aberrant lo-

cations, a variety of filters are used to remove unrealistic locations (e.g. Jessopp et al. 2013, Thiers et al. 2014). To model measurement error rather than discard large amounts of data, we developed 2 versions of a state-space model similar to that of Winship et al. (2012). Our model versions were designed to handle measurement errors typical of light-based geolocation, through techniques such as modelling the error of each coordinate (latitude, longitude) separately. However, similarly to Winship et al. (2012), we needed to remove some of the most aberrant geolocations (see Supplement 2 at www.int-res.com/articles/suppl/m565p237_suppl/ for details). Our filter removed on average 3% of the longitudes and 16% of the latitudes, which is lower than rates reported by others (23−35% of both coordinates removed; Takahashi et al. 2008, Gaston et al. 2011, Thiers et al. 2014).

Similar to our polar bear analysis, we used a 1-behaviour DCRW based on the model of Jonsen et al. (2005). Because the auklets were captured to place and retrieve the loggers, we know the exact locations of the individuals at the beginning and end of the time series $(\mathbf{x}_0 = \mathbf{x}_n = \left[\begin{smallmatrix} 130.616667°W \\ 54.294167°N \end{smallmatrix}\right])$. We used this information when initialising the model. Because we had locations at regular time intervals (daily), and we assumed that the animal continued in the same direction (equivalent of fixing $\theta = 0$) and that there was no correlation in process stochasticity (equivalent of fixing $\rho = 0$; see Breed et al. 2012), we simplified the model of Eqs. (1)−(7) to:

Initialisation eqn $\quad \mathbf{x}_1 = \mathbf{x}_0 + \boldsymbol{\epsilon}_1$ (9)

Process eqn $\quad \mathbf{x}_t = \mathbf{x}_{t-1} + \gamma(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}) + \boldsymbol{\epsilon}_t$ (10)

Measurement eqn $\quad \mathbf{y}_t = \mathbf{x}_t + \boldsymbol{\eta}_t$ (11)

where

$$\boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Sigma}_\epsilon) \quad \boldsymbol{\Sigma}_\epsilon = \begin{bmatrix} \sigma^2_{\epsilon,lon} & 0 \\ 0 & \sigma^2_{\epsilon,lat} \end{bmatrix} \quad (12)$$

$$\boldsymbol{\eta}_t = \begin{bmatrix} \eta_{t,lon} \\ \eta_{t,lat} \end{bmatrix} \quad (13)$$

and

$$\sigma_{\epsilon,lon} > 0, \quad \sigma_{\epsilon,lat} > 0, \quad 0 < \gamma < 1 \quad (14)$$

As a result, the process equation had only 3 parameters to estimate: γ, $\sigma_{\epsilon,lon}$ and $\sigma_{\epsilon,lat}$.

To explore how to best model the measurement errors of light-based geolocation, we made 2 versions of this model. Each version differed in its description of $\eta_{t,c}$, where c represents the coordinate (i.e. latitude and longitude). While both versions used the error information provided by the tag for the coordinates of

Table 1. Error structures of the 2 versions of the model for the rhinoceros auklet *Cerorhinca monocerata* data. Both versions used the standard deviation (SD) provided by the tag ($\phi_{t,c}$) at each time step (t) and for each coordinate (c), and estimated correction factors (α_c). Note that α is the equivalent to $\frac{1}{\sqrt{\psi}}$ in Eq. (6). One version used the normal distribution (M_N), while the other used a t-distribution (M_T), which requires estimating the degrees of freedom (df_c), which are set to be >3 (see Albertsen et al. 2015). The t-distribution becomes a normal distribution when $df_c \rightarrow \infty$, in which case $\alpha_c \phi_{t,c}$ would be equivalent across distributions. k represents the number of parameters estimated for the complete model

Model	Measurement error distribution	k
M_N	$\eta_{t,c} \sim N(0, (\alpha_c \phi_{t,c})^2)$	5
M_T	$\eta_{t,c} \sim tdist(\alpha_c \phi_{t,c}, df_c)$	7

each location, one version (M_N) used the normal distribution (as in Winship et al. 2012) and the other (M_T) used the t-distribution, because its heavy tails allow for large measurement errors. Table 1 shows a formal description of the measurement error component of each version.

To compare the 2 model versions, we used a common information-theoretic metric: AIC. Because TMB can be used to estimate the negative log likelihood value at the MLE, we could calculate the AIC value (AIC $= -2\log(\mathcal{L}) + 2k$, where $\mathcal{L}$ is the likelihood value at the MLE and k is the number of parameters estimated, see Burnham & Anderson 2002). We chose AIC, but other metrics such as the Bayesian information criterion (BIC) could also be calculated (Burnham & Anderson 2002).

To verify that we could adequately select the best version with AIC, we simulated each model version 100 times. We created movement paths with 300 time steps, which is close to the number of locations gathered for the auklets (297−299). For the model parameters, we used the values estimated by the best version for auklet 1 (Table S2.1 in Supplement 2). To simulate the error values returned by the tags, we sampled directly from the tag error values from auklet 1. To reproduce the filtering, we removed 3% of the longitudes and 16% of the latitudes. We fitted the 2 versions to each simulation with the goal to verify that the lowest AIC value would be associated with the underlying simulated model version.

Case study 3 – Fastloc-GPS tracks from grey seals

We used the Fastloc-GPS movement data of 4 pregnant female grey seals captured on Sable Island,

Nova Scotia, Canada, in June 2013. The seals were equipped with MK10-AF tags (Wildlife Computers) that were programmed to archive location information every 15 min. The average duration of a grey seal dive is approximately 5 min (Beck et al. 2003). Thus, we expected to obtain 1 location every 2 to 3 surfacings. The data were downloaded when the seals were recaptured in December 2013 or January 2014. To preserve battery life, the unit switched to a less frequent transmission schedule of every 8 d when the unit was dry for >12 h. As preprocessing, we removed any location with tag residual values >35 or based on less than 5 satellites (as suggested by Dujon et al. 2014). This quality filtering removed a total of 6 to 23% of the locations of each time series. All 4 time series had >7000 locations and the time intervals between consecutive locations were often close to 15 min. However, the 4 time series had instances of locations separated by more than 40 h. The number of observed locations was 43 to 64% of that expected by the 15 min programming schedule. See Supplement 3 at www.int-res.com/articles/suppl/m565p237_supp/ for more details on the data of each seal.

Our third goal was to use TMB to quantify temporal changes in movement behaviour. A popular method to quantify animal behaviour with movement data is using a switching DCRW (Jonsen et al. 2005, Breed et al. 2009). This model is based on the DCRW model described in the previous case studies, but allows for 2 sets of mean turning angle and behavioural persistence parameters, with each set associated with a distinct behaviour. However, the Laplace approximation, which is one of the main tools TMB uses to accelerate the estimation process (Albertsen et al. 2015, Kristensen et al. 2016), is only appropriate for differentiable likelihood functions and cannot be used on discrete states (Kass et al. 1990, Fournier et al. 2012, Bolker et al. 2013). This complicates the implementation of the switching DCRW in TMB. To be able to use the Laplace approximation, we constructed a model similar to that of Breed et al. (2012), which was also based on the DCRW but considers behavioural states as continuously changing in intensity rather than as a limited set of discrete states (see Whoriskey et al. 2017 for another solution). As per Lidgard et al. (2014), we assumed that Fastloc-GPS locations were sufficiently accurate to ignore measurement error. The resulting model was written as:

Behaviour eqn $\quad \gamma_i = \gamma_{i-1} + \nu_i$ (15)

Movement eqn $\quad \mathbf{x}_i = \mathbf{x}_{i-1} + \gamma_i \frac{\Delta t_i}{\Delta t_{i-1}} (\mathbf{x}_{i-1} - \mathbf{x}_{i-2}) + \boldsymbol{\epsilon}_i$ (16)

where

$$\nu_i \sim N(0, \Delta t_i^2 \sigma_\nu^2) \quad (17)$$

$$\boldsymbol{\epsilon}_i \sim N(0, \boldsymbol{\Sigma}_i) \qquad \boldsymbol{\Sigma}_i = \begin{bmatrix} \Delta t_i^2 \sigma_{lon}^2 & 0 \\ 0 & \Delta t_i^2 \sigma_{lat}^2 \end{bmatrix} \quad (18)$$

and

$$\sigma_{lon} > 0, \quad \sigma_{lat} > 0 \quad (19)$$

Note that the movement equation (Eq. 16) was similar to the process equation in the 2 previous case studies (Eqs. 2 & 10). However, there are 3 important differences. First, the true locations, $\mathbf{x}_i$, were directly observed. Second, while these observations were collected at semi-regular intervals (i.e. the tags were programmed to take a location every 15 min), sometimes consecutive locations were separated by multiple hours or, in extreme cases, by a few days. As such, the locations were indexed by i, $\mathbf{x}_i$, and we used the time difference between subsequent locations, $\Delta t_i = t_i - t_{i-1}$, to account for the effect of the time discrepancies. In particular, we expected the animal to move farther during longer time periods, thus we multiplied the distance moved in each direction at the previous step ($\mathbf{x}_{i-1} - \mathbf{x}_{i-2}$) by the time ratio $\frac{\Delta t_i}{\Delta t_{i-1}}$. Because we expected locations that are taken long time periods apart to be less informed by the previous movement, we increased the variance of the movement equation with Δt_i. Third, the behavioural persistence parameter, γ_i, was time-varying and represented the continuous behavioural state. As in Eq. (10), there was no turning angle parameter. The tendency to continue in the same direction and move at the same speed as the previous step was controlled simultaneously with γ_i. The continuous change in this behavioural state was modelled by a simple random walk (Eq. 15), which assumed temporal correlation. Note that this behaviour equation also accounted for the time discrepancies by increasing the variance with Δt_i. The value of γ_i can be used as a continuous representation of the changes in behaviour or can be further categorized using threshold values for γ_i. Because animal movement paths are often discretized into 2 categories (e.g. Jonsen et al. 2005, Breed et al. 2009), we used a threshold value of $\gamma_i < 0.7$ to identify areas of tortuous movement. The choice of the threshold value was arbitrary, but we chose 0.7 because both with the simulations and seal data, $\gamma_i \geq 0.7$ most adequately captured directed movement and this value fell within the values associated with directed movement (52–98, see Jonsen et al. 2005, Breed et al. 2009).

To verify that we could adequately estimate the behavioural states of this model, we simulated 100

movement paths similar to the seal movement paths. As in Breed et al. (2012), we simulated changes in γ_i using a sine-wave, rather than as in Eq. (15). We used Eq. (16) to simulate the true locations of the seal for 10 000 time steps. As an estimate of the movement variability, we used the observed standard deviation in the movement of the seals: $\boldsymbol{\Sigma}_\epsilon = \begin{bmatrix} 0.730^2\,\text{km}^2 & 0 \\ 0 & 0.784^2\,\text{km}^2 \end{bmatrix}$. To ensure that the small error associated with Fastloc-GPS would not affect our analysis, we added measurement error in the simulation, using: $\mathbf{y}_i = \mathbf{x}_i + \boldsymbol{\eta}_i$, $\boldsymbol{\eta}_i \sim N(0, \boldsymbol{\Sigma}_\eta)$. We used the standard deviation calculated from the data set of differences between Fastloc-GPS locations and known locations found by Dujon et al. (2014): $\boldsymbol{\Sigma}_\eta = \begin{bmatrix} 0.038^2\,\text{km}^2 & 0 \\ 0 & 0.055^2\,\text{km}^2 \end{bmatrix}$. Finally, to account for the effect of preprocessing and missing locations associated with the behaviour of the seal, we removed 45 % of the simulated observations, $\mathbf{y}_i$, which resulted in a time-series of length 6500. We then applied the model described in Eqs. (15)–(19), where we assumed that the simulated observed locations, $\mathbf{y}_i$, were the directly observed true locations in the model: $\mathbf{x}_i$.

To assess how well we estimated the continuous behavioural states, γ_i, we identified whether the simulated values fell within the estimated 95 % confidence interval (CI). For this, we used the estimated standard errors provided by TMB and multiplied them by the 2.5 and 97.5th percentiles of the normal distribution (as in Auger-Méthé et al. 2016a).

All code is available in Supplement 4 at www.int-res.com/articles/suppl/m565p237_supp/.

RESULTS

Case study 1 – Argos tracks from polar bears

Our simulations indicated that TMB was at least 1 order of magnitude faster than bsam (Fig. 1A). The mean computational time for TMB stayed below 2 min, even for the highest data frequency (5 locations per day, see Fig. 1 and Table S1.6 in Supplement 1). In contrast, the mean computational time of bsam increased from 15 to 35 min with increasing data frequency. While TMB was faster, the accuracy of the 2 methods was similar (Fig. 1B). TMB was slightly less accurate when estimating the states and parameters ρ and ψ, but was generally more accurate when estimating the parameters θ, γ, $\sigma_{\epsilon,\text{lon}}$ and $\sigma_{\epsilon,\text{lat}}$ (Tables S1.6 & S1.7). The accuracy of both methods increased with increasing data frequency (Fig. 1B and Tables S1.6 & S1.7).
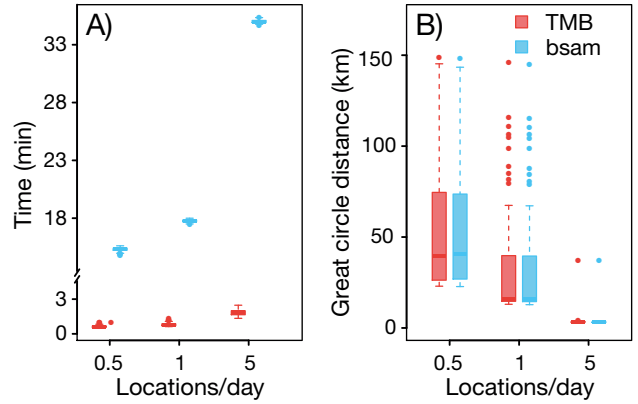

Fig. 1. Comparison of the efficiency and accuracy of Template Model Builder (TMB) and bsam for different data frequency scenarios. (A) Computational time of TMB (in red) and bsam (in blue). Note the break in the *y*-axis. (B) Accuracy in terms of great circle distance between the estimated and simulated locations of the animal

For our polar bear data, TMB was more than 30 times faster than bsam (Table S1.3). While the methods differed in terms of computational time, most parameter estimates were similar (Table S1.4). In addition, the estimated true locations were almost identical (Fig. 2, Fig. S1.2), but TMB was slightly less accurate than bsam (60–470 m, Table S1.3).
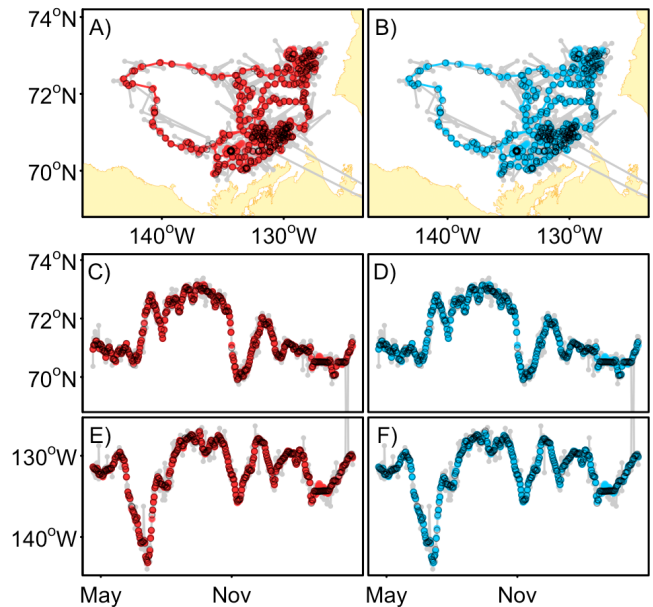

Fig. 2. Comparison of the estimated paths by Template Model Builder (TMB) and bsam for polar bear *Ursus maritimus* PB1. The grey points and lines represent the Argos movement path. The colour points and lines represent the estimated locations (red: TMB, blue: bsam). The black circles represent the GPS locations. (A, B) Movement paths, (C–F) time series of each coordinate. See Fig. S1.2 in Supplement 1 for the other bears

TMB returned a false convergence message for 1 of the 4 bears (Table S1.3). The fact that the accuracy of TMB is close to that of bsam for this bear, and that the accuracy results from this bear fall in the range of those from the 3 other individuals (Table S1.3), indicates that the false convergence did not affect the results significantly. However, the false convergence message returned (Table S1.3) often indicates that there is a discontinuity in the function, or its gradient, near the last set of parameter values evaluated, suggesting that there might be a problem with the model.

Case study 2 – light geolocation tracks from rhinoceros auklets

Our simulations indicated that AIC selected the correct model version in most cases (≥97/100 for both M_N and M_T, see Table S2.2). However, when the version with the normal distribution was simulated, a t-distribution version was sometimes selected (3/100). In these cases, the estimates for the degrees of freedom of the t-distribution (df_c) were large (>1.1×10^8). The t-distribution becomes a normal distribution when $df_c \rightarrow \infty$ and thus M_T becomes almost equivalent to M_N when df_c parameters are large.

For the auklet data, the best version (M_T) used the t-distribution (Table 2). This indicates that the tag error information was not always reliable, something that is also evident from the fact that M_N produced such a poor movement track for auklet 4 (Fig. S2.2). The movement paths produced by the 2 versions differed (Fig. 3 and Fig. S2.2), indicating that choices made when modelling measurement errors will affect ecological inferences. The movement paths produced by the best version mostly avoided land, even though we did not incorporate land as a covariate in the model (Fig. 3 and Fig. S2.2). The parameter estimates associated with M_T indicated that the heavy tails of the t-distribution, which allowed for large measurement errors, were useful (df_c close to 3; Table S2.1).

Table 2. Comparison of the 2 versions of the model (M_N, M_T) used to represent the light-based geolocation data of 4 rhinoceros auklets *Cerorhinca monocerata*. The versions differ in their measurement error structure (see Table 1). The AIC results are presented for each individual

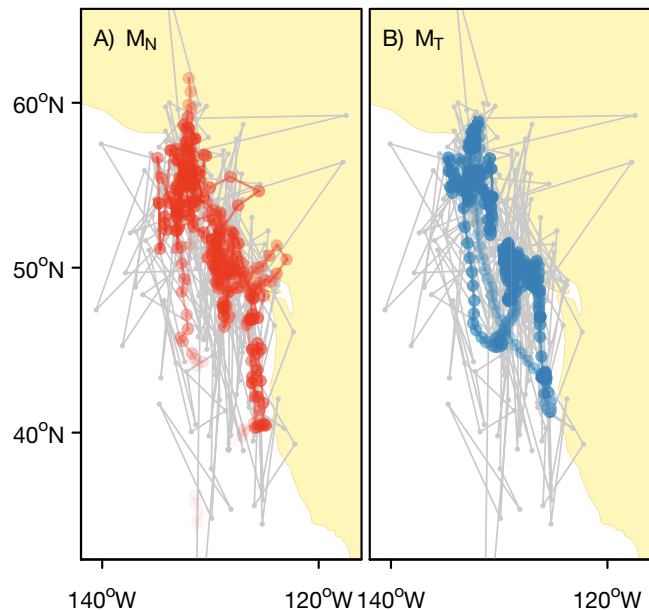| Model	∆ AIC			
	Auklet 1	Auklet 2	Auklet 3	Auklet 4
M_N	114.7	129.0	129.0	209.0
M_T	0.0	0.0	0.0	0.0

Fig. 3. Comparison of the fit of the 2 model versions to one rhinoceros auklet *Cerorhinca monocerata* movement track. The grey tracks represent the observed light-based geolocation data, and the colour tracks show the estimated movement paths. (A) Model with normal measurement errors (M_N); (B) model with t-distributed measurement errors (M_T). The intensity of the colour represents the confidence in the location estimate, and any location with $(SE(x_{\text{lat},t}) + SE(x_{\text{lon},t})) \geq 5$ is not displayed. For the other rhinoceros auklets, see Fig. S2.2 in Supplement 2

Case study 3 – FastLoc-GPS tracks from grey seals

Our simulations demonstrated that the model was capable of recovering the behavioural states (Fig. 4A) and that the model could be used to adequately identify the areas associated with different behaviours (Fig. 4B,C). On average, 93.2% (range: 79.9–100.0%) of the states of a simulation fell within the estimated CI. The grey seal movement paths showed that seals varied their movement behaviour during the course of the time series and that, when we used a threshold of $\gamma_t < 0.7$, we could identify the less correlated movement behaviour often associated with area-restricted search (Fig. 5).

DISCUSSION

Our results demonstrated that TMB is a useful tool to model marine movement data. We showed that TMB could increase the efficiency of fitting state-space models to Argos data without substantially decreasing the accuracy of the estimates. We demon-
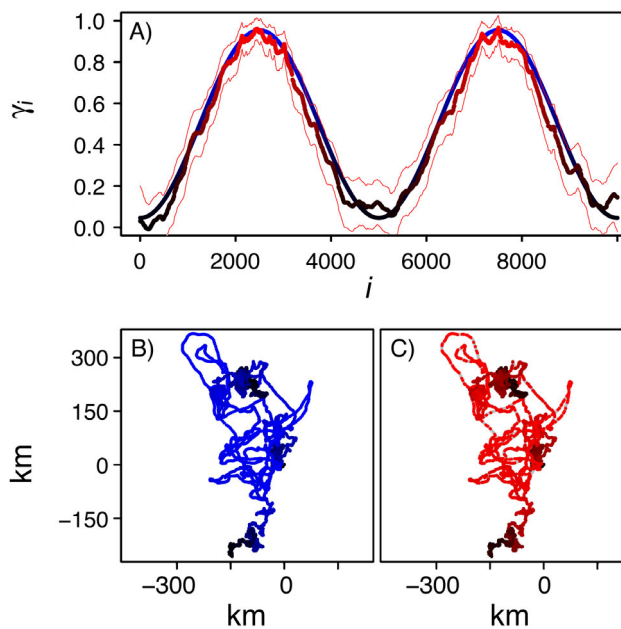
Fig. 4. Simulated and estimated changes in behaviour of modelled grey seal *Halichoerus grypus* movement. (A) Temporal changes in simulated γ_i values in blue, as well as the estimated γ_i values in red and associated 95 % CI. The colour intensity represents the value of the behavioural persistence parameter, γ_i. Lighter colours represent high γ_i values and thus movement with directional and speed persistence, often associated with travelling behaviour. Darker colours represent low γ_i values and thus movement with less directional and speed persistence, sometimes associated with foraging behaviour. (B) Simulated movement path, which can be used to compare to (C) the estimated movement path in red (light grey shows the locations that were removed before fitting the model)

strated that TMB's efficiency and likelihood-based framework promotes model comparison, an approach we used to identify how to best model the measurement error associated with light-based geolocators. To maximize TMB's efficiency, we used the Laplace approximation to model behavioural changes. The Laplace approximation restricts the use of discrete states (Fournier et al. 2012, Bolker et al. 2013, Kristensen et al. 2016). However, we showed that changes in behaviour can be modelled using a continuous behavioural state that can be discretized afterward (Fig. 5). This approach also allowed us to directly incorporate the effects of irregularly sampled locations on behavioural transitions and observed movement tracks. Overall, our results showed that TMB is a fast, accurate, and flexible framework to model marine data and that our new models can extract useful information from light-based geolocation and FastLoc-GPS data.

Efficiency and accuracy comparison

Both our simulation and polar bear data analyses showed that TMB was at least an order of magnitude faster than bsam. Because we timed the compile and load functions of TMB when applying the model to each individual movement path, our analysis slightly underestimated TMB's computational efficiency. These actions required 4 to 14 % of the time needed to fit the model to a polar bear movement path, but could have been executed a single time before repeatedly fitting the model to data. Albertsen et al. (2015) also found that TMB was much faster than bsam. Their best TMB model was 58 to 152 times faster than bsam, a gain that can be mostly attributed to TMB but may also be influenced by the continuous formulation of their TMB model. Regardless, as it can take days to fit complex models to telemetry data with the common Bayesian tools used in movement ecology (McClintock et al. 2015), such increases in efficiency will change how movement analyses can be approached.

The parameter and state estimates of TMB and bsam were similar, and both packages provided accurate estimates of the true locations of the animals. The accuracy of both packages increased sig-
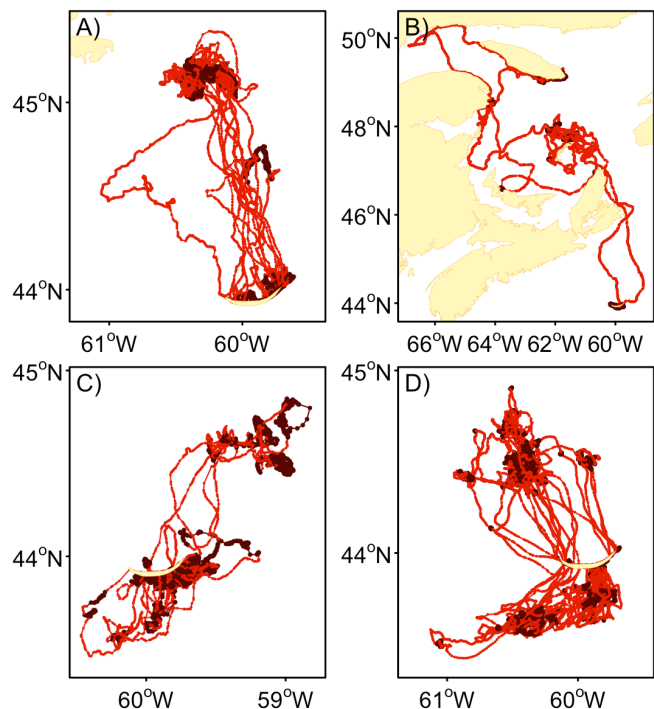

Fig. 5. Estimated changes in behaviour during the movement paths of 4 grey seals *Halichoerus grypus*. (A–D) Seals 1, 2, 3 and 4, respectively. Large dark points represent $\gamma_t \ll 0.7$, and thus area-restricted movement. Smaller light red points represent $\gamma_t \geq 0.7$, and thus directed movement

nificantly with increasing data frequency. While TMB was slightly better at estimating most parameter values, bsam was marginally better at estimating the true locations and 2 of the parameters. This discrepancy could be due to bsam's use of priors (see Supplement 1) and TMB's use of the Laplace approximation. In addition, the false convergence message for one of the bears and the jagged likelihood profiles (Supplement 1) indicated that TMB had difficulties estimating the parameters. We recently highlighted that ecological state-space models may have estimation problems that are easier to diagnose with non-Bayesian tools such as TMB (Auger-Méthé et al. 2016a). For the polar bear data, the estimation should have been helped by the temporally clustered Argos locations, which almost served as replicates for the same true location. However, 3 aspects of the DCRW formulation may have triggered estimation problems that potentially affect both packages, but may be more easily diagnosed with TMB.

First, the degrees of freedom from the *t*-distribution of some Argos categories were fixed to 2, which is lower than the minimum value used by others to ensure that the *t*-distribution has adequate statistical properties (e.g. defined skewness, see Albertsen et al. 2015). Second, the measurement error associated with each geographic coordinate was modelled independently. However, the sizes of Argos errors are often correlated in the latitude and longitude directions (Costa et al. 2010). Third, the fixed parameters for the measurement error may not be adequate for this data set, and the single correction factor provided by the model (ψ in Eq. 6) may not be sufficient to account for data-specific differences. Estimated values for these parameters differ between studies, and differences are inconsistent across Argos quality categories (Jonsen et al. 2005, Hoenner et al. 2012, Lowther et al. 2015). In addition, the bears studied were more than 20° north of the seals used to estimate these parameters (Vincent et al. 2002, Jonsen et al. 2005), and the modelled error structure does not account for the decreasing distance covered by a degree of longitude with increasing latitude (Lowther et al. 2015). The inability of a single correction factor to account for data-driven discrepancies across quality categories and geographic coordinates may also help explain why our estimates of the true locations were slightly less accurate than those of others that compared bsam estimates to GPS locations (Hoenner et al. 2012, Lowther et al. 2015). Indeed, Lowther et al. (2015) replaced the fixed measurement parameters from bsam with values directly estimated from the animals they studied and applied longitudinal

corrections, which may have contributed to their better estimates of the true locations of the animals.

To directly compare TMB to bsam, we limited the scope of our analysis and simply reproduced the model from Jonsen et al. (2005). However, TMB's flexibility allows researchers to model the measurement error more adequately. For example, we could easily use a bivariate *t*-distribution that represents the directional correlation in measurement error (see McClintock et al. 2015), or estimate the measurement parameters rather than use fixed values (see Albertsen et al. 2015).

Facilitating model comparison

One of the main advantages of TMB's increase in computational efficiency is that it allows researchers to develop and fit multiple models to the same data set. Through a simulation study, we showed that we could use AIC to select the model that best represents the data. Using this approach, we showed that the measurement error of light-based geolocation was best modelled with *t*-distributions. Many raw observations were on land, including many locations far inland. However, rhinoceros auklets spend most of their winter migration at sea (Gaston & Dechesne 1996). The best model estimated few locations on land, which further suggests it adequately represented the movement of these animals. Because TMB is also flexible, many alternate models could be explored. As the accuracy of latitude estimates of light-based geolocations is closely related to daylength, some have modelled measurement error as a function of time since the equinox (e.g. Nielsen et al. 2006). Land masks have been incorporated in state-space models to further improve location estimates of seabirds (e.g. Thiebot & Pinaud 2010). Many light-based geolocators also collect sea surface temperature data (e.g. Takahashi et al. 2015), and many state-space models incorporate such information (e.g. Nielsen et al. 2006, Lam et al. 2010). Implementing such alternative models in TMB could streamline model assessment and comparison, which will enable us to understand the strengths and weaknesses of different models and help us improve them.

Modelling discrete behaviours as continuous states

The Laplace approximation, one of the tools that maximizes the efficiency of TMB, cannot be used on discrete states (Bolker et al. 2013). However, this con-

straint can be overcome by using a continuous behavioural state, which in turn also allowed us to incorporate the effects of irregularly sampled observations into our movement model. Using simulations, we showed that the continuous behaviour model could adequately estimate behavioural states. Using data from a grey seal, we demonstrated that this approach can identify areas of tortuous movement. Some of these areas overlapped with those associated with high probability of area-restricted search behaviour according to hidden Markov models and seal–fish encounters (Lidgard et al. 2014). While a more in-depth comparison of the 2 methods is needed, the geographical overlap indicates that the continuous behaviour model may be adequate at identifying area-restricted search.

We used a model with continuous behaviour to maximize the use of TMB's main tools, but it is possible to implement discrete state models in TMB by using algorithms, such as the Viterbi algorithm (see Whoriskey et al. 2017). Such methods are promising, but by-passing the Laplace approximation is likely to reduce computational efficiency. TMB could be used to further refine and explore a wide range of behavioural models. Other aspects such as circadian cycles, environmental covariates and diving information could be incorporated in the modelling framework.

CONCLUSION

We showed that TMB's efficiency, accuracy and flexibility make it a promising tool to model marine movement data. In particular, we demonstrated how it can be used to model the measurement errors associated with a range of telemetry data. The models presented are only a small sample of what could be implemented with TMB. TMB could be used to model a wide variety of error structures and animal behaviours, and we have highlighted a few potential avenues of future development. In particular, models with environmental and biological covariates should be easy to implement in TMB and model comparison could be used to identify the most important covariates. Such approaches would help deepen our understanding of the underlying mechanisms behind the diverse movement patterns exhibited by marine animals.

While terrestrial telemetry is not as heavily affected by large inaccuracies in geolocations as marine telemetry, TMB could also become a useful framework to model the movement of land-dwelling animals. While conventional GPS systems can be used to study large terrestrial animals, small species are often studied with telemetry systems such as light-based geolocation (e.g. song birds, Bridge et al. 2013). In addition, the recent technological advancements in telemetry allow ecologists to collect high-resolution movement paths of terrestrial animals with a panoply of associated environmental and physiological data (Kays et al. 2015). While these increasingly large datasets will allow us to answer a breadth of new questions, their growing volume has pushed movement ecology into the realm of big data (Kays et al. 2015). TMB, which is much faster than many of the Bayesian tools used to fit complex hierarchical models to data, should provide a solution to some of these challenges and be useful for studying the movement of both marine and terrestrial animals. Thus, TMB has the potential to become an important framework in movement ecology, as well as in the many other fields that rely on complex hierarchical models.

LITERATURE CITED

Albertsen CM, Whoriskey K, Yurkowski D, Nielsen A, Mills Flemming J (2015) Fast fitting of non-Gaussian state-space models to animal movement data via Template Model Builder. Ecology 96:2598–2604

Auger-Méthé M, Field C, Albertsen CM, Derocher AE, Lewis MA, Jonsen ID, Mills Flemming J (2016a) State-space models' dirty little secrets: even simple linear Gaussian models can have estimation problems. Sci Rep 6:26677

Auger-Méthé M, Lewis MA, Derocher AE (2016b) Home ranges in moving habitats: polar bears and sea ice. Ecography 39:26–35

Beck CA, Bowen WD, McMillan JI, Iverson SJ (2003) Sex differences in the diving behaviour of a size-dimorphic capital breeder: the grey seal. Anim Behav 66:777–789

Block BA, Jonsen ID, Jorgensen SJ, Winship AJ and others (2011) Tracking apex marine predator movements in a dynamic ocean. Nature 475:86–90

Bolker BM, Gardner B, Maunder M, Berg CW and others (2013) Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. Methods Ecol Evol 4: 501–512

Bost CA, Thiebot JB, Pinaud D, Cherel Y, Trathan PN (2009) Where do penguins go during the inter-breeding period? Using geolocation to track the winter dispersion of the macaroni penguin. Biol Lett 5:473–476

Breed GA, Jonsen ID, Myers RA, Bowen WD, Leonard ML (2009) Sex-specific, seasonal foraging tactics of adult grey seals (*Halichoerus grypus*) revealed by state-space analysis. Ecology 90:3209–3221

Breed GA, Costa DP, Goebel ME, Robinson PW (2011) Electronic tracking tag programming is critical to data collection for behavioral time-series analysis. Ecosphere 2: art10

Breed GA, Costa DP, Jonsen ID, Robinson PW, Mills Flemming J (2012) State-space methods for more completely capturing behavioral dynamics from animal tracks. Ecol Model 235-236:49–58

Bridge ES, Kelly JF, Contina A, Gabrielson RM, MacCurdy RB, Winkler DW (2013) Advances in tracking small migratory birds: a technical review of light-level geolocation. J Field Ornithol 84:121–137

Bryant E (2007) 2D location accuracy statistics for Fastloc®-cores running firmware versions 2.2 & 2.3. Tech Rep TR01. Wildlife Telemetry Systems Ltd. Available at http://www.wildtracker.com/results_files/Technical%20 Report%20TR01.pdf

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach, 2nd edn. Springer-Verlag, New York, NY

Costa DP, Robinson PW, Arnould JPY, Harrison AL and others (2010) Accuracy of ARGOS locations of pinnipeds at-sea estimated using fastloc GPS. PLOS ONE 5:e8677

Dujon AM, Lindstrom RT, Hays GC (2014) The accuracy of Fastloc-GPS locations and implications for animal tracking. Methods Ecol Evol 5:1162–1169

Ekstrom PA (2004) An advance in geolocation by light. Mem Natl Inst Polar Res 58:210–226

Ekstrom PA (2007) Error measures for template-fit geolocation based on light. Deep-Sea Res II 54:392–403

Fitzpatrick R, Thums M, Bell I, Meekan MG, Stevens JD, Barnett A (2012) A comparison of the seasonal movements of tiger sharks and green turtles provides insight into their predator-prey relationship. PLOS ONE 7: e51927

Fournier DA, Skaug HJ, Ancheta J, Ianelli J and others (2012) AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim Methods Softw 27:233–249

Gaston AJ, Dechesne SB (1996) Rhinoceros auklet (*Cerorhinca monocerata*). In: Poole A (ed) The birds of North America online. Cornell Lab of Ornithology, Ithaca, NY

Gaston AJ, Smith PA, Tranquilla LM, Montevecchi WA and others (2011) Movements and wintering areas of breeding age thick-billed murre *Uria lomvia* from two colonies in Nunavut, Canada. Mar Biol 158:1929–1941

Greene C, Block BA, Welch D, Jackson G, Lawson GL, Rechisky EL (2009) Advances in conservation oceanography: new tagging and tracking technologies and their potential for transforming the science underlying fisheries management. Oceanography 22:210–223

Hammerschlag N, Gallagher AJ, Lazarre DM (2011) A review of shark satellite tagging studies. J Exp Mar Biol Ecol 398:1–8

Hoenner X, Whiting SD, Hindell MA, McMahon CR (2012) Enhancing the use of Argos satellite data for home range and long distance migration studies of marine animals. PLOS ONE 7:e40713

Hussey NE, Kessel ST, Aarestrup K, Cooke SJ and others (2015) Aquatic animal telemetry: a panoramic window into the underwater world. Science 348:1255642

Jessopp MJ, Cronin M, Doyle TK, Wilson M, McQuatters-Gollop A, Newton S, Phillips RA (2013) Transatlantic migration by post-breeding puffins: a strategy to exploit a temporarily abundant food resource? Mar Biol 160: 2755–2762

Johnson DS, London JM, Lea MA, Durban JW (2008) Continuous-time correlated random walk model for animal telemetry data. Ecology 89:1208–1215

Jonsen ID, Mills Flemming J, Myers RA (2005) Robust state-space modeling of animal movement data. Ecology 86: 2874–2880

Jonsen I, Basson M, Bestley S, Bravington M and others (2013) State-space models for bio-loggers: a methodological road map. Deep-Sea Res II 88-89:34–46

Kass RE, Tierney L, Kadane JB (1990) The validity of posterior expansions based on Laplace's method. In: Geisser S, Hodges JS, Press SJ, Zellner A (eds) Bayesian and likelihood methods in statistics and econometrics: essays in honor of George A. Barnard, Vol 7. Elsevier Science Publishers B.V./North-Holland, Amsterdam, p 473–488

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

Kennedy AS, Zerbini AN, Rone BK, Clapham PJ (2014) Individual variation in movements of satellite-tracked humpback whales *Megaptera novaeangliae* in the eastern Aleutian Islands and Bering Sea. Endang Species Res 23: 187–195

Kristensen K, Nielsen A, Berg CW, Skaug H, Bell BM (2016) TMB: automatic differentiation and Laplace approximation. J Stat Softw 70:1–21

Kuhn CE, Johnson DS, Ream RR, Gelatt TS (2009) Advances in the tracking of marine species: using GPS locations to evaluate satellite track data and a continuous-time movement model. Mar Ecol Prog Ser 393:97–109

Lam CH, Nielsen A, Sibert JR (2010) Incorporating sea-surface temperature to the light-based geolocation model TrackIt. Mar Ecol Prog Ser 419:71–84

Lidgard DC, Bowen WD, Jonsen ID, Iverson SJ (2014) Predator-borne acoustic transceivers and GPS tracking reveal spatiotemporal patterns of encounters with acoustically tagged fish in the open ocean. Mar Ecol Prog Ser 501: 157–168

Lowther AD, Lydersen C, Fedak MA, Lovell P, Kovacs KM (2015) The Argos-CLS Kalman filter: error structures and state-space modelling relative to Fastloc GPS data. PLOS ONE 10:e0124754

McClintock BT, London JM, Cameron MF, Boveng PL (2015) Modelling animal movement using the Argos satellite telemetry location error ellipse. Methods Ecol Evol 6:266–277

Musyl MK, Brill RW, Curran DS, Gunn JS and others (2001) Ability of archival tags to provide estimates of geographical position based on light intensity. In: Sibert JR, Nielsen JL (eds) Electronic tagging and tracking in marine fisheries, Vol 1. Kluwer Academic Press, Dordrecht, p 343–367

Nielsen A, Bigelow KA, Musyl MK, Sibert JR (2006) Improving light-based geolocation by including sea surface temperature. Fish Oceanogr 15:314–325

Patterson TA, McConnell BJ, Fedak MA, Bravington MV, Hindell MA (2010) Using GPS data to evaluate the accuracy of state-space methods for correction of Argos satellite telemetry error. Ecology 91:273–285

Phillips RA, Xavier JC, Croxall JP (2003) Effects of satellite transmitters on albatrosses and petrels. Auk 120: 1082–1090

Phillips RA, Silk JRD, Croxall JP, Afanasyev V, Briggs DR (2004) Accuracy of geolocation estimates for flying seabirds. Mar Ecol Prog Ser 266:265–272

Pollet IL, Ronconi RA, Jonsen ID, Leonard ML, Taylor PD, Shutler D (2014) Foraging movements of Leach's storm-petrels *Oceanodroma leucorhoa* during incubation. J Avian Biol 45:305–314

Pütz K, Trathan PN, Pedrana J, Collins MA, Poncet S, Lüthi B (2014) Post-fledging dispersal of king penguins (*Aptenodytes patagonicus*) from two breeding sites in the South Atlantic. PLOS ONE 9:e97164

Reisinger RR, Keith M, Andrews RD, de Bruyn PJN (2015) Movement and diving of killer whales (*Orcinus orca*) at a Southern Ocean archipelago. J Exp Mar Biol Ecol 473: 90–102

Silva MA, Jonsen I, Russell DJF, Prieto R, Thompson D, Baumgartner MF (2014) Assessing performance of Bayesian state-space models fit to Argos satellite telemetry locations processed with Kalman filtering. PLOS ONE 9: e92277

Skaug HJ, Fournier DA (2006) Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. Comput Stat Data Anal 51:699–709

Takahashi A, Ochi D, Watanuki Y, Deguchi T and others (2008) Post-breeding movement and activities of two streaked shearwaters in the north-western Pacific. Ornitholog Sci 7:29–35

Takahashi A, Ito M, Suzuki Y, Watanuki Y and others (2015) Migratory movements of rhinoceros auklets in the north-western Pacific: connecting seasonal productivities. Mar Ecol Prog Ser 525:229–243

Thiebot JB, Pinaud D (2010) Quantitative method to estimate species habitat use from light-based geolocation data. Endang Species Res 10:341–353

Thiers L, Delord K, Barbraud C, Phillips RA, Pinaud D, Weimerskirch H (2014) Foraging zones of the two sibling species of giant petrels in the Indian Ocean throughout the annual cycle: implication for their conservation. Mar Ecol Prog Ser 499:233–248

Tomkiewicz SM, Fuller MR, Kie JG, Bates KK (2010) Global positioning system and associated technologies in animal behaviour and ecological research. Philos Trans R Soc Lond B Biol Sci 365:2163–2176

Vincent C, McConnel BJ, Ridoux V, Fedak MA (2002) Assessment of Argos location accuracy from satellite tags deployed on captive grey seals. Mar Mamm Sci 18: 156–166

Whoriskey K, Auger-Méthé M, Albertsen CM, Whoriskey FG, Binder TR, Krueger CC, Mills Flemming J (2017) A hidden Markov movement model for rapidly identifying behavioral states from animal tracks. Ecol Evol (in press)

Winship AJ, Jorgensen SJ, Shaffer SA, Jonsen ID, Robinson PW, Costa DP, Block BA (2012) State-space framework for estimating measurement error from double-tagging telemetry experiments. Methods Ecol Evol 3:291–302

Young HS, Maxwell SM, Conners MG, Shaffer SA (2015) Pelagic marine protected areas protect foraging habitat for multiple breeding seabirds in the central Pacific. Biol Conserv 181:226–235