Modeling the Diving Behavior of Whales: A Latent-Variable Approach with Feedback and Semi-Markovian Components

Roland LANGROCK, Tiago A. MARQUES, Robin W. BAIRD, and Len THOMAS

Recent years have seen a fast-growing body of literature concerned with the statistical modeling of animal movement in the two horizontal dimensions. On the other hand, there is very little statistical work that deals with animal movement in the vertical dimension. We present an approach that provides an important step in analyzing such data. In particular, we introduce a hidden Markov-type modeling approach for time series comprising the depths of a diving marine mammal, thus modeling movement in the water column. We first develop a baseline Markov-switching model, which is then extended to incorporate feedback and semi-Markovian components, motivated by the observations made for a particular species, Blainville's beaked whale (*Mesoplodon densirostris*). The application of the proposed model to the beaked whale data reveals both strengths and weaknesses of the suggested modeling framework. The framework is general enough that we anticipate that it can be used for many other species given minor changes in the model structure.

Key Words: Behavioral state; Distance sampling; Hidden Markov model; Maximum likelihood; Movement model.

1. INTRODUCTION

Understanding and being able to quantitatively describe the way animals dive is important for various reasons. For example, characterizing via a statistical model the movement process of diving animals in the absence of a disturbance such as anthropogenic noise, and measuring the change in the model parameters in the presence of distur-

Roland Langrock (✉) is a Research Fellow (E-mail: *roland@mcs.st-and.ac.uk*), Tiago A. Marques is a Senior Research Fellow, Len Thomas is a Reader, Centre for Research into Ecological and Environmental Modelling (CREEM) and School of Mathematics and Statistics, University of St Andrews, St Andrews, UK. Robin W. Baird is a Research Biologist, Cascadia Research Collective, Olympia, WA, USA.

bance, might be an effective way to quantify impacts (cf. Tyack et al. 2011). Furthermore, a good mathematical description of diving aids in the understanding of biological processes such as gas exchange and physiology, and how these processes may be affected by anthropogenic disturbance (cf. Hooker et al. 2009). Thirdly, diving and surfacing patterns can have direct influence on the performance of methods used to estimate animal abundance, since they govern the availability for detection (Borchers et al. 2013; Langrock et al. 2013). Finally, and more generally, the quantitative description of an animal's movement may contribute to our understanding of the animal's movement decisions (e.g., Schick et al. 2008).

While there is a vast literature on modeling animal movement in the two horizontal dimensions, there is relatively little work that discusses statistical inference for movement in the vertical dimension. For dive data in particular, an exception is a recent paper by Higgs and Ver Hoef (2012), which deals with categorical and thus somewhat simplified data related to depths. McClintock et al. (2013) also incorporate categorical depth data in their movement model. Bailleul et al. (2008) apply a first-passage time analysis in the depth dimension to detect behavioral changes of southern elephant seals. Walker et al. (2011) use splines to classify dive shapes, but do not attempt to model an entire movement path. Dowd and Joy (2011) use an autoregressive process of order two to model vertical velocity data collected on seals, but for example do not take into account the bounding nature of the sea surface, such that simulations from their model would lead to unrealistic movement patterns. The following comment, made by Houser (2006), is still accurate: "[...] whereas movement and behavior models for terrestrial mammals enjoy a rich history of development and application, similar models for marine mammals are all but nonexistent." In this paper, we propose and illustrate a flexible modeling framework for the analysis of such data.

In particular, we suggest a hidden Markov-type modeling approach for time series comprising the depths of a diving whale, thus modeling movement in the water column. Hidden Markov models (HMMs) are obvious candidates for modeling animal movement because they allow movement patterns to be correlated over time, but in a way such that the pattern occasionally changes according to switches in an underlying, non-observable behavioral state process. In recent years, HMMs and related models have been used several times to model movement of animals in the two horizontal dimensions, neglecting the vertical dimension (see, e.g.: Holzmann et al. 2006; Patterson et al. 2009; Langrock et al. 2012). Here, we focus specifically on movement of marine mammals in the vertical dimension.

While there does not seem to be a consistent use of terminology in the literature, the label "HMM" usually refers to processes that obey the assumption of the observations being conditionally independent, given the states. Models that involve additional dependencies between observations are often referred to as "Markov-switching models" (see Zucchini and MacDonald 2009). We consider models that involve such dependencies, and will hence mostly refer to them as Markov-switching models. However, the distinction is somewhat artificial since the methods we consider, including the model fitting strategy, exploit (standard and nonstandard) HMM inference machinery. In addition to the extra dependencies at the observation level, and motivated by movement data obtained for a beaked whale, we

consider two other important extensions to the basic Markov-switching model: (1) incorporating feedback from the observed process on the subsequent generation of states, and (2) allowing for some state dwell-time distributions (describing the time spent in a state before changing to a different state) to be non-geometric.

Our motivating example concerns Blainville's beaked whale (*Mesoplodon densirostris*), a species that has been the focus of a considerable amount of recent research. This has been in part motivated by several mass strandings associated with naval active sonar operations (Cox et al. 2006; D'Amico et al. 2009), and the resulting conjecture, recently confirmed (e.g., McCarthy et al. 2011; Tyack et al. 2011), that the species might be particularly sensitive to such disturbance. The behavioral characteristics of these whales make them very difficult to survey using visual methods, as they spend a rather small proportion of their time near the surface. They make some of the most extreme dives recorded for cetaceans (Baird et al. 2006; Tyack et al. 2006). The data set we use was collected using a time-depth recorder, and represents the longest high resolution depth series recorded for this species, ranging over 78 hours in duration.

The paper is structured as follows. Section 2 describes the motivating and method-illustrating data set, which is the series of depths recorded for an individual Blainville's beaked whale. The general modeling framework that we suggest is introduced in Section 3. Section 4 develops a model of the suggested type and applies it to the beaked whale data. A discussion of the modeling approach, its potential applications and future work is given in Section 5.

2. MOTIVATING BEAKED WHALE DATA

We consider the time series of depths observed for a single adult female Blainville's beaked whale, tagged off the west side of the island of Hawaii. A suction-cup attached tag, containing an Mk9 time-depth recorder (Wildlife Computers, Washington, USA), was attached to the whale. The tag detached after 78.73 hours (an unusually long period of time for this type of attachment), floated to the surface and was located and recovered using the signals from a built-in VHF transmitter. Depth was measured every second, but this resolution seems overly fine with respect to the modeling of behavioral states. Thus, we coarsened the resolution by a factor of 10 to decrease the computational effort, while still maintaining a fairly high resolution in order to get a detailed picture of the whale's movement pattern. The resulting 28343 depth observations are illustrated in Figure 1. More detailed information on the data and on the behavior and physiology of the studied species is given in Baird et al. (2008).

Blainville's beaked whales present two types of dives, shallow dives and deep dives, which have very distinct characteristics, given their also very distinct purposes. Deep dives have been associated with foraging, and have been recorded to be 835 meters deep and 47 minutes long, on average (Tyack et al. 2006), which puts these animals amongst those who dive deepest and longest among the air breathing creatures. On the other hand, shallow dives are believed to have no foraging purposes, with no echolocation-mediated foraging recorded during those (Tyack et al. 2006), and mean depths around 70 meters deep and 9
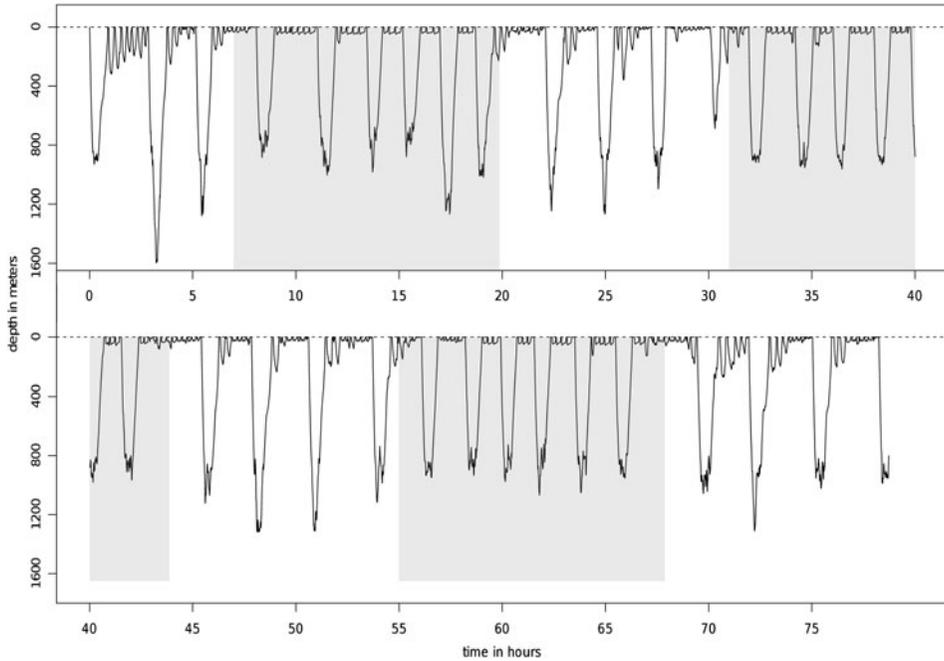
Figure 1. Beaked whale depth profile. The shaded periods correspond to "night" periods, between 17:44 h (sunset) and 06:37 h (sunrise).

minutes long. Tyack et al. (2006) have associated these with a period of recovery from the build-up of metabolites of anaerobic metabolism.

3. A LATENT-VARIABLE MODELING FRAMEWORK FOR DIVE DATA

3.1. BASELINE MARKOV-SWITCHING MODEL

We initially develop a baseline model for time series of depths related to a diving marine mammal. Clearly, the movement pattern of a marine mammal changes according to changes of its behavioral state. For example, on the descent to its foraging grounds an animal will move differently than when it actually is foraging. This motivates the use of mixture models, and in particular of dependent mixtures such as Markov-switching models. In a basic N-state Markov-switching model, the underlying state process, denoted by S_t, is Markovian, such that at any time $t + 1$ the state the animal is in depends only on what state the animal was in at time t. We denote the state transition probabilities at time t by $\omega_{ij}^{(t)}$, $i, j = 1, \ldots, N$, with $\omega_{ij}^{(t)} = \Pr(S_{t+1} = j | S_t = i)$, and summarize them in the transition probability matrix (t.p.m.) $\mathbf{\Omega}^{(t)}$, with $\omega_{ij}^{(t)}$ the entry in row i and column j. The Markov chain S_t represents the non-observable (behavioral) state process which determines the distribution at the observation level, i.e., that of the depths.

Since depths are continuous-valued and non-negative, we consider gamma distributions for the state-dependent process, which we denote by Z_t. Clearly, the distribution of the

depth observed at time t depends on both the depth at time $t-1$ and on the state the animal is in. For example, at time t the expected depth of an animal on the ascent from a dive is expected to be slightly smaller than its depth at time $t-1$. Thus, we reparameterize the gamma distribution in terms of a mean and a standard deviation parameter, μ and σ, respectively (such that shape $= \mu^2/\sigma^2$ and scale $= \sigma^2/\mu$), and assume that

$$Z_t \mid S_t = i, Z_{t-1} = z_{t-1} \sim \text{Gamma}\big(\mu = \max(z_{t-1} + \theta_i, 1), \sigma = \nu_i\big). \tag{1}$$

Here θ_i and ν_i depend on the state the animal is in, with θ_i giving the expected change in depth when in state i, and ν_i describing the uncertainty in the corresponding change. The max operator constrains the mean of the depth distribution to be greater than or equal to one, which is done in order to avoid ending up with an impossible negative mean (if the drift θ_i is negative and exceeds z_{t-1} in absolute value). In the application to the beaked whale depth time series, we will consider models with $N = 7$ states, corresponding to "surfacing" periods and "descent", "bottom" and "ascent" phases for shallow and for deep dives, respectively. In the following we let $f_i(z_t|z_{t-1})$ denote the conditional probability density function of the gamma distribution for depth at time t, given the previous depth z_{t-1} and with parameters θ_i and ν_i, as defined in (1). An alternative way to model a series of depths would be to model depth displacements, rather than actual depths. However, such an approach would require the state-dependent distributions to be truncated flexibly according to the current distance to the surface.

Conditional on an initial depth, $Z_0 = z_0$, the likelihood of this model is given by

$$\mathcal{L} = f(z_1, \ldots, z_T | z_0)$$

$$= \sum_{i_1=1}^{N} \cdots \sum_{i_T=1}^{N} f(z_1, \ldots, z_T | S_1 = i_1, \ldots, S_T = i_T, z_0) f(S_1 = i_1, \ldots, S_T = i_T)$$

$$= \sum_{i_1=1}^{N} \cdots \sum_{i_T=1}^{N} \prod_{t=1}^{T} f_{i_t}(z_t | z_{t-1}) \Pr(S_1 = i_1) \prod_{t=1}^{T-1} \omega_{s_{i_t}, s_{i_{t+1}}}^{(t)},$$

where $z_0, z_1, \ldots, z_T$ denote the observed depths. While the observations often directly reveal the underlying behavioral state, the states are in general not directly observed, such that within the likelihood we need to sum over all possible state sequences that may have given rise to the observed depths. This brute force evaluation of the likelihood involves N^T function evaluations, which clearly makes a numerical maximization impossible even for moderate values of N and T. However, there is an alternative, much more efficient way to evaluate the likelihood, given by a recursive scheme called the forward algorithm. Applying the forward algorithm, the likelihood can be expressed as

$$\mathcal{L} = \delta \mathbf{P}(z_1|z_0)\mathbf{\Omega}^{(1)}\mathbf{P}(z_2|z_1)\mathbf{\Omega}^{(2)} \cdot \ldots \cdot \mathbf{\Omega}^{(T-2)}\mathbf{P}(z_{T-1}|z_{T-2})\mathbf{\Omega}^{(T-1)}\mathbf{P}(z_T|z_{T-1})\mathbf{1}^t, \tag{2}$$

where $\mathbf{P}(z_t|z_{t-1}) = \text{diag}(f_1(z_t|z_{t-1}), \ldots, f_N(z_t|z_{t-1}))$, $\mathbf{1} \in \mathbb{R}^N$ is a row vector of ones and δ denotes the initial distribution of the Markov chain, i.e., $\delta_i = \Pr(S_1 = i)$. Regarding the choice of δ, one may simply model a dive profile that starts with a movement pattern that can clearly be assigned to one of the states, such that δ is effectively known and can

thus be fixed correspondingly. Alternatively, one can either estimate $\boldsymbol{\delta}$ or use the stationary distribution of the Markov chain S_t, if it exists. The representation (2) is that characteristic of HMMs, with the computational cost now linear in the number of observations, T. The model parameters can be estimated by direct numerical likelihood maximization, subject to well-known technical issues (see Zucchini and MacDonald 2009, for a detailed account of these and also for details on how to compute the likelihood of an HMM).

While within the likelihood evaluation the hidden states are essentially treated as nuisance variables, it is often of interest to also estimate the most likely states, given some parameter estimates (both for diagnostic purposes and also as the states are sometimes of interest in their own right). To do this, an efficient dynamic programming algorithm called the Viterbi algorithm can be applied, yielding the sequence of states $s_1^*, \ldots, s_T^*$ that is most likely to have given rise to the observed sequence of depths, under the fitted model:

$$\left(s_1^*, \ldots, s_T^*\right) = \underset{(s_1, \ldots, s_T) \in \{1, \ldots, N\}^T}{\operatorname{argmax}} \Pr(S_1 = s_1, \ldots, S_T = s_T \mid Z_0 = z_0, \ldots, Z_T = z_t);$$

for details see Chapter 5 in Zucchini and MacDonald (2009).

3.2. INCORPORATING FEEDBACK

Whether or not a whale switches, say, away from a descent phase, does not depend only on what state the whale is currently in. Instead, a whale will typically switch away from a descent phase when it arrives at its foraging grounds (i.e., at some relevant depths). Likewise, an animal will switch away from an ascent phase when it reaches the vicinity of the surface. Thus, the sequence of behavioral states will usually not be well-described by a simple Markov chain; dependence on the actual depths is key and should be taken into account. This can be accomplished by considering feedback models in the spirit of Zucchini et al. (2008) and Langrock (2012). In the given context, we suggest modeling some of the state transition probabilities as functions of the current depth, z_t. For example, we can use a polynomial logistic regression of order d to link $\omega_{ij}^{(t)}$ to z_t:

$$\omega_{ij}^{(t)}(z_t) = \Pr(S_{t+1} = j \mid S_t = i, Z_t = z_t) = \operatorname{logit}^{-1}\left(\sum_{k=0}^{d} \lambda_k z_t^k\right). \tag{3}$$

For example, we expect the probability of leaving the descent phase in a deep dive cycle to be very close to zero for small values of z_t, and to increase considerably when the animal approaches the depths that are relevant for foraging (>700 meters, roughly, in case of the example beaked whale). For such a model, the likelihood structure (2) remains valid, with $\boldsymbol{\Omega}^{(t)}$ replaced by $\boldsymbol{\Omega}^{(t)}(z_t)$, for $t = 1, \ldots, T-1$, where the t.p.m. $\boldsymbol{\Omega}^{(t)}(z_t)$ contains the transition probabilities at time t, some of which are depth-dependent. While the likelihood structure is still that characteristic of HMMs, such feedback models are neither Markov-switching nor hidden Markov models in the strict sense, since the state process is not Markovian. Thus, we refer to such a model as a latent-variable model with feedback, closely following Zucchini et al. (2008) who used the term "latent-state model" to refer to such feedback models.

3.3. Semi-Markovian Components

Another issue with the baseline model is that in a basic Markov-switching model the state dwell times (i.e., the durations of stays in the different states) are necessarily geometrically distributed. This can be unrealistic, for example for the bottom phase of deep dives: having spent a lot of physiological capital going down hundreds of meters to the depths at which the prey exists, it would make no sense for an animal to almost immediately begin its ascent. Other marine mammals, such as, e.g., sperm whales and short-finned pilot whales, spend long periods resting at the surface, such that the dwell time of the surface state may not be well-described by a geometric distribution.

So-called hidden *semi*-Markov models (HSMMs) relax the corresponding condition: the dwell time in an HSMM state is explicitly modeled by some distribution on the positive integers (Guédon 2003). Langrock and Zucchini (2011) demonstrate that specially structured HMMs can capture the 'semi'-property of HSMMs, i.e., that HMMs can accommodate—at least approximately—any desired state dwell-time distribution. This means that HSMMs can conveniently be fitted using the HMM inference machinery. The idea is to expand any semi-Markovian state into a large set of Markovian states sharing the same state-dependent distribution, structuring the transition probabilities between those states such that the desired dwell-time distribution is accurately approximated. This approach allows to consider arbitrary dwell-time distributions while preserving the Markov property of the latent process. Thus, we maintain the likelihood. Therefore, this extension leads to an increased computational effort (see remarks in Section 5).

3.4. Remarks on Fine-Tuning of the Model

Putting all these components together, we obtain a natural and flexible modeling framework, in which we allow for movement patterns to vary across different behavioral states and where we accommodate autocorrelation of the observations. In general, the exact way the feedback and semi-Markovian extensions need to be set up in order to make a suitable model, and also the exact configuration of the state structure, highly depend on both the specific movement patterns exhibited by the species that is considered and on the purpose of the analysis (e.g., obtaining a realistic pattern of times at the surface for use in estimating the "availability bias" in visual abundance estimation methods). In Section 4, we discuss the particular case represented by the beaked whale data described in Section 2. While the details of the model structure are thus developed specifically for this species and the given series, we believe that the suggested type of approach, using the baseline model discussed in Section 3.1 and potentially incorporating the extensions outlined in Sections 3.2 and 3.3, can be applied to many other marine mammal species, typically after some straightforward modifications of details of the model structure. Because of the simple structure of the HMM likelihood, it is relatively easy to fit a suite of different models thought to be biologically plausible, and use standard model selection, goodness-of-fit, and diagnostic tools to select among them.

4. CASE STUDY: BEAKED WHALE DATA

4.1. MODEL DEVELOPMENT AND INFERENCE

For the beaked whale series described in Section 2, with depth observations every 10 s, it seems plausible to distinguish three primary behavioral states: "surfacing" (state 1 in the following), "deep-diving" and "shallow-diving". Since the movement pattern within each of the latter two states clearly is not homogeneous over time, we further subdivide the state associated with deep dives into states 2 ("descent"), 3 (subsequently referred to as "foraging", although this nominal state does not necessarily imply that the animal is actually feeding) and 4 ("ascent"), and analogously the state associated with shallow dives into states 5, 6, and 7 (with state 6 referred to as the "bottom" of a shallow dive). Thus, we consider a model with seven states, with some natural constraints on the state transitions. In particular, we assume that a stay in state 1 is always followed by the animal passing either through the states 2, 3, 4, and 1 (in that order, thus performing a deep dive), or through the states 5, 6, 7, and 1 (in that order, thus performing a shallow dive). Some transitions are thus impossible, and hence there are considerably fewer parameters to be estimated than usually would be the case for a seven-state Markov-switching model.

To account for the fact that the state switching dynamics to some extent depend on the actual depths, we incorporate several feedback mechanisms in the model. We used the Akaike Information Criterion to choose the orders of the feedback mechanisms—d in (3)—trying a maximum order of two in each case. We model the transitions to the bottom parts of the dives, i.e., the conditional probability of the animal switching to state 3 ("foraging") at time $t + 1$, given it is in state 2 ("deep dive descent") at time t, and the conditional probability of the animal switching to state 6 ("bottom of a shallow dive") at time $t + 1$, given it is in state 5 ("shallow dive descent") at time t, as functions of depth at time t, z_t:

$$
\omega_{23}^{(t)}(z_t) = \Pr(S_{t+1} = 3 | S_t = 2, Z_t = z_t)
$$

$$
= \begin{cases} \text{logit}^{-1}(\alpha_0 + \alpha_1 z_t + \alpha_2 z_t^2) & \text{if } \alpha_2 \geq 0 \text{ or } z_t < -\frac{\alpha_1}{2\alpha_2}; \\ \text{logit}^{-1}(\alpha_0 - \frac{\alpha_1^2}{4\alpha_2}) & \text{otherwise}; \end{cases} \tag{4}
$$

$$
\omega_{56}^{(t)}(z_t) = \Pr(S_{t+1} = 6 | S_t = 5, Z_t = z_t) = \text{logit}^{-1}(\beta_0 + \beta_1 z_t + \beta_2 z_t^2),
$$

with coefficients $\alpha_0, \alpha_1, \alpha_2, \beta_0, \beta_1 \in \mathbb{R}$. Transition probabilities involving a quadratic predictor tend to zero as the covariate value increases if the coefficient associated with the quadratic term is negative. In such a case an animal may in simulations be "trapped" in the corresponding state: e.g., on a descent in a deep dive cycle an animal may pass the relevant layers of the water column without leaving the state, such that at some point the probability of leaving the state becomes approximately zero. To avoid this undesirable feature, we used the slight modification given in (4), specifying the quadratic predictor in $\omega_{23}^{(t)}(z_t)$ to be constant for values higher than the value at which the maximum is attained, which results in an alternative smooth continuation of the predictor function. The logistic regression used to model $\omega_{56}^{(t)}(z_t)$ is, in principle, susceptible to the same problem, but in our application the coefficient β_2 was estimated to be positive, such that no differentiation between two cases,

as for $\omega_{23}^{(t)}(z_t)$, was necessary. Future work could explore alternative functional forms for these predictors. For example, one could consider mixtures of nondecreasing functions to obtain flexible forms that are less susceptible to the problems with quadratic terms. In particular, one could consider linear combinations of I-splines (Ramsay 1988). However, it can be too restrictive to consider nondecreasing functions only (cf. our results for $\omega_{56}^{(t)}(z_t)$ given below).

For the conditional probabilities of switching to state 1 ("surfacing") at time $t+1$, given state 4 ("deep dive ascent") or given state 7 ("shallow dive ascent"), respectively, at time t, we assume

$$\omega_{41}^{(t)}(z_t) = \omega_{71}^{(t)}(z_t) = \text{logit}^{-1}(\gamma_0 + \gamma_1 z_t),$$

with $\gamma_0, \gamma_1 \in \mathbb{R}$. We assume common parameters for the transitions from state 4 ("deep dive ascent") to state 1 ("surfacing") and from state 7 ("shallow dive ascent") to state 1, respectively, since clearly the whale switches from an "ascent" mode to the "surfacing" mode when it reaches the vicinity of the surface, and it is not expected that there is a relevant difference in this mechanism between shallow dive and deep dive cycles (although the different velocities in the states could in principle make a minor difference). Finally, after a deep dive, the considered beaked whale seems to remain relatively near the surface for some time, which we want to explicitly incorporate in the model. Thus, we extend the feedback mechanisms suggested in Section 3.2 by considering—in addition to the state process, S_t, and the state-dependent depth process, Z_t—a third process, R_t, giving the time since the whale was last more than 500 meters deep. In the following we refer to R_t as the "time since the last deep dive". The probability of beginning a deep dive given a surfacing period finishes, denoted by $\pi_t(r_t)$, is then modeled as a function of the time that has passed since the last deep dive:

$$\pi_t(r_t) = \text{Pr}(S_{t+1} = 2 | S_t = 1, S_{t+1} \neq 1, R_t = r_t) = \text{logit}^{-1}(\zeta_0 + \zeta_1 r_t),$$

with $\zeta_0, \zeta_1 \in \mathbb{R}$. In this model, if the whale switches away from state 1 in the time interval $(t, t+1]$, then it will start a deep dive with probability $\pi_t(r_t)$ and a shallow dive with probability $1 - \pi_t(r_t)$.

Figure 2 illustrates the dependence structure of a Markov-switching model that involves feedback of the suggested type (but no semi-Markovian states). The feedback mechanisms are represented by the arrows from the processes R_t and Z_t to the state process S_t. State transition probabilities not affected by feedback (in one of the ways described above) are assumed to be homogeneous over time, such that in the following we omit the superscript in $\omega_{ij}^{(t)}$ for corresponding i, j. Due to the assumptions made concerning the dive cycle, most of the state transition probabilities necessarily equal zero. We summarize the transition
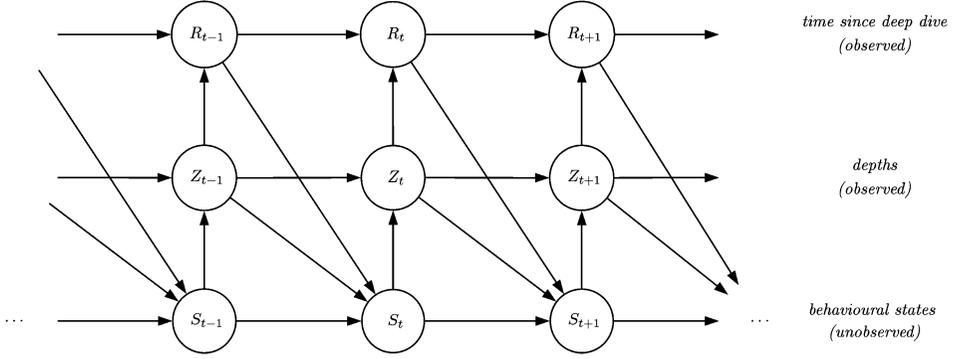
Figure 2.   Dependence structure of a feedback Markov-switching model for the time series of depths.

probabilities at time t in the t.p.m. $\mathbf{\Omega}^{(t)}(z_t, r_t)$, with

$$
\mathbf{\Omega}^{(t)}(z_t, r_t)
$$
$$
= \begin{pmatrix}
\omega_{11} & (1-\omega_{11})\pi_t(r_t) & 0 & 0 & (1-\omega_{11})(1-\pi_t(r_t)) & 0 & 0 \\
0 & 1-\omega_{23}^{(t)}(z_t) & \omega_{23}^{(t)}(z_t) & 0 & 0 & 0 & 0 \\
0 & 0 & \omega_{33} & 1-\omega_{33} & 0 & 0 & 0 \\
\omega_{41}^{(t)}(z_t) & 0 & 0 & 1-\omega_{41}^{(t)}(z_t) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1-\omega_{56}^{(t)}(z_t) & \omega_{56}^{(t)}(z_t) & 0 \\
0 & 0 & 0 & 0 & 0 & \omega_{66} & 1-\omega_{66} \\
\omega_{71}^{(t)}(z_t) & 0 & 0 & 0 & 0 & 0 & 1-\omega_{71}^{(t)}(z_t)
\end{pmatrix}.
$$

Conditional on an initial depth, $Z_0 = z_0$, and on an initial time since the last deep dive, $R_0 = r_0$, the likelihood of the considered feedback Markov-switching model still has the typical HMM matrix product structure given in (2), only $\mathbf{\Omega}^{(t)}$, $t = 1, \ldots, 28341$, needs to be replaced by $\mathbf{\Omega}^{(t)}(z_t, r_t)$ due to the inclusion of the feedback components.

Finally, for the considered beaked whale dive series, the assumption of a geometric state dwell-time distribution seems unrealistic for the "foraging" mode (state 3); see remarks above. We thus use the approach described in Section 3.3 to extend the seven-state feedback model by assuming a negative binomial state dwell-time distribution in state 3, with mean parameter η and size parameter q. The associated probability mass function, expressed in terms of q and $\phi = \eta/(\eta + q)$, is given by

$$
p(k) = \frac{\Gamma(k + q - 1)}{(k-1)!\,\Gamma(q)} \phi^q (1 - \phi)^{k-1}, \quad k = 1, 2, 3, \ldots.
$$

The geometric distribution is a special case ($q = 1$). Considering a more flexible class of dwell-time distributions for state 3 is particularly important if one is interested in the foraging behavior, since it is this deep-diving bottom phase that has been associated with active search for an acquisition of prey (Tyack et al. 2006). For example, one component of an energetics model may be the distribution of time in this state; alternatively a prey characterization study may require a distribution of depths at which active foraging takes place.

By implementing the approach outlined in Section 3.3, we maintain the likelihood structure given in (2), but due to state 3 being expanded into a set of m states, the dimensions

of both $\mathbf{\Omega}^{(t)}(z_t, r_t)$ and $\mathbf{P}(z_t | z_{t-1})$ become $(m + 6) \times (m + 6)$ (cf. Langrock et al. 2012). The value chosen for m needs to be sufficiently large to ensure a close approximation to the HSMM—ideally such that the interval $[1, m]$ contains almost all the mass of the state dwell-time distribution, and at least such that the p.m.f. of the state dwell-time distribution monotonically decreases for values larger than m (cf. Langrock et al. 2012). We used $m = 250$ in the given application, which turns out to be a conservative choice, as the interval $[1, 250]$ contains $\approx 99.8\%$ of the mass of the fitted state dwell-time distribution.

The described seven-state model, including feedback components and one semi-Markovian component, was fitted to the series of depths described in Section 2 via numerical maximum likelihood estimation using the optimizer `nlm` in R (R Core Team 2012). Confidence intervals were obtained based on the Hessian of the log-likelihood. We reparameterized all constrained parameters in terms of unconstrained parameters. In particular, we used log transformations to enforce θ_2 and θ_5 to be positive and θ_4 and θ_7 to be negative, such that the mean depth displacement was restricted to be positive in the descent phases and to be negative in the ascent phases. The likelihood was maximized with respect to the unconstrained parameters. For transformed parameters we obtained approximate standard errors (and hence confidence intervals) for the parameters of interest, i.e., the constrained parameters, using the delta method. The example time series was assumed to start with the animal being in state 2 ("deep dive descent"), so we fixed the initial distribution $\boldsymbol{\delta}$ accordingly. This also means that the initial value used for the time since the last deep dive, r_0, has no effect on the likelihood (because it only influences transitions away from the surfacing state, and such transitions do not occur before the animal has finished the first deep dive, at which point the time since the last deep dive is known). The R code used to estimate the model parameters is provided in the online supplementary material to this manuscript.

4.2. RESULTS

The parameter estimates and associated 95 % confidence intervals, given in Table 1, reveal some intriguing features of the analyzed time series. First, the deviation of the fitted dwell-time distribution in state 3 ("foraging") from a geometric is striking, with the mode of the fitted distribution being at 131, and the 95 % confidence interval for the size parameter not containing the value one (which corresponds to the special case of a geometric distribution, and note that the value one is actually far outside the 95 % confidence interval, which has the lower boundary 10.19). This gives evidence for the inappropriateness of geometric dwell-time distributions, and hence simple Markov-switching or hidden Markov models, in this case. Second, the estimated state-dependent mean displacements indicate that both the descent and the ascent phase in deep dive cycles involve a vertical speed that is considerably higher, on average, than that in the corresponding phases in a shallow dive cycle. The estimated depth displacement parameters in states 2, 4, 5, and 7 correspond to vertical speeds of 1.5, 0.8, 0.4 and 0.4 meters per second, respectively, figures that match the findings of Tyack et al. (2006) remarkably well: after manually classifying observations into states, those authors obtain corresponding values of 1.6, 0.7, 0.3 and 0.3.

Figure 3 illustrates the various feedback mechanisms for the fitted model. The corresponding limits of the pointwise 95 % confidence intervals occasionally lie outside the

Table 1. Parameter estimates for the model with feedback and a semi-Markovian component, fitted to the beaked whale data, and 95 % confidence intervals of the estimates.

Related to state	State process			State-dep. process		
	Parameter	Estimate	95 % CI	Parameter	Estimate	95 % CI
1	ω_{11}	0.93	[0.92, 0.94]	θ_1	−4.84	[−5.10, −4.58]
("surfacing")	ζ_0	−3.83	[−4.79, −2.87]	ν_1	0.49	[0.48, 0.50]
	$\zeta_1 \cdot 10^3$	3.28	[1.82, 4.74]			
2	α_0	−23.71	[−36.31, −11.11]	θ_2	14.87	[14.61, 15.13]
("deep dive	$\alpha_1 \cdot 10^3$	46.97	[16.24, 77.70]	ν_2	5.24	[5.04, 5.45]
descent")	$\alpha_2 \cdot 10^3$	−0.03	[−0.04, −0.01]			
3	q	27.22	[10.19, 44.26]	θ_3	0.10	[−0.18, 0.38]
("foraging")	η	135.43	[125.43, 145.44]	ν_3	8.76	[8.56, 8.95]
4,7	γ_0	4.59	[3.52, 5.66]	θ_4	−7.72	[−7.86, −7.57]
("deep/shallow	γ_1	−0.75	[−0.91, −0.58]	ν_4	4.24	[4.13, 4.35]
dive ascent")				θ_7	−3.73	[−3.91, −3.55]
				ν_7	3.73	[3.61, 3.86]
5	β_0	−1.48	[−1.71, −1.26]	θ_5	4.04	[3.85, 4.24]
("shallow dive	$\beta_1 \cdot 10^3$	−22.26	[−28.09, −16.44]	ν_5	3.52	[3.42, 3.63]
descent")	$\beta_2 \cdot 10^3$	0.07	[0.04, 0.09]			
6	ω_{66}	0.97	[0.96, 0.97]	θ_6	−0.13	[−0.16, −0.11]
("bottom of a				ν_6	1.04	[1.02, 1.06]
shallow dive")						

Figure 3. Top left: estimated probability of switching to "foraging", given "deep dive descent". Top right: estimated probability of switching to "shallow dive bottom", given "shallow dive descent". Bottom left: estimated probability of switching to "surfacing", given "deep dive ascent" (or "shallow dive ascent"). Bottom right: estimated conditional probability of beginning a deep dive cycle, given the animal switches away from "surfacing". Each of these probabilities is a function either of depth, z_t, or of the time since the last deep dive, r_t. The solid lines correspond to the maximum likelihood estimates, and the dashed lines indicate the 95 % pointwise confidence intervals.
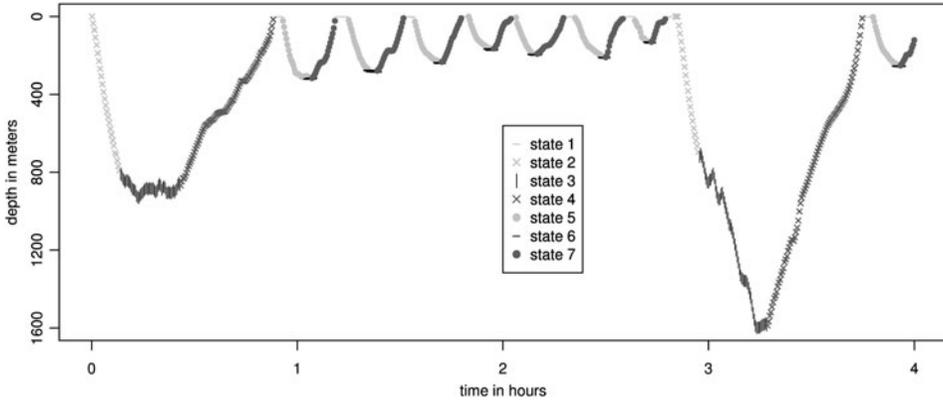
Figure 4. Sequence of states that under the fitted model is most likely to have given rise to the depths observed in the first four hours after deployment of the tag (found using the Viterbi algorithm).

corresponding parameter space, which is due to the application of the delta method, with the asymptotic distribution of the quantity of interest being approximated by a normal distribution. An alternative would be to derive the confidence intervals based on a large Monte Carlo sample from the Hessian-based approximate multivariate normal distribution of the (constrained) estimators. Notably, the probability of switching from state 5 to state 6, $\omega_{56}(z_t)$, is slightly higher for small depths ($z_t < 50$) than for moderate depths ($100 < z_t < 200$). This corresponds well to the observed diving pattern, where most of the shallow dives are to depths less than ∼50 meters, but those that do go further down than this are to depths up to ∼350 meters.

For the fitted model, we applied the Viterbi algorithm to find the sequence of states that is most likely to have given rise to the observed sequence of depths. Some example results of applying Viterbi are illustrated in Figure 4, which shows the decoded states only for the first four hours of observation in order to facilitate the visualization. Based on visual inspection, the assignments of the observations to the seven states of the model seem plausible. However, the second deep dive depicted in Figure 4 shows that with the given model specification, where states 2, 3, and 4 are visited in this order only, there might still be notable descents and ascents that are allocated to state 3 ("foraging").

Simulating from the fitted model provides a means of (informal) model checking, since it allows us to visually check whether or not the model captures the main features of the whale's diving pattern. Figure 5 displays a sequence of depths that has been simulated from the fitted model. In addition, Figure 6 gives the empirical and the model-derived depth distributions, with the latter obtained through Monte Carlo simulation (using 100 simulated series, each of the same length as the observed series).

The fitted model captures the movement pattern fairly well, but also shows some mismatches, which depending on the question of interest may or may not be irrelevant. For example, the observed dives are "more regular" than implied by the fitted model, with the beaked whale often apparently repeatedly diving to the same depth. Furthermore, the usage of the water column shows some mismatch at depths around 800 meters, and at depths between 50 and 200 meters. Finally, at the given time scale with observations every 10
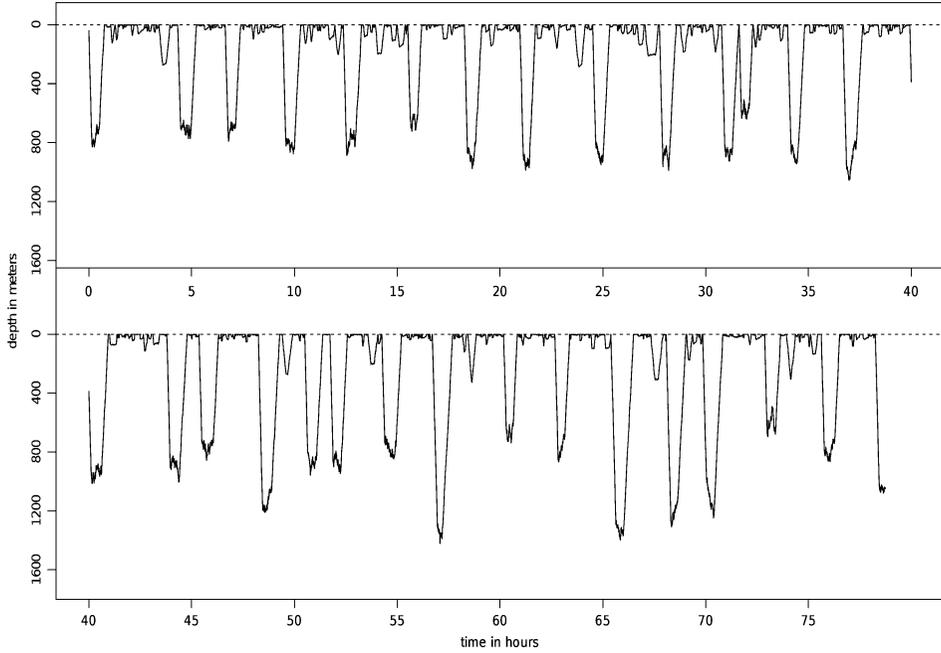
Figure 5. Depth profile simulated from the fitted model.
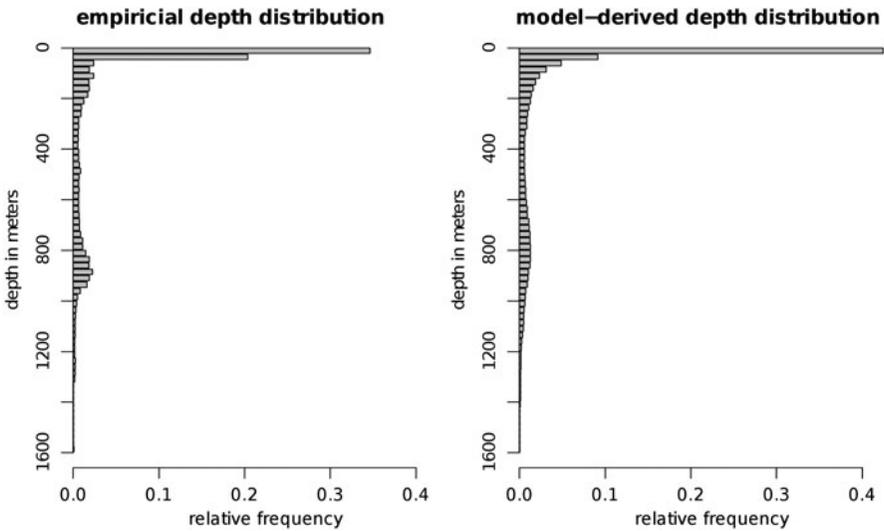
Figure 6. Empirical and model-derived depth distribution (i.e., usage of the water column).

seconds, the model does not fully capture the fine-scale correlation of movement. This is manifested in the estimated standard deviations of the state-dependent gamma distributions (see Table 1). For example, according to the fitted model, an animal that is in state 3 and makes a depth displacement of, say, 10 meters in the current 10-second interval, could well make a depth displacement of, say, -10 meters in the subsequent 10-second

interval, which is physically unrealistic. This problem arises because we essentially assume a biased random walk within each state, with directional bias determined by the drift parameter. While this approach does imply that successive vertical movement directions are correlated, the correlation is modeled only indirectly via the state process. Biased and correlated random walks seem more realistic in this context, as they involve a trade-off between directional bias and directional persistence (i.e., the tendency to maintain the previous movement direction), and thus comprise a component that explicitly addresses the fine-scale correlation. We tried to fit corresponding models to the beaked whale data, but encountered immense difficulties during optimization with local likelihood maxima, which we were not able to resolve in a satisfactory manner. This seems to be an identifiability issue, likely stemming from the fact that the distinction between correlated random walks and biased and correlated random walks is notoriously difficult (see, e.g., Codling, Plank, and Benhamou 2008). In any case, the extent of this issue is related to the time scale considered: the finer the resolution of the observations is in time, the higher the correlation of successive displacements will be. From a modeling perspective, we expect that it will often be easier to model data with a relatively coarse resolution in time, because it seems difficult to adequately model the fine-scale correlation. However, the resolution should be chosen such that it does not hinder addressing the (biological) question of interest.

5. DISCUSSION

The approach presented here allowed us to model the movement of a beaked whale for a period of over 3 days. We were able to recover patterns of vertical speeds obtained by others (Tyack et al. 2006) who used a manual state assignment process. In particular, the Viterbi-based assignment of (behavioral) states to the observations was plausible, such that we believe our approach to be a useful tool for automated allocation of segments of dive time series to behavioral states. The fitted model also enabled us to simulate realistic dive paths (at least if regarded at a more coarse time scale than the one considered), which can be useful in scenario planning, for example as part of mitigation strategies (see, e.g., van der Hoop, Vanderlaan, and Taggart 2012).

We note that the modeling should always be implemented keeping in mind the specific objectives being considered, as well as the spatial and temporal scale at which inferences are to be made. As an example, if the aim is to estimate availability bias in population assessment surveys, the model complexity might be reduced to focus on the states which have a direct impact on when the animal is available to be detected. In the case of visual surveys (e.g., visual line transects), state 1 ("surfacing") would be the most important to be described accurately. On the other hand, if passive acoustic methods were used, all but the persistence in state 3 ("foraging") might be irrelevant. For a given biological problem, it may also be useful to replace the visual assessment of fit that we applied here by more quantitative criteria, tailored to the problem at hand.

As with many other movement modeling techniques, care needs to be taken when choosing the temporal scale of analysis (Yackulic et al. 2011; Harris and Blackwell 2013). Any interpretation of a discrete-time model of the suggested type must be made relative to the

considered resolution in time, as the model only allows behavioral switches to occur at the time scale the model operates on. For a temporal resolution as fine as in our application, with only 10 s between successive observations, this issue is unlikely to be problematic in terms of the interpretations being made. We tested our approach also on a slightly coarser temporal scale, using the beaked whale data set but considering depth observations only every 30 s (results not shown). In this additional modeling exercise our approach identified the same movement pattern, but with parameter estimates rescaled as expected due to the different resolution in time: e.g., the expected length of a stay in the semi-Markovian state 3 ("foraging") changed from 135 time units ($\cong$1350 s) to 43 time units ($\cong$1290 s). However, many of the existing data sets on animal movement involve a very much coarser temporal scale, e.g., with observations made every two hours. Such coarse resolutions substantially limit the chances of drawing meaningful conclusions on the behavioral state process, and require different modeling strategies than those discussed here.

Modeling at a very fine temporal scale clearly carries a computational cost. However, applying the forward algorithm to evaluate the likelihood renders the computational effort linear in the number of observations and makes numerical likelihood maximization feasible even for very high numbers of observations. In the application to the beaked whale depth series, with observations every 10 s, leading to a total of 28,342 observations, we were still able to fit highly complex latent-variable models within a reasonable time. Fitting the model described in Section 4 but without the semi-Markovian component took about 3 hours on an i7 CPU, at 2.7 GHz and with 4 GB RAM, and fitting the same model including the semi-Markovian component took about 21 hours. In the latter case the computing time could be reduced to about 7 hours by using the parameter estimates obtained for the simpler Markovian model as initial values in the numerical maximization.

While being illustrated on a single animal's dive profile, the procedure also represents a flexible approach for modeling depth profiles of other diving species. As a possible approach to model development, one could start by specifying how many states seem to be required to describe a given species' movement. Then, given biological constraints, transition probabilities between some of these might be set to zero, and others would be modeled incorporating feedback mechanisms. For states in which a non-geometric dwell time is suspected, a semi-Markovian component could be added. Our approach might also be used under a different context to model other movement/displacement processes, namely those with a clear boundary. For example, the sea surface for a deep-diving species represents a similar barrier to movement as the earth's surface for a bird of prey. The methods described comprise a set of building blocks providing a versatile modeling approach. In order to (at least partly) investigate the general applicability of the approach, we applied this modeling strategy to a second time series, corresponding to the movement of a second Blainville's beaked whale (results not shown). The model fitted to this second data set revealed features very similar to those discussed in Section 4.2. According to the fitted models, the major differences between the two data sets were that the second whale performed less extreme deep dives, in terms of the maximum depths reached, and that after a deep dive the animal on average needed much less time to rest at the surface before starting the next deep dive.

There are a number of possible extensions of the considered approach. First, there are different ways of extending the framework in order to simultaneously model multi-

ple movement paths, accounting for possible heterogeneity. One can, e.g., consider hierarchical formulations of the models, which usually are relatively straightforward to implement yet accompanied by large scale increases in computational time (Altman 2007; Schliehe-Diecks, Kappeler, and Langrock 2012). A computationally less demanding way of accounting for heterogeneity across individuals is to incorporate individual-specific covariates in the model (if available). Second, the considered class of models can be extended in order to allow for modeling of movement in three-dimensional space. The easiest and perhaps most feasible strategy for accomplishing this task is to assume contemporaneous conditional independence between vertical and horizontal movement, given the behavioral state, thus modeling the horizontal and vertical movement separately in the observation process (cf. Chapter 8 in Zucchini and MacDonald 2009, for a discussion of contemporaneous conditional independence). An alternative is to consider displacement vectors, rather than actual positions, then modeling step lengths using some distribution on the non-negative real numbers, and directions using a suitable three-dimensional directional distribution. This would be a natural extension of the existing work on hidden Markov-type modeling of horizontal movement (e.g., Morales et al. 2004; Langrock 2012, but is technically more challenging, particularly due to the bounding nature of the sea surface. Third, the suggested approach could be extended in a relatively straightforward way to include time-varying parameters, allowing for diel variation or response to changing environmental conditions. Fourth, the ocean floor also is a natural barrier to vertical movement and, while here we did not take this into account explicitly, this could for example be accomplished by modeling the probability of transitioning from the "deep dive descent" state to the "foraging" state as a function of the distance to the ocean floor, if that information is available.

Finally, it would be worthwhile to explore the practicality of a Bayesian inference approach. In particular, it may be useful to include prior information on the parameters, which could aid in overcoming numerical problems (e.g., those we encountered when trying to incorporate biased and correlated random walks; see remarks at the end of Section 4.2). A Bayesian approach would also provide a more efficient and natural way for dealing with uncertainty in the parameter estimates, avoiding the problems we encountered with regard to confidence intervals where the boundaries occasionally lie outside the possible range of values (a consequence of the application of the delta method). Posterior predictive checks would also enable the uncertainty to be acknowledged within model checking, whereas we based our assessments essentially only on the point estimates. Lastly, a Bayesian approach would readily extend to hierarchical modeling of data from multiple individuals. On the other hand, a Bayesian analysis can prove prohibitively computer intensive, at least for the very large samples of data often associated with dive record data.

In conclusion, the methods presented and implemented here represent an important step in creating biologically realistic models for the vertical component of animal movement. We anticipate that these methods, together with extensions like those outlined above, will have broad utility and stimulate further research in this area.

ACKNOWLEDGEMENTS

REFERENCES

Altman, R. (2007), "Mixed Hidden Markov Models: an Extension of the Hidden Markov Model to the Longitudinal Data Setting," *Journal of the American Statistical Association*, 102, 201–210.

Bailleul, F., Pinaud, D., Hindell, M., Charrassin, J.-B., and Guinet, C. (2008), "Assessment of Scale-Dependent Foraging Behaviour in Southern Elephant Seals Incorporating the Vertical Dimension: a Development of the First Passage Time Method," *Journal of Animal Ecology*, 77, 948–957.

Baird, R. W., Webster, D. L., McSweeney, D. J., Ligon, A. D., Schorr, G. S., and Barlow, J. (2006), "Diving Behaviour of Cuvier's (*Ziphius Cavirostris*) and Blainville's (*Mesoplodon Densirostris*) Beaked Whales in Hawai'i," *Canadian Journal of Zoology*, 84, 1120–1128.

Baird, R. W., Webster, D. L., Schorr, G. S., Mcsweeney, D. J., and Barlow, J. (2008), "Diel Variation in Beaked Whale Diving Behavior," *Marine Mammal Science*, 24, 630–642.

Borchers, D. L., Zucchini, W., Heide-Jørgensen, M. P., Cañadas, A., and Langrock, R. (2013), "Using Hidden Markov Models to Deal with Availability Bias on Line Transect Surveys," *Biometrics*. doi:*10.1111/biom.12049*.

Codling, E. A., Plank, M. J., and Benhamou, S. (2008), "Random Walk Models in Biology," *Journal of the Royal Society Interface*, 5, 813–834.

Cox, T. M., et al. (2006), "Understanding the Impacts of Anthropogenic Sounds on Beaked Whales," *Journal of Cetacean Research and Management*, 7, 177–187.

D'Amico, A., Gisiner, R. C., Ketten, D. R., Hammock, J. A., Johnson, C., Tyack, P. L., and Mead, J. (2009), "Beaked Whale Strandings and Naval Exercises," *Aquatic Mammals*, 35, 452–472.

Dowd, M., and Joy, R. (2011), "Estimating Behavioral Parameters in Animal Movement Models Using a State-Augmented Particle Filter," *Ecology*, 92, 568–575.

Guédon, Y. (2003), "Estimating Hidden Semi-Markov Chains from Discrete Sequences," *Journal of Computational and Graphical Statistics*, 12, 604–639.

Harris, K. J., and Blackwell, P. G. (2013), "Flexible Continuous-Time Modelling for Heterogeneous Animal Movement," *Ecological Modelling*, 255, 29–37.

Higgs, M. D., and Ver Hoef, J. M. (2012), "Discretized and Aggregated: Modeling Dive Depth of Harbor Seals from Ordered Categorical Data with Temporal Autocorrelation," *Biometrics*, 68, 965–974.

Holzmann, H., Munk, A., Suster, M., and Zucchini, W. (2006), "Hidden Markov Models for Circular and Linear-Circular Time Series," *Environmental and Ecological Statistics*, 13, 325–347.

Hooker, S. K., Baird, R. W., and Fahlman, A. (2009), "Could Beaked Whales Get the Bends?: Effect of Diving Behaviour and Physiology on Modelled Gas Exchange for Three Species: *Ziphius Cavirostris*, *Mesoplodon Densirostris* and *Hyperoodon Ampullatus*," *Respiratory Physiology & Neurobiology*, 167, 235–246.

Houser, D. S. (2006), "A Method for Modeling Marine Mammal Movement and Behavior for Environmental Impact Assessment," *IEEE Journal of Oceanic Engineering*, 31, 76–81.

Langrock, R. (2012), "Flexible Latent-State Modelling of Old Faithful's Eruption Inter-Arrival Times in 2009," *Australian & New Zealand Journal of Statistics*, 54, 261–279.

Langrock, R., and Zucchini, W. (2011), "Hidden Markov Models with Arbitrary Dwell-Time Distributions," *Computational Statistics & Data Analysis*, 55, 715–724.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., and Morales, J. M. (2012), "Flexible and Practical Modeling of Animal Telemetry Data: Hidden Markov Models and Extensions," *Ecology*, 93, 2336–2342.

Langrock, R., Borchers, D. L., and Skaug, H. J. (2013), "Markov-Modulated Nonhomogeneous Poisson Processes for Modeling Detections in Surveys of Marine Mammal Abundance," *Journal of the American Statistical Association*. doi:*10.1080/01621459.2013.797356*.

McCarthy, E., Moretti, D., Thomas, L., DiMarzio, N., Morrissey, R., Jarvis, S., Ward, J., Izzi, A., and Dilley, A. (2011), "Changes in Spatial and Temporal Distribution and Vocal Behavior of Blainville's Beaked Whales (*Mesoplodon Densirostris*) During Multiship Exercises with Mid-Frequency Sonar," *Marine Mammal Science*, 27, E206–E226.

McClintock, B. T., Russell, D. J. F., Matthiopoulos, J., and King, R. (2013), "Combining Individual Animal Movement and Ancillary Biotelemetry Data to Investigate Population-Level Activity Budgets," *Ecology*, 94, 838–849.

Morales, J. M., Haydon, D. T., Frair, J. L., Holsinger, K. E., and Fryxell, J. M. (2004), "Extracting More Out of Relocation Data: Building Movement Models as Mixtures of Random Walks," *Ecology*, 85, 2436–2445.

Patterson, T. A., Basson, M., Bravington, M. V., and Gunn, J. S. (2009), "Classifying Movement Behaviour in Relation to Environmental Conditions Using Hidden Markov Models," *Journal of Animal Ecology*, 78, 1113–1123.

R Core Team (2012), *R: a Language and Environment for Statistical Computing*, Vienna: R Foundation for Statistical Computing. URL: *http://www.R-project.org/*.

Ramsay, J. O. (1988), "Monotone Regression Splines in Action," *Statistical Science*, 3, 425–461.

Schick, R. S., Loarie, S. R., Colchero, F., Best, B. D., Boustany, A., Conde, D. A., Halpin, P. N., Joppa, L. N., McClellan, C. M., and Clark, J. S. (2008), "Understanding Movement Data and Movement Processes: Current and Emerging Directions," *Ecology Letters*, 11, 1338–1350.

Schliehe-Diecks, S., Kappeler, P. M., and Langrock, R. (2012), "On the Application of Mixed Hidden Markov Models to Multiple Behavioural Time Series," *Interface Focus*, 2, 180–189.

Tyack, P. L., Johnson, M., Aguilar de Soto, N., Sturlese, A., and Madsen, P. T. (2006), "Extreme Diving of Beaked Whales," *Journal of Experimental Biology*, 21, 4238–4253.

Tyack, P. L., Zimmer, W. M. X., Moretti, D., Southall, B. L., Claridge, D. E., Durban, J. W., Clark, C. W., D'Amico, A., DiMarzio, N., Jarvis, S., McCarthy, E., Morrissey, R., Ward, J., and Boyd, I. (2011), "Beaked Whales Respond to Simulated and Actual Navy Sonar," *PLoS ONE*, 6, e17009.

van der Hoop, J. M., Vanderlaan, A. S. M., and Taggart, C. T. (2012), "Absolute Probability Estimates of Lethal Vessel Strikes to North Atlantic Right Whales in Roseway Basin, Scotian Shelf," *Ecological Applications*, 22, 2021–2033.

Walker, C. G., MacKenzie, M. L., Donovan, C. R., Kidney, D., Quick, N. J., and Hastie, G. D. (2011), "Classification of Animal Dive Tracks Via Automatic Landmarking, Principal Components Analysis and Clustering," *Ecosphere*, 2, 92.

Yackulic, C. B., Blake, S., Deem, S., Kock, M., and Uriarte, M. (2011), "One Size Does Not Fit All: Flexible Models Are Required to Understand Animal Movement Across Scales," *Journal of Animal Ecology*, 80, 1088–1096.

Zucchini, W., and MacDonald, I. L. (2009), *Hidden Markov Models for Time Series: an Introduction Using R*, London: Chapman & Hall.

Zucchini, W., Raubenheimer, D., and MacDonald, I. L. (2008), "Modeling Time Series of Animal Behavior by Means of a Latent-State Model with Feedback," *Biometrics*, 64, 807–815.