PLOS | **SUBMISSION**

Understanding narwhal diving behaviour using Hidden Markov Models with dependent state distributions and long range dependence

Manh Cuong Ngô[1], Mads Peter Heide-Jørgensen[1], Susanne Ditlevsen[2*],

1 Greenland Institute of Natural Resources, c/o Greenland Representation, Strandgade 91,2, 1401 Copenhagen K, Denmark
2 Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

* susanne@math.ku.dk

Abstract

Diving behaviour of narwhals is still largely unknown. We build three-state Hidden Markov models (HMM) to describe the diving behaviour of a narwhal and fit the models to a three-dimensional response vector of maximum dive depth, duration of dives and post-dive surface time of 8,609 dives measured in East Greenland over 83 days, an extraordinarily long and rich data set. In particular, diurnal patterns in diving behaviour for a marine mammal is being inferred, by using periodic B-splines with boundary knots in 0 and 24 hours. Several HMMs with covariates are used to characterize dive patterns. Narwhal diving patterns have not been analysed like this before, but in studies of other whale species, response variables have been assumed independent. We extend the existing models to allow for dependence between state distributions, and show that the dependence has an impact on the conclusions drawn about the diving behaviour. It is thus paramount to relax this strong and biologically unrealistic assumption to obtain trustworthy inferences.

Author summary

Narwhals live in pristine environments. However, the increase in average temperatures in the Arctic and the concomitant loss of summer sea ice, as well as increased human activities, such as ship traffic and mineral exploration leading to increased noise pollution, are changing the environment, and therefore probably also the behavior and well-being of the narwhal. Here, we use probabilistic models to unravel the diving and feeding behavior of a male narwhal, tagged in East Greenland in 2013, and followed for nearly two months. The goal is to gain knowledge of the whales' normal behavior, to be able to later detect possible changes in behavior due to climatic changes and human influences. We find that the narwhal uses around two thirds of its time searching for food, it typically feeds during deep dives (more than $350\,m$), and it can have extended periods, up to 3 days, without feeding activity.

1   Introduction

The narwhal (*Monodon monoceros*) is a long-lived toothed whale that primarily inhabit cold waters of the Atlantic sector of the Arctic. The largest abundances are found in

East and West Greenland and in the Canadian High Arctic. Physical maturity is reached at an age of ~ 30 years and at a body length and mass of $4-5\,m$ and 1,500-1,800 kg [8]. The narwhal belong to the group of the deepest diving cetaceans with the maximum exceeding $1,800\,m$ [10]. Their diet consists of few prey species including Greenland halibut (*Reinhardtius hippoglossoides*), polar cod (*Boreogadus saida*), capelin (*Ammodytes villosus*) and squids (*Gonatus sp.*) [9,13]. Predation on Greenland halibut constitutes a potential competition with fisheries, especially in winter, at the same time as narwhal skin and meat as well as their valuable tusks for centuries have been important hunting products for Inuit cultures in Greenland and northern Canada. Anthropogenic factors like underwater noise are also a concern for a species that, with decreasing sea ice coverage, is increasingly exposed to underwater noise from shipping and seismic exploration [21]. It is therefore important to understand and quantitatively describe the diving activities of narwhals, by robust statistical methods, and assess how these are changed due to potential conflicts with anthropogenic activities, and to ensure the long-term conservation of one of the most specialized species in the North Atlantic.

The first step is to understand the diving patterns of narwhals under natural conditions, which we address in this study. Diving behaviour is however cryptic since it includes both physiological constraints, energetic demands and habitat and environmental regimes. Modelling of the observed diving behaviour is one way of gaining insight to the overall diving patterns, and changes in model parameters is a way to compare and estimate quantitatively impacts from anthropogenic activities (see [4,5]).

In this study we apply multivariate Hidden Markov Models (HMMs) with covariates [22], to describe the diving dynamics in the vertical dimension of an individual narwhal. These types of models for similar diving data of Blainville's beaked whales (*Mesoplodon densirostris*) were first introduced in [15]. A HMM assumes an underlying unobserved process, which governs the dynamics of the observed variables. The assumption is that the observed behaviour in a dive will depend on the present state, and introduces autocorrelation in the model [22], in agreement with the observed data, where the narwhal tends to repeat the same type of dives in a series before switching to another type of dives. Similar to the blue whales (*Balaenoptera musculus*) data analysed in [5] and the short-finned pilot whales (*Globicephala macrorynchus*) data analysed in [19], our narwhal data suggest three distinct states: state 1 covers the shortest and shallowest dives which we interpret as near-surface travelling, state 2 covers the longer and deeper dives which we interpret as feeding states, and state 3 covers the deepest and longest dives which we interpret as intense deep feeding behaviours. Pohle et al [18] recommended against using more than four states in biological modelling like this, in order to avoid the complexity of the correspondence between states of the model and the biological phenomenon. DeRuiter et al [5] suggested three states for their data, even if a formal model selection procedure would point to a more complex model, because models with more underlying states might obscure patterns in the data and provide less insight in the underlying biological process, even if they might perform better in terms of forecasting. Biological knowledge should guide the choice of number of states. They also argue that model misspecifications, such as too inflexible state dependent distributions, variations over time, missing covariate information or outliers might cause model selection criteria to favour models with more complex structures than warranted. Therefore, we choose the three-states HMM model.

These HMMs have shown powerful for modelling animal movement by taking into account the correlation over time between different movement patterns, with an extensive research in two horizontal dimensions (see, e.g., [14,16,17]), and recently, in one vertical dimension [5,15], possibly including further information on vertical movements. In this study, we use vertical depth data, and the three response variables are the maximum depth reached in a dive, the duration of a dive, and the post-dive

⬦PLOS | **SUBMISSION**

surface time before initiating a new dive.   56

The data set covers 1,995 hours ($\sim$ 83 days) and is extraordinarily long, and thus   57
provides a unique opportunity to obtain detailed information on diving behaviour. An   58
example of the data is shown in Fig 1. Such data are usually only on the order of a   59
couple of days or less, for example, the time series of short-finned pilot whales analysed   60
in [19] cover up to 18 hours and 64 dives, whereas the time series of blue whales   61
analysed in [5] cover up to 6 hours and 67 dives, and [14] analyses 79 hours of a single   62
Blainville's beaked whale. However, here we only have data from a single narwhal   63
limiting the generalizability of the analysis.   64
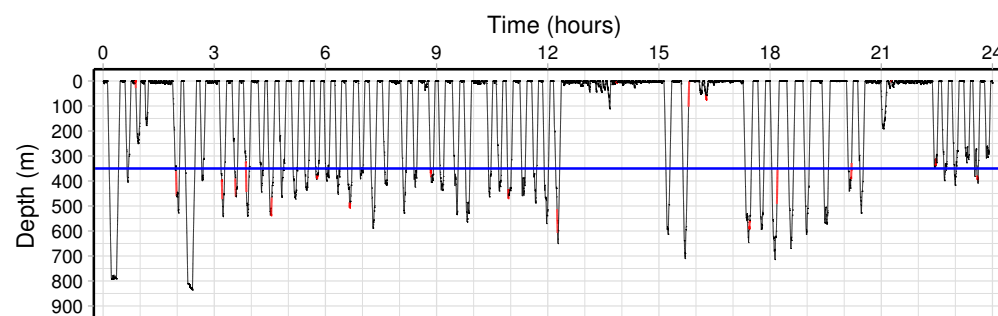

Fig 1. Diving data. Representative part of the narwhal diving data, covering 24
hours of dives on August 15th 2013. The red parts are where a lower temperature in the
stomach has been registered, indicating that the narwhal has swallowed a prey. The
blue line indicates a depth of $350m$, the threshold for a *deep dive* used in the definition
of the covariates.

In all previous studies, *contemporaneous conditional independence* was assumed,   65
meaning that the state dependent processes are independent given the underlying state.   66
This is a strong and biologically unrealistic assumption, since deeper dives will typically   67
take longer, also when conditioning the dive to be either shallow, medium or deep. A   68
positive correlation is still expected, beyond the correlation implied by the hidden states.   69
DeRuiter et al [5] argued for the assumption of conditional independence because unless   70
a multivariate normal distribution can be assumed, there is usually no simple candidate   71
multivariate distribution to specify the correlation structure. This is partly due to some   72
of their response variables being discrete. In this study, we will relax the assumption of   73
conditional independence, taking advantage of the continuity of the response variables.   74
They are all restricted to be positive and with right skewed distributions. Previous   75
studies have therefore used conditionally independent gamma distributions for these   76
variables. Here, we will assume dependent log-normal distributions, such that their   77
log-transforms follow a multivariate normal distribution. We only assume dependence   78
between maximum depth and dive duration, which has a high correlation in the full   79
data set (0.864), as well as high correlations within the three estimated states for the   80
chosen model (0.538, 0.777 and 0.458, respectively), whereas we assume both of these   81
variables independent of post-dive surface time, since the correlations in the full data set   82
were low (0.046 and 0.042, respectively). We also do the analysis with the standard   83
choice of the gamma distributions, and compare the results.   84

Covariates were included in [5, 15, 16], appearing in the transition probabilities   85
between hidden states, whereas no covariates were included in [19]. Here we include   86
covariates in all elements of the transition matrix, trying out different covariate process   87
models and select the optimal model by the Akaike Information Criterion (AIC). We   88
consider two covariates related to the recent deep dives performed by the narwhal.   89

Dives can reach $> 1,800\,m$, and deeper dives are usually related to intense feeding [10]. We define a *deep dive* as a dive to a depth of at least $350\,m$. Note that this definition is only used to define the covariates, and is not related to the decoding of states. One covariate is the time passed since the last deep dive, which was also used in [15]. The hypothesis is that the longer the time passed since last deep dive, the higher the narwhal's propensity for initiating a deep dive will be. Another covariate counts the number of consecutive deep dives that the narwhal has performed. The hypothesis is that the more dives in a row and more time spent at great depths, the higher the narwhal's propensity for changing diving pattern to shallower depth or near-surface travelling. By introducing such history dependent covariates, the model allows a longer dependence structure than the one implied by the Markov property. These models with dependencies between observables caused by the underlying state, as well as including feedback from the observed process, were introduced in [15] to model Blainville's beaked whale. The last covariate is time of day at initiation of the dive, modelled by a periodic B-spline with boundary knots in 0 and 24 hours. Diurnal effects on marine mammal diving patterns are difficult to estimate in this type of models because the time series are typically too short, but this data set is extraordinarily long, making this inference possible. However, here we only analyse a single whale, and results might not generalize.

2   Materials & Methods

We analyse the time series of depth measurements of a mature male narwhal (420 cm, estimated mass 950 kg) tagged in East Greenland from August 13th until November 6th 2013. Permission for capturing, handling, and tagging of narwhals was provided by the Government of Greenland (Case ID 2010–035453, document number 429 926). The tag (Mk10 time-depth recorder from Wildlife Computers, Redmond, WA, USA) was attached to the whale and retrieved one year later with 1994.83 hours of dive data (approximately 83 days and 2 hours), see [12]. In this time interval the narwhal performed 8,609 dives to depths of at least $20\,m$. Depth was measured every second at a resolution of $0.5m$, and summarized in three variables within each dive to describe the behaviour: dive duration, maximum depth and post-dive surface time, as also used in [5,19]. Here we define a dive as one in which the whale reach a depth of at least $20\,m$, otherwise it is considered time spent at the surface. This threshold was chosen in order to avoid creating too many shallow dives near the surface, see [1]. The surface and dive durations also enter in the model as part of the covariate counting the time since last deep dive. The total time the whale spent at the surface was 37 days and 5.6 hours, i.e., the narwhal spent 44.8% of its time at the surface.

The first week of tagging, the narwhal also had the temperature of the stomach measured, see [11]. A temperature drop indicates that a prey has entered the stomach. Fig 1 shows an example part of the data, and the red parts indicate temperature drops. These typically happen during deep dives, and support the assumption that deep dives are related to foraging.

In this study, the observed response variable, denoted by $\boldsymbol{X}_t$, is three-dimensional, describing the diving behaviour related to each dive, where t indicates the dive number, $t = 1, 2, \ldots, T$, with $T = 8,609$ dives. The first response variable, $X_{1,t}$, is the maximum depth reached in dive number t, and takes real positive values between 20 and $910.5\,m$. The second response variable, $X_{2,t}$, is the duration of dive number t, and takes real positive values between 33 seconds and 28 minutes. The third response variable, $X_{3,t}$, is the time spent at the surface after the dive before initiating a new dive, and takes real positive values between 1 second and 209.7 minutes. We assume that the diving behaviour depends on an underlying unobserved process, which we denote by $C_t, t = 1, 2, \ldots$, with a number m of unobserved behavioural states, $C_t \in \{1, \ldots, m\}$,

which govern the dynamics of the observed variables. The assumption is that the distributions of the observed maximum depth, duration of dive, and post-dive duration of dive number t depend on the state. We choose the three-states HMM model, $m = 3$, and identify the three states with three different diving patterns in the water column; near-surface travelling, medium depth dives, and deep dives probably related to feeding.

2.1 Hidden Markov Model

An m-dimensional hidden Markov model assumes that the distribution of the p-dimensional response vector $\boldsymbol{X}_t$ depends on a hidden state C_t, where $\{C_t : t = 1, 2, \dots\}$ is an unobserved underlying process satisfying the Markov property:

$$P(C_t = j \mid C_{t-1} = i, \dots, C_1 = l) = P(C_t = j \mid C_{t-1} = i),$$

where $C_t \in \{1, \dots, m\}$ for $t = 2, 3, \dots$. Denote the state transition probabilities for $m = 3$ at time t by $\omega_{ij}(t), i, j = 1, 2, 3$, where $\omega_{ij}(t) = P(C_{t+1} = j \mid C_t = i)$. The transition probability matrix $\Omega(t)$ is then

$$\Omega(t) = \begin{bmatrix} \omega_{11}(t) & \omega_{12}(t) & \omega_{13}(t) \\ \omega_{21}(t) & \omega_{22}(t) & \omega_{23}(t) \\ \omega_{31}(t) & \omega_{32}(t) & \omega_{33}(t) \end{bmatrix} \tag{1}$$

where $\omega_{ij}(t) \geq 0$ and $\sum_{j=1}^{3} \omega_{ij}(t) = 1$. Here, we let $\omega_{ij}(t)$ depend on t to allow time varying covariates to affect the transition probabilities, see Section 2.3. The distribution of $\boldsymbol{X}_t$ is conditionally independent of everything else given C_t:

$$f(\boldsymbol{X}_t | \boldsymbol{X}_{t-1}, \dots, \boldsymbol{X}_1, C_t, C_{t-1}, \dots, C_1) = f(\boldsymbol{X}_t | C_t), \, t = 1, 2, \dots \tag{2}$$

where f denotes a probability density function, i.e., the distribution of $\boldsymbol{X}_t$ depends only on the current state C_t and not on previous states or observations. The model is illustrated in Fig 2.
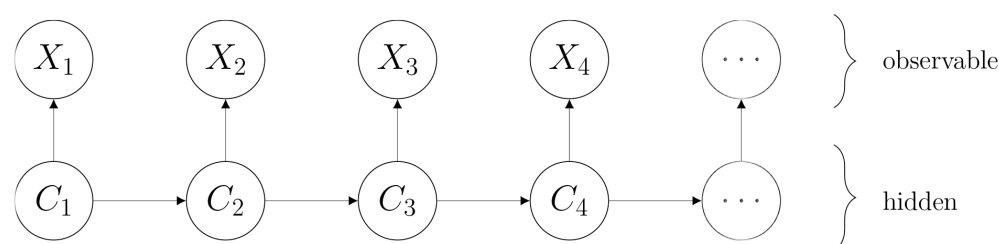

Fig 2. Hidden Markov Model. The hidden states C_t represent behavioural states that influence the distribution of the observed variables X_t.

2.2 State dependent distributions

The state-dependent distributions are the probability density functions of $\boldsymbol{X}_t$ associated with state i. Under the *contemporaneous conditional independence* assumption, the p different components of the response vector $\boldsymbol{X}_t$ are assumed independent given the hidden state, and the probability density can be decomposed as

$$f(\boldsymbol{X}_t \mid C_t = i) = f_i(\boldsymbol{X}_t) = \prod_{k=1}^{p} f_{i,k}(X_{k,t}), \tag{3}$$

where $X_{k,t}$ is the kth observed component of $\boldsymbol{X}_t$. Here we have $p = 3$, the components being maximum depth (MD), duration of dives (DT), and post-dive duration (PD). Thus, $\boldsymbol{X}_t = (X_{MD,t}, X_{DT,t}, X_{PD,t})^T$, where ^{T} denotes transposition. *Contemporaneous conditional independence* implies that the state dependent processes $X_{MD,t}, X_{DT,t}$ and $X_{PD,t}$ are independent given the underlying state C_t. This assumption has been used in [5] and [19] because in general, there is no simple way to address the correlation between variables within states, and the dependence induced by the Markov chain is often sufficient to fit the data. However, in this paper, we will relax this assumption, and let f_i be a joint distribution function, allowing for dependent coordinates, which for our data turn out to improve the fit considerably.

All three response variables are positive right-skewed variables, so natural candidates for $f_{i,k}$ are gamma distributions, as used in [5] and [19], or log-normal distributions, i.e., the logarithm of the response variables follow a 3-dimensional normal distribution. Here, we will try three different distributions. The first candidate is independent gamma distributions, to compare with the usual approach. The gamma distribution is parametrized by shape parameter μ and scale parameter σ, with mean $\mu\sigma$ and variance $\mu\sigma^2$, and the state dependent probability density functions are given by

$$f_i(\boldsymbol{X}_t) = \prod_{k \in \{MD,DT,PT\}} f_{i,k}(X_{k,t}) = \prod_{k \in \{MD,DT,PT\}} \Gamma(\mu_i^k)^{-1}(\sigma_i^k)^{-\mu_i^k} X_{k,t}^{\mu_i^k - 1} e^{-\frac{X_{k,t}}{\sigma_i^k}}, \quad (4)$$

for $i = 1, 2, 3$.

We will also assume both independent and correlated log-normal distributions, such that $\log \boldsymbol{X}_t$ is multivariate normal, where $\log \boldsymbol{X}_t = (\log X_{MD,t}, \log X_{DT,t}, \log X_{PT,t})^T$, taking advantage of the computational convenience of the normal distribution. The log-normal distribution is parametrized by log-mean μ and log-variance σ^2. Thus, given $C_t = i$ and k, the mean and variance of $\log \boldsymbol{X}_{k,t}$ is μ_i^k and $(\sigma_i^k)^2$, and the mean and variance of $\boldsymbol{X}_{k,t}$ is $\exp(\mu_i^k + (\sigma_i^k)^2/2)$ and $(\exp((\sigma_i^k)^2) - 1)\exp(2\mu_i^k + (\sigma_i^k)^2)$. The log-correlation is denoted by ρ_i. The correlation between the first two components is $(\exp(\rho_i\sigma_i^{MD}\sigma_i^{DT}) - 1)/\sqrt{(\exp((\sigma_i^{MD})^2) - 1)(\exp((\sigma_i^{DT})^2) - 1)}$, where $(\sigma_i^{MD})^2$ and $(\sigma_i^{DT})^2$ are the log-variances of maximum depth and dive duration, respectively. The correlation is approximately equal to the log-correlation ρ_i when $(\sigma_i^{MD})^2$ and $(\sigma_i^{DT})^2$ are small. We assume post-dive duration independent of the other two response variables, based on the low marginal correlation in the data. Thus, the state dependent probability density functions are given by

$$f_i(\boldsymbol{X}_t) = \frac{1}{(2\pi)^{3/2}\sqrt{|\Sigma_i|}} e^{-\frac{1}{2}(\log \boldsymbol{X}_t - \mu_i)^T \Sigma_i^{-1}(\log \boldsymbol{X}_t - \mu_i)}, \quad (5)$$

where $|\cdot|$ denotes the determinant of a matrix, $\mu_i = (\mu_i^{MD}, \mu_i^{DT}, \mu_i^{PD})^T$,

$$\Sigma_i = \begin{bmatrix} (\sigma_i^{MD})^2 & \rho\sigma_i^{MD}\sigma_i^{DT} & 0 \\ \rho\sigma_i^{MD}\sigma_i^{DT} & (\sigma_i^{DT})^2 & 0 \\ 0 & 0 & (\sigma_i^{PD})^2 \end{bmatrix}$$

and $\rho = 0$ in the independent case.

2.3 Covariates

To allow for a longer memory in the model beyond the autocorrelation induced by the hidden process, we incorporate feedback mechanisms by letting the state transition probabilities depend on the history. We consider two covariates related to the recent deep dives performed by the narwhal. One covariate is the continuous variable τ_t, defined as time passed since the last deep dive before dive number t, where a *deep dive*

is defined as a dive to a depth of at least $350\,m$. Note that this definition is only used to define the covariates, and is not related to the decoding of states. Thus, the algorithm might decode dives of lesser depth to state 3, and deeper dives not to be in state 3. The value chosen is somewhat arbitrary, and we tried different values between 250 and $450\,m$, without much effect on the results. The other covariate is the discrete variable d_t taking non-negative integer values, counting the number of consecutive deep dives that the narwhal has performed before dive number t. Finally, we consider the covariate of the hour at which the dive is initiated. More specifically, we define the covariate processes $\mathcal{T}_t$, the time since the last deep dive, D_t, the number of consecutive deep dives up to dive number t, and H_t, the hour of initiation of dive t, and denote the measured covariates by τ_t, d_t and h_t. Thus, the short term memory is modelled by the hidden states, and the long term memory is modelled by modulation of the transition probabilities as a function of past dynamics. The model is illustrated in Fig 3, and Fig 4 illustrates the response variables and the three covariates for 60 consecutive dives.
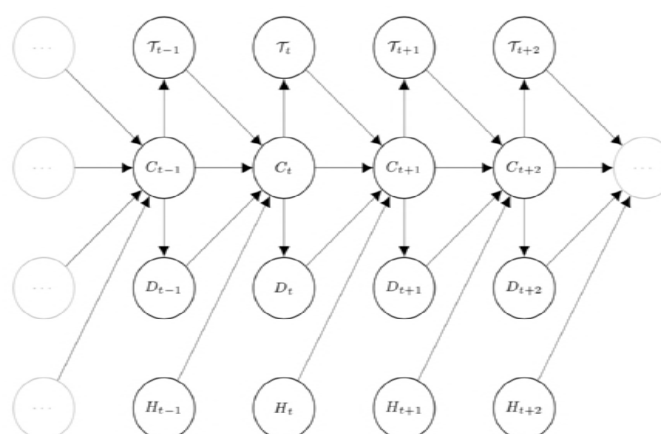

Fig 3. Hidden Markov Model with feedback processes. The distribution of the transitions between hidden states C_t depends on the observed covariate processes $\mathcal{T}_t$, D_t and H_t.

The covariates enter the transition probabilities $\omega_{ij}(t) = \omega_{ij}(\eta_{ij}(t))$ in Eq. (1) through a *predictor*, $\eta_{ij}(t)$, see Eq. (7) below. We consider several models. If there are no covariates for a given predictor, then $\eta_{ij}(t) = \eta_{ij}$ does not depend on t. Table 1 lists the different models, where $\alpha_{ij}, \beta_{ij}, \gamma_{ij}, \delta_{ij}, \theta_{ij}$ and ζ_{ij} are real parameters. We tried more models, but only include these for illustration. Covariates d_t and τ_t were incorporated as natural cubic splines with three degrees of freedom. The covariate d_t counts number of deep dives in a row, and is therefore around 0 when not in state 3. This covariate therefore carries no information unless in state 3, and only enters in η_{31} and η_{32}. Likewise, τ_t is expected to be around 0 when in state 3, and therefore only enters η_{ij} for $i = 1$ or 2. The effect of time of day is modelled by a periodic B-spline with three degrees of freedom, with boundary knots in 0 and 24 hours.
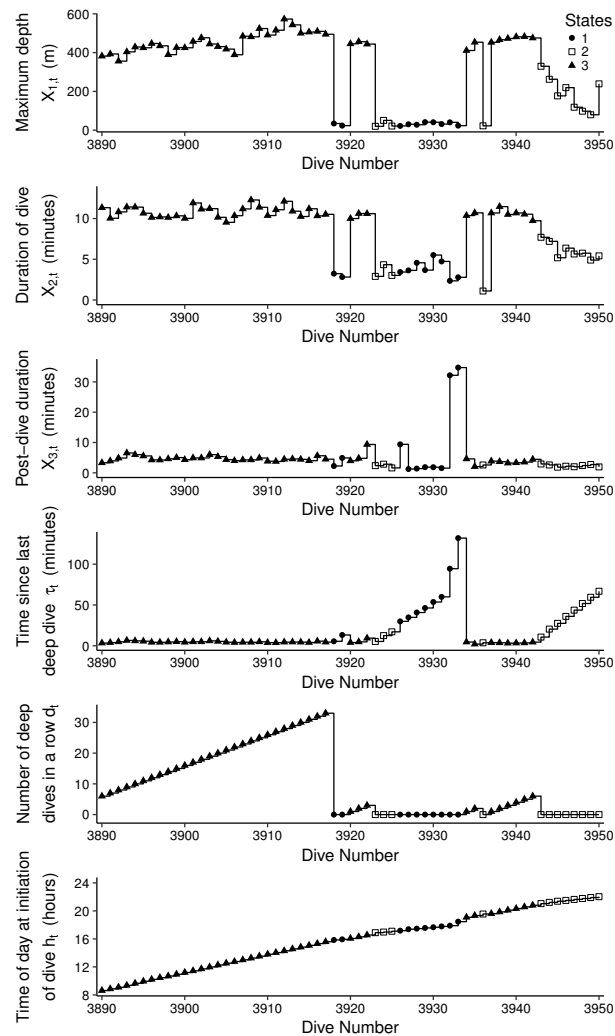
Fig 4. Response variables and covariate processes. Time series plot of maximum depth, duration of dive, and post-dive duration from dive number 3890 to 3950 and the covariate processes counting the time since last deep dive, number of deep dives in a row, and the hour at initiation of dive. The symbols indicate the decoded hidden states from a model fitted to a dependent log-normal distribution (Model 1). State 1 (circles) represents near-surface travelling, State 2 (squares) represents a feeding state of medium long and deep dives, and State 3 (triangles) represents an intense deep feeding state of the deepest and longest dives.

2.4 The likelihood function and optimization

The likelihood L_T of $x_1, x_2, \ldots, x_T$, where x_t is the observation of $\boldsymbol{X}_t$, assumed to be generated by an m-state HMM, can in general be computed recursively in only $O(Tm^2)$ operations by the forward algorithm [22]. The likelihood is expressed as

$$L_T = \delta \mathcal{P}(x_1)\Omega(\tau_1, d_1, h_1)\mathcal{P}(x_2)\cdots\Omega(\tau_{T-1}, d_{T-1}, h_{T-1})\mathcal{P}(x_T)\mathbf{1}, \qquad (6)$$

where $\mathcal{P}(x_t) = \text{diag}(f_1(x_t), f_2(x_t), f_3(x_t))$ is a diagonal matrix with diagonal elements $f_i(x_t)$ given in Eq. (4) when the gamma distribution is used, or Eq. (5) when the log-normal distribution is used, Ω is given by Eq. (1) and $\mathbf{1} \in \mathbb{R}^m$ is a column vector of

Table 1. Different models for covariate effects on the transition probabilities between behavioural states.

Model	$\eta_{12}(t)$	$\eta_{13}(t)$	$\eta_{21}(t)$	$\eta_{23}(t)$	$\eta_{31}(t)$	$\eta_{32}(t)$
			Predictors in the transition probabilities			
1	$\alpha_{00} + T_{12}^t + H_{12}^t$	$\alpha_{01} + T_{13}^t + H_{13}^t$	$\beta_{00} + T_{21}^t + H_{21}^t$	$\beta_{01} + T_{23}^t + H_{23}^t$	$\gamma_{00} + D_{31}^t + H_{31}^t$	$\gamma_{01} + D_{32}^t + H_{32}^t$
2	$\alpha_{00} + H_{12}^t$	$\alpha_{01} + H_{13}^t$	$\beta_{00} + H_{21}^t$	$\beta_{01} + H_{23}^t$	$\gamma_{00} + H_{31}^t$	$\gamma_{01} + H_{32}^t$
3	$\alpha_{00} + T_{12}^t$	$\alpha_{01} + T_{13}^t$	$\beta_{00} + T_{21}^t$	$\beta_{01} + T_{23}^t$	$\gamma_{00} + D_{31}^t + H_{31}^t$	$\gamma_{01} + D_{32}^t + H_{32}^t$
4	$\alpha_{00} + T_{12}^t$	$\alpha_{01} + T_{13}^t$	$\beta_{00} + T_{21}^t$	$\beta_{01} + T_{23}^t$	$\gamma_{00} + D_{31}^t$	$\gamma_{01} + D_{32}^t$

The predictors η_{ij} relate to the transition probabilities as given in equation (7). Denote the spline effects of hour by $H_{ij}^t = \sum_k \delta_{ij}^{(k)} h_k^t$, of τ_t by $T_{ij}^t = \sum_k \theta_{ij}^{(k)} s_k^t$, and of d_t by $D_{ij}^t = \sum_k \zeta_{ij}^{(k)} d_k^t$ for $k = 1, 2, 3$ and $i, j = 1, 2, 3; i \neq j$.

ones. The initial state distribution is denoted by δ, which is an m-dimensional row vector; $\delta_i = P(C_1 = i)$. For δ, we choose the uniform distribution, $\delta_i = 1/m$. Alternatively, it can be estimated, but there is no need for this extra computational effort, since our dataset is large and the influence of δ will be negligible. Furthermore, δ has no particular biological relevance.

The transition parameters in Eq. (1) are constrained to be between 0 and 1 with row sums equal to 1, and thus, even if there are $3^2 = 9$ entries, there are only $3 \cdot 2 = 6$ free parameters. To obtain an unconstrained optimization problem, we reparametrise to working parameters, as also done in [5, 16, 19], see also [22], by defining

$$\omega_{ij}(t) = \frac{\exp(\eta_{ij}(t))}{\sum_{j=1}^m \exp(\eta_{ij}(t))} \tag{7}$$

where $\eta_{ij}(t)$ is the predictor for dive t for $1 \leq i, j \leq 3, i \neq j$, and $\eta_{ii} = 0$ for $i = 1, 2, 3$. This assures positive entries and that rows sum to 1.

We used the direct numerical Newton-Raphson algorithm `nlm` (`optim` in case `nlm` failed) in R [20] to estimate the parameters of the model by maximizing the log-likelihood, $\mathcal{L}_T := \log L_T$, where L_T is given in Eq. (6). The procedure `ns` from the package `splines` (version 3.5.0) was used to calculate the periodic splines. The procedure `pbs` from the package `pbs` (version 1.1) was used to calculate the periodic splines.

Using a combination of R and Rcpp [6] for calculating the log-likelihood function $\mathcal{L}_T$ improved the runtime considerably. To mitigate the problem of local maxima, we ran the optimization algorithm thousands times with different starting values for the parameters of the state-dependent distributions, as well as for the parameters of the transition matrix. The final result was chosen as the one giving the maximum log-likelihood. For the best fits, we employed the jittering procedure used in [5], but it did not improve the log-likelihood value.

Once the optimal model was selected and parameters of the model were estimated, it was of interest to decode the most likely state sequence $c_1^*, \ldots, c_T^*$. The Viterbi algorithm [7, 22] is suitable to estimate the hidden states given the observed depths and durations:

$$(c_1^*, \ldots, c_T^*) = \underset{(c_1, \ldots, c_T) \in \{1,2,3\}}{\text{argmax}} \Pr(C_1 = c_1, \ldots, C_T = c_T \mid x_0, \ldots, x_T).$$

We evaluate by AIC which model provides the best fit, and study if final conclusions on diving behaviour differ if dependence is assumed.

PLOS | **SUBMISSION**

3   Results

The estimation algorithm classifies each dive to one of the three hidden states. The classification depends on the model, but all models roughly group dives according to maximum depth. The algorithm allocates labels arbitrarily, so to compare across models we relabeled the states, such that State 1 represents the shortest and shallowest dives, which we define as near-surface travelling, State 2 represents medium long and deep dives, which we define as a feeding state, and State 3 represents the deepest and longest dives, which we define as an intense deep feeding state. Thus, one goal of comparing models is to access if conclusions on diving behaviour expressed through the decoded classes of the dives differ between models. If they all classify the same, it does not matter which model we use, maybe except for the estimation of covariate effects. If the classification differ from model to model, it is important to choose the statistically best model.

Table 2 lists the model selection results from the optimization. We use AIC to select the best model, which is highlighted in bold. The correlated log-normal model is clearly preferred above the independent models, with huge AIC differences, and the log-normal distribution is clearly preferred above the gamma distribution. Models with ΔAIC larger than 10 have essentially no support in the data compared to the best model [2]. Model 1 is the best among the tested models for all state distribution models, which balance accuracy and complexity of the model. It has diurnal effects on all transition probabilities, and nonlinear effects of τ_t and d_t on some of the transition probabilities. The marginal fit is illustrated in Fig 5 for correlated log-normal state distributions. The fits look convincing.

Table 2. Model selection results.

	Independent Gamma distribution		Independent Log-normal distribution		**Correlated Log-normal distribution**	
Model	np	ΔAIC	np	ΔAIC	np	ΔAIC
1	60	5050.51	60	2309.12	**63**	**0**
2	42	5386.94	42	2652.4	45	255.97
3	48	5096.47	48	2353.17	51	34.34
4	42	5194.44	42	2451.91	45	166.88

Differences in AIC values, ΔAIC, between the different models, where all AIC values were subtracted the AIC value of the model with the lowest AIC. For all the tested state distributions, covariate model 1 was preferred. The best fit is given by the minimum AIC, which occurs for the dependent log-normal state distribution, highlighted in bold. np: number of parameters.

A deep dive is defined to be a dive of at least $350\,m$, and therefore d_t will typically be 0 once the narwhal is in state 1 or 2, and thus, we only expect d_t to influence η_{3j}. This is exactly what we see; models with no dependence of d_t in transitions from states 1 and 2 are preferred over the other models (results not shown). Furthermore, we would expect that low or moderate values of d_t would make the propensity to stay in state 3 large, and thus, decrease the probabilities for transiting to other states, whereas large values of d_t are expected to increase the probabilities of changing, because the narwhal might need to rest. Thus, a non-linear relation might be adequate. This is explored in Models 1, 3 and 4, and effectively, the optimal model 1 does include a non-linear relationship of d_t. Fig 6 illustrates the estimated covariate effects, and parameter
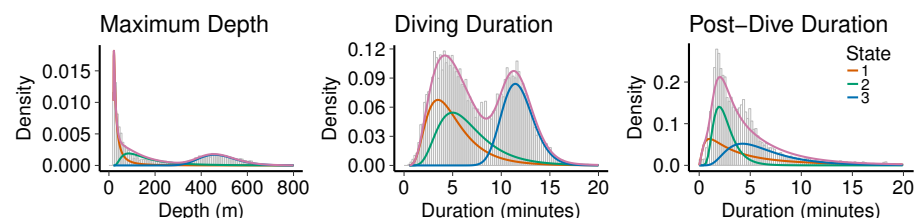
Fig 5. Model fit. Histograms of response variables Maximum Depth, Duration of Dives and Post-dive Duration. The fit of Model 1 with correlated log-normal distributions is indicated with a magenta curve, and the distribution of the fitted states are indicated with colours as given in the legend. State 1 corresponds to near-surface travelling, state 2 is feeding at medium depths, and state 3 is intense feeding at large depths.

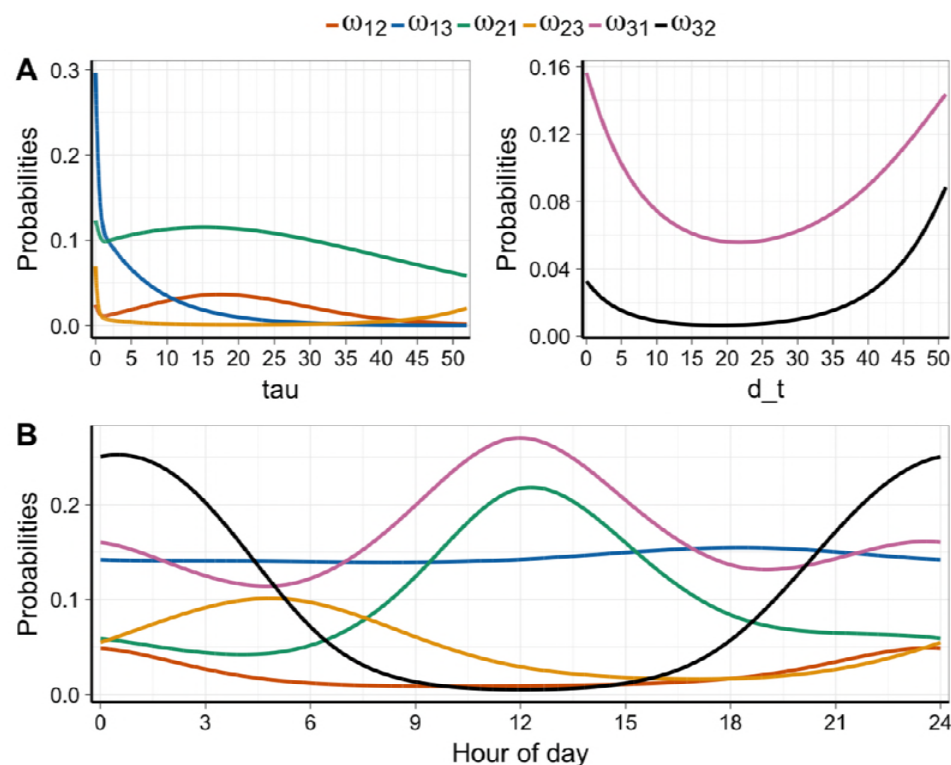estimates and confidence intervals can be found in Tables 4 and 5 in the Supplementary Material.

Fig 6. Covariate effects. A: Transition probabilities between behavioural states depending on covariates of correlated log-normal model 1, at approximately 12 pm. B: Transition probabilities depending on diurnal effects in model 1 with correlated log-normal state distributions, calculated for $\tau_t = 0.58$ and $d_t = 0$ (the medians).

The covariate τ_t indicates the time passed since last deep dive. We expect that τ_t has impacts on states 1 and 2, but not on state 3 (which is the case for the selected model). In the left panel of Fig 6A the effect of τ_t is illustrated. The transition probabilities do not seem to depend much on τ_t, except for the probability of changing

from surface travelling to deep dives. The probability is much higher for very small   296
values of τ_t, and decreasing fast towards 0 for larger values. This is not what was   297
expected, but might reflect the following. When short time has passed since last deep   298
dive, it is probably also short time since the whale was in state 3 of intense feeding.   299
Thus, it reflects that the whale is still in an overall behavioral state of intense feeding,   300
but just had a short break with surface travelling. This phenomenon can be seen in Fig   301
7 where the state decoding is shown for 12 representative hours. It is seen that after (at   302
least) six dives in state 3, the whale changes to a few shallow dives for a short time, and   303
then continues with another three dives in state 3. When a little longer time passes, the   304
whale has effectively stopped its intense foraging, and the probability of a change to   305
state 3 becomes smaller. Then, when long time has passed, we expect the transition   306
probability to increase, which is not what is estimated. However, there are very few   307
large observations of τ_t: 75% of the values are belov 2.8 hours, and 90% are below 7.8   308
hours. Therefore effect estimates for large values are unreliable. The same is true for   309
covariate d_t: more than half are 0, 75% are 2 or smaller, and 90% are 8 or lower.   310
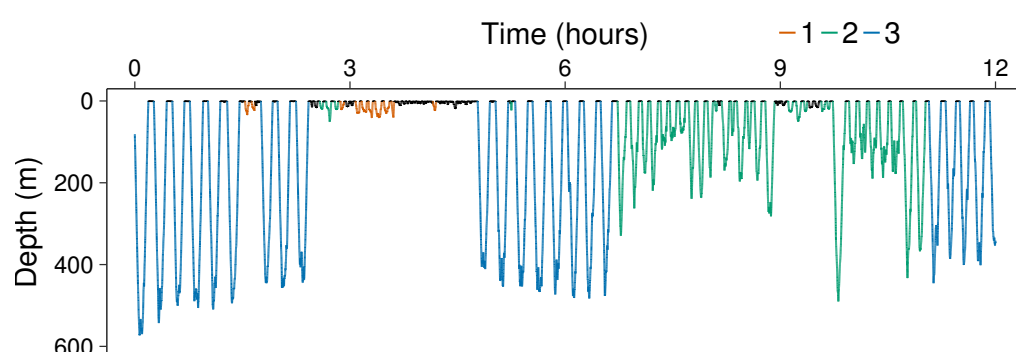

Fig 7. State decoding close-up. The estimated hidden state per dive for 12 hours of the data.

The effect of d_t is illustrated in the right panel of Fig 6A. As expected, for values   311
above 20 dives in a row, the probabilities to exit state 3 increase with increasing d_t. The   312
probability of changing to shallow waters (state 1) is much higher than the probability   313
of changing to medium deep dives (state 2) after a period of intense feeding (state 3).   314
Fig 6B shows the diurnal effects on the transition probabilities. Changing from state   315
3 to 2 has highest probability around midnight, whereas changing from state 2 to 3 has   316
highest probability around 6 pm. Changing to state 1 has highest probability around   317
noon. The transition probabilities from state 1 do not depend much on diurnal effects.   318
Table 3 lists the estimated means and standard deviations of the state distributions   319
for correlated log-normal and independent gamma distributions, respectively. Means   320
and standard deviations of maximum depth and diving time are estimated larger for   321
both shallow dives (state 1) and medium dives (state 2) with the correlated model   322
compared to the independent model, whereas the two models estimate mean and   323
variances approximately the same for deep dives (state 3). Thus, taking into account   324
the dependence between the two state variables reveals more variable diving patterns   325
(i.e., larger variance within states), unless the narwhal is doing intense feeding at deep   326
dives, where the need for regular breathing do not allow the whale to digress.   327
Fig 8 shows the decoded hidden states for Model 1 with dependent log-normal state   328
distribution. The correlated model estimates that the narwhal spends around 33.5% of   329
its dives, corresponding to 34.4% of the time, in State 1 of near-surface travelling, which   330
encompasses dives down to $793\,m$ of durations up to 28 minutes. This is a large value   331
for the surface state, but it is only the extreme tail of the distribution, and is   332

Table 3. Summary measures of Model 1.

	State 1	State 2	State 3
Correlated Log-normal distribution			
E_{MD}	34.40	160.44	459.41
SD_{MD}	79.66	143.22	83.14
E_{DT}	5.08	6.63	11.78
SD_{DT}	2.71	3.02	1.65
E_{PD}	6.65	2.58	6.49
SD_{PD}	10.61	1.16	3.73
Corr	0.56	0.81	0.46
Independent Log-normal distribution			
E_{MD}	25.95	135.55	458 .00
SD_{MD}	52.94	123.59	85.86
E_{DT}	4.39	7.01	11.81
SD_{DT}	2.16	2.20	1.67
E_{PD}	6.85	2.65	6.64
SD_{PD}	11.59	1.13	4.18
Independent Gamma distribution			
E_{MD}	20.46	113.71	455.46
SD_{MD}	22.09	95.05	87.47
E_{DT}	4.10	6.84	11.78
SD_{DT}	1.75	2.21	1.70
E_{PD}	8.24	2.56	7.32
SD_{PD}	9.5	1.03	5.09

Means and standard deviations of Model 1 based on correlated Log-normal, independent Log-normal and independent Gamma distribution. MD: Maximum Depth; DT: Diving Time; PD: Post-Dive duration. E: mean; SD: standard deviation; Corr: Correlation.

represented by a single dive. It reflects that the log-normal distribution has heavier tails than the gamma distribution. Of the time spent in state 1, only 36.4% of the time is spent diving, the rest of the time the whale is in the surface. The narwhal spends around 32.5% of its dives, corresponding to 19.9% of the time, in medium depths of between $20\,m$ and $836\,m$ and durations between 0.7 and 22 minutes. Also here, a few deep dives are decoded as medium dives. Of the time spent in state 2, 69.9% of the time is spent diving, the rest of the time the whale is in the surface. Finally, 34.0% of dives, corresponding to 45.9% of the time, are spent during intense feeding at depths between $231.5\,m$ and $910.5\,m$ and durations between 7.2 and 19.5 minutes. Of the time spent in state 3, 62.8% of the time is spent diving, the rest of the time the whale is in the surface. Fig 7 illustrates a close-up of the decoding of dives for an example period of 12 hours. The correlated model thus decodes a few of the deep dives as pertaining to states 1 and 2, probably because of these dives taking longer time than the deep dives decoded as state 3.

Model 1 with an independent gamma state distribution estimates that the narwhal spends around 28.9% of its dives in State 1 of near-surface travelling, which encompasses dives up to $258\,m$ of durations up to 14.8 minutes. It spends around 34.8% of its dives in medium depths of between $20\,m$ and $584\,m$ and durations between 1.5 and 28 minutes. Finally, 36.3% of dives are spent during intense feeding at depths between $172.5\,m$ and $910.5\,m$ and durations between 6.5 and 22 minutes. The independent model seems to be mostly guided by the maximum depth.

The decoding of behavioural states is thus different depending on whether
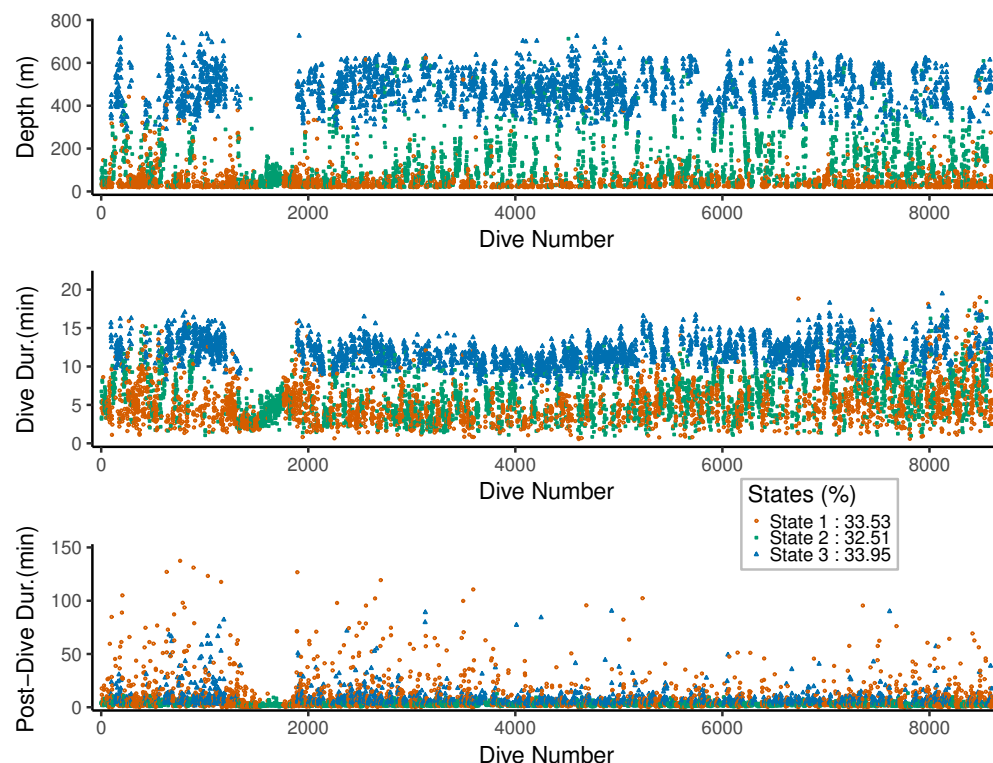
Fig 8. State decoding. The estimated hidden state per dive for each of the three observed variables under covariate model 1 and state distribution the correlated log-normal. The longest pause of no deep dives starts from the 1345th dive until the 1894th dive, and it lasts approximately 2 days and 17.5 hours.

correlation between observed variables are accounted for or not. The independent gamma model estimated that the narwhal spent 71.1% of its dives for feeding, whereas when correlation is taken into account, this estimate drops to 66.5%. These numbers are not too different, but to explore the differences further, we plot the distributions of the state variables and the covariates as a function of the decoding for the independent gamma and the correlated log-normal models in Fig 9. The orange density plots show the dives that are decoded the same in the two models, whereas the blue density plots show the dives that are differently decoded in the two models. In the correlated model, the distributions of dive characteristics in State 1 and 2 are broader, whereas they are more peaked in state 3, compared to the independent model.

To check the fit of the model more than what is given in Fig 5, we calculated the pseudo-residuals [22], qq-plots can be found in Fig 10. The fit is acceptable, maybe except for a too small lower tail for the Maximum Depth variable. This is probably due to the threshold of a depth of $20m$ in the definition of a dive.

4 Discussion

In this study, we investigate different multivariate HMMs with covariate effects for modelling the diving activity of a narwhal in the vertical dimension in the water column. The data set is extraordinarily long, which has made it possible to describe diurnal patterns in diving behaviour.

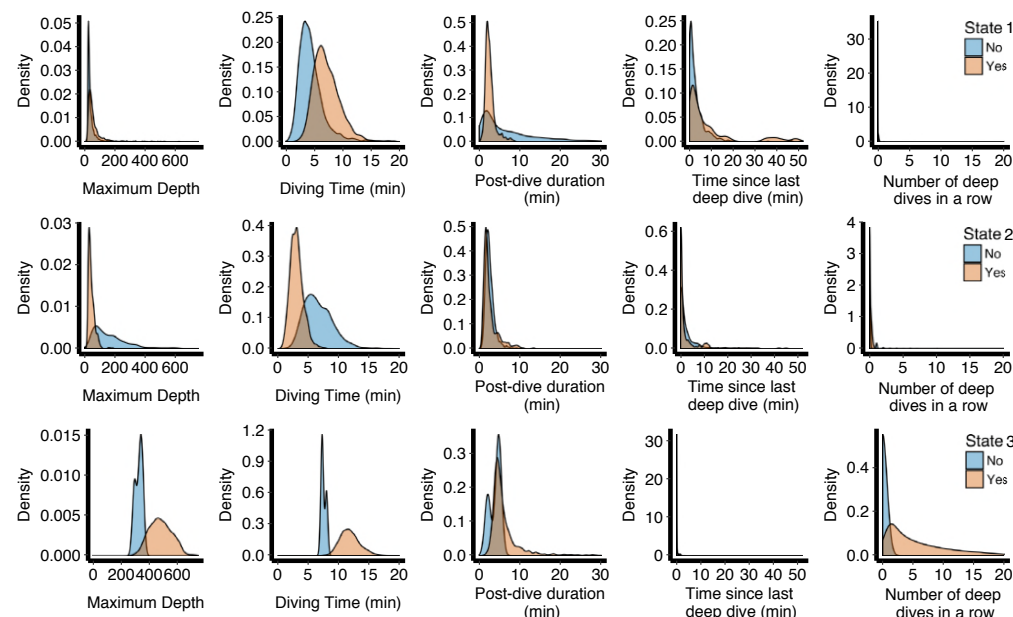We extend the usual HMM models for diving behaviour of marine mammals to allow

Fig 9. Distribution of hidden states. The empirical distributions of the response variables and the covariates stratified by decoded state in the correlated log-normal model 1. Within each panel, the orange density plots depict the distribution where the independent gamma and the correlated log-normal distributions agree on the decoded states, the blue density plots are those where the two distributions disagree.
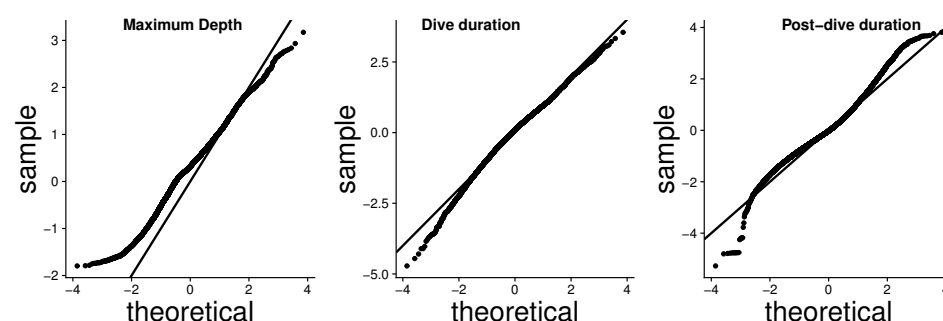

Fig 10. Quantile-quantile residual plots. QQ-plots of forecast pseudo-residuals from covariate model 1 with correlated log-normal state distribution.

for dependence between state distributions, and show that the dependence has some impact on the conclusions drawn about the diving behaviour. We find that statistically the correlated model clearly outperforms the independent model, and more importantly, conclusions on the diving behaviour differ between the two models. The main differences are that the correlated model estimates more variable state distributions of maximum depth and dive duration, and that 66.5% of the dives are for feeding, compared to 71.1% in the independent model, under the assumption that states 2 and 3 in fact are representing feeding states in both models. Furthermore, the estimated means and variances of depth and duration of shallow and medium dives are estimated larger in the correlated model. Finally, ignoring the dependence between response variables usually leads to too narrow confidence intervals on parameter estimates.

PLOS | SUBMISSION

Direct observations of feeding events were limited to the first week of the diving data but the depths where feeding events were detected served as a valid proxy for the depth threshold between behavioural state 2 and state 3. The observation that feeding events involve deep dives ($\geq 350m$) is also supported by studies of the buzzing activity during dives to different depths for narwhals travelling in the same area and time of the year as the whale included in this study [3].

Apparently the whale could stay in state 1 and 2 for long periods (> 24 hours) without transiting to state 3, and it even showed a pause of almost 3 days without deep dives. This indicates that feeding occur infrequently and that narwhals at least during summer and fall may have extended periods without feeding activity (see also [13]). However, the median of these pauses without state 3 dives was 44 minutes and the mean was 2 hours.

Transition from state 1 to presumed feeding activity is more likely to be to state 3 with deep dives, and rarely goes to state 2 from state 1. Diving activity in state 3 usually last for a series of dives (5-10) perhaps indicating that specific layers of prey is being detected and explored for a series of dives before the whale needs to spend an extended period at the surface. The post dive time is typically around 6.5 minutes after a state 3 dive, whereas it is typically only 2.6 minutes after a state 2 dive. The whale probably needs to spend more time at the surface to recover from nitrogen tissue tension following a longer breath-hold diving activity.

Even though detailed dive information supplemented by data on feeding events have been available for this analysis it may still not be adequate for describing the important drivers of diving behaviour. Both physiological constrains and reproductive state as well as environmental conditions may influence the diving activity to an extent that cannot be fully discerned in HMM analysis of dive series. For logistical reasons it is very difficult if not impossible to obtain information on all factors that affect the diving behaviour. However, the analysis of dive series provides a minimal insight into the integrated effect of the various factors driving the diving behaviour and the major advantage of the HMM analysis probably relies in the objective inter- and intra-specific comparison of diving activity.

Acknowledgements

References

1. Aguilar Soto, N., Johnson, M. P., Madsen, P. T., Díaz, F., Domínguez, I., Brito, A., and Tyack, P. Cheetahs of the deep sea: deep foraging sprints in short-finned pilot whales off tenerife (canary islands). Journal of Animal Ecology. 2008:936–947.

2. Burnham KM, Anderson DR. Model Selection and Multimodel Inference. 2nd ed. Springer, New York; 2002.

3. Blackwell, S., Tervo, O., Conrad, A., Sinding, M. H. R., Ditlevsen, S., and Heide-Jørgensen, M. (2018). Spatial and temporal patterns of sound production in East Greenland narwhals. PLoS ONE 13(6): e0198295. https://doi.org/10.1371/journal.pone.0198295

4. DeRuiter SL, Southall BL, Calambokidis J, Zimmer WMX, Sadykova D, Falcone EA, et al. First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. Biology Letters. 2013;9(4). doi:10.1098/rsbl.2013.0223.

5. DeRuiter SL, Langrock R, Skirbutas T, Goldbogen JA, Calambokidis J, Friedlaender AS, et al. A multivariate mixed hidden Markov model for blue whale behaviour and responses to sound exposure. Annals of Applied Statistics. 2017;11(1):362–392. doi:10.1214/16-AOAS1008.

6. Eddelbuettel D, François R. Rcpp: Seamless R and C plus plus Integration. Journal of Statistical Software. 2011;40(8):1–18.

7. Forney G. Viterbi algorithm. Proceedings of the IEEE. 1973;61(3):268–278.

8. Garde E, Hansen SH, Ditlevsen S, Tvermosegård KB, Hansen J, Harding KC, et al. Life history parameters of narwhals (Monodon monoceros) from Greenland. Journal of Mammalogy. 2015;96(4):866–879.

9. Heide-Jørgensen, M., Dietz, R., and Leatherwood, S. A note on the diet of narwhals (Monodon monoceros) in Inglefield Bredning (NW Greenland). Meddr Grønland, Biosci. 1994; 39:213–216.

10. Heide-Jørgensen MP. Narwhal Monodon monoceros. In: Perrin, WF and Wursig, B and Thewissen, JGM, editor. Encyclopedia of Marine Mammals, 2nd Edition; 2009. p. 754–758.

11. Heide-Jørgensen, M. P., Nielsen, N. H., Hansen, R. G., and Blackwell, S. B. Stomach temperature of narwhals (Monodon monoceros) during feeding events. Animal Biotelemetry. 2014:2–9.

12. Heide-Jørgensen MP, Nielsen NH, Hansen RG, Schmidt HC, Blackwell SB, Jørgensen OA. The predictable narwhal: satellite tracking shows behavioural similarities between isolated subpopulations. Journal of Zoology. 2015;297(1):54–65.

13. Laidre K, Heide-Jørgensen M. Winter feeding intensity of narwhals (Monodon monoceros). Marine Mammal Science. 2005;21(1):45–57.

14. Langrock R, King R, Matthiopoulos J, Thomas L, Fortin D, Morales JM. Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. Ecology. 2012;93(11):2336–2342.

15. Langrock R, Marques TA, Baird RW, Thomas L. Modeling the Diving Behavior of Whales: A Latent-Variable Approach with Feedback and Semi-Markovian Components. Journal of Agricultural Biological and Environmental Statistics. 2014;19(1):82–100.

16. Michelot T, Langrock R, Bestley S, Jonsen ID, Photopoulou T, Patterson TA. Estimation and simulation of foraging trips in land-based marine predators. Ecology. 2017;98(7):1932–1944.

17. Patterson TA, Basson M, Bravington MV, Gunn JS. Classifying movement behaviour in relation to environmental conditions using hidden Markov models. Journal of Animal Ecology. 2009;78(6):1113–1123.

18. Pohle J, Langrock R, van Beest FM, Schmidt NM. Selecting the Number of States in Hidden Markov Models: Pragmatic Solutions Illustrated Using Animal Movement. Journal of Agricultural Biological and Environmental Statistics. 2017;22(3):270–293.

19. Quick NJ, Isojunno S, Sadykova D, Bowers M, Nowacek DP, Read AJ. Hidden Markov models reveal complexity in the diving behaviour of short-finned pilot whales. Scientific Reports. 2017;7.

20. R Core Team. R: A Language and Environment for Statistical Computing; 2017. Available from: https://www.R-project.org/.

21. Reeves, R. R., Ewins, P. J., Agbayani, S., Heide-Jørgensen, M. P., Kovacs, K. M., Lydersen, C., Suydam, R., Elliott, W., Polet, G., van Dijk, Y., and Blijleven, R. Distribution of endemic cetaceans in relation to hydrocarbon development and commercial shipping in a warming Arctic. Marine Policy. 2014;**44,**:375–389.

22. Zucchini W, MacDonald IL, Langrock R. Hidden Markov Models for Time Series: An Introduction using R. 2nd ed. Chapman & Hall/CRC, FL, Boca Raton; 2016.

PLOS | **SUBMISSION**

Supplementary Materials

Table 4. Estimates of the model parameters (1st part) in model 1 and their 95% confidence intervals for correlated Log-normal distribution.

Correlated log-normal distribution		
	Estimate	95% CI
μ_1^{MD}	2.61	[2.56, 2.66]
μ_2^{MD}	4.78	[4.73, 4.84]
μ_3^{MD}	6.11	[6.11, 6.12]
σ_1^{MD}	1.36	[1.33, 1.39]
σ_2^{MD}	0.77	[0.72, 0.81]
σ_3^{MD}	0.18	[0.17, 0.19]
μ_1^{DT}	1.50	[1.48, 1.52]
μ_2^{DT}	1.80	[1.77, 1.82]
μ_3^{DT}	2.46	[2.45, 2.46]
σ_1^{DT}	0.50	[0.49, 0.51]
σ_2^{DT}	0.43	[0.41, 0.46]
σ_3^{DT}	0.14	[0.14, 0.14]
μ_1^{PD}	1.26	[1.22, 1.3]
μ_2^{PD}	0.86	[0.83, 0.88]
μ_3^{PD}	1.73	[1.71, 1.75]
σ_1^{PD}	1.13	[1.10, 1.15]
σ_2^{PD}	0.43	[0.41, 0.45]
σ_3^{PD}	0.53	[0.52, 0.55]
ρ_1	0.56	[0.53, 0.58]
ρ_2	0.81	[0.78, 0.83]
ρ_3	0.46	[0.43, 0.50]

In state i, μ_i and σ_i are the log-mean and log-standard deviation of the correlated log-normal distribution. Index MD stands for Maximum Depth, DT stands for Dive Duration and PD stands for Post-Dive time. The depth is measured in meters, and time in seconds. The confidence intervals were computed from the Hessian of the negative log-likelihood function, i.e., based on the inverse of the observed Fisher information (to be continued on next page).

Table 5. Estimates of the model parameters (2nd part).

	Correlated log-normal distribution	
	Estimate	95% CI
α_{00}	-3.82	$[-5.30, -2.34]$
α_{01}	-0.97	$[-1.57, -0.37]$
β_{00}	0.08	$[-1.18, 1.33]$
β_{01}	-1.74	$[-3.22, -0.263]$
γ_{00}	1.88	$[1.10, 2.66]$
γ_{01}	-3.08	$[-5.15, -1.02]$
$\theta_{12}^{(1)}$	1.95	$[0.11, 3.80]$
$\theta_{12}^{(2)}$	-1.71	$[-3.77, 0.36]$
$\theta_{12}^{(3)}$	-2.48	$[-5.48, 0.52]$
$\theta_{13}^{(1)}$	-3.11	$[-4.39, -1.82]$
$\theta_{13}^{(2)}$	-6.60	$[-8.28, -4.92]$
$\theta_{13}^{(3)}$	-7.08	$[-10.5, -3.61]$
$\theta_{21}^{(1)}$	0.19	$[-1.58, 1.97]$
$\theta_{21}^{(2)}$	-0.57	$[-2.39, 1.25]$
$\theta_{21}^{(3)}$	-0.51	$[-4.10, 3.08]$
$\theta_{23}^{(1)}$	-5.04	$[-8.82, -1.26]$
$\theta_{23}^{(2)}$	-5.39	$[-7.13, -3.66]$
$\theta_{23}^{(3)}$	-0.48	$[-1.93, 0.98]$
$\zeta_{31}^{(1)}$	-2.83	$[-3.75, -1.92]$
$\zeta_{31}^{(2)}$	-5.35	$[-6.61, -4.1]$
$\zeta_{31}^{(3)}$	-0.84	$[-2.51, 0.83]$
$\zeta_{32}^{(1)}$	-4.83	$[-6.38, -3.28]$
$\zeta_{32}^{(2)}$	-8.80	$[-10.30, -7.28]$
$\zeta_{32}^{(3)}$	-0.38	$[-2.76, 2.00]$
$\delta_{12}^{(1)}$	0.41	$[-1.78, 2.60]$
$\delta_{12}^{(2)}$	2.53	$[1.33, 3.73]$
$\delta_{12}^{(3)}$	-0.06	$[-2.34, 2.22]$
$\delta_{13}^{(1)}$	0.21	$[-0.82, 1.23]$
$\delta_{13}^{(2)}$	0.08	$[-0.42, 0.58]$
$\delta_{13}^{(3)}$	0.02	$[-1.06, 1.10]$
$\delta_{21}^{(1)}$	-2.77	$[-4.88, -0.66]$
$\delta_{21}^{(2)}$	-2.29	$[-3.34, -1.25]$
$\delta_{21}^{(3)}$	-3.46	$[-5.37, -1.55]$
$\delta_{23}^{(1)}$	-1.24	$[-3.48, 1.01]$
$\delta_{23}^{(2)}$	0.63	$[-0.93, 2.19]$
$\delta_{23}^{(3)}$	1.54	$[-0.37, 3.45]$
$\delta_{31}^{(1)}$	-2.32	$[-3.41, -1.23]$
$\delta_{31}^{(2)}$	-0.60	$[-1.14, -0.06]$
$\delta_{31}^{(3)}$	-2.49	$[-3.59, -1.39]$
$\delta_{32}^{(1)}$	4.45	$[0.99, 7.90]$
$\delta_{32}^{(2)}$	6.28	$[4.05, 8.51]$
$\delta_{32}^{(3)}$	4.94	$[1.47, 8.41]$