On the surfacing rate in minke whales in the northeastern Atlantic.

Tore Schweder and Nils Øien
CEES, University of Oslo, Box 1066 Blindern, N-0316 Oslo, Norway;
Institute of Marine Research, Box 1870 Nordnes, N-5817 Bergen, Norway.

April 27, 2007

Abstract

Ten minke whales have been VHF-tagged and the time of successive surfacings recorded. From distributional considerations dives are classified as short (<35s), long (>140s) or of intermediate duration. Logistic regression models with random effects for whale, and for batches of successive dives are fitted for the probability of a dive being short, and the conditional probability of the dive being long given that it is not short. The number of surfacings in the minute before, and the five minutes period before the current surfacing are used as covariates.

1 Introduction

The aim of this paper is to establish a statistical model for successive dive times for minke whales in the Northeast Atlantic. We are particularly interested in characterizing the variability in short term dive pattern between whales, and also within whale over time.

The methods used for estimating abundance of minke whales in the Northeast Atlantic (Schweder et al. 1997, Skaug et al. 2004) depend on the availability of data on surfacing rates and patterns of minke whales. In this context a number of dive time series have been collected by means of VHF tagging over the period 2001-2006. In these experiments the surfacing behavior of individual minke whales has been studied over periods of about 3 hrs to 64 hours (Øien et al. 2007) and different patterns have emerged although a general aspect of a mixture of short and long dives is apparent. While general descriptive results have been reported (Øien et al. 2003, 2007), no explorative modeling has been done. Statistical modeling of successive dive times for minke whales is of interest for abundance estimation by the hazard probability method (Skaug et al. 2004, Okamura et al. 2006), and might also reveal possible biological and physiological dependencies.

2 Material and methods

2.1 Material

The dive series have been collected during sightings surveys and dedicated instrumentation activities as described in Øien et al. (2003, 2007). Radio tags were applied to minke whales by means of the launching system ARTS (Arial Rocket Transmission System, Heide-Jørgensen et al. 2001). Although not usually observed directly, some disturbance of the animals caused by the application process might be expected but this could not be verified because the whales were usually visually lost after a short time. The instrumented whales were followed from the survey vessel by means of a receiver with antenna and direction finder and a PC-GPS unit where all surfacing were recorded along with the survey vessel's position. Due to the fact that we are not able to follow minke whales visually over significant periods, no considerations have been done with respect to behavioral aspects, but sea state has usually been recorded. The resulting data include date and time of surfacing, corresponding vessel position, and Beaufort sea state.

2.2 Statistical method

The recorded data contain information on time and location of successive surfacings over observational periods of differing length for ten whales. Here we concentrate on the process of dives and surfacings in time, and disregard spatial and other aspects of the process.

The tagging might impact the whale and influence the dive pattern for some time after tagging. This is informally investigated in two ways. Detrended plots of surfacing time by surfacing number are inspected for possible initial periods that stands out in contrast to variation later in the series. Mean dive time during the first 30 dives and also the first 60 dives after tagging are also compared to mean dive time for each series.

Next we look for a distinction between short and longer dives. This is done within series, and in the pooled material by way of distribution plots. From these we argue that the longer dives are of two types: dives of intermediate length, and long dives. The distinction between them is also found by looking at distribution plots.

We end up with a three component mixture model for the marginal distribution of dive length. The dynamics in the dive process is modeled by the mixture probabilities being stochastically determined by the dive activity in a short period preceding a given dive.

Since short dives primarily are made in order to draw enough air, the probability of the next dive being short should depend on the amount of breathing the individual has done in the period shortly before the current surfacing. This amount of breathing could be characterized in several ways, say by the length of the previous couple of dives, or by the number of surfacings in some short previous periods. By some graphical analysis we decided to use two covariates. One is the number of surfacings done in the minute before the current surfacing, with 3 or more surfacings coded as 3. This variable is x_1. The other covariate x_2 is an indicator for the whale having had one or more surfacings between 1 and 5 minutes before the current surfacing. Both the covariates are regarded as categorical, although both are expected to have a monotonous impact on the probability p of the coming dive being short.

We use linear logistic regression models for this relationship. The logistic mixed effects models for the probability p of the next dive being short are of the form

$$
\begin{aligned}
\log(p/(1-p)) &= whale + x_1 * x_2 + b && \text{Ms1} \\
&= x_1 * x_2 + b && \text{Ms2} \\
&= x_1 * x_2 + w && \text{Ms3} \\
&= x_1 * x_2 + w + b && \text{Ms4}
\end{aligned}
$$

The first of these models, Ms1 ("M" for mixed, "s" for short) has a random batch effect b, while the effects of the categorical variables x_1 and x_2 with interactions are fixed, as is the intercept depence on whale. In the bottom two models the whale effect w is random. Fixed effects models Fs1 and Fs2 are obtained by loosing the random batch effect in models Ms1 and Ms2. Batches consists of 30 consecutive dives in these four models. Model Ms4 is turned into model Ms5 by changing the batch size to 60.

Next, the conditional probability of a dive being long given that it is not short, $\pi = r/(1-p)$ is considered. Here, r is the unconditional probability of a dive being long. The same models as for short dives are fitted, but now denoted Ml1, Ml2 etc.

Do whales take several rapid breaths to prepare for long dives, or do they rather take more rapid breaths after long dives? That is, do they tend to prepare for long dives, or are they responding to having taken log dives by rapid breathing to resupply oxygen? To study this question we regress the long dives with at least one long dive both before and after in the series on the numbers of successive short dives before, and after the long dive.

3 Results and discussion

Some summary statistics for the 10 observed dive series are given in Table 1. In two of the ten series, mean dive time is less for the 30 first dives after tagging than in the remaining period, while the mean diving time in the first 60 dives is very similar to the long term mean for all individuals. This might be due to some stress by being tagged, lasting for the first 30 dives, but not lasting much further. This is also confirmed by considering linearly detrended plots of surfacing time by surfacing number. Figure 1 shows the plot for whale 10. For this whale the first 30 minutes or so might be a period of some stress. For most other whales the pattern at the start of the detrended graph is hard to distinguish from the long term variation. We decided to remove the first 30 dives from further analysis.

The mean dive times d_{30+} varies considerably between whales (Table 1). When splitting the observed period in five minute periods, and counting the number of surfacings in each, the mean count varies also considerably between whales. In all but one of the series, the variance in the counts is less than the mean, and the counts are thus under-dispersed relative to the Poisson distribution.

From Figures 2 and 3 an approximate model for individual dives times is obtained by distinguishing between short, medium length and long dives. We found 35s and 140s to be good division points between short, medium length and long dives. The short and the long dives are approximately log-normally distributed, while those of intermediate length are approximately uniformly distributed. From the qq-plots at the bottom of Figure 4 the log-normal distribution is seen to approximate the distribution of both short and long dives, except for the extreme tails. The pooled set of dive times (time in seconds between subsequent recorded surfacings) have empirical cumulative distribution function as shown in Figure 5, with the fitted mixture cdf as a dotted curve.

The modeled marginal distribution of dive time has density

$$f(d) = p\frac{1}{\sigma_s}\phi\left(\frac{\log(d) - \mu_s}{\sigma_s}\right)/d + q\frac{1}{105}u\left(\frac{d - 35}{105}\right) + r\frac{1}{\sigma_l}\phi\left(\frac{\log(d) - \mu_l}{\sigma_l}\right)/d, \quad (1)$$

where ϕ is the standard normal density, u the standard uniform density, and non-negative mixing proportions, $p + q + r = 1$, and the parameters μ and σ given in Figure 5 for the pooled sample.

3.1 Logistic regressions

Surfacings are strongly serially correlated. Figure 6 shows the surfacings in 20 randomly chosen 5-minutes periods. The picture looks pretty much like this for all the whales except for whale 9 (Figure 7) which tend to have more short dives (see also Øien et al 2007).

Our dynamic dive model is based on the three mixture components in (1), but with mixture probabilities depending on covariates calculated from the period preceding the current surfacing. The short term auto-correlation in the dive series suggests logistic regression models for the mixing probabilities. After some search for covariates to characterize the breathing state of the whale at a given surfacing, we chose the pair x_1 and x_2 where the first is the number of surfacings in the previous minute, with 3 or more coded as 3, and where x_2 is an indicator for whether the whale had a surfacing between 1 and 5 minutes previous to the surfacing in question. Figure 8 shows box plots of dive time by levels of x_1 and x_2 for the pooled sample, while Figure 9 shows the same but also by whale. There is clearly a strong effect of x_1 on the dive time, while the effect of x_2 is also clear.

Table 2 shows regression results for fixed and mixed effects logistic regression models for the probability of the coming dive being short (<35s). The covariates x_1 and x_2 are categorical, with contrasts of the type treatment. Fixed effects of whale are not shown, but the standard variations of the estimated intercept by whale are given as sd(whale) when whale has fixed effects.

The fixed effects models Fs1 and Fs2 are fitted by glm, while the mixed effects models Ms1 and Ms2 are fitted by maximum likelihood by using the package glmmML in the system R. The mixed effects models are fitted by the restricted maximum likelihood method by way of glmmPQL in R, but where residual deviance and AIC is not produced. In the models with batch b as a random effect, batch size is 30 consecutive dives in models Ms2 and Ms4, while it is 60 in model Ms5. Corresponding logistic models are fitted for the conditional probability of a dive being long (>140) given that it is not short.

The results show considerable variability in these probabilities both between and within whales, and that these variabilities are roughly of the same magnitude. It is interesting that the variability within whale does not increase when batch size is doubled.

3.2 Simulating dive sequences

When estimating abundance of minke whales by the hazard probability model, surfacings are traditionally regarded as following a Poisson point process with constant rate (Skaug et al. 2004). This is indeed an approximate model for the surfacing process. Our model for the dive series is also a model for the surfacing process, and it provides a more representative picture than the Poisson model.

Models Ms4 and Ml4 seems the most appropriate for simulating dive series for a random period of say 30 minutes for a randomly chosen minke whale in the northeastern Atlantic. The numbers below are taken from these models.

Consecutive dives for a randomly encountered whale are simulated in batches of 30 as follows. First a random whale effect is drawn from the $\mathrm{N}(0, .66^2)$ distribution for the choice probability p, and from the $\mathrm{N}(0, .95^2)$ distribution for the conditional choice probability π. Then random batch effects are drawn for each batch of 30 dives from the $\mathrm{N}(0, .77^2)$ and $\mathrm{N}(0, 1.01^2)$ distributions for p and π respectively. For given values of x_1 and x_2 the probability p is calculated, and a Bernoulli trial is performed to determine whether the next dive is short. If so, its length is drawn from the common log-normal distribution of short dives. If not short, the probability π is calculated and a Bernoulli experiment followed by a draw from either the uniform distribution of medium length dives or the log-normal distribution of long dives is carried out. The covariates are then updated, and the process moved forward.

The simulation might be initiated by drawing values for x_1 and x_2 from Table 4, for a couple of initial dives if necessary. Dive times are then simulated sequentially as described above. Since the process has rather short memory, disposing of the first 30 dives or so is sufficient to have the process approximately in stationary state.

In some cases, a succession of dives from a randomly chosen surfacing is required. In other situations however, e.g. when the simulation is part of the abundance estimation process, the time points of surfacings over a randomly chosen time window is of interest.

3.3 Is rapid surfacing a response to breathing need, or preparation for a long dive?

Consider long dives dl with completed sequences of successive short dives both before and after the long dive, let be and af be the number of successive short dives before and after the long dive respectively. Frequency distributions of these numbers are given in Table 5. A gam analysis of the form $dl \sim whale + s(be) + s(af)$ with $whale$ categorical and s a smoothing spline, shows that the length of long dives is related to both covariates (Figure 10). The effect of number of successive short dives on the length of the long dive is linear at the start, peaks at $be = af = 3$ and falls slightly thereafter. The effect of be is about 60% that of af. Table 5 shows the frequency distributions for be and af.

In other words, the series of more rapid blows do not seem to be very dependent on the longer dives. This may be in line with the findings of Croll et al. (2001) that Balaenopteridae whales seem to have considerable shorter dives than their theoretical aerobic dive limit based on body size. The usual finding is that there is a significant relationship between body mass and dive duration (Schreer & Kovacs 1997, Mori 2002). In a study of bottlenose dolphins, Fortuna et al. (1999) found that correlations between short and long dives varied between different behavioural states. The strongest correlations appeared to be when the bottlenose dolphins were diving; then both the previous long dive and the number of ventilations as well as the duration of the short dives' sequence and the duration of long dive were correlated.

References

[1] Croll, D.A., Avecedo-Gutiérrez, A., Tershy, B.R., Urbán-Ramírez, J. 2001. The diving behavior of blue and fin whales: is dive duration shorter than expected based on oxygen stores? *Comparative Biochemistry and Physiology Part A* **129**:797-809.

Series	whale	hours	surfacings	d_{30}	d_{60}	d_{30+}	n	mean	variance
mi_2001_01	1	6.52	352	56.10	68.00	67.85	72	4.40	2.13
mi_2001_02	2	63.88	2319	77.67	99.56	99.49	758	3.02	1.17
mi_2002_01	3	3.88	236	64.70	57.61	58.66	40	5.13	1.96
mi_2002_02	4	50.56	2499	69.93	73.01	72.90	599	4.11	3.26
mi_2002_03	5	47.84	1394	112.70	124.27	123.88	562	2.42	1.32
mi_2003_01	6	23.82	1445	64.60	59.16	59.29	279	5.07	4.38
mi_2004_01	7	2.89	136	73.00	78.01	78.28	27	3.81	1.93
mi_2004_02	8	10.46	428	87.53	87.95	88.22	116	3.41	2.82
mi_2006_01	9	15.39	1107	21.07	51.67	50.91	182	5.91	9.53
mi_2006_02	10	39.64	1743	73.13	82.01	82.08	468	3.66	2.08

Table 1: Hours of observation and number of dives observed; mean dive time in the 30 first dives after tagging d_{30}, d_{60} is mean diving time for the 60 first dives, while the dives after the first 30 have mean dive time d_{30+}; number of 5-minute periods after the first 30 dives n, and mean and variance of the number of surfacings in these 5-minute periods.

[2] Heide-Jørgensen, M.-P., Kleivane, L., Øien, N., Laidre, K.L. and Jensen, M.V. 2001. A new technique for deploying satellite transmitters on baleen whales: Tracking a blue whale (Balaenoptera musculus) in the North Atlantic. *Marine Mammal Science* **17(4)**:949-954.

[3] Mori, Y. 2002. Optimal diving behaviour for foraging in relation to body size. *J.Evol.Biol.* **15**:269-276.

[4] Øien.N., Bøthun, G., Kleivane, L. 2003. Update on available data on surfacing rates of northeastern Atlantic minke whales. Paper SC/55/NAM 7 presented to the IWC Scientific Committee, May 2003 (unpublished). 7pp.

[5] Øien.N., Bøthun, G., Kleivane, L. 2007. Update on available data on surfacing rates of Northeastern Atlantic minke whales. Paper SC/59/RMP 6 presented to the IWC Scientific Committee, May 2007 (unpublished). 8pp.

[6] Okamura, H., Minamikawa, S., Kitakado, T. 2006. Effect of surfacing patterns on abundance estimates of long-diving animals. *Fisheries Science* **72**:631-638.

[7] Schreer, J.F., Kovacs, K.M. 1997. Allometry of diving capacity in air-breathing vertebrates. *Canadian Journal of Zoology/Revue Canadien de Zoologie* **75**:339-358.

[8] Schweder, T., Skaug, H.J., Dimakos, X.K., Langaas, M. and Øien, N. 1997. Abundance of northeastern Atlantic minke whales, estimates for 1989 and 1995. *Rep.int.Whal.Commn* **47**: 453-483.

[9] Skaug, H.J., Øien, N., Schweder, T. and Bøthun, G. 2004. Abundance of minke whales (Balaenoptera acutorostata) in the Northeast Atlantic: variability in time and space. *Can.J.Fish.Aquat.Sci.* **61**: 870-886

Model	Fs1	Ms1	Fs2	Ms2	Ms3	Ms4	Ms5
intercept			2.78	3.21	3.24	3.38	3.46
$x_1 = 1$	-1.26	-1.29	-1.25	-1.29	-1.26	-1.26	-1.28
$x_1 = 2$	-3.28	-3.61	-3.18	-3.62	-3.28	-3.52	-3.55
$x_1 = 3$	-4.07	-4.72	-3.54	-4.64	-4.05	-4.59	-4.59
$x_2 = 1$	-1.89	-1.71	-1.75	-1.64	-1.89	-1.67	-1.82
$x_1 : x_2 = 1 : 1$	0.16	-0.04	0.30	-0.02	0.16	-0.04	-0.01
$x_1 : x_2 = 2 : 1$	1.68	1.43	1.94	1.49	1.68	1.40	1.15
$x_1 : x_2 = 3 : 1$	2.41	2.15	2.48	2.17	2.40	2.08	2.22
sd(whale)	0.76	0.74					
sd(w)					0.55	0.66	0.63
sd(b)		0.77		1.00		0.77	0.68
Residual deviance	13673	13220	14272	13320			
df	11327	11326	11336	11335			
AIC	13707	13260	14288	13340			

Table 2: Logistic regression results for the probability of a dive being short. Fixed effects models (Fs1, Fs2) are fitted by glm in R. Mixed effects models Ms1 and Ms2 are fitted by maximum likelihood (glmmML in R). The remaining mixed effects models (Ms3, Ms4, Ms5) are fitted by restricted maximum likelihood (glmmPQL in R). The random batch effect is for batches of 30 dives in all but Ms5, where the batch size is 60. sd(b) is estimated standard deviation of random batch effect, sd(w) is estimated standard deviation of random whale effect, while sd(whale) is standard deviation of estimated fixed effects of whale. Fixed effects for whale not shown.

Model	Fl1	Ml1	Fl2	Ml2	Ml3	Ml4	Ml5
intercept			-0.89	-1.82	-1.29	-1.84	-1.93
$x_1 = 1$	1.44	1.88	1.61	1.92	1.44	1.65	1.79
$x_1 = 2$	2.89	3.10	2.93	3.19	2.88	2.89	2.96
$x_1 = 3$	3.15	3.52	2.84	3.50	3.14	3.26	3.38
$x_2 = 1$	0.09	1.10	0.07	1.11	0.09	0.93	0.90
$x_1 : x_2 = 1 : 1$	0.02	-0.40	-0.40	-0.53	0.01	-0.27	-0.36
$x_1 : x_2 = 2 : 1$	-0.82	-1.09	-1.40	-1.38	-0.82	-1.04	-0.99
$x_1 : x_2 = 3 : 1$	-1.14	-1.41	-1.40	-1.66	-1.15	-1.32	-1.26
sd(whale)	0.84	1.06					
sd(b)		1.25		1.47		1.01	0.98
sd(w)					0.79	0.95	0.97
Residual deviance	5270	5080	5820	5170			
df	4694	4693	4703	4702			
AIC	5300	5110	5840	5188			

Table 3: Logistic regression results for the probability of a dive being long given that it is not short. Otherwise same as previous table.

	0	1	2	3
0	0.036	0.047	0.046	0.022
1	0.282	0.276	0.170	0.121

Table 4: Joint relative frequency distribution of x_1 (in the columns) and x_2, n=11344.

number of short dives	0	1	2	3	4+
before	0.18	0.30	0.17	0.18	0.17
after	0.14	0.30	0.19	0.20	0.17

Table 5: Relative frequencies of number of successive short dives before and after a long dive, n=2807.

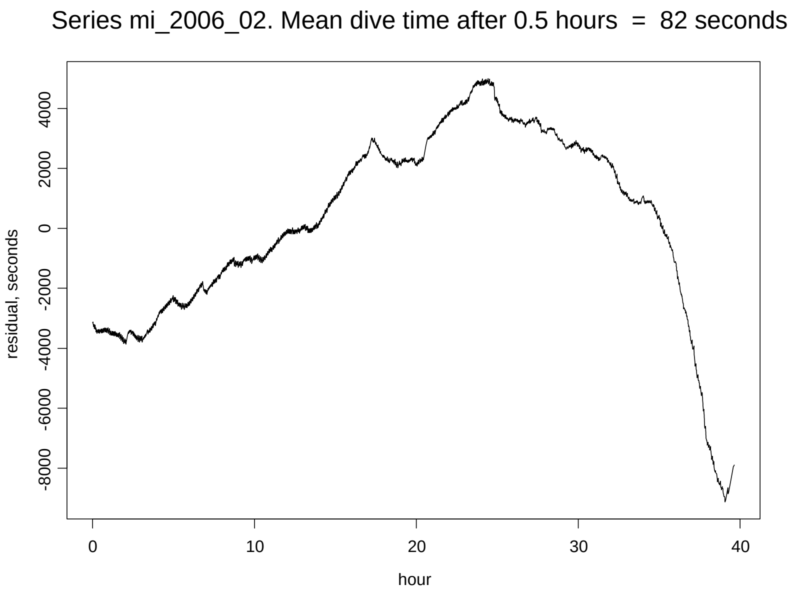

Figure 1: Linearly detrended graph of time of surfacing by number of surfacing for whale 10.

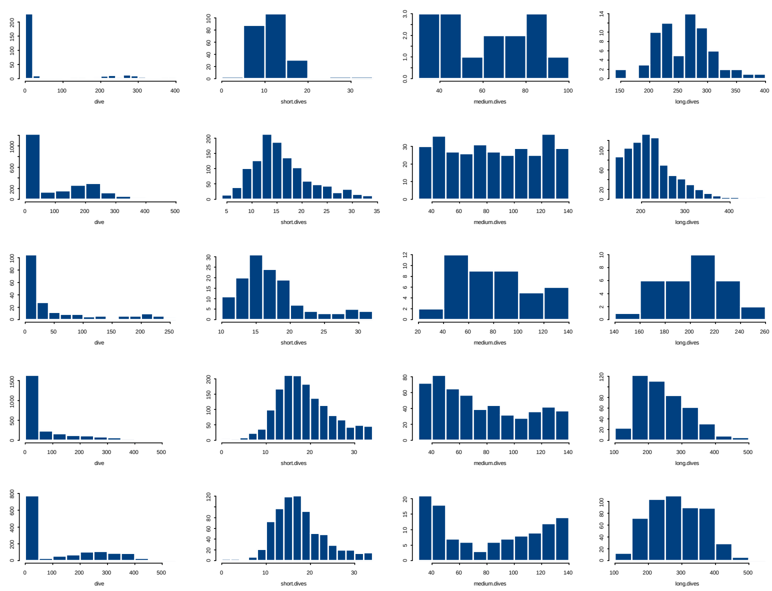

Figure 2: Histograms for all dives (first column), and for short (second column), medium and long dives, by series. Top row is for whale 1, etc.

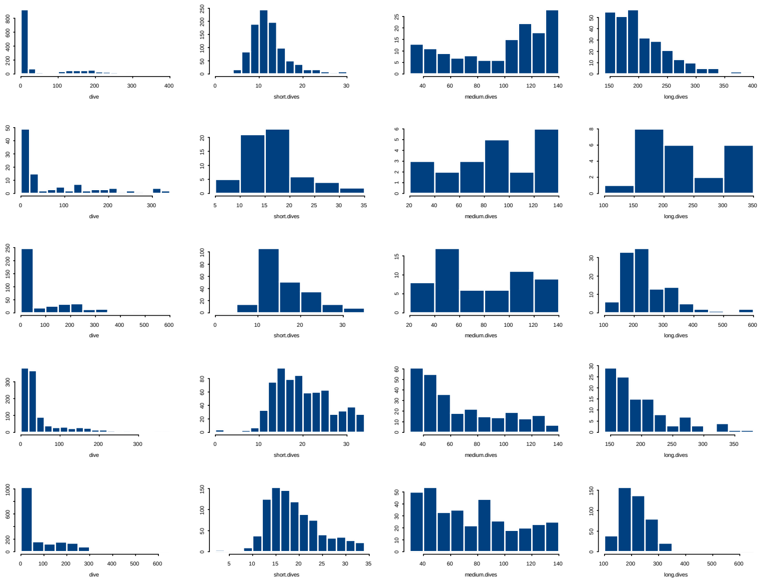

Figure 3: Histograms for whale 6 to 10, as in previous figure.

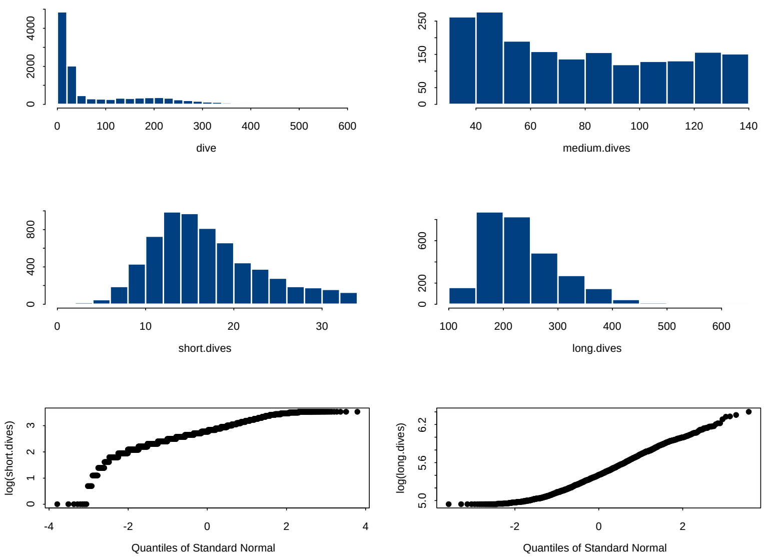

Figure 4: Histograms of pooled dive times (upper left); short dives, $d < 35s$ (middle left); medium dives, $35s \leq d \leq 140s$ (upper right); long dives $140s < d$ (middle right). The bottom panels show qq-plots of logarithmic short and long dive times respectively against the normal distribution.

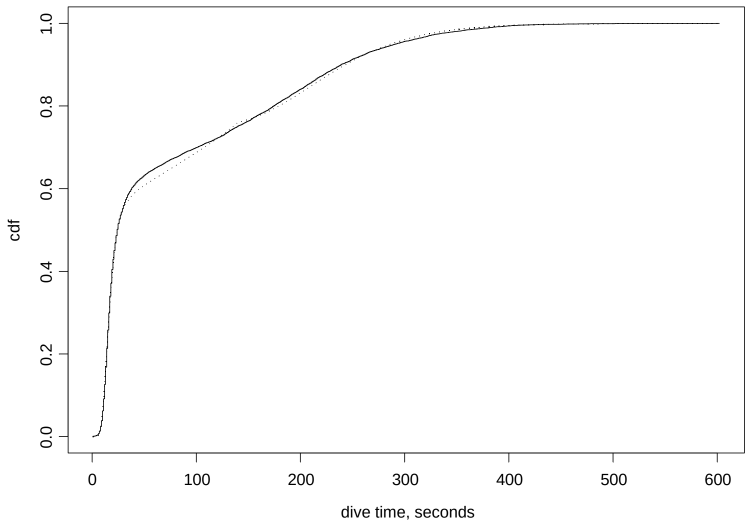

Figure 5: Empirical cdf for pooled dive times (solid line), and for a mixture of log normal ($\mu_1 = 2.79, \sigma_1 = 0.38$), $p = 0.58$, uniform $(35, 140)$, $q = 0.17$ and log normal ($\mu_2 = 5.42, \sigma_2 = 0.28$), $r = 0.25$.

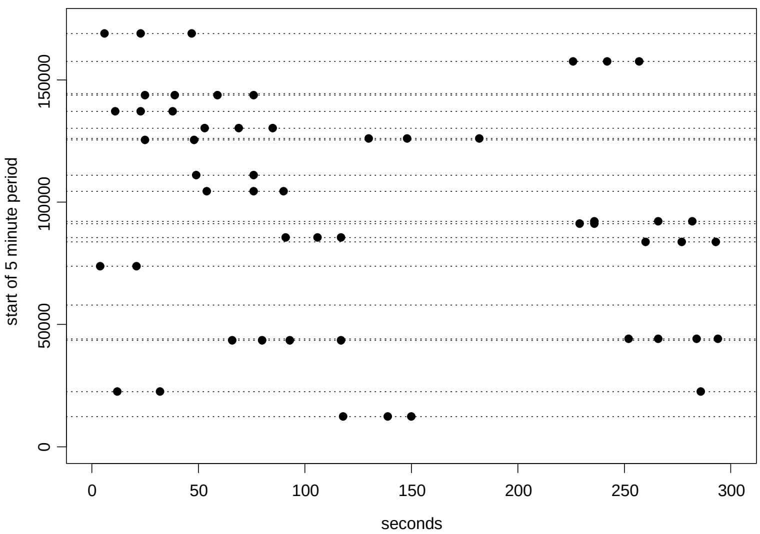

Figure 6: Surfacing times in 20 randomly chosen 5-minute periods for whale 5, series mi_2002_03.

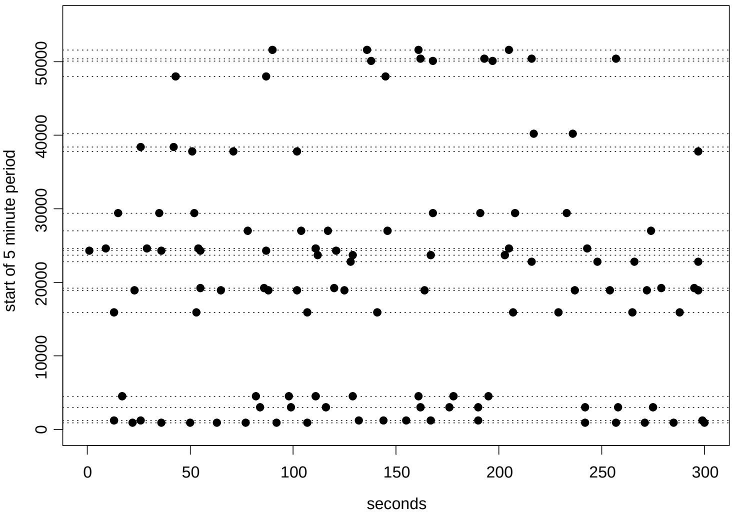

Figure 7: Surfcings in 20 randomly chosen 5-minute periods, whale 9 series mi_2006_01.

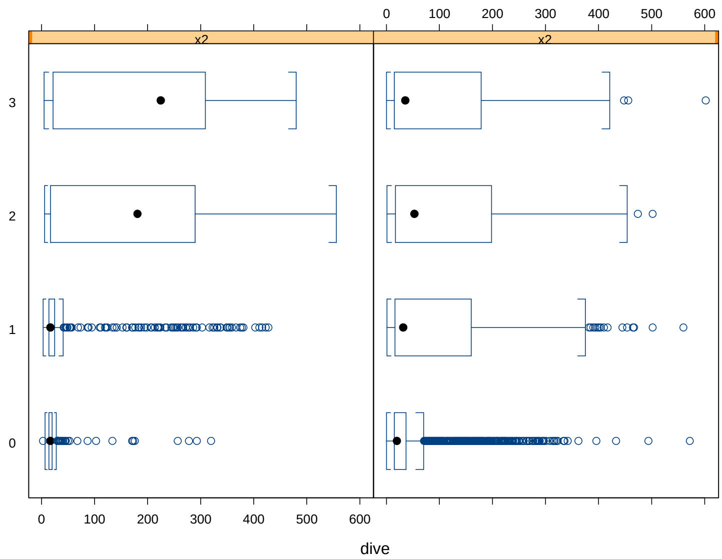

Figure 8: The pooled time dive times distributed by x_1 = number of surfacings in the previous minute, 3 or more coded as 3, and x_2 indicating whether there were any surfacings between 1 and 5 minutes previuos to the surfacing. The first 30 dives excluded from each series.

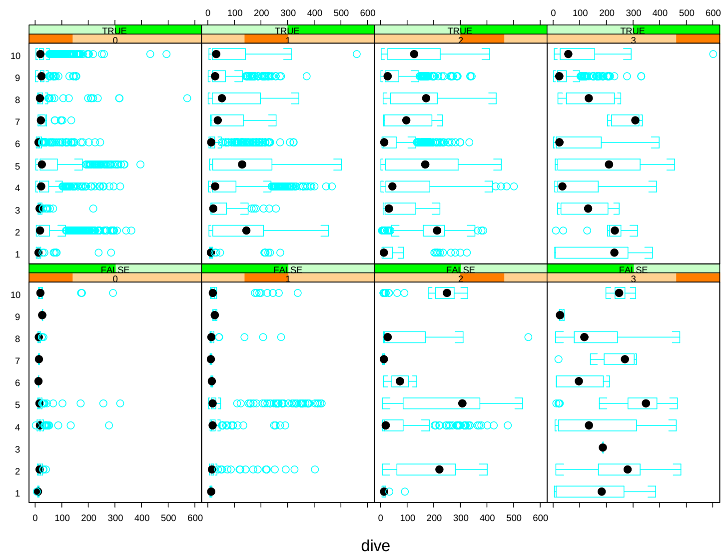

Figure 9: Dive times by whale and covariate levels.

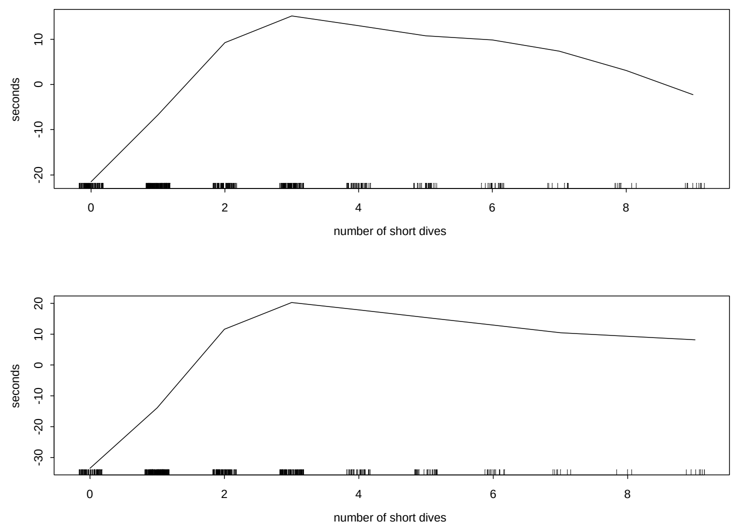

Figure 10: Effects on dive length by number of successive short dives before (upper panel) and after a long dive. Ten or more short dives is lumped into the 9+ cell. Gam analysis.