Fitting Models of the Population Consequences of Acoustic Disturbance to Data from Marine Mammal Populations

James S. Clark
H.L. Blomquist Professor of the Environment, Professor of Biology, Professor of Statistics
Nicholas School of the Environment
Duke University
Durham, NC 27708
phone: (919) 613-8036 fax: (919) 660-7425 email: jimclark@duke.edu

Robert S. Schick, Postdoctoral Associate
Nicholas School of the Environment
Duke University
Durham, NC 27708
phone: (626) 403-0728 fax: (919) 660-7425 email: rss10@duke.edu

Award Number: N000141010516

LONG TERM GOALS

To build an ecological modeling framework that facilitates understanding of the at-sea health of various species of marine mammals. We will use the results from these models to explore and quantify the impact of different types of disturbance (both environmental and anthropogenic) on these species. Modeling will in a Bayesian framework, which will allow us to fully account for uncertainty in the data, the biological processes, and in model output.

OBJECTIVES

Our scientific objectives are to build a statistical framework for understanding the at-sea health of (initially) three species of marine mammals: southern and northern elephant seals, and northern right whales. For elephant seals our goal is to build a hierarchical Bayesian model that provides daily estimates of lipid status, as lipid status of the mother is directly linked to pup survival. This model will use the drift dive behavior of elephant seals (Crocker et al. 1997) as the link to the underlying true, yet unmeasurable, lipid state.

We are just starting to analyze the right whale data, but our initial scientific objective is to build a demographic model that provides spatially and temporally explicit estimates of individual survival and calving.
### Title and Subtitle
**Fitting Models of the Population Consequences of Acoustic Disturbance to Data from Marine Mammal Populations**

### Author(s)
Duke University, Nicholas School of the Environment, Durham, NC 27708

### Performing Organization Name(s) and Address(es)

Duke University, Nicholas School of the Environment, Durham, NC 27708

### Distribution/Availability Statement
Approved for public release; distribution unlimited

### Security Classification of:
- a. Report unclassified
- b. Abstract unclassified
- c. This Page unclassified

### Limitation of ABSTRACT
Same as Report (SAR)

### Number of Pages
26

---

*Form Approved
OMB No. 0704-0188*

*Standard Form 298 (Rev. 8-98) Printed by ANSS Std Z9-18*
APPROACH

Work on the PCAD project at Duke University is led by Jim Clark with the assistance of one Postdoctoral Researcher, Rob Schick (started May 15, 2010). To date this work has focused primarily on the elephant seal drift dive/lipid status model. In addition to active work on this model, both Clark and Schick have participated in both PCAD working group meetings in 2010 (March in Santa Barbara, and September in Woods Hole). Clark and Schick also traveled in June to St. Andrews, Scotland to work the other members of the PCAD modeling sub-group: John Harwood, Len Thomas, and Leslie New. In addition to colleagues at St. Andrews, we are working closely with Mark Hindell, Clive McMahon, Dan Costa, Patrick Robinson (elephant seal colleagues), as well as Scott Kraus, Philip Hamilton, Peter Tyack, Bob Kenney (right whale colleagues).

Elephant Seals - Model Formulation

The model for at-sea lipid gain/loss in elephant seals contains two primary components: an observation model for the drift dives, and a process model for the lipid gain/loss process experienced at sea. The observation model is:

\[ D_{i,t} = \alpha_1 + \alpha_2 \frac{L_{i,t}}{R_{i,t}} + \varepsilon, \]

where \( D_{i,t} \) is the observed mean daily drift rate (in m/s) of individual \( i \), \( \alpha \) are the parameters governing the link between lipid status and drift, \( L_{i,t} \) is the estimated daily lipid status of individual \( i \) (in kg), and \( R_{i,t} \) is the fixed daily lean mass value of individual \( i \) (in kg). \( R_{i,t} \) is not currently estimated; instead we hold this value fixed in the model as follows. We use three observations for each individual: 1) lean mass at departure; 2) lean mass at arrival; and 3) pup mass. We interpolate a linear daily sequence from departure lean mass to (arrival lean mass – pup mass). At \( t = 2/3 \times T \), i.e. 2/3 of the way through the post-moult trip, we interpolate a linear daily sequence from \( R_{i,t=2/3} \) to \( R_{i,t=T} + \text{pup mass} \) (Figure 1). In this way we account for our basic understanding of the biology of the pregnant females.

The process model for lipid gain assumes the following form:

\[ L_{i,t+1} = L_{i,t} + X_{i,t}\beta + \varepsilon_{i,t}, \]

where \( L_{i,t} \) is the estimate of lipid content in the animal, and \( X_{i,t}\beta \) are the dynamic covariates and the parameters that govern the effect of covariates on lipid gain/loss. Examples of these covariates for southern can include, daily transit, number of drift dives, number of foraging dives, distance to ice edge, bathymetric depth, etc. For northerns we can extend this to include dynamic covariates at depth for additional biological variables like density of mesozooplankton, density of phytoplankton, etc. The full model summary is given by:

\[
p(\theta | X, Y, \text{priors}) \propto \prod \prod \prod N(D_{i,t}, | Z_{i,t}\alpha, \sigma_D^2)N(\alpha | V_\alpha, V_\alpha)IG(\sigma_D^2, | a_\sigma, b_\sigma) \]

\[
N(L_{i,t+1} | L_{i,t} + X_{i,t}\beta + \sigma_L^2)N(\beta | b, V_\beta)IG(\sigma_L^2 | a_{\sigma_L}, b_{\sigma_L}),
\]
where
\[ \theta = \left( L, \beta, \alpha_1, \alpha_2, \sigma_D^2, \sigma_L^2 \right) \]
are parameters and latent states to be estimated, \( X \) are covariates,
\[ Y = \left( L_D, L_A, D_i \right) \]
are inputs, \( Z_{i,t} \) is the ratio of lipid to lean mass, and \( L_D, L_A, R_D, R_A \) are the departure and arrival lipid and lean measurements for each seal.

In summary, the model provides estimates of the lipids time series (\( L \)), the regression parameters that govern lipid gain/loss (\( \beta \)), missing covariate data (\( X \)), the \( \alpha \) parameters governing the link between the lipid/lean ratio and the observed drift rate (\( D_{i,t} \)), the observation error (\( \sigma_D^2 \)), and the process error (\( \sigma_L^2 \)).

The Data
To date we have worked with a small ‘best’ subset of data from Mark Hindell’s group at University of Tasmania. The so-called ‘best’ group contains 1) the best delineations of candidate dives into labeled drift dives, i.e. the tags included swim-speed indicators, 2) each animal in the group has measured departure and arrival mass, and 3) each animal in the group has a measured pup weight. In geographic space, three approximate strategies are apparent: 1) a shelf-edge strategy; 2) a Ross Sea strategy; and 3) a pelagic strategy (Figure 2). A detailed example of one individual, in this case a Ross Sea animal, highlights the relationship between the number of drift dives and the buoyancy throughout the trip (Figure 3).

We are in the process of fitting the model to a subset of the northern elephant seal data (Figure 4), and it is interesting to note that on average, the drift rate signal is quite different. (We just show the subset of tracks that has the full suite of body composition data.) Most northern elephant seals seem to abruptly flip between negative and positive buoyancy and remain positively buoyant for the duration of the post-mount trip (Figure 5).

Elephant Seals - Fitting the Model to Data
Following methods outlined in (Clark 2007), we fit the model to data as follows. We 1) initialize the \( L \) and \( R \) time series for each animal, 2) setup the design matrix containing the covariates governing change in lipid, 3) set the prior distributions, and 4) initialize each of the Gibbs chains. Each Gibbs chain will contain posterior estimates of a given parameter at each step through the Gibbs sampler. The Gibbs sampler allows us to factor the above high dimensional model into a series of lower dimension conditional distributions. At each step through the Gibbs sampler, we sample new values for each parameter conditioned on all other parameters, as well as the current value of the parameter being estimated (Clark 2007). Using this approach over many steps through
the Gibbs sampler we build up the marginal posterior distribution for each of the parameters or quantities to be estimated.

We initialize a value for each of the parameters to be estimated, and then update each in turn at each step. We use graphical methods to check for convergence, discard burnin, and derive quantitative estimates of the parameters, e.g. median and 95% BCI, from the remaining sub-sample.

We have experimented with a variety of model formulations. Variations in formulation have included using discrete or continuous states in the model to account for different segments of the post-moult trip, i.e. migrating away from the colony, foraging, and returning to the colony. We have experimented with a variety of different covariates in the process model – both dynamic and static. Examples include: # drift dives, states, time since departure, depth, sea surface temperature, distance to ice edge, ice concentration (from remotely sensed data), distance to shelf break, and a body weight maintenance term.

To facilitate model selection, we calculate marginal model likelihood for each of these different models (Clark 2007). That is, we tally the likelihood of the data \( y \) given the set of parameters \( \theta_m \) associated with a given model \( M_m \). Given this probability, \( p(y \mid M_m) \), for each model, and a penalty for model size (the model prior), we find the model with the highest posterior probability.

Lastly we use the parameters \( \theta_m \) associated with a given model \( M_m \) to generate predictive intervals for a response \( y \) to a given covariate, i.e., we determine the predicted response of lipid gain/loss to changes in the covariates in model \( M_m \). This is a useful way to see the response between the covariates and the predicted process in a way that accounts for both parameter and model uncertainty (Clark 2007).

In summary, following model fitting exercises, we obtain the parameters as mentioned above as well as an index of model selection, and a way to plot the predictive distributions that account for uncertainty in the model and the parameters contained therein.

**Elephant Seals - Prior Sensitivity Analysis**

Informative priors are needed for process and observation error. We have run a 3 factorial sensitivity analysis to gain inference on the effect of varying the strength of these priors. We presented the results from this sensitivity analysis at the WHOI meeting, which engendered discussion on min/max at-sea weight loss or gain. Unfortunately the only physiological data on this process are from beach measurements (Costa et al. 1986; Deutsch et al. 1990; Deutsch et al. 1994; Crocker et al. 2001), so while we can use these to inform the prior, what happens at sea might be quite different. The working group felt that on average gain/loss of about 6 kg/day was biologically reasonable, with an upper limit of about 12 kg/day.
Right Whales – Initial Data Summary

Subsequent to the WHOI meeting, we have focused our initial efforts on exploring the available right whale data to ascertain their utility in answering the PCAD questions regarding health and population consequences. At the meeting we explored at least three possible research questions: 1) the effects of body status (or health) on population demography; 2) the effects of acute acoustic exposure to both playback sounds and shipping sounds on short term activity budgets; and (briefly) 3) the effect of conspecific attraction on moving right whales.

We have concentrated our initial data exploration efforts on #1, as this will also help inform the feasibility of #3. Peter Tyack’s group is doing initial processing of the dTag data to assess the feasibility of #2.

WORK COMPLETED

We have completed the following tasks:
1. Built the statistical model to estimate at-sea lipid status
2. Tested the model using simulated data
3. Fit the model to a subset of southern elephant seals
4. Fit the model to a subset of northern elephant seals
5. Sampled static and dynamic covariates for southern and northern elephant seals
6. Begun data exploration for northern right whales

RESULTS

One of the primary goals of the analysis is the at-sea estimate of daily lipid content, and based on feedback from the working group, we were able to successfully identify biologically meaningful results (Figure 6).

In addition to being biologically plausible, elephant seal experts were able to infer foraging strategies by merely looking at the lipids time series (Figure 7). What emerged from the initial analysis is that departure and arrival mass are not necessarily the best indicators of foraging success. Some animals depart the colony at similar weights, yet return at starkly different weights, e.g. c69902pm and f99305pm (Figure 6). Similarly, foraging strategy is not necessarily an indication of how well an animal will do. For example Ross Sea animals c16202pm, c16204pm, and c16205pm all have similar Cartesian paths to/from the Ross Sea, yet their at-sea lipid gain/loss process is very different (Figure 6). c162 is the same animal, tagged in two different years (2002, 2004), and exhibits markedly similar lipids time series, both in terms of departure/arrival lipid content as well as at sea gain and loss (Figure 6). This fits with current knowledge of foraging strategies across years among elephant seals from this colony (Bradshaw et al. 2004).

While the estimates are closely tied to the drift and measured lipids data (Figure 6), model results are sensitive to prior knowledge, and the less informed the prior, the wider the confidence intervals are (Figure 8). We are waiting on a regression equation from
Mark Hindell that will help inform this relationship slightly, especially for estimates of observation error around neutral buoyancy.

Using the marginal model likelihoods, we arrived at the following set of covariates for southern elephant seals: transit, # drift dives (linear and quadratic), body weight penalty ($L/R$), and time. Using this model, we derived the predictive distributions, and the transit and the # drift dives covariates have a stronger relationship with predicted lipid change $y$ (Figure 9).

**Right Whales – Initial Data Summary**

We have had several discussions with Philip Hamilton (NEAq) and Bob Kenney (URI) about the Photo-ID Catalog and the sightings database. While the Photo-ID data were summarized in the PCAD/WHOI summary document, visual exploration of them has yielded a better understanding of how the data are distributed across individuals, survey platform, survey region, and time.

**Sightings Data (from URI)**

The sightings data, broadly speaking, encompass all spatial sightings of right whales as well as effort (aerial or ship-based) that went into the collection of sightings. As of March 2010, there are 3.2 million records in this database, 55,000 of which are right whale sightings. The date range is 1762-2009; they exist for parts of 2010, but database integration is still ongoing. The spatial range is essentially the entire North Atlantic Ocean, yet the overwhelming majority of sightings are in the nearshore waters of the Northwest Atlantic Ocean from Florida to Nova Scotia.

There are three main types of data available: 1) Line-transect Surveys; 2) Platform of Opportunity Surveys (POP); and 3) Opportunistic Sightings. Both of the survey types (line-transect, and POP) are included in the SPUE calculations. Effort is derived by calculating the linear length of time-specific survey lines within each grid cell. SPUE is simply the number of right whale sightings per amount of effort. Visually the data show effort over a broad spatial range with accompanying sightings in known hotspot areas (Figure 10), which progress over the year (Figure 11). Surveys, and sightings are in the ‘sightings’ database, and include both on and off-effort sightings from Line-transect surveys as well as POP surveys (Figure 12).

**Catalog Data (From NEAq)**

The Photo-ID data are curated by NEAq, and are quantitatively summarized in the working group summary from the WHOI meeting. Graphically, we can see the extent of the Photo-ID data by tallying sightings over space and time (Figure 13, Figure 14). Lastly, there are several ways to view the data at the individual level. One way is as a heat map of sightings per area aggregated over time (Figure 15), which at present is uncorrected for effort. A second way is to simply show the individual resights for an animal in Cartesian space, also uncorrected for effort (Figure 16).

**Calving Data**
The calving data provided by Philip Hamilton at the WHOI meeting clearly show the low-calving event that occurred in the late 1990’s (Figure 17).

**Health Data**

The health data are similar in structure to the calving data in that unique ‘health-status’ assignments occur at unique intervals. As enumerated by Scott Kraus at the meeting, pictures are collected for an animal over a unit of time (this unit usually corresponds to a specific habitat). These pictures are used to assess and score the health of the animal. There are three health levels, not-thin, thin, very thin, with the 5,104 sightings distributed as follows: not-thin (4,123), thin (962), very thin (19). Philip made clear that fat individuals, i.e. ‘not-thin,’ are easy to see and discern; the ‘very thin’ ones are also easy to see, and are often the last sighting of that individual - either known dead or are often correlated with a presumed dead status.

An example plot depicting health for Adult males and females (Figure 18) highlights a few relevant biological issues. First, within any given habitat males are typically seen less than females because they are more nomadic. Second, females are generally thinner owing to the demands of lactation.

**IMPACT/APPLICATION**

To date, no one has been able to estimate the at-sea lipid status of foraging seals. Though previous efforts have used the drift dive signal in relation to environmental features (Biuw et al. 2007; Robinson et al. 2010), this is the first attempt to actually estimate the hidden physiological process of lipid gain/loss in pregnant females. Because the link between arriving maternal mass and future juvenile survival is strong (McMahon et al. 2000), this represents a significant advance in our understanding of the at sea physiology of elephant seals.

For right whales, while previous efforts have quantified survival in this population (Caswell et al. 1999; Fujiwara and Caswell 2001), these attempts have neither been Bayesian nor spatial. We feel that within the PCAD project goals, understanding where and when animals do better/worse will represent a significant contribution.

**RELATED PROJECTS**

None

**REFERENCES**


Figure 1. Cartoon depicting the assumption about lean mass in the model. We start with a straight line between measured values of departure lean mass ($R_D$) and arrival lean mass ($R_A$). 2/3 of the way through the trip, we interpolate between $R_{2/3}$ and $R_A +$ pup mass.
Figure 2. Geographic extent of the southern elephant seal data colony at Macquarie Island (blue dot). Each dot along the path represents a mean daily position, and is colored according to the mean drift rate observed for all drift dives within the day. Redder dots indicate higher positive buoyancy. Shelf-edge animals are those at the lower left of the map; Ross Sea animals are at the lower right. Pelagic animals are those animals with considerable time in open ocean, e.g. the animal at the top left.
Figure 3. Detailed example of one track - in this case an animal that forages almost exclusively in the Ross Sea. Color is as in Figure 2; size of each daily position reflects the number of drift dives recorded on that day. Inset map is simply the drift dive time series for this same animal with the same color scheme.
Figure 4. Map of northern elephant seal tracks to which we have preliminarily fit the model. These are exclusively from Año Nuevo (northerly blue dot), though the fuller set contains Isla San Benito tracks as well (southerly blue dot).
Figure 5. Example path and drift rate time series for one animal tagged at Año Nuevo in 2004. Color and symbol scheme is as previous. Note how the animal suddenly gets positively buoyant around 120 days at sea and then stays buoyant for the duration of the trip.
Figure 6. Posterior estimates of the lipid time series for each of the 11 adult female elephant seals tagged at Macquarie Island. Solid black line is the median estimate; dashed lines are the 95% confidence limits.
Figure 7. Four-panel view of one seal shows (clockwise from top-left): 1) the geographic data color coded by drift rate; 2) the start-end weights of this (red) and all other seals; 3) the estimates of the lipids; and 4) the drift rate time series symbol coded by # drift dives/day. Prior to seeing this plot, Mark Hindell correctly inferred the foraging strategy of this animal simply by looking at the lipids time series.
Figure 8. Results from a prior sensitivity analysis for one seal (c16104pm - a 'shelf-edge' animal). The center panel uses the same parameters as in Figure 6. Arrows in the plot indicate the change in prior knowledge for observation error (vertical) and process error (horizontal).
Figure 9. Predictive distributions depicting the relationship between the covariates and the lipid change that accounts for both parameter and model uncertainty. Note transit and # drift dives have the strongest relationship with lipid gain/loss; (Robinson et al. 2010) have noted transit is the strongest predictor of change in drift rate for northern elephant seals.
Figure 10. Sightings Per Unit Effort (SPUE) data for one year (2000) across the entire surveyed range. Small grey pixels indicate effort in a cell (10') but no sightings.
Figure 11. The same data as Figure 10, color coded by month to see the spatial progression over the year. The approximate habitat use progression is CCB --> GSC --> BOF --> RB ---> SEUS. Note that the last connection is dashed because some animals will not go to the SEUS (note, for example, one sees wintertime animals in CCB and the SEUS). Recently, Jordan Basin has been found to be of increased importance, and paternity information suggests that this is the actual breeding ground (not BOF as previously thought) (Frasier et al. 2007). Roseway Basin is an important habitat that seems to alternate in importance with BOF depending on food presence.
Figure 12. Example sightings derived from flight and ship surveys. The on-effort sightings along the with space-time tally of survey lines contribute to the SPUE calculations mapped in Figure 10 and Figure 11. Note that there are 5 different surveys depicted here, but I’ve only color-coded the sightings – not the lines as well.

You can graphically see the distinction between line-transect surveys, and POP surveys. Here sightings observed on p100133 were from a targeted right whale survey from a plane, but were not part of a density estimation procedure.
Figure 13. Sightings by year and region. One biological pattern of note was the apparent 'abandonment' of RB in the mid to late 1990's, which is thought to have been driven by food limitation out on the Scotian Shelf.
Figure 14. Though this misses some of the interannual signal, for example whales leaving RB for BOF, it does show the approximate progression of animals through the main regions across their range. One habitat not represented here, but of increasing importance is Jordan Basin, where it’s thought whales go after BOF. This is not an easily accessible area of the GOM, and this signal is being picked up through opportunistic sightings and paternity analysis (Frasier et al. 2007).
Figure 15. Sightings 'heat map' for adult females in the major regions. Darker blue colors represent more sightings (aggregated across multiple years and uncorrected for effort). Note the two examples of 'non-Bay of Fundy' females in this plot (1007, 1204); these are females who have been sighted on more than one occasion with a calf in SEUS, but never in BOF.
Figure 16. Sample plot of movements and health status through space and time, respectively. EGNo 1507 is an adult male, first seen as a calf in 1985 (note sighting off Georgia), and last seen in June 2009 (n\text{total}=353). Sightings in top panel are aggregated across years, but colored by month to depict progression across habitats.
Figure 17. All calving events for individual cows, post 1970. Each calving event with a known interval is colored with longer intervals being redder. The relative lack of calves in the late 1990’s is immediately apparent.
Figure 18. Health status for adult right whales (males and females). Females in general are thinner than males owing to lactation requirements. The apparent influx of males (i.e. many more sightings) in the late 1990's is likely due to the BOF/RB shift. Size of the circle indicates the number of pictures used to make the assessment. Females are typically resighted more, because in any given habitat they are more resident than the males.