

**Len Thomas and John Harwood**

**A comparison of grey seal population models incorporating density dependent pup survival and fecundity**

NERC Sea Mammal Research Unit and Centre for Research into Ecological and Environmental Modelling,  
University of St Andrews, St Andrews KY16 8LB

NOTE: THIS PAPER AND ITS CONTENTS SHOULD NOT BE REFERENCED WITHOUT PRIOR PERMISSION OF THE AUTHORS

**Summary**

We fitted state-space models for density dependent pup survival and density dependent fecundity to regional estimates of British grey seal pup production. The results provide inconclusive evidence for one model over the other, but this could be because of limitations in the fitting algorithm used. Nevertheless, the results illustrate the very large difference in estimated total population size in 2003 under the two models: 96,000 under the density dependent pup survival model and 394,000 under the density dependent fecundity model.

**Introduction**

In this paper, we compare two spatially-explicit, stochastic models for density dependent population regulation in British grey seals. In the first, the density dependent parameter is first-year survival, while in the second it is fecundity. Our models are fit to the 1984-2003 annual pup production data at the level of region (North Sea, Inner Hebrides, Outer Hebrides and Orkney). Both models also include fitness-dependent movement of recruiting females between regions. The fitting algorithm is a modification of the computer-intensive Bayesian procedure presented at the last SCOS meeting (Thomas and Harwood 2003).

Our primary goals are: (1) to determine whether the pup production data alone provide evidence for one model over the other, and (2) to quantify differences in estimated total seal numbers under the two models. An alternative approach to question (1) using auto-regression (Thomas et al. 2004) found that the annual rate of change in pup production was negatively affected by pup production in the previous year, as would be predicted if density dependent fecundity were the dominant mechanism.

**Materials and Methods**

*Models*

The density dependent survival (DDS) model is identical to that presented in Thomas and Harwood (2003), and the density dependent fecundity (DDF) model is a small modification of it. Both models are formulated as state-space models (Buckland *et al.* 2004). In essence, this means that they are composed of a state process, which models the true but unknown state of the population (i.e., the number of animals in each age group and region in each time period), and an observation process, which models how the survey data are generated given the true states.

In constructing the state processes, we divide the seal population in each region into 7 age classes: pups (age 0), age 1 – age 5 adult females (pre-breeding), and age 6 and older females. Note that our models do not include adult males.

The time step for the process models is 1 year, beginning just after the breeding season. The models are made up of four sub-processes: survival, age incrementation, movement of recruiting females and breeding.

Survival is modelled as a binomial random process. For the DDS model, we assume that pup survival follows a Beverton-Holt function of the form:

$$\phi_{p,r,t} = \frac{\phi_{p\max}}{1 + \beta_r n_{0,r,t-1}}$$

where  $n_{0,r,t-1}$  is the number of pups born in region  $r$  in year  $t-1$ ,  $\phi_{p,r,t}$  is survival rate of these pups,  $\phi_{p\max}$  is maximum pup survival rate, and  $1/\beta_r$  reflects the carrying capacity of the region. For the DDF model, we assume pup survival is constant across regions and times, i.e.,  $\phi_{p,r,t} = \phi_p$ .

Since half of the pups born will be male, the expected number of female pups surviving in

both models will be  $0.5 \phi_{p,r,t} n_{0,r,t-1}$ . For both models, we assume that adult female survival rate,  $\phi_a$  is constant across regions and time.

Age incrementation is deterministic – all seals age by one year (although those in the age 6+ category remain there).

To model movement, we assume that only females breeding for the first time may move from their natal region. Once a female has started breeding she remains faithful to that region. We assume that movement is fitness dependent (Ruxton and Rohani 1998), such that females will only move if the value of the density dependent parameter (pup survival or fecundity) is higher elsewhere, and the probability of movement is proportional to the difference in the density dependent parameter between regions. In addition, we assume that females are more likely to move among regions that are close together, and that females show some degree of site fidelity – that is, they may not move even if conditions for their offspring will be better elsewhere. We model movement from each region as a multinomial random variable where probability of movement from region  $r$  to region  $i$  at time  $t$  is:

$$\rho_{r \rightarrow i,t} = \begin{cases} \frac{\theta_{r \rightarrow i,t}}{\sum_{j=1}^4 \theta_{j \rightarrow i,t}} & : \sum_{j=1}^4 \theta_{j \rightarrow i,t} > 0 \\ I_{i=r} & : \sum_{j=1}^4 \theta_{j \rightarrow i,t} = 0 \end{cases}$$

where  $I_{i=r}$  is an indicator that is 1 when  $i=r$  and 0 otherwise, and

$$\theta_{r \rightarrow i,t} = \begin{cases} \gamma_{sf} & : i = r \\ \frac{\gamma_{dd} \max(\Delta_{i,r,t}, 0)}{\exp(\gamma_{dist} d_{r,i})} & : i \neq r \end{cases}$$

where  $\gamma_{sf}$ ,  $\gamma_{dd}$ , and  $\gamma_{dist}$  are three movement parameters that index the strength of the site fidelity, density dependence and distance effects respectively,  $\Delta_{i,r,t}$  is the difference in the density dependent parameter between regions  $i$  and  $r$  (see below), and  $d_{r,i}$  is the 20% trimmed mean of the distances between colonies in regions  $r$  and those in region  $i$  (standardized so that the largest distance is 1.0). For the DDS model,

$$\Delta_{i,r,t} = \phi_{p,i,t} - \phi_{p,r,t}$$

while for the DDF model,

$$\Delta_{i,r,t} = \alpha_{i,t} - \alpha_{r,t}$$

where  $\alpha_{r,t}$  is the fecundity rate in region  $r$  at time  $t$ , as defined below.

We model breeding by assuming that the number of pups produced is a binomial random variable, with rate  $\alpha_{r,t}$ . For the DDS model, we assume this value is constant across regions and times, i.e.,  $\alpha_{r,t} = \alpha$ . For the DDF model, we assume this value follows a Beverton-Holt function of the form:

$$\alpha_{r,t} = \frac{\alpha_{max}}{1 + \beta_r n_{0,r,t-1}}$$

This implies that the probability of a female breeding in a particular year is influenced by the conditions during the previous breeding season.

For the observation process, we assume that pup production estimates follow a normal distribution with a constant coefficient of variation (CV) which we assume to be a known value.

In summary, both models contain 10 parameters. The models share 8 parameters: adult survival  $\phi_a$ , one carrying capacity parameter for each region  $\beta_1 - \beta_4$ , and three movement parameters  $\gamma_{sf}$ ,  $\gamma_{dd}$ , and  $\gamma_{dist}$ . They differ in two parameters: the DDS model has maximum pup survival  $\phi_{pmax}$  and constant fecundity  $\alpha$ , while the DDF model has constant pup survival  $\phi_p$  and maximum fecundity  $\alpha_{max}$ .

#### Data and Priors

Our input data were the pup production estimates for 1984-2003 from Duck (2004), aggregated into regions. Some new colonies have been added to the regional totals, as described by Duck (2004). These made only minor differences to the estimates of pup production for the years 1984-2002, compared with the data used in Thomas and Harwood (2003).

We also included the Helmsdale colony in the North Sea region. This colony was previously excluded from the analysis, but the 2003 pup production estimate of 947 was too big to ignore. Helmsdale is not surveyed every year, so we estimated production for missing years by fitting a loess smooth (function `loess` in SPlus, with span set to 0.75) to the years where data were available and predicting the missing values.

Prior distributions for each parameter are given in Table 1, and are shown on Figures 3 and 4. Prior distributions for the states were generated using the priors for the parameters in conjunction with the 1984 data, as described by Thomas *et al.* (in press). We then fit the models to the data for 1985-2003.

Table 1. Prior parameter distributions

Parameter	Prior	Expected value
$\phi_a$	Beta(22.05,1.15)	0.95
$\phi_{p\max}, \phi_p$	Beta(14.53,6.23)	0.7
$\beta_1$	Gamma(4,2.07x10 <sup>-4</sup> )	8.29x10 <sup>-4</sup>
$\beta_2$	Gamma(4, 2.96x10 <sup>-4</sup> )	1.18x10 <sup>-3</sup>
$\beta_3$	Gamma(4,7.40x10 <sup>-5</sup> )	2.96x10 <sup>-4</sup>
$\beta_4$	Gamma(4,5.76x10 <sup>-5</sup> )	2.30x10 <sup>-4</sup>
$\gamma_{sf}$	Gamma(2.25,1.33)	3
$\gamma_{dd}$	Gamma(2.25,0.49)	ln(3)
$\gamma_{dist}$	Gamma(2.25,0.22)	0.5
$\alpha, \alpha_{\max}$	Beta(22.05,1.15)	0.95

*Fitting Method*

We used a modified version of the particle filtering algorithm that was outlined in Thomas and Harwood (2003) and described in detail in Thomas *et al.* (in press). The algorithm (which is also called sequential importance sampling or SIS) is a computer-intensive method for estimating the posterior distribution of the parameters and states of a state-space model. It is well suited to the analysis of time series data, as data point are introduced one at a time into the algorithm, making it potentially more efficient than other computer-intensive techniques such as Markov chain Monte Carlo (MCMC). Particle filtering methods were first developed for engineering applications and have only recently been applied to biological problems. Consequently, much methodological work is still required.

Our alterations to the 2003 algorithm were designed to address two issues: (1) high Monte Carlo variation in results (i.e., differences in estimates of parameters and states between multiple runs using different random number seeds) caused by particle depletion (see Thomas and Harwood 2003); (2) possible biases in estimated state values caused by kernel

smoothing at each time period. The new algorithm is simple and reliable but inefficient, because a large amount of computer time is required to produce reliable estimates.

We start by defining prior distributions on the parameters and the states (i.e., the numbers of seals in each region and age class before the first time period). We simulate a large number parameter and state vectors from these priors. Each pair of parameter and state vectors is called a ‘particle’. We stochastically project each particle forward to the first time period using the state process (i.e., our model of the population dynamics), and calculate the likelihood of the simulated pup production generated for each particle, given the observed pup production in the first year and the observation model. These likelihoods form weights for each particle. Many of these weights are very small, indicating that the starting state and parameter combinations are unfeasible, given the observed pup production in the first year. To avoid wasting computer memory and time on these very unlikely particles, we implement rejection control (Lui 2001), which probabilistically discards particles below a critical value (in our case we used the mean of the particle weights), and reweights the remaining particles so that they estimate the same distribution as before the rejection control step. This reduced set of particles is then projected forwards to the end of the time series, and the weights re-calculated given the likelihood of the pup production estimates for the second and subsequent years of data. Many of these weights will be very small, so we again implement rejection control to reduce the number of particles that need to be stored. The distribution of particles at the end is a weighted estimate of the posterior distribution of the parameters and states.

We can calculate the effective sample size of the remaining particles as

$$ESS = \frac{N_T}{1 + [CV(w)]^2}$$

where  $N_T$  is the number of particles after rejection control and  $CV(w)$  is the coefficient of variation of the weights of these particles. We have found that for these models, reliable inferences require an ESS of around 1000. This is rarely achieved in one run of the above algorithm, but there is no problem in adding to the final particle numbers by making multiple runs, so long as the same critical values are used

for rejection control in all runs (i.e., the mean particle weights from the first run). In the results reported here, we used 50 runs each starting with 100,000 particles, giving an ESS of 893 for the DDS model and 914 for the DDF model. The observation CV was fixed to a relatively high value (25%) to avoid a prohibitively large number of runs

#### *Model outputs and comparison*

For both models, we present posterior estimates of the model parameters and estimated pup production from 1984-2003. The models also estimate adult female numbers, but do not include adult males. We therefore calculated total pre-breeding population sizes by assuming that the number of adult males is 73% of the number of adult females (Hiby and Duck, unpublished).

To compare the models, we calculated the mean posterior Akaike Information Criterion (MPAIC) using the same method as Thomas and Harwood (2003). Since the two models have the same number of parameters, the ranking according to this criterion will be the same as that using the mean posterior likelihood.

## Results

Posterior estimates of true pop production for the two models are shown in Figures 1 and 2 (these estimates are known technically as smoothed estimates; see Thomas *et al.* in press). The estimates are very similar, although the DDS model is a slightly better fit (Table 2). There is some evidence of poor choice of prior for the Outer Hebrides region in the DDF model, causing some initial oscillations in pup production numbers (see Discussion). Neither model does a good job of tracking the rapid increase in pup production in the Inner and Outer Hebrides in the late 1980s and early 1990s, and the subsequent stabilization in pup numbers in both regions. In this sense, neither model can be said to fit the data well.

Although the models produce almost identical estimates of pup production, they give substantially different estimates of total pre-breeding population size, with the estimates from the DDF model being four times higher than those from the DDS model (Table 3).

*Table 2. Mean posterior log-likelihood, AIC and Akaike weights*

Model	LnL	AIC	Akaike weight
DDS	-658.8	1337.5	0.68
DDF	-659.4	1339.0	0.32

*Table 3. Estimated size, in thousands, of the British grey seal population at the start of the 2003 breeding season, from density dependent survival (DDS) and fecundity (DDF) models. Numbers are posterior means with 95% confidence limits in brackets.*

Region	DDS	DDF
North sea	11.7 (9.6-14.7)	46.6 (33.1-61.8)
Inner Hebrides	8.3 (6.8-10.2)	37.5 (25.7-49.3)
Outer Hebrides	31.4 (25.2-39.3)	152.0 (102.4-209.4)
Orkney	44.9 (25.5-56.6)	157.4 (113.1-209.7)
Total	96.2 (77.1-120.8)	393.6 (274.4-530.2)

Posterior parameter estimates for the models are given in Figures 3 and 4. Adult survival ( $\phi_a$ ) is estimated to be 0.98 for the DDS model and 0.99 for the DDF model, rather higher than the prior of 0.95. The juvenile survival and fecundity parameters ( $\phi_j$  and  $\alpha$ ) are almost unchanged relative to the prior in both models. Similarly, the movement parameters ( $\gamma$ s) are little changed. Posterior distributions of the density dependence parameters ( $\beta$ s) are slightly tighter than the priors, and are similar between models.

## Discussion

For the runs reported here, we fixed the CV of the pup production estimates at 25%. This value is much higher than that estimated by Hiby and Duck (unpublished) of 7% at the colony level, and less for regional aggregations. The effect of using a higher CV is to reduce the influence of the data on the posterior states and parameter, relative to the priors. We therefore regard our results as preliminary, pending improvements in the fitting algorithm.

We are actively working on improving the fitting methods. The current algorithm is simple (and therefore reliable) but inefficient. The relatively high CV set on pup production produces a

relatively flat likelihood surface and this results in relatively little particle depletion. The current algorithm ran for 17 hours on our machine; if the CV had been set at 10%, it would have required approximately two weeks to achieve the same precision. We expect to be able to improve efficiency, while at the same time maintaining reliability, using tools such as auxiliary particle filtering, simulated annealing and tempering, and limited kernel smoothing (Doucet *et al.* 2001, Lui 2001, Newman *et al.* in press, Thomas *et al.* in press). We are also working with K. Newman and C. Fernandez on an MCMC implementation of the DDS model, which we hope to compare with the particle filtering algorithm.

One other caveat regarding our results is the oscillations in estimated pup production in the Outer Hebrides under the DDF model. We suspect this is due to the way we set priors on initial states for this model, and plan to investigate this further.

The populations estimates from the DDS model are comparable to those produced last year (Thomas & Harwood 2003). However, those from the DDF model are very much higher because of the low levels of fecundity that are predicted in the final years of the time series. It seems clear that these models cannot be distinguished using time series of pup production estimates aggregated at a regional level.

Neither model provides a particularly convincing fit to the time series of pup production, and they predict levels of survival or fecundity that are very much lower than those that have been observed at individual grey seal colonies. The poor fit may be a result of the specific functional form used to model density dependence in both models. In the case of the DDS model it is based on empirical observations. However, there are no data that can be used to specify alternate forms for the DDF model. In addition, we show elsewhere (Thomas and Harwood 2004) that the observed changes in pup production could also be a consequence of an increase in the numbers of seals being killed to protect salmon farms. However, the numbers that would be required seem unfeasibly high. Until these inconsistencies have been resolved, there will remain considerable uncertainty about the current size of the British grey seal population.

## References

- Buckland, S.T., K.B. Newman, L. Thomas and N.B. Koesters. 2004. State-space models for the dynamics of wild animal populations. *Ecological Modelling* 171: 157-175.
- Doucet, A., N. de Freitas and N. Gordon [eds.]. 2001. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York.
- Duck, C.D. 2004. Grey seal pup production in Britain in 2003. *SCOS Briefing Paper 04/2*.
- Hiby, L. and C.D. Duck. Unpublished. Estimates of the size of the British grey seal *Halichoerus grypus* population and levels of uncertainty.
- Lui, J.S. 2001. *Monte Carlo Strategies in Scientific Computing*. Springer-Verlag, New York.
- Newman, K.B., S.T. Buckland, S.T. Lindley, L. Thomas and C. Fernández. In press. Hidden process models for animal population dynamics. *Ecological Applications*.
- Ruxton, G.D. and P. Rohani. 1998. Fitness-dependent dispersal in metapopulations and its consequences for persistence and synchrony. *Journal of Animal Ecology* 67: 530-539.
- Thomas, L., T. van Lamsweerde and J. Harwood. The nature of density dependence in British grey seal populations. *SCOS Briefing Paper 04/4*.
- Thomas, L. and J. Harwood. 2003. Estimating grey seal population size using a Bayesian state-space model. *SCOS Briefing Paper 03/3*.
- Thomas, L. and Harwood, J. 2004. Possible impacts on the British grey seal population of deliberate killing related to salmon farming. *SCOS Briefing Paper 04/7*
- Thomas, L., S.T. Buckland, K.B. Newman, and J. Harwood. In press. A unified framework for modelling wildlife population dynamics. *Australian and New Zealand Journal of Statistics*.

Figure 1. Estimates of pup production from the density dependent survival model (DDS). Input data are shown as circles, while the lines show the weighted mean of the particle values, bracketed by 2.5th and 97.5th percentiles.

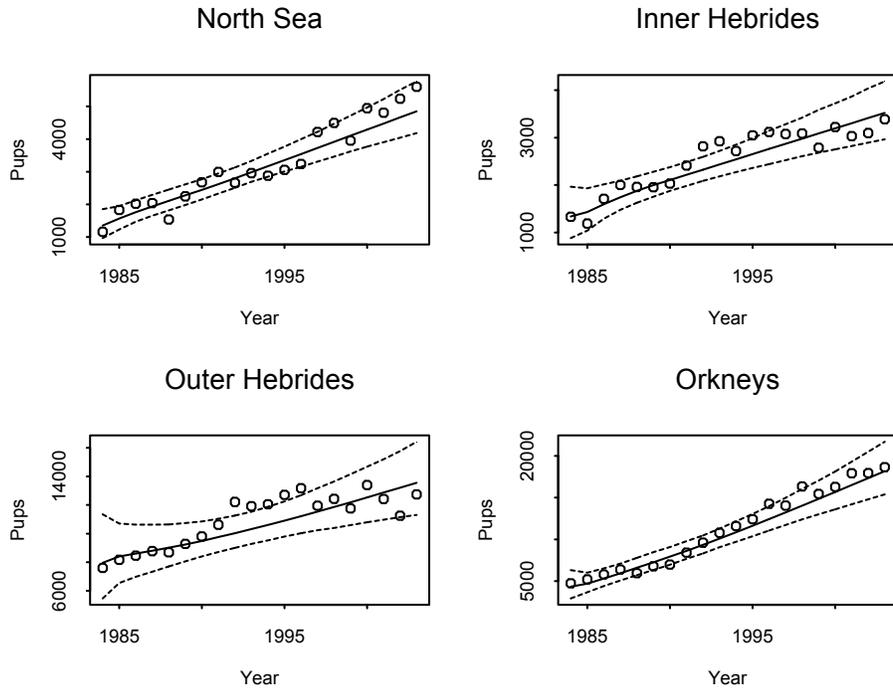


Figure 2. Estimates of pup production from the density dependent fecundity model (DDF).

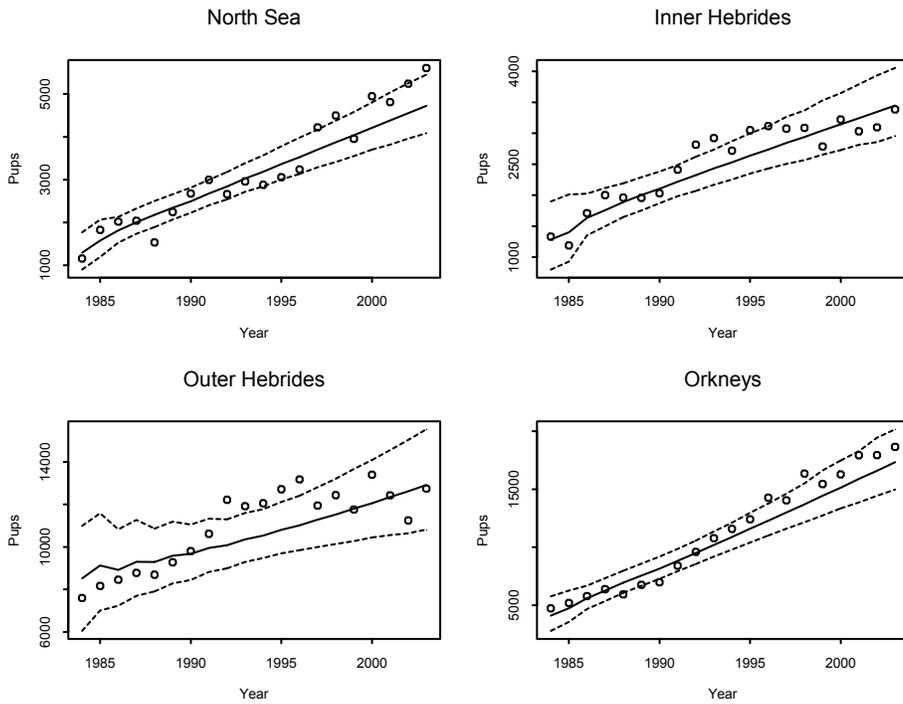


Figure 3. Posterior parameter estimates (histograms) and priors (solid lines) from the density dependent survival model (DDS). The vertical line shows the posterior mean, its value is given in the title of each plot after the parameter name.

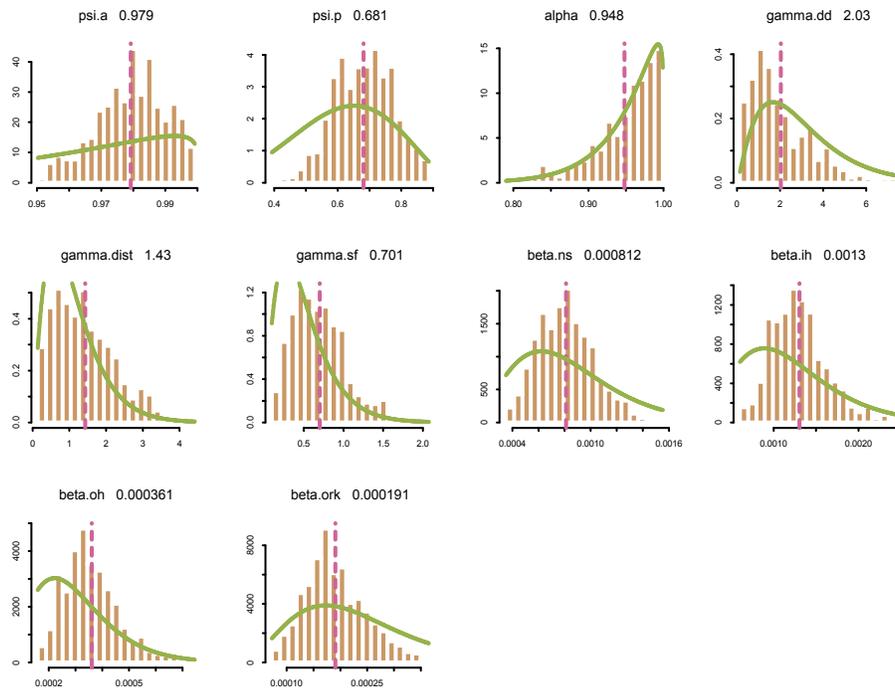


Figure 4. Posterior parameter estimates (histograms) and priors (solid lines) from the density dependent fecundity model (DDF).

