
Len Thomas and John Harwood

Estimating the size of the UK grey seal population between 1984 and 2006.

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 fit and compared four Bayesian state-space models of British grey seal dynamics, based on regional estimates of pup production from 1984 to 2006. The models were the same as those used in our 2005 and 2006 briefing papers: they allowed for a number of different forms of density dependence in either pup survival or fecundity, as well as fitness-dependent movement of recruiting females between regions. The fitting algorithms we employed were more sophisticated, allowing us to use a more appropriate observation error value than in previous years. However, Monte Carlo variation in results was also higher. We found relatively small differences in model selection criterion values between the model with simple density dependent survival and that with simple density dependent fecundity, although the former was slightly preferred. The estimated adult population size in 2006 for these two models was 116,000 (95% CI 90-154,000) and 248,000 (190-377,000) respectively, with the other two models taking intermediate values.

Introduction

In this paper, we present estimates of population size and related demographic parameters using the modelling framework of Thomas and Harwood (2005), fit to pup production data from 1984-2006. The biological system is represented using a state-space model: a stochastic time-series model that includes a “state process” for the evolution of the true but unknown state of the population through time, and an “observation process” that describes the measurements taken on the population (Buckland *et al.* 2004, Thomas *et al.* 2005, Newman *et al.* 2006).

We fit and compare the same four models used by Thomas and Harwood (2005, 2006). Two models allow for density dependent pup survival (DDS) and density dependent fecundity (DDF). In both cases, the density dependent relationship follows a Beverton-Holt function. Two further models extend this function by adding an extra

parameter that allows the effect of density dependence to be lessened until the population is close to carrying capacity (see Thomas and Harwood 2005). We refer to these as extended density dependent pup survival (EDDS) extended density dependent fecundity (EDDF). The models are formulated within the Bayesian statistical framework, and informative priors were specified on the model parameters and initial states (the 1984 population numbers).

As in previous reports, the models were fit using a computer-intensive algorithm called a Monte Carlo particle filter (Liu 2001). However, in previous years we had employed a relatively simple particle filter and this forced us to assume that the observation error was fixed at an implausibly large value (25% coefficient of variation). For this report, we developed a more sophisticated fitting algorithm that allowed us to estimate the observation error and use this estimated value when fitting and comparing the models.

Materials and Methods

Models

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 $1/\beta_r$ is proportional to the carrying capacity of the region. The EDDS model includes an extra parameter, ρ , that can alter the shape of the relationship between pup survival and pup numbers:

$$\phi_{p,r,t} = \frac{\phi_{p\max}}{1 + (\beta_r n_{0,r,t-1})^\rho} \quad (1)$$

For the DDF and EDDF models, 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 all models, we assume that adult female survival rate, ϕ_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 γ_{sf} , γ_{dd} , and γ_{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 and EDDS models,

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

while for the DDF and EDDF models,

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

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 and EDDS models, 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:

$$\alpha_{r,t} = \frac{\alpha_{\max}}{1 + \beta_r n_{6+,r,t}}$$

The EDDF model is similar, with

$$\alpha_{r,t} = \frac{\alpha_{\max}}{1 + (\beta_r n_{6+,r,t})^\rho} \quad (2)$$

For the observation process, we assume that pup production estimates follow a normal distribution with a constant coefficient of variation (CV), the value of which is governed by a model parameter, ψ , where $CV = 1/\sqrt{\psi}$ (i.e., ψ is a precision parameter). We estimated the value of ψ in an initial model run, and applied this estimate to other runs – see *Model outputs and comparison*, below.

In summary, the DDS and DDF models have 11 parameters. They share 9: adult survival ϕ_a , one carrying capacity parameter-related parameter for each region $\beta_1 - \beta_4$, three movement parameters γ_{sf} , γ_{dd} , and γ_{dist} , and the

observation precision parameter ψ . They differ in two parameters: the DDS model has maximum pup survival $\phi_{p\max}$ and constant fecundity α , while the DDF model has constant pup survival ϕ_p and maximum fecundity α_{\max} . The EDDS and EDDF models have one additional parameter, ρ , for the shape of the density-dependent response.

Data and Priors

Our input data were the pup production estimates for 1984-2006 from Duck and Mackey (2007), aggregated into regions.

Prior distributions for each parameter are given in Table 1, and are shown on Figure 3. We followed Thomas and Harwood (2005) in using a re-parameterization of the model to set priors on the numbers of pups at carrying capacity in each region, denoted χ_r for region r , rather than directly on the β s.

Table 1. Prior parameter distributions

Param	Distribution	Mean	Stdev
ϕ_a	Be(22.05,1.15)	0.95	0.04
$\phi_{p\max}, \phi_p$	Be(14.53,6.23)	0.7	0.1
χ_1	Ga(4,2500)	10000	5000
χ_2	Ga(4,1250)	5000	2500
χ_3	Ga(4,3750)	15000	7500
χ_4	Ga(4,10000)	40000	20000
ρ	Ga(4,2.5)	10	5
γ_{sf}	Ga(2.25,1.33)	0.5	0.33
γ_{dd}	Ga(2.25,0.49)	3	2
γ_{dist}	Ga(2.25,0.22)	ln(3)	ln(2)
α, α_{\max}	Be(22.05,1.15)	0.95	0.04
ψ	Ga(2.1, 66.67)	140	96.61

Prior distributions for the states in the DDS and EDDS models were generated using the priors for the parameters in conjunction with the 1984 data, as described by Thomas *et al.* (2005). Prior states for the DDF and EDDF model were generated in a similar manner, as described by Thomas and Harwood (2005). The prior distribution on ψ implies a prior mean on observation CV of 0.10 and prior standard deviation of 0.05.

Fitting Method

We used an extension of the particle filtering algorithm of Thomas and Harwood (2004, 2005, 2006), implemented in the C programming language. A particle filter is an algorithm that produces a set of weighted samples (particles) taken from the prior distributions on the parameters and states and projected forward stochastically through the time series. The weights relate to the manner in which the particles were sampled, how they were projected forward and the likelihood of the observed pup production given the simulated pup numbers. An introduction to particle filtering algorithms in the context of wildlife studies is given by Newman *et al.* (2006), and a more detailed description of the algorithm used here, applied to a similar model of seals, is given by Newman *et al.* (submitted). An outline of the main features of the algorithm is given below, for completeness – it is not necessary to read the rest of this sub-section to understand the results that follow.

Integrating out the observation error parameter.

We have found that it is not practical to estimate both the observation precision parameter ψ simultaneously with the states and other model parameters (because of the strong influence of ψ on the likelihood and hence particle weights for a given set of state values). Instead, we integrate ψ out of the model, and estimate the marginal posterior of this parameter conditional on estimates of the states and other model parameters (Newman *et al.* submitted). Given a gamma prior on ψ and a normal observation model, ψ has a gamma posterior distribution. The resulting likelihood weights in the integrated model are then based on the t -distribution.

Initial rejection control. The aim of this procedure is to weed out at an early stage sets of parameter and state combinations that are simulated from the prior but clearly have very low density in the posterior, so that computer time can be focussed on areas of parameter and state space that have higher posterior density. We simulated sets of 1,000,000 particles from the prior distributions, projected them forwards from 1984 to 1985 and calculated likelihood weights based on the 1985 data. We then applied rejection control, an algorithm that probabilistically removes particles with low weight (Liu 2001), using the mean of the

particles weights as the rejection control criterion. This typically resulted in about a quarter of the particles being retained. We repeated this process until we had at least 1,000,000 particles surviving the initial rejection control stage.

Auxiliary particle filter (Liu and West 2001). With this procedure, we projected forward one time step at a time, starting in 1985, initially deterministically. We then resampled the particles using the deterministic weights – i.e., according to the expected pup production in the next time period – thereby producing a set of “promising” particles. Resampled copies of the same ancestor particle will have the same parameter values, so to maintain parameter diversity we used kernel smoothing to jitter the parameter values (see Liu and West 2001 for details). This can cause bias (Newman submitted), so we kept the amount of kernel smoothing to a minimum, using a discount value of 0.997 (a value of 1.0 results in no jittering at all). After kernel smoothing, particles were then projected forward stochastically to the next time period, and weights were adjusted to take account of the initial resampling.

Final rejection control. At the last time period, rejection control was used to reduce the number of particles that must be stored. The rejection control criterion was the mean of the particle weights. This reduced the number of particles stored per run from 1,000,000 to between 200,000 and 600,000, depending on the model used.

Multiple runs. The above procedures generated samples based on 1,000,000 particles (although fewer were stored after the final rejection control). However, even this many samples gave a very imprecise estimate of the posterior distributions of interest for all models. Hence, many multiple runs (between 100 and 300) were required to reduce Monte Carlo error to acceptable levels. To reduce the resulting outputs down to a manageable level for post-processing (i.e., calculating posterior distributions on quantities of interest), it was necessary to apply further rejection control, this time using a rejection control criterion of the 99.99th percentile of the particle weights from all of the multiple runs for a particular model.

Model outputs and comparison

In previous reports, we have used the effective sample size of particles as a metric of reliability of the particle filter. This is not tenable with the new algorithm because the resampling ensures a high effective sample size of particles, but many of these particles are derived from the same initial (or “ancestral”) particle generated from the prior distributions. Hence, here we report the number of unique ancestral particles in our final sample, and we would ideally like this to be 1000 or more for each model.

To compare the models, we calculated the mean posterior Akaike Information Criterion (AIC) using the same method as Thomas and Harwood (2004, 2005, 2006). This criterion is a form of penalized likelihood, which recognizes the fact that models with more parameters are expected to fit better *a priori* by adding a penalty proportional to the number of model parameters. It is similar in spirit to the Bayesian Deviance Information Criterion (Spiegelhalter et al. 2002). Models were compared using Akaike weights (Burnham and Anderson 1998, p124), which can be thought of in the Bayesian context as the posterior probability of each model being the best approximating model.

Since the observations are assumed to be normally distributed random variables, there is an argument for using the bias-adjusted version of AIC, denoted AICc (Burnham and Anderson, 1998, p51). This criterion contains an extra term that imposes a stronger penalty on models with more parameters, with the effect of this extra term decreasing as the number of observations increases.

It is not useful to compare models where the observation error parameter has been estimated independently for each model. To understand why, consider two competing models, one of which fits the data well and the other poorly. Because the “good” model closely fits the data, the estimated error CV will be small, so the likelihood surface will be quite peaked, and small deviations in predicted pup counts from the observations will give rise to low negative log likelihood values, and so high AIC. On the other hand, because the “poor” model does not fit the data, the estimated error CV will be large, so the likelihood surface will be quite flat, and even large deviations in predicted pup counts from the observations may be assigned relatively large negative likelihood values and so low AIC.

Hence, a poor model can end up with a lower (i.e., “better”) AIC than a good one.

To resolve this paradox, we first estimated the measurement error parameter by fitting the data to the DDS model alone. We then took the estimated posterior mean and fit all four models using this fixed value. We present model selection statistics for these four models using the fixed observation error value.

For all four models, we also present posterior estimates of the model parameters and estimated pup production from 1984-2006. 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).

Results

Unique ancestral particle numbers

Including the estimated observation error in our algorithms meant that much more computer time was required for comparable accuracy, and we did not achieve our goal of generating at least 1000 unique ancestral particles for all models, despite simulating almost one billion particles in total. We therefore consider our results below to be somewhat provisional, pending further runs.

Table 2. Number of particles simulated (K), number saved after final rejection control step (K) and number of unique ancestral particles (U).*

Model	K (x10 ⁷)	K* (x10 ⁷)	U
<i>ψ</i> estimated			
DDS	300	4.90	1011
<i>ψ</i> fixed			
DDS	100	2.64	311
EDDS	100	1.08	176
DDF	100	0.16	187
EDDF	300	0.76	194

Estimate of observation precision parameter

Posterior estimates of both states and parameters from the DDS model run where *ψ* is estimated are shown in Figure 1. In this case, we achieved our target of >1000 unique ancestral particles, and the histograms showing parameter estimates are quite smooth, with little evidence of multi-modality that would indicate substantial Monte Carlo variability.

Estimated pup production generally tracks the observations well (Figure 1a), although there are several periods where the model consistently over- or under-predicts – for example in the North Sea colony, estimated pup production is higher than the observed value for 1992-1996 and then lower than the observed value for 8 out of the subsequent 10 years.

Histograms of marginal parameter estimates (Figure 1b) indicate that the posterior estimates are almost identical to the priors for ϕ_{pmax} and α , indicating that essentially nothing has been learnt from the data about these parameters. In contrast, posteriors for the other parameters are somewhat modified relative to the priors. This is similar to the findings of Thomas and Harwood (2005), who used a more formal sensitivity analysis.

The estimated posterior mean of *ψ* is 151, and we used this value in the 4 model runs reported in the next sub-section. As an aside, posterior mean observation CV can easily be estimated, as the weighted average of the CV for each particle:

$$CV^{\hat{}} = \frac{\sum_{k=1}^{K^*} w^{[k]} \left(1/\sqrt{\psi^{[k]}} \right)}{\sum_{k=1}^{K^*} w^{[k]}}$$

where $w^{[k]}$ is the weight associated with particle k , $\psi^{[k]}$ is the observation precision parameter for that particle, and K^* is the total number of particles. Using this method, the computed estimate of CV is 0.083 (corresponding 95% credibility interval 0.068, 0.101).

Comparison of models for density dependence

Smoothed posterior estimates of pup production (Thomas *et al.* 2005) for the four models, run with *ψ* fixed at 151, are shown in Figure 2. For all of these models we did not achieve a sufficient sample of unique particles (Table 1).

The DDS model estimates (Figure 2a) are almost identical to those from the analysis where *ψ* was estimated (Figure 1a), which is re-assuring despite the low number of unique particles. The state estimates are quite similar among the other models, although subjectively the extended density dependence models appear to do a better job of capturing the recent levelling-off of pup production in the Inner and Outer Hebrides. All of the models show runs of positive and negative

residuals, as noted for the DDS model earlier. Also, they are all rather better fits than previous analyses (e.g., Thomas and Harwood 2006, Figure 1) – something we discuss later.

There is some evidence of a lack of smoothness in the estimates – e.g., for the DDF model, Outer Hebrides 1986, and EDDF model, Outer Hebrides around 1995. These may be artefacts caused by Monte Carlo error, due to the low sample size of ancestral particles used in producing these results.

According to the AIC and AICc statistics (Table 3), the DDS and DDF models were strongly favoured over the EDDS and EDDF models. The DDS model was also slightly favoured over DDF (1.92 AIC and AICc points lower). However, we caution that these results are likely subject to high Monte Carlo error, and are potentially biased by the fitting method used (see Discussion).

Posterior parameter estimates are shown in Figure 3. There is clear evidence of multimodality in many of the histograms, particularly for the DDF and EDDF models, indicating high Monte Carlo variation. Hence we do not attempt to make any inferences about the estimated values.

Table 4. Estimated size, in thousands, of the British grey seal population at the start of the 2006 breeding season, derived from models fit to data from 1984-2006. Numbers are posterior means with 95% credibility intervals in brackets.

	DDS	EDDS
North Sea	13.8 (11 17.9)	14.6 (10.1 20.3)
Inner Hebrides	9.5 (7.1 12.7)	9.6 (7.6 13)
Outer Hebrides	33.8 (26.5 44.6)	36.1 (28.6 52.6)
Orkney	58.5 (45.8 79.1)	64.3 (40.3 100.6)
Total	115.7 (90.4 154.4)	124.6 (86.6 186.4)
	DDF	EDDF
North Sea	28.3 (21.6 41.1)	26.9 (22.3 30.5)
Inner Hebrides	23.1 (16.8 34.9)	17.7 (15.9 21.9)
Outer Hebrides	86.8 (63.7 151.2)	67.4 (60.1 83.5)
Orkney	109.8 (88 150.4)	98.7 (87.3 108.8)
Total	248 (190.1 377.7)	210.6 (185.6 244.8)

Table 3. Mean posterior log-likelihood, AIC, AICc and Akaike weights for models with fixed observation precision of 151 fit to data from 1984-2006.

Model	LnL	AIC	ΔAIC	Akaike (AIC) weight	AICc	ΔAICc	Akaike (AICc) weight
DDS	-681.18	1386.36	0.00	0.72	1390.32	0.00	0.72
EDDS	-687.10	1400.12	13.82	0.00	1404.86	13.83	0.00
DDF	-682.15	1388.29	1.92	0.28	1392.24	1.92	0.28
EDDF	-686.65	1399.30	12.93	0.00	1402.97	12.93	0.00

Discussion

Reliability of results

One aspect of reliability is Monte Carlo variation – i.e., variability in results that would be obtained by repeatedly running the fitting algorithm on the same data. We set an *ad hoc* target of 1000 unique ancestral particles, and this seems justified from our results: the initial DDS run that had >1000 unique particles appeared to show low Monte Carlo variation (as evidenced by smooth estimated pup production trajectories and posterior parameter distributions), while the runs with fixed ψ value, which had <1000

unique particles showed evidence of high Monte Carlo variation. A more formal approach was taken by Newman *et al.* (submitted), who used replicate runs of the algorithm on replicate simulated datasets to quantify both bias and Monte Carlo variation – although they only looked at parameter estimates and not states. For their model (similar to the DDS model here, with estimated ψ), 300 million particles appeared to be enough to reduce Monte Carlo variation to negligible levels. In our runs, the same number of particles appeared adequate for the DDS model, but clearly more will be required for the other models, particularly DDF and EDDF. For

example, for EDDF, 300 million particles produced 194 unique ancestral particles (Table 1), implying that 1546 million particles will be required to produce 1000 unique ancestral particles.

A second aspect of reliability is bias induced by the fitting algorithm. The kernel smoothing of parameters that is employed within the auxiliary particle filter preserves the first two moments of the parameter distributions, but does not preserve the relationship between parameters and states. Using simulated data and comparisons with Markov chain Monte Carlo estimates, Newman *et al.* (submitted) found that using the discount parameter of 0.997, as employed here, led to negligible bias in the marginal posterior parameter estimates. However, they did not investigate bias in the state estimates. A preliminary study comparing state estimates for the DDS model with and without kernel smoothing (Thomas, unpublished) shows that the better fit reported in this report relative to last years' report may be largely due to kernel smoothing bias rather than the better observation error model. This matter needs further investigation.

Comparison with previous estimates

Despite differences in algorithm and observation error models, the estimates of total population size are rather similar to those from last year. For example, the 2005 estimate from Thomas and Harwood (2006) under the DDS model was 105.2 (95%CI 79.3-141.7), compared with estimates for 2005 in the current report (Appendix 1) of 113.0 (88.9-149.7). Compared with Thomas and Harwood (2006), the DDS estimates are slightly higher, the DDF estimates slightly lower, and the EDDS and EDDF estimates are closer to the DDS and DDF estimates, respectively. There is still a nearly 2-fold difference in population size estimates between the DDS and DDF model, and although the model selection statistics in Table 3 cannot be considered reliable due to Monte Carlo variation, there appears still to be little to choose among these competing biological models.

One might expect that, since the estimated observation error is much lower than that assumed by Thomas and Harwood, confidence intervals on the population size estimates might be narrower. While there is some tendency for this, the effect is not great.

Future work

We are currently engaged in work aimed at further testing and improving the fitting algorithm. We are also working, in collaboration with others at SMRU, to improve the prior distributions on model parameters, based on intensive mark-recapture studies of seals at the Isle of May and North Rhona. However, by far the biggest source of uncertainty comes from which is the appropriate demographic parameter to model density dependence in. We have previously demonstrated the utility of obtaining a single additional estimate of total population size (Thomas and Harwood 2005, Matthiopoulos *et al.* 2006), and we hope to obtain such an estimate in the foreseeable future.

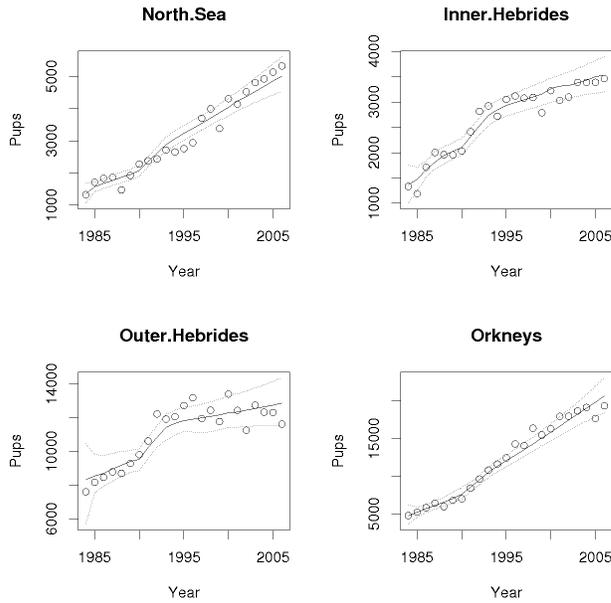
References

- 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.
- Duck, C.D. and B.L. Mackey 2007. Grey seal pup production in Britain in 2006. *SCOS Briefing Paper 07/2*.
- Hiby, L. and C.D. Duck. Unpublished. Estimates of the size of the British grey seal *Halichoerus grypus* population and levels of uncertainty.
- Jamieson, L.E. and S.P. Brooks. 2004. Density dependence in North American ducks. *Animal Biodiversity and Conservation* 27:113-128.
- Liu, J.S. 2001. *Monte Carlo Strategies in Scientific Computing*. Springer-Verlag, New York.
- Liu, J., and West, M. 2001. Combined parameter and state estimation in simulation-based filtering. In *Sequential Monte Carlo Methods in Practice.*, eds. Doucet, A., de Freitas, N., and Gordon, N. 197-223. Springer-Verlag, Berlin.
- Matthiopoulos, J., L. Thomas, B. McConnell, C.D. Duck, D. Thompson, P. Pomeroy, J. Harwood, E.J. Milner-Gulland, N. Wolf and M. Mangel. 2006. Putting long-term, population monitoring data to good use: the causes of density dependence in UK grey seals. *SCOS Briefing Paper 06/8*
- Newman, K.B., C. Fernández, L. Thomas., and S.T. Buckland. Submitted. Inference for state-space models for wild animal populations.
- Newman, K.B., S.T. Buckland, S.T. Lindley, L. Thomas and C. Fernández. 2006. Hidden

- process models for animal population dynamics. *Ecological Applications* 16:74-86.
- 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.
- Spiegelhalter D. J., Best N. G., Carlin B. P. and van der Linde A. 2002. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society B.* 64, 583-640.
- Spiegelhalter D. J., A. Thomas, N.G. Best and D. Lunn. 2005. WinBUGS 2.10.
- Thomas, L. and J. Harwood 2004. A comparison of grey seal population models incorporating density dependent fecundity and pup survival. *SCOS Briefing Paper 04/6*
- Thomas, L. and J. Harwood. 2005. Estimating the size of the UK grey seal population between 1984 and 2004: model selection, survey effort and sensitivity to priors. *SCOS Briefing Paper 05/3*
- Thomas, L. and J. Harwood. 2006. Estimating the size of the UK grey seal population between 1984 and 2005, and related research. *SCOS Briefing Paper 06/3*
- Thomas, L., S.T. Buckland, K.B. Newman, and J. Harwood. 2005. A unified framework for modelling wildlife population dynamics. *Australian and New Zealand Journal of Statistics* 47: 19-34.
- Thomas, L., T. van Lamsweerde and J. Harwood. 2004. The nature of density dependence in British grey seal populations. NERC Special Committee on Seals Briefing Paper 04/4.

Figure 1. Posterior estimates from a DDS model of grey seal population dynamics where the observation precision parameter ψ is estimated, fit to pup production estimates from 1984-2006.

(a) Estimates of true pup production (solid lines), together with 95% credibility interval (dotted lines) and observed pup production (circles).



(b) Parameter estimates (histograms) and priors (solid lines). The vertical line shows the posterior mean; its value is given in the title of each plot after the parameter name, with the associated standard error in parentheses.

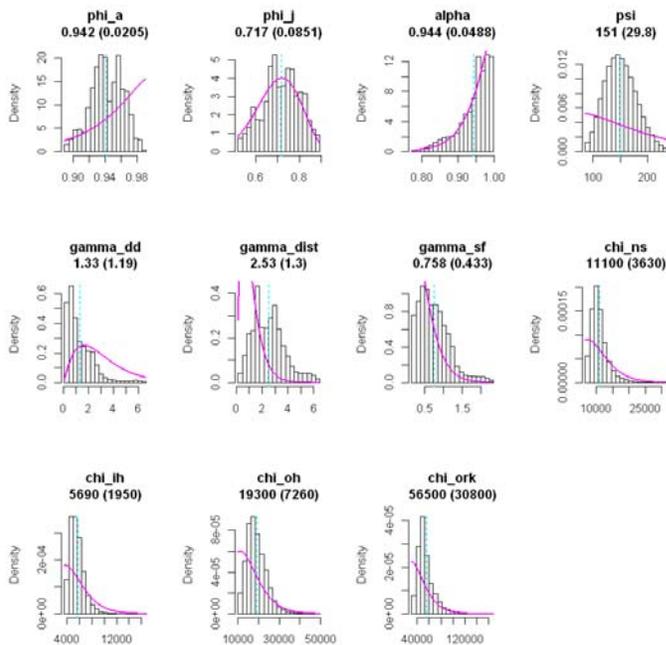
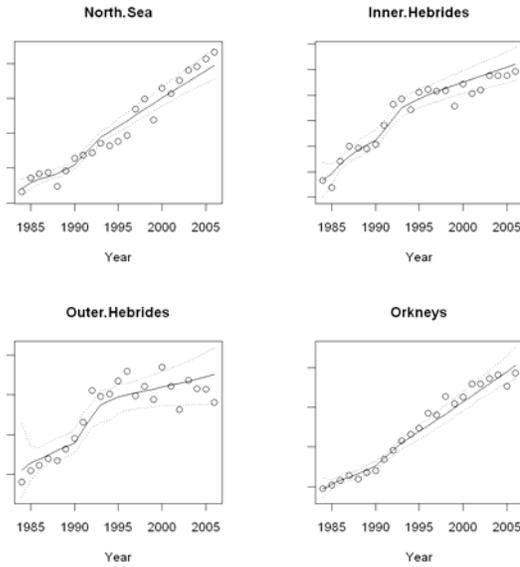
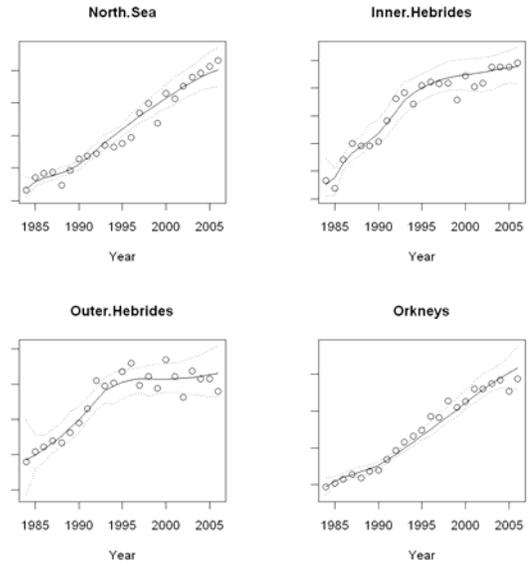


Figure 2. Estimates of true pup production from four models of grey seal population dynamics, where the observation precision parameter ψ is fixed at 151, fit to pup production estimates from 1984-2006. Input data are shown as circles, while the lines show the posterior mean bracketed by the 95% credibility interval.

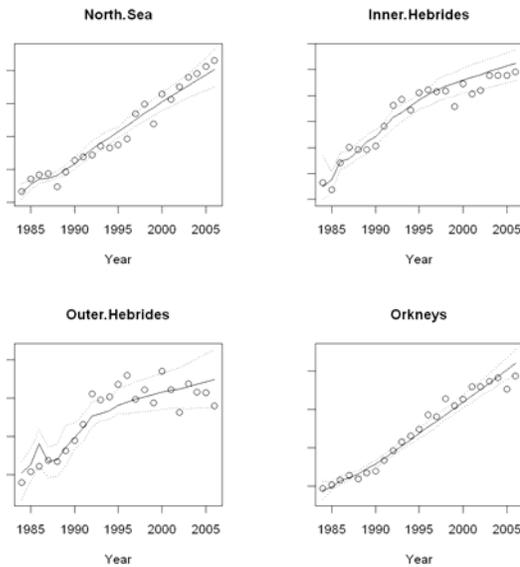
(a) Density dependent survival (DDS)



(b) Extended density dependent survival (EDDS)



(c) Density dependent fecundity (DDF)



(d) Extended density dependent fecundity (EDDF)

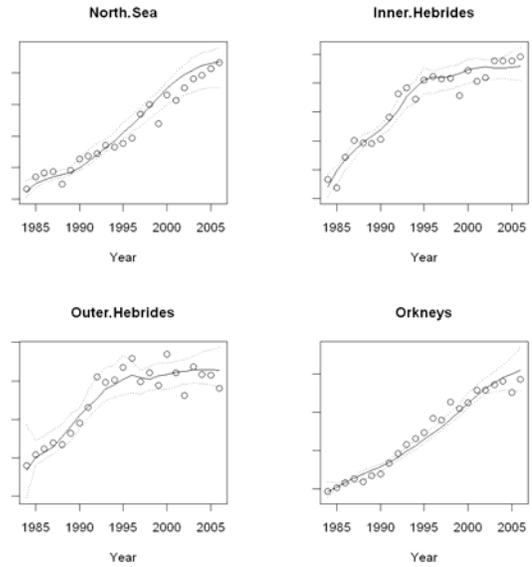
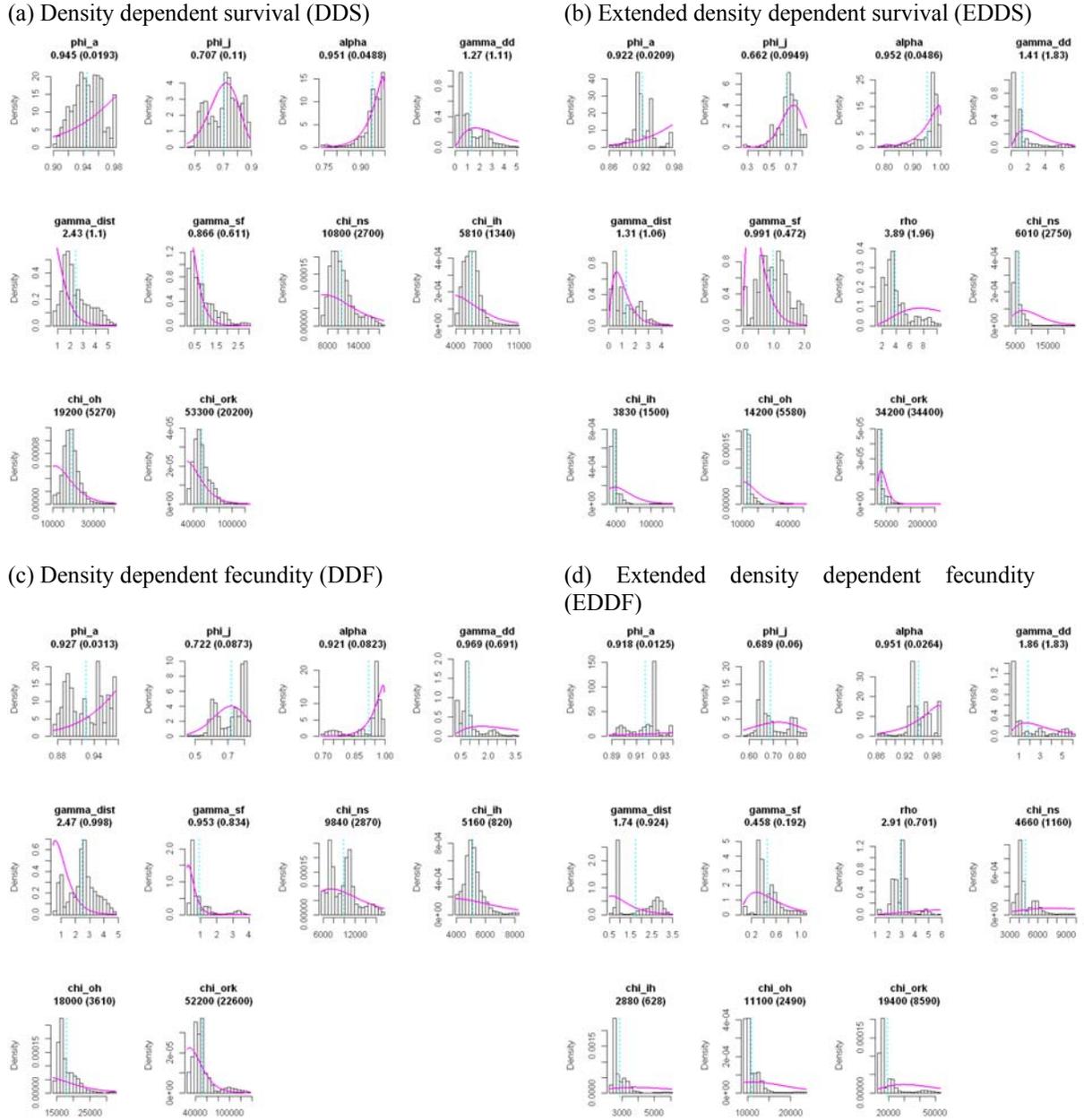


Figure 3. Posterior parameter estimates (histograms) and priors (solid lines) from four models of grey seal population dynamics where the observation precision parameter ψ is fixed at 151, fit to pup production estimates from 1984-2006. The vertical line shows the posterior mean, its value is given in the title of each plot after the parameter name, with the associated standard error in parentheses.



Appendix 1

Estimates of total population size, in thousands, at the beginning of each breeding season from 1984-2005, made using four models of British grey seal population dynamics. Numbers are posterior means followed by 95% credibility intervals in brackets.

Density dependent survival model

Year	North Sea	Inner Hebrides	Outer Hebrides	Orkneys	Total
1984	4.6 (3.8 5.6)	4.7 (3.6 5.7)	24.4 (19.2 31.6)	16.7 (13.7 22.2)	50.3 (40.3 65.2)
1985	5 (4 6.1)	4.9 (3.9 6.1)	24.5 (19.4 31.8)	18 (14.9 23.6)	52.4 (42.3 67.6)
1986	5.5 (4.7 6.7)	5.4 (4.5 6.7)	26.4 (20.9 33.9)	19.9 (17 26.3)	57.3 (47.1 73.6)
1987	6.1 (5.3 7.3)	5.9 (5 7.5)	28.1 (22.8 36.5)	22 (18.9 29.2)	62.1 (52 80.4)
1988	6.6 (5.8 8.1)	6.4 (5.4 8.2)	29.8 (24.6 38.7)	24.1 (20.9 31.9)	66.9 (56.6 86.9)
1989	7 (6.1 8.6)	6.7 (5.5 8.6)	30.2 (24.6 39.6)	25.8 (22.4 34.2)	69.7 (58.7 91)
1990	7.4 (6.4 9.2)	7 (5.6 9.1)	30.3 (24.6 40.2)	27.6 (24 36.6)	72.2 (60.6 95.1)
1991	7.9 (6.8 9.7)	7.2 (5.7 9.5)	30.2 (24.6 39.7)	29.4 (25.5 38.9)	74.7 (62.7 97.8)
1992	8.3 (7.2 10.3)	7.4 (5.8 9.8)	30.3 (24.7 39.4)	31.3 (27.2 41.2)	77.2 (64.9 100.6)
1993	8.7 (7.5 10.8)	7.6 (5.9 10.1)	30.5 (25 39.2)	33.2 (28.8 43.5)	79.9 (67.2 103.5)
1994	9.1 (7.8 11.3)	7.8 (6 10.4)	30.7 (25.2 39.1)	35.1 (30.1 45.9)	82.7 (69.1 106.6)
1995	9.5 (8.1 11.8)	8 (6.1 10.6)	30.9 (25.3 39.3)	37 (31.6 48.3)	85.5 (71.2 110)
1996	9.9 (8.4 12.4)	8.1 (6.2 10.9)	31.2 (25.5 39.5)	39 (33 50.7)	88.3 (73.1 113.4)
1997	10.4 (8.7 12.8)	8.3 (6.3 11)	31.4 (25.6 39.7)	41 (34.5 53.1)	91 (75.1 116.6)
1998	10.8 (9 13.3)	8.4 (6.4 11.1)	31.7 (25.7 40.1)	42.9 (35.8 55.6)	93.8 (76.9 120.1)
1999	11.2 (9.3 13.8)	8.6 (6.5 11.2)	31.9 (25.8 40.5)	44.9 (37.2 58.3)	96.6 (78.8 123.9)
2000	11.6 (9.6 14.4)	8.7 (6.6 11.4)	32.2 (25.9 41.1)	46.9 (38.5 61)	99.3 (80.5 127.8)
2001	12 (9.8 14.9)	8.9 (6.7 11.6)	32.5 (26 41.6)	48.8 (39.7 63.7)	102.1 (82.3 131.9)
2002	12.3 (10.1 15.6)	9 (6.8 11.9)	32.7 (26.1 42.2)	50.8 (41 66.5)	104.9 (84 136.1)
2003	12.7 (10.3 16.2)	9.1 (6.8 12.1)	33 (26.2 42.8)	52.7 (42.3 69.6)	107.6 (85.6 140.6)
2004	13.1 (10.6 16.8)	9.3 (6.9 12.3)	33.3 (26.3 43.3)	54.7 (43.5 72.7)	110.3 (87.3 145.1)
2005	13.5 (10.8 17.3)	9.4 (7 12.5)	33.6 (26.4 44)	56.6 (44.7 75.9)	113 (88.9 149.7)
2006	13.8 (11 17.9)	9.5 (7.1 12.7)	33.8 (26.5 44.6)	58.5 (45.8 79.1)	115.7 (90.4 154.4)

Extended density dependent survival model

Year	North Sea	Inner Hebrides	Outer Hebrides	Orkneys	Total
1984	5.2 (4.3 6.7)	5.5 (4.4 6.8)	26.7 (22 33.6)	19.5 (14.7 22.9)	56.9 (45.5 70)
1985	5.6 (4.6 6.9)	5.8 (4.8 7.1)	27.9 (22.3 34.4)	20.7 (16.1 24.5)	60.1 (47.8 72.9)
1986	6 (5.1 7.3)	6.1 (5.2 7.4)	29.8 (25 36.3)	22.1 (18.4 26.1)	64.2 (53.7 77.1)
1987	6.5 (5.6 7.8)	6.6 (5.6 7.8)	32 (27.3 38.6)	23.8 (19.5 27.6)	68.9 (58 81.8)
1988	7 (6.2 8.4)	7.1 (6 8.3)	34.1 (28.9 41.3)	25.5 (20.8 29.2)	73.7 (61.9 87.2)
1989	7.5 (6.6 9)	7.6 (6.3 8.8)	35.1 (29.6 41.8)	27.5 (22.2 31.2)	77.6 (64.8 90.9)
1990	8.1 (7.1 9.6)	8.1 (6.8 9.4)	35.7 (29.9 42.6)	29.5 (24.1 33.6)	81.4 (67.8 95.1)
1991	8.6 (7.4 10.3)	8.5 (7.1 10.1)	36.1 (29.7 43.7)	31.7 (26 36.1)	84.9 (70.2 100.2)
1992	9.2 (7.8 10.9)	8.9 (7.2 10.8)	36.1 (29.5 43.9)	34.1 (28.1 39)	88.3 (72.6 104.6)
1993	9.8 (8.2 11.7)	9.1 (7.3 11.1)	36.1 (29.4 44.6)	36.6 (30.7 42.2)	91.6 (75.6 109.6)
1994	10.4 (8.4 12.4)	9.2 (7.5 11.1)	35.9 (29.2 45.3)	39.3 (33.3 45.6)	94.8 (78.4 114.5)
1995	11 (8.8 13.1)	9.3 (7.7 11.2)	35.7 (29.2 44.9)	42.1 (35.1 49.3)	98 (80.8 118.5)
1996	11.6 (9.1 14.2)	9.3 (7.6 11.3)	35.5 (29.4 45.9)	44.8 (36.7 53)	101.2 (82.7 124.4)
1997	12.1 (9.3 15)	9.4 (7.7 11.3)	35.3 (29.1 47.2)	47.5 (38.4 56.5)	104.4 (84.5 129.9)
1998	12.6 (9.8 15.7)	9.4 (7.7 11.4)	35.3 (28.5 48.3)	50.1 (39.9 60)	107.4 (86 135.5)
1999	13.1 (10.1 16.5)	9.4 (7.7 11.7)	35.3 (28.3 49.3)	52.6 (41.8 63.4)	110.3 (87.8 141)
2000	13.4 (10.4 17.2)	9.4 (7.6 12.1)	35.3 (28.3 50)	54.7 (44.1 67.6)	112.9 (90.4 146.8)
2001	13.8 (10.7 17.8)	9.4 (7.5 12.3)	35.5 (28.3 50.6)	56.7 (45.9 72.2)	115.4 (92.4 152.8)
2002	14.1 (10.8 18.3)	9.5 (7.5 12.5)	35.6 (28.5 51.1)	58.4 (45.5 77.2)	117.6 (92.3 159.1)
2003	14.3 (10.8 18.9)	9.5 (7.5 12.6)	35.7 (28.6 51.6)	60 (45.4 82.5)	119.6 (92.4 165.6)
2004	14.4 (10.7 19.5)	9.6 (7.6 12.8)	35.9 (28.7 51.9)	61.5 (44 88.2)	121.4 (90.9 172.4)
2005	14.5 (10.5 20)	9.6 (7.6 12.9)	36 (28.6 52.3)	62.9 (42.1 94.2)	123 (88.8 179.3)
2006	14.6 (10.1 20.3)	9.6 (7.6 13)	36.1 (28.6 52.6)	64.3 (40.3 100.6)	124.6 (86.6 186.4)

Density dependent fecundity model

Year	North Sea	Inner Hebrides	Outer Hebrides	Orkneys	Total
1984	5.8 (5 7.1)	6.4 (4.8 9.9)	44.6 (33.9 55.1)	19.6 (17.5 22.1)	76.4 (61.2 94.3)
1985	6.3 (5.6 7.8)	7 (6 10.9)	45.9 (35.2 58.8)	21.4 (18.7 23.9)	80.7 (65.4 101.3)
1986	6.9 (6.2 8.6)	7.7 (6.6 11.6)	47.7 (37.8 62.1)	23.3 (20.3 26.5)	85.7 (70.8 108.8)
1987	7.7 (6.8 9.6)	8.6 (7.1 12.5)	49.9 (40.6 66.4)	25.8 (22.2 29.4)	92 (76.7 117.9)
1988	8.4 (7.2 10.5)	9.4 (7.9 13.4)	51.4 (41.7 70.4)	28.5 (24.3 32.7)	97.6 (81.1 127.1)
1989	9.1 (7.7 11.5)	10.2 (8.5 14.3)	52.9 (42.7 73.6)	31.1 (26.6 36.1)	103.3 (85.6 135.5)
1990	9.9 (8.2 12.7)	11.1 (9.2 15.4)	54.8 (44 77.7)	34.1 (29.1 39.8)	109.8 (90.6 145.6)
1991	10.7 (8.9 13.8)	11.9 (9.9 16.5)	56.8 (45.5 81.6)	37.3 (32 43.8)	116.7 (96.4 155.7)
1992	11.7 (9.6 15.1)	12.8 (10.4 17.6)	59.1 (46.8 85.9)	40.8 (35.1 48.5)	124.4 (101.9 167)
1993	12.7 (10.3 16.5)	13.7 (11 18.7)	61.6 (48.2 90.8)	44.7 (38 53.8)	132.7 (107.4 179.7)
1994	13.8 (11 17.9)	14.6 (11.6 19.7)	64 (49.5 95.3)	48.8 (41.1 59.4)	141.1 (113.2 192.3)
1995	14.9 (11.8 19.5)	15.5 (12.2 20.9)	66.1 (50.9 99.9)	52.9 (44.5 65.6)	149.4 (119.4 206)
1996	16 (12.6 21.1)	16.3 (12.8 22.2)	68.2 (52.2 104.8)	57.3 (47.9 71.9)	157.8 (125.5 220)
1997	17.1 (13.4 22.9)	17.1 (13.4 23.5)	70.2 (53.6 109.7)	61.9 (51.5 78.5)	166.3 (131.9 234.6)
1998	18.3 (14.3 24.8)	17.8 (13.9 24.8)	72.2 (55 114.4)	66.7 (55.1 85.3)	175 (138.2 249.3)
1999	19.5 (15.1 26.7)	18.5 (14.4 26.1)	74.1 (56.3 119.2)	71.6 (58.9 92.3)	183.6 (144.7 264.3)
2000	20.7 (16 28.6)	19.2 (14.9 27.4)	75.9 (57.6 123.9)	76.6 (62.8 99.7)	192.4 (151.3 279.6)
2001	21.9 (16.9 30.6)	19.9 (15.3 28.7)	77.7 (58.9 128.6)	81.7 (66.9 107.3)	201.3 (158 295.2)
2002	23.2 (17.9 32.7)	20.6 (15.7 29.9)	79.6 (60.2 133.2)	87.1 (71.1 115.3)	210.4 (164.8 311)
2003	24.4 (18.8 34.7)	21.2 (16.1 31.2)	81.4 (61.5 137.7)	92.5 (75.3 123.5)	219.6 (171.7 327.1)
2004	25.7 (19.7 36.8)	21.9 (16.3 32.4)	83.2 (62.7 142.2)	98.1 (79.5 132.1)	228.9 (178.3 343.5)
2005	27 (20.7 38.9)	22.5 (16.6 33.7)	85 (63.4 146.7)	103.9 (83.7 141.1)	238.4 (184.4 360.4)
2006	28.3 (21.6 41.1)	23.1 (16.8 34.9)	86.8 (63.7 151.2)	109.8 (88 150.4)	248 (190.1 377.7)

Extended density dependent fecundity model

Year	North Sea	Inner Hebrides	Outer Hebrides	Orkneys	Total
1984	5.2 (4.8 5.9)	5.7 (5 6.5)	27.9 (23.9 34.2)	20.7 (18.1 22.4)	59.4 (51.8 69)
1985	5.5 (5.1 6.4)	6 (5.3 6.8)	29.9 (25.8 34.8)	21.9 (19.1 23.3)	63.3 (55.3 71.3)
1986	6 (5.6 6.8)	6.4 (5.8 7.1)	32.1 (28.7 36.6)	23.4 (21 24.9)	67.9 (61 75.4)
1987	6.5 (6.1 7.3)	6.9 (6.3 7.6)	34.2 (31.7 38.5)	24.9 (23.1 26.6)	72.6 (67.2 80.1)
1988	7 (6.6 8)	7.5 (6.9 8.3)	36.4 (34.3 41.2)	26.7 (24.9 28.6)	77.6 (72.7 86.1)
1989	7.5 (7.1 8.6)	8.1 (7.4 8.8)	38.5 (36.4 42.9)	28.7 (26.7 31.1)	82.9 (77.6 91.4)
1990	8 (7.6 9.2)	8.8 (7.9 9.4)	40.7 (38.3 44.8)	30.9 (28.5 33.6)	88.5 (82.2 97.1)
1991	8.6 (8.1 10.1)	9.4 (8.6 10.1)	42.9 (40.1 46.9)	33.2 (30.4 36.1)	94.2 (87.2 103.2)
1992	9.3 (8.7 11)	10.2 (9.4 10.8)	45.2 (41.7 49.3)	35.7 (32.8 39.1)	100.4 (92.7 110.2)
1993	10.2 (9.5 12)	10.9 (10.3 11.7)	47.4 (43 52.4)	38.6 (35.6 42.5)	107.1 (98.5 118.4)
1994	11.1 (10.3 13.1)	11.7 (10.8 12.6)	49.5 (44.9 55.6)	41.7 (38.5 45.9)	114 (104.6 127.3)
1995	12.1 (11 14.4)	12.5 (11.2 13.3)	51.6 (46 59.1)	45 (41.4 50)	121.2 (109.7 136.8)
1996	13.3 (11.8 15.6)	13.2 (11.7 14.5)	53.5 (47.2 62.3)	48.8 (44.4 54.7)	128.8 (115.2 147.1)
1997	14.6 (12.7 16.8)	13.8 (12 15.6)	55.1 (48.4 65.2)	53.1 (47.4 59.3)	136.6 (120.5 156.9)
1998	16 (13.6 18.1)	14.4 (12.4 16.5)	56.5 (49.1 67.5)	57.4 (50.8 64.4)	144.3 (125.9 166.5)
1999	17.5 (14.6 19.3)	14.8 (12.5 17.4)	57.8 (49.8 69.8)	62.2 (54.5 69.7)	152.3 (131.4 176.3)
2000	18.9 (15.6 20.9)	15.3 (12.7 18.3)	59.1 (50.8 72.1)	67.1 (58.4 75.1)	160.4 (137.5 186.4)
2001	20.4 (16.6 22.6)	15.7 (13 19.1)	60.4 (52.2 74.2)	72.4 (62.5 80.7)	168.8 (144.3 196.7)
2002	21.8 (17.8 24.3)	16.1 (13.3 19.9)	61.8 (53.9 76.4)	77.8 (66.9 86.3)	177.4 (151.9 206.9)
2003	23.2 (18.9 26)	16.5 (13.7 20.5)	63.2 (56 78.4)	83.2 (71.6 92)	186 (160.2 216.9)
2004	24.5 (20.1 27.6)	16.8 (14.6 21)	64.6 (57.6 80.1)	88.5 (76.5 97.7)	194.4 (168.8 226.5)
2005	25.8 (21.2 29.1)	17.3 (15.5 21.5)	66 (58.8 81.9)	93.6 (81.8 103.4)	202.7 (177.3 235.9)
2006	26.9 (22.3 30.5)	17.7 (15.9 21.9)	67.4 (60.1 83.5)	98.7 (87.3 108.8)	210.6 (185.6 244.8)