Northwest Atlantic   Fisheries Organization

SCIENTIFIC COUNCIL MEETING – NOVEMBER 2001

Assessment Models for the West Greenland Shrimp Stock
Using a Bayesian Approach. First Results.

by

C. Hvingel and M.C. S. Kingsley

Pinngortitaleriffik, Greenland Institute of Natural Resources
Box 570, DK-3900 Nuuk, Greenland

Abstract

This paper presents first results of a Bayesian approach to shrimp stock assessments. Non-age structured models based on the logistic function with and without an explicit term for cod predation, is used to describe shrimp population dynamics. Biomass indices from survey and standardised commercial catch-per-unit-effort series, catch, cod stock size estimates and prior estimates of model parameters provided information to the models. The inadequacy of a model to fully describe population behaviour and the inevitable errors associated with the observed data series, i.e. observation and process error, was incorporated simultaneously using a state-space modelling framework.

The models can provide posterior probability distributions of management related parameters as desired. E.g. the model predicts that in the absence of cod predation a fishery of 100 000 tons annually will have 75% probability of being below 'Maximum sustainable yield' – or 25% risk of being above.

Introduction

Annual landings of northern shrimp (*Pandalus borealis*) in the North Atlantic have increased from about 100000 tons in the early 1980s to more than 350000 tons in 2000 (Skúladóttir, pers. com.). Although shrimp has become one of the most important target species in these waters, little progress has been made in developing a standardised predictive tool for stock assessment (cf. Stefánsson *et al.*, 1994; Cadrin and Clark, 1999; Savard *et al.*, 1991; Hvingel and Kingsley, 2000; Anon., 2000a).

Management advice for most shrimp stocks in the North-west Atlantic is formulated by qualitative assessment of recent trends in biomass indices, in the size composition of the stock, and in the catch history (Anon., 2000b; ref Can. assessment etc.), and by a qualitative evaluation of the short-term prospects for recruitment. Advice is given as a Total Allowable Catch (TAC) one year ahead as set by consensus of the assessment board. The method of arriving at a certain TAC value is not explicitly stated and the uncertainty associated with the process is not quantified. Such methods lack predictive rigour, including formal statements of uncertainty, and are therefore unable to satisfy increasing demands for longer-term predictions and for quantitative comparisons between alternative management options.

A quantitative assessment may be reached through construction of mathematical models describing stock dynamics and their links to collected data and ancillary information. Each model represents one hypothesis about "the state of nature". From the infinite collection of possible models a subset of the most plausible may be chosen based on expert knowledge. An integrated evaluation of inferences about management options from that subset may be

considered taking into account the "model uncertainty" of the assessment process. Within each model uncertainty arises from its inability to fully capture the dynamics of the stock and from having to interpret erroneous data. Thus, to capture the full range of assessment uncertainty, three levels of uncertainty need to be considered: 1. model uncertainty, 2. process error and 3. observation error.

Following the ideas of Punt and Hilborn (1997) the first step in the construction of an assessment framework is specifying the alternative hypotheses. This is done through exploring models within three levels of model complexity: 1. the logistic (This paper), 2. the delay-difference (Hvingel and Kingsley, in prep.) and 3. the age-based (Hvingel and Kingsley, in prep.). Within these model-types different versions is investigated by including predation and environmental effects. Process and observation error is incorporated by using a state-space modelling framework (Schnute, 1994) and Bayesian inference (e.g. Gelman *et al.*, 1995) to estimate probability distributions of model parameters.

Age-structured population models that account in great detail for age-dependent growth, growth-dependent recruitment to the fishery, and age-specific fishing and natural mortality are now the standard tool for the assessment of most fish stocks. Such models are critically dependent on accurate annual information on the age and length composition of the stock and the catches. But shrimp have no permanent hard parts to accumulate annual layers and ages cannot be measured. Modal analysis of length distributions (e.g. McDonald and Pitcher, 1979) seems to be inevitably dependent on assumptions about growth rate that are largely unverifiable and leaves large residual uncertainty about converting size distributions into age distributions. This is particularly true for the relatively slow-growing shrimp in northern waters and for the important larger size classes.

The use of non-age-structured models may therefore be an alternative. Such models are usually of the stock-production type (Graham, 1935; Schaefer, 1954; Fox, 1970; Pella and Tomlinson, 1969) and describe stock dynamics simply in terms of rates of change of total biomass, rather than by the detail of age-specific growth and mortality of individuals (see Hilborn and Walters, 1992, chap 8. for a review of production models).

Predation is an important factor in shrimp stock dynamics. Several fish and marine mammal species prey on shrimp (Pedersen and Riget, 1993; Magnússon and Pálsson, 1989; Grunwald, 1998, ref on marine mammals). The Atlantic cod (*Gadus morhua*) co-occurs with northern shrimp, is a bottom-feeding predatory fish, and can attain high biomass densities, qualities which give it the potential for being a controlling predator (Lilly *et al.*, 2000; Berenboim *et al.*, 2000; Stefánsson *et al.*, 1998). The rapid development of the West Greenland shrimp fisheries during the 1970s is thought to be connected to a major decline in the cod stock in that area (Carlsson and Smith, 1978); shrimp fisheries on the Labrador banks appear similarly to have thriven after the collapse of the 'northern' cod stocks in NAFO 2J3KL. Predation by cod predation should therefore be taken into account in attempting to model the stock dynamics of northern shrimp, even at an aggregate stock level.

The physical environment also affects shrimp populations (e.g. Koeller, 2000; Anderson, 2000) and may be the cause of the rapid changes in abundance seen in some stocks (Anderson, 2000; Apollonio *et al.*, 1986). However, at this stage of modelling, environmental effects were not considered as explicit variables.

This paper is a first draft of the first article in a planned series of three on shrimp population dynamic modelling, with the purpose of developing a first set of models based on the logistic function. These models constitute hypotheses about "the state of nature" intended along with the models developed in following papers (Hvingel and Kingsley, in prep. a and b) to enter into an integrated shrimp stock assessment and decision framework able to make use of all available information pertaining to the dynamics of the stock and to provide realistic estimates of uncertainty of the assessment process.

Methods

Modelling framework

The models were built in a state-space framework (e.g. Schnute, 1994). The state-space conception of a stock dynamic model regards the series of state values defining the stock trajectory through time as existing simultaneously, related to each other and to the parameters that control stock dynamics by a set of simultaneous,

possibly stochastic, equations. The models were thus defined in terms of sets of parameters (q), the values of which defined the stock dynamics of the shrimp stock.

A Bayesian approach was used to construct a "posterior" distribution of likelihoods of possible values of model parameters. The posterior gives the probability of a hypothetical set of values for unobserved model parameters (q), given the data and a "prior" probability distribution for q which defines the information available about q before the experiment was carried out or the observation made. Bayes (1763) restated a general probability identity in the form:

$$p(q \mid data) \propto p(data \mid q) \, p(q) \, ,$$

where $p(q \mid data)$ is the posterior probability distribution, $p(q)$ is the estimated or assumed prior probability distribution of the unobservables, and $p(data \mid q)$ is the likelihood of the data as a function of the values of q .

In applying this equation, the posterior probability distribution of q is derived numerically by sampling many possible trial sets of values. Obtaining large samples, with satisfactory statistical properties, from the posterior distributions of multiple parameters presents complex computational problems. The programming framework WinBUGS v.1.3, made available by the Medical Research Council and the Imperial College of Science, Technology and Medicine in England (www.mrc-bsu.cam.ac.uk/bugs; Gilks *et al.*, 1994; Spiegelhalter *et al.*, 2000), provided a means of specifying and analysing Bayesian models, including selection and implementation of appropriate algorithms. For numerical integration WinBUGS uses "Metropolis-Hastings within Gibbs sampling" (Gilks *et al.*, 1996, see http://www.mrc-bsu.cam.ac.uk/bugs/).

State equations

The basic stock-dynamic model was a generalisation of the logistic model of population growth (Richards, 1959; Pella and Tomlinson, 1969). Its differential form is:

$$\frac{dB}{dt} = Br\left(1 - \left(\frac{B}{K}\right)^{m}\right),$$

where B is biomass, K is carrying capacity, and r is the intrinsic rate of growth. m is a shape parameter for the stock-recruitment curve; a value of 1 gives the standard logistic trajectory. If m takes high values, it implies that density-dependent reductions in population growth do not become significant until higher stock levels.

A discrete and stochastic form of this model, also modified to include fishing mortality and predation, was expressed in the form:

$$P_{t+1} = \left(P_t - \frac{C_t}{K} - \frac{V_t(P_t, Cod_t)}{K} + P_t \frac{MSY(m+1)^{(m+1)/m}}{Km}\left(1 - P_t^{m}\right)\right) \cdot \exp(n_t)$$

where P is the stock biomass relative to carrying capacity $(P=B/K)$ and *MSY* is the annualised value of the instantaneous maximum sustainable yield rate. C_t , the catch, and V_t, the predation, in year t, are also normalised on the carrying capacity. It was found that this reparameterization gave slightly more tractable solution characteristics (cf. Meyer and Millar, 1999). The 'process errors', v_t are normally, independently and identically distributed with mean 0 and variance s^{2} .

The basic stock-dynamic model in the absence of fishing and explicit predation has the characteristics that the MSY level of biomass relative to carrying capacity is equal to:

$$P_{MSY} = (m+1)^{-1/m}$$

and the MSY itself is given by

$$MSY = P_{MSY} r \frac{m}{m+1} = rK \frac{m}{(m+1)^{\frac{m+1}{m}}}$$

V is a predation term depending simultaneously on the shrimp stock biomass and on the biomass of the predator, Atlantic cod (*Gadus morhua*). Shrimp is not regarded as the primary forage for cod if fish, e.g. capelin, is available (ref. Bereinboom and Icelandic). The relative abundance of shrimp may therefore be expected to have to reach a certain level for cod to switch to this food type. A Holling type III functional response (Holling, 1959) was therefore used for inclusion of cod predation in the model. This function predicts a sigmoidal response of predation rate to increasing prey density, which essentially states that at intermediate shrimp density the consumption *per predator* increases linearly with prey density, but that at low prey density it approaches zero and at high prey density it has an asymptotic maximum. It can most simply be expressed in the form:

$$O(P) = \frac{O_{max} P^2}{P^2 + P_{0.5}^2}$$

where O_{max} is the maximum consumption of prey per predator (kg/kg) reached at large prey biomass, and P_{05} is the prey biomass index at which the consumption is half of the maximum. The total consumption at time t is then given by:

$$V_t(P_t, Cod_t) = Cod_t O(P_t) \exp(\textbf{\textit{t}}_t) = Cod_t \frac{O_{max} P_t^2}{P_t^2 + P_{0.5}^2} \exp(\textbf{\textit{t}}_t)$$

Inclusion of predation in this form requires including in the stock-dynamics model the two defining constants O_{max} and $P_{0.5}$ and an error term $\textbf{\textit{t}}$ which is normally, independently and identically distributed with mean 0 and variance $\textbf{\textit{e}}^2$.

Priors

Bayesian statistics perceives a set of observations as updating or correcting a prior model of reality, and a Bayesian model must be provided with information from two sources. One source is the study itself, i.e. the data acquired therein; the other is ancillary knowledge about the parameters, independent of the study, conceptually pre-existing, and summarised in "prior" distributions of model parameter values. The prior distribution for a parameter should incorporate all external information available. If no prior knowledge is available a low-informative "reference" prior (Kass and Wasserman, 1996) is used.

Prior probability distributions of parameter values may in data scarce situations (i.e. in most marine population dynamic modelling) often have a strong effect on the posterior distribution and therefore on the conclusions drawn from a model. Selection of appropriate priors is therefore an important phase of building a Bayesian model. Distributions of the different parameters in the model interact, and what appears to be a reference prior for one variable may through its interactions with other variables act as an undesirably informative prior for them. Some guidelines for developing priors are given in Punt and Hilborn, 1997; and Gelman *et al.*, 1995.

q_s, q_c (catchability)

The catchability coefficients, q_c and q_s, scaled the biomass indices of the standardised CPUE and survey series respectively to the absolute biomass estimates. Little prior information about these parameters was available except that the survey index was a *minimum* estimate of stock biomass (ref.) so values of q_s greater than 1 are unlikely. The CPUE indices were expressed relative to a value of 1 in the most recent year, and therefore q_c had to be smaller than 1 as well.

Thus Reference priors were used. Because the catchability coefficients interacted with the carrying capacity, K, a uniform distribution was not non-informative, and a prior distribution uniform on a log scale from $-\infty$ to 0 was preferred (Punt and Hilborn, 1997; McAllister and Kirkwood, 1997; Gelman *et al.*, 1995). In practice this was approximated by using $\ln(q_c)\sim U(20,0)$ and $\ln(q_s)\sim U(20,0)$ ('$\sim$' means distributed as; U is the uniform distribution).

MSY (Maximum sustainable yield)

Few estimates of the intrinsic rate of increase, r, for other shrimp stocks have been published, which makes it hard to define an informative prior for this parameter. We therefore parameterised the model in terms of *MSY* instead of *r*.

The prior distribution for MSY was modelled conservatively by a log-normal distribution, defined by ln(x) having mean 4.5 and standard deviation 1. The mode of this distribution was at 35 thousand tons somewhat less than the 50 thousand tons, which is the average catches of the time series. This prior corresponds in this set-up to a prior for r with 10th percentile at 0.1 and 90th percentile at 4.1.

m (the shape parameter)

We had no prior information on the shape-parameter, *m,* for this stock. The ordinate at the point of inflection of Richards curve is given by $P_{MSY} = (m+1)^{-1/m}$. The curve of *m* as a function of P_{MSY} in the interval $0.35 < p < 0.65$ was approximated by $m = a\exp(bP_{MSY}^{c})$ with appropriate values of *a* (0.4333), *b* (4.8121), and *c* (1.6734). P_{MSY} was assigned a uniform prior on the interval 0.35 to 0.65, and this approximate function was used to convert this to a non-informative prior for *m.*

K (carrying capacity)

The largest catch taken in a year is about 100 thousand tons and this may be taken as a lower bound on K. It was hard to come up with an upper limit to K. However, it seemed at least in this case where little information was available in the data on K, to be important to prevent sampling of very high and unrealistic values. Thus an upper limit was constructed as follows: some of the highest densities ever recorded in the survey are around 0.5kg/m^2. The total distributional area of the stock is about 120 000 km^2, much of which has densities well below commercial interest. If 5% of the total area had 0.5 kg/m^2 it yields some 3 000 thousand tons. This was chosen to provide what was thought as a safe upper limit. Instead of using a uniform distribution U (100,3000) the prior was uniform on a log scale with limits at 4.6 and 8 in order to reduce the tendency of the sampler to wander off to extreme values.

P_0 (starting biomass level)

There is usually little information about the historic size of a stock preceding the time series of biomass indices. Commonly the assumption of Pzero = 1 i.e. the stock was, initially, at carrying capacity is used to partly overcome this difficulty. However, this may have the effect of imposing an informative prior on K, it is not easy to defend in any circumstances, and it is especially questionable when, as in the present case, the stock had already been fished for some years before the start of the observations and/or was subject to predation by a fluctuating cod stock. Instead we used the less informative normally distributed prior P~N (1,0.2) in the model without a predation effect, and P~N(0.8,0.2) when predation was included – the decreased mean to acknowledge the presence of a large cod stock historic to the time series.

n, t, k, w (error terms)

The priors for the error terms associated with the biomass indices were based on an estimated CV of around 10% for the CPUE-series (Hvingel *et al.*, 2000) and about 17% for the survey series (Kingsley *et al.*, 2000). The probability distribution of their standard deviations therefore had modes at 0.10 and 0.17 respectively, i.e. their precision were gamma distributed: $1/sd_c^2\sim G(2,0.03)$ and $1/sd_s^2\sim G(2,0.0675)$. The standard deviations of the stock size, P, and the predation, V, were given priors similar to that for sd_c^2.

Omax (maximum predation rate)

Omax provides an upper limit to the amount of shrimp consumed by cod. The data itself could not be expected to contain much information about this parameter. An informative prior for the largest possible quantity of shrimp (kg) eaten pr. kg of cod was therefore constructed using estimates of growth and food conversion coefficients for cod feed to satiation in laboratory experiments (Björnsson *et al.*, 2001; Björnsson and Steinarsson, in press). A maximum annual weight gain pr. kg fish for each age group were calculated inserting mean weight at age (Anon. 2000, NWWG) and the average water temperature on the banks off West Greenland of 2°C (Buch, 2000) in the growth model (Björnsson and Steinarsson, in press). A food conversion coefficient of 3 (Björnsson *et al.*, 2001) was then applied to convert the weight gain to maximum annual shrimp consumption. Assuming an average population structure (1975-89) a level for Omax was estimated to 2.5kg shrimp/kg cod/year. The prior took the form Omax~N(2.5,0.14) truncated at 1.5 and 3.5.

P~50~

The parameter P_{50} is the size of the shrimp stock at which 50% of Omax is reached. A prior uniform on a log-scale between 0.1 and 5 were chosen for the P_{50} to be non-informative with respect to the predation rate, O(P).

Data and link functions

Two series of biomass indices were available (Table 1). A standardised and combined series of annual commercial vessel catch rates (Hvingel *et al.*, 2000; Hvingel, 2000) and a biomass index obtained by the Greenland trawl survey for the years 1988-2000 (Kingsley *et al.*, 2000; Carlsson *et al.*, 2000).

These indices were assumed scaled to real biomass by a catchability constant. Lognormal observation error were applied, ω and κ, and hence these data took the distributions:

$$CPUE_t \sim \log N(\ln(q_c KP_t), \boldsymbol{w}) \quad , \quad surv_t \sim \log N(\ln(q_s KP_t), \boldsymbol{k})$$

Total reported catch in Subarea 1 + Div. 0A 1970-2000 (Hvingel, 2000) was used as yield data (Table 1). These data are considered very reliable without major discarding problems or misreportings and were therefore entered in the model as error-free.

A series of cod biomass 1955-1992 were obtained from the assessments done by the NWWG (Anon., 1996). Absolute estimates of stock size were only available for West and East Greenland stocks combined. However, a separate series for the West Greenland component was constructed as an estimated 70% fraction of the total stock (Table x). In 1992 the cod fishery stopped due to a very low stock size and the VPA estimates done by the NWWG could not be continued. The cod stock was set to 5000 tons for 1993-2000. The series was entered deterministically in the model and the error associated was left to accumulate along with the process error in the error term for predation, V.

Convergence diagnostics

Conclusions based on an MCMC output depend on the assumption that the chain of sampled values for each parameter can be considered to be an unbiased random sample from the target (posterior) distribution. A chain is usually serially correlated, and at the start is affected by the initial value assigned to each parameter. Initial values are not necessarily, or even usually, random samples from the posterior distributions. To check that the influence of the initial value had decayed and that the chain had 'converged' to a stationary sampling from a constant distribution before data collection began, several diagnostic tools were applied. This was done in S+ (ref) using the Bayesian Output Analysis programme (BOA) v. 0.5.0 by Brian Smith (www.public-health.uiowa.edu/boa/), which provides a suite of processing and diagnostic tools for MCMC outputs.

A number of parallel chains with different starting points and with different random number seeds were analysed by the Brooks, Gelman and Rubin convergence diagnostic (Gelman and Rubin, 1992; Brooks and Gelman, 1998) to evaluate if the samples could be considered to have arisen from the target distribution. A stationarity test

(Heidelberger and Welch 1983) was applied to individual chains. If evidence of non-stationarity is found this process discards iterations from the beginning of the chain until the remaining chain passes the test. A convergence diagnostic (Geweke 1992), which especially tests for convergence of the *mean* of the sampled parameter was also used along with Raftery and Lewis's (1992) tests for convergence to the stationary distribution and estimation of the run-lengths needed to accurately estimate quantiles.

Model check

In order to check whether the model was a 'good' fit to the data, the observations were compared with their predicted distributions. Different goodness-of-fit statistics were computed.

Firstly, we calculated the simple difference between each observed data point and its trial value in each MCMC sampling step. The summary statistics of the distributions of these residuals indicated by their central tendency whether the modelled values were biased with respect to the observations. The dispersion of the residuals was compared with both the classical statistical estimate of the SE of the respective observed value and the modelled collective estimate of the precision of the observation sequence.

Secondly, the overall posterior distribution was investigated for potential effects of model deficiencies by comparing each data point with its posterior predictive distribution (Posterior Predictive Checks—PPC; Gelman *et al.* 1995, 1996). Intuitively this is equivalent to running the model backwards to see the data needed to get the suggested result and then judge if that is a realistic scenario.

Simulated sets of observed data, $data_{rep}$, were drawn from the sampling distributions for the observations as output from the model. Therefore, $data_{rep}$ has distribution:

$$P(data_{rep} \mid data_{obs}) = \int P(data_{rep} \mid \boldsymbol{q}) P(\boldsymbol{q} \mid data_{obs}) d\boldsymbol{q}$$

In this expression the term $P(data_{rep} \mid \boldsymbol{q})$ represents the sampling of observations from the distributions set up in the model to define them, and the term $P(\boldsymbol{q} \mid data_{obs})$ represents the sampling of the parameters of those distribution from their own *posterior* distributions. If the model fitted the observed data well, the observed data should, collectively, be a 'likely' drawing from its simulated distribution, i.e. the observed data and the replicate data should look alike. The degree of similarity between the original and the replicate data points were summarised in a vector of p.values, calculated as the proportion of N simulations in which a sampling of the posterior distribution for an observed value exceeded the value in the input data:

$$p.value = \frac{1}{N} \sum_{j=1}^{N} I((data_{j}^{rep}, \boldsymbol{q}_{j}) - (data^{obs}, \boldsymbol{q}_{j})) \quad,$$

where $I(x)$ is 1 if x is true, 0 if x if false. Values close to 0 or 1 in the vector *p-value* would indicate that the observed data point was an unlikely drawing from its posterior distribution.

Thirdly, the 'Conditional Predictive Ordinate' (Gelfand and Dey, 1994) was calculated as the inverse of the posterior mean of the inverse of the likelihood for observation i:

$$CPO_{i} = \left[\frac{1}{N} \sum_{j=1}^{N} \frac{1}{p(data_{i} \mid \boldsymbol{q}_{j})} \right]^{-1}$$

where N is the number of MCMC samples. This statistic indicated by small values if the relevant data points were a poor fit to the model.

Fourthly, to compare the fit of the two models we calculated the deviance (Gilks *et al.*, 1992) as the sample mean of the log-likelihoods times –2.

$$\overline{DEV} = \frac{1}{N} \sum\nolimits_{j=1}^{N} (-2) \sum_i \ln\left(p(data_i \mid \boldsymbol{q}_j) \right)$$

and the negative cross-validatory log-likelihood (Gelfan and Dey, 1994):

$$NCL = \sum_i \ln\left(\left[\frac{1}{N} \sum\nolimits_{j=1}^{N} \frac{1}{p(data_i \mid \boldsymbol{q}_j)} \right]^{-1} \right)$$

again N is the number of MCMC samples. These statistics both give small values the better the fit of the model.

Results

Two models were investigated. In each run BUGS was set up to do 5 mill. iterations, however only recording every 100th to compensate for high serial correlation within the chains of some of the sampled parameter values. The first 1000 samples of the recorded chain were discarded for burn-in leaving 49,000 samples as the final result. Convergence diagnostics were calculated for all parameters listed in Table 3 to confirm that the model and sampling set-up was appropriate to ensure convergence.

Model 1 (without explicit predation effect)

A Bayesian model, similar to that described above, but omitting all observations, parameters, equations and distributions concerning predation on shrimp by Atlantic cod, was fitted to the data series on CPUE, survey and catch. This model was similar in concept to that fitted by likelihood methods in Hvingel and Kingsley (2000).

The model was able to produce a relative tight simulation of the observed data (Figure 1). The probabilities of getting a more extreme observation than the realised ones given in the two data series were in the range of 0.12 to 0.5 (Table 2) i.e. the observations did not lie in the tails of their posterior distributions. For the CPUE series, the model was having slight problems capturing the peaks in 1979, 1982, 1987, 1994 and 2000 as also indicated by these points having the largest log-residuals and smallest COP (Table 2). The survey series was generally less well estimated with the values of 1991 and 1994 being the most obvious deviations.

Some of the parameters showed high linear correlation (Table 3). The catchabilities, i.e. the parameters that scale the biomass indices to the real biomass were highly negatively correlated with the parameter for carrying capacity, K, and showed some correlation with MSY and the estimated biomass level of 2001, P_{2001}, as well.

For the parameters P_{msy}, K and P_{zero} the marginal posterior distributions tended to approximate the input priors (Figure 2). However, the model was able to add some precision to the prior estimate of MSY. The posterior was positively skewed and showed a mode at 95 000 tons and upper and lower quartiles at 85 000 and 155 000 tons giving a more optimistic view of MSY than assumed in the prior. The catchabilities, q_s and q_c, showed marked peaks at 0.13 and 0.00036 respectively but had relatively wide posterior distributions. The estimated CV of the observed CPUE series had a median at about 0.09 and for the survey series at 0.14. The process error, given as CV_P, had a median of 11.4%.

The estimated time series of mean biomass level, P_t, ranged from 1.31 to 2.05 times mean Pmsy and the probability that each successive annual shrimp biomass ratio had been greater than the P at MSY was calculated to be larger than 0.9, i.e. it seemed likely that the stock had been maintained above its MSY level throughout the history of the fishery.

Model 2 (cod predation effect included)

Entering a predation effect in the model only had minor influence on the ability of the model to fit the data as judged from the diagnostics presented in table 2 and 4. Actually the model without an explicit cod predation effect showed

a slightly better fit as indicated by smaller values for DEV and NCL (Table 2 and 4). Large correlations among parameters still prevailed (Table x)

The probabilities of getting a more extreme observation than the realised ones were in the range of 0.15 to 0.5 (Table 4). For the CPUE series it was still the peaks in 1979, 1982, 1987, 1994 and 2000 that had the largest log-residuals and smallest COP's (Table 4). Similarly, the greater variability of the survey series was still less well estimated with the values of 1991 and 1994 being the most obvious deviations. However, the model was now able to provide a historic perspective of a likely development of the stock since 1956 (figure 4), which agrees with the contemporary perception of the amount of shrimp at West Greenland (Horsted, pers. comm.).

Compared to the results from model 1, the marginal posterior for the parameter K now had a more peaked distribution with a mode at 650000 t. The posterior distribution for P_{msy}, approximately uniform in the absence of predation, changed to favour values at the lower end of its range i.e. values defining the Fox form of the logistic model (Fox, 1970). The prior for Pzero was updated to a slightly lower mean of 0.72 however no improvement of precision was evident. In shape and precision the posterior for MSY looked similar to the one estimated in the model 1. The mode was at the lower quartile at 100 000 t - the upper quartile at 172 000 t. The catchabilities, q_s and q_c, were less precisely estimated than in model 1 and had modes slightly higher at 0.15 and 0.00048 respectively. The estimated CV's of the observed biomass indices and of the estimated biomass, CV_P, were similar to those estimated in model 1. The posterior for CV of estimated consumption, V, showed the same mode as the prior however, is precision was increased.

As expected there was no information in the data about Q_{max} and the posterior was a copy of the prior. The parameter P_{50} was markedly updated with a posterior showing a peak at 2. The estimated median annual consumption by cod 1955-2000 (figure 7) ranged from 2 000-118 000 tons. When the offshore fishery began in 1976 the amount of shrimp eaten by cod was of the same magnitude as the catches. It then declined with the declining cod stock until a final strong year class of cod showed up and made consumption increase in the mid to late 1980's. After that it declined to almost zero as the cod finally disappeared in the beginning of the 1990's.

References

Anderson, P. J. In press. Pandalid shrimp as indicators of marine ecological regime shift. Pp. 00-00 in P. Koeller, J. Boutillier and S. Tveite, eds. Pandalid shrimp fisheries–science and management at the millennium. Proc. NAFO/ICES/PICES Pandalid Shrimp Symposium, Halifax N.S. 8–10 Sept. 1999. J. Northw. Atl. Fish. Sci. 00.

Apollonio, S., D.K. Stevenson and E.E. Dunton, Jr.. 1986. Effects of temperature on the biology of the northern shrimp, *Pandalus borealis*, in the Gulf of Maine. NOAA Tech. Rep. NMFS 42, 22 pp.

Beverton, R. J. H. and S. J. Holt. 1957. On the dynamics of exploited fish populations. MAFF London. Series 2. 19(12): 172-178.

Anon. 2000a. ACFM report 2000. ICES, Copenhagen, Denmark. http://www.ices.dk. Section 3.1.9 and 3.4.6.

Anon. 2000b. Scientific Council reports 2000. NAFO, Halifax, Canada. http://www.nafo.ca/ publications/redbook/2000/index.htm

Berenboim, B. I., A. V. Dolgov, V. A. Korzhev and N. A. Yaragina. 2000. The impact of cod on the dynamics of Barents Sea shrimp (Pandalus borealis) as determined by multispecies models. *J. Northw. Atl. Fish. Sci. vol. 27: 69-75*

Brooks, S. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics, 7(4), 434-455.

Cadrin, S. X. and S. H. Clark. 1999. Application of catch-survey models to the northern shrimp fishery in the Gulf of Maine. North American J. of Fisheries Management 19:551-568.

Deriso, R. B. 1980. Harvesting strategies and parameter estimation for an age- structured mo del. *Can. J. Fish. Aquat. Sci. 37:268-282*.

Fournier, D. A. and I. J. Doonan. 1987. A length-based stock assessment method utilising a generalized delay-difference model. *Can. J. Fish. Aquat. Sci. 44:422-437*.

Fox, W.W. 1970. An exponential surplus-yield model for optimising exploited fish populations. *Trans. Am. Fish. Soc. 99: 80–88*.

Gelman, A. and Rubin, D. B. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science, 7, 457-511*.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4*, (ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith). Clarendon Press, Oxford, UK.

Gilks, W.R., Thomas, A., and Spiegelhalter, D.J. 1994. A language and program for complex Bayesian modelling. Statistician 43: 169–178

Gompertz, B. On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. Phil. Trans. Roy. Soc. London 123, 513–585, 1825.

Graham, M. 1935. Modern theory of exploiting a fishery and application to North Sea trawling. J. Cons. int. Expl. Mer. 10:264–274.

Heidelberger, P. and Welch, P. 1983. Simulation run length control in the presence of an initial transient. Operations Research, 31: 1109– 1144.

Hilborn, R. and C. J. Walters, 1992. Quantitative fisheries stock assessment. Chapman and Hall, New York, 570 pp.

Holling, C. S. 1959. Some characteristics of simple types of predation and parasitism. *Can. Entomol. 91:385-398*.

Hvingel, C. and O. Folmer. 1998. The Greenlandic Fishery for Northern Shrimp (*Pandalus borealis*) off West Greenland, 1970-1998. *NAFO SCR Doc. 98/111 Serial No. 4020*.

Hvingel, C., H. Lassen, and D. G. Parsons, 2000. A biomass index for northern shrimp (*Pandalus borealis*) in Davis Strait based on multiplicative modelling of commercial catch-per-unit-effort data (1976 - 1997). *J. Northw. Atl. Fish. Sci. vol. 26: 25-36*.

Hvingel, C. and M.C.S. Kingsley (2000). The uncertainty of an assessment procedure for the West Greenland stock of northern shrimp, *Pandalus borealis* (Krøyer). *J. Northw. Atl. Fish. Sci. 27:183–192*.

Kass, R. E. and L. Wasserman. 1996. The selection of prior distributions by formal rules. J. Am. Stat. Ass. Sept. 1996, vol. 91, No. 435, p.1345– 1370.

Koeller, P. A. 2000. The relative importanceof abiotic and biotic factors to the management of northern shrimp (Pandalus borealis) fishery on the Scotian shelff. *J. Northw. Atl. Fish. Sci. vol. 27: 21-33*.

Lilly, G. R., D. G. Parsons and D. W. Kulka. 2000. Was the increase in shrimp biomass on the Norteast Newfoundland shelf a consequence of a release in predation pressure from cod? *J. Northw. Atl. Fish. Sci. vol. 27: 45-61*.

Macdonald, P. D. M. and T.J. Pitcher. 1979. Age-groups from size-frequency data: A versatile and efficient method of analysing distribution mixtures. *J. Fish. Res. Board Can., 36: 987-1011*.

McCrary, J. A., 1971. Sternal spines as a characteristic for differentiating between females of some Pandalidae. *J. Fish. Res. Board Can., 28: 98-100*.

Meyer, R. and R. B. Millar. 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. *Can. J. Fish. Aquat. Sci. 56:37-52*.

Pedersen, S. A. and F. Riget. 1993. Feeding habits of redfish (*Sebastes* spp.) and Greenland halibut (*Reinhardtius hippoglossoides*) in West Greenland waters. *ICES J. Mar Sci. Vol. 50, No. 4, Nov. 1, 1993. pp. 445-459*.

Punt, A. and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries 7, 35-63 (1997).

Raftery, A. L. and Lewis, S. (1992). How many iterations in the Gibbs sampler? In Bayesian Statistics 4, (ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith), pp. 763-774. Oxford University Press, Oxford.

Richards, F.J. 1959 A flexible growth function for empirical use. J. Exp. Bot. 10: 290–300.

Savard, L., I. H. McQuinn and D.G. Parsons. 1991. Investigating the potential of sequential population analysis for northern shrimp (*Pandalus borealis*) in Division 0A. NAFO SCR Doc. 91/71, Serial No. N1955, 12 pp

Schaefer, M. B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Bull. Inter-Am. Trop. Tuna Comm. 1(2): 27–56*.

Schnute, J. 1987. A general firhery model for size structured fish populations. *Can. J. Fish. Aquat. Sci. 44:924-940*.

Schnute, J. 1994. A general framework for developing sequential fisheries models. *Can. J. Fish. Aquat. Sci. 51:1676-1688*.

Shumway, S.E., Perkins, H.C., Shick, D.F., and Stickney, A.P. 1985. Synopsis of biological data on the pink shrimp, Pandalus borealis Krøyer, 1838. *NOAA Technical Report NMFS 30 (FAO Fisheries Synopsis No. 144). 57 p*.

Spiegelhalter, D.J., A. Thomas, N. Best. 2000. WinBUGS version 1.3 User Manual. MRC Biostatistics Unit, Inst. of Public Health, Cambridge, England.

Stefánsson, G., U. Skúladóttir and G. Pétursson. 1994. The use of a stock production type model in evaluating the offshore Pandalus borealis stock of north Icelandic waters, including the predation of northern shrimp by cod. *ICES CM 1994/K:25, 13p. Ref. D,G.*

Table 1. Input data series: catch in 1000 tons, standardised catch-per-unit-effort in the fishery (CPUE), biomass index from research survey and estimates of total cod (G. morhua) biomass.

Year	Catch	CPUE	Survey	Cod
1955				2018
1956				1940
1957				1500
1958				1555
1959				1510
1960				1854
1961				1857
1962				1703
1963				1690
1964				1700
1965				1573
1966				1618
1967				1449
1968				1024
1969				625
1970	8.6			458
1971	9.4			391
1972	9.7			265
1973	12.6			160
1974	22.0			100
1975	37.9			73
1976	50.1	0.949		155
1977	42.1	0.890		143
1978	34.5	0.703		140
1979	35.2	0.636		158
1980	46.0	0.766		125
1981	44.8	0.741		121
1982	44.6	0.932		112
1983	46.8	0.812		87
1984	43.4	0.761		65
1985	54.5	0.795		42
1986	63.1	0.828		43
1987	63.7	1.059		175
1988	60.3	0.786	229.8	234
1989	65.7	0.631	228.0	230
1990	69.4	0.605	228.3	117
1991	75.9	0.591	172.6	38
1992	86.8	0.646	225.1	18
1993	75.6	0.646	256.1	14
1994	76.6	0.609	270.6	7
1995	70.7	0.664	217.1	7
1996	69.2	0.692	248.9	7
1997	64.5	0.664	206.2	7
1998	66.1	0.742	293.3	7
1999	76.5	0.790	287.4	7
2000	76.5	1.000	350.0	7

Table 2. Log-residuals, probability of getting a more extreme observation and conditional predictive ordinate (CPO) in model 1.

Year	CPUE-series resid.(%)	p.extreme	CPO	Survey-series resid.(%)	p.extreme	CPO
1976	4.1	0.37	6.4			
1977	3.9	0.35	7.4			
1978	-4.0	0.35	9.2			
1979	-7.8	0.24	6.0			
1980	1.7	0.46	10.5			
1981	-3.9	0.36	10.1			
1982	7.5	0.25	4.6			
1983	-0.2	0.49	11.2			
1984	-2.6	0.40	10.5			
1985	-1.2	0.45	6.4			
1986	-1.0	0.45	10.8			
1987	13.2	0.14	0.4			
1988	4.1	0.35	9.8	-14.2	0.19	0.017
1989	-4.1	0.34	12.4	-1.0	0.47	0.033
1990	-2.9	0.39	14.8	4.4	0.40	0.031
1991	-0.5	0.49	16.3	-18.8	0.12	0.014
1992	1.0	0.46	15.5	0.2	0.50	0.034
1993	-2.1	0.42	14.4	10.5	0.26	0.021
1994	-7.2	0.25	7.8	16.3	0.15	0.011
1995	0.8	0.47	15.1	-6.2	0.36	0.031
1996	1.1	0.46	14.5	3.7	0.42	0.029
1997	-1.0	0.45	14.5	-13.2	0.21	0.020
1998	-1.8	0.44	13.2	10.1	0.26	0.018
1999	-2.5	0.40	11.6	1.1	0.48	0.026
2000	7.4	0.25	4.0	7.2	0.33	0.017
Dev. mean	-126.6			96.3		
Dev. min.	-165.1			81.5		
NCL	-54.7			49.7		

Table 3. Significant correlations (p<0.01) among parameters in model 1.

	K	MSY	P2001	Pmsy	Pzero	sd c	sd P	sd s	q
MSY	0.25								
P2001	0.32	0.18							
Pmsy	0.04	-0.18	0.02						
Pzero	0.03	ns	0.07	ns					
sd c	ns	-0.01	0.04	0.03	ns				
sd P	0.05	ns	-0.04	ns	ns	-0.18			
sd s	ns	0.01	0.02	ns	ns	ns	-0.07		
qc	-0.84	-0.29	-0.41	-0.06	-0.04	ns	-0.07	-0.01	
qs	-0.84	-0.30	-0.41	-0.06	-0.04	ns	-0.07	-0.01	1.00

Table 4. Mean, standard deviation and 25, 50, and 75 percentiles of MCMC samples from the posterior distribution for selected parameters in model 1.

Parameter	Mean	sd	25%	Median	75%
K	1519	708.4	917.1	1426	2081
MSY	128.4	64.99	85.63	113.5	155.5
P2001	0.9455	0.1613	0.8384	0.9353	1.039
Pmsy	0.5011	0.08284	0.4283	0.4966	0.5724
Pzero	1.007	0.1967	0.8745	1.009	1.14
precCPUE	139.2	56.18	98.34	130.2	169.9
precP	82.68	31.97	59.99	77.25	99.32
precsurv	52.74	20.73	37.78	49.71	64.46
q	8.00E-04	4.92E-04	4.33E-04	6.44E-04	0.001021
qs	0.2808	0.1737	0.1516	0.2252	0.3571

Table 5. log-residuals, probability of getting a more extreme observation and conditional predictive ordinate (CPO) in model 2

Year	CPUE-series resid.(%)	p.extreme	CPO	Survey-series resid.(%)	p.extreme	CPO
1976	5.6	0.31	5.1			
1977	5.5	0.31	6.0			
1978	-3.9	0.37	10.5			
1979	-8.4	0.23	6.0			
1980	1.6	0.44	11.2			
1981	-3.7	0.37	10.4			
1982	8.0	0.24	3.8			
1983	0.0	0.49	11.3			
1984	-2.9	0.40	10.8			
1985	-1.5	0.45	11.4			
1986	-2.2	0.42	10.4			
1987	12.1	0.15	0.7			
1988	3.9	0.37	10.4	-14.1	0.19	0.017
1989	-3.9	0.37	13.4	-0.5	0.49	0.033
1990	-2.0	0.43	16.0	5.6	0.37	0.030
1991	0.0	0.50	17.3	-18.1	0.12	0.014
1992	0.9	0.46	15.7	0.5	0.49	0.034
1993	-2.1	0.43	14.7	10.7	0.25	0.020
1994	-7.8	0.24	7.8	16.1	0.15	0.012
1995	0.1	0.49	15.2	-6.6	0.33	0.030
1996	0.4	0.48	14.8	3.3	0.42	0.030
1997	-2.1	0.42	14.7	-14.0	0.19	0.019
1998	-2.5	0.40	12.4	9.7	0.26	0.019
1999	-3.4	0.38	11.3	0.5	0.48	0.026
2000	7.1	0.26	4.5	7.2	0.33	0.017
Dev. mean	-70.0			125.9		
Dev. min.	-103.5			112.2		
NCL	-25.4			64.5		

Table 6. Significant correlations (p<0.01) among parameters in model 2.

	K	MSY	Omax	P2001	P50	Pmsy	Pzero	sd c	sd P	sd V	sd s	q
MSY	0.57											
Omax	ns	ns										
P2001	0.17	0.45	ns									
P50	-0.22	-0.33	ns	0.10								
Pmsy	-0.07	-0.12	ns	0.16	0.24							
Pzero	0.25	0.12	-0.01	0.13	0.22	0.06						
sd c	-0.03	-0.02	ns	0.03	-0.01	ns	-0.02					
sd P	-0.18	-0.13	ns	-0.11	-0.07	-0.04	-0.11	-0.15				
sd V	ns	ns	ns	0.02	ns	0.01	-0.02	ns	-0.02			
sd s	ns	0.03	ns	0.04	-0.02	ns	ns	ns	-0.07	ns		
q	-0.82	-0.59	ns	-0.40	0.15	0.02	-0.29	0.04	0.17	-0.01	-0.02	
qs	-0.82	-0.59	ns	-0.40	0.15	0.02	-0.29	0.04	0.17	-0.01	-0.02	1.00

Table 7. Mean, standard deviation and 25, 50, and 75 percentiles of MCMC samples from the posterior distribution for selected parameters in model 2.

| Parameter	Mean	sd	25%	Median	75%
K	1259	669.7	716.3	1067	1686
MSY	148.1	73.66	99.75	125.3	171.8
Omax	2.494	0.1415	2.399	2.493	2.589
P2001	0.849	0.1576	0.756	0.8505	0.9466
P50	2.315	0.8679	1.69	2.162	2.79
Pmsy	0.4735	0.07623	0.408	0.4601	0.5298
Pzero	0.7199	0.2035	0.5796	0.719	0.8581
precCPUE	140.9	56.03	100.7	131.8	171.4
precP	98.67	40.14	70.19	91.73	119.4
precV	66.28	46.79	31.95	55.4	89.13
precsurv	53.04	20.78	38.03	49.92	64.71
q	0.00112	0.00061	0.00062	0.001	0.001509
qs	0.3927	0.2134	0.2172	0.3482	0.529
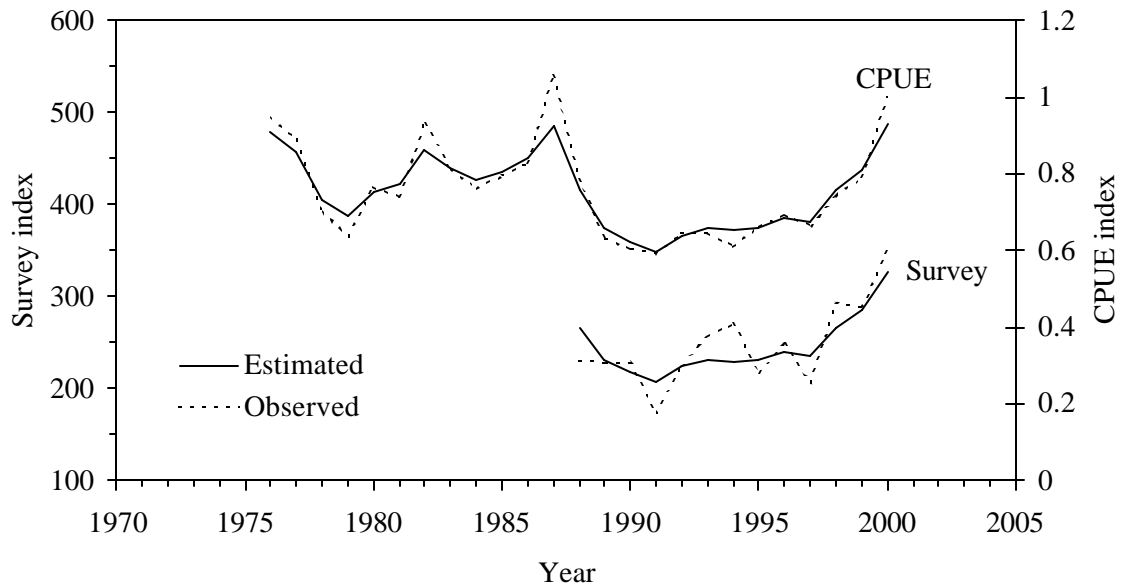
Fig. 1. Observed biomass indices based on standardised catch-per-unit-effort data (CPUE) from the fishery, swept area biomass index estimates from the Greenland trawl survey and corresponding medians of their posterior distributions estimated by model 1.
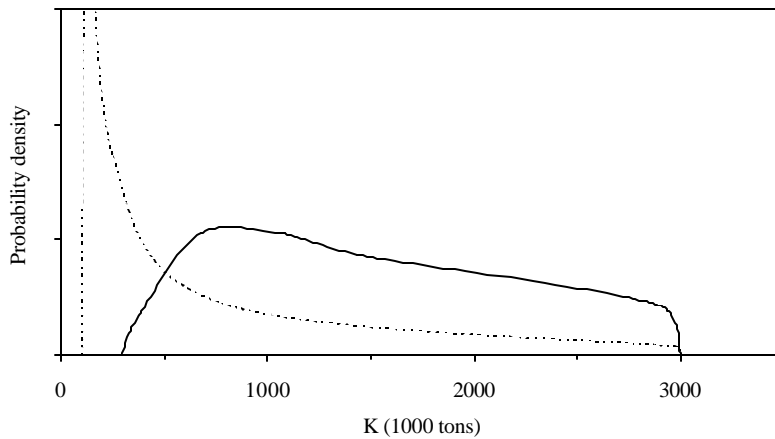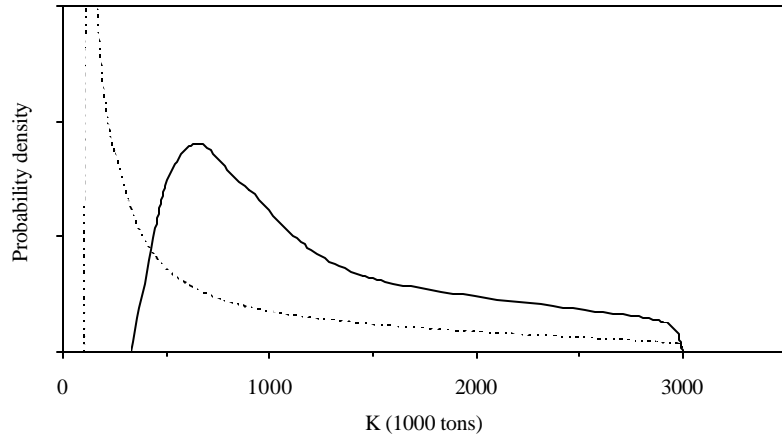

Fig. 2a-j. Prior (broken lines) and posterior (solid lines) density functions of parameters in model 1 as obtained from the MCMC samples.
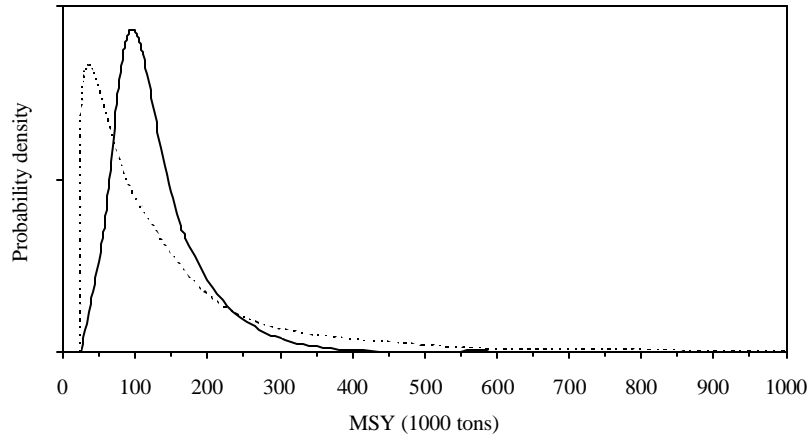
Fig. 2b.
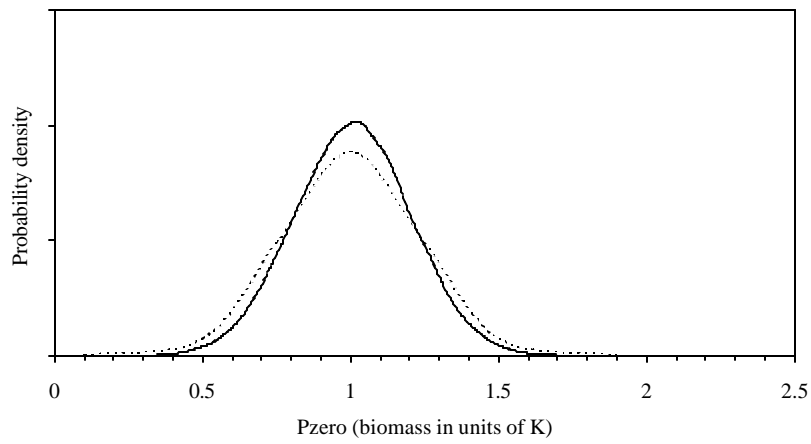
Fig. 2c.
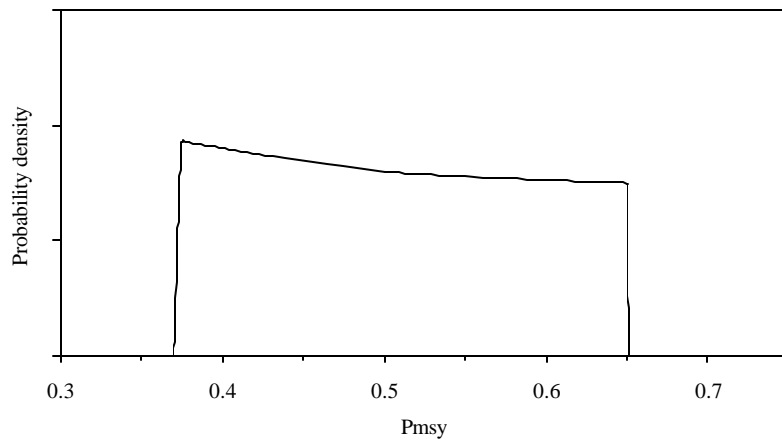
Fig. 2d.
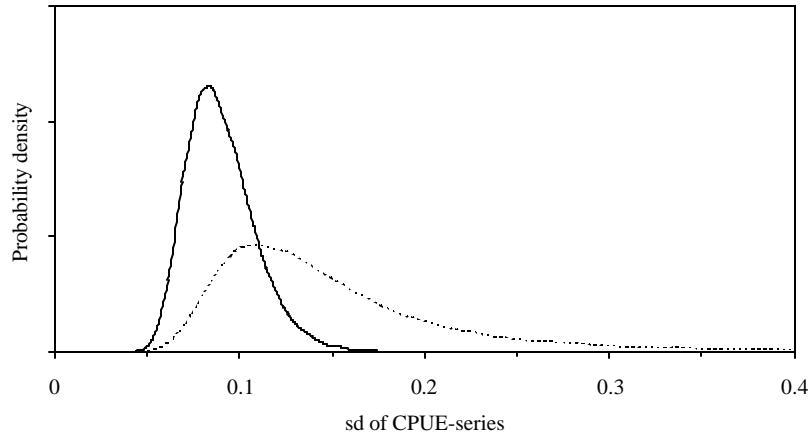
Fig. 2e.
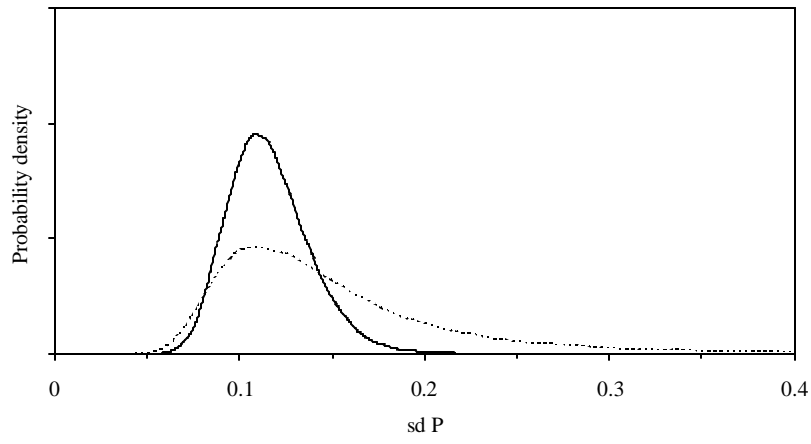
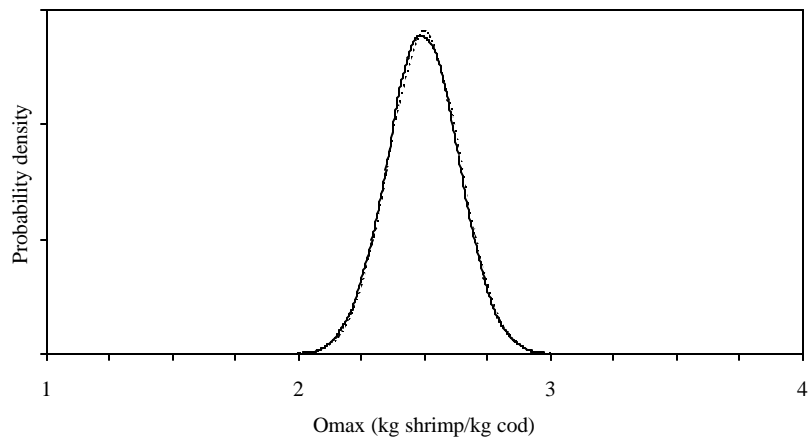**Fig. 2f.**
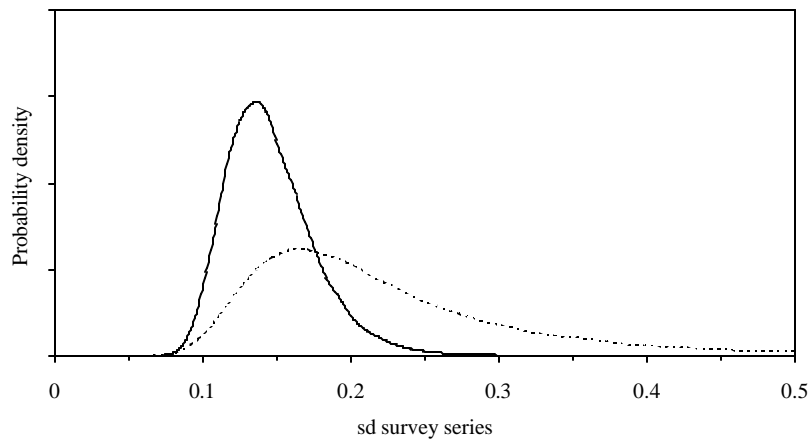
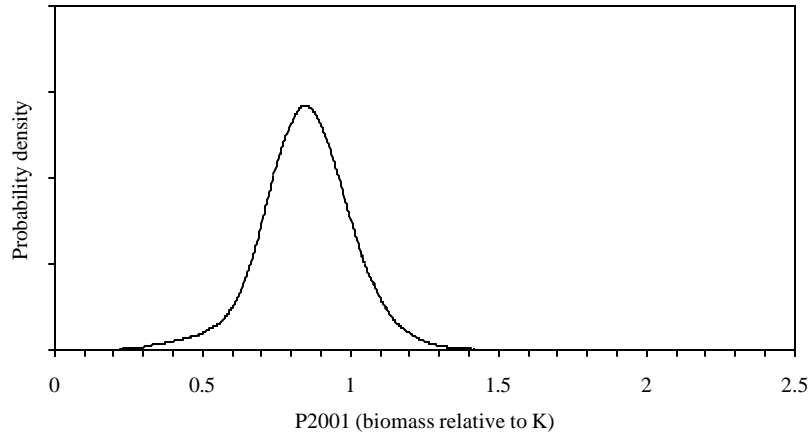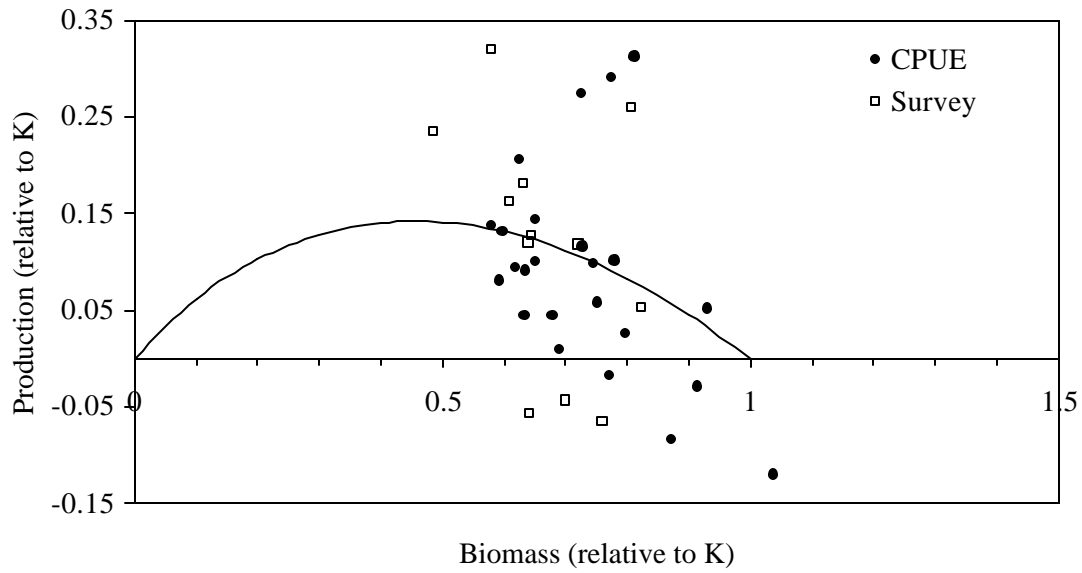**Fig. 2g.**
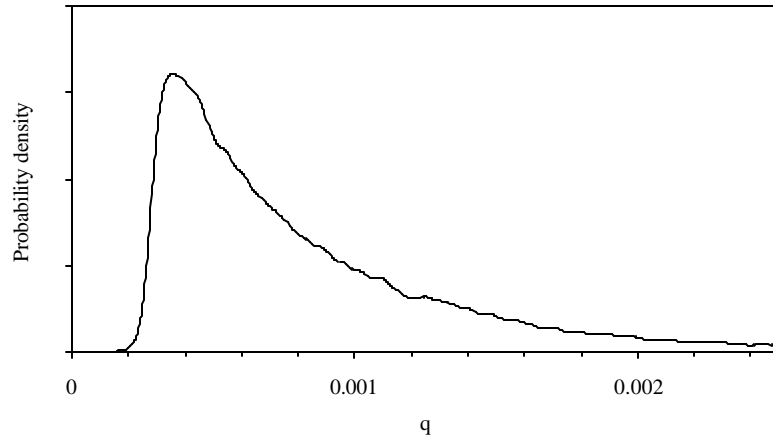
Fig. 2h.
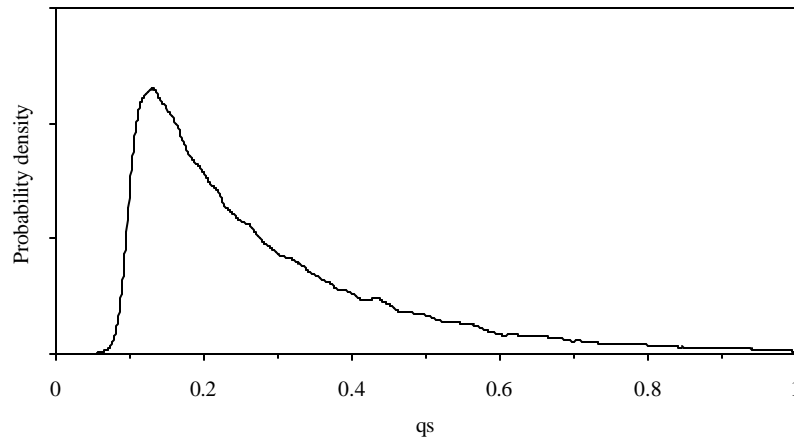
Fig. 2i.
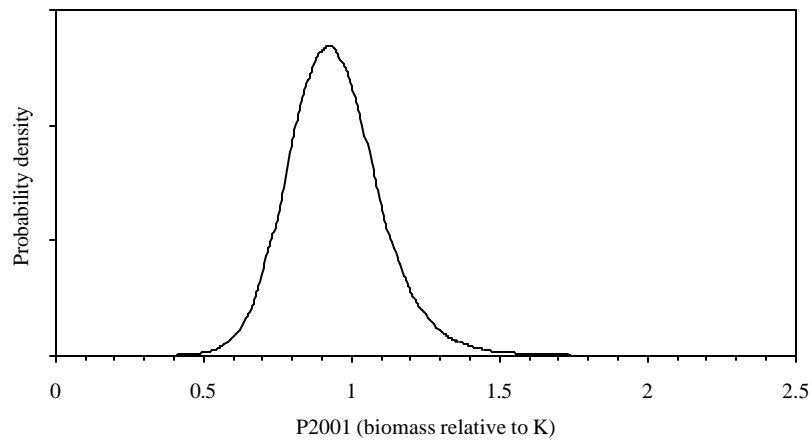
Fig. 2j.
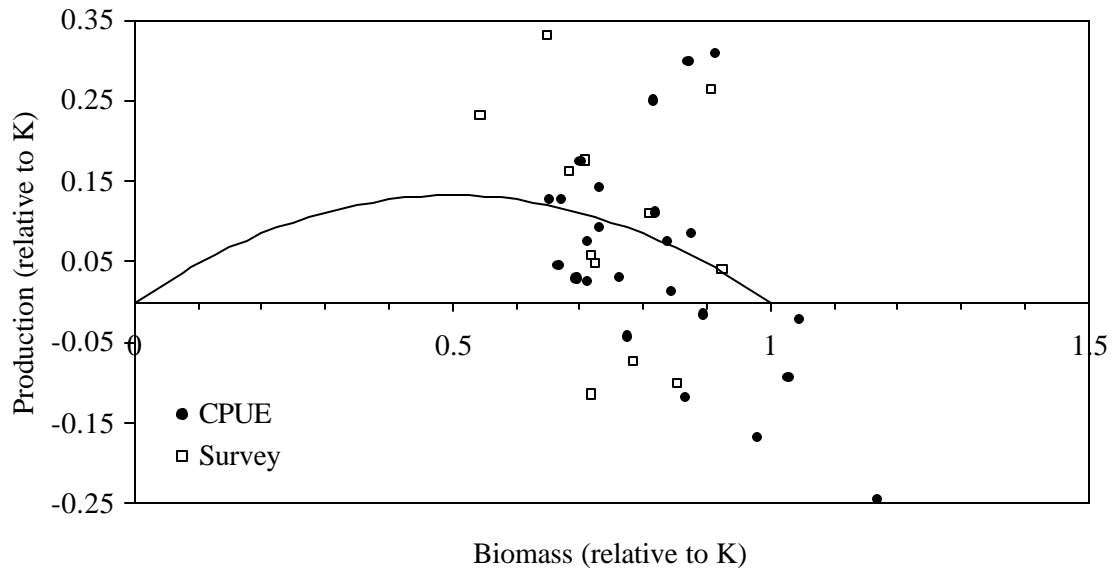
Fig. 3. Production *vs.* stock size as predicted from the indices using model 1 and a "best fit" (least squares) stock-recruitment curve.
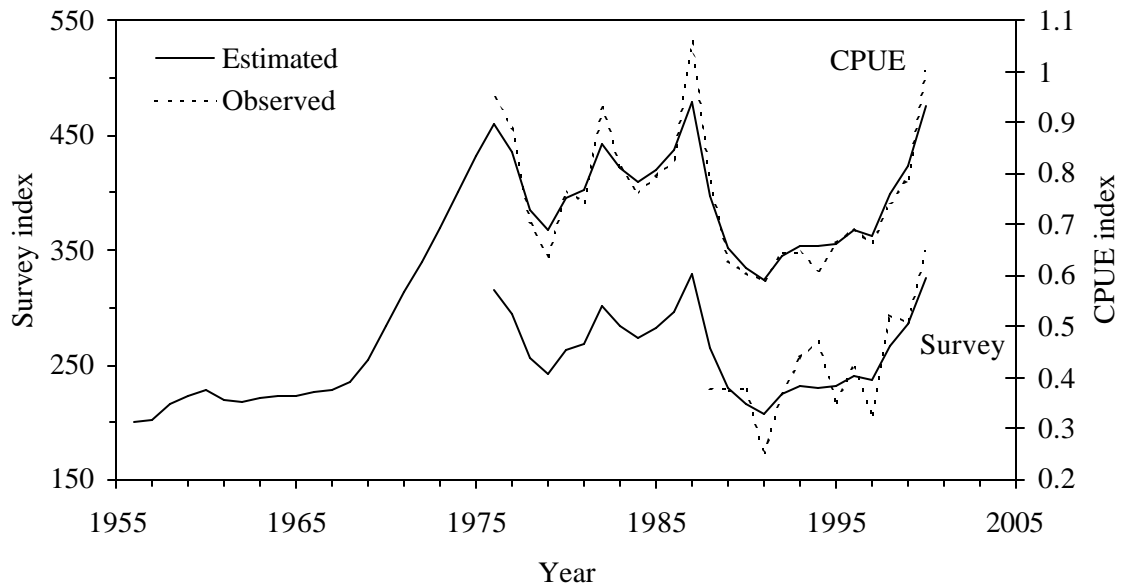
Fig. 4. Observed biomass indices based on standardised catch-per-unit-effort data (CPUE) from the fishery, swept area biomass index estimates from the Greenland trawl survey and corresponding medians of their posterior distributions estimated by model 2.

Fig. 5a-m. Prior (broken lines) and posterior (solid lines) density functions of parameters in model 1 as obtained from the MCMC samples.
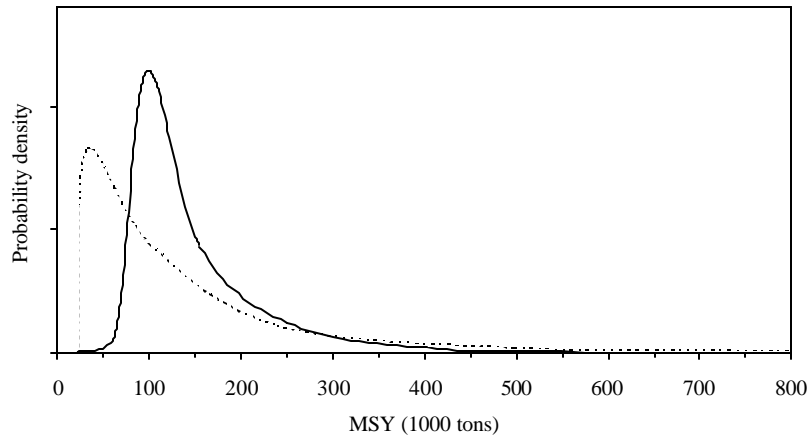
Fig. 5b.
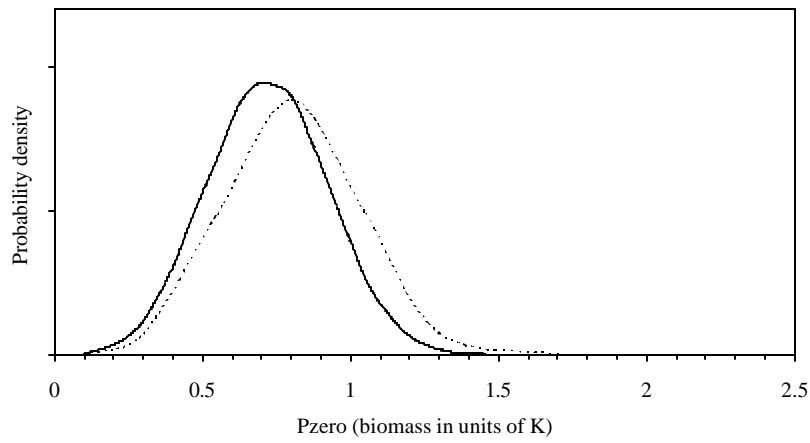
Fig. 5c.

Fig. 5d.

Fig. 5e.

Fig. 5f.

23
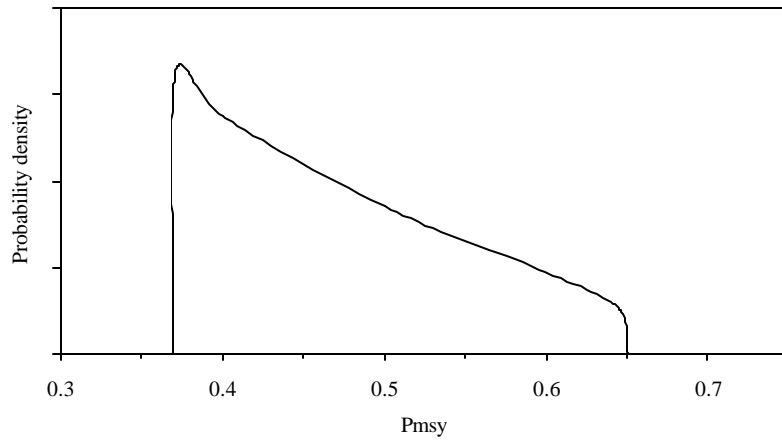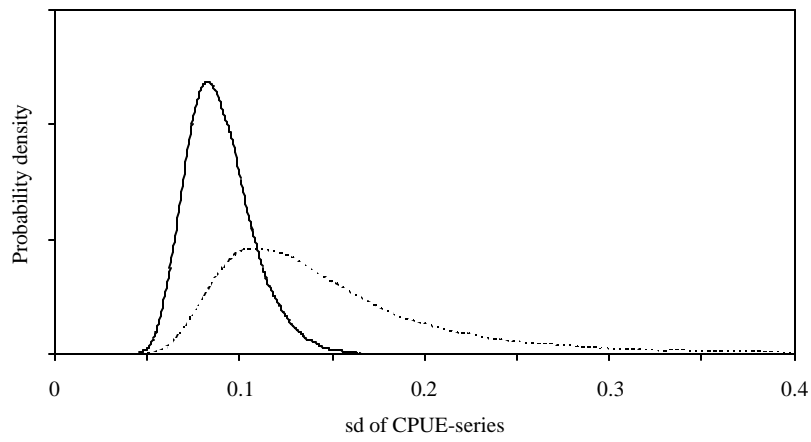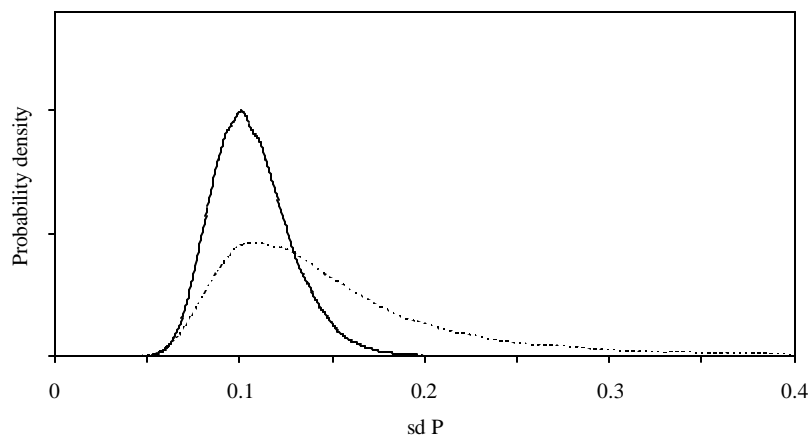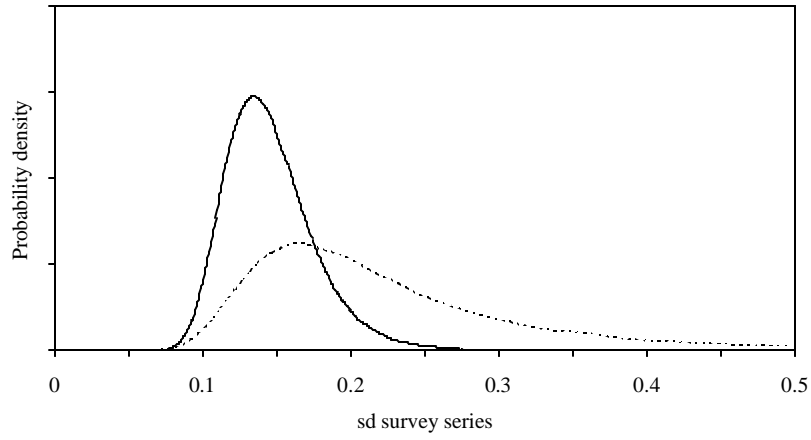
Fig. 5g.
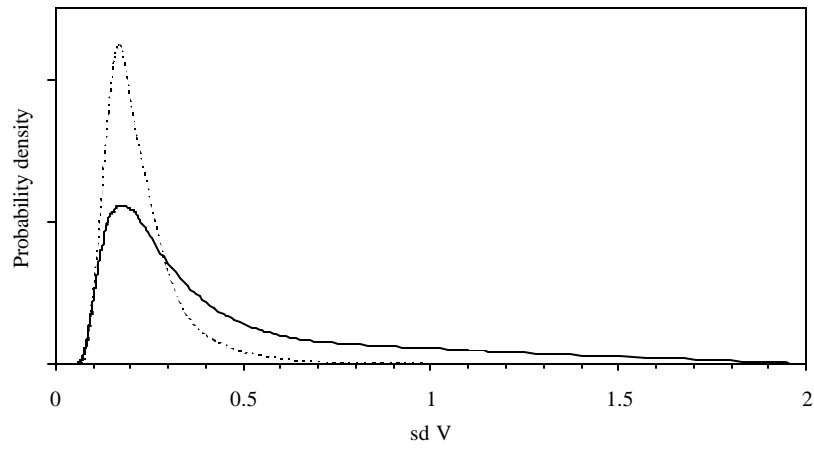
Fig. 5h.
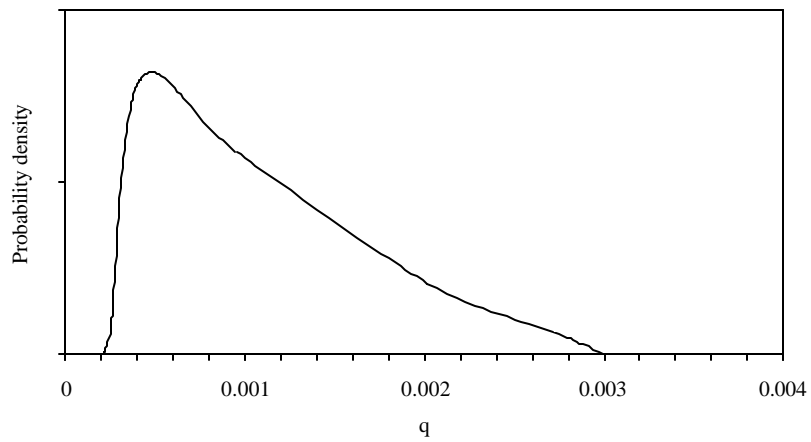
Fig. 5i.
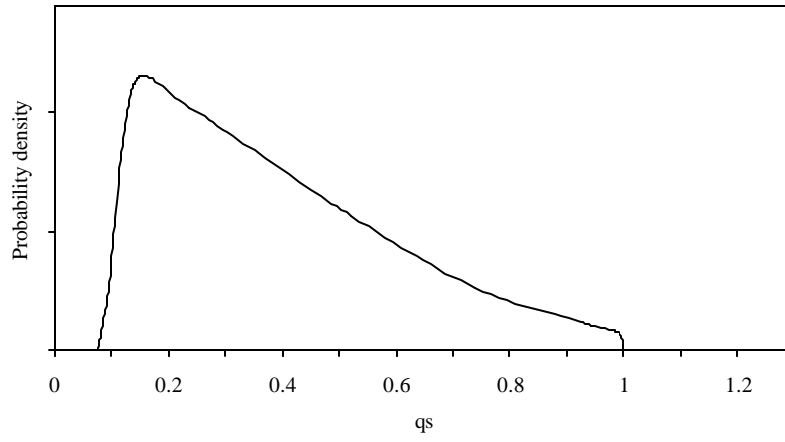
Fig. 5j.
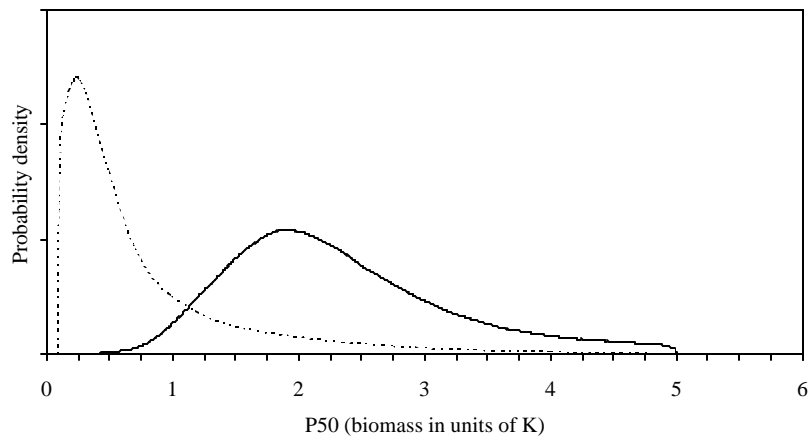
Fig. 5k.

Fig. 5l.

Fig. 5m.

Fig. 6. Production *vs.* stock size as predicted from the indices using model 2 and a "best fit" (least squares) stock-recruitment curve.
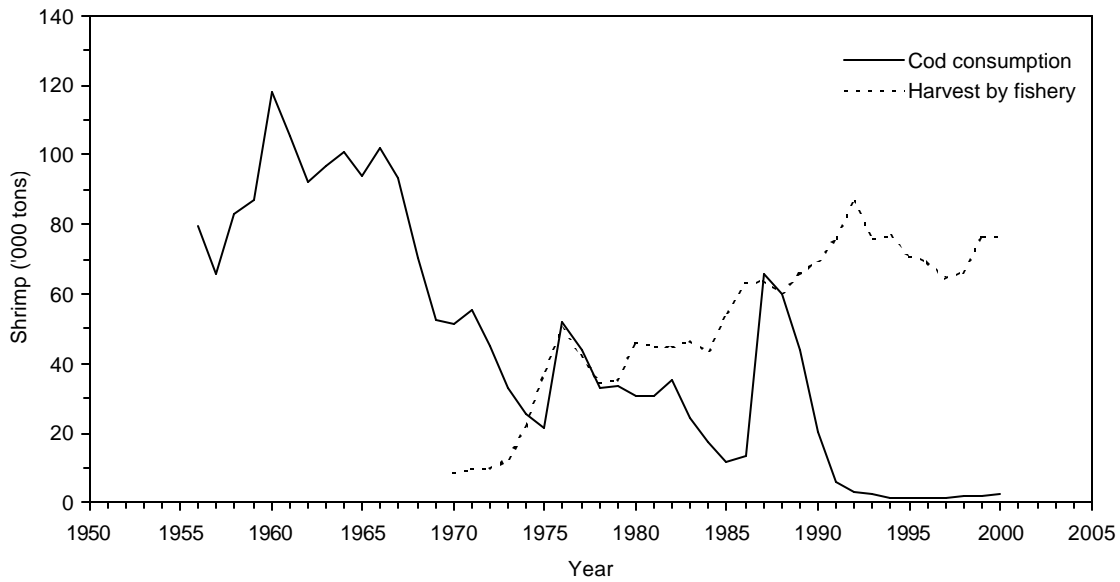
Fig. 7. Annual removal (tons) of shrimp from the West Greenland stock by cod and by the fishery. Cod consumption is the median of the posteriors as estimated by model 2
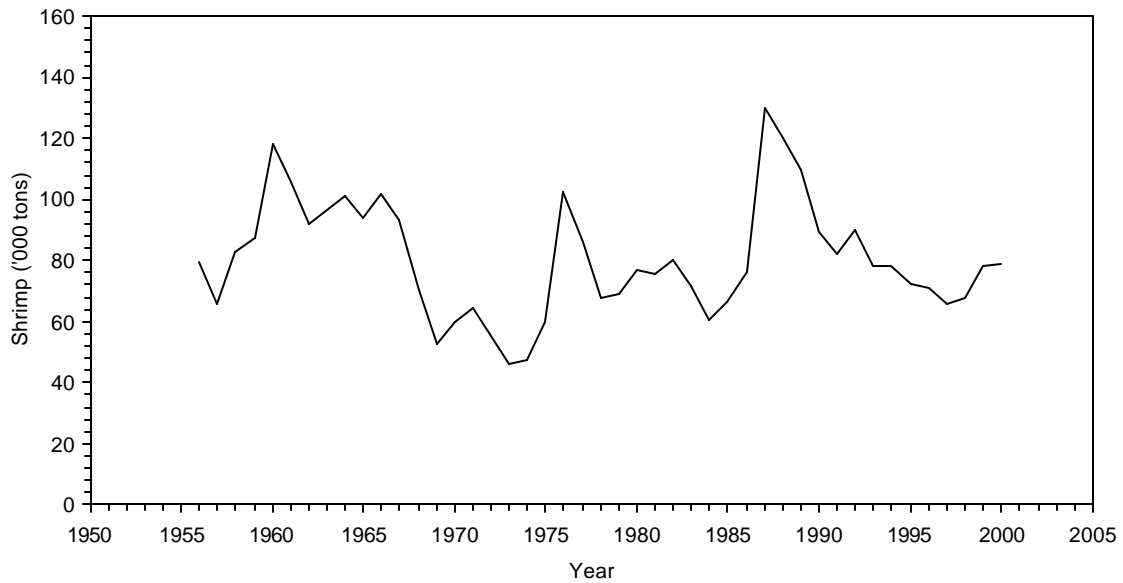

Fig. 8. Summed annual removal (tons) of shrimp from the West Greenland stock by cod and the fishery. (Cod consumption is the median of the posteriors as estimated by model 2)