Nonlinear state-space modeling of fisheries biomass dynamics using Metropolis-Hastings within Gibbs sampling

By Russell B. Millar and Renate Meyer

University of Auckland, New Zealand

SUMMARY

State-space modeling and Bayesian analysis are both active areas of applied research in fisheries stock assessment. Combining these two methodologies facilitates the fitting of state-space models that may be nonlinear and have non-normal errors, and hence it is particularly useful for the modeling of fisheries dynamics. Here, this approach is demonstrated by fitting a non-linear surplus production model to data on South Atlantic albacore tuna (*Thunnus alalunga*), The state-space approach allows for random variability in both the data (measurement of relative biomass) and in annual biomass dynamics of the tuna stock. Sampling from the joint posterior distribution of the unobservables was achieved using Metropolis-Hastings within Gibbs sampling.

Keywords: Bayesian analysis; Fish stock assessment; Markov-chain Monte-carlo; Nonlinear state-space models; Surplus production models; Tuna

Address for correspondence: Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand

1 Introduction

Bayesian methodology and state-space modeling have both been prominent in the recent fisheries literature (e.g., Sullivan, 1992; Pella, 1993; Hilborn *et al.*, 1994; McAllister *et al.*, 1994; Schnute, 1994; Walters and Ludwig, 1994; Reed and Simons, 1996; Kinas, 1996; Punt and Hilborn, 1997) and these two approaches are regarded by some as the future for stock assessment methodology (e.g., Hilborn, 1992; Kimura *et al.*, 1996).

The Bayesian approach to fisheries stock assessment is promoted as a natural way to portray uncertainty about key population parameters, and to express the risks associated with alternative management decisions (e.g., Hilborn, 1992). Moreover, it also permits knowledge about other populations of the same (or similar) species to be incorporated as prior knowledge. This is viewed by many fisheries scientists as a coherent way to utilize the vast amount of existing information that is held in fisheries databases throughout the world (e.g., Liermann and Hilborn, 1997; Hilborn and Liermann, 1998; Myers *et al.*, in press).

In the context of fisheries population dynamics, the state-space paradigm explicitly models the randomness in both the dynamics of the population and in the observations made on the population. Depending on the type of data collected, the relevant state of the population will typically be the total biomass of all fish above the minimum legal size, or the biomass (or numbers) of fish at a range of age- or length-classes. In the former case, the state equation might specify the biomass in the following year as a function of the current biomass, additions due to fish growth and recruitment of new individuals to the minimum legal size, and removals from fishing or natural mortality. The observed variable would often be a relative measure of biomass obtained from catch rates of commercial and/or research fishing.

The majority of fisheries population models in current use are observation-error models (Hilborn and Walters, 1992; Polacheck *et al.*, 1993), that is, they model only random error in the observations and assume that the state equation is deterministic. With this assumption, the entire history of the states of the fishery are found deterministically from specification of the parameters of the state equation and of the state of the population at the onset of fishing. It is then straightforward to specify a likelihood for the observed data. Some fisheries modelers have also considered process-error models (e.g., Breen, 1991) in which random error in the state equation is modeled, but the observations are assumed to be deterministic given the states. The general consensus appears to be that if only one of these sources of randomness can be modeled then observation-error models are preferable (Hilborn and Walters, 1992; Polacheck *et al.*, 1993).

The Kalman filter (Kalman, 1960) has recently been used to incorporate both observation and process error in likelihood-based models of catch-at-length (Sullivan, 1992), catch-at-age (Schnute, 1994), delay-difference biomass (Kimura *et al.*, 1996), and special cases where linear biomass dynamics could be posed (Freeman and Kirkwood, 1995; Reed and Simons, 1996). A number of these authors have acknowledged that use of the Kalman filter sacrifices some realism by requiring the state and observation equations to be linear and the errors to be normally distributed. The extended Kalman filter uses linear approximation to fit nonlinear state-space models, and has been used by Pella (1993) to model total biomass, and by Gudmundsson (1994, 1995) to model catch-at-age and catch-at-length data. A common conclusion of these authors is that it is difficult to obtain reliable maximum likelihood estimates of the process error and observation error variances, and it is usual to include additional information into the model by taking the ratio of these variances to be known.

Penalized likelihood has been used (e.g., Ludwig *et al.*, 1988; Schnute, 1994; Richards and Schnute, 1998) as an alternative to the Kalman filter. This approach treats process errors as fixed parameters to be estimated, and hence fitting of the model reduces to a conceptually simple maximization problem because the likelihood for the observables is easily determined for any specified values of the model parameters and process errors. This methodology has the advantage of being very generally applicable, but the disadvantage of undesirable properties under common circumstances. For example, in the context of using penalized likelihood to fit generalized linear mixed models, the estimates of fixed effects are not consistent when there are limited data per random effect (which is typically the case), and asymptotic bias correction formula have been provided by Lin and Breslow (1996)

This paper combines the Bayesian and state-space techniques for purposes of stock assessment and demonstrates that nonlinear equations and non-normal errors are easily accommodated. It generalizes the approach of Carlin *et al.* (1992) by using Metropolis-Hastings within Gibbs sampling (Gilks, 1996) to sample from arbitrary full conditional densities. The example uses a Schaefer surplus production model (Schaefer, 1954; Ricker, 1975). This model is conceptually simple and has a state equation that is nonlinear with respect to two biologically interpretable parameters. Moreover, it is very commonly used in fisheries stock assessment worldwide.

2 Surplus Production Model

In the surplus production model, the (deterministic) state equation for total biomass is

$$B_y = B_{y-1} + g(B_{y-1}) - C_{y-1} \ , \tag{1}$$

where B_y is biomass at the start of year y, C_y is catch during year y, and the surplus production function, $g(B)$, quantifies the overall change in biomass due to growth, recruitment, and natural mortality (Ricker, 1975). If year 1 is the year in which fishing commenced then it is usual to assume $B_1 = K$, where B_1 is virgin biomass and K is the carrying capacity of the stock habitat.

Carrying capacity is assumed to be the level at which additions due to growth and recruitment are balanced by removals due to natural mortality, that is, $g(K) = 0$. Surplus production is assumed to be positive at stock levels below K because this enables the biomass to rebuild (in the absence of fishing) toward its carrying capacity.

The (deterministic) observation equation is typically

$$I_y = qB_y \ , \tag{2}$$

where I_y is a relative biomass index, and q is the "catchability coefficient". In practice, this index is often catch-per-unit-effort (CPUE) calculated as the total catch divided by total fishing effort. It may be calculated from commercial fishing data or from research surveys.

In a recent critique, Polacheck *et al.* (1993) compared three popular methods for fitting the model defined by equations (??) and (??). None of these methods allowed for random error in both equations, but they did include the process-error model (error in equation (??) only) and observation-error model (error in equation (??) only). Polacheck *et al.* (1993) used the simple quadratic form of surplus production proposed by Schaefer (1954),

$$g(B_y) = rB_y\left(1 - \frac{B_y}{K}\right) \ . \tag{3}$$

The unknowns, r (intrinsic growth rate of population) and K (carrying capacity), are of immediate relevance to fisheries managers. For example, the maximum surplus production (MSP) of $rK/4$ occurs when $B = K/2$. When the biomass indices are CPUE's from commercial fishing then the Schaefer surplus production model can be used to determine optimal effort (E_{opt}), defined to be the level of commercial fishing effort (e.g. number of hooks to be deployed) required to harvest MSP when $B = K/2$. From equation (??), $\mathrm{MSP}/\mathrm{E}_{opt} = qK/2$, giving $\mathrm{E}_{opt} = r/2q$.

Surplus production models require only time series of catches and relative biomass indices and hence are widely used in fisheries assessment. When additional knowledge about the stock dynamics (e.g., growth of individuals, mortalities, fecundity, recruitment) is available, or when more complete data are obtained (e.g. catch-at-age), then a more complex model could be considered.

State-space version

Suggested error structures for the state and observation equations include additive normal with fixed variance, additive normal with fixed coefficient of variation (CV), and multiplicative lognormal error (Polacheck *et al.*, 1993). When such analyses are not restricted by computational considerations, fisheries modelers tend to choose multiplicative lognormal errors. (For example, Kimura *et al.* (1996) used lognormal errors in an observation-error analysis, but switched to normal errors to implement a Kalman filter analysis of the same data.) Multiplicative lognormal error is assumed here. Using the Schaefer surplus production model (equation (**??**)), the stochastic form of the process equations (**??**) can be written

$$\begin{aligned} \log(B_1)|K,\sigma^2 &= \log(K) + u_1 \\ \log(B_y)|B_{y-1},K,r,\sigma^2 &= \log(B_{y-1} + rB_{y-1}(1 - B_{y-1}/K) - C_{y-1}) + u_y \ , \\ & \qquad\qquad\qquad\qquad\qquad\qquad y = 2, 3, ... \end{aligned} \tag{4}$$

and the stochastic form of the observation equations (**??**) is

$$\log(I_y)|B_y,q,\tau^2 = \log(q) + \log(B_y) + v_y \ , \qquad y = 1, 2, ... \tag{5}$$

where the u_y's and v_y's are iid $N(0, \sigma^2)$ and iid $N(0, \tau^2)$ random variables, respectively.

3 Application to South Atlantic Albacore Tuna

The above model was applied to the South Atlantic albacore tuna (*Thunnus alalunga*) data (Table 1) analysed in Polacheck *et al.* (1993). The biomass index is catch-per-unit-effort measured as kg of tuna caught per 100 hooks deployed.

Specification of priors

The unobservables are $(K, r, \sigma^2, q, \tau^2, B_{1967}, ..., B_{1989})$. It is enough to specify the prior on $(K, r, \sigma^2, q, \tau^2)$ because the joint prior for $(K, r, \sigma^2, q, \tau^2, B_{1967}, ..., B_{1989})$ is then determined using the specification of the conditional distributions in equation (**??**) (see Appendix).

With regard to specification of priors for Bayesian stock assessment, the prevailing recommendation (e.g., Walters and Ludwig, 1994; Punt and Hilborn, 1997; Hilborn and Liermann, 1998) is to use non-informative priors except when informative priors can be obtained by formal means. Indeed, considerable effort is now focused on using hierarchical models to obtain formal prior distributions for certain key population parameters (e.g., Liermann and Hilborn, 1997; Myers *et al.*, 1997; Myers *et al.*, in press). Not all parameters are equally amenable to this approach. For example, biologically meaningful parameters such as r (intrinsic growth rate of the population) could be assumed exchangeable over various stocks of the same species or family (Gelman *et al.*, 1995). However, quantities such as K (carrying capacity) will depend on stock-specific covariates such as habitat range.

Here, a prior distribution for r was obtained from hierarchical modeling using twelve other tuna stocks (see Myers *et al.*(in press) for details of this methodology). Work is currently proposed to obtain formal prior distributions for other population parameters such as K and σ^2 (Ransom Myers, pers. comm.). In the meantime, the first model herein uses weakly informative priors for K and σ^2 that were informally derived from existing information and are described below. The sensitivity to these two priors is examined and reported in the Results section.

Prior for carrying capacity, K

Punt *et al.* (1995) applied a Bayesian analysis of an (observation error) age-structured model to this stock, and specified a Uniform[80,300] prior on the carrying capacity K (in 1000s t). However, this was not a good choice and indeed, Walters and Ludwig (1994) warned against using uniform prior distributions on finite intervals if it assigns zero prior probability to feasible values of the unknown. Punt and Hilborn (1997) subsequently supported this recommendation. For application here, the values of 80 and 300 were therefore taken to express an interval of (moderately) high prior probability for K, and were taken to be the 10 and 90 percentile points (respectively) of a lognormal distribution. These percentiles equate to a lognormal random variable with mean and standard deviation of 5.04 and 0.516 (respectively) on the log scale.

Prior for intrinsic growth rate of population, r

The hierarchical analysis of Myers *et al.* (in press) induces a lognormal prior for r with mean of -1.38 and standard deviation of 0.51. These correspond to 10 and 90 percentiles for r of 0.13 and 0.48, respectively.

Prior for process error variance, σ^2

Much of the process variability will arise from recruitment variability. Examination of recruitment data on South Pacific albacore for the years 1959 to 1990 (available from http://www.mscs.dal.ca/~myers/welcome.html) gave a CV of 0.34 on recruitment at age 3. These age-3 fish correspond to approximately 12% of the overall biomass (calculated using the growth curve and natural mortality specified in Punt *et al.* (1995)), giving a CV of 0.04 on total biomass due to recruitment variability alone. To allow for the additional variability of natural mortality and growth rates, the upper bound on CV was taken to be twice this, 0.08. The values 0.04 and 0.08 were taken to be the bounds on an interval of (moderately) high prior probability for the coefficient of variation, and an inverse gamma prior on σ^2 was specified such that the 10 and 90 percentiles on coefficient of variation were 0.04 and 0.08 respectively.

Prior for catchability, q

A uniform prior was chosen for $\log(q)$. The quantity $\log(q)$ acts as an intercept term in equation (**??**) and hence this can be considered a non-informative prior (Kass and Wasserman, 1996). This prior for q has previously been used by McAllister *et al.* (1994), Walters and Ludwig (1994), and Punt *et al.* (1995).

Prior for observation error variance, τ^2

In practice, the observed CPUE will typically be obtained from analysing the log-books of selected fishing vessels, and hence knowledge of the magnitude of sampling variability can be deduced. These log-book data were not available for this analysis, and so for purposes of this example a CV on CPUE of around 10% was used. Specifically, an inverse gamma prior on τ^2 was specified such that the 10 and 90 percentiles on coefficient of variation were 0.05 and 0.15 respectively.

Joint prior for $(K, r, \sigma^2, q, \tau^2)$

Punt and Hilborn (1997) use biological considerations to argue that the priors on K and r can reasonably be assumed independent, and moreover, Walters and Ludwig (1994) and Kinas (1996) use mutually independent priors on K, r, q and any variance parameters. This practice is followed here, and using $\pi(\cdot)$ to denote prior densities, the joint prior is taken to be $\pi(K, r, \sigma^2, q, \tau^2) = \pi(K)\pi(r)\pi(\sigma^2)\pi(q)\pi(\tau^2)$.

Sampling from the posterior using Metropolis-Hastings within Gibbs sampling

The Gibbs sampler (Gelfand and Smith, 1990) is a numerical technique for sampling from the joint posterior distribution, $f(\theta_1, \theta_2, ..., \theta_n|\boldsymbol{x})$, where $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ are the unknowns and $\boldsymbol{x}$ denotes the observables. Given a starting vector $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \ldots, \theta_n^{(0)})$ the Gibbs sampler proceeds by sampling from the univariate full-conditional posteriors as follows

$$\begin{aligned}
\text{simulate } \theta_1^{(1)} &\sim f(\theta_1|\theta_2^{(0)}, \ldots, \theta_n^{(0)}, \boldsymbol{x}) \\
\text{simulate } \theta_2^{(1)} &\sim f(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, \ldots, \theta_n^{(0)}, \boldsymbol{x}) \\
&\vdots \\
\text{simulate } \theta_n^{(1)} &\sim f(\theta_n|\theta_1^{(1)} \ldots, \theta_{n-1}^{(1)}, \boldsymbol{x})
\end{aligned}$$

and yields $\boldsymbol{\theta}^{(m)} = (\theta_1^{(m)}, \ldots, \theta_n^{(m)})$ after m such cycles. This defines a Markov chain with transition kernel $k(\boldsymbol{\theta}^{(m+1)}, \boldsymbol{\theta}^{(m)}) = \prod_{i=1}^n f(\theta_i^{(m+1)}|\theta_1^{(m+1)}, \ldots, \theta_{i-1}^{(m+1)}, \theta_{i+1}^{(m)}, \ldots, \theta_n^{(m)}, \boldsymbol{x})$, that, under mild conditions, converges to the joint posterior as its equilibrium distribution (see Gilks *et al.*, 1996). More generally, it is enough to just sample each full conditional using a Metropolis-Hastings step (Gilks, 1996) which is convenient if the full conditionals are of non-standard form. This technique is known as Metropolis-Hastings within Gibbs (MH-Gibbs) sampling, or alternatively, as single-component Metropolis-Hastings.

With the state-space implementation of the Schaefer surplus production model defined in equations (**??**) and (**??**) the MH-Gibbs sampler exhibited extremely slow mixing (see Gilks and Roberts, 1996). The performance of the Gibbs sampler is highly dependent upon the parameterization of the model and there are known to be computational difficulties (see Results section) with the parameterization in (**??**) and (**??**). After some experimenting with different parameterizations it was found that the problem of slow mixing was eliminated by using the states $P_y = B_y/K$ rather than B_y.

These new states are the ratio of biomass to carrying capacity, and upon replacing B_y by KP_y the state equations become

$$\begin{aligned} \log(P_{1967})|\sigma^2 &= u_{1967} \\ \log(P_y)|P_{y-1}, K, r, \sigma^2 &= \log(P_{y-1} + rP_{y-1}(1 - P_{y-1}) - C_{y-1}/K) + u_y \ , \end{aligned} \tag{6}$$

for $y = 1968, ..., 1989$, and the observation equations become

$$\log(I_y)|P_y, Q, \tau^2 = \log(Q) + \log(P_y) + v_y \ , \qquad y = 1967, ..., 1989, \tag{7}$$

where $Q = qK$.

The full-conditional posterior distributions for this model are given in the Appendix. Some of these distributions are not of standard form, and hence the need for the Metropolis-Hastings step within Gibbs sampling. For this purpose we used *adaptive rejection Metropolis sampling, ARMS* (Gilks *et al.*, 1995; Gilks and Neal, 1997). A subroutine written in the programming language C is available from Gilks *et al.* (1995).

Two main runs of 250 000 iterations of the MH-Gibbs sampler were performed, using as starting values the observation-error model and process-error model fits of Polacheck *et al.* (1993), respectively. Each run took approximately 75 min. on a 233MHz PC. Output was produced at the rate of almost 1MB per minute and hence, to keep the resulting computer files at a manageable size, a thinning of 25 was used. That is, only every 25th sample was saved, resulting in 10 000 samples from each run being written to disc.

For each of the two main runs, the CODA software of Best *et al.* (1995) was used to produce convergence diagnostics for the eight unobservables $K, r, \sigma^2, Q, \tau^2, B_{1989}$, MSP, and E_{opt}. Fifteen of these 16 sequences passed the stationarity test of Heidelberger and Welch (1983). The sole exception passed the second phase of the Heidelberger and Welch test, that is, with the first 10% (1000) of the samples removed. In what follows, the first 1000 thinned iterations of the MH-Gibbs sampler were considered as burn-in and are subsequently ignored, reducing the sequences to a length of 9000. (Several lesser runs were also used to note any problems with burn-in, none were observed.)

Two-sample Kolmogorov-Smirnov tests were applied to the eight pairs of sequences to determine whether the two main runs were sampling the same distribution. (This test is approximate due to lack of independence within each sequence, however, lag 1 autocorrelation did not exceed 0.08 in magnitude for any of the sixteen sequences.) All had p-value>0.1, and consequently, the two main runs were considered to be equivalent to one run of length 18000.

Posterior predictive checks (Gelman *et al.*, 1995) were conducted to assess whether discrepancies between the observed CPUE and posterior model were accordant with predicted discrepancies. This required sampling from the posterior predictive distribution for CPUE, which was easily accomplished by randomly generating a new sequence of CPUE's (equation (**??**)) for each of the 18000 values sampled from the joint posterior distribution of $\boldsymbol{\theta} = (K, r, \sigma^2, Q, \tau^2, P_{1967}, ..., P_{1989})$. Discrepancies, which may be functions of both observables and unobservables, were then calculated using the observed CPUE's (Table 1) and using the randomly generated CPUE's. Gelman *et al.* (1996) recommend using several different discrepancies, which might be a mixture of problem specific and omnibus measures of model discrepancy. The first three discrepancies below can be considered specific to the example herein, and the fourth is the omnibus chi-square measure.

Denoting sample i from the posterior by $\boldsymbol{\theta}^{(i)}, i = 1, \ldots, 18000$, the corresponding generated CPUE sequence by $\boldsymbol{I}^{(i)}$, and the actual observed CPUE sequence (Table 1) by $\boldsymbol{I}$, the following discrepancies were calculated:

1. Auto-correlation: $T_1(\boldsymbol{I}, \boldsymbol{\theta}^{(i)})$ and $T_1(\boldsymbol{I}^{(i)}, \boldsymbol{\theta}^{(i)})$ where

$$T_1(\boldsymbol{I}, \boldsymbol{\theta}) = \sum_{y=1968}^{y=1989} \frac{v_{y-1} v_y}{21\tau^2} \ ,$$

   where $v_y = \log(I_y) - \log(QP_y)$ is the observation error in equation (**??**).

2. Constant variance: $T_2(\boldsymbol{I}, \boldsymbol{\theta}^{(i)})$ and $T_2(\boldsymbol{I}^{(i)}, \boldsymbol{\theta}^{(i)})$ where $T_2(\boldsymbol{I}, \boldsymbol{\theta})$ is the Spearman rank-order correlation coefficient to test for an association between the predicted value, $\log(QP_y)$ and the observation error, v_y.

3. Normality: $T_3(\boldsymbol{I}, \boldsymbol{\theta}^{(i)})$ and $T_3(\boldsymbol{I}^{(i)}, \boldsymbol{\theta}^{(i)})$ where $T_3(\boldsymbol{I}, \boldsymbol{\theta})$ is the Kolmogorov-Smirnov statistic for the test of normality

$$H_0 : \log(I_y) \sim N(\log(QP_y), \tau^2), \qquad y = 1967, \ldots, 1989.$$

4. Chi-square goodness of fit: $T_4(\boldsymbol{I}, \boldsymbol{\theta}^{(i)})$ and $T_4(\boldsymbol{I}^{(i)}, \boldsymbol{\theta}^{(i)})$ where

$$T_4(\boldsymbol{I}, \boldsymbol{\theta}) = \sum_{y=1967}^{y=1989} \frac{v_y^2}{\tau^2} \ .$$

These quantities were calculated within Splus (Becker *et al.*, 1988).

Scatterplots were used to make graphical comparisons between the realized discrepancies $T_j(\boldsymbol{I}, \boldsymbol{\theta}^{(i)})$ and the predicted values $T_j(\boldsymbol{I}^{(i)}, \boldsymbol{\theta}^{(i)})$ $(i = 1, ..., 18000,\ j = 1, 2, 3, 4)$. If the model is reasonable then the points will be scattered symmetrically around the 45^o line. The posterior predictive p-value (Gelman *et al.*, 1996) for test j $(j = 1, 2, 3, 4)$ is calculated as the proportion of points for which $T_j(\boldsymbol{I}^{(i)}, \boldsymbol{\theta}^{(i)})$ exceeds $T_j(\boldsymbol{I}, \boldsymbol{\theta}^{(i)})$.

4 Results and Sensitivity to Priors

The posterior predictive p-values for $T_j, j = 1, 2, 3, 4$, were 0.69, 0.27, 0.50, and 0.42, respectively. These correspond to the proportion of points lying above the 45^o line in the plots of predicted versus realized discrepancies (Fig. 1). These p-values indicate that the discrepancies between observed CPUE and posterior model are very typical of those predicted.

A comparison between the observed CPUE's (Table 1) and the posterior predictive distribution of the CPUE's was made by overlaying the 95% posterior predictive intervals for log(CPUE)'s on to a plot of the observed log(CPUE)'s (Fig. 2). The observed CPUE's lie entirely within the predictive intervals. It should be noted that this particular model check, and also the chi-square discrepancy, T_4, will be insensitive to basic model inadequacies because the observation variance is a function of the residuals $\log(I) - \log(QP)$. (Hence, the importance of the three model-specific discrepancies.) However, Figure 2 and T_4 would be sensitive to a prior on τ^2 that was incompatible with the likelihood.

The posterior distribution of MSP (1000's t) has mean of 19.4, and 2.5, 50, and 97.5 percentiles of 13.9, 19.6, and 24.1, respectively (Fig. 3). For optimal effort these values are, in units of 10^6 hooks, 61.2, 44.8, 60.8, and 79.0, respectively. The catch and fishing effort in the five most recent years of data (Table 1) are well in excess of the 97.5 percentile on MSP and optimal effort.

The posterior modes of K, r, q, MSP, and E_{opt} are similar to the maximum likelihood estimates obtained by Polacheck *et al.* (1993) under the observation error model.

This may hold in general when priors are diffuse and observation error variance is substantially larger than process error variance. Here, the posterior means of observation error variance, τ^2, and process error variance, σ^2, were 0.012 and 0.0037, respectively. It should be borne in mind that the similarity of the point estimates does not imply that use of the observation error model and the present model would lead to similar management decisions (Hilborn, 1997).

Hilborn and Walters (1992) and Prager (1994) comment that the relative biomass index data used in surplus production models typically contain only limited information for inference about carrying capacity, K, because of high correlation with the other parameters. In effect, K scales the biomass sequence and hence is highly confounded with q (Fig. 4). It is also the case that K and r can be highly confounded because large stocks having low productivity (high K, small r) can give similar expected relative biomass trajectories as obtained from small stocks having high productivity (low K, high r). Furthermore, the sequence of biomasses B_y, $y = 1967, ...1989$ will be highly autocorrelated (Fig. 4). Gilks and Roberts (1996) show that slow mixing of the Gibbs sampler may result when high correlation is present in the joint posterior distribution. These considerations prompted the reparameterization obtained by dividing the biomass states by K, leading to equations (**??**) and (**??**) and reducing correlation in the posterior (Fig. 4).

Sensitivity to prior specification

For parameters r, q and τ^2, the prior specifications have formal justification in accordance with the published fisheries literature, and these justifications will be valid for many other fisheries. The priors on carrying capacity, K, and process error, σ^2, were vaguely informative and could have been derived in a variety of alternative manners. Hence, it is particularly important to assess the sensitivity of the results to these two priors.

The analysis was repeated with a uniform prior on $\log(K)$. The posterior on K now puts more mass on larger values of K, and somewhat less mass on lower values of K (Fig. 5). Larger values of K suggest that the CPUE data arise from the fishing down of an unproductive (low r) stock. Hence, use of this alternative prior on K results in more posterior mass on lower values of MSP and E_{opt}. In particular, the mean of the posterior for MSP drops from 19.4 to 18.7. This is a small change in relation to the posterior uncertainty about MSP.

The analysis was also repeated with a non-informative prior on σ^2 (uniform on $\log(\sigma^2)$). Only the posteriors for σ^2 and τ^2 show much change, with the posterior mean for σ^2 increasing to 0.0071 and that for τ^2 decreasing to 0.011. The higher posterior probability on larger values of σ^2 corresponds to more variability in the population dynamics and hence greater risk under a given harvesting scheme.

5 Discussion

A formal Bayesian stock assessment requires careful expert consideration of prior information. For some parameters, such as q, a standard non-informative prior (uniform on $\log(q)$) can be obtained. For other parameters there is some guidance in the existing fisheries literature, and the use of hierarchical modeling for development of formal priors for certain key population parameters is well underway. It could also be useful to inspect the prior induced on functions of the model parameters. For example, the independent lognormal priors on K and r resulted in a diffuse lognormal prior on MSP (Fig. 3). The prior on MSP is a reasonable choice in the sense that it shows little prior preference over the range of MSP values supported by the likelihood (e.g., Spiegelhalter *et al.*, 1994; Gelman *et al.*, 1995).

Applications of the Kalman filter or penalized likelihood to state-space stock assessment models have required the modeler to specify the ratio between process and observation error variances (e.g., Ludwig *et al.*, 1988; Kimura *et al.*, 1996; Richards and Schnute, 1988). Not surprisingly, the sensitivity analysis performed above shows that for the Bayesian implementation used here, inference about these variances is also quite dependent on the priors placed on them. In contrast, the posteriors of other quantities were little affected by this. Rather than simply assuming that the ratio of these variances is known, the Bayesian approach permits a coherent expression of prior knowledge of these variances. In practice this could come from assessment of observation error through calculation of the variability in fishing vessel log-books (say), and from hierarchical analysis of process error variability.

The combination of the MH-Gibbs sampler and ARMS proved to be a quick and reliable method for sampling from the posterior distribution. A C-program running on a 233 MHz PC generated 3300 samples per minute and the generated sequence showed quick burn-in and good mixing. The Metropolis algorithm could be employed to sample directly from the joint posterior density, subject to specification of a good jumping rule that takes into account the correlation structure of the parameters (typically assessed by using the Hessian matrix of the posterior evaluated at the posterior mode.) A further possibility is sampling-importance resampling (SIR). The prior distribution (Appendix) could be used as the initial choice of importance function, and be adapted if found to be inefficient (West, 1993; Kinas, 1996). Sampling from the prior can proceed by sampling sequentially from $\pi(K, r, q, \sigma^2, \tau^2)$, $f(P_{1967}|\sigma^2)$, and $f(P_y|P_{y-1}, K, r, \sigma^2), y = 1968, \ldots, 1989$.

The International Commission for the Conservation of Atlantic Tunas (ICCAT) has subsequently made a slight revision of the data in Table 1, but more importantly, it is now known that the stock had been fished quite heavily prior to the beginning of this data series in 1967 (Punt *et al.*, 1995), and it is suspected that the earlier CPUE data may be unreliable. Hence, this application cannot be considered a formal stock assessment of South Atlantic albacore.

The quantity that we have called "maximum surplus production" (MSP) is more commonly known as "maximum sustainable yield" (MSY) in the fisheries literature, despite the fact that sustained fishing at this level will likely drive the stock to extinction due to the random variability in biomass dynamics (Hilborn and Walters, 1992). The state-space approach to fisheries dynamics will help to change perceptions concerning "sustainable fishing", and implemented within the Bayesian framework the consequences of random stock dynamics and effects of alternative management decisions can be assessed as risks to the well-being of the stock and to those fishing it.

Acknowledgements

This manuscript benefited greatly from the helpful comments of three referees and an associate editor. We also thank Professor Ram Myers (Univ. of Dalhousie) for use of his worldwide stock and recruitment database.

References

Becker, R. A., J. M. Chambers, and A. R. Wilks. (1988) *The new S language.* California: Wadsworth and Brooks/Cole. 702 p.

Best, N. G., Cowles, M. K. and Vines, S. K. (1995) *CODA Manual Version 0.30.* Cambridge, UK: MRC Biostatistics Unit.

Breen, P. A. (1991) *Assessment of the red rock lobster (Jasus edwardsii) North and South Island stock, November 1991.* New Zealand Fisheries Assessment Research Document 91/16.

Carlin, B. P., Polson, N. G. and Stoffer, D. S. (1992) A Monte Carlo approach to nonnormal and nonlinear state-space modeling. *J. Amer. Stat. Assoc.*, **87**, 493-500.

Freeman, S. N. and Kirkwood, G. P. (1995) On a structural time series method for estimating stock biomass and recruitment from catch and effort data. *Fish. Res.*, **22**, 77-98.

Gelfand, A. E. and Smith, A. F. M. (1990) Sampling-based approaches to calculating marginal densities. *J. Amer. Stat. Assoc.*, **85**, 398-409.

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (1995) *Bayesian Data Analysis.* London: Chapman and Hall.

Gelman, A., Meng, X. and Stern, H. (1996) Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statistica Sinica.*, **6**, 733-807.

Gilks, W. R. (1996) Full conditional distributions. *In* Gilks *et al.* eds. *Markov Chain Monte Carlo in Practice.* p. 75-88. London: Chapman and Hall.

Gilks, W. R., Best, N. G. and Chan, K. K. C. (1995) Adaptive rejection Metropolis sampling within Gibbs sampling. *Appl. Statist.*, **44**, 455-472.

Gilks, W. R. and Neal, R. M. (1997) Corrigendum: adaptive rejection Metropolis sampling. *Appl. Statist.*, **46**, 541-542.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996) *Markov Chain Monte Carlo in Practice.* London: Chapman and Hall.

Gilks, W. R. and Roberts, G. O. (1996) Strategies for improving MCMC. *In* Gilks *et al.* eds. *Markov Chain Monte Carlo in Practice.* p. 89-114. London: Chapman and Hall.

Gudmundsson, G. (1994) Time series analysis of catch-at-age observations. *Appl. Statist.*, **43**, 117-126.

Gudmundsson, G. (1995) Time series analysis of catch-at-length observations. *ICES J. mar. Sci.*, **52**, 781-795.

Heidelberger, P. and Welch, P. (1983) Simulation run length control in the presence of an initial transient. *Operations Res.*, **31**, 1109-1144.

Hilborn, R. (1992) Current and future trends in fisheries stock assessment and management. *S. Afr. J. mar. Sci.*, **12**, 975-988.

Hilborn, R. (1997) Statistical hypothesis testing and decision theory in fisheries science. *Fisheries.*, **22**, 19-20.

Hilborn, R. and M. Liermann. (1998) Standing on the shoulders of giants: learning from experience in fisheries. *Rev. Fish. Bio. Fish.*, **8**, 273-283.

Hilborn, R., Pikitch, E. K. and McAllister, M. K. (1994) A Bayesian estimation and decision analysis for an age-structured model using biomass survey data. *Fish. Res.*, **19**, 17-30.

Hilborn, R. and Walters, C. J. (1992) *Quantitative fisheries stock assessment: choice, dynamics and uncertainty.* New York: Chapman and Hall.

Kalman, R.E. (1960) A new approach to linear filtering and prediction problems. *J. Basic Eng.*, **82**, 34-45.

Kass, R. E. and Wasserman, L. (1996) The selection of prior distributions by formal rules. *J. Amer. Stat. Assoc.*, **91**, 1343-1370.

Kimura, D. K., Balsiger, J. W. and Ito, D. H. (1996) Kalman filtering the delay difference equation: practical approaches and simulations. *Fish. Bull.*, **94**, 678-691.

Kinas, P. G. (1996) Bayesian fishery stock assessment and decision making using adaptive importance sampling. *Can. J. Fish. Aquat. Sci.*, **53**, 414-423.

Liermann, M. and R. Hilborn. (1997) Depensation in fish stocks: a hierarchic Bayesian meta-analysis. *Can. J. Fish. Aquat. Sci.*, **54**, 1976-1984.

Lin, X., and N. E. Breslow. (1996) Bias correction in generalized linear mixed models with multiple components of dispersion. *J. Amer. Stat. Assoc.*, **91**, 1007-1016.

Ludwig, D., C. J. Walters, and J. Cooke. (1988) Comparison of two models and two estimation methods for catch and effort data. *Nat. Resource Mod.*, **2**, 457-498.

McAllister, M. K., Pikitch, E. K., Punt, A. E. and Hilborn, R. (1994) A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. *Can. J. Fish. Aquat. Sci.*, **51**, 2673-2687.

Myers, R. A., Mertz, G. and Fowlow, P. S. (1997) Maximum population growth rates and recovery times for Atlantic cod, *Gadus morhua*. *Fish. Bull.*, **95**, 762-772.

Myers, R. A., Bowen, K. G. and N. J. Barrowman. (In press) The maximum reproductive rate of fish at low population sizes. *Can. J. Fish. Aquat. Sci.*, **xx**, xxx-xxx.

Pella, J. J. (1993) Utility of structural time series models and the Kalman filter for predicting consequences of fishery actions. *In* Kruse *et al.* eds. *Proceedings of the international symposium on management strategies for exploited fish populations.* p. 571-593. Alaska Sea Grant College Program Report 93-02, University of Alaska, Fairbanks.

Polacheck, T., Hilborn, R. and Punt, A. E. (1993) Fitting surplus production models: comparing methods and measuring uncertainty. *Can. J. Fish. Aquat. Sci.*, **50**, 2597-2607.

Prager, M. H. (1994) A suite of extensions to a nonequilibrium surplus-production model. *Fish. Bull.*, **92**, 374-389.

Punt, A. E., Butterworth, D. S. and Penney, A. J. (1995) Stock assessment and risk analysis for the South Atlantic population of albacore *Thunnus alalunga* using an age-structured production model. *S. Afr. J. mar. Sci.*, **16**, 287-310.

Punt, A. E. and Hilborn, R. (1997) Fisheries stock assessment and decision analysis: the Bayesian approach. *Rev. Fish. Bio. Fish.*, **7**, 35-63.

Reed, W. J. and Simons, C. M. (1996) Analyzing catch-effort data by means of the Kalman filter. *Can. J. Fish. Aquat. Sci.*, **53**, 2157-2166.

Richards, L. J., and Schnute, J. T. (1998) Model complexity and catch-age analysis. *Can. J. Fish. Aquat. Sci.* **55**, 949-957.

Ricker, W. E. (1975) Computation and interpretation of biological statistics of fish populations. *Bull. Fish. Res. Board. Can.*, **191**, 382 p.

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**, 25-56.

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

Spiegelhalter, D. J., Freedman, L. S., and Parmar, M. K. B. (1994) Bayesian approaches to randomized trials. With discussions. *J. Roy. Statist. Soc. Ser. A.*,

157, 357-416.

Sullivan, P. J. (1992) A Kalman filter approach to catch-at-length analysis. *Biometrics*, **48**, 237-257.

Walters, C. and Ludwig, D. (1994) Calculation of Bayes posterior probability distribution for key population parameters. *Can. J. Fish. Aquat. Sci.*, **51**, 713-722.

West, M. (1993) Approximating posterior distributions by mixtures. *J. R. Stat. Soc. B*, **55**, 409-422.

Appendix

The Metropolis-Hastings within Gibbs sampler was used here to sample from the joint posterior distribution of the 28 unknowns $(K, r, \sigma^2, Q, \tau^2, P_{1967}, ..., P_{1989})$, where $P_y = B_y/K$ and $Q = qK$ are parameters used in the reparameterized equations (**??**) and (**??**). This requires each of the 28 univariate full-conditional posterior distributions to be sampled in turn, and it is enough that this sample be obtained by a single Metropolis-Hastings step. These full-conditional distributions are given below.

Prior and posterior densities

Prior

The joint prior density $\pi(K, r, \sigma^2, Q, \tau^2, P_{1967}, ..., P_{1989})$ is obtained from the prior $\pi(K, r, \sigma^2, Q, \tau^2)$ and the distribution of $(P_{1967}, ..., P_{1989}|K, r, \sigma^2)$ determined from the state equations (**??**). That is,

$$\begin{aligned}
&\pi(K, r, \sigma^2, Q, \tau^2, P_{1967}, ..., P_{1989}) \\
&\quad = \pi(K, r, \sigma^2, Q, \tau^2) f(P_{1967}, ..., P_{1989}|K, r, \sigma^2) \\
&\quad = \pi(K, r, \sigma^2, Q, \tau^2) f(P_{1967}|\sigma^2) \prod_{y=1968}^{y=1989} f(P_y|P_{y-1}, K, r, \sigma^2) \ .
\end{aligned}$$

Note that the $f(P_y|P_{y-1}, K, r, \sigma^2)$ terms in this prior are implicitly conditioning on the catches, C_y. In surplus production models the data are the relative biomass indices and the catch data are assumed to provide no additional information. Indeed, in many fisheries the permitted catch will have been set in advance by the managers of that fishery.

Posterior

The posterior distribution of $(K, r, \sigma^2, Q, \tau^2, P_{1967}, ..., P_{1989})$ is proportional to the joint density of the data and unknowns. That is,

$$
\begin{aligned}
&f(K, r, \sigma^2, Q, \tau^2, P_{1967}, ..., P_{1989} | I_{1967}, ..., I_{1989}) \\
&\quad \propto \pi(K, r, \sigma^2, Q, \tau^2, P_{1967}, ..., P_{1989}) \prod_{y=1967}^{y=1989} f(I_y | P_y, Q, \tau^2) \\
&\quad = \pi(K, r, \sigma^2, Q, \tau^2) f(P_{1967} | \sigma^2) \prod_{y=1968}^{y=1989} f(P_y | P_{y-1}, K, r, \sigma^2) \prod_{y=1967}^{y=1989} f(I_y | P_y, Q, \tau^2) \ .
\end{aligned}
$$

Full-conditional densities

The full-conditional density for an unobservable, θ (say), is determined by the terms in the joint posterior that are functions of θ. The other terms in the posterior simply contribute to the normalizing constant. Each of the full conditionals given below has support $\mathbb{R}^+$.

Full conditional of P_y.

For $1968 \leq y \leq 1988$,

$$
\begin{aligned}
&f(P_y | K, r, \sigma^2, Q, \tau^2, P_{1967}, ..., P_{y-1}, P_{y+1}, ..., P_{1989}, I_{1967}, ..., I_{1989}) \\
&\propto \quad f(P_y | P_{y-1}, K, r, \sigma^2) f(I_y | P_y, Q, \tau^2) f(P_{y+1} | P_y, K, r, \sigma^2) \qquad (8) \\
&\propto \quad \exp\left\{ -\log(P_y) - \frac{(\log(P_y) - \log(P_{y-1} + rP_{y-1}(1 - P_{y-1}) - C_{y-1}/K))^2}{2\sigma^2} \right. \\
&\qquad \left. - \frac{(\log(I_y) - \log(QP_y))^2}{2\tau^2} - \frac{(\log(P_{y+1}) - \log(P_y + rP_y(1 - P_y) - C_y/K))^2}{2\sigma^2} \right\} .
\end{aligned}
$$

When $y = 1967$ the first factor in (**??**) is $f(P_{1967} | \sigma^2)$. When $y = 1989$ the third factor in (**??**) is omitted.

Full conditionals of K and r.

$$
\begin{aligned}
& f(K|r,\sigma^2,Q,\tau^2,P_{1967},...,P_{1989},I_{1967},...,I_{1989}) \\
& \quad \propto \quad \pi(K,r,\sigma^2,Q,\tau^2)\prod_{y=1968}^{1989} f(P_y|P_{y-1},K,r,\sigma^2) \\
& \quad \propto \quad \pi(K,r,\sigma^2,Q,\tau^2)\exp\left\{\frac{-1}{2\sigma^2}\sum_{y=1968}^{1989}(\log(P_y)-\log(P_{y-1}+rP_{y-1}(1-P_{y-1})-C_{y-1}/K))^2\right\} .
\end{aligned}
$$

The full conditional for r is given by the same formula (up to a constant of proportionality) by fixing all other parameters.

Full conditional of σ^2.

$$
\begin{aligned}
& f(\sigma^2|K,r,\sigma^2,Q,\tau^2,P_{1967},...,P_{1989},I_{1967},...,I_{1989}) \\
& \quad \propto \quad \pi(K,r,\sigma^2,Q,\tau^2)f(P_{1967}|\sigma^2)\prod_{y=1968}^{1989} f(P_y|P_{y-1},K,r,\sigma^2) \\
& \quad \propto \quad \frac{\pi(K,r,\sigma^2,Q,\tau^2)}{(\sigma^2)^{\frac{n}{2}}} \\
& \qquad \times\exp\left\{\frac{-1}{2\sigma^2}\left[\log(P_{1967})+\log\sum_{y=1968}^{1989}(\log(P_y)-\log(P_{y-1}+rP_{y-1}(1-P_{y-1})-C_{y-1}/K))^2\right]\right\} ,
\end{aligned}
$$

where $n=23$ is the number of states, $P_y, y=1967,...,1989$.

Full conditional of Q *and* τ^2.

$$
\begin{aligned}
& f(Q|K,r,\sigma^2,\tau^2,P_{1967},...,P_{1989},I_{1967},...,I_{1989}) \\
& \quad \propto \quad \pi(K,r,\sigma^2,Q,\tau^2)\prod_{y=1967}^{1989} f(I_y|P_y,Q,\tau^2) \\
& \quad \propto \quad \frac{\pi(K,r,\sigma^2,Q,\tau^2)}{(\tau^2)^{\frac{n}{2}}}\exp\left\{\frac{-1}{2\tau^2}\sum_{y=1967}^{1989}(\log(I_y)-\log(Q)-\log(P_y))^2\right\} ,
\end{aligned}
$$

and this formula also gives the full conditional for τ^2 (up to a constant of proportionality).

Table 1. Catch and effort data for South Atlantic albacore tuna.

Year	Catch (1000's t)	Effort (10^8 hooks)	CPUE (kg/100 hooks)
1967	15.9	0.257	61.89
1968	25.7	0.325	78.98
1969	28.5	0.513	55.59
1970	23.7	0.531	44.61
1971	25.0	0.439	56.89
1972	33.3	0.870	38.27
1973	28.2	0.833	33.84
1974	19.7	0.545	36.13
1975	17.5	0.417	41.95
1976	19.3	0.527	36.63
1977	21.6	0.595	36.33
1978	23.1	0.595	38.82
1979	22.5	0.656	34.32
1980	22.5	0.598	37.64
1981	23.6	0.694	34.01
1982	29.1	0.905	32.16
1983	14.4	0.536	26.88
1984	13.2	0.361	36.61
1985	28.4	0.944	30.07
1986	34.6	1.125	30.75
1987	37.5	1.605	23.36
1988	25.9	1.158	22.36
1989	25.3	1.155	21.91

Figure 1. Scatter plots of realized versus predicted discrepancies. The posterior predictive p-value is given by the proportion of points lying above the 45^o line.

Figure 2. Observed log(CPUE) (solid line) and the 2.5, 50, and 97.5 percentiles from the posterior predictive distribution of log(CPUE) (dashed lines).

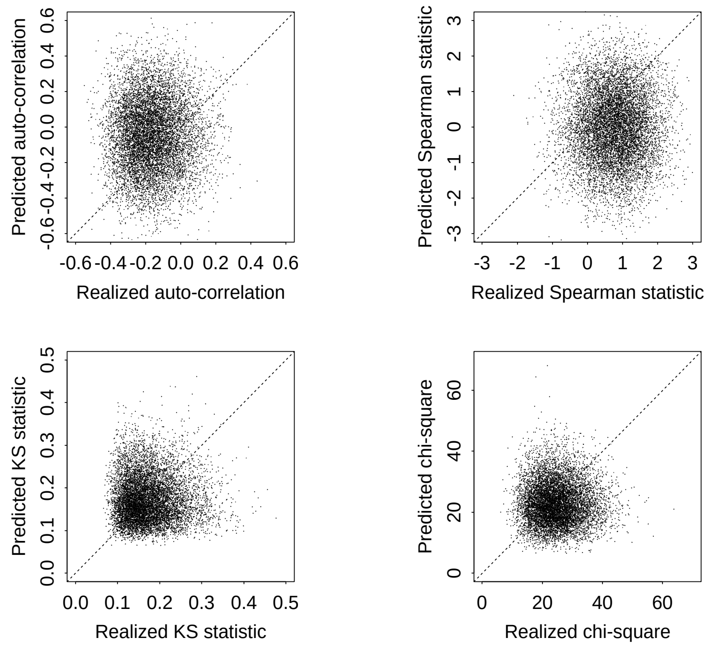
Figure 3. Posterior densities (solid lines) obtained using the priors specified in section **??**. Proper prior densities are given by the dashed lines.

Figure 4. Scatter plots of samples from the posterior distribution of (K, q) and (B_{1967}, B_{1968}) from the original parameterization (top row), and of (K, Q) and (P_{1967}, P_{1968}) from the reparameterized model (bottom row).

Figure 5. Posterior densities (long-dash lines) obtained from using a uniform prior on $\log(K)$. Proper prior densities are given by the short-dash lines and the posterior densities of Fig. 3 are shown by solid lines.

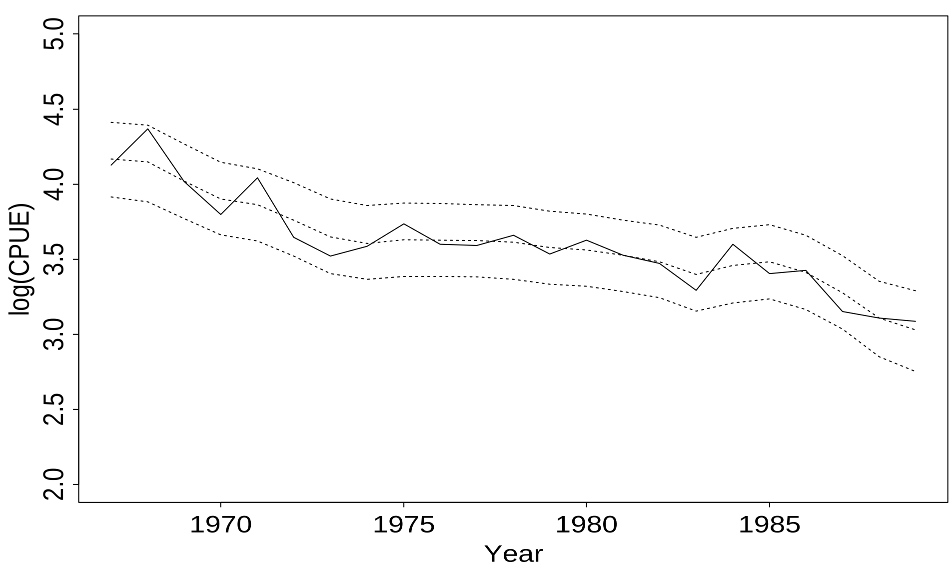
Figure 6. Posterior densities (long-dash lines) obtained from using a uniform prior on $\log(\sigma^2)$. Proper prior densities are given by the short-dash lines and the posterior densities of Fig. 3 are shown by solid lines.

Fig. 1



Predicted auto-correlation
Realized auto-correlation
-0.6 -0.4 -0.2 0.0 0.2 0.4 0.6
Predicted Spearman statistic
Realized Spearman statistic
-3 -2 -1 0 1 2 3
Predicted KS statistic
Realized KS statistic
0.0 0.1 0.2 0.3 0.4 0.5
Predicted chi-square
Realized chi-square
0 20 40 60

Fig. 2:



log(CPUE)
2.0 2.5 3.0 3.5 4.0 4.5 5.0
1970 1975 1980 1985
Year

Fig. 3

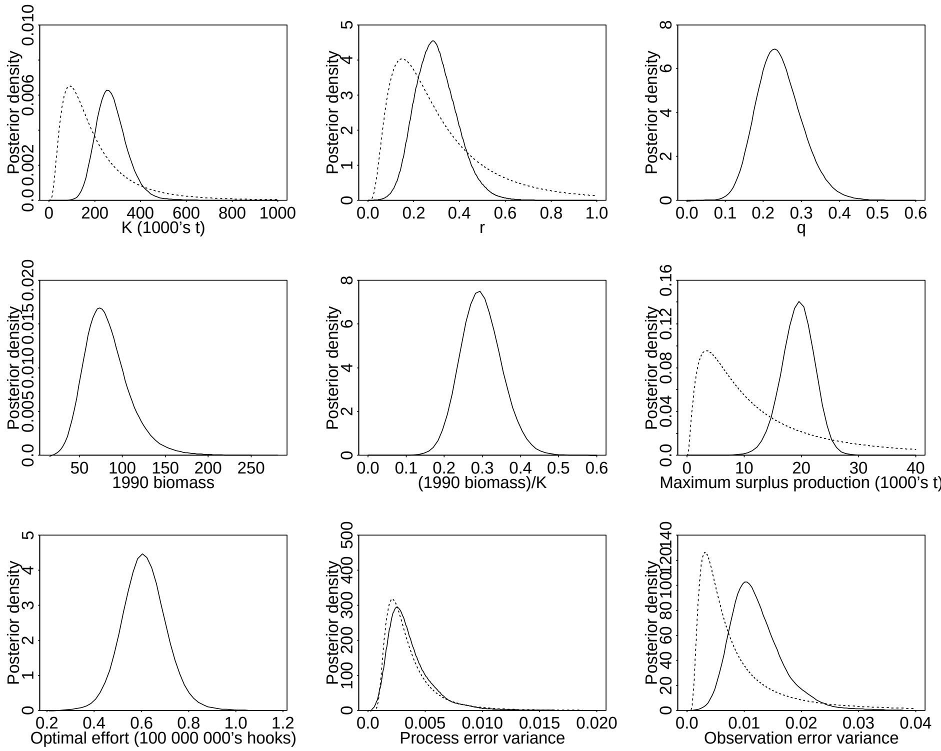

Posterior density
K (1000's t)
r
q
1990 biomass
(1990 biomass)/K
Maximum surplus production (1000's t)
Optimal effort (100 000 000's hooks)
Process error variance
Observation error variance

Fig. 4:

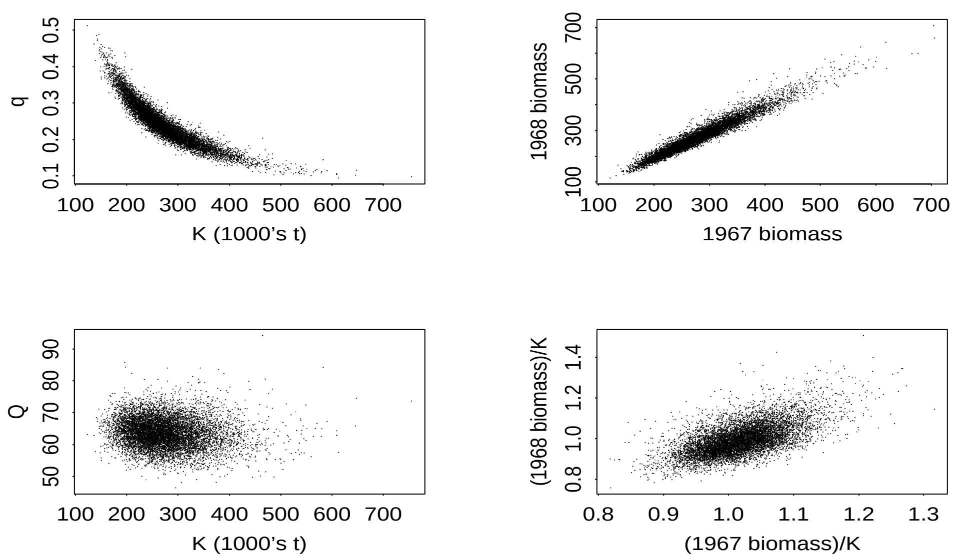

q
K (1000's t)
1968 biomass
1967 biomass
Q
K (1000's t)
(1968 biomass)/K
(1967 biomass)/K

Fig. 5

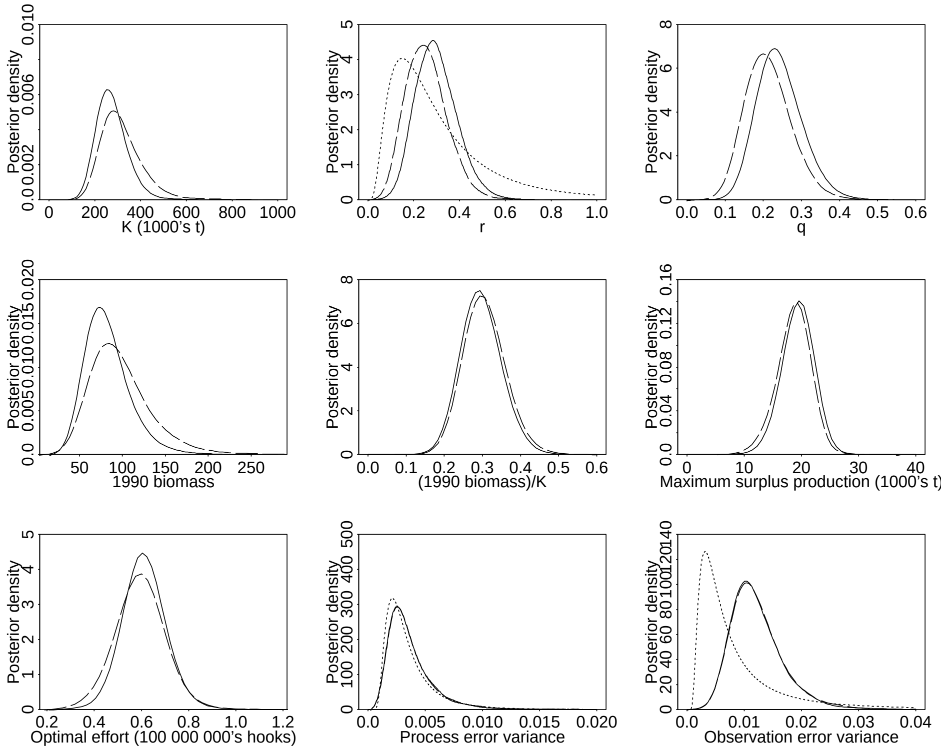
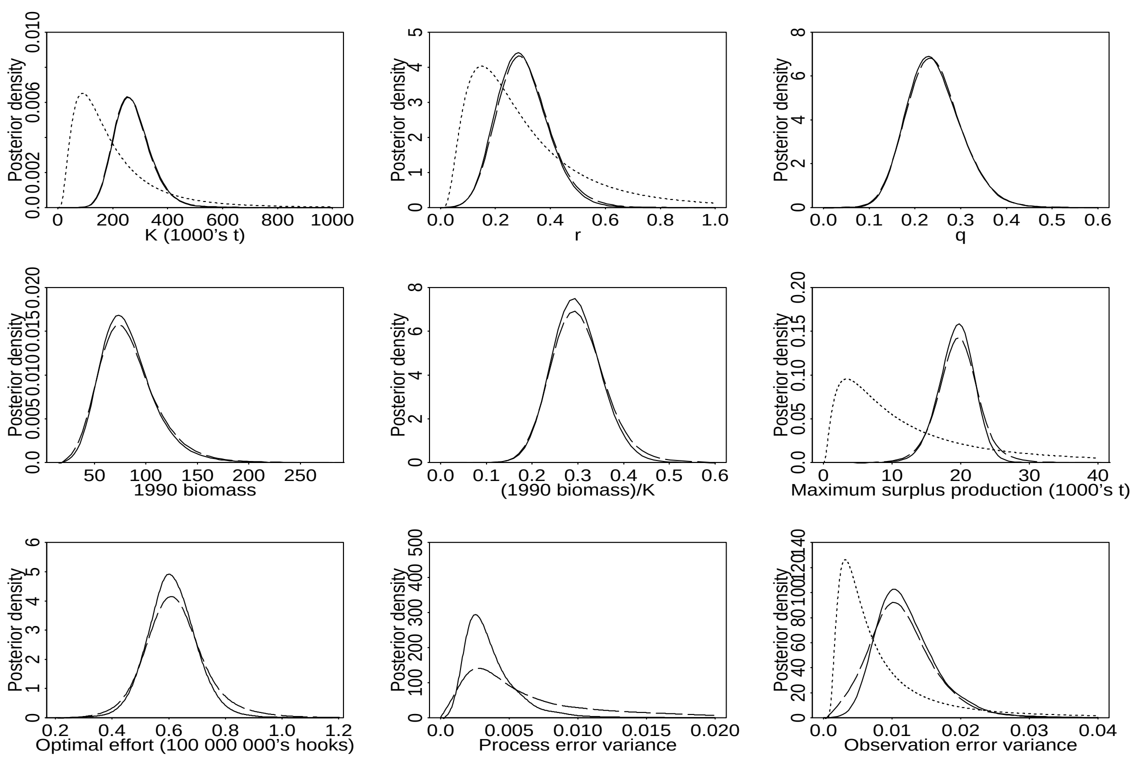

Posterior density
K (1000's t)
r
q
1990 biomass
(1990 biomass)/K
Maximum surplus production (1000's t)
Optimal effort (100 000 000's hooks)
Process error variance
Observation error variance

Fig. 6:


Posterior density
K (1000's t)
r
q
1990 biomass
(1990 biomass)/K
Maximum surplus production (1000's t)
Optimal effort (100 000 000's hooks)
Process error variance
Observation error variance