Bayesian stock assessment using a state–space implementation of the delay difference model

Renate Meyer and Russell B. Millar

Abstract: This paper presents a Bayesian approach to fisheries stock assessment using the delay difference model to describe nonlinear population dynamics. Given a time series of annual catch and effort data, models in the Deriso–Schnute family predict exploitable biomass in the following year from biomass in the current and previous year and from past spawning stock. A state–space model is used, as it allows incorporation of random errors in both the biomass dynamics equations and the observations. Because the biomass dynamics are nonlinear, the common Kalman filter is generally not applicable for parameter estimation. However, it is demonstrated that the Bayesian approach can handle any form of nonlinear relationship in the state and observation equations as well as realistic distributional assumptions. Difficulties with posterior calculations are overcome by the Gibbs sampler in conjunction with the adaptive rejection Metropolis sampling algorithm.

Résumé : Cet article présente une analyse bayésienne appliquée à l'évaluation des stocks de poissons utilisant le modèle à différences retardées pour décrire une dynamique des populations non linéaire. Avec une série temporelle de données de prises et d'effort, les modèles de la famille Deriso–Schnute prévoient la biomasse exploitable de l'année suivante à partir de la biomasse de l'année en cours et de l'année précédente ainsi qu'à partir du stock passé de reproducteurs. On utilise un modèle d'espace–d'états qui permet d'incorporer les erreurs aléatoires dans les équations de la dynamique de la biomasse et dans les observations. Comme la dynamique de la biomasse est non linéaire, on ne peut généralement appliquer le filtre de Kalman commun pour l'estimation des paramètres. Cependant, on démontre que l'analyse bayésienne peut traiter toute forme de relation non linéaire dans les équations d'état et d'observation de même que les hypothèses de distribution réalistes. On peut résoudre les difficultés liées aux calculs postérieurs en recourant à l'échantillonneur de Gibbs avec l'algorithme d'échantillonnage de Metropolis à rejet adaptatif.

[Traduit par la Rédaction]

Introduction

Delay difference models bridge the gap between complex fully age-structured models and simple surplus production models (e.g., for a review, see Hilborn and Walters 1992). Whereas the age-structured methods model the population cohorts in numbers-at-age and are usually applied to catch-at-age data, the simpler surplus production models can be fit when only catch and effort data are available, e.g., in situations where it is impractical, difficult, or expensive to age the species. However, the surplus production models sacrifice biological realism in the description of biomass dynamics for the sake of mathematical simplicity.

Population dynamics models in general relate exploitable biomass in year $t + 1$ to biomass, growth, recruitment, natural mortality, and catch in the previous year t. Surplus production models (Schaefer 1954; Pella and Tomlinson 1969; Fox 1970) in particular aggregate the terms for growth, recruitment, and natural mortality into one combined term called "surplus production". They have been popular and widely used in practical stock assessment because the models are parsimonious in the number of parameters, simple fitting procedures are readily available (Polachek et al. 1993), and they produce "management" parameters such as maximum sustainable yield and virgin biomass. On the other hand, the complex age-structured models (Fournier and Archibald 1982; Megrey 1989) enjoy greater credibility in the stock assessment community, and more detailed management decisions, such as size limitations and gear mesh size restrictions, can be based on their outcomes (Hilborn and Walters 1992; Punt et al. 1995). However, this is typically at the expense of obtaining comprehensive age information (but also see Francis 1992; McAllister et al. (1994) for applications where fully age-structured models are fit to relative abundance data) and of making assumptions on vulnerability and growth.

The delay difference model was developed by Deriso (1980), and its important contribution to fisheries population theory was realized immediately (Walters 1980). Deriso's (1980) model was generalized by Schnute (1985) and extended to length-structured data by Fournier and Doonan (1987) and Schnute (1987). Delay difference models can be placed conceptually midway between surplus production and catch-at-age models. They do not attempt to aggregate terms for recruitment, growth, and natural mortality into surplus production but retain individual parameters for these in the model. Thus, in contrast with the surplus production models, the parameters of the delay difference model preserve their biological interpretability and practical significance. Like

Received January 12, 1998. Accepted July 28, 1998.
J14384

R. Meyer[1] and R.B. Millar. Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand.

[1]Author to whom all correspondence should be addressed. e-mail: meyer@stat.auckland.ac.nz

© 1999 NRC Canada

catch-at-age methods, they can be termed "age structured", as they make initial detailed assumptions on the biomass dynamics in each age-class. These age-specific equations, however, are collapsed into just one equation for the entire population. Hence, no age data are required for fitting delay difference models, making them applicable to most fisheries. The delay difference model thus captures the advantages of both classes: like the fully age-structured model, it has a sound footing in biological theory, yielding parameters of immediate practical biological relevance while at the same time retaining the one-dimensional biomass dynamics and data requirements of the ad hoc surplus production models. Information on growth, age at recruitment, and natural mortality that it needs is often available from independent sources.

As has been pointed out in the literature (Pella 1993; Polachek et al. 1993), two stochastic components have to be taken into account when fitting population dynamics models: natural variability underlying the annual biomass dynamics transitions (process error) and uncertainty in the observed abundance indices due to measurement and sampling error (observation error). Both process and observation error can be incorporated in a very natural way by casting the problem into the framework of dynamic state–space modeling of modern time series methodology (e.g., Fahrmeir and Tutz 1994). State–space models relate time series observations (here the relative abundance indices $\{I_t\}$, e.g., catch per unit effort (CPUE) from commercial fisheries) to unobserved states (here the biomasses $\{B_t\}$) by a stochastic observation model for I_t given B_t. The states are assumed to follow a stochastic transition model (here the delay difference model). Given the observations $\{I_t\}$ the estimation of the states B_t, $t = 1,..., N$, is the primary goal of statistical inference, termed "filtering" for $t = N$, "smoothing" for $t < N$, and "prediction" for $t > N$. Thus, by treating the annual biomasses as unknown states, one explicitly allows for stochasticity in the populations dynamics through specification of their conditional distribution given previous states, further unknown model parameters, and explanatory variables (such as catch). At the same time, the observations are linked to the biomasses and measurement error is encompassed by specifying the conditional distribution of each observed relative abundance index given the state of the stock in that year.

State–space models have been successfully applied in engineering, economics, management science, and biology (West and Harrison 1997). A number of prominent time series models, e.g., autoregressive moving-average, structural time series, and dynamic regression models, can be described and dealt with in a flexible and unifying way. State–space models have only recently been introduced and applied in fisheries modeling (Sullivan 1992; Pella 1993; Schnute 1994; Kimura et al. 1996; Reed and Simons 1996) using the classical maximum likelihood (ML) approach to parameter estimation. For approximately normally distributed data and linear state transitions, the famous Kalman filter (Kalman 1960) has found numerous applications in the ML analysis of time series data (e.g., West and Harrison 1997). The Kalman filter depends crucially on the linearity of state–space equations, which is generally not warranted in delay difference or surplus production models. Even the extended Kalman filter for nonlinear models (e.g., Pella 1993; Gudmundsson 1994) relies heavily on linear approximations.

Although the ML approach has been well explored, a Bayesian approach to state–space modeling in fisheries is still outstanding. There have been recent Bayesian stock assessment analyses using population dynamics models (McAllister et al. 1994; Raftery et al. 1995; Kinas 1996; McAllister and Ianelli 1997), but none of these have employed a state–space model and none have successfully incorporated both process and observation error, as detailed in the discussion at the end of this paper.

We will show that a Bayesian approach (Carlin et al. 1992; Gamerman 1997) can handle any form of nonlinearity in the state and observation equations as well as realistic distributional assumptions. The biologically meaningful interpretations of the parameters in a delay difference model permit one to make efficient use of auxiliary information on growth or natural mortality in the elicitation of prior distributions. This makes a Bayesian approach particularly attractive for fitting delay difference models. More compelling reasons why a Bayesian technique should be the method of choice for stock assessment are given in Punt and Hilborn (1997). Trying to avoid an engagement in philosophical debates about the foundations of statistics (the interested reader is referred to Berger and Wolpert 1988), we would like to emphasize the pragmatic advantages of Bayesian inference (Gelman et al. 1995). Its flexibility and generality allow one to cope with very complex problems through the quantification of uncertainty and application of the Bayes' rule (eq. 5). Fitting complicated, highly nonlinear multiparameter models, such as those in fisheries population dynamics, is possible within the Bayesian paradigm. Consequently, we propose a fully Bayesian approach to statistical inference in the delay difference model by an integration into the scheme of nonlinear state–space models. Following Carlin et al. (1992), we advocate the Gibbs sampler for posterior computation in nonlinear state–space models for its ease of implementation, and we demonstrate its effectiveness. The Gibbs sampler, as described for instance in Casella and George (1992), Tierney (1994), and O'Hagan (1994), is a special Markov chain Monte Carlo (MCMC) method that uses each of the one-dimensional full conditional posterior distributions in turn to generate a sample from the joint posterior distribution of all the unknowns.

The paper is organized as follows. In the first section, we set up the assumptions and specifications for the delay difference equations that describe the population dynamics. The second section introduces the general Bayesian setup for parameter estimation and the computational techniques to calculate the posterior distributions. Prior and likelihood specifications for a specific delay difference model are given in the third section, and the fourth section describes the Gibbs sampling approach including the calculation of the full conditional densities. The fifth section illustrates the feasibility of the proposed Bayesian analysis using the data set on yellowfin tuna (*Thunnus albacares*) previously analyzed by Kimura et al. (1996) and contrasts the results obtained by the two distinct approaches. It is also compared with results obtained from a surplus production model, implemented as described in Millar and Meyer (1998). As the ultimate objective of Bayesian stock assessment is to pro-

© 1999 NRC Canada

vide fisheries management with risks of alternative management options, we point out how the results obtained from the Bayesian analysis can be used for decision analysis. This is the topic of the sixth section. The final section closes with a discussion on the general merits, flexibility, and potential of the novel approach proposed.

Delay difference population dynamics

We proceed from data information given in the form of a time series of catches $\{C_t\}_{t=1}^N$ and relative abundance indices $\{I_t\}_{t=1}^N$ in years $t = 1,\ldots N$. To derive the deterministic delay difference equations (Hilborn and Walters 1992), specific assumptions on the stock's weight–age relationship and natural mortality are needed and will be specified in the following. We make the usual assumption of so-called "knife-edged recruitment" of fish at age $a = k$ years but note that Deriso (1980) provided an approach for also allowing incomplete recruitment to be modeled. The mean weight of an a-year-old fish is assumed to be a linear function of the weight-at-age $a - 1$:

$$(1) \qquad \omega_a = \alpha + \rho\omega_{a-1} \text{ for } a \geq k.$$

The model implies that the annual weight increment decreases by the factor ρ as $\omega_{a+1} - \omega_a = \rho(\omega_a - \omega_{a-1})$, where $0 < \rho < 1$ is Ford's growth coefficient. Some straightforward mathematical manipulations show the equivalence to Schnute's (1985) generalization of Deriso's (1980) growth model:

$$\omega_{k+j} = \omega_{k-1} + (\omega_k - \omega_{k-1})\frac{1-\rho^{1+j}}{1-\rho} \text{ for } j \geq 0$$

where ω_{k+j} is the weight of a $(k+j)$-year-old fish and ω_{k-1} is the prerecruitment weight (assumed to be 0 in Deriso's (1980) original formulation). In the following application, the growth parameters ρ and $\omega = \frac{\omega_{k-1}}{\omega_k}$ are assumed to be known, i.e., assessed from independent sources of information, to compare results with those in Kimura et al. (1996). In general, however, these may well be regarded as unknown model parameters.

Let $N_{a,t}$ denote the population number at age a in year t. Assuming that all ages of recruited fish have equal natural mortality and are equally vulnerable to the fishery, the annual change for each cohort is described by

$$(2) \qquad N_{a+t,t+1} = s_t N_{a,t}$$

where s_t denotes the total survival probability in year t. We assume that fishing takes place in a pulse at the start of each year so that under independence of fishing and natural mortality:

$$s_t = s_t^N s_t^F$$

where $s_t^F = \frac{B_t - C_t}{B_t}$ and the natural survival probability $s_t^N = e^{-M}$ is constant (and assumed to be known in the practical application discussed later on in order to make results comparable with those of Kimura et al. (1996), but M can in general be regarded as an additional parameter to be estimated in the Bayesian framework). Let R_t denote total biomass of recruits assumed to enter in a pulse in year t.

Under the above assumptions on growth, mortality, and the fact that the total biomass B_t of the stock that is fully vulnerable to fishing at the start of year t can be written as

$$(3) \qquad B_t = \sum_{a=k}^{\infty} \omega_a N_{a,t},$$

the cohort population dynamics equation can be collapsed into one delay difference equation, predicting the biomass in the next year as a linear function of the biomass in the current and the previous year and the biomass of new recruits added to the stock. For details on the actual derivation, the reader is referred to Schnute (1985) or Hilborn and Walters (1992, pp. 332–335). The stochastic version in terms of expected biomass (suppressing the conditioning on ρ, ω, M, q, R_t, R_{t+1}, and fixed catches C_t and C_{t-1}) is expressed as

$$(4) \qquad \mathcal{E}[B_{t+1} | B_t, B_{t-1}] = (1+\rho)e^{-M}(B_t - C_t) - \rho e^{-2M}\frac{(B_t - C_t)}{B_t}(B_{t-1} - C_{t-1}) - \rho\omega e^{-M}\frac{(B_t - C_t)}{B_t}R_t + R_{t+1} \text{ for } t = 2,\ldots,N$$

$$\mathcal{E}[I_t | B_t] = qB_t \text{ for } t = 1,\ldots,N$$

assuming that the observed relative abundance index is proportional to the total biomass. This is a simplifying assumption that is often made but often not realistic, as catchability might increase with decreasing stock size, for instance. If this is suspected, time-dependent catchability parameters should be specified in the model. We presume that prior to fishing, the stock is in dynamic equilibrium about its carrying capacity, K. This complements the state equations above with those for the first 2 years:

$$\mathcal{E}[B_1] = K$$

$$\mathcal{E}[B_2 | B_1] = e^{-M}(1 + \rho - \rho e^{-M})(B_1 - C_1) - \rho\omega e^{-M}\frac{(B_1 - C_1)}{B_1}R_1 + R_2.$$

The first mathematical techniques employed for parameter estimation in the delay difference model were ML under the assumption of a normal distribution for process error only (see Hilborn and Walters 1992) and nonlinear least squares, allowing only for observation error by treating the biomass dynamics delay difference equations as deterministic (Punt 1988; Ludwig and Walters 1989). However, the extreme influence of the presumed ratio of the measurement and process error variances on parameter estimates is well known (Ludwig and Walters 1981; Schnute 1989, 1991; Pella 1993; Polachek et al. 1993). Using a state–space implementation, we will be able to account for both observation and process error.

Schnute (1994) commented on two problems associated with ML estimation in state–space fisheries models: the necessity to solve high-dimensional integrals when calculating the likelihood function and the inability to estimate the unknown states, although they have as much biological impor-

© 1999 NRC Canada

Table 1. Catch (millions of pounds) and CPUE (pounds per boat-day) data from Pella and Tomlinson (1969).

Year	Catch	CPUE
1934	60.9	10 361
1935	72.3	11 484
1936	78.4	11 571
1937	91.5	11 116
1938	78.3	11 463
1939	110.4	10 528
1940	114.6	10 609
1941	76.8	8 018
1942	42.0	7 040
1943	50.1	8 441
1944	64.9	10 019
1945	89.2	9 512
1946	129.7	9 292
1947	160.2	7 857
1948	207.0	8 353
1949	200.1	8 363
1950	224.8	7 057
1951	186.0	10 108
1952	195.3	5 606
1953	140.0	3 852
1954	140.0	5 339
1955	140.9	8 191
1956	177.0	6 507
1957	163.0	6 090
1958	148.5	4 768
1959	140.5	4 982
1960	244.3	6 817
1961	230.9	5 544
1962	174.1	4 120
1963	145.5	4 368
1964	203.9	4 844
1965	180.1	4 166
1966	182.3	4 513
1967	178.9	5 292

tance as the unknown model parameters. He came to the conclusion that both problems can be addressed by shifting to a Bayesian perspective. Even as early as 1980, Deriso mentioned in his discussion that it might be possible to improve on parameter estimates by taking a Bayesian approach in assigning prior probability distributions to certain parameters based on information about growth, recruitment, and catchability, which is often available from independent assessments. Yet, at that early stage, Bayesian approaches were hindered by computational problems of the multiparameter integrations needed to calculate posterior probability distributions. These impediments have been overcome by the immense progress made within the last decade in Bayesian computational technology via MCMC methods (see Gilks et al. 1996 for an introduction). MCMC approaches to dynamic models represent some of the currently critical research frontiers in Bayesian time series modeling. To encourage thier use in fisheries, the Bayesian approach to the analysis of nonlinear state–space models is explained in the next section.

General framework for Bayesian stock assessment

The Bayesian paradigm enables the stock assessment scientist to include substantive knowledge as well as subjective opinion into the analysis through the elicitation of informative priors for the model parameters. Thus, he or she can make full use of the whole collection of historical experience and incorporate basic biological knowledge, expert judgment, and information from inferences for related species and stocks.

The Bayesian approach to stock assessment in general consists of two conceptually and practically distinct steps: (*i*) constructing a full probability model that consists of a joint probability distribution for all observable (here the CPUEs) and unobservable quantities (here the biomasses and model parameters) and (*ii*) by conditioning on the observed data, calculating the posterior distribution, i.e., the conditional probability distribution of the unobservables of interest, given the observed data.

In the first step, the joint probability density $p(\mathbf{y}, \boldsymbol{\theta})$ of the observations $\mathbf{y} = (y_1,...,y_N)$ and the unobservables $\boldsymbol{\theta} = (\theta_1,..., \theta_n)$ can be written as the product of two densities, referred to as the prior density $p(\boldsymbol{\theta})$ and the sampling density or likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$:

$$p(\mathbf{y}, \boldsymbol{\theta}) = p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta}).$$

In light of the data, our opinion as to the state of nature is then updated to the posterior distribution. Conditioning on the known value of $\mathbf{y}$ and using the fundamental Bayes' rule yields the posterior

$$(5) \qquad p(\boldsymbol{\theta}|\,\mathbf{y}) = \frac{p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})}{p(\mathbf{y})} \propto p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})$$

because $p(\mathbf{y})$ is independent of $\boldsymbol{\theta}$ and regarded as a normalization constant. One could then marginalize the joint posterior over certain components of $\boldsymbol{\theta}$ to obtain characteristics of $\boldsymbol{\theta}$ of interest, such as the posterior mean, median, or mode of a specific component θ_i. If one is interested in a certain function $\phi = g(\boldsymbol{\theta})$, its posterior density can be derived by multiplying $p(g^{-1}(\phi)|\mathbf{y}))$ by the absolute value of the determinant of the Jacobian of $g^{-1}(\phi)$.

Thus, the second step, although conceptionally easy, is indeed a formidable problem in general because it requires high-dimensional integration to obtain the normalization constant $p(\mathbf{y})$ and to calculate one-dimensional characteristics, and possibly further differentiation to compute posterior distributions of transformations. Before the development of MCMC, there were essentially three different approaches to handle this multidimensional integration (for an overview, see Evans and Swartz 1995): (*i*) asymptotic approximations like the normal approximations based on Taylor series expansion of the logarithm of the posterior density around its mode, and the more precise Laplace approximations (Tierney and Kadane 1986), (*ii*) numerical integration via Gaussian quadrature techniques (Gamerman 1997, chap. 3.3), and (*iii*) Monte Carlo integration via the hit-and-miss method (Rubinstein 1981), the more efficient importance sampling (O'Hagan 1994), and sampling/importance resampling (SIR) (Rubin 1987, 1988; Smith and Gelfand

© 1999 NRC Canada

1992). Laplace expansion relies on large sample asymptotics and the approximations can be very bad in small-sample situations. Gaussian quadrature suffers from the curse of dimensionality in that the amount of computation rises exponentially with the number of parameters. Monte Carlo integration via importance sampling substitutes the deterministic integration by a statistical estimation problem, that of estimating the mean of a certain multivariate distribution. This can be done by drawing a random sample and estimating the expectation by the sample mean. Although applicable to high-dimensional problems, the conventional Monte Carlo methods can be very inefficient in certain situations. For a more detailed criticism of Monte Carlo integration from a Bayesian point of view, see O'Hagan (1987). The efficiency of the SIR algorithm depends heavily on the importance density, which should be chosen close to the joint posterior density and have heavier tails to ensure an adequate coverage of the relevant posterior regions. This will be hard to achieve in high dimensions. Asymptotic approximations and numerical integration not only require a high degree of mathematical sophistication of the data analyst but also customizing the estimation routine to each specific problem. A major breakthrough for the routine implementation of Bayesian inference was the realization that any high-dimensional integration can be performed by using MCMC methods of which the Gibbs sampler is an important special case.

Instead of generating a sequence of independent samples from the joint posterior, in MCMC, a Markov chain is constructed whose equilibrium distribution is just the joint posterior. Thus, after running the Markov chain for a certain "burn-in" period, one obtains (correlated) samples from the limiting distribution (provided that the Markov chain has reached convergence). Asymptotic theory ensures that averaging of a function of interest over realizations from a single run of the chain provides a consistent estimate of its expectation.

The Gibbs sampler is a specific MCMC method where in a cycle, we sample from each of the full conditional distributions

$$p(\theta_i|\theta_1,\ldots,\theta_{i-1},\theta_{i+1},\ldots,\theta_n)$$

suppressing the dependence of the conditional posteriors upon **y**.

Given an arbitray set of starting values $\theta_1^{(0)},\ldots,\theta_n^{(0)}$ the algorithm proceeds as follows:

$$\text{Simulate } \theta_1^{(1)} \sim p(\theta_1|\theta_2^{(0)},\ldots,\theta_n^{(0)})$$

$$\text{Simulate } \theta_2^{(1)} \sim p(\theta_2|\theta_1^{(1)},\theta_3^{(0)},\ldots,\theta_n^{(0)})$$

$$\vdots$$

$$\text{Simulate } \theta_n^{(1)} \sim p(\theta_n|\theta_1^{(1)}\ldots,\theta_{n-1}^{(1)})$$

and yields $(\theta_1^{(m)},\ldots,\theta_n^{(m)})$ after m such cycles. This defines a Markov chain with transition kernel $k(\boldsymbol{\theta}^{(m+1)}, \boldsymbol{\theta}^{(m)}) = \Pi_{i=1}^n p(\theta_i^{(m+1)} \mid \theta_1^{(m+1)}, \ldots, \theta_{i-1}^{(m+1)}, \theta_{i+1}^{(m)}, \ldots, \theta_n^{(m)})$ that converges to the joint posterior as its equilibrium distribution (see Gilks et al. 1996).

Consequently, if all the full conditional distributions are available, all that is required is sampling iteratively from these. Once a sample $\{(\theta_1^{(m)}, \ldots, \theta_n^{(m)}, m = 1, \ldots, k\}$ from the joint posterior is available, note the ease with which a sample from the marginal posterior distribution of θ_1, say, and the posterior distributions of any function $g(\boldsymbol{\theta})$ can be obtained. These are simply given by $\{\theta_1^{(m)}, m = 1, \ldots, k\}$ and $\{g(\boldsymbol{\theta}^{(m)}, m = 1, \ldots, k$, respectively, and there is no need for high-dimensional integration or differentiation, respectively.

We will follow these two fundamental steps in a Bayesian stock assessment of yellowfin tuna in the eastern tropical Pacific Ocean using a state–space implementation of the delay difference model. The data, consisting of catch in millions of pounds and CPUE in pounds per boat-day for the years 1934–1967, are taken from Pella and Tomlinson (1969) and are listed in Table 1. The data set has been previously analyzed by Kimura et al. (1996) using the Kalman filter. We chose this historical data set to demonstrate the viability of our nonlinear state–space approach via Gibbs sampling and to compare results with those of ML estimation via Kalman filtering.

Constructing the joint probability model

In order to make results comparable with the analysis by Kimura et al. (1996), we make the same assumptions on the fixed model parameters. We assume linear growth in weight-at-age, i.e., $\rho = 1$, $\omega = 0$, and an instantaneous natural mortality rate of $M = 0.6$. Recruitment is assumed to be constant for all years, i.e., $R_t = R$ for all $t = 1,\ldots,N = 34$, and not equal to the equilibrium recruitment (Kimura et al. 1996) to allow for the possibility of a change in recruitment with the onset of fishing.

As an initial Gibbs sampling implementation in terms of total biomass exhibited extremely slow convergence, we divided the state equations by K to avoid high correlations between the states and carrying capacity. Thus, expressing the unknown total and recruitment biomasses as proportions of carrying capacity K, i.e., by the transformations

$$P_t = \frac{B_t}{K},\ k = \frac{1}{K},\ r = \frac{R}{K}, \text{ and } Q = qK$$

in eq. 4, we obtain a nonlinear state–space model with state transition equations

$$(6)\qquad P_1 = 1 + u_1$$

$$P_2 = e^{-M}(1+\rho-\rho e^{-M})(P_1 - kC_1) + r\left(1-\rho\omega e^{-M}\frac{(P_1-kC_1)}{P_1}\right) + u_2$$

$$P_{t+1} = (1+\rho)e^{-M}(P_t - kC_t) - \rho e^{-2M}\frac{(P_t-kC_t)}{P_t}(P_{t-1}-kC_{t-1}) + r\left(1-\rho\omega e^{-M}\frac{(P_t-kC_t)}{P_t}\right) + u_{t+1} \text{ for } t = 2,\ldots,N$$

and observation equations

$$(7)\qquad I_t = QP_t + v_t \quad \text{for } t = 1,\ldots,N.$$

© 1999 NRC Canada

We assume independent normal errors for $\{u_t\}$ and $\{v_t\}$. Specifically, $u_t \sim N(0, \sigma^2)$ and the CPUEs are given approximately constant coefficient of variation by assuming the v_t to be $N(0, w_t\tau^2)$ with weights w_t proportional to the squared fitted values obtained from a nonlinear robust smoothing of the CPUE time series by means of running medians (using the SPLUS function "smooth"). The weights are standardized by $w_N = 1$ so that $v_N \sim N(0, \tau^2)$.

The unobservables in the delay difference model are (k, r, Q, σ^2, τ^2, P_1,..., P_N). The joint prior density is given (using a repeated application of Bayes' theorem) by

$$(8) \quad p(k, r, Q, \sigma^2, \tau^2, P_1, \ldots, P_N) = p(k, r, Q, \sigma^2, \tau^2)p(P_1)p(P_2 | P_1, k, r, \sigma^2) \times \prod_{t=3}^{N} p(P_t | P_{t-1}, P_{t-2}, k, r, \sigma^2).$$

The specification of prior distributions in a Bayesian analysis is one of the most controversial issues. Walters and Ludwig (1994) and Punt and Hilborn (1997) gave some general guidelines for selecting priors for parameters of stock assessment models. Following the advice of Punt and Hilborn (1997) of a pragmatic choice between informative and noninformative priors, we put informative priors on r and τ^2 and noninformative priors on K, Q, and σ^2. As our prior information does not permit the specification of a prior correlation structure among the parameters, we assumed them to be independent.

The K, Q, and σ^2 can be regarded as scale parameters and a noninformative prior is therefore uniform on a log scale. An informative prior could in principle be constructed for r from recruitment data of related species and stocks. Here, we choose to construct a vague prior on r by using information on natural mortality of yellowfin tuna. With the assumption of linear growth-at-age, i.e., $\rho = 1$ and $\omega = 0$ in eq. 1, and using eqs. 2 and 3, one would expect a proportion of recruitment biomass to carrying capacity

$$r(M) = \frac{1}{\Sigma_{a=1}^{\infty} a \exp(-M(a-1))}$$

depending on the instantaneous natural mortality rate M if the stock were in equilibrium. Hennemuth (1961) analyzed length–frequency data of yellowfin tuna in the eastern Pacific from 1954 to 1959 to obtain an estimate of M of 0.77 with 95% confidence limits of 0.64 and 0.9. Schaefer (1967) considered the lowest and highest probable values to be 0.55 and 1.05. Until recently, the value $M = 0.8$ was used by the Inter-American Tropical Tuna Commission in age-structured assessment models; however, Francis (1977) concluded from simulations that 0.8 was too high and recommended $M = 0.6$. We therefore use $M = 0.2$ and $M = 1.0$ to construct a lognormal prior for r with 10% quantile equal to $r(0.2) = 0.155$ and 90% quantile equal to $r(1.0) = 0.445$.

In practice, informative priors can be constructed for the observation error variance based on information about the sampling design. Here, as in Carlin et al. (1992), we choose a vague conjugate inverse gamma distribution for τ^2 with mean and standard deviation equal to 250 000.

The sampling distribution (i.e., the likelihood function) is given by

$$(9) \quad p(I_1, \ldots, I_N | P_1, \ldots, P_N, Q, k, r, \sigma^2, \tau^2) = \prod_{t=1}^{N} p(I_t | P_t, Q, \tau^2).$$

In the following section, we describe an MCMC technique to sample from the posterior distribution.

Sampling from the posterior distribution

The large number of unobservables (34 states plus five parameters) clearly rules out numerical techniques to do the required integrations. In their recent review on the Bayesian approach to fisheries stock assessment and decision analysis, Punt and Hilborn (1997) described three Monte Carlo methods for posterior calculations: grid search, the Metropolis–Hastings (MH) algorithm, and the SIR algorithm. They favoured the last two algorithms because of their superior performance. The SIR algorithm has been successfully applied in Bayesian stock assessment using relative abundance data by McAllister et al. (1994), Raftery et al. (1995), and Kinas (1996) and using catch–age data by McAllister and Ianelli (1997). However, its efficiency depends crucially on the development of a good importance function that gives a reasonable approximation to the posterior. A frequently used candidate is the prior, a choice that can be very inefficient. McAllister and Ianelli (1997) suggested a multivariate t density, which requires a nonlinear minimization to find the posterior modes and an estimate of the inverse of the Hessian matrix of the posterior density. This makes the use of the SIR algorithm rather complicated and computationally expensive. Therefore, we suggest the Gibbs sampler, which requires generating from merely univariate densities without the necessity of calculating derivatives and the use of numerical optimization procedures.

A Gibbs sampling approach to dynamic nonlinear state–space models has been proposed by Carlin et al. (1992). This is the approach taken here for the implementation of state–space delay difference models.

First, we have to calculate the univariate full conditional posterior densities for all 39 unobservables in the model. The full conditional posterior density (up to a constant of proportionality) of a certain parameter θ_i can be constructed from the joint posterior of $\boldsymbol{\theta}$, the product of eqs. 8 and 9, by simply extracting the terms that involve θ_i. The forms of the full conditional densities for the unobserved quantities are given in the Appendix.

Since we use a conjugate prior for τ^2, its full conditional posterior density is again inverse gamma. Similarly, the full conditional posterior density of σ^2 is inverse gamma. Due to nonlinearity in the state equations, the full conditionals for the states P_t and the parameters k, r, and Q are rather complex. A simple rejection method (Devroye 1986), as proposed by Carlin et al. (1992), is no longer feasible because the acceptance probability would be too close to zero to be efficient. Therefore, we used the recently developed MCMC method of adaptive rejection Metropolis sampling (ARMS) (Gilks et al. 1995; Gilks and Neal 1997) to sample from an arbitrarily complex and not necessarily log-concave density.

© 1999 NRC Canada

This is a "Metropolized" version of adaptive rejection sampling, which uses rejection sampling to drive a Markov chain that converges to the full conditional posterior. Similar to rejection sampling, it only requires the full conditional density to be known up to a normalization constant. Like the derivative-free version of adaptive rejection sampling for log-concave densities, ARMS constructs an envelope function without specifying derivatives. A subroutine written in programming language C is available in Gilks et al. (1995).

We use ARMS to simulate not only from the full conditionals for P_t, k, r, and Q but for all other unobservables. The use of fast generators for normal and inverse gamma variables (e.g., Devroye 1986) could accelerate the implementation; however, we observed adequate performance with the universal method.

Comparison of results

We performed 250 000 cycles of the Gibbs sampler and thinned the chain by taking every 25th observation to avoid highly correlated values. For the remaining 10 000 samples, we used a burn-in of 1000, which yielded a final chain of length 9000.

Extensive convergence diagnostics were calculated for the states P_1, P_{34}, and P_{35} and the parameters K, r, Q, σ^2, and τ^2 using the CODA software of Best et al. (1995). All chains passed the Heidelberger and Welch (1983) stationarity and halfwidth test. The Raftery and Lewis (1992) convergence diagnostics confirmed that the thinning, burn-in period, and minimum sample size were sufficient. Lags and autocorrelations within each chain were reasonably low. Geweke's (1992) Z scores do not fall within the extreme tails of a standard normal distribution, suggesting that the chain fully converged. Trace plots and kernel estimates for the marginal posterior densities for the above unobservables are listed in Fig. 1. The prior density is shown as well for those parameters with a proper prior distribution. Kernel density estimates of the marginal prior densities for the states P_1, P_N, and $P_{N=1}$ were obtained by using the Gibbs sampler to sample from the joint prior (eq. 8). Summary statistics including mean, standard deviation, and the 25, 50, and 75% quantiles are given in Table 2.

As can be seen from the kernel density plot in Fig. 1, the posterior distribution of carrying capacity K is positively skewed, with posterior mean close to the upper quartile. An interquartile range from 852 × 10^6 to 1890 × 10^6 lb (1 lb = 0.454 kg) for K captures the two virgin biomass estimates of 1300 × 10^6 and 950 × 10^6 lb given by Kimura et al. (1996) under two different assumptions on the error variances. Pella's (1993) analysis using a Kalman filter implementation of a surplus production model gives a carrying capacity estimate of 1415 × 10^6 lb, which is midway between the posterior median and mean of K in Table 2.

The posterior medians of the biomasses vary from 453 × 10^6 to 1180 × 10^6 lb over the period from 1934 to 1967 and compare with a biomass range of 400 × 10^6 to 1000 × 10^6 lb from Kimura et al.'s (1996) analysis (under the assumption of predominantly measurement error) and a biomass range of 500 × 10^6 to 1500 × 10^6 lb from Pella's (1993) analysis (under the assumption of mainly process error). As for the forecasting problem, the delay difference model predicts a biomass with posterior mean equal to 51.1 ± 14.3% of carrying capacity for the following year, 1968.

Kimura et al. (1996) analyzed this data set by ML using the Kalman filter. This requires that all parameters enter linearly in the process and observation equations. However, the delay difference equations (written as in Kimura et al. 1996)

$$B_{t+1} = (1+\rho)s_t B_t - \rho s_t s_{t-1} B_{t-1} + R_{t+1} - \rho\omega s_t R_t$$

are nonlinear in B_{t-1} and B_t (even though not apparent at first sight) through the dependence of the annual survival rates s_t on B_t. Kimura et al. (1996) eluded nonlinear dependence of the annual survival rates s_t (= e^{-M-F_t}, where F_t is the instantaneous fishing mortality) on the biomass B_t by numerically solving the Baranov catch equations

$$C_t = B_t F_t \frac{(1-e^{-M-F_t})}{M+F_t}$$

sequentially for F_t, $t = 1,..., N$, once the biomass B_t was obtained through forward projection via the Kalman filter. This amounts to treating the unknown biomasses deterministically instead of as random. In their delay difference model, process error is interpreted strictly as the variation of recruitment only, ignoring variability due to growth, survival, and possibly other environmental sources. Therefore, the unknown parameters to be estimated in Kimura et al.'s (1996) model reduce to virgin biomass B_0, recruitment R, and the catchability q.

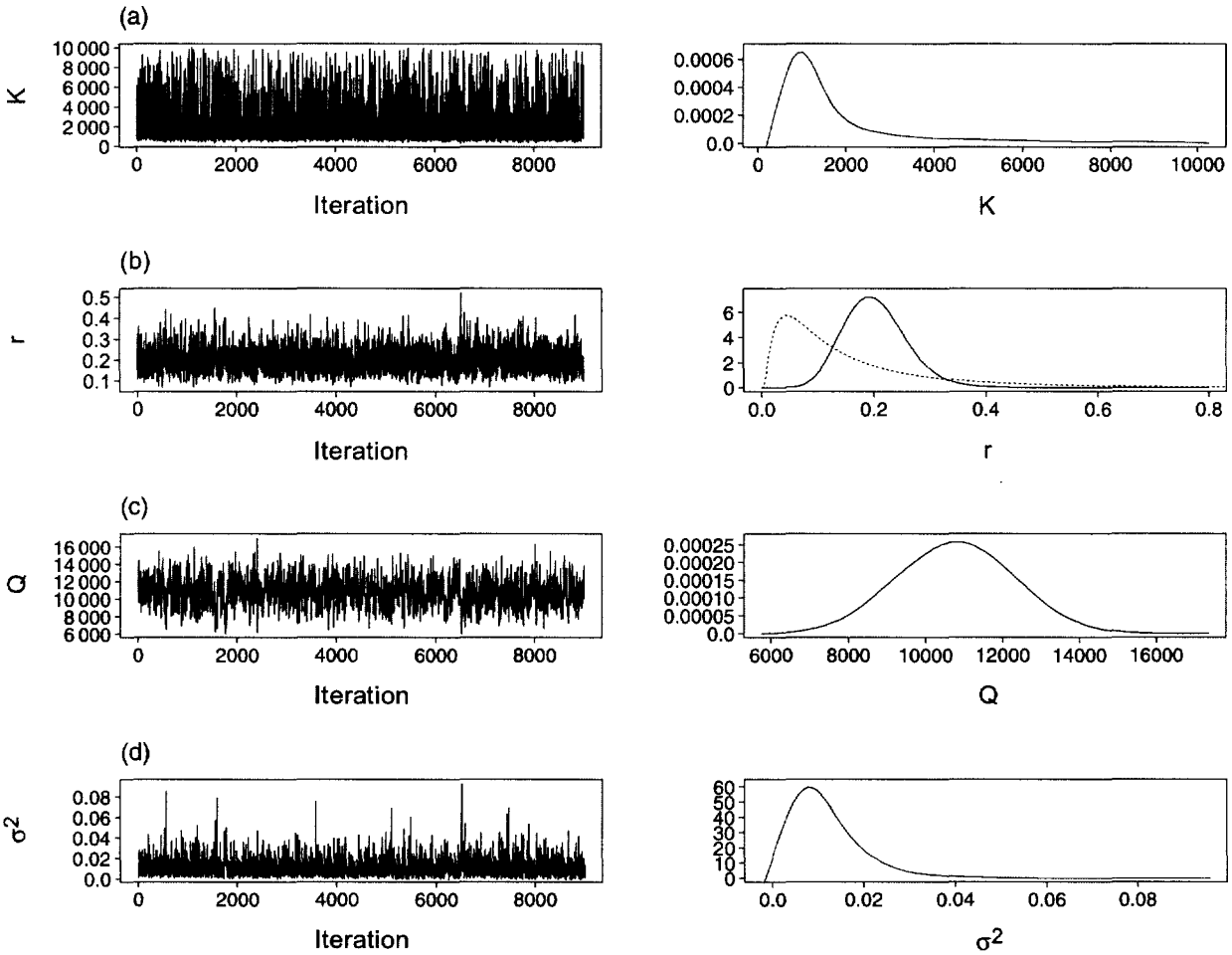
Our Bayesian approach gives estimates of the variability of unknown process and measurement errors in the form of posterior densities for the process and measurement error variances σ^2 and τ^2, respectively. In contrast, both σ^2 and τ^2 were assumed to be known in the Kalman filter implementation of Kimura et al. (1996). Although all parameters including the process and measurement error variances can be estimated using the Kalman filter when their ratio σ^2/τ^2 is known, Kimura et al. (1996) observed that the estimation of the ratio from the data is impractical, as the likelihood function appears to be insensitive to this ratio. Pella (1993) gave ML estimates of σ^2 = 32 775 and τ^2 = 46 but commented that measurement error might be underestimated by the Kalman filter methodology.

Kimura et al. (1996) analyzed the data under two different variance assumptions: (1) σ^2 = 32 775 and τ^2 = 46 (using Pella's (1993) estimates) and (2) σ^2 = 1000 and τ^2 = 5000. These values for τ^2 seem far too low in the light of our Bayesian analysis, which gives a posterior mean for τ^2 of 260 000. The two scenarios therefore are not quite adequately reflecting the situation of mainly process error (assumption 1) and mainly measurement error (assumption 2) as intended. The posterior mean of τ^2 is slightly higher than its prior mean of 250 000, but its posterior standard deviation is considerably reduced from 250 000 to 160 000. Furthermore, we observed a negative posterior correlation of −0.48 between τ^2 and σ^2.

We also fitted a surplus production model to the data. For the surplus production term that aggregates growth, recruitment, and natural mortality, we used the two-parameter Schaefer (1954) function $g(B) = sB\left(1-\frac{B}{K}\right)$, where K is the carrying capacity and s the so-called "intrinsic growth rate".

© 1999 NRC Canada

Fig. 1. Graphs on the left show trace plots of the MCMC simulations for (*a*) carrying capacity K, (*b*) recruitment r, (*c*) catchability Q, (*d*) process error variance σ^2, (*e*) observation error variance τ^2, (*f*) depletion in year 1, P_1, (*g*) depletion in year N, P_N, and (*h*) depletion in year $N + 1$, P_{N+1}. Graphs on the right show the corresponding prior marginal densities (broken lines) and the kernel estimates of posterior marginal densities (solid lines) obtained from MCMC samples using the Bayesian state–space approach to parameter estimation in the delay difference model. (Figure 1 concluded on next page.)



Note that the Schaefer (1954) surplus production function is a symmetric function of the previous biomass; it is zero if the previous biomass is zero or if the previous biomass is at its carrying capacity, i.e., when the population is in equilibrium where production due to growth and recruitment equals natural mortality. Using the same transformation $P_t = B_t/K$ as in the delay difference model, the corresponding nonlinear state–space model has state equations

$$P_1 = 1 + u_1$$

$$P_{t+1} = P_t + sP_t(1 - P_t) - \frac{1}{K}C_t + u_t \text{ for } t = 1, \ldots, N - 1$$

and observation equations as in eq. 4. We used a Bayesian approach with the same prior assumptions on the common parameters K, Q, σ^2, and τ^2 as for the delay difference model and a rather vague lognormal prior for s (as derived in Millar and Meyer 1998) with 10 and 90% quantiles equal to 0.1 and 1.0, respectively. As already mentioned in the introduction, biological interpretability of the parameters in the delay difference model facilitates the elicitation of informative priors. This is a much harder task for compound parameters such as the intrinsic growth rate in the surplus production model. The results are given in Table 3. With a posterior median of 2250 × 10^6 lb, this surplus production model gives a higher estimate of carrying capacity than the delay difference model and Pella's (1993) implementation of the surplus production model via Kalman filtering.

We plotted the posterior medians of the total biomasses $B_1, \ldots, B_N$ in Fig. 2 and included the ML/Kalman filter biomass estimates from Kimura et al. (1996) under the two different assumptions on the error variances. Unlike Kimura et al.'s (1996) predictions under assumption 1 of mainly process error, our predicted CPUEs do not follow strictly the observed CPUEs. It is not surprising that the Bayesian fit is in between that of the Kalman filter fit under assumption 1 of process error only and assumption 2 of when at least some measurement error is assumed. The predicted CPUEs

© 1999 NRC Canada

Fig. 1 (*concluded*).

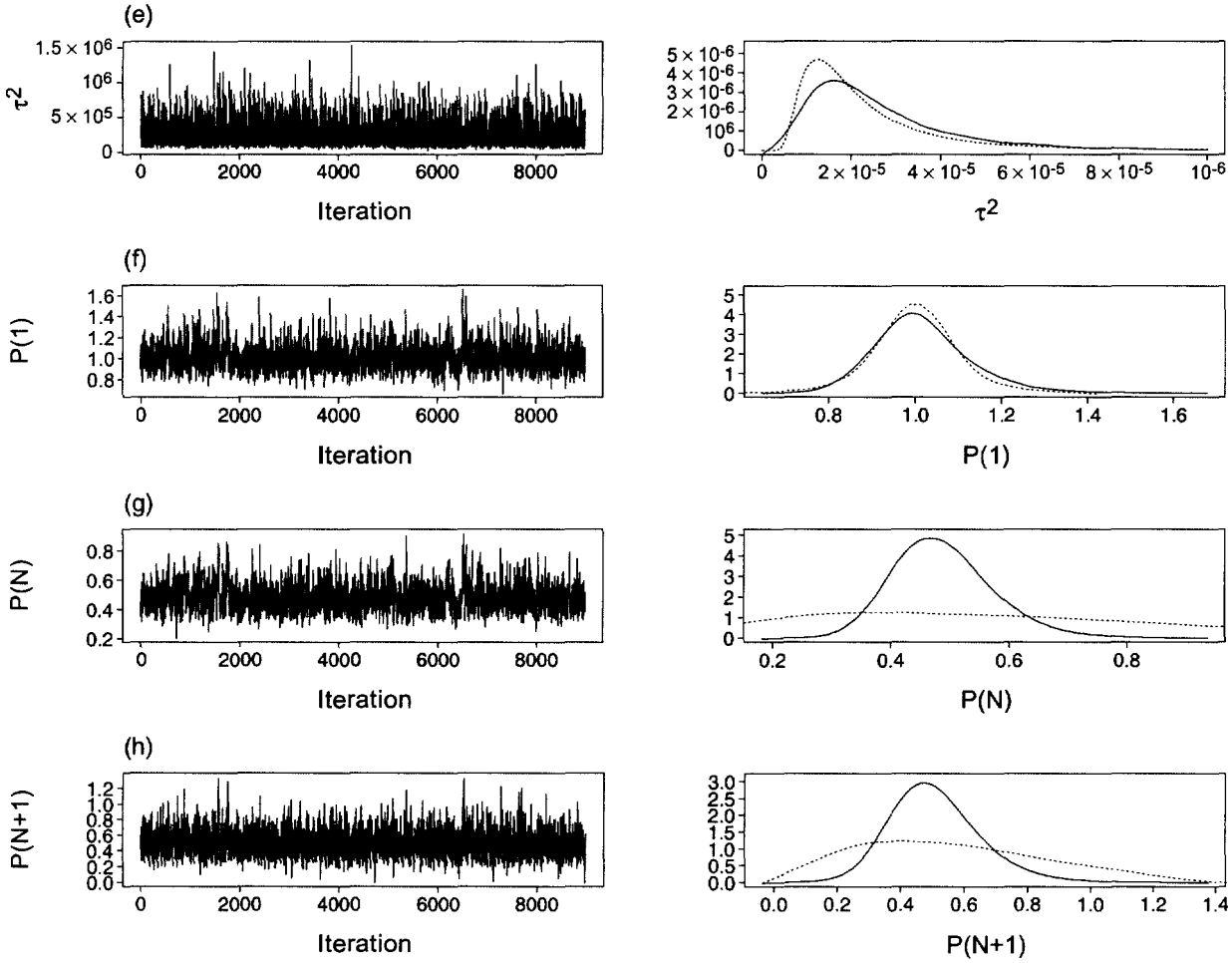


Table 2. Summary for sample size of 9000 from posterior density from the delay difference model.

Parameter	Mean	SD	25%	Median	75%
P_1	1.020	0.110	0.945	1.00	1.08
P_{34}	0.490	0.0845	0.440	0.481	0.539
P_{35}	0.511	0.143	0.414	0.496	0.592
K	1 800	1 680	852	1 150	1 890
Q	10 800	1 490	9 760	10 800	11 800
r	0.203	0.0517	0.166	0.199	0.234
σ^2	0.0114	0.00792	0.006	0.00975	0.0148
τ^2	260 000	160 000	148 000	216 000	324 000

Table 3. Summary for sample size of 9000 from posterior density from the surplus production model.

Parameter	Mean	SD	25%	Median	75%
P_1	1.03	0.104	0.957	1.02	1.08
P_{34}	0.500	0.0719	0.452	0.494	0.543
P_{35}	0.499	0.146	0.399	0.490	0.590
K	2 840	1 870	1 530	2 250	3 540
Q	10 400	1 080	9 740	10 500	11 100
s	0.350	0.207	0.196	0.308	0.457
σ^2	0.0142	0.00799	0.00907	0.0131	0.0181
τ^2	224 000	162 000	120 000	172 000	268 000

$Q{\cdot}P_t$ (= $q{\cdot}B_t$) for both the delay difference and the surplus production model, overlaid by the observed CPUEs, are shown in Fig. 3.

Management decision analysis

In their recent review on the Bayesian approach for fisheries stock assessment and decision analysis, Punt and Hilborn (1997) elaborated on how the results of a Bayesian analysis can be used to quantify risks associated with alternative management actions. The interested reader is also referred to Francis and Shotton (1997) for a recent review on risk in fisheries management.

For expository purposes, we will indicate here the typical way in which the output of the previous Bayesian stock assessment can be used as input to a decision analysis. Let us suppose that the alternative management options consist of setting different constant catch quotas for the next 5 years, a fixed annual total allowable catch of 150, 180, 200, and 220 × 10^6 lb. But note that more complex feedback-control

© 1999 NRC Canada

Fig. 2. Posterior medians of annual biomass of yellowfin tuna (1934–1967) obtained from the MCMC simulations using the Bayesian state–space approach to parameter estimation in the delay difference model. This is compared with the Kalman filter fit by Kimura et al. (1996) under assumption 1 (process error only) and assumption 2 (process and measurement error) as explained in the text under "Comparison of results".

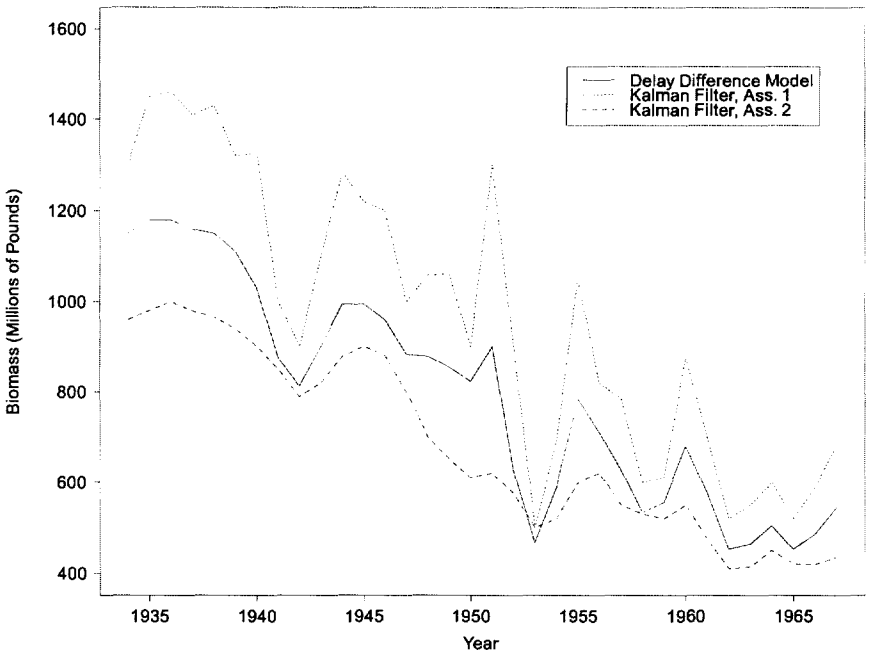
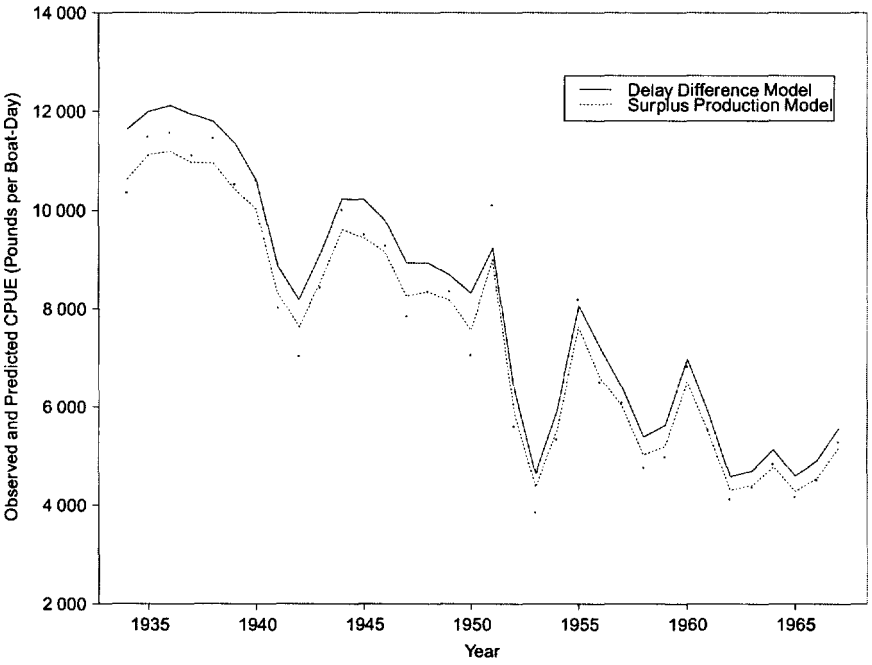


Fig. 3. Observed CPUEs and posterior means of the predicted CPUEs for yellowfin tuna (1934–1967) obtained from MCMC samples using the Bayesian state–space approach to parameter estimation in the delay difference model and in the surplus production model.



decision rules that depend on the current estimated state of the stock can easily be dealt with in the Bayesian state–space framework. Fisheries management will seek to assess the potential impact of these options. In practice, these impacts will be measured by various performance indicators such as decline/increase in stock size, minimum population size, or the variation in the exploitation rate, depending on management objectives. The performance measure chosen in this exemplary decision analysis is the probability of the biomass at the end of the management period, i.e., the begin-

© 1999 NRC Canada

Table 4. Posterior median of predicted annual depletion from 1968 to 1973 and posterior risk of $B(1973) < 0.3K$ under four different management options, i.e., constant 5-year total allowable catch (TAC) quotas (millions of pounds).

TAC option	$P(1968)$	$P(1969)$	$P(1970)$	$P(1971)$	$P(1972)$	$P(1973)$	Risk
150	0.502	0.528	0.553	0.569	0.581	0.590	0.077
180	0.502	0.506	0.510	0.512	0.514	0.519	0.151
200	0.502	0.493	0.482	0.475	0.469	0.464	0.235
220	0.502	0.477	0.454	0.430	0.417	0.401	0.340

ning of year 1973, falling below the threshold of 30% of virgin biomass.

It is straightforward to estimate this probability for a given set of catch quotas using MCMC simulation. The task is to generate a random sample from the marginal posterior distribution of P_{N+6}. Because this has a representation as a product of conditional posterior distributions of the previous states

$$p(P_{N+6}|\{I_t\}_{t=1}^N) \propto p(P_{N+6}|P_{N+5}, P_{N+4}, k, r, \sigma^2) \times \ldots \times p(P_{N+1}|P_N, P_{N-1}, k, r, \sigma^2) \times p(P_t, t=1,\ldots,N, k, r, Q, \sigma^2, \tau^2|\{I_t\}_{t=1}^N),$$

we only have to sample sequentially from the conditional posterior distribution of $P_{N+1},\ldots, P_{N+6}$ for each sample value that we already obtained from the posterior distribution of all the other unobservables. In practice, this is done by adding these six unknowns and their state equations to the set of unobservables in the existing Gibbs sampling program. Then, a consistent estimate of the probability that $P_{N+6} < 0.3$ is given by the corresponding relative frequency in the sample.

A clearly represented summary that could be presented to managers is in the form of a table that gives the estimate of this risk (and possibly other performance indicators) and the posterior medians of the predicted annual depletions (possibly together with corresponding interquartile ranges) for each management option. This is illustrated in Table 4. In addition, the posterior density of the future annual biomasses could be plotted. A plot would be preferred to a single summary statistic in this case because the biomass distributions are skewed.

Because a decision analysis uses parameter estimates from a population dynamics model, its quality depends crucially on that of the estimates and the ability of the stock assessment model to capture and quantify most of the uncertainty, in particular to account for both process and observation error. As this is the strength of Bayesian state–space methodology, a risk assessment based on its outcomes will ultimately provide an improvement to management advice.

Discussion

The nonlinear state–space approach to stock assessment within the Deriso–Schnute model family handles the problem of parameter estimation, smoothing, filtering, and forecasting very efficiently. This is accomplished using a Bayesian approach to statistical inference via the Gibbs sampler.

This approach offers a superior alternative to the existing techniques of nonlinear least squares and ML via Kalman filtering. Application of the Kalman filter suffers from some severe restrictions. Unless the (more complicated) extended Kalman filter is employed, it requires linearity of state and observation equations in all the model parameters (as also in Sullivan 1992; Reed and Simons 1996). The normal distribution assumption is crucial and the variances of process and observation errors (or at least their ratio) must be known. As stated in Pella (1993), it is not possible to "cleanly disentangle process from measurement error". To estimate an unknown state at time t, only the previously estimated biomasses of the preceding time points enter into the forward projections, thus not making full use of all the information provided in the data. This can be overcome, however, using fixed-interval smoothing, as outlined in Pella (1993). In the special application to Deriso–Schnute models, the assumptions of the Kalman filter are not met when $w > 0$, although Kimura et al. (1996) concluded from simulation studies that the Kalman filter results are not severely biased when applied in this general case. When trying to find the parameter values that maximize the likelihood function, one has to rely on numerical optimization procedures that are guaranteed only to find a local, not the global, maximum. Furthermore, measures of precision of the ML estimates cannot be provided, as the information matrix in most biomass dynamics applications is singular or its determinant is close to zero, making asymptotic variance estimates obtained by matrix inversion doubtful.

In contrast with these constraints, the Bayesian approach can handle arbitrary distributional assumptions as well as any form of nonlinear relationships in the state and observation equations. Process error of the biomass dynamics equations is not only confined to recruitment variability but comprises variation due to growth, mortality, and environmental influences. There is no need for a restriction to constant recruitment. A three-parametric recruitment function (of the spawning stock S_t in year t) (e.g., Schnute 1985)

$$(10) \quad R_t = R(S_t) = \alpha S_t(1-\beta\gamma S_t)^{1/\gamma}$$

that includes the constant productivity, Ricker, Beverton–Holt, and Schaefer functions as special cases could be specified. As pointed out by a referee, there is no reason that the conventional assumption that all of the process error is due to variation in recruitment could not be incorporated into the biomass dynamics equations. Furthermore, different assumptions on fishing mortality such as those given in Schnute (1985) can be made. Important extensions would be to consider stochastic natural mortality, which, for comparative purposes, was assumed to be known in this application. Stochastic historical catches can be incorporated by distinguish-

© 1999 NRC Canada

ing between observed catches, $\hat{C}_t$, and true catches, C_t, and adding the observation equation $\hat{C}_t = C_t + v_t^C$. In the Bayesian analysis, measurement and process errors are clearly separated and the precision of error variance estimates (as of all the other parameters) can be assessed in detail from the posterior densities.

Apart from ease and standardization, one of the major advantages of our implementation using the all-purpose subroutine ARMS to sample from any non-log-concave density is that it allows specifying more realistic distributions than the normal for measurements and states. This does not cause any further complications, as all that is required is the specification of the logarithm of the density up to an additive constant; no derivatives or modes of the densities are necessary (as in the tuning of the multivariate t distribution in implementations of SIR).

So far, none of the Bayesian stock assessment models have used state–space methodology. McAllister et al. (1994) and McAllister and Ianelli (1997) fitted a fully age-structured model to relative abundance and catch–age data, respectively, allowing for observation error with known variance but using deterministic annual transitions of the numbers of fish in each age-class, treating only annual recruitment as subject to random fluctuations. As noted above, this case is easily subsumed in a state–space model. Similarly, Raftery et al. (1995) considered a deterministic population dynamics model for bowhead whales and developed a pseudo-Bayesian approach for "deterministic simulation models," which has been heavily criticized by pure Bayesians (Wolpert 1995) for being prone to the Borel paradox and violating fundamental statistical principles such as the likelihood and stopping-rule principle. Raftery et al. (1995) regarded the population dynamics model as a deterministic simulation model with input parameters (in their application, consisting of age-specific natural mortality and fertility rates, kill records, and initial population size) and output parameters (the population size for each year broken down by age and sex). They proposed to specify all premodel information on input and output parameters through a "premodel" distribution. This encompasses all available information including that of the data but not including that from the simulation model itself. In analogy to the Bayesian paradigm that updates the prior distribution to the posterior distribution by conditioning on the data, the premodel distribution is updated to the "postmodel" distribution through the simulation model. Like Wolpert (1995), we do not agree that this may be viewed as a generalization of standard Bayesian inference to deal with simulation models. Instead, we think that a more natural approach within the coherent Bayesian paradigm would be to integrate these into the framework of state–space models, regarding the model outputs as unknown states, incorporating any prior information about these in a stochastic distribution of the states, and relating these unknown states to the data through stochastic observation equations.

Kinas (1996) used a Bayesian approach to estimate the parameters of a surplus production model using adaptive importance sampling and SIR. The specification of his model comes probably closest to a state–space model, allowing for both process and observation error. However, only the parameters of the surplus production model and the observation error variance were estimated. Although alluding to the biomass dynamics equations as stochastic, Kinas (1996) did not estimate the unknown states, the annual biomass of orange roughy (*Hoplostethus atlanticus*), in a coherent Bayesian way by treating them as unobservables, like the model parameters, and specifying a joint prior distribution. Instead, he employed an adhoc time series fitting procedure for predicting these states. Our Bayesian state–space methodology could be readily applied to the model specified in Kinas (1996) using the Gibbs sampler for posterior computation.

Depending on the parametrization and the correlation among the parameters, mixing in the Gibbs sampler could be slow and subsequent observations could be highly correlated. High autocorrelations can be overcome by a reasonable thinning of the chain, as suggested by the Raftery and Lewis (1992) diagnostics on a small test chain. Reparametrizations and joint updating of several parameters may have to be considered.

The practical construction of the full conditional posterior densities is not a substantial task. It only requires calculating the joint posterior density up to a normalization constant, i.e., the product of prior and likelihood. The full conditional posterior density of a certain component θ_i, say, can then be obtained by simply ignoring all terms in the joint density that do not depend on θ_i, as these are subsumed in the normalization constant. This procedure can be automated. It compares with the task in a general MH algorithm to specify a proposal density $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{(m)}$ (depending on the current state $\boldsymbol{\theta}^{(m)}$) to generate a new random variate $\boldsymbol{\theta}^*$ that will then be accepted with probability

$$\alpha(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{(m)}) = \min\left\{1, \frac{p(\boldsymbol{\theta}^*|\,\mathbf{y})q(\boldsymbol{\theta}^{(m)}|\boldsymbol{\theta}^*)}{p(\boldsymbol{\theta}^{(m)}|\,\mathbf{y})q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{(m)})}\right\}.$$

Following the "divide et impera" principle, the Gibbs sampler reduces the problem of generating an n-dimensional random variate $\boldsymbol{\theta}^*$ to that of drawing n univariate random variates, a comparatively simple task. Moreover, the MH algorithm, just like the SIR algorithm, suffers from the general difficulty of choosing an n-dimensional proposal density that is both easy to sample from and close to the target density. Because of large sample asymptotics, a multivariate normal or t distribution is often the proposal density of choice, which requires finding the mode of the posterior as well as differentiation to calculate the inverse of the Hessian, needed to specify its mean and covariance matrix. The efficiency of the MH algorithm depends crucially on the choice of these parameters, like the center, scale, and degrees of freedom of the t distribution, and these should be tuned so that the acceptance probability is roughly 0.3 (for recommendations, see Casella and George 1992; Gamerman 1997). This is because an MH chain that makes only small jumps has a high acceptance rate but mixes slowly, as it takes a long time to traverse the whole parameter space, whereas a chain with large proposal steps will have low acceptance probabilities, making the chain stay at its current value for long periods and thereby resulting in slow mixing as well. Most likely, the costs of calculating the proposal density and fine-tuning the MH algorithm will balance those of specifying the full conditionals for the Gibbs sampler and finding reparametrizations with low posterior correlations. A

© 1999 NRC Canada

practical comparison as to the relative performance of the MH algorithm and the Gibbs sampler in stock assessment applications is beyond the scope of this paper but an important issue that will have to be investigated.

The computation time needed to run a large chain has not been a problem at all in our implementations of surplus production and delay difference models. For a problem with 40 unobservables the run of a C program generating 250 000 cycles of the Gibbs sampler using ARMS for all full conditionals took about 90 min on a SUN ULTRA. This suggests that the Gibbs sampler might be a feasible option in a Bayesian analysis of more complicated age- or size-structured models such as those of Schnute (1987), Fournier and Doonan (1987), Raftery et al. (1995), or McAllister and Ianelli (1997) and a contender for the SIR algorithm.

Despite all of the pros reviewed in the introduction, delay difference models have been used only rarely in practical stock assessments (Gallaway et al. 1983; Quinn et al. 1984; Deriso 1985; Jacobson et al. 1987; Wankowski and Williams 1987; Punt 1988; Zheng and Walters 1988; Sampson 1990; Collie and Walters 1991; Polovina 1991; Hall and Brown 1995), especially when compared with the prominent surplus production models. Reasons for this may be the strong assumptions that delay difference models make on linear growth of weight-at-age (eq. 1) and on equality of spawning and exploitable biomasses when modeling recruitment as a function of spawning stock size as through eq. 10. These may often be violated in practical applications. However, the latter is of no concern in this particular application where recruitment is independent of the spawning biomass, and in general, the delay difference model allows for an annual recruitment rate that can be any arbitrary time series (eq. 4). Another reason that could have deterred from their application may be the mathematically more complicated form of the delay difference equations that incorporates time lags and recruitment functions and makes parameter estimation more involved. However, they have been widely used in simulation studies and proved to be valuable tools that provide insight into population dynamics, temporal changes, and the understanding of age-structure effects in exploited stocks (Hilborn and Walters 1992). We hope that with the new methodology for parameter estimation provided here, delay difference models will come to enjoy greater popularity in future practical stock assessments, in tune with their biological relevance.

Acknowledgements

This work was supported by the University of Auckland Research Committee. We would like to thank Andre Punt and an anonymous referee for their helpful suggestions and constructive criticism on an earlier version of this paper.

References

Berger, J.O., and Wolpert, R.L. 1988. The likelihood principle. 2nd ed. IMS Lect. Notes Monogr. Ser. No. 9. Institute of Mathematical Statistics, Hayward, Calif.

Best, N.G., Cowles, M.K., and Vines, S.K. 1995. CODA manual version 0.30. MRC Biostatistics Unit, Cambridge, U.K.

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

Casella, G., and George, E.I. 1992. Explaining the Gibbs sampler. Am. Stat. **46**: 167–174.

Collie, J.S., and Walters, C.J. 1991. Adaptive management of spatially replicated groundfish populations. Can. J. Fish. Aquat. Sci. **48**: 1273–1284.

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

Deriso, R.B. 1985. Stock assessment and new evidence of density-dependence. *In* Fisheries dynamics: harvest management and sampling. *Edited by* P.R. Mundy, T.J. Quinn, and R.B. Deriso. Tech. Rep. Wash. Sea Grant, Seattle, Wash. pp. 49–60.

Devroye, L. 1986. Non-uniform random variate generation. Springer, New York.

Evans, M., and Swartz, T. 1995. Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. Stat. Sci. **10**: 254–272.

Fahrmeir, L., and Tutz, G. 1994. Multivariate statistical modelling based on generalized linear models. Springer, New York.

Fournier, D.A., and Archibald, C. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. **39**: 1195–1207.

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

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

Francis, R.C. 1977. TUNPOP: a simulation of the dynamics and structure of the yellowfin tuna stock and surface fishery of the eastern Pacific Ocean. Inter-Am. Trop. Tuna Comm. Bull. **16**: 233–279.

Francis, R.I.C.C. 1992. Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. **49**: 922–930.

Francis, R.I.C.C., and Shotton, R. 1997. "Risk" in fisheries management: a review. Can. J. Fish. Aquat. Sci. **54**: 1699–1715.

Gallaway, B.J., Griffiths, W.B., Craig, P.C., Gazey, W.J., and Helmericks, J.W. 1983. An assessment of the Colville River Delta (Alaska, U.S.A.) stock of arctic cisco (*Coregonus autumnalis*): migrants from Canada? Biol. Pap. Univ. Alaska, **0**(21): 4–23.

Gamerman, D. 1997. Markov chain Monte Carlo. Stochastic simulation for Bayesian inference. Chapman and Hall, London, U.K.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 1995. Bayesian data analysis. Chapman and Hall, London, U.K.

Geweke, J. 1992. Evaluation of the accuracy of sampling-based approaches to calculating posterior moments. *In* Bayesian statistics 4. *Edited by* J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith. Oxford University Press, Oxford, U.K. pp. 169–193.

Gilks, W.R., and Neal, R.M. 1997. Corrigendum: adaptive rejection Metropolis sampling. Appl. Stat. **46**: 541–542.

Gilks, W.R., Best, N.G., and Chan, K.K.C. 1995. Adaptive rejection Metropolis sampling within Gibbs sampling. Appl. Stat. **44**: 455–472.

Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. 1996. Markov chain Monte Carlo in practice. Chapman and Hall, London, U.K.

Gudmundsson, G. 1994. Time series analysis of catch-at-age observations. Appl. Stat. **43**: 117–126.

© 1999 NRC Canada

Hall, N.G., and Brown, R.S. 1995. Delay-difference models for the western rock lobster (*Panulirus cygnus*) fishery of Western Australia. ICES Mar. Sci. Symp. **199**: 399–410.

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

Hennemuth, R.C. 1961. Year class abundance, mortality and yield-per-recruit of yellowfin in the eastern Pacific Ocean, 1954–1959. Inter-Am. Trop. Tuna Comm. Bull. **6**:1–51.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, London, U.K.

Jacobson, D.L., MacCallum, W.R., and Spangler, G.R. 1987. Biomass dynamics of Lake Superior lake herring (*Coregonus artedii*): application of Schnute's difference model. Can. J. Fish. Aquat. Sci. **44**(Suppl. 2): 275–288.

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

Kimura, D.K., Balsiger, J.W., and Ito, D.H. 1996. Kalman filtering the delay-difference equation: practical approaches and simulations. Fish. Bull. U.S. **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.

Ludwig, D., and Walters, C.J. 1981. Measurement errors and uncertainty in parameter estimates for stock and recruitment. Can. J. Fish. Aquat. Sci. **38**: 711–720.

Ludwig, D., and Walters, C.J. 1989. A robust method for parameter estimation from catch and effort data. Can. J. Fish. Aquat. Sci. **46**: 137–144.

McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. **54**: 284–300.

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.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Am. Fish. Soc. Symp. **6**: 4–48.

Millar, R.B., and Meyer, R. 1998. Nonlinear state–space modeling of fisheries biomass dynamics using the Gibbs sampler. Tech. Rep. STAT9802. Available from the authors upon request or as a postscript file from http://www.stat.aukland.ac.nz/reports/techrpt.html

O'Hagan, A. 1987. Monte Carlo is fundamentally unsound. Statistician, **36**: 247–249.

O'Hagan, A. 1994. Bayesian inference. Vol. 2B of Kendall's Advanced theory of statistics. Edward Arnold, London, U.K.

Pella, J.J. 1993. Utility fo structural time series models and the Kalman filter for predicting consequences of fishery actions. *In* Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. *Edited by* G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II. Alaska Sea Grant College Program Rep. 93-02, University of Alaska, Fairbanks, Alaska. pp. 571–593.

Pella, J.J., and Tomlinson, P.K. 1969. A generalized stock production model. Inter-Am. Trop. Tuna Comm. Bull. **13**: 421–496.

Polachek, 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.

Polovina, J.J. 1991. Evaluation of hatchery releases of juveniles to enhance rockfish stocks, with application to Pacific ocean perch *Sebastes alutus*. Fish. Bull. U.S. **89**: 129–136.

Punt, A.E. 1988. Model selection for the dynamics of southern African hake resources. M.S. thesis, Department of Applied Mathematics, University of Cape Town, Cape Town, South Africa.

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

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.

Quinn, T.J., Deriso, R.B., Hoag, S.H., and Myhre, R.J. 1984. A summary of methods of estimating annual surplus production for the Pacific halibut fishery. Int. North Pac. Fish. Comm. Bull. **42**: 73–81.

Raftery, A.L., and Lewis, S. 1992. How many iterations in the Gibbs sampler? *In* Bayesian statistics 4. *Edited by* J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith. Oxford University Press, Oxford, U.K. pp. 763–774.

Raftery, A.E., Givens, G.H., and Zeh, J.E. 1995. Inference from a deterministic population dynamics model for bowhead whales. J. Am. Stat. Assoc. **90**: 402–416.

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.

Rubin, D.R. 1987. A noniterative sampling importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: the SIR algorithm, comment to a paper by Tanner and Wong. J. Am. Stat. Assoc. **82**: 543–546.

Rubin, D.R. 1988. Using the SIR algorithm to simulate posterior distributions (with discussion). *In* Bayesian statistics 3. *Edited by* J.M. Bernardo et al. Oxford University Press, Oxford, U.K. pp. 395–402.

Rubinstein, R.Y. 1981. Simulation and the Monte Carlo method. Wiley, New York.

Sampson, D.B. 1990. A length-structured population model for southern fin whales and a test for density dependence. J. Cons. Cons. Int. Explor. Mer, **46**: 249–268.

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

Schaefer, M.B. 1967. Fishery dynamics and present status of the yellowfin tuna population of the eastern Pacific Ocean. Inter-Am. Trop. Tuna Comm. Bull. **12**: 87–136.

Schnute, J.T. 1985. A general theory for analysis of catch and effort data. Can. J. Fish. Aquat. Sci. **42**: 414–429.

Schnute, J.T. 1987. A general fishery model for a size-structured fish population. Can. J. Fish. Aquat. Sci. **44**: 924–940.

Schnute, J.T. 1989. The influence of statistical error on stock assessment: illustrations from Schaefer's model. *In* Effects of ocean variability on recruitment and an evaluation of parameters in stock assessment models. *Edited by* R.J. Beamish and G.A. McFarlane. Can. Spec. Publ. Fish. Aquat. Sci. No. 108. pp. 101–109.

Schnute, J.T. 1991. The importance of noise in fish population models. Fish. Res. (Amst.), **11**: 197–223.

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

Smith, A.F.M., and Gelfand, A.E. 1992. Bayesian statistics without tears: a sampling–resampling perspective. Am. Stat. **46**: 84–88.

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

Tierney, L. 1994. Markov chains for exploring posterior distributions (with discussion). Ann. Stat. **22**: 1701–1762.

© 1999 NRC Canada

Tierney, L., and Kadane, J.B. 1986. Accurate approximations for posterior moments and marginal densities. J. Am. Stat. Assoc. **81**: 82–86.

Walters, C. 1980. Comment on Deriso's delay-difference population model. Can. J. Fish. Aquat. Sci. **37**: 2365.

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

Wankowski, J.W.J., and Williams, S. 1987. Yield estimates for the eastern Bass Strait continental shelf demersal trawl fishery. Tech. Rep. Mar. Sci. Lab. No. 65, Department of Conservation, Forests and Lands, Queenscliff, Victoria, Australia.

West, M., and Harrison, P.J. 1997. Bayesian forecasting and dynamic models. 2nd ed. Springer, New York.

Wolpert, R. 1995. Comment on Raftery et al. 1995: Inference from a deterministic population dynamics model for bowhead whales. J. Am. Stat. Assoc. **90**: 426–427.

Zheng, J., and Walters, C.J. 1988. Population dynamics and stock assessment of Wanshan spring decapterus-maruadsi T. and S. in the south China Sea. Fish. Res. (Amst.), **6**: 217–232.

Appendix

In the following, let

$$g(P_t) = \begin{cases} 1 & \text{for } t = 1 \\ e^{-M}(1+\rho-\rho e^{-M})(P_1 - kC_1) + r\left(1 - \rho\omega e^{-M}\dfrac{(P_1 - kC_1)}{P_1}\right) & \text{for } t = 2 \\ (1+\rho)e^{-M}(P_{t-1} - kC_{t-1}) - \rho e^{-2M}\dfrac{(P_{t-1} - kC_{t-1})}{P_{t-1}}(P_{t-2} - kC_{t-2}) + r\left(1 - p\omega e^{-M}\dfrac{(P_{t-1} - kC_{t-1})}{P_{t-1}}\right) & \text{for } t = 3,\ldots,N+1. \end{cases}$$

Full conditional posterior density of P_t, $t = 3,\ldots, N-2$

$$\begin{aligned} &p(P_t|P_1,\ldots,P_{t-1},P_{t+1},\ldots,P_N,k,r,Q,\sigma^2,\tau_2) \\ &\quad \propto p(P_t|P_{t-1},P_{t-2},k,r,\sigma^2)\times p(P_{t+1}|P_t,P_{t-1},k,r,\sigma^2)\times p(P_{t+2}|P_{t+1},P_t,k,r,\sigma^2)\times p(I_t|P_t,Q,\tau^2) \\ &\quad \propto \exp\left(-\frac{1}{2\sigma^2}\{(P_t - g(P_t))^2 + (P_{t+1} - g(P_{t+1}))^2 + (P_{t+2} - g(P_{t+2}))^2\}\right)\times \exp\left(-\frac{1}{2w_t\tau^2}(I_t - QP_t)^2\right). \end{aligned}$$

Similar expressions are obtained for P_1, P_2, P_{N-1}, P_N, and P_{N+1} by dropping respective terms.

Full conditional posterior density of k

$$p(k|P_1,\ldots,P_N,r,Q,\sigma^2,\tau^2) \propto p(k)\prod_{t=2}^{N} p(P_t|P_{t-1},P_{t-2},k,r,\sigma^2) \propto \begin{cases} \dfrac{1}{k}\exp\left(-\dfrac{1}{2\sigma^2}\displaystyle\sum_{t=2}^{N}(P_t - g(P_t))^2\right) & k > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Full conditional posterior density of r

$$p(r|P_1,\ldots,P_N,k,Q,\sigma^2,\tau^2) \propto p(r)\prod_{t=2}^{N} p(P_t|P_{t-1},P_{t-2},k,r,\sigma^2) \propto \begin{cases} \dfrac{1}{r}\exp\left(-\dfrac{(\log r - \mu_r)^2}{2\sigma_r^2} - \dfrac{1}{2\sigma^2}\displaystyle\sum_{t=2}^{N}(P_t - g(P_t))^2\right) & r > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Full conditional posterior density of Q

$$p(Q|P_1,\ldots,P_N,k,r,\sigma^2,\tau^2) \propto p(Q)\prod_{t=2}^{N} p(I_t|P_t,Q,\tau^2) \propto \begin{cases} \dfrac{1}{Q}\exp\left(\dfrac{-1}{2\tau^2}\displaystyle\sum_{t=1}^{N}\dfrac{(I_t - QP_t)^2}{w_t}\right) & Q > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Full conditional posterior density of σ^2

The full conditional distribution for σ^2 is IG(α, β), where

$$\alpha = \frac{N}{2}, \qquad \beta = \frac{1}{2}\sum_{t=1}^{N}(P_t - g(P_t))^2.$$

© 1999 NRC Canada

Full conditional posterior density of τ^2

The IG(3, 500 000) prior distribution is conjugate family, so that the full conditional distribution for τ^2 is again IG(α, β), where

$$\alpha = 3 + \frac{N}{2}, \qquad \beta = 500\,000 + \frac{1}{2}\sum_{t=1}^{N} \frac{(I_t - QP_t)^2}{w_t}.$$

© 1999 NRC Canada