

Extending production models to include process error in the population dynamics

Andre E. Punt

Abstract: Four methods for fitting production models, including three that account for the effects of error in the population dynamics equation (process error) and when indexing the population (observation error), are evaluated by means of Monte Carlo simulation. An estimator that represents the distributions of biomass explicitly and integrates over the unknown process errors numerically (the NISS estimator) performs best of the four estimators considered, never being the worst estimator and often being the best in terms of the medians of the absolute values of the relative errors. The total-error approach outperforms the observation-error estimator conventionally used to fit dynamic production models, and the performance of a Kalman filter based estimator is particularly poor. Although the NISS estimator is the best-performing estimator considered, its estimates of quantities of management interest are severely biased and highly imprecise for some of the scenarios considered.

Résumé : Une simulation de Monte Carlo permet d'évaluer quatre méthodes d'ajustement des modèles de production, dont trois qui prennent en compte les effets des erreurs dans la dynamique de population (erreurs de processus) et dans l'indexation de la population (erreurs d'observation). Des quatre estimateurs examinés, l'estimateur NISS, qui représente la distribution de la biomasse de façon explicite et qui fait numériquement l'intégration des erreurs de processus inconnues, fonctionne le mieux; il n'est jamais le pire et est très souvent le meilleur en ce qui concerne les médianes des valeurs absolues des erreurs relatives. La méthode de l'erreur totale fonctionne mieux que celle de l'erreur d'observation couramment utilisée dans l'ajustement de modèles dynamiques de production; la méthode qui utilise un filtre de Kalman fonctionne particulièrement mal. Bien que la méthode NISS soit la meilleure des méthodes étudiées, les estimations d'intérêt pour la gestion qu'elle produit sont particulièrement faussées et très imprécises dans certains scénarios.

[Traduit par la Rédaction]

Introduction

Production (or biomass dynamics) models have a long history in quantitative fisheries science and have provided a key component of the basis that led to the popularity of maximum sustainable yield (MSY), and its associated biomass (B_{MSY}), as biological reference points for fisheries management. Production models are virtually the only method for stock assessment in situations in which the only data available are a time series of catches and some index of abundance. Historically, the index of abundance was fishery catch-per-unit-effort (CPUE), but the use of the results from fishery-independent resource surveys in production model assessments is more common now given the known limitations of CPUE as an index of abundance (e.g., Cooke and Beddington 1984).

Production models remain a useful tool for providing fisheries management advice in situations in which (a) there are insufficient age- and size-composition data to permit the use of assessment methods based on age- or size-structured models or (b) there is interest in biomass levels and surplus production before the years for which age- and size-composition

data are available. The latter is particularly important when efforts are being made to rebuild overfished stocks to earlier levels that have not been observed since age- and size-composition data were collected.

The production models used for management purposes have changed over time from those assuming that each data point represents a population in dynamic equilibrium through those averaging the observed data over several years to approximate an equilibrium situation (e.g., Gulland 1961; Fox 1975) to those that do not make the assumption that the population is in equilibrium, i.e., dynamic production models (e.g., Pella and Tomlinson 1969).

Despite considerable evidence that methods based on the assumption that the population is in equilibrium are likely to lead to estimates of biological reference points that are substantially biased (e.g., Sissenwine 1978; Polacheck et al. 1993; Williams and Prager 2002), these methods are nevertheless still used today to some extent (e.g., Moses 2000; International Commission for the Conservation of Atlantic Tunas (ICCAT) 2000). The development of production models has, since the 1980s, been focused primarily on dynamic production models. Three main classes of dynamic production model have emerged. Conceptually, these three classes are special cases of the following generic production model:

$$(1a) \quad B_{y+1} = (B_y + g(B_y) - C_y)e^{\varepsilon_y}$$

$$(1b) \quad I_y = qB_y e^{\eta_y}$$

Received 5 November 2002. Accepted 31 August 2003.
Published on the NRC Research Press Web site at
<http://cjfas.nrc.ca> on 3 November 2003.
J17170

A.E. Punt. School of Aquatic and Fishery Sciences, Box
355020, University of Washington, Seattle, WA 98195-5020,
U.S.A. (e-mail: aepunt@u.washington.edu).

where B_y is the biomass at the start of year y ($y = 1, 2, \dots, n + 1$); $g(B)$ is surplus production as a function of biomass; ε_y is the extent of process error during year y , $\varepsilon_y \sim N(0; \sigma_\varepsilon^2)$; I_y is the index of abundance for year y ; q is the catchability coefficient; and η_y is the extent of measurement (or observation) error during year y , $\eta_y \sim N(0; \sigma_\eta^2)$.

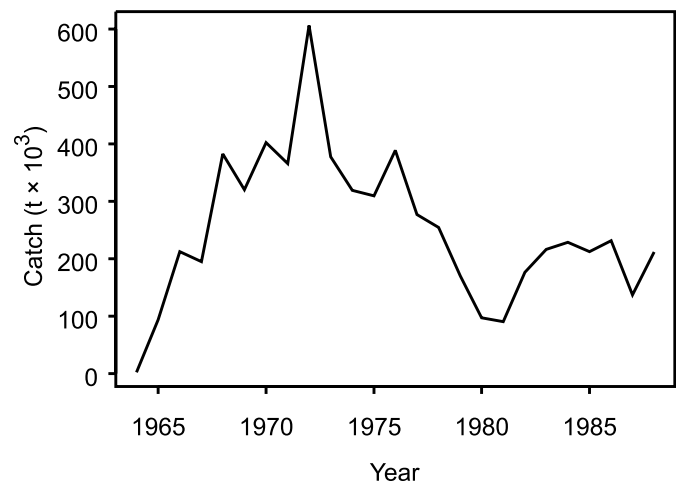
The three classes of production-model estimator on which actual assessments have been based are (i) observation-error estimators ($\sigma_\varepsilon = 0$), (ii) process-error estimators ($\sigma_\eta = 0$), and (iii) total-error estimators (also known as errors-in-variables estimators; $\sigma_\varepsilon/\sigma_\eta = C$, where C is a prespecified constant). There are a variety of sources for process error, including variability in recruitment, growth, and natural mortality, whereas observation error arises from, for example, sampling error and (random) variation in the availability of the population to the sampling gear at the time of sampling. Although Hilborn (2001) calculates surplus production directly, most production model estimators assume $g(B)$ to be one of a small number of unimodal functions, the most common of which are the Schaefer (1954), Fox (1970), and Pella–Tomlinson (1969) forms.

Initially, process-error estimators were most popular because, for some choices for $g(B)$, parameter estimation could be achieved by means of multiple linear regression rather than nonlinear regression techniques (e.g., Schnute 1977; Hilborn 1979). However, simulation studies (e.g., Punt 1989a; Polacheck et al. 1993) have shown that process-error estimators tend to provide less precise estimates of model outputs such as current biomass and MSY than observation-error estimators. Although the total-error approach of Ludwig et al. (1988) has formed the basis for actual stock assessments (e.g., Punt 1989b), the bulk of recent applications of dynamic production models have been based on the observation-error formulation (Prager 1994, 2002; Williams and Prager 2002).

Simulation studies of the performance of production model estimators have shown that observation-error estimators provide biased and imprecise estimates in the presence of process error (Punt 1989a), which raises the question whether approaches that attempt to consider both observation and process error simultaneously might provide more satisfactory estimates. Apart from the approach of Ludwig et al. (1988), account can be taken when fitting production models of process and observation error by Kalman filter (Reed and Simons 1986; Pella 1993; Kimura et al. 1996) and, more generally, state-space models (de Valpine 2002). The latter two approaches differ from observation-error and total-error estimators in that a distribution for the biomass in year $y + 1$ is determined from a distribution for the biomass in year y by forecasting the distribution of biomass in year $y + 1$ using a population dynamics model and then updating this distribution based on the observed abundance index for year $y + 1$. Given a distribution for the initial biomass, this forecasting–updating process can be applied recursively to compute distributions of biomass for each year of interest. In essence, therefore, unlike observation-error and total-error estimators, the Kalman filter and general state-space models integrate over the unknown population biomasses and (or) process errors.

This paper outlines how production models can be based on each of the four types of estimators referred to above. Previous production model type estimators based on Kalman

Fig. 1. Catch history for northern Namibian hake (*Merluccius capensis*).



filters and general state-space models have tended to be based on population dynamics models that are linear in their parameters (possibly after transformation). The presentation below deals with a broader class of models for which no transformation exists that will make the population dynamics model linear. The paper then evaluates these estimators in terms of their likely bias and precision using simulations based on the 25-year catch history for hake (*Merluccius capensis*) off northern Namibia (Fig. 1).

Materials and methods

The alternative estimators

The following sections outline production model estimators based on each of the four frameworks (observation error, total error, Kalman filter, and general state-space model) referred to above. Readers interested in a more general comparison of the four frameworks can consult de Valpine (2002).

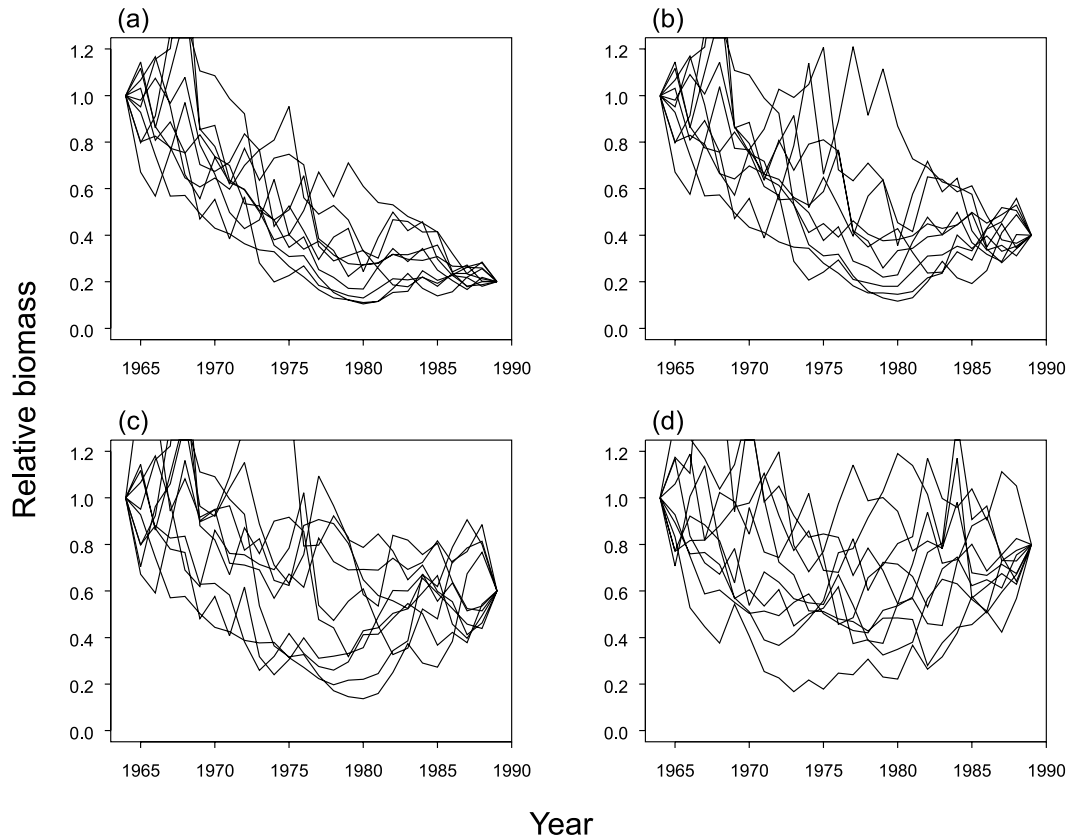
The four estimators are all based on the population dynamics model and observation model given in eq. 1, where the function $g(B)$ is taken to be the Schaefer (or logistic) form of the production function, i.e., $g(B) = rB(1 - B/K)$, where r is the intrinsic growth rate and K is pre-exploitation population size (or carrying capacity). Applying these estimators to other forms for the production function or to delay-difference models (Deriso 1980; Schnute 1985) would involve straightforward modifications to the equations provided below.

For ease of comparison, it is assumed by all of the estimators that the population was at its pre-exploitation level at the start of the first year for which catches are available, i.e., $B_1 = K$. The parameters of interest are r , K , MSY, the ratio of the exploitation rate in the last year for which catches are available to the exploitation rate at which MSY is achieved, $F_n/F_{MSY} = (C_n/B_n)/(r/2)$, the biomass at the end of the modeled time horizon, B_{n+1} , and the relative biomass at the end of the modeled time horizon, $D_{n+1} = B_{n+1}/K$.

For all of the estimators, the likelihood of the abundance-index datum for year y , given the biomass for year y , is

Table 1. The specifications of the simulations trials.

Scenario	r	D_{n+1}	σ_ϵ	σ_η
Baseline	0.35	0.2, 0.4, 0.6, 0.8	0.2	0.2
No process error	0.35	0.2, 0.4, 0.6, 0.8	0.0	0.283
No observation error	0.35	0.2, 0.4, 0.6, 0.8	0.283	0.0
Alternative r values	0.15, 0.25, 0.45	0.4	0.2	0.2

Fig. 2. Time trajectories of “true” population size expressed relative to the pre-exploitation level for four choices for the 1989 relative biomass: (a) 0.2, (b) 0.4, (c) 0.6, and (d) 0.8.

$$(2) \quad L(I_y/q, B_y, \sigma_\eta) = \frac{1}{\sqrt{2\pi}\sigma_\eta I_y} e^{-\frac{(\ln I_y - \ln(qB_y))^2}{2\sigma_\eta^2}}$$

The observation-error estimator

The estimates of r , q , and K for this estimator are obtained by minimizing the negative of the sum of the logarithms of eq. 2 over all years for which abundance-index information is available, where the biomass time series is obtained by projecting from $B_1 = K$ using eq. 1a under the assumption of deterministic dynamics, i.e., $\{\epsilon_y = 0: y = 1, 2, \dots, n\}$.

The total-error estimator

The estimates of r , q , and K for this estimator are obtained by minimizing a negative log-likelihood function that includes contributions from the abundance indices and the process errors (which are treated as nuisance parameters when applying this estimator, i.e., the parameter vector for this estimator is $\{r, q, K, [\epsilon_y: y = 1, 2, \dots, n], \sigma_\epsilon, \sigma_\eta\}$):

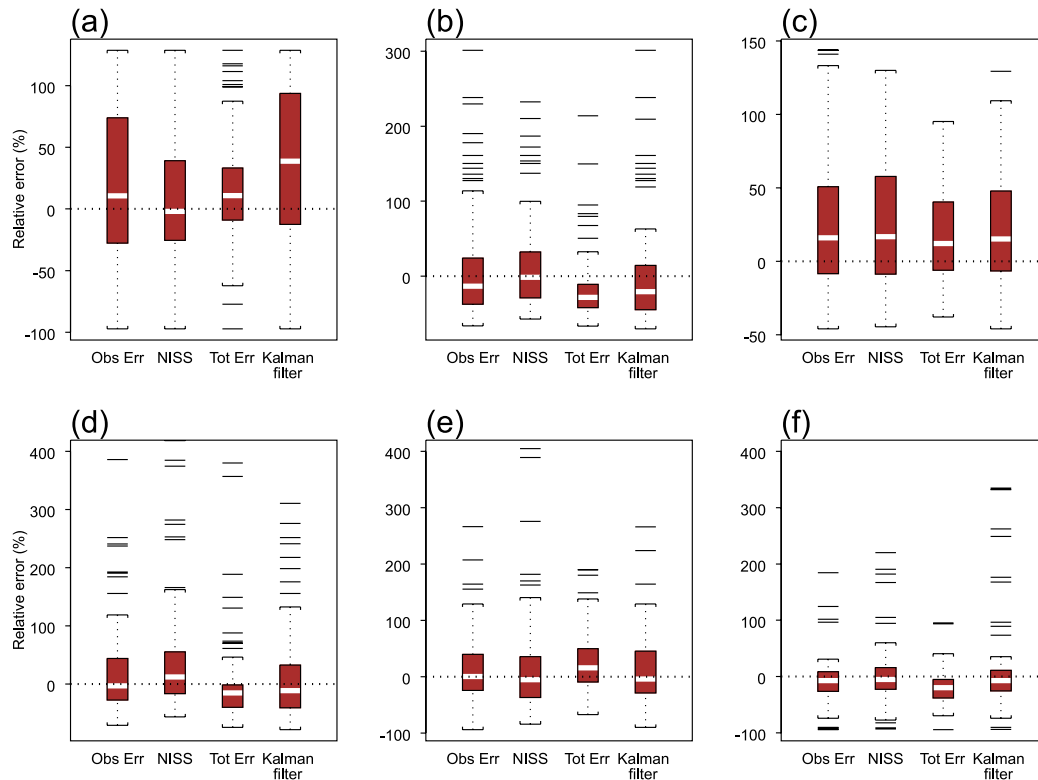
$$(3) \quad -\ln L \propto n \ln \sigma_\eta + \frac{1}{2\sigma_\eta^2} \sum_y (\ln I_y - \ln(qB_y))^2 + n \ln \sigma_\epsilon + \frac{1}{2\sigma_\epsilon^2} \sum_y \epsilon_y^2$$

It is necessary to specify the relationship between σ_ϵ and σ_η to apply this estimator. For the purposes of this paper, $\sigma_\epsilon = \sigma_\eta$. This assumption is arbitrary, but the results of applications of total-error production models to real data sets are seldom very sensitive to the ratio of σ_ϵ to σ_η unless this ratio strongly favors observation error or process error (Ludwig et al. 1988).

General state-space model

The general state-space model considered in this paper is implemented using the numerical integration state-space (NISS) approach of de Valpine (2002) and de Valpine and Hastings (2002). Unlike the total-error estimator, the NISS estimator and the Kalman filter estimator outlined below

Fig. 3. Relative error distributions for (a) intrinsic growth rate, (b) carrying capacity, (c) relative biomass in 1989, (d) biomass in 1989, (e) F_n/F_{MSY} , and (f) MSY (maximum sustainable yield) for four alternative estimators (observation error (Obs Err), NISS (numerical integration state-space), total error (Tot Err), Kalman filter).



integrate over the process errors, as is appropriate given that the process errors are random effects in a nonlinear time-series model. Note that even though the process errors are integrated out, this is not a Bayesian estimator.

For this approach, therefore, the quantity maximized to determine the estimates for the model parameters and outputs (ignoring the conditioning on the model parameters) is

$$(4) \quad L(\mathbf{I}^{(n)}) = P(I_1) \prod_{y>1} P(I_y | \mathbf{I}^{(y-1)})$$

where $\mathbf{I}^{(y)}$ is the vector (I_1, I_2, \dots, I_y) .

The y th factor in the likelihood $P(I_y | \mathbf{I}^{(y-1)})$ is computed using the formula

$$(5) \quad P(I_y | \mathbf{I}^{(y-1)}) = \int \frac{1}{\sqrt{2\pi}\sigma_\eta I_y} e^{-\frac{(\ln I_y - \ln(gB_y))^2}{2\sigma_\eta^2}} P(B_y | \mathbf{I}^{(y-1)}) dB_y \\ = \int P(I_y | B_y) P(B_y | \mathbf{I}^{(y-1)}) dB_y$$

The distribution for the biomass at the start of year y , conditional on the data for years $y-1$ and earlier, $P(B_y | \mathbf{I}^{(y-1)})$, is updated from the distribution for the biomass at the start of year $y-1$ according to the “filtering” equation:

$$(6) \quad P(B_y | \mathbf{I}^{(y-1)}) = \int P(B_y | B_{y-1}) P(B_{y-1} | \mathbf{I}^{(y-1)}) dB_{y-1}$$

where $P(B_{y-1} | \mathbf{I}^{(y-1)}) \propto P(I_{y-1} | B_{y-1}) P(B_{y-1} | \mathbf{I}^{(y-2)})$ so that

$$(7) \quad P(B_y | \mathbf{I}^{(y-1)}) =$$

$$\int P(B_y | B_{y-1}) \frac{P(I_{y-1} | B_{y-1}) P(B_{y-1} | \mathbf{I}^{(y-2)})}{P(I_{y-1} | \mathbf{I}^{(y-2)})} dB_{y-1}$$

and $P(B_y | B_{y-1})$ is the probability

$$(8) \quad P(B_y | B_{y-1}) = \frac{1}{\sqrt{2\pi}\sigma_\epsilon} e^{-\frac{[\ln B_y - \ln(B_{y-1} + g(B_{y-1}) - C_y)]^2}{2\sigma_\epsilon^2}}$$

The integrations (eqs. 5 and 7) cannot be evaluated analytically. Instead they are evaluated numerically by discretizing the range for the annual biomass.

The initial distribution of biomass, $P(B_1)$, is obtained by projecting an arbitrary probability distribution forward with zero catches until it reaches equilibrium. Determining the initial distribution of biomass therefore involves conducting integrals (eqs. 5 and 7) with $C_y = 0$ numerically until convergence occurs.

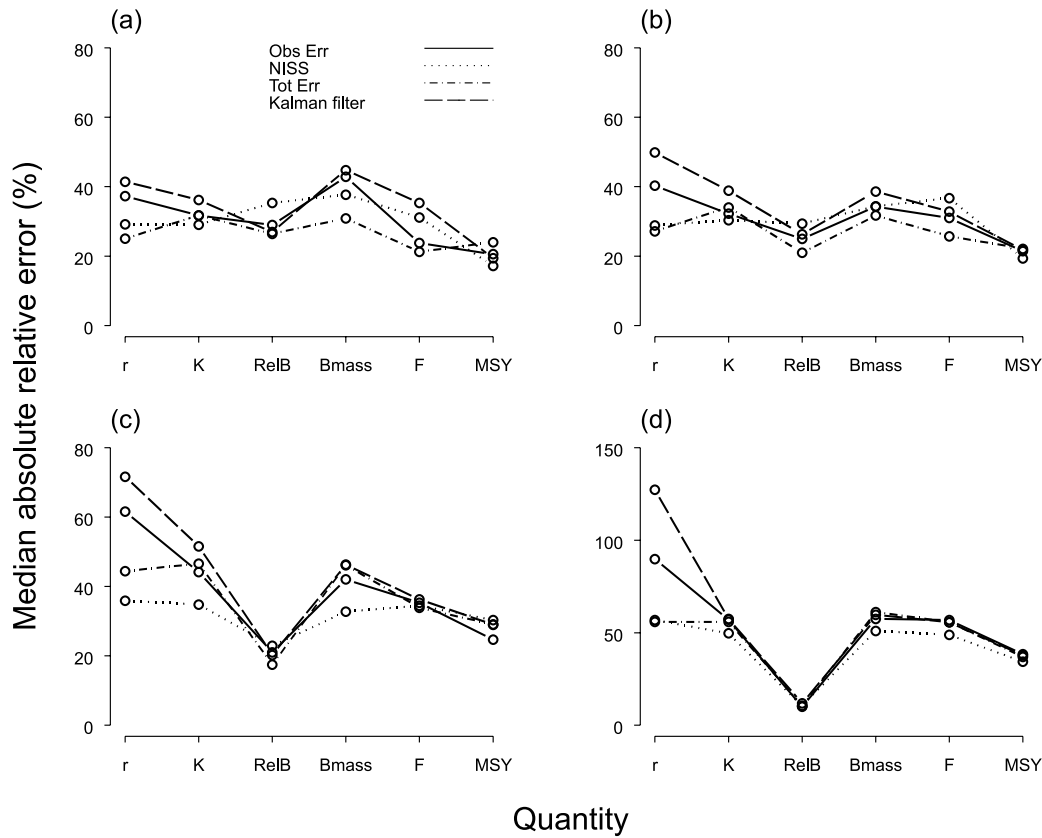
The expected biomass at the start of year y , $E(B_y)$, is computed using the equation

$$(9) \quad E(B_y) = \int B_y P(B_y | \mathbf{I}^{(y-1)}) dB_y$$

The Kalman filter

The traditional Kalman filter (e.g., Meinhold and Singpurwalla 1983; Harvey 1989) is a special case of the general state-space approach outlined above in which the state equation (i.e., the population dynamics model) is linear in its parameters. However, the Schaefer production model is nonlinear so it is necessary to apply an extended Kalman

Fig. 4. Median absolute relative errors for four estimators (observation error (Obs Err), NISS (numerical integration state-space), total error (Tot Err), Kalman filter) and six model outputs (r , K , relative biomass (RelB), biomass (Bmass), F_n/F_{MSY} (F), MSY) for variants of the baseline trial in which the “true” 1989 relative biomass is modified. Relative biomass values depicted are (a) 0.2, (b) 0.4, (c) 0.6, and (d) 0.8.



filter approach in which the nonlinear dynamics model is linearized (Pella 1993; Dereksdóttir and Magnusson 2003). In the case of a production model, the Kalman filter used here updates the expected value and variance of the logarithm of population biomass, i.e., $E(\ln B_y)$ and $\text{Var}(\ln B_y)$ from the expected value and variance of the logarithm of the population biomass at the start of the previous year and the information contained in the abundance index for year $y - 1$. The initial state, year $y = 1$, corresponds to a population distributed about its pre-exploitation level, i.e.,

$$(10) \quad E(\ln B_1) = \ln K \text{ and } \text{Var}(\ln B_1) = \frac{\sigma_\epsilon^2}{1 - (1 - r)^2}$$

The equations to update the expected value and variance of the logarithm of the population biomass are (Dereksdóttir and Magnusson 2003)

$$(11a) \quad E(\ln B_{y+1}) = \ln \tilde{B}_{y+1} + K_y [\ln(I_y/q) - \ln \tilde{B}_{y+1}]$$

$$(11b) \quad \text{Var}(\ln B_{y+1}) = F_y^2 (1 - K_y) \text{Var}(\ln B_y) + \sigma_\epsilon^2$$

where K_y is the Kalman gain factor for year y :

$$(12) \quad K_y = \frac{\text{Var}(\ln B_y)}{\text{Var}(\ln B_y) + \sigma_\eta^2}$$

\tilde{B}_{y+1} is the biomass at the start of year $y + 1$ based on the deterministic component of the population dynamics model, i.e., the result of replacing B_y by $\exp(E(\ln B_y))$ in eq. 1a, and

$$(13) \quad F_y = \frac{\partial(\ln \tilde{B}_{y+1})}{\partial \ln B_y} = \frac{\exp(E(\ln B_y)) + r \exp(E(\ln B_y))(1 - 2 \exp(E(\ln B_y))/K)}{\tilde{B}_{y+1}}$$

It is straightforward to show that the initial state (eq. 10) is the solution to eqs. 11a and 11b when $\sigma_\eta^2 \rightarrow \infty$.

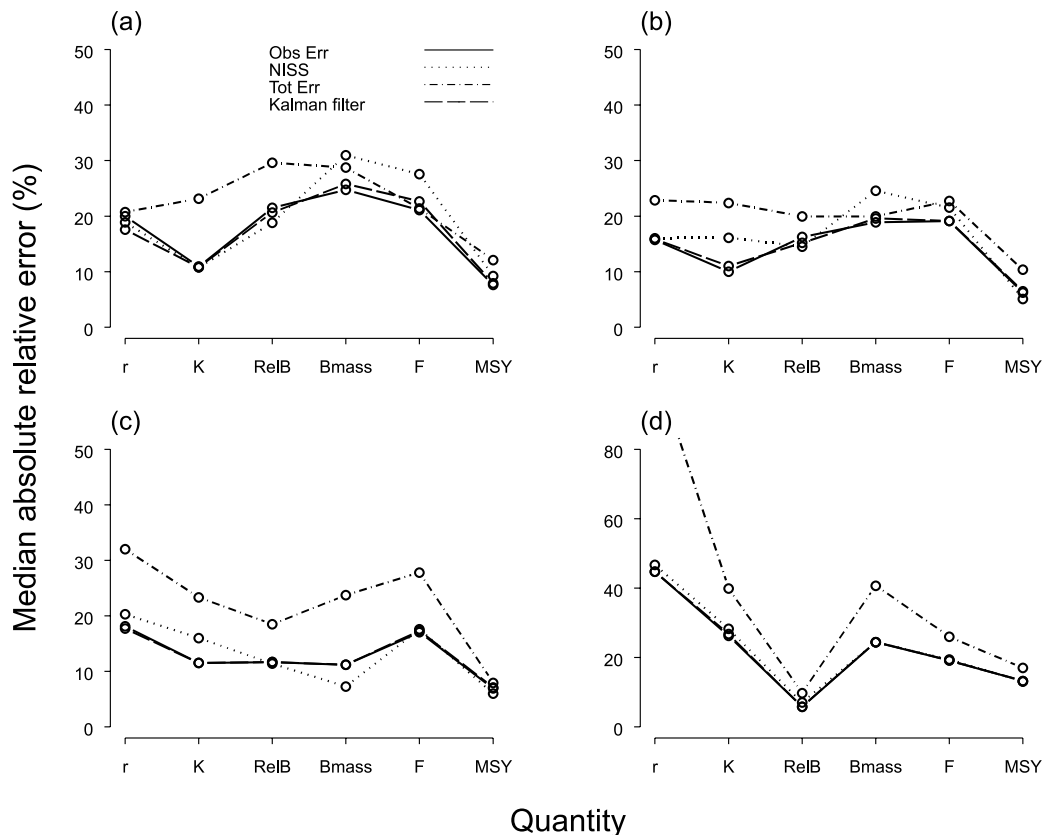
The function maximized to obtain the estimates for the model parameters and outputs is

$$(14) \quad L(I^{(n)}) = \prod_y \frac{1}{\sqrt{2\pi(\text{Var}(\ln B_y) + \sigma_\eta^2)} I_y} \times \exp\left(-\frac{(\ln(I_y/q) - E(\ln B_y))^2}{2(\text{Var}(\ln B_y) + \sigma_\eta^2)}\right)$$

Evaluating the alternative estimators

In common with several previous studies that have evaluated the performances of production model estimators (e.g., Polacheck et al. 1993; Kimura et al. 1996), the four estimators outlined above are evaluated by means of Monte Carlo simulation. Each simulation trial consists of 100 simulations (results with a larger number of simulations were not qualitatively different from those based on 100 simulations). The “true” dynamics of the population are assumed to be gov-

Fig. 5. Median absolute relative errors for four estimators (observation error (Obs Err), NISS (numerical integration state-space), total error (Tot Err), Kalman filter) and six model outputs (r , K , relative biomass (RelB), biomass (Bmass), F_n/F_{MSY} (F), MSY) for simulation trials in which there is no process error. Relative biomass values depicted are (a) 0.2, (b) 0.4, (c) 0.6, and (d) 0.8.



erned by eq. 1a and the historical sequence of catches is taken to be that for hake off northern Namibia (Fig. 1). In general, when evaluating estimators by means of simulation, consideration is given to models of the “truth” that are more complicated than those underlying the estimators being evaluated (e.g., Punt 1989a; Punt et al. 2002). However, for the purposes of this study, the “true” dynamics are taken to be the same as those underlying the four estimators so that any differences between the estimated and true values for the model parameters and outputs are due to the impact of the estimator rather than, say, model misspecification error.

Although the parameters of the Schaefer model are r and K , the “truth” for the simulation trials of this paper is based on specifying r and the relative biomass at the start of year $n + 1$, D_{n+1} , and solving for K . As the result, the initial relative biomass and that for year $n + 1$ are the same for all of the simulations, but because of the differences in process errors, the relative biomasses for the intervening years and the value of K differ among simulations. Alternative scenarios for the simulation experiment were created by varying the values for r , D_{n+1} , σ_ϵ , and σ_η (Table 1).

The results are summarized by the distribution for the relative error, i.e., for a quantity X :

$$(15) \quad R_i^X = \frac{X_i^{\text{True}} - X_i^{\text{Est}}}{X_i^{\text{True}}}$$

where R_i^X is the relative error for quantity X and simulation i ,

X_i^{True} is the “true” value of quantity X for simulation i , and X_i^{Est} is the estimate of quantity X for simulation i .

Results and discussion

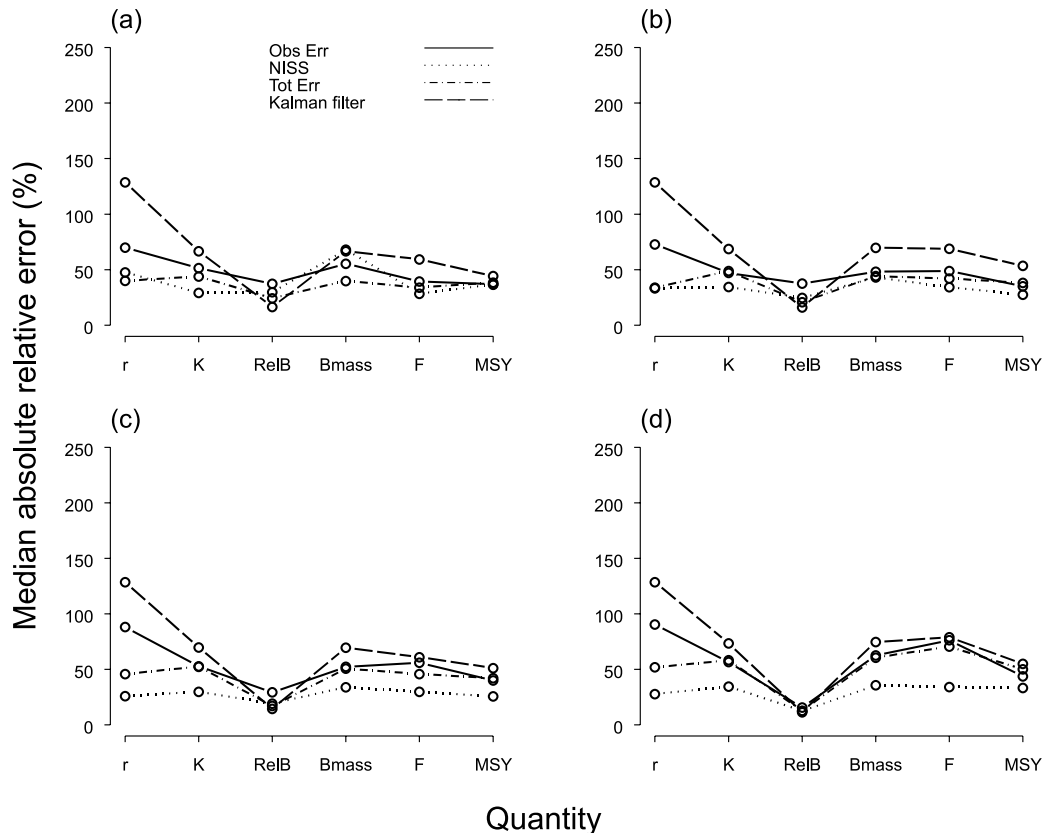
The simulation scenarios

Ten time trajectories of “true” population biomass (expressed relative to the pre-exploitation equilibrium level (K)) are shown for each of the four baseline simulation trials (Fig. 2). The initial relative biomass (year 1964) is always 1, whereas the 1989 relative biomass differs among simulation trials but is the same for each simulation within a simulation trial. The information content of the “true” biomass trajectory is clearly determined to a large extent by the pre-specified value for D_{n+1} , but there is also (substantial) intersimulation variability in the time trajectories of biomass due to the impact of process error. The time trajectories for the abundance index (not shown here) exhibit even greater variability due to the impact of measurement error.

Detailed results for a single simulation trial

The relative error distributions for the baseline trial in which the 1989 relative biomass equals 0.4 are plotted (Fig. 3). The “truth” underlying this simulation trial most closely mimics the results of the actual production model based assessments of Namibian hake in the late 1980s (e.g., Punt 1989b). However, those assessments were highly un-

Fig. 6. Median absolute relative errors for four estimators (observation error (Obs Err), NISS (numerical integration state-space), total error (Tot Err), Kalman filter) and six model outputs (r , K , relative biomass (RelB), biomass (Bmass), F_n/F_{MSY} (F), MSY) for simulation trials in which there is no observation error. Relative biomass values depicted are (a) 0.2, (b) 0.4, (c) 0.6, and (d) 0.8.



certain and so the sensitivity of the performance of the estimators to a variety of alternative scenarios needs to be examined.

The possibility of extreme estimates (relative errors >100%) is much greater for r (21% of simulations for the Kalman filter estimator), 1989 biomass, and F_n/F_{MSY} than for the 1989 relative biomass, K , and MSY. In terms of minimizing bias, the NISS estimator performs best for r , F_n/F_{MSY} , and MSY, whereas the total-error estimator performs best for 1989 biomass and 1989 relative biomass. The Kalman filter estimator provided the least biased estimates of K . The estimates provided by the total-error estimator exhibit the least inter-simulation variability, whereas those provided by the observation-error estimator vary the most among simulations.

It is clearly necessary to combine the effects of the bias and the imprecision of the estimates of the model outputs in any statistic for comparing overall performance among estimators. The root-mean-square-error (RMSE) is such a statistic. However, the RMSE is very sensitive to outliers. Therefore, a more robust performance statistic, the median of the absolute values of the relative errors (MARE), forms the basis for the comparisons among the four estimators in this paper.

For the baseline simulation trial (Fig. 4b), in terms of the MARE, the total-error estimator performs “best” for r , F_n/F_{MSY} , 1989 biomass, and 1989 relative biomass, whereas the NISS estimator performs best for MSY and K . The observation-error estimator performs better than the Kalman filter for all of the model outputs except MSY in terms of the MARE. Although the NISS and total-error estimators

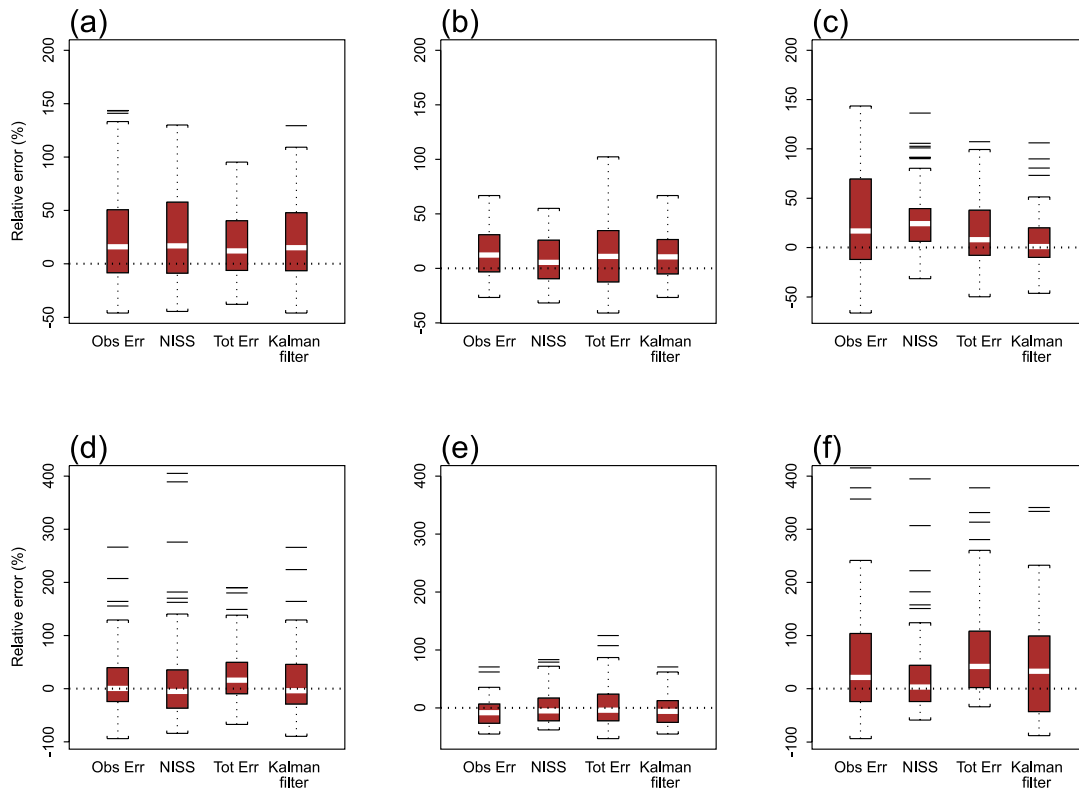
perform best, their estimates of absolute biomass are still relatively poor (MAREs of 32% and 34%, respectively). This result is, however, not unexpected given the outcomes from previous investigations into the performance of production model based stock assessment methods (e.g., Punt and Butterworth 1993). One of the implications of the uniformly poor performance of the estimators in terms of their ability to estimate absolute biomass is that were catch limits to be based on the estimates of biomass from any of these estimators, it would be appropriate to apply a precautionary reduction factor to reduce the probability of catch limits exceeding sustainable levels.

The estimates of 1989 biomass from the Kalman filter are negatively biased (Fig. 3d). One possible reason for this is that the estimate of the biomass in year y is taken to be the median of its distribution (i.e., $\hat{B}_y = \exp(E(\ln B_y))$) which is less than the mean of its distribution ($\hat{B}_y = \exp(E(\ln B_y) + 0.5\text{Var}(\ln B_y))$). However, using the means of the distributions for biomass rather than medians did not reduce the extent of bias for 1989 biomass noticeably.

Sensitivity to alternative “true” model specifications

The MARE for each estimator and model output for various “true” levels for the 1989 relative biomass can be used to examine sensitivity to the value of this parameter (Fig. 4). The results for a 1989 relative biomass of 0.2 are qualitatively the same as those for a 1989 relative biomass of 0.4 in terms of the best-performing estimator. However, the NISS

Fig. 7. Distributions of relative error for relative biomass in 1989 (upper panels) and F_n/F_{MSY} (lower panels) for trials in which (a and d) $\sigma_\epsilon = \sigma_\eta = 0.2$, (b and e) $\sigma_\epsilon = 0$, and (c and f) $\sigma_\eta = 0.0$ for four estimators (observation error (Obs Err), NISS (numerical integration state-space), total error (Tot Err), Kalman filter).



estimator outperforms the other three estimators for all but the 1989 relative biomass for “true” relative biomasses of 0.5 and larger. The size of the MARE increases as the “true” 1989 relative biomass increases. This is, however, not unexpected because higher 1989 relative biomasses correspond to lesser “contrast” in the data (Fig. 2), which, in turn, reduces the ability to adequately estimate quantities of management interest (Hilborn 1979). In contrast to when the 1989 relative biomass is 0.4 (Fig. 3), the estimates can be highly biased when the 1989 relative biomass is high; r tends to be positively biased, whereas K and 1989 biomass are negatively biased.

The observation-error estimator is based on the assumption that the amount of process error is negligible compared with the amount of observation error. Therefore, the performance of this estimator relative to those of the other estimators should be expected to improve markedly for simulation trials in which there are no process errors. As expected, the observation-error estimator generally does perform best when there is only observation error (Fig. 5). However, the performances of the Kalman filter and NISS estimators are virtually identical to that of the observation-error estimator. This occurs because the parameters related to the amounts of process and observation error are estimable parameters for these latter two estimators. In contrast, the performance of the total-error estimator is generally poorer (particularly for a 1989 relative biomass of 0.8). It should be noted that the MARE is lower in Fig. 5 than in Fig. 4 for similar scenarios, confirming the expectation from previous studies that, irrespective of the choice of estimation framework, increased

process error leads to poorer performance of production-model estimators.

The NISS estimator is consistently the best performing of the four estimators when there is only process error (Fig. 6). The performance of the Kalman filter in this case is consistently poorest for quantities other than 1989 relative biomass. Although the NISS estimator performs best in Fig. 6, its performance is, however, very poor when compared with Figs. 4 and 5.

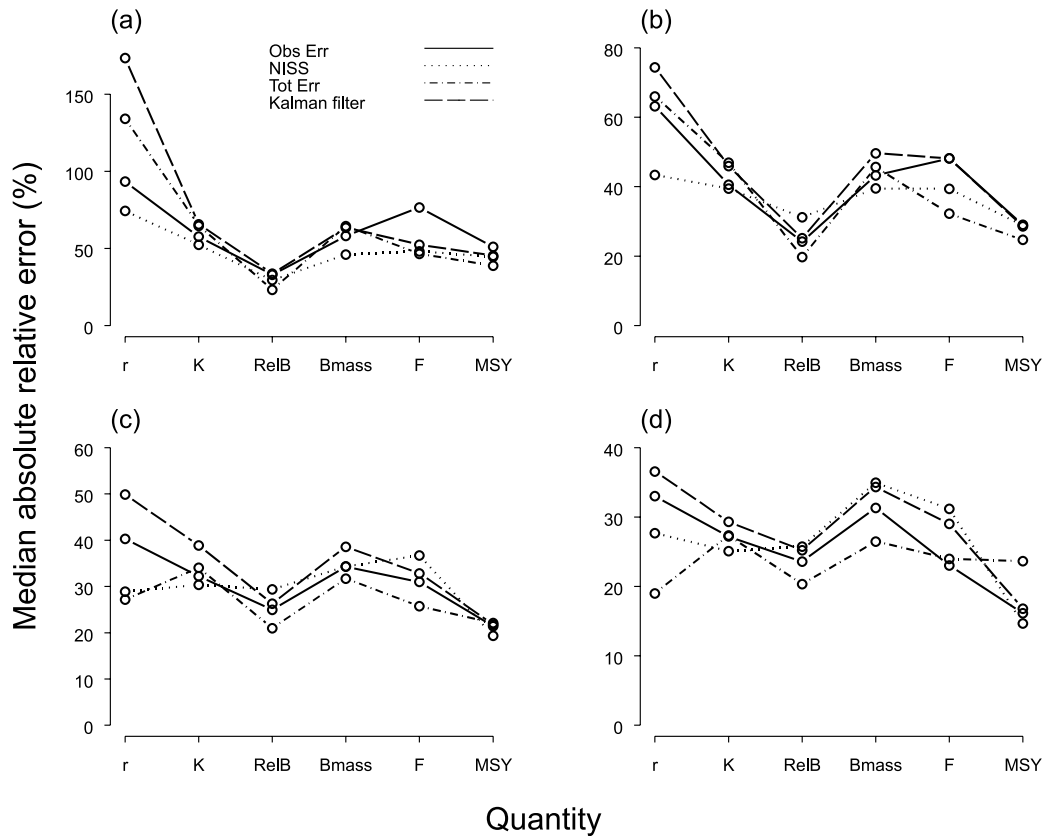
The impact of the magnitude of process and observation error can be seen more clearly from the relative error distributions for 1989 relative biomass and F_n/F_{MSY} (Fig. 7). The estimates are less biased and more precise (leading to lower MAREs) when there is only observation error (Figs. 7b and 7e), and the chances of extreme estimates are markedly lower for this case. The estimates of 1989 relative biomass from the NISS estimator are positively biased when there is only process error (Fig. 7c), but the estimates from this estimator are also much more precise, which leads to the NISS estimator having the lowest overall MARE.

The relative performances of the different estimators are sensitive to the “true” value for the intrinsic growth rate (r) (Fig. 8). The NISS estimator performs best for the lower values for r (although performance is generally poorer the lower the “true” value for r) and the total-error estimator performs best for the higher values for r .

Other estimation methods and structural models

The approaches considered in this paper are not Bayesian. It is, however, possible to construct Bayesian production

Fig. 8. Median absolute relative errors for four estimators (observation error (Obs Err), NISS (numerical integration state-space), total error (Tot Err), Kalman filter) and six model outputs (r , K , relative biomass (RelB), biomass (Bmass), F_n/F_{MSY} (F), MSY) for simulation trials in which the “true” value of r is (a) 0.15, (b) 0.25, (c) 0.35, and (d) 0.45. The 1989 relative biomass is 0.4 for all of the simulations.



models (e.g., McAllister and Kirkwood 1998; Meyer and Millar 1999; McAllister et al. 2001). Although the ability of Bayesian estimators to provide more accurate and precise estimates of fisheries management related quantities than frequentist estimators has yet to be evaluated in depth (Nielsen and Lewy (2002) being a notable exception), the ability of Bayesian approaches to incorporate additional information through prior distributions makes this a key issue for future research.

Observation-error and total-error estimators are commonly used to fit age- and size-structured population dynamics models (e.g., Hilborn 1990; Methot 2000), whereas Kalman filter approaches have been applied to fit size-structured population dynamics models (e.g., Sullivan 1992). The ability to apply general state-space modeling techniques to fully age- and (or) size-structured fisheries models requires the development of more efficient numerical methods for computing the likelihood function than that used for the NISS estimator. This is because the computational requirements of eqs. 5–8 become prohibitive when there are many state variables. The use of Monte Carlo methods (e.g., Carlin et al. 1992; Shumway and Stoffer 2000) is therefore probably necessary to implement such models.

Choices among estimators

The four estimators considered in this paper are all candidates for use in actual stock assessments. All four estimators

can make use of multiple time series of abundance indices, as well as series of absolute estimates of abundance. None of the estimators uniformly outperformed all of the others for all of the scenarios considered. However, overall, the NISS estimator is probably best as it is never the worst estimator and is often the best. The total-error estimator appears to be the next-best estimator. Although this points to the use of the NISS estimator in future production-model assessments, the very poor performance of this estimator in some of the simulation trials highlights that even the best-performing production model estimators may provide estimates of quantities of management interest that are severely biased and highly imprecise. It should also be recalled that the “true” model for the simulations was the same as that underlying the estimators. The level of error would have been even greater had a “true” model differing from the Schaefer model been used to represent reality.

One of the key assumptions underlying the simulations is that the biomass at the start of the first year for which catches are available equals K , i.e., $B_1 = K$, and that the estimators are aware of this. Although this assumption is unlikely to be greatly violated for hake off northern Namibia (Fig. 1), the sensitivity of the results to this assumption is examined by a simulation trial in which the “true” relative biomass at the start of the first year for which catches are available is set equal to 0.6 and the “true” relative biomass in year $n + 1$ is set equal to 40% of this, i.e., $D_{n+1} = 0.24$.

Fig. 9. Relative error distributions for (a) intrinsic growth rate, (b) carrying capacity, (c) relative biomass in 1989, (d) biomass in 1989, (e) F_n/F_{MSY} , and (f) MSY (maximum sustainable yield) for four alternative estimators (observation error (Obs Err), NISS (numerical integration state-space), total error (Tot Err), Kalman filter) when the relative biomass at the start of 1964 is 0.6 rather than 1.

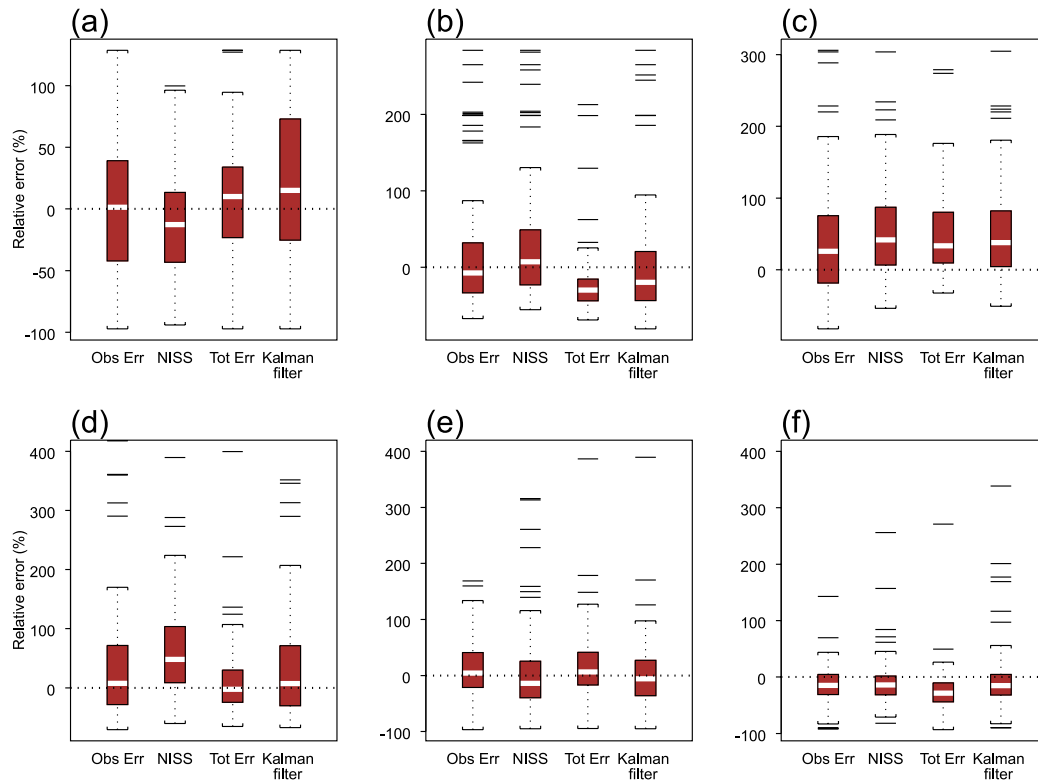
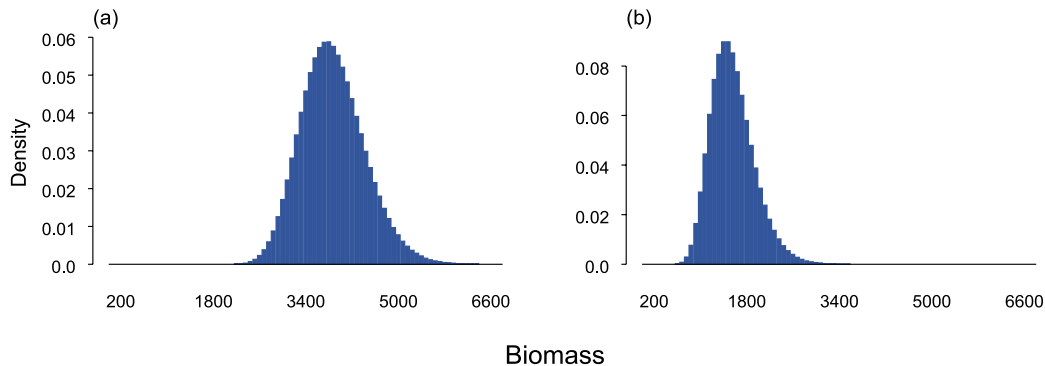


Fig. 10. Distribution of biomass for (a) 1964 and (b) 1989 from an application of the numerical integration state-space (NISS) estimator to one of the simulated data sets.



The relative error distributions for the four estimators when the initial relative biomass is treated as an estimable parameter by all four estimators rather than being assumed equal to 1 are shown (Fig. 9). The estimates of 1989 biomass and 1989 relative biomass are more biased and variable than is the case for the baseline trial. There is also a notable deterioration in the performance of the NISS estimator in terms of its ability to estimate 1989 biomass and F_n/F_{MSY} (Figs. 9d, 9e). The observation-error estimator again outperforms the Kalman filter estimator.

The relatively poor performance of the Kalman filter is perhaps surprising given that it would have been equivalent to the NISS estimator had the population dynamics model been linear. The extended Kalman filter considered in this

paper is based on the assumption that the distribution of biomass is lognormal. However, the distributions of biomass from the NISS approach are not as skewed as might have been anticipated had biomass been lognormally distributed (see Fig. 10 for examples of biomass distributions from the NISS estimator), suggesting that, although beyond the scope of the present paper, a different formulation of the Kalman filter may have performed somewhat better. Another reason for the relatively poor performance of the Kalman filter may lie in the attempt to estimate the observation- and process-error variances and the error associated with linearizing the population dynamics model (see eq. 13).

It should be noted that there are factors other than performance under simulated conditions that need to be considered

when selecting an estimator for use in production-model assessments. In particular, the computational demands of the NISS estimator are substantial compared with those of the other three estimators. Also, this estimator involves approximating integrals numerically. The accuracy of the approximations depends on how fine the numerical grid used for integration is. However, there is a tradeoff between increased accuracy of integration and the time demands of the calculations.

Acknowledgements

This work was funded through NMFS grant NA07FE0473. The comments by two anonymous reviewers are greatly acknowledged.

References

- 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.
- Cooke, J.G., and Beddington, J.R. 1984. The relationship between catch rates and abundance in fisheries. *IMA J. Math. Appl. Med. Biol.* **1**: 391–405.
- Dereksdóttir, E., and Magnusson, K.G. 2003. A strike limit algorithm based on adaptive Kalman filtering with an application to aboriginal whaling of bowhead whales. *J. Cetacean Res. Manag.* **5**: 29–37.
- Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structured model. *Can. J. Fish. Aquat. Sci.* **37**: 268–282.
- de Valpine, P. 2002. Review of methods for fitting time-series models with process and observation error and likelihood calculations for nonlinear, non-gaussian state-space models. *Bull. Mar. Sci.* **70**: 455–471.
- de Valpine, P., and Hastings, A. 2002. Fitting population models with process noise and observation error. *Ecol. Monogr.* **72**: 57–76.
- Fox, W.W. 1970. An exponential surplus-yield model for optimizing exploited fish populations. *Trans. Am. Fish. Soc.* **99**: 80–88.
- Fox, W.W. 1975. Fitting the generalized stock production model by least-squares and equilibrium approximation. *Fish. Bull.* **73**: 23–37.
- Gulland, J.A. 1961. Fishing and the stocks of fish at Iceland. *Fish. Invest. Min. Agric. Fish. Food U.K. (Ser. 2)*, **23**: 1–52.
- Harvey, A.C. 1989. Forecasting, structural time series models, and the Kalman filter. Cambridge University Press, Cambridge.
- Hilborn, R. 1979. Comparison of fisheries control systems that utilize catch and effort data. *J. Fish. Res. Board Can.* **36**: 1477–1489.
- Hilborn, R. 1990. Estimating the parameters of full age-structured models from catch and abundance data. *Bull. Int. North Pac. Fish. Comm.* **50**: 207–213.
- Hilborn, R. 2001. Calculation of biomass trend, exploitation rate, and surplus production from survey and catch data. *Can. J. Fish. Aquat. Sci.* **58**: 579–584.
- International Commission for the Conservation of Atlantic Tunas. 2000. Report of the Standing Committee on Research and Statistics (SCRS). International Commission for the Conservation of Atlantic Tunas (ICCAT), Madrid, Spain.
- 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.
- Ludwig, D., Walters, C.J., and Cooke, J.G. 1988. Comparison of two models and two estimation methods for catch and effort data. *Nat. Resour. Model.* **2**: 457–498.
- McAllister, M.K., and Kirkwood, G.P. 1998. Bayesian stock assessment: a review and example application using the logistic model. *ICES. J. Mar. Sci.* **55**: 1031–1060.
- McAllister, M.K., Pikitch, E.K., and Babcock, E.A. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. *Can. J. Fish. Aquat. Sci.* **58**: 1871–1890.
- Meinhold, R.J., and Singpurwalla, N.D. 1983. Understanding the Kalman filter. *Am. Stat.* **37**: 123–127.
- Methot, R.D. 2000. Technical description of the Stock Synthesis Assessment Program. NOAA Tech Memo NMFS-NWFSC-43.
- Meyer, R., and Millar, R.B. 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. *Can. J. Fish. Aquat. Sci.* **56**: 37–52.
- Moses, B.S. 2000. A review of artisanal marine and brackishwater fisheries of south-eastern Nigeria. *Fish. Res.* **47**: 81–92.
- Nielsen, A., and Lewy, P. 2002. Comparison of the frequentist properties of Bayes and maximum likelihood estimators in an age-structured fish stock assessment model. *Can. J. Fish. Aquat. Sci.* **59**: 136–143.
- Pella, J.J. 1993. Utility of 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 Prog. Rep. No. 92-02, University of Alaska, Fairbanks. pp. 571–593.
- Pella, J.J., and Tomlinson, P.K. 1969. A generalized stock production model. *Inter-Am. Trop. Tuna Comm. Bull.* **13**: 421–458.
- 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.
- Prager, M.H. 2002. Comparison of logistic and generalized surplus-production models applied to swordfish, *Xiphias gladius*, in the North Atlantic Ocean. *Fish. Res.* **58**: 41–57.
- Punt, A.E. 1989a. A further preliminary comparative investigation into the performance of catch-effort model estimation procedures used in the management of the ICSEAF hake stocks. *Coll. Sci. Pap. Int. Comm. SE. Atl. Fish.* **16**: 63–122.
- Punt, A.E. 1989b. $f_{0,1}$ quotas for the ICSEAF hake stocks based on seven model-estimation procedures which perform best in simulation trials. *Coll. Sci. Pap. Int. Comm. SE. Atl. Fish.* **16**: 123–149.
- Punt, A.E., and Butterworth, D.S. 1993. Variance estimates for fisheries assessment: their importance and how best to evaluate them. *In* Risk evaluation and biological reference points for fisheries management. Edited by S.J. Smith, J.J. Hunt, and D. Rivard. *Can. Spec. Publ. Fish. Aquat. Sci.* No. 120. pp. 145–162.
- Punt, A.E., Smith, A.D.M., and Cui, G. 2002. Evaluation of management tools for Australia's South East Fishery. 2. How well do commonly used stock assessment methods perform? *Mar. Freshw. Res.* **53**: 631–644.
- Reed, W.J., and Simons, C.M. 1986. Analyzing catch-effort data by means of the Kalman filter. *Can. J. Fish. Aquat. Sci.* **53**: 2157–2166.
- 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.
- Schnute, J. 1977. Improved estimates from the Schaefer production model: theoretical considerations. *J. Fish. Res. Board Can.* **34**: 583–603.
- Schnute, J.T. 1985. A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sci.* **42**: 414–429.

- Shumway, R.H., and Stoffer, D.S. 2000. Time series analysis and its applications. Springer Verlag, New York.
- Sissenwine, M.P. 1978. Is MSY an adequate foundation for optimum yield? *Fisheries*, **3**: 22–42.
- Sullivan, P.J. 1992. A Kalman filter approach to catch-at-length analysis. *Biometrics*, **48**: 237–257.
- Williams, E.H., and Prager, M.H. 2002. Comparison of equilibrium and nonequilibrium estimators for the generalized production model. *Can. J. Fish. Aquat. Sci.* **59**: 1533–1552.