Fisheries Research 172 (2015) 325–334

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Autocorrelated error in stock assessment estimates: Implications for management strategy evaluation

John Wiedenmann[a,*], Michael J. Wilberg[b], Andrea Sylvia[b,1], Thomas J. Miller[b]

[a] Department of Ecology, Evolution and Natural Resources, Rutgers University, United States
[b] Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, United States

ARTICLE INFO

Article history:
Received 22 April 2015
Received in revised form 28 July 2015
Accepted 29 July 2015
Available online 24 August 2015

Keywords:
Management strategy evaluation
Stock assessment error
Lag-1 autocorrelation

ABSTRACT

Management strategy evaluation (MSE) is often used in fisheries science to evaluate the effects of different management options. MSE models typically include a stock assessment component to estimate population size and management reference points based on data generated using the model, but including a full assessment within an MSE can be computationally intensive. A commonly used alternative to the full assessment approach is to simulate the error about the stock assessment outcomes as a stochastic process with an assumed level of autocorrelated estimation error. There is little guidance on what might be a reasonable assumed amount of autocorrelation, and what factors might influence this amount. We estimate the amount of temporal autocorrelation in errors of estimated biomass and recruitment from statistical catch at age stock assessment models over a series of scenarios spanning life histories, exploitation levels, recruitment variability, and data quality. Autocorrelation in the error in biomass estimates (ϕ_S) was positive and relatively high, with median estimates ranging between 0.7 and 0.9. Estimates were highest for the long-lived life history and lowest for the short-lived life history. Exploitation level also affected the amount of autocorrelation, with higher values for lightly exploited populations. On average, however, estimates of ϕ_S did not change over time as more data were included in the assessment, and were independent of whether or not a harvest policy was applied. Recruitment variability and data quality had relatively minor effects on autocorrelation of errors.

1. Introduction

Simulation modeling is often used in fisheries science to evaluate the effects that management decisions have on a resource (e.g. a population, assemblage, or community) and on stakeholders (Milner-Gulland et al., 2010). This class of simulation models is referred to as management strategy evaluation (MSE) or the management procedure (MP) approach (Butterworth and Punt, 1999). MSE has become a widely used tool to aid fisheries managers in variety of areas. For example, MSE has been used to identify robust harvest control rules in both data-rich (Punt et al., 2008; A'Mar et al., 2009) and data-poor (Wiedenmann et al., 2013; Carruthers et al., 2014) situations, as well as for selecting effective regulations for controlling recreational harvests (Miller et al., 2010).

An MSE typically has three components, an operating model, an assessment model, and a management model, and these components are designed to mimic the resource dynamics and scientific assessment process, and how these interact with the management options being tested (Punt et al., 2015). In the operating model, the population of interest is projected through time, and the true status is known. Data are generated using the operating model based on the true state and an observation process, usually with some level of observation error. These data are then used in the assessment model to estimate population status. The estimated status derived from the assessment model informs the management model and is used in conjunction with a harvest strategy to determine the total allowable catch and possibly the regulations to achieve that catch. The catch is then removed from the population in the following time step, and this loop is repeated for several years and model iterations to account for uncertainty in the population, assessment and management dynamics.

For the stock assessment portion of the MSE, two approaches are typically used, termed the "full" or "stochastic process" approaches (ICES, 2013). The full approach implements a complete stock

---

* Corresponding author.
E-mail address: john.wiedenmann@gmail.com (J. Wiedenmann).

[1] Present address: Department of Natural Resource Ecology and Management, Iowa State University, United States.

http://dx.doi.org/10.1016/j.fishres.2015.07.037
0165-7836/

assessment model, such as a statistical catch at age model (SCAA) that estimates a suite of parameters and produces estimates over the entire time series of data availability (e.g., A'Mar et al., 2009; Punt, 2003). Depending upon the assessment model being used and the data being generated, many parameters may be estimated (100+), which can lead to lengthy run times. For example, most SCAA models require numerical solutions and search over the parameter space to find the best parameter estimates, often requiring hundreds or thousands of iterations. This can lead to run times to be 100–1000 times longer than an MSE without an integrated assessment model. Additionally, it is not always possible to determine whether an assessment has reached the best solution (i.e., reached the true minimum of the objective function).

A commonly used alternative to the full approach is to simulate the error from the stock assessment as a stochastic process. The stochastic process approach greatly reduces the computation time of the MSE, which allows for a greater exploration of management options and uncertainty scenarios. As one common example of the stochastic process approach, the time series of estimated biomass (S_{est}) is modeled around the true biomass (S) with lag-1 autocorrelated error:

$$S_{est}\left(t\right)=Se^{\epsilon_S(\mathrm{t})-0.5\sigma_S^2}$$

$$\epsilon_S(t)=\phi_S\epsilon_S\left(\mathrm{t}-1\right)+\sqrt{1-\phi_S^2}\varphi_S\left(t\right) \tag{1}$$

$$\phi_s\left(t\right)\sim N\left(0,\sigma_s^2\right)$$

where ϕ_S determines the degree of autocorrelation in the estimates (Punt et al., 2008; Irwin et al., 2008; Wilberg et al., 2008). All the error dynamics are controlled by specifying σ_S^2 and ϕ_S in the stochastic process approach. Drawbacks of the stochastic process approach are that it does not produce the full range of output of an assessment and may not capture complex feedbacks between the state of the system and the variance, bias, and correlation of errors. Additional assumptions are required if additional assessment output is needed in the MSE loop (e.g. fishery selectivity, recruitment time series). For example, Irwin et al. (2008) generated estimates of abundance at age using the same error structure used to estimate biomass, such that an overestimate of biomass of 10% was the result of an overestimate of abundance of 10% in all age classes. The stochastic process approach also requires specifying values for σ_S^2 and ϕ_S. Simulation studies will often explore a range of values for σ_S^2, and either a single value for ϕ_S (e.g. Irwin et al., 2008) or a range of values (e.g. Punt et al., 2008). In general, the assumed values for ϕ_S are positive and high (>0.7). While the assumption of high positive autocorrelation of assessment errors seems reasonable based on the multi-year effects that are produced in age-structured models (e.g., Mohn, 1999), guidance to assist researchers in choosing the appropriate values is lacking.

It would be valuable to have a more formal basis for implementing the stochastic approach in MSEs given the increasing importance of MSE in fisheries management (Milner-Gulland et al., 2010; Punt et al., 2015), and the potential impact the assumed values of ϕ_S and σ_S can have on the results when using the stochastic process approach. We use a simulation model to estimate the degree of temporal autocorrelation and uncertainty in final year (terminal) biomass estimates from a full SCAA assessment model. We also estimate the autocorrelation in terminal recruitment estimates for comparison with biomass estimates. The simulation model was run over a range of species life histories and exploitation intensities to identify potential factors controlling the amount of autocorrelation.

2. Methods

To understand the temporal autocorrelation of errors in stock assessment estimates, the estimated values from a stock assessment model must be compared to the true values. Because the true dynamics (e.g. biomass, recruitment) are unknown for real world systems, we conducted a simulation study in which we simulated the true population dynamics and applied a stock assessment model over a range of scenarios encompassing different life histories, exploitation histories, and levels of data quality. The simulation model was developed in AD Model Builder (Fournier, 2011, and contains three main components defined in detail below and summarized here. The foundation of the simulation is the operating model, which determines the true population dynamics of the stock and how data are generated. Data generated in the operating model are based on the true dynamics within the model with some specified amount of observation error. The operating model generates data on fishery harvests, as well as a fishery-independent index of abundance. These data are then used in the assessment model to estimate stock status and biological reference points. The assessment model is an SCAA model, and output from the assessment is used in the management model to determine a catch limit using a harvest policy (we also explored the effects of not using a harvest policy; see Section 2.2). The catch limit estimated in the management model is removed from the population, without implementation error, and the simulation loop continued for a set number of years. This process is repeated many times for each model specification (e.g. life history) to account for the variability in the population dynamics, data generation, and assessment estimation. At the end of each run, the true and estimated values of biomass and recruitment are used to calculate the amount of autocorrelation in the error in these estimates. Our simulation model evaluated the effects of two management models to determine how management may affect the quality of assessment estimates.

2.1. Operating, assessment, and management models

The population dynamics followed an age-structured model (Quinn and Deriso, 1999) with the equations governing the dynamics in Table 1. Equations used in the model are referenced by their number in Table 1, such that the numerical abundance-at-age is referred to as Eq. (T1.1). The population began at unfished equilibrium abundance at age in year 1 of the simulation. Recruitment was determined from the abundance of that cohort the previous year, decreased by continuous time-varying natural and fishing mortality (Eq. (T1.1)). Recruitment to the population followed the Beverton–Holt stock-recruit relationship, with bias-corrected log-normal stochasticity and autocorrelated deviations (Eq. (T1.2)). Parameters controlling the degree of autocorrelation and variability in recruitment (Table 2) were based on the recruitment meta-analysis of Thorson et al. (2014). Parameters for the Beverton–Holt model were derived from the unfished spawning biomass, unfished recruitment, and the steepness parameter (Eq. (T1.2)), where steepness represents the fraction of unfished recruitment that results when the spawning biomass is reduced to 20% of the unfished level. Total spawning biomass in a given year was calculated by summing the product of the maturity at age, weight at age and abundance at age over all recruited age classes (Eq. (T1.3)). Weight at age was an allometric function of length at age, which followed a von Bertalanffy growth function (Eqs. (T1.5) and (T1.6)). The proportion mature at age was calculated using a logistic function (Eq. (T1.7)).

The model contained a single fishery, with a logistic selectivity function (Eq. (T1.8)). The selectivity ogive varied over time as the parameter that determines the age at 50% selectivity varies annually in an autocorrelated manner (Eq. (T1.8)), although the source

Table 1
Equations governing the dynamics in the operating model.

	Equation	Description
	Population dynamics	
1	$N(a,t)=\begin{cases} R(t) a=a_R \\ N(a-1,t-1)e^{-z(a-1,t-1)} a_R<a<a_{\max} \\ N(a-1,t-1)e^{-z(a-1,t-1)}+ a=a_{\max} \\ N(a,t-1)e^{-z(a,t-1)} \end{cases}$	Numerical abundance at age
2	$R(t)=\frac{S(t-a_R)}{a+\beta S(t-a_R)}e^{\epsilon_R-0.5\sigma_R^2}$ $\alpha=\frac{S_0(1-h)}{4hR_0}\beta=\frac{5h-1}{4hR_0}$ $\epsilon_R(t)=\phi_R\epsilon_R(t-1)+\sqrt{1-\phi_R^2}\varphi_R(t)$ $\varphi_R(t)\sim N\left(0,\sigma_R^2\right)$	Stock-recruit relationship
3	$S(t)=\sum_a m(t)w(a)N(a,t)$	Spawning biomass
4	$Z(a,t)=M(t)+\gamma(a,t)F(t)$ $M(t)=\bar{M}e^{\epsilon_M(t)-0.5_M^2}$ $\epsilon_M(t)=\phi_M\epsilon_M(t-1)+\sqrt{1-\phi_M^2\varphi_M(t)}$ $\varphi_M(t)\sim N\left(0,\sigma_M^2\right)$	Total mortality with time-varying natural mortality
	Life history	
5	$L(a)=L_\infty(1-e^{-k(a-a_0)})$	Length at age
6	$w(a)=bL(a)^c$	Weight at length
7	$m(a)=\frac{1}{1+e^{-\frac{a-m_{50}}{m_{\mathrm{slope}}}}}$	Maturity at age
	Fishing dynamics	
8	$\gamma(\mathrm{a},t)=\frac{1}{1+\mathrm{e}^{-\mathrm{a}-\gamma_{50}(t)/\gamma_{\mathrm{slope}}}}$ $\gamma_{50\%}(t)=\bar{\gamma}_{50\%}e^{\epsilon}\gamma^{(\mathrm{t})-0.5\sigma_\gamma^2}$ $\epsilon_\gamma=\phi_\gamma\epsilon_\gamma(t-1)+\sqrt{1-\phi_\gamma^2}\varphi(t)$ $\varphi(t)\sim N\left(0,\sigma_\gamma^2\right)$	Time-varying selectivity at age in the fishery
9	$C(a,t)=\frac{\gamma(a,t)F(t)}{Z(a,t)}w(a)N(a,t)\left(1-\mathrm{e}^{-z(a,t)}\right)$ $C(t)=\sum_a C(a,t)$	Total catch
	Data-generating dynamics	
10	$C_{\mathrm{obs}}(t)=C(t)^{\epsilon_C(t)-0.5\sigma_C^2}$ $\epsilon_C(t)\sim N\left(0,\sigma_C^2\right)$	Observed catch
11	$g(\mathrm{a})=\frac{1}{1+e^{-\mathrm{a}-\mathrm{g}_{50}/\mathrm{g}}}$	Selectivity at age in the survey
12	$I(\mathrm{a},\mathrm{t})=q(t)g(a)N(\mathrm{a},\mathrm{t})$ $I(t)=\sum_a I(\mathrm{a},\mathrm{t})$ $q(t)=qe^{\epsilon_q(t)-0.5\sigma_q^2}$ $\epsilon_q(t)\sim N\left(0,\sigma_q^2\right)$	True index of abundance
13	$I_{\mathrm{obs}}(t)=I(t)^{\epsilon_I(t)-0.5\sigma_I^2}$ $\epsilon_I(t)\sim N\left(0,\sigma_I^2\right)$	Observed index of abundance
14	$p_{\mathrm{obs}}(t)=\frac{1}{n}\Theta(t)$ $\Theta(t)\sim\mathrm{Multinomial}(n,p(t))$ $p(t)=\frac{1}{I(t)}((Ia_R,t)),\ldots,I(a_{\max},t)$	Observed vector of proportion-at-age in fishery f

for the changes was not modeled explicitly. Because both natural (M) and fishing mortality (F) occurred continuously throughout the year, catch was calculated using the Baranov catch equation (Quinn and Deriso, 1999; Eq. (T1.9)).

Each model run was divided into two periods. The initial period covers 80 years, followed by a subsequent management period covering the next 45 years. For the first 50 years, the population remains unfished but with stochastic recruitment. A single fishery developed in year 50 of the initial period, which was described by a linear increase in fishing mortality (F) until year 65, followed by a constant at the peak fishing mortality for the remainder of the initial period. The intensity of fishing (F = 0.5, 1.0, 2.5 × F_{MSY} for the light, moderate, or heavy exploitation scenarios) at the plateau during this period, along with the pattern of recruitment determined the population abundance at the start of the management period.

At the start of management period (year 81), the population was first assessed using data generated during the initial period, starting in year 60, and with a 1-year lag between the last year of the data collected and when the assessment is conducted. Thus, the estimation model did not include the full fishing history for the stock. Fishery catch data (both total and proportions-at-age) and a fishery-independent survey-derived index of abundance (both total and proportions-at-age) were generated annually. These data were generated by applying observation error to the true values (Eqs. (T1.10)–(T1.14)) using lognormal distributions for the total catch and the index of abundance (with time-varying catchability) and multinomial distributions for the proportions at age. We included two scenarios of coefficient of variation for the total catch and index data, time-varying catchability, and input sample sizes for the proportions at age to explore the interactions between data quality and the autocorrelation in assessment estimates (Table 2).

The time series of catch and survey data were input into the SCAA to estimate abundance at age and fishing mortality rates in each year. The parameters estimated in the SCAA were the initial abundance (associated with the first year of data), recruitments and fishing mortality rates (across years), fishery selectivity parame-

Table 2
Parameters values used in the model. Parameters not associated with life history are presented at the top, with multiple values explored for the "good" and "poor" assessment cases. Life history parameters are presented below. In case where the same value was used across life histories, only a single value is presented. Two values for steepness were explored, but were not tied to the assessment uncertainty runs.

Parameter	Description	Value
σ_R	Standard deviation of stock-recruit relationship	0.77, 1.25
ϕ_R	Autocorrelation in recruitment	0.44
σ_M	Standard deviation of time-varying M	0.15
ϕ_M	Autocorrelation in M	0.3, 0.9
σ_γ	Standard deviation of age at 50% selectivity	0.1
ϕ_γ	Autocorrelation in selectivity	0.3, 0.9
σ_C	Standard deviation of catch estimates	0.15
σ_q	Standard deviation of catchability	0.01, 0.05
σ_I	Standard deviation of survey estimates	0.29, 0.63
E_C	Input sample size of the catch	200, 50
E_I	Input sample size of the survey	200,50

Parameter	Description	Long-lived	Medium-lived	Short-lived
a_R	Age at recruitment (to population; in years)	5	2	1
$a_{\max}$	Maximum age (years)	20	12	7
M	Mean natural mortality rate (year^{-1})	0.1	0.2	0.4
R_0	Virgin recruitment (number or fish)		1×10^6	
h	Steepness		0.45,0.9	
a_0	Age at length = 0		0	
L_∞	Maximum length (cm)		90	
k	Growth rate (year^{-1})	0.07	0.13	0.27
b_1	L – W scalar (kg cm^{-1})		3.0×10^{-6}	
b_2	L – W exponent		3	
m_{50}	Age at 50% maturity (years)	7	3.5	1.75
γ_{50}	Mean age at 50% selectivity in fishery (years)	7	3.5	1.75
g_{50}	Mean age at 50% selectivity in survey (years)	5.3	2.6	1.3
m,γ,and g_{slope}	Slope of maturity, and selectivity (fishery and survey) functions		1	

Table 3
Likelihood functions used in the statistical catch at age analysis (SCAA).

Equation	Description
$\mathscr{L}_{SCAA} = \sum_i \ell(t)$	Full likelihood for SCAA model
$\ell(1) = 0.5n\log\left(\sigma_c^2\right) + \frac{1}{2\sigma_c^2}\sum_t (\log(C_{\text{obs}}(t)) - \log(C_{\text{est}}(t)))^2$	Likelihood component for annual catches
$\ell(2) = 0.5n\log\left(\sigma_I^2\right) + \frac{1}{2\sigma_I^2}\sum_t (\log(I_{\text{obs}}(t)) - \log(I_{\text{est}}(t)))^2$	Likelihood component for the annual index of abundance
$\ell(3) = -E_C \sum_t \sum_a p_{\text{obs},C}(a,t)\log(p_{\text{est},C}(a,t))$	Likelihood component for the annual proportion-at-age in the catch
$\ell(4) = -E_I \sum_t \sum_a p_{\text{obs},I}(a,t)\log(p_{\text{est},I}(a,t))$	Likelihood component for the annual proportion-at-age in the index

ters, and the survey catchability. Parameters were estimated using a maximum likelihood approach with lognormal likelihood functions for the total catch and total index of abundance and multinomial likelihood functions for the proportions at age in the catch and index of abundance (Table 3). The selectivity and survey catchability parameters that varied over time in the operating model were assumed to be constant over time in the SCAA, and natural mortality was assumed to be constant at the true mean value. All other required SCAA inputs (i.e., maturity- and weight-at-age) were set to the true values specified in the operating model. The SCAA also estimated the spawning potential ratio (SPR) based reference points to calculate a target catch (Clark, 1991; NEFSC, 2002; Haltuch et al., 2008). The target fishing mortality rate was specified at $F_{35\%}$ for all life histories. Assessments were conducted annually during years 31–55.

We explored a constant fishing mortality rate scenario and one in which the target fishing mortality rate was $F_{35\%}$. In the first management scenario, no harvest policy is used and the true F throughout the management period is fixed at the plateau value from the initial period (0.5, 1.0 or 2.0 × F_{MSY}). A harvest policy was applied in the second management scenario, whereby the annual catch was estimated using the abundance in the terminal year and the $F_{35\%}$ from the assessment model; this level of catch was removed from the population the following year by calculating the resulting F using the Baranov catch equation (Quinn and Deriso, 1999).

2.2. Parameterization and model runs

We explored several scenarios to identify factors affecting the level of autocorrelation in the estimation error from the assessment model. We explored three life histories, three exploitation histories (0.5, 1.0 or 2.0 × F_{MSY}), two management scenarios, two levels of data quality (σ_I), two levels of productivity at low stock sizes (h), and two levels of recruitment variability (σ_R; Table 4). The life histories explored were 'short-lived', 'medium-lived' and 'long-lived'. The short-lived life history had a low maximum age, high natural mortality rate, rapid growth, and early age at maturation. In contrast, the long-lived life history had a high maximum age, low natural mortality, slow growth, and late age at maturation. The medium-lived life history was between the short- and long-lived life histories. For each life history, we set the maximum age (7, 12, and 20 years for the short-, medium- and long-lived life histories, respectively), and mean natural mortality rate (0.4, 0.2, and 0.1 year^{-1}). The maximum age was an aggregate age class. For each life history we explored two values for steepness of the stock-recruitment function (0.45 and 0.9), representing low and high productivity at low stock sizes. All other life history parame-

Table 4
List of model runs explored for each life history and exploitation level. Harvest policy refers to whether a target fishing mortality rate ($F_{35\%}$) was used in the management period.

Model run	Sample size (E)	Survey error (σ_1)	ϕ_M	ϕ_Y	Recruitment variability (σ_R)	Harvest policy?	Stock-recruit steepness (h)
1	200	0.29	0.3	0.3	0.77	No	0.45
2	50	0.63	0.9	0.9	0.77	No	0.45
3	200	0.29	0.3	0.3	1.25	No	0.45
4	50	0.63	0.9	0.9	1.25	No	0.45
5	200	0.29	0.3	0.3	0.77	Yes	0.45
6	50	0.63	0.9	0.9	0.77	Yes	0.45
7	200	0.29	0.3	0.3	1.25	Yes	0.45
8	50	0.63	0.9	0.9	1.25	Yes	0.45
9	200	0.29	0.3	0.3	0.77	No	0.9
10	50	0.63	0.9	0.9	0.77	No	0.9
11	200	0.29	0.3	0.3	1.25	No	0.9
12	50	0.63	0.9	0.9	1.25	No	0.9
13	200	0.29	0.3	0.3	0.77	Yes	0.9
14	50	0.63	0.9	0.9	0.77	Yes	0.9
15	200	0.29	0.3	0.3	1.25	Yes	0.9
16	50	0.63	0.9	0.9	1.25	Yes	0.9

ters were either fixed across life histories (L_∞ and the length-weight parameters b and c) or determined from the other parameters. The mean natural mortality, M, was used to determine growth rate, $k = M/1.5$, and age at 50% maturity, $m_{50\%} = M/1.4$ (Charnov and Berrigan, 1991; Charnov et al., 1993; Frisk et al., 2001), which then determined the initial age at 50% selectivity in the fishery ($\gamma_{50\%}(t = 1) = m_{50\%}$). Both M and $s_{50\%}$ varied through time in an autocorrelated manner; Eqs. (T1.4) and (T1.8). For the survey, age at 50% selectivity was lower than that of the fishery, $g_{50\%} = 0.75\ \gamma_{50\%}\ (t = 1)$, and was rounded down to the nearest integer to determine the age at recruitment to the population, $a_R = \lfloor g_{50\%} \rfloor$.

We explored "good" and "poor" cases for data quality, whereby several factors were adjusted to affect assessment performance (Table 4). For each case we varied the CV of the observation error in the survey (lower for the good scenario), the number of samples collected to generate age structured data (higher for the good case), and the amount of autocorrelation in the time-varying parameters (lower in the good scenario). The degree of autocorrelation in M and $s_{50\%}$, and the standard deviation for M, $s_{50\%}$, q, and the survey index varied based on the assessment quality scenarios (Table 4). Values for these parameters were based on preliminary model runs to confirm that the values chosen influenced uncertainty in assessment estimates, without resulting in a large number of failed assessment estimations (determined by the maximum gradient component of the objective function in AD Model Builder). Uncertainty in total catch estimates was fixed for the good and poor assessment scenarios (Table 2), and was based on estimates for a range of groundfish stocks (NEFSC, 2008). In addition, we explored two levels of recruitment variability, with the levels of variability based on the meta-analysis of Thorson et al. (2014).

For each scenario, 1000 iterations were run. At the end of each run, the terminal estimate of biomass and recruitment from each assessment was stored along with the true values, and we calculated the amount of lag-1 autocorrelation in the error of biomass and recruitment estimates using a maximum likelihood approach. If $V(t)$ and $V^*(t)$ represent the true and estimated values (for biomass or recruitment) in year t, respectively, then the log-likelihood is:

$$\mathscr{L}_\phi = -\frac{n}{2}\log(2\pi) - n\log(\sigma) - \sum_t V(t) + 0.5\log(1-\phi^2) - \frac{1}{2\sigma^2}\sum_{t1} [\log(V(t)) - \phi\log(V(t-1)) - \log(V^*(t)) + \phi\log(V^*(t))\,]^2 - \frac{1-\phi^2}{2\sigma^2}[\ \log(V(1)) - \log(V^*(1))\,]^2 \quad (2)$$

and σ and ϕ were estimated by minimizing $-\mathscr{L}_\phi$.

3. Results

Estimates of the lag-1 autocorrelation (ϕ_S) in biomass errors were almost always positive, with the majority of values between 0.5 and 1.0. Life history and exploitation history had clear effects on ϕ_S (Fig. 1), with life history having the largest effect overall across all model scenarios considered (Table 5). Across life histories, ϕ_S was highest for the long-lived life history and lowest for the short-lived life history, with the medium one in between. Median estimates of ϕ_S ranged between 0.7 and 0.82 for the short-lived life history, between 0.8 and 0.9 for the medium-lived life history, and between 0.88 and 0.94 for the long-lived life history (Fig. 1). For a particular life history, estimates of ϕ_S increased as the fishing mortality rate decreased. However, the magnitude of the differences across exploitation scenarios varied with the species life history. The largest differences in ϕ_S across fishing mortality rates occurred for the short-lived life history, and the smallest were for the long-lived life history (Fig. 1).While life history and exploitation history had the largest effects on ϕ_S, steepness of the stock-recruit relationship, and whether or not a harvest policy was applied in the management period also affected estimates of ϕ_S. Slightly higher estimates of ϕ_S, on average, occurred for the model runs with a lower steepness and when no harvest policy was applied (Table 5). Recruitment variability and assessment uncertainty had the smallest effects on estimates of ϕ_S.

To determine if ϕ_S changed over time as more data were included in the assessment, we split the time series in half and calculated ϕ_S for each half, then determined the difference between the estimates (late ϕ_S – early ϕ_S). The median difference was centered around 0 for all model runs, but the variability depended upon exploitation history and whether or not a harvest policy was applied. Using a harvest policy increased the range of ϕ_S for the light exploitation scenario and decreased the range for the heavy exploitation scenario (Fig. 2).

The impact of exploitation history on estimates of ϕ_S was explored by relating the estimates of ϕ_S with the mean fishing mortality rate (relative to the true $F_{35\%}$) over the entire estimation period. There was a negative relationship between the mean $F/F_{35\%}$ and the estimated ϕ_S, such that increasing fishing pressure resulted in lower estimates of ϕ_S (Fig 3), although this relationship only explained 22% of the variability in ϕ_S.

We also calculated the autocorrelation in the recruitment error, ϕ_R, across model runs for each time series of stock assessment estimates (Fig. 1 and Table 5). Life history had the largest effect on estimates of ϕ_R, followed by recruitment variability (Table 5). The pattern of ϕ_R across life histories followed the opposite trend com-

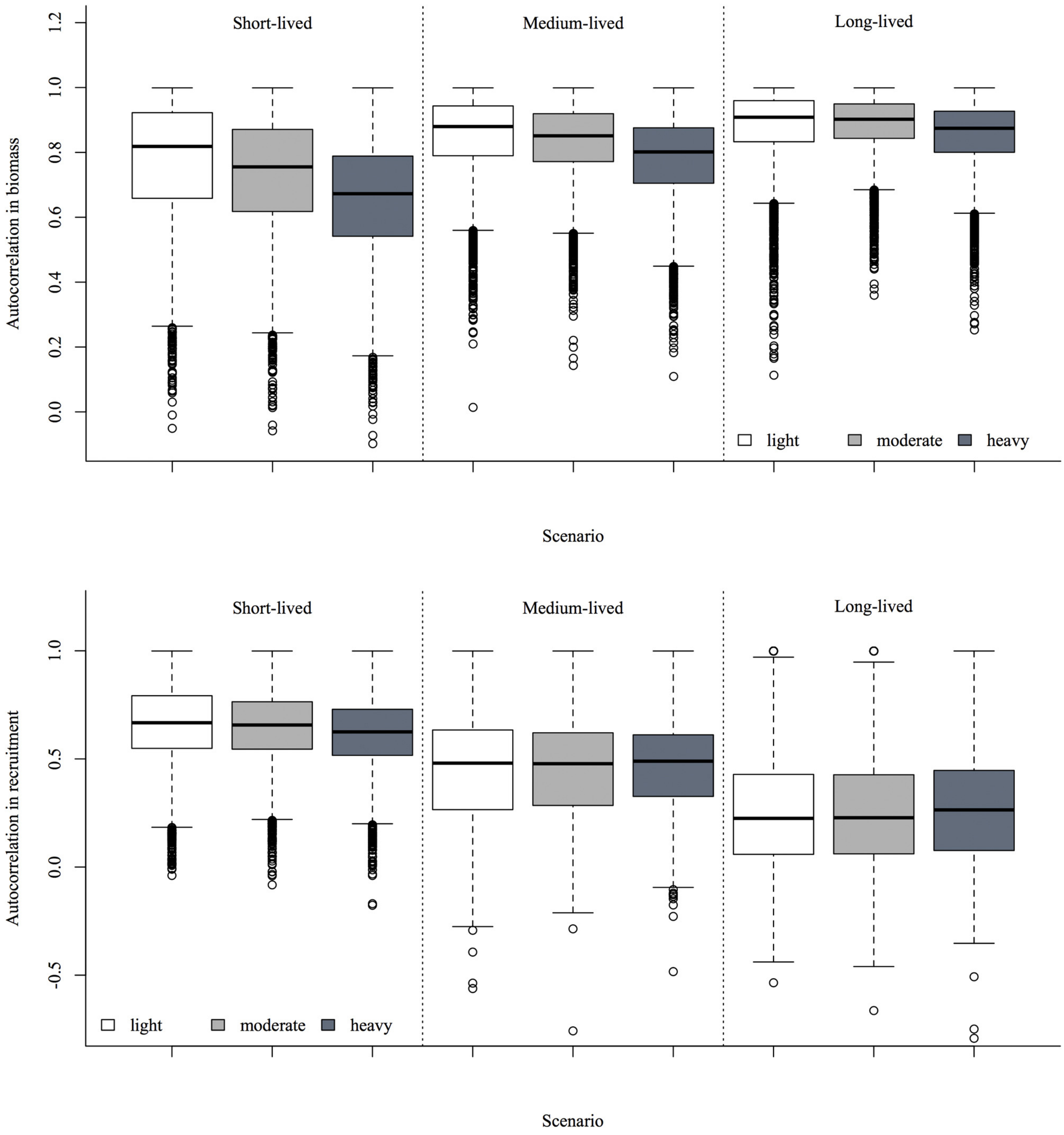

Fig. 1. Estimated lag-1 autocorrelation in biomass (ϕ_S; top panel) and recruitment (ϕ_R; bottom panel). Results are aggregated across model runs listed in Table 4 for the different life histories and exploitation histories explored.

pared to ϕ_S, with generally higher estimates of ϕ_R for the short-lived life history and the lowest estimates for the long-lived life history. For the recruitment variability scenarios, estimates of ϕ_R were slightly higher, on average for runs with lower variability (Table 5).

Estimated ϕ_R was positively correlated with ϕ_S, but ϕ_R was usually less than ϕ_S (Fig. 4). For the short-lived life history, estimates of ϕ_R and ϕ_S were scattered around the 1:1 line, although the slope of a linear regression through the estimates was below 1. For both the medium- and long-lived life histories, the majority of estimates are well below the 1:1 line, with slopes well below 1 (Fig. 4).

Across model scenarios, the effect on estimates of σ_S was generally small, with median estimates ranging between 0.3 and 0.38. The exception to this pattern was for the model runs with low and high assessment uncertainty, which resulted in median estimates of σ_S 0.2 and 0.5, respectively (Table 5). Although the effect was not as large compared to the assessment uncertainty runs, median estimates of σ_S were slightly higher for the low steepness runs, and for the light exploitation history runs (Table 5).

4. Discussion

We estimated the amount of temporal autocorrelation in errors of estimated biomass and recruitment from SCAA stock assessment models over a series of scenarios spanning life histories, exploitation levels, stock productivity, recruitment variability, and data quality. Autocorrelation in the error in biomass estimates (ϕ_S)

Table 5

Comparison of the median estimates of the autocorrelation and standard deviation in biomass assessment error (ϕ_S and σ_S, respectively) and the autocorrelation in recruitment assessment error (ϕ_R), across model run categories. The median for each category (e.g., longevity) is an aggregate across all other categories (e.g., exploitation, steepness, etc.). The range in estimates for a given category was calculated, and the largest ranges are shown in bold.

Model Run		ϕ_S	σ_S	ϕ_R
Life history	Short-lived	0.74	0.37	0.65
	Medium-lived	0.84	0.31	0.48
	Long-lived	0.89	0.34	0.24
	Range	**0.15**	0.06	**0.41**
Exploitation	Light	0.88	0.38	0.49
	Moderate	0.85	0.33	0.49
	Heavy	0.80	0.31	0.49
	Range	0.08	0.07	0.01
Stock-recruit steepness	Low	0.87	0.38	0.48
	High	0.82	0.30	0.49
	Range	0.05	0.08	0.01
Assessment uncertainty	Low	0.83	0.21	0.50
	High	0.86	0.50	0.47
	Range	0.02	**0.29**	0.03
Recruitment variability	Low	0.84	0.33	0.52
	High	0.85	0.35	0.45
	Range	0.01	0.02	0.07
Harvest policy applied?	Yes	0.81	0.35	0.49
	No	0.87	0.33	0.49
	Range	0.06	0.02	0.00

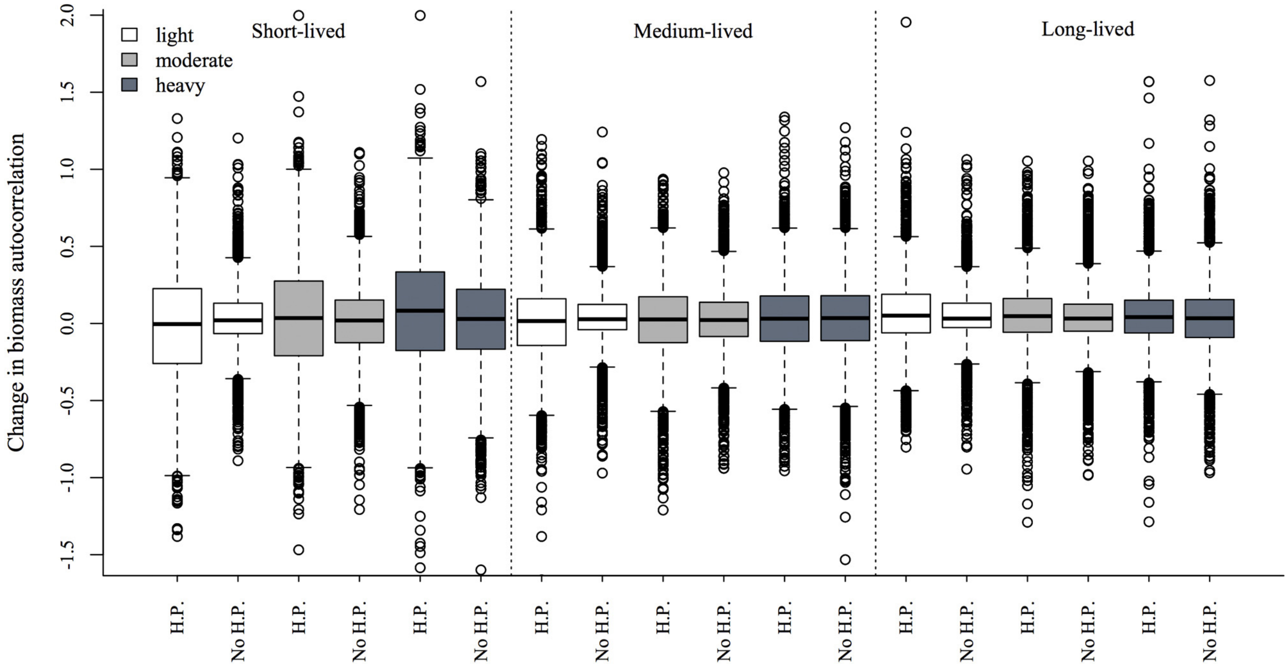

Fig. 2. Change in the estimated autocorrelation in biomass error (ϕ_S) over time. Estimates of ϕ_S were calculated for the first and second halves of the time period, and change was calculated as the difference between these estimates. Results are shown across life histories, exploitation histories, and whether or not a harvest policy was used (denoted H.P. and No H.P).

was positive and relatively high, with median estimates ranging between 0.7 and 0.9. Estimates were highest for the long-lived life history and lowest for the short-lived life history. Exploitation level also affected the amount of autocorrelation, with higher values for lightly exploited populations. On average, however, estimates of ϕ_S did not change over time as more data were included in the assessment, and were not affected substantially by the assumed level of steepness in the stock-recruit relationship, or whether or not a harvest policy was applied. In contrast, recruitment variability and data quality had relatively minor effects on autocorrelation of biomass errors.

In general, higher autocorrelation in the error in biomass estimates indicates poorer estimation by the stock assessment model, as estimates are more consistently above or below the true value. Therefore it is not surprising that estimation was poorer for the scenarios with light exploitation, as there was reduced contrast in the data to help with the estimation. It is well known that increased contrast in data such as the time series of relative abundance is more informative and improves parameter estimation in a range of assessment models (e.g., Hilborn and Mangel, 1997; Magnusson and Hilborn, 2007). For SCAA models, Magnusson and Hilborn (2007) showed that data with the additional contrast of a return trip provide no additional information to the model, and do not

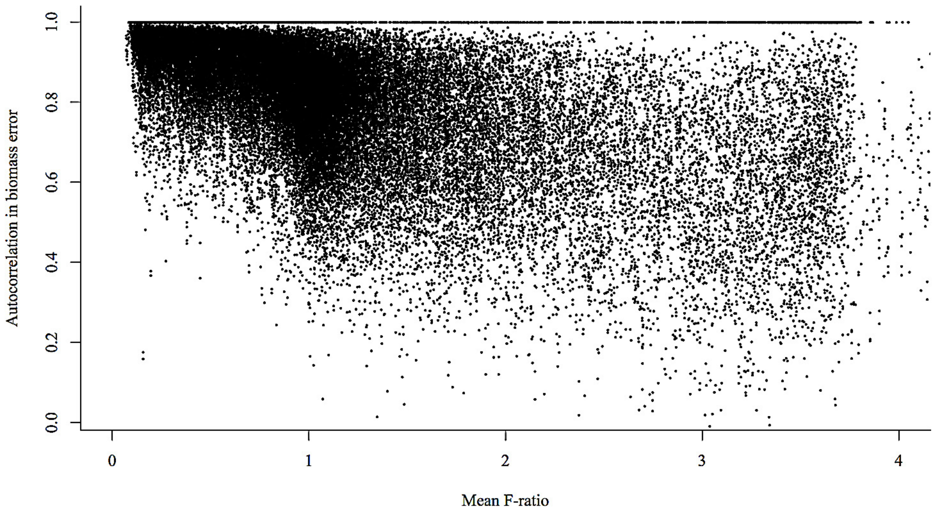

Fig. 3. Estimated lag-1 autocorrelation (ϕ_S) in biomass estimates across all model runs as a function of the mean fishing mortality ratio ($F_{\text{ratio}} = F/F_{\text{target}}$) over the entire time period.

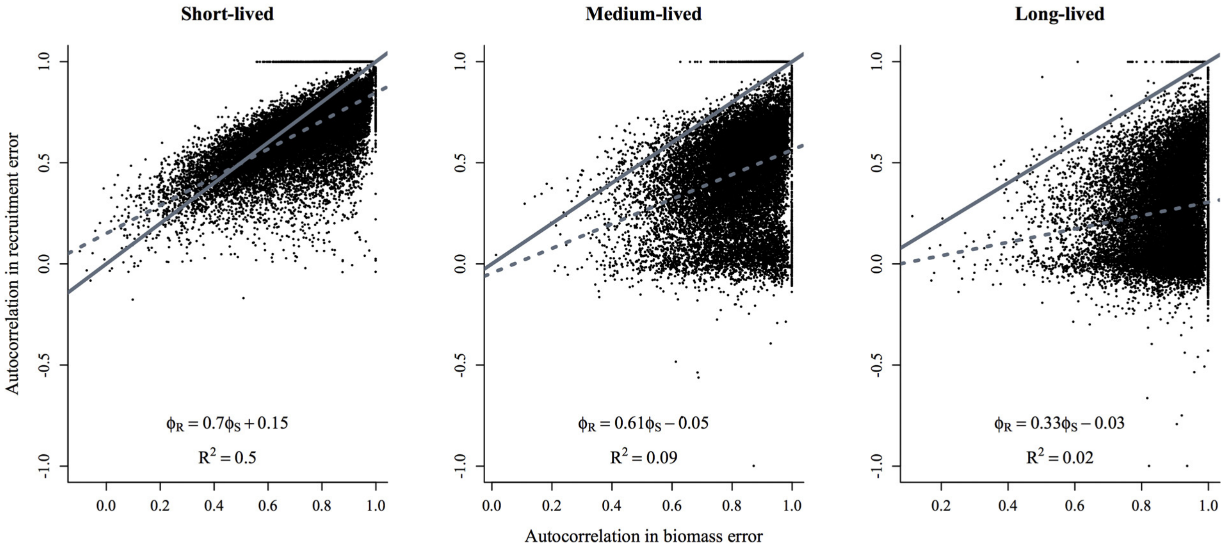

Fig. 4. Estimated lag-1 autocorrelation in recruitment estimates across model scenarios compared to the estimated autocorrelation in biomass estimates across life histories. The solid gray line it the 1:1 line, while the dotted gray line represents the best fitting line to the estimates.

result in an increased ability to estimate parameters. Additionally, our results agree with other simulation studies that found SCAA models have substantially lower accuracy in low fishing mortality rate scenarios than in high fishing mortality rate scenarios (Bence et al., 1993; Wilberg and Bence, 2006). Although higher levels of fishing mortality improved estimation (Fig. 4), a reduction in fishing mortality, which occurred for the heavy exploitation scenarios with a harvest policy applied, resulted in comparable or slightly higher estimates of ϕ_S.

Life history also had an important effect on ϕ_S. Estimates of ϕ_S increased with increasing longevity. A key difference across life history scenarios is the relative contribution of recruits to the total population biomass. Recruits comprise a greater proportion of the population biomass for the short-lived life history given it has fewer age classes, higher growth rates and earlier age of entry into the fishery (and survey), compared to the medium- and long-lived life histories. Recruitment estimates generally had lower autocorrelation than biomass estimates (Fig. 4), so it follows that for cases where recruits comprise a sizeable proportion of the biomass (the short-lived life history and the heavy exploitation scenario) that ϕ_S would be lower. Additionally, the fishing mortality rates were lower, on average, for the long-lived life history than for the medium- or short-lived life histories.

Our results have a number of implications for fisheries management simulations and development of MSE models. First, for MSE studies relying on the stochastic process method of simulating assessments, including autocorrelation of assessment errors is necessary to replicate the outcomes of actual assessment model. Our results provide a range of estimates that can inform choices of the ϕ_S parameter. Higher values of ϕ_S than estimated here have typically been assumed in previous studies. For example, Irwin et al. (2008) and Wilberg et al. (2008) fixed $\phi_S = 0.7$ in their study of harvest policies for yellow perch (*Perca flavescens*) in Lake Michigan. Punt et al. (2008) explored a range of values $\phi_S = 0$, 0.71, and 0.87 in their study of threshold control rules for groundfish along the western U.S., and they found that the level of ϕ_S had an effect on the interannual variability in catches resulting from a particular management policy, an important factor for consideration when selecting a harvest option. While the estimates of autocorrelation used in these studies are within the range of values identified in our simulations, a broader range of values, both below 0.7 and above 0.9 may be warranted under some conditions. The selection of a particular value should be tied to the life history and exploitation level of the species being modeled, with lower values of ϕ_S in scenarios with high exploitation rates and short-lived life histories.

Another important implication for MSE studies is that the error in biomass and recruitment estimates does not show the same level of autocorrelation. Our study found that recruitment error autocorrelation, ϕ_R, was generally lower than biomass error autocorrelation ϕ_S, particularly for the medium- and long-lived life histories. When using the stochastic process MSE assessment approach, one way to estimate recruitments is to use the true age structure and the biomass estimated using Eq. (1) (e.g. Irwin et al., 2008), but this method assumes the same level of autocorrelation in biomass and recruitment estimates. Our analyses show that an alternative approach to estimating recruitment may be more appropriate when using the stochastic process assessment method. One possibility is to draw recruitment and biomass errors from a multivariate distribution with a positive correlation between the errors depending upon the life history of the species being modeled.

In addition to estimates of ϕ_S, investigators wishing to use the stochastic approach must also specify σ_S in Eq. (1). Estimates of σ_S from our simulation were largely controlled by the specified uncertainty in the annual survey estimates, the number of samples taken to determine the age composition of both the survey and the fishery, and the error and autocorrelation in M and $\gamma_{50\%}$. Because the accuracy of the assessment models was largely controlled by values we set for the low and high assessment uncertainty runs, the values presented here (between 0.2 and 0.5) should only be used a guide when using the stochastic process MSE approach. Whenever possible a wide range of values for σ_S should be explored to determine their effect on MSE results, and values below and above our range of estimates are warranted.

Ideally, parameters of a stochastic process could be chosen such that the full stock assessment and stochastic process approaches would result in the same general predictions for a given case study. The results of studies that have compared full and shortcut assessment approaches in MSE models indicate that the approaches can lead to differing predictions for what the optimal harvest policy may be under certain conditions (ICES, 2013). This result does not necessarily invalidate the use of the stochastic process approach in MSE. Rather, it emphasizes the importance of carefully choosing the parameters for the stochastic process approach. For example, a full age-structured assessment will provide estimates (with error) of recruitments, selectivities, and possibly biological reference points. If the management system being modeled requires short term projections, then these estimates can be used in the projections, and error will propagate through time, potentially influencing the performance of a particular harvest policy. In contrast, the stochastic process approach does not produce these estimates; some assumptions must be made if projections are to be conducted as part of the management model. For example, selectivity at age may be fixed at the true values, and recruitments may be generated using the estimated biomass and the true proportions at age (Irwin et al., 2008). If a stochastic process approach is going to be used in an MSE, investigators should conduct some simulation studies to identify levels of autocorrelation in errors that allow the stochastic process to closely match the pattern of errors from full assessments.

Although we explored a number of model scenarios in this study (Table 4), many other scenarios involving changes in population dynamics (e.g. regime shifts), data uncertainty or bias (e.g., ageing error), or model misspecification (e.g., incorrectly assumed form of stock-recruit relationship) may impact estimates of $\sigma_{S,}$ ϕ_S, and ϕ_R. Simulation studies involving full assessment models are common (e.g., Magnusson and Hilborn, 2007; Deroba and Schueller, 2013; Hurtado-Ferro et al., 2015), and the information needed to calculate the autocorrelated error in assessment estimates is often readily available. Having estimates of $\sigma_{S,}$ ϕ_S, and ϕ_R (and for other assessment-estimated quantities; Punt et al., 2015) from a range of studies exploring different population and assessment dynamics would increase our understanding of the factors influencing estimates from stock assessment models.

Managing fish stocks in the face of uncertainty is a key challenge for fisheries managers, and MSE is an essential tool to help identify robust management practices across a range of uncertain outcomes. Punt et al. (2015) recommended using a full stock assessment model if possible in MSEs. However, MSEs that include a full stock assessment are limited in the number of scenarios that can be explored due to the sometimes lengthy computation time. Using the stochastic process approach within an MSE is a useful alternative, particularly when a large number of scenarios must be explored, and the results of this paper can be used as a guide in the selection of appropriate levels of autocorrelation in error in biomass and recruitment estimates. Future work, however, is needed to identify if either the full stock assessment or stochastic process approach is more robust in identifying optimal management policies because both approaches rely on substantial simplifications of complex processes.

Acknowledgements

We thank Rick Methot and Rich Seagraves for insightful discussions on this topic, and André Punt and an anonymous reviewer for comments on this manuscript. This work was funded by the Mid-Atlantic Fisheries Management Council. This is contribution no. 5054 of the University of Maryland Center for Environmental Science.

References

A'Mar, Z.T., Punt, A.E., Dorn, M.W., 2009. Evaluation of two management strategies for the Gulf of Alaska walleye Pollock fishery under climate change. ICES J. Mar. Sci. 66, 1614–1632.

Bence, J.R., Gordoa, A., Hightower, J.E., 1993. Influence of age-selective surveys on the reliability of stock synthesis assessments. Can. J. Fish. Aquat. Sci. 50, 827–840.

Butterworth, D.S., Punt, A.E., 1999. Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56, 985–998.

Carruthers, T.R., Punt, A.E., Walters, C.J., MacCall, A., McAllister, M.K., Dick, E.J., Cope, J., 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fish. Res. 153, 48–68.

Charnov, E.L., Berrigan, D., 1991. Dimensionless numbers and the assembly rules for life histories. Philos. Trans. R. Soc. B Biol. Sci. 332, 41–48.

Charnov, E.L., Berrigan, D., Shine, R., 1993. The M/k ratio is the same for fish and reptiles. Am. Nat. 142, 707–711.

Clark, W.G., 1991. Groundfish exploitation rates based on life history parameters. Can. J. Fish. Aquat. Sci. 48, 734–750.

Deroba, J.J., Schueller, A.M., 2013. Performance of stock assessments with misspecified age- and time-varying natural mortality. Fish. Res. 146, 27–40.

Fournier, D., 2011. An Introduction to AD Model Builder for Use in Nonlinear Modeling and Statistics. Version 10. http://admb-project.org/documentation/manuals/admbuser-manuals

Frisk, M.G., Miller, T.J., Fogarty, M.J., 2001. Estimation and analysis of biological parameters in elasmobranch fishes: a comparative life history study. Can. J. Fish. Aquat. Sci. 58, 969–981.

Haltuch, M.A., Punt, A.E., Dorn, M.W., 2008. Evaluating fishery management reference points in a variable environment. Fish. Res. 94, 290–303.

Hilborn, R., Mangel, M., 1997. The Ecological Detective: Confronting Models with Data. Princeton University Press, Princeton, N.J.

Hurtado-Ferro, F., Szuwalski, C., Valero, J.L., Anderson, S.C., Cuningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES J. Mar. Sci 72, 88–110.

ICES, 2013. Report of the Workshop on Guidelines for Management Strategy Evalua-tions (WKGMSE), 21–23 January 2013. ICES HQ, Copenhagen, Denmark.

Irwin, B.J., Wilberg, M.J., Bence, J.R., Jones, M.L., 2008. Evaluation of harvest policy performance for yellow perch in southern Lake Michigan. Fish. Res. 94, 267–281.

Magnusson, A., Hilborn, R., 2007. What makes fisheries data informative? Fish Fish 8, 337–358.

Miller, T.J., Blair, J.A., Ihde, T.F., Jones, R.M., Secor, D.H., Wilberg, M.J., 2010. Fishsmart: An innovative role for science in stakeholder-centered approaches to fisheries management. Fisheries 35, 424–433.

Milner-Gulland, E.J., Arroyo, B., Bellard, C., Blanchard, J., Bunnefeld, N., Delibes-Mateo, M., Edwards, C., Nuno, A., Palazy, L., Reljic, S., Riera, P., Skrbinsek, T., 2010. New directions in management strategy evaluation through cross-fertilization between fisheries science and terrestrial conservation. Biol. Lett., http://dx.doi.org/10.1098/rsbl.2010.0588

Mohn, R., 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56, 473–488.

Northeast Fisheries Science Center (NEFSC), 2002. Final Report of the Working Group on Re-Evaluation of Biological Reference Points for New England Groundfish. U.S. Department of Commerce, Northeast Fisheries Science Center Reference Document 2–4; 254 p. Available at: http://www.nefsc.noaa.gov/nefsc/publications/. (July 2010).

Northeast Fisheries Science Center (NEFSC), 2008. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4–8, 2008. US Dept Commer, Northeast Fish Sci Cent Ref Doc 8–15; 884 p + xvii. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 2, 543–1026.

Punt, A.E., 2003. Evaluating the efficacy of managing West Coast groundfish resources through simulations. Fish. Bull. 101, 860–873.

Punt, A.E., Dorn, M.W., Haltuch, M.A., 2008. Evaluation of threshold management strategies for groundfish off the U. S. West Coast. Fish. Res. 94, 251–266.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., Haddon, M., 2015. Management strategy evaluation: best practices. Fish Fish, http://dx.doi.org/10.1111/faf.12104 (in press).

Quinn, T., Deriso, R., 1999. Quantitative Fish Dynamics. Oxford University Press, New York.

Thorson, J.T., Jensen, O.P., Zipkin, E., 2014. How variable is recruitment for exploited marine fishes? A hierarchical model for testing life history theory. Can. J. Fish. Aquat. Sci. 71, 973–983.

Wiedenmann, J., Wilberg, M., Miller, T., 2013. Evaluation of harvest control rules for data poor fisheries. N. Am. J. Fish. Man. 33, 845–860.

Wilberg, M.J., Bence, J.R., 2006. Performance of time-varying catchability estimators in statistical catch-at-age analysis. Can. J. Fish. Aquat. Sci. 63, 2275–2285.

Wilberg, M.J., Irwin, B.J., Jones, M.L., Bence, J.R., 2008. Effects of source-sink dynamics on harvest policy performance for yellow perch in southern Lake Michigan. Fish. Res. 94, 282–289.