Fisheries Research 230 (2020) 105636

Contents lists available at ScienceDirect

Fisheries Research

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

On the dangers of including demographic analysis in Bayesian surplus production models: A case study for Indian Ocean blue shark

Zhe Geng[a,b], André E. Punt[c], Yang Wang[a,b], Jiangfeng Zhu[a,b,*], Xiaojie Dai[a,b]

[a] College of Marine Sciences, Shanghai Ocean University, 999 Hucheng Huan Road, Shanghai 201306, China
[b] Key Laboratory of Sustainable Exploitation of Oceanic Fisheries Resources, Ministry of Education, Shanghai Ocean University, Shanghai 201306, China
[c] Schoolof Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98105, United States

ARTICLE INFO

Handled by S Xavier Cadrin

Keywords:
Population dynamics model
Simulation testing
Demographic analysis
Observation error
Stock assessment

ABSTRACT

The Schaefer and Pella-Tomlinson production models (LPM and PTPM) can be used to provide management advice in data poor situations, as they require only a time-series of catches and an index of abundance. These models are commonly fit using Bayesian methods, with the prior for the intrinsic rate of growth (r) set based on the results of a demographic analysis. We used simulations based on blue shark *Prionace glauca* in the Indian Ocean to evaluate the performances of estimation methods that reflect different assumptions regarding the form of the production function and the basis for the prior for r. Nine age-structured operating models reflected different levels of productivity (determined by the steepness of stock-recruitment relationship [$h = 0.4$; $h = 0.6$; $h = 0.79$] and the pattern of historical catches (increasing, stable and declining). As expected, estimation performance was poorer for greater extents of observation error, and better when there was more 'contrast' in biomass. However, the PTPM usually performed worse than the LPM, particularly for high levels of observation error. Surprisingly, the prior for r with mean set to of the estimate of the r inferred from the demographic analysis combined with the LPM performed best for an increasing catch series (a one-way trip in biomass) and high uncertainty in the abundance index. Additional analyses revealed that the poor performance of the PTPM was due to the additional estimation variance associated with the estimation of the shape parameter, while the better performance for the 'wrong prior' occurred because the Schaefer model assumes a linear relationship between growth rate and population depletion whereas an age-structured model implies a non-linear relationship. Given poor data, r is not updated much, leading the LPM to overestimate productivity. This paper highlights the dangers of naively integrating demographic analysis into Bayesian surplus production models, and the value of including simulation analysis as a part of the standard set of diagnostics used when selecting an estimation method on which to base stock assessments. We also recommend use of JABBA-Select or a prior for r from a demographic analysis that accounts for the status of the population when the data on which the demographic parameters are based as well as the form of the production function.

1. Introduction

Considerable progress has been made in fisheries stock assessment methods over the last 30 years, owing to improvements in computer power and computational algorithms (Maunder and Punt, 2013). In particular, complex age-structured models and integrated analysis methods are now more frequently and widely used as the basis for stock assessments to support fisheries management (Zhang, 2013; Walters et al., 2008). Methods such as MULTIFAN-CL (Fournier et al., 1990) and Stock Synthesis (Methot and Wetzel, 2013) are based on complex population dynamics models that account for, *inter alia*, age-based migration, relationships between recruitment and environmental factors, and tagging data (Lee et al., 2017; Maunder and Watters, 2003). However, such models are hard to apply when data are limited (Kelly and Codling, 2006). For example, the cost of age identification and collection of length-frequency data means that only a few fish species have stock assessments based on complex models (Costello et al., 2012).

Surplus production models (SPMs) can provide time-trajectories of population size and associated fisheries reference points with few parameters (Winker et al., 2018). These models represent all biological processes, including natural mortality, individual growth, and fecundity using a production function. The Schaefer/logistic production model (LPM) (Schaefer, 1954) estimates time-trajectories of population size based on estimates of the intrinsic growth rate (r) and carrying

---

[*] Corresponding author at: College of Marine Sciences, Shanghai Ocean University, 999 Hucheng Huan Road, Shanghai 201306, China.
E-mail address: jfzhu@shou.edu.cn (J. Zhu).

https://doi.org/10.1016/j.fishres.2020.105636
Received 27 February 2020; Received in revised form 9 May 2020; Accepted 11 May 2020

capacity (K), while the Pella–Tomlinson production model (PTPM) (Pella and Tomlinson, 1969) includes a parameter p that allows the population size at which production is maximized, B_{MSY}/K, to differ from 0.5K, i.e. $B_{MSY}/K = (p + 1)^{-1/p}$. Given a catch time-series, the LTM and PTPM require only a time-series of abundance indices for parameter estimation, which makes them useful for fisheries where age- or size-composition data are unavailable. Some data-limited methods of stock assessment can determine catch limits with fewer data, but these methods lead to highly uncertain estimates of stock status and often lack transparent assumptions about historical population dynamics (Dichmont and Brown, 2010).

The Bayesian approach to stock assessment takes account of prior information, and model robustness can be improved markedly by using informative priors (Punt and Hilborn, 1997). However, in the context of the PTPM, it is challenging to develop informative priors for p and particularly K given only biological information or limited survey data, and the priors for these parameters are consequently often selected with the intent of being uninformative (e.g., Aires-da-Silva and Gallucci, 2007). Demographic analysis has been suggested as a standard way to create informative priors for r for use in production model analyses by several authors (e.g., McAllister et al., 2001; Cortés, 2008; Tsai et al., 2014) because the relationship between r and life-history parameters can be determined using demographic analysis (McAllister et al., 2001). The uncertainty in life-history parameters will be reflected in the prior for r. For example, the prior for r for blue shark (*Prionace glauca*) would differ by a factor of two depending on whether a one- or two-year reproduction cycle is assumed (Takeuchi et al., 2004; Aires-da-Silva and Gallucci, 2007).

Simulation has been used widely to evaluate the performances of assessment methods (e.g., Cope, 2013; Cope et al., 2015; Carruthers et al., 2014), and is used here to evaluate the implications of prior choices using blue shark in the Indian Ocean as a case-study species. Specifically, we: (a) compared four ways for constructing a prior for r, and (b) evaluated which of the LPM and PTPM has better estimation performance given observation error in catches and the index of abundance. The results of our simulations illustrate some concerns with using demographic analysis as the basis for priors for r in uninformative situations and we consequently use additional analyses to conduct a deeper exploration of the use of demographic analysis as the basis for a prior for r.

2. Materials and methods

The simulation study involved using Stock Synthesis (Methot and Wetzel, 2013) parameterized for Indian Ocean blue shark as an operating model (OM) and production models fitted using Bayesian methods as the estimation methods (EMs). The simulations accounted for uncertainty in productivity and the quality of the available data. The operating model and estimation methods assume deterministic dynamics (i.e., no process error) for simplicity and because recruitment variation for sharks is likely low, and certainly low compared to that of most teleosts (Walker, 1998). The OM is started at unexploited equilibrium and this assumption also forms the basis for the EMs.

2.1. Operating model

A simplified version of the 2017 stock assessment for the Indian Ocean blue shark conducted using Stock Synthesis (Rice, 2017; Coelho et al., 2018) was used as the operating model (OM). Natural mortality was sex- and age-specific, growth was modeled using the 3-parameter Richards growth model (Schnute, 1981; Richard et al., 2013; Methot and Wetzel, 2013), and the stock-recruitment relationship was modeled using the Beverton–Holt form, with biological parameter values set to those in the 2017 assessment (Table 1). For computational ease, and to more easily identify the reasons for the patterns in the results, all the fisheries in the 2017 assessment were merged into a single fishery with a logistic selectivity pattern equal to the maturity ogive. The assumption that the exploitable biomass is the same as the mature biomass and the assumption of deterministic dynamics increased the consistency between the production models and the OM, resulting in differences between the OM and EMs being primarily a result of differences in the assumed population dynamics model and observation error. The simulated CPUE data were based on the expected CPUE in the OM under the assumption of lognormal error, and the simulated catch data were likewise assumed to be lognormally distributed. The simulated data were generated using the bootstrap feature of Stock Synthesis, and the analyses were based on 100 simulated data sets.

Table 1
Life history parameters and model assumptions used in the operating model.

Parameter	Value	
	Female	Male
Natural mortality-at-age	Supplementary Fig. S1	
Age-at-maturity(m_t)	5 yr	N/A
Reference age-1 (a_1)	0.5 yr	
Maximum age-2 (a_2)	26 yr	
Length at age a_1 (L_1)	49.8 cm	
Length at age a_2 (L_2)	283.2 cm	
Growth rate (K)	0.129 yr^{-1}	0.131 yr^{-1}
Weight-at-length	$W = 5.39\text{e–}006 \times L^{3.10}$	$W = 3.29\text{e–}006 \times L^{3.23}$
Size-at-50%-maturity	145cm	N/A
The slope of the maturity curve	−0.138	N/A
Stock-recruitment relationship	Beverton–Holt	
Steepness (h)	0.4, 0.6, 0.79	
Log of recruitment in an unexploited state log(R_0)	7.6721	
Reproduction cycle	1 yr	N/A
Fecundity (litter size)	38	N/A
Sex ratio at birth	0.5	0.5
Fishery selectivity (Inflection point and slope)	$\beta_1 = 145$ cm; $\beta_2 = 21.3365\ \text{cm}^{-1}$	

2.2. Estimation methods

2.2.1. The Pella–Tomlinson and Schaefer production models

The PTPM is defined by Polacheck et al. (1993):

$$B_t = \begin{cases}\varphi K \\ B_{t-1} + \frac{r}{p}B_{t-1}(1 - (B_{t-1}/K)^p) - C_t\end{cases}\text{if } t = 1$$

otherwise (1)

where B_t is the biomass at the start of year t, and C_t is the total catch during year t. The value of ϕ was set to 1 to match the OM, and setting $p = 1$ leads to the Schaefer model. The annual growth rate for the PTPM (and the LPM for $p = 1$) is given by:

$$r_t = \frac{r}{p}(1 - (B_t/K)^p) \tag{2}$$

The relationship between the modeled biomass and the CPUE indices, $\widehat{I}_t$ is:

$$\widehat{I}_t = qB_te^{\varepsilon_t} \tag{3}$$

where q is the catchability coefficient, I_t is observed abundance index, and ε_t is independent and normally distributed observation error with mean 0 and precision τ (i.e., 1/variance). The negative log-likelihood likelihood function for the CPUE data is, therefore:

$$LL = -\frac{n}{2}[\ln(2\pi) - 2\ln(\tau)] + \frac{1}{2\tau^2}\sum\nolimits_{t=1}^{n}[\ln(I_t) - \ln(\widehat{I}_t)]^2 \tag{4}$$

where n is the number of CPUE data points.

A Markov Chain Monte Carlo (MCMC) algorithm based on R2jags version 0.5–7 (Su, 2015) implemented within R (R Development Core Team, 2018) was used to sample the parameter space. Four chains, each with 85,000 iterations were generated. The first 5000 iterations for each chain were eliminated as a burn-in, and a thinning rate of 40 was applied. This led to 8000 samples (2000 from each chain) from the posterior distribution. The Brooks–Gelman–Rubin statistic (BGRs; Gelman and Rubin, 1992; Brooks and Gelman, 1998) was used to evaluate convergence (< 1.1 for all parameters), and only converged results were used for inference.

2.2.2. Demographic analysis

Demographic analysis, using a Leslie population projection matrix (Caswell, 2002), was used to develop an informative prior for r for use in the EMs:

$$N_{t+1} = \mathbf{M}N_t \tag{5}$$

where N_t is the vector of numbers-at-age (females only) at the start of year t, and $\mathbf{M}$ is a Leslie population projection matrix based on a birth-pulse population and a post-breeding census (Caswell, 2002), i.e.:

$$\mathbf{M} = \begin{bmatrix} f_0 & f_1 & f_2 & \dots & f_x \\ s_0 & 0 & 0 & 0 & 0 \\ 0 & s_1 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & s_{x-1} & 0 \end{bmatrix} \tag{6}$$

where s_x is the annual natural survivorship for age x, and the f_x represents the age-specific per-capita fecundity rates. The first age class (age 0) represents newborn pups, and the fecundity (f_x) terms include the probability that a pregnant female delivers pups at the end of the year ($f_x = p_0 m_x$ where m_x is the average number of female eggs per female, and p_0 is the proportion of eggs that survive to pups[1]). The m_x terms were calculated as the product of the eggs-per-female and the sex ratio of litters, divided by the length of the reproductive cycle in years.

In a stable state, $\mathbf{M}N_t = \lambda N_t$, where λ is the finite rate of population increase ($r = \ln\lambda$), and is computed as the dominant eigenvalue of $\mathbf{M}$ (Simpfendorfer et al., 2005). The demographic model was parameterized based on the values for the biological parameters in the OM. The OM is based on the Beverton–Holt stock-recruitment relationship (Rice, 2017):

$$R = \frac{4hR_0 P/P_0}{(1-h) + (5h-1)P/P_0} \tag{7}$$

where R is the number of age-0 animals (pups), P is the number of eggs produced, R_0 is the number of pups at unexpected equilibrium, and P_0 is the number of eggs produced at unexploited equilibrium. According to Brooks et al. (2010), the value for p_0 can be derived from the pups/egg ratio in the limit of zero population size, i.e.:

$$p_0 = \frac{4hR_0}{SB_0(1-h)} \tag{8}$$

2.2.3. Priors

The parameters of the EM are r, q, K and p. B_{MSY}/K is seldom larger than 0.5 for the Beverton–Holt stock-recruitment relationship (Punt et al., 2014). The prior for p was therefore set to U [0,1], which corresponds to a range for B_{MSY}/K from 0.37 to 0.5. The prior for q was selected to be uninformative (U[−20,0] on log-q; Millar and Meyer, 2000), as was the prior for the precision parameter $\tau \sim$ Gamma (0.001,0.001). It is seldom possible to obtain independent information about K, and it is therefore common to place a uniform prior on this parameter based on the maximum historical annual catch (e.g., Froese et al., 2016). However, Millar and Meyer (2000) argued that this assumption could lead to a flat posterior distribution in some cases. Hence, following McAllister (2011), a uniform prior was placed on log-K − U[log$C_{\max}$, log50$C_{\max}$], where $C_{\max}$ is the maximum historical annual catch.

Four priors are considered for r:

- normal with a mean given by the value of r from the demographic analysis [$r = 0.1$, 0.16 and 0.25 for $h = 0.4$, 0.6 and 0.79], and a standard deviation of 0.2 (e.g., Takeuchi et al., 2004) [denoted "DEM"];
- normal with a mean that is half of the value of r from the demographic analysis [$r = 0.05$, 0.08 and 0.13 for $h = 0.4$, 0.6 and 0.79], and a standard deviation of 0.2 [denoted "DEML"];
- normal with a mean based on a demographic analysis that assumes $p_0 = 1$ (0.38), and a standard deviation of 0.2 [denoted "DEMH"];
- U[0,0.5], which contains the values of r for low and very low productivity species (Froese et al., 2017) [denoted "NON"].

Some demographic analyses (e.g., Aires-da-Silva and Gallucci, 2007) are based on the assumption $p_0 = 1$ owing to lack of information regarding the recruitment process and the value of steepness. p_0 should be less than 1 so the third prior is based on an over-estimation of r. In addition, the second prior is speculative, developed primarily to explore the consequences of a 'mis-specified' and underestimated prior. Table 2 summarizes the prior specifications for the eight EMs considered.

2.3. Scenarios

The scenarios relate to the priors on which the EM is based (see section 2.2.3) and the specifications of the OM. Three values for the steepness of the stock-recruitment relationship ($h = 0.4$; $h = 0.6$; $h = 0.79$), where $h = 0.79$ is the value assumed in the 2017 assessment of Indian Ocean blue shark, were considered to explore robustness to productivity. Three catch series (Fig. 1): "Increasing" (the actual catch series for blue shark in the Indian Ocean), a time-series that decreases to zero during the last 20 years of the modeled period (denoted "Declining"), and a time-series that declines and is relatively stable over the last 30 years of the modeled period (denoted "Stable") were considered to explore the impact of data contrast on estimation performance (Fig. 2). Lognormal observation error was applied to the historical catch and CPUE, with CVs of 0.1, 0.4, and 0.7 (CPUE) and 0.1, 0.5 and 1 (catch). These values capture a broad range of scenarios from very informative to very imprecise, and the higher values are not unreasonable given the lack of information for blue shark in the Indian Ocean.

2.4. Performance evaluation and expectations

Estimation performance was evaluated in terms of the means (over simulations) of the absolute relative errors (*MAREs*) of MSY and current stock status relative to unfished biomass (referred to as 'depletion'):

$$MARE = \frac{|E-T|}{T} \tag{9}$$

where E and T are respectively the estimated and true values of the quantities of interest. The estimated value is set to the mean of the posterior distribution. The simulations involve many factors so a linear model of the form:

$$\log(MARE) = \mathrm{EM}^{*}\,h^{*}\text{catch-series}+\mathrm{CV(catch)}^{*}\mathrm{CV(cpue)} \tag{10}$$

where * denotes an interaction of variables, was fitted to the results of the simulations. Stepwise model selection was applied separately to the *MAREs* for MSY and depletion to enable patterns to be detected quantitatively. The most parsimonious model was selected using AIC (Akaike, 1974).

The expectation for this analysis would have been: (a) the *MAREs* would be higher for larger extents of CPUE and catch observation error,

---

[1] Or equivalently for this case a multiplier on the pup survival given the OM does not allow for fishery impacts on pups.

Table 2
The eight estimation methods and the priors on which they are based.

Abbreviation	r			p	log(K)
	h = 0.4	h = 0.6	h = 0.79		
DEM_LPM	N[0.1,0.2]	N[0.16,0.2]	N[0.25,0.2]	Fixed at 1	U[$C_{\max}$,50*$C_{\max}$]
DEM_PTPM	N[0.1,0.2]	N[0.16,0.2]	N[0.25,0.2]	U[0,1]	
NON_ LPM	U[0,0.5]			Fixed at 1	
NON_PTPM	U[0,0.5]			U[0,1]	
DEML_LPM	N[0.05,0.2]	N[0.08,0.2]	N[0.13,0.2]	Fixed at 1	
DEML_PTPM	N[0.05,0.2]	N[0.08,0.2]	N[0.13,0.2]	U[0,1]	
DEMH_LPM	N[0.38,0.2]			Fixed at1	
DEMH_PTPM	N[0.38,0.2]			U[0,1]	

Gamma denotes the gamma distribution; U denotes the uniform distribution; N[x,y] denotes a normal distribution with expectation and standard deviation equal to x and y respectively.

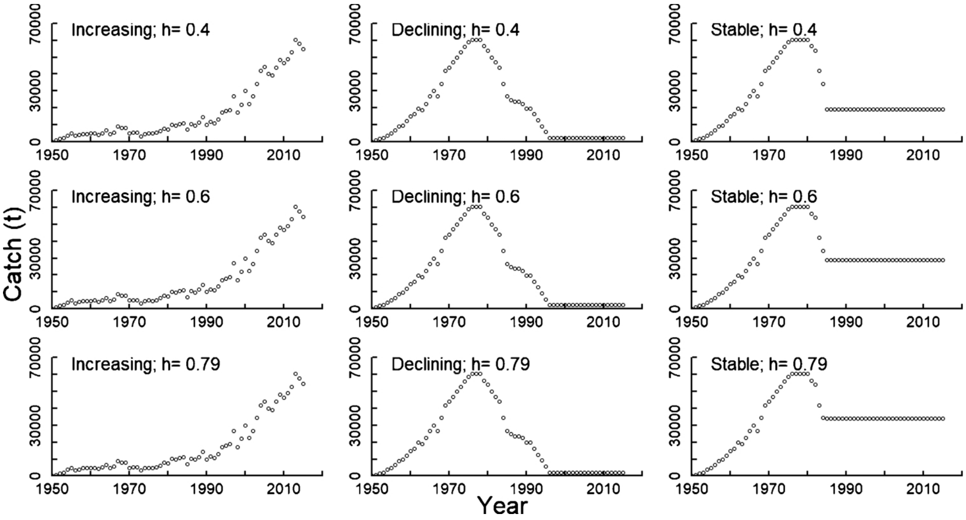

Fig. 1. The catch series (1950–2015) in the operating models.

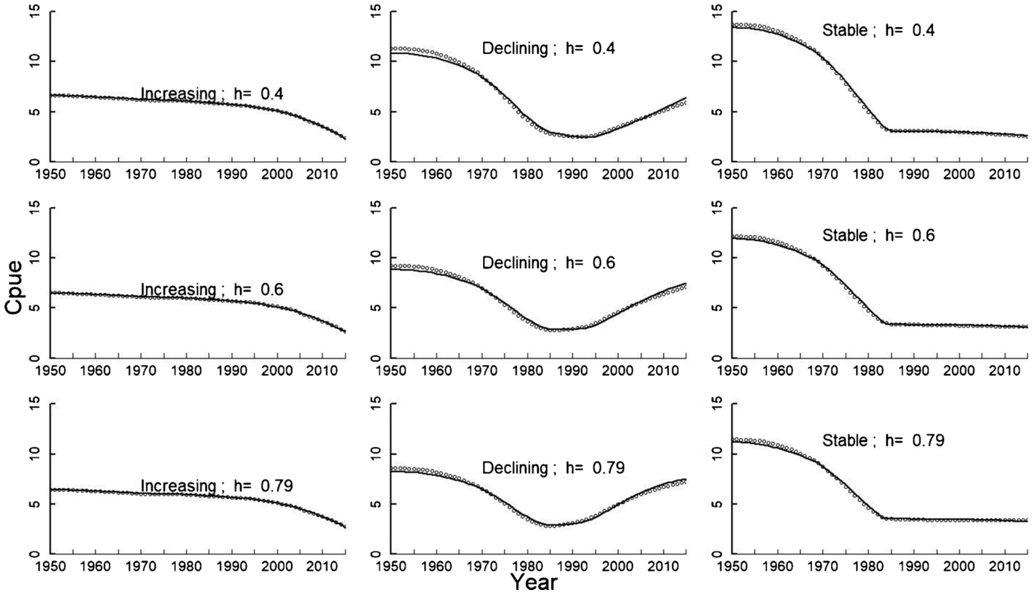

Fig. 2. Expected operating model CPUE (1950–2015) and the fit of the LPM and PTPM models when the catch and CPUE data are not subject to observation error.

(b) more contrast in biomass would lead to better estimation performance, (c) the PTPM would lead to lesser bias and more variation than the LPM, and (d) the results for the EMs based on the prior for r centered on the results of the demographic analysis would outperform the other EMs.

3. Results and discussion

3.1. Results of the simulations

The values for the *MAREs* for all scenarios are listed in

Table 3
Summary of the application of the linear model to determine which factors most influence the *MAREs* for MSY. Significant results ($p < 0.05$) are indicated by bold underlined typeface. The impact of using the LPM, steepness = 0.4, the declining catch series and assuming the DEM prior are included in the intercept. The estimates indicate the difference between the impacts of the other factors and the intercept.

Factors	Estimate	Standard error	t value	p value
(Intercept)	−2.89	0.03	−83.07	**< 0.001**
Catch(Increasing)	1.17	0.04	26.94	**< 0.001**
Catch(Increasing)*Model(PTPM)	0.03	0.04	0.70	0.49
Catch(Increasing)*Model (PTPM)*Steepness(0.6)	−0.01	0.05	−0.20	0.84
Catch(Increasing)*Model (PTPM)*Steepness(0.79)	0.11	0.05	2.05	**0.04**
Catch(Increasing)*Steepness(0.6)	0.07	0.06	1.09	0.28
Catch(Increasing)*Steepness(0.79)	0.14	0.06	2.29	**0.02**
Catch(Stable)	−0.07	0.04	−1.63	0.10
Catch(Stable)*Model(PTPM)	0.03	0.04	0.82	0.41
Catch(Stable)*Model (PTPM)*Steepness(0.6)	0.16	0.05	2.92	**0.04**
Catch(Stable)*Model (PTPM)*Steepness(0.79)	−0.32	0.05	−5.91	**< 0.001**
Catch(Stable)*Steepness(0.6)	−0.64	0.06	−10.49	**< 0.001**
Catch(Stable)*Steepness(0.79)	−0.16	0.06	−2.54	**0.01**
cv_catch	0.82	0.02	34.73	**< 0.01**
cv_catch*cv_cpue	0.25	0.05	4.99	**< 0.001**
cv_cpue	0.91	0.03	27.64	**< 0.001**
Model(PTPM)	0.04	0.03	0.03	0.97
Model(PTPM)*Steepness(0.6)	0.02	0.04	0.61	0.54
Model(PTPM)*Steepness(0.79)	0.06	0.04	1.48	0.14
Prior_r(DEMH)	0.09	0.04	2.12	**0.03**
Prior_r(DEMH)*Catch(Increasing)	0.22	0.05	4.03	**< 0.001**
Prior_r(DEMH)*Catch (Increasing)*Steepness(0.6)	−0.04	0.08	−0.46	0.65
Prior_r(DEMH)*Catch (Increasing)*Steepness(0.79)	0.01	0.08	0.14	0.89
Prior_r(DEMH)*Catch(Stable)	0.02	0.05	0.41	0.68
Prior_r(DEMH)*Catch (Stable)*Steepness(0.6)	−0.13	0.08	−1.63	0.10
Prior_r(DEMH)*Catch (Stable)*Steepness(0.79)	−0.08	0.08	−1.04	0.30
Prior_r(DEMH)*Model(PTPM)	0.08	0.03	2.93	**< 0.01**
Prior_r(DEMH)*Steepness(0.6)	−0.03	0.05	−0.53	0.60
Prior_r(DEMH)*Steepness(0.79)	−0.07	0.05	−1.31	0.19
Prior_r(DEML)	−0.01	0.04	0.27	0.79
Prior_r(DEML)*Catch(Increasing)	−0.01	0.05	−0.21	0.83
Prior_r(DEML)*Catch (Increasing)*Steepness(0.6)	−0.05	0.08	−0.66	0.51
Prior_r(DEML)*Catch (Increasing)*Steepness(0.79)	−0.19	0.08	−2.52	**0.01**
Prior_r(DEML)*Catch(Stable)	−0.01	0.05	−0.11	0.91
Prior_r(DEML)*Catch (Stable)*Steepness(0.6)	0.04	0.08	0.52	0.60
Prior_r(DEML)*Catch (Stable)*Steepness(0.79)	0.03	0.08	0.35	0.73
Prior_r(DEML)*Model(PTPM)	−0.01	0.03	−0.52	0.61
Prior_r(DEML)*Steepness(0.6)	−0.02	0.05	−0.32	0.75
Prior_r(DEML)*Steepness(0.79)	−0.02	0.05	−0.40	0.69
Prior_r(NON)	0.03	0.04	0.65	0.52
Prior_r(NON)*Catch(Increasing)	0.03	0.05	0.50	0.61
Prior_r(NON)*Catch (Increasing)*Steepness(0.6)	0.00	0.08	−0.06	0.95
Prior_r(NON)*Catch (Increasing)*Steepness(0.79)	−0.15	0.08	−1.97	0.05
Prior_r(NON)*Catch(Stable)	0.01	0.05	0.16	0.87
Prior_r(NON)*Catch (Stable)*Steepness(0.6)	−0.01	0.08	−0.15	0.88
Prior_r(NON)*Catch (Stable)*Steepness(0.79)	−0.02	0.08	−0.21	0.84
Prior_r(NON)*Model(PTPM)	0.00	0.03	0.14	0.89
Prior_r(NON)*Steepness(0.6)	−0.03	0.05	−0.47	0.64
Prior_r(NON)*Steepness(0.79)	−0.02	0.05	−0.33	0.74
Steepness(0.6)	0.02	0.04	0.42	0.67
Steepness(0.79)	−0.02	0.04	−0.41	0.68

Table 4
As for Table 3, but for depletion.

Factors	Estimate	Standard error	*t* value	*p* value
(Intercept)	−3.91	0.03	−144.98	**< 0.001**
Catch(Increasing)	1.73	0.03	55.58	**< 0.001**
Catch(Increasing)*Model(PTPM)	−0.17	0.04	−4.85	**< 0.001**
Catch(Increasing)*Model (PTPM)*Steepness(0.6)	−0.01	0.05	−0.11	0.91
Catch(Increasing)*Model (PTPM)*Steepness(0.79)	0.06	0.05	1.19	0.24
Catch(Increasing)*Steepness(0.6)	0.01	0.04	0.37	0.71
Catch(Increasing)*Steepness (0.79)	−0.12	0.04	−3.33	**< 0.001**
Catch(Stable)	0.70	0.03	22.43	**< 0.001**
Catch(Stable)*Model(PTPM)	0.06	0.04	1.75	0.08
Catch(Stable)*Model (PTPM)*Steepness(0.6)	−0.13	0.05	−2.52	**0.01**
Catch(Stable)*Model (PTPM)*Steepness(0.79)	−0.15	0.05	−3.00	**< 0.01**
Catch(Stable)*Steepness(0.6)	0.70	0.04	19.56	**< 0.001**
Catch(Stable)*Steepness(0.79)	1.11	0.04	30.69	**< 0.001**
cv_catch	0.12	0.02	5.45	**< 0.001**
cv_catch*cv_cpue	0.14	0.05	3.08	**< 0.01**
cv_cpue	0.87	0.03	28.42	**< 0.001**
Model(PTPM)	0.14	0.03	4.76	**< 0.001**
Model(PTPM)*Steepness(0.6)	0.02	0.04	0.51	0.61
Model(PTPM)*Steepness(0.79)	−0.06	0.04	−1.54	0.12
Prior_r(DEMH)	0.02	0.02	0.65	0.52
Prior_r(DEMH)*Catch (Increasing)	0.08	0.03	2.55	**0.01**
Prior_r(DEMH)*Catch(Stable)	−0.01	0.03	−0.25	0.81
Prior_r(DEMH)*Model(PTPM)	−0.03	0.02	−1.04	0.30
Prior_r(DEML)	0.00	0.02	0.02	0.99
Prior_r(DEML)*Catch(Increasing)	−0.05	0.03	−1.72	0.08
Prior_r(DEML)*Catch(Stable)	−0.01	0.03	−0.33	0.74
Prior_r(DEML)*Model(PTPM)	0.03	0.02	1.37	0.17
Prior_r(NON)	0.02	0.02	0.79	0.43
Prior_r(NON)*Catch(Increasing)	−0.02	0.03	−0.55	0.58
Prior_r(NON)*Catch(Stable)	−0.01	0.03	−0.25	0.80
Prior_r(NON)*Model(PTPM)	0.01	0.02	0.44	0.66
Steepness(0.6)	−0.03	0.03	−1.18	0.24
Steepness(0.79)	0.06	0.03	2.38	**0.02**

Supplementary Tables S1–S9. The linear model supports that the main effects as well as several 2- and 3-way interactions are consequential (Tables 3 and 4). However, the reduction in the variance in log (*MARE*) explained by the interactions that are evaluated to be “significant” is quite small (adding these interactions only increased the explained variation by 2% for both MSY and depletion [e.g., the Residual Sum of Squares for depletion decreased from 75,617 to 75,551]; Supplementary Tables S10 and S11). Although measures of significance are provided in Tables 3 and 4, they should only be interpreted qualitatively because the ‘data’ is model output, and significance could be increased by increasing the number of replicates.

The values for the parameters of the selected sub-model of Eq. 10 for MSY are listed in Table 3. The magnitude of the *MARE* for MSY differs among the catch series (highest for ‘increasing’, intermediate for ‘declining’, and lowest for ‘stable’). For the stable catch series, the *MARE* for MSY is lower for the higher values of steepness (0.6 and 0.79). The *MARE* for MSY increases with increases to both the extent of catch and (particularly) CPUE observation error, with a strong (and positive) interaction (Table 3). The *MAREs* for MSY differ among the EMs but there is a complex interaction among the various factors in relation to the impact of EM choice. However, in general, the *MARE* is higher for the DEMH prior and the use of the PTPM. The effect of selecting the DEML prior (negatively biased prior mean for *r*) leads to slightly lower *MAREs* for MSY, but significantly for the increasing catch series and higher steepness (a reduction in *MARE* of almost 19% all things being equal for the increasing catch series for both production model types).

The values for the parameters of the selected sub-model of Eq. 10 for depletion are listed in Table 4. The main factors again influence the *MAREs* but fewer interactions are among the factors that are important for depletion than was the case for MSY. The most influential factor is again the catch series but, in this case, the lowest *MAREs* occur for the 'declining' series and the highest (again) for the 'increasing' series. Increasing the extent of observation error again leads to larger *MAREs* but the size of the interaction of catch observation error is less consequential than was the case for MSY. OM steepness is again important, but performance is now best for the lowest value for steepness (0.6), with the poorest performance for the highest value for steepness (0.79). Basing the EM on the PTPM again leads to higher *MAREs*.

3.2. Comparison with a priori expectations

Expectation (a) is supported by the results of the simulation analysis because the *MAREs* increase with increasing extents of observation error for catch and CPUE. Lower CPUE observation error (particularly the lowest level) somewhat reduced the likelihood of extreme over-estimation caused by high catch observation (Supplementary Figs S2–S4). Expectation (b) is also supported because the *MAREs* are highest for the increasing catch series (a one-way trip in biomass). The *MAREs* for the most informative catch series ('declining') were either the lowest (depletion) or equally lowest. In relation to expectation (c), the *MAREs* for the PTPM are higher than those for the LPM, presumably because the reduction in bias is more than offset by the increase in variation. Expectation (d) is generally strongly supported for depletion, with the "DEM" EMs usually leading to the best performance (although the differences in *MAREs* among EMs are seldom large). The situation for MSY is less clear, with quite marked differences in *MAREs* amongst EMs, in particular, the 'wrong prior' for *r* (DEML) leads to the lowest *MAREs*. To help summarize the results for MSY, the best EM for each combination of catch series, steepness value, and combination of observation error CV is indicated in Supplementary Tables S1–S9. This highlights that when one of the EMs substantially outperforms the others, it is the DEML prior combined with the Schaefer production model (abbreviation DEML_LTM) when the catch series is increasing.

3.3. Why does the 'wrong' prior sometimes outperform the 'correct' prior for r?

Two sets of analyses were undertaken to explore how the prior for *r* interacts with the choice of population dynamics model. The first set of analyses involved fitting the LPM and PTPM when there was no observation error based on maximum likelihood, and the second set involved fitting these models using a Bayesian estimation framework when there was observation error in catch and CPUE.

The LPM and PTPM can fit the CPUE data almost perfectly when there was no observation error (Fig. 2), suggesting that both models have the flexibility to mimic the data. However, it is evident that the LPM and PTPM do not capture the dynamics exactly when the results of the fits are expressed as the estimated relationship between population growth rate and population depletion (Fig. 3) and the estimated production function (Fig. 4), although the PTPM fits better. This occurs because of model mis-specification (for the LPM) and because neither the LPM or PTPM allows for time-lags and transient age-structure effects, as noted by Punt and Szuwalski (2012). In particular, the LPM assumes a linear relationship between growth rate and depletion, the PTPM allows for a non-linear relationship between growth rate and depletion, and the OM is based on an age-structured model. This leads to a marked difference between the 'true' (OM) relationship between growth rate and population depletion, particularly for the LPM.

Accounting for observation error (see Fig. 5–7 for a sample of the scenarios for the DEM_LPM, DEM_PTPM, DEML_LPM, and DEML_PTPM estimation methods for a steepness of 0.6; Supplementary Fig. S2 for all scenarios for the increasing catch series) reveals that the PTPM usually mimics the relationship between growth rate and population depletion better than the LPM in terms of central tendency, and particularly at low levels of depletion. However, the PTPM leads to much greater inter-simulation variation in this relationship than the LPM. The lack of information caused by observation error means that the prior for *r* is not updated substantially for the fairly uninformative increasing catch series, (Fig. 8). This lead, for higher observation error CVs for catch and CPUE, to the situation in which the true relationship between growth rate and depletion is not included in the samples from the posterior (second and third rows of Fig. 5–7), and to frequently over-estimate *K* and hence MSY. The over-estimation of growth rate occurs under the DEM prior because the LPM is forced to have a growth-rate intercept equal to the value for *r* generated from the prior for *r* that declines to zero at *K*.

The relationship between growth rate and depletion is linear for the LPM (Eq. 2) so *a priori* this relationship will over-estimate productivity compared to the true situation even though the base level for *r* is correct for the OM. This "problem" can be rectified by either the availability of better data or basing the prior on a value for *r* that is more representative of the full range of depletion than the growth rate in the limit of zero population size. Overall therefore, the DEML_LPM EM performs best for estimating MSY for the increasing catch series when the data are uninformative because (a) using the DEML prior leads to an *a priori* relationship between growth rate and depletion that better mimics the true relationship over most of the range of depletion, and (b) using LPM leads to lesser variance.

4. Conclusion and recommendations

The results of this paper highlight the danger of integrating demographic analysis into Bayesian surplus production models without careful consideration of the behavior and consequences of the prior for *r* and the implied relationship between growth rate and population depletion. These consequences, and hence the risk of over-estimation of productivity when the prior for r is based on information of demography in the limit of zero population size as is the intent, appears to be greater in data-poor (high uncertainty in catch and CPUE) and uninformative (one-way trip in catch and biomass) cases. While simulation is not only the diagnostic tool that should be used to select among estimation methods, it is an effective way to do so and should be part of the standard model selection toolbox. This is particularly highlighted by Fig. 2–4 that show that a model can provide a near-perfect fit to a data set yet lead to quite erroneous results. We recommend that simulation be used more routinely, particularly in data-poor cases, to examine the consequences of limited data, and hence the implications of different prior choices. Examining sensitivity to different priors is an important first step, but cannot quantify the effects of model mis-specification.

Simulation has been used extensively to evaluate the performances of assessment methods (e.g., Punt and Szuwalski, 2012; Zhang, 2013; Carruthers et al., 2014), and this paper continues this work. We wish to highlight the importance of not simply conducting simulations but also developing analysis methods to attempt to synthesize the vast amount of output from a simulation study (in this paper using graphical summaries and linear models) and conducting additional analyses to assess why the results are as they are. The most important conclusions of this work are not the numerical results, but the insight gained on how production models represent population dynamic processes.

The productivity of the OM is related to several parameters such as the growth curve, the rate of natural mortality, and the shape (and steepness) of the stock-recruitment relationship (e.g., Sissenwine and Shepherd, 1987; Punt and Cope, 2017). The use of an informative prior for *r* (DEM) from the demographic analysis led to better performance than the use of the non-informative prior (NON). However, it is also clear that an SPM can only capture the equilibrium production dynamics given a correct assumption about the shape of the production function. *r* in LPM and that in PTPM are related according to:

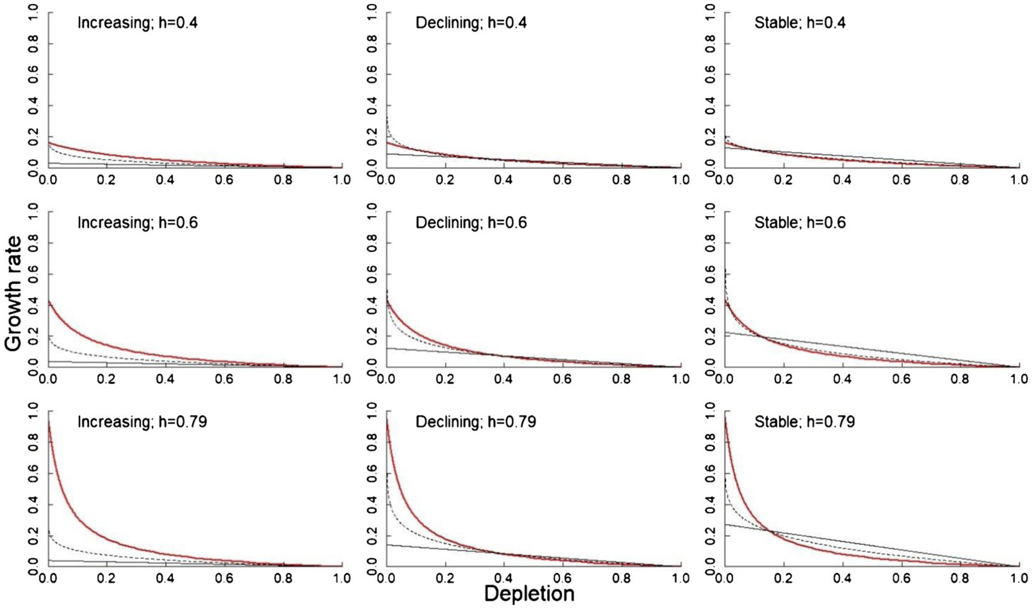

Fig. 3. True (operating model; red line) and estimated relationships between the growth rate and population depletion. The estimates are based on fitting the LPM and PTPM to the catch and CPUE data sampled without error (black solid and black dotted lines respectively). For interpretation of the references to color in this figure citation, the reader is referred to the web version of this article.)

$$r_{PTPM} = \frac{(p + 1)^{\frac{p+1}{p}}}{4} r_{LPM} \tag{11}$$

from which is evident that r_{PTPM} is lower than r_{LPM} when p is less than 1, confirming that, all things being equal, using the same (demographic analysis-based) prior for r will have different consequences for the two SPMs (particularly given uninformative data). The form of relationship between growth rate and population depletion does not explicitly account for in demographic analysis. JABBA-Select (Winker et al., 2020) extends standard Bayesian SPMs by accounting for stock-recruitment steepness and changes over time in selectivity using an age-structure equilibrium model (ASEM) that applies the methods of Sissenwine and Shepherd (1987) to create an informative prior for m (shape parameter) $[m = p + 1]$ and H_{MSY} (harvest rate at MSY). H_{MSY} is a function of r and m:

$$H_{MSY} = \frac{r}{m} \tag{12}$$

JABBA-Select has yet to be fully tested, but initial simulation testing shows that it is less biased and more accurate than convention SPMs, and can perform as well as some age-structured production models particularly for data-poor cases (Winker et al., 2020). However, JABBA-Select requires additional information in the form of a prior for steepness. This is amongst the most difficult parameters to estimate in data-poor situations (Lee et al., 2012; Zhu et al., 2012; and approximately 80% of the global catch comes from such fisheries; Costello et al., 2012). Similar to demographic analysis, steepness can be derived analytically from life-history information, but in common with a

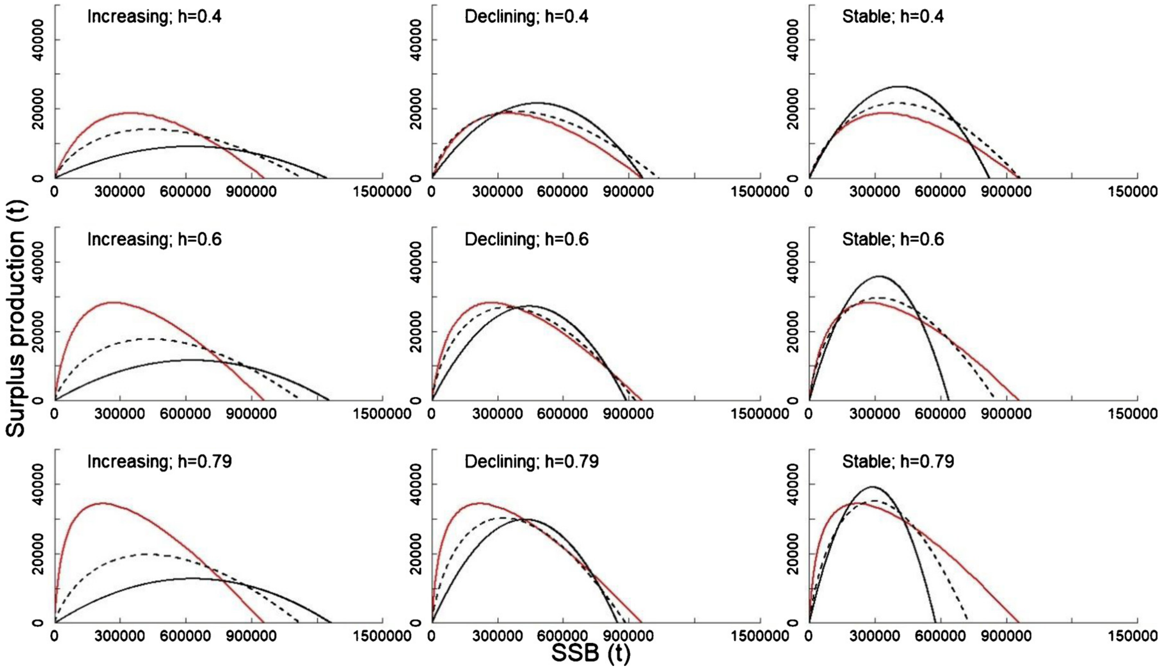

Fig. 4. True (operating model; red line) and estimated relationships between the equilibrium production and spawning stock biomass (SSB). The estimates are based on fitting the LPM and PTPM to the catch and CPUE data sampled without error (black solid and black dotted lines respectively). For interpretation of the references to color in this figure citation, the reader is referred to the web version of this article.)

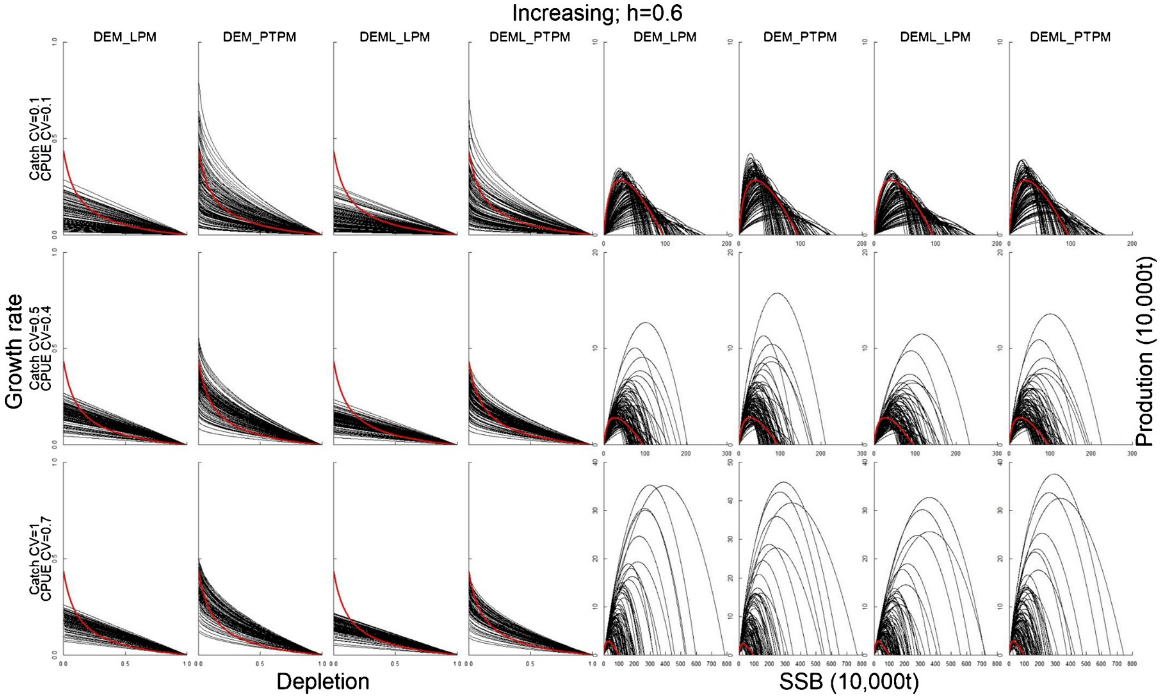

Fig. 5. True (operating model; red line) relationship between the growth rate and population depletion and summaries of posteriors from 100 stochastic simulations (left panels), and true production functions and 100 production functions from posteriors (right panels). The results in the figure pertain to the increasing catch series and three selected sets of levels of observation error. For interpretation of the references to color in this figure citation, the reader is referred to the web version of this article.)

demographic analysis requires values for the biological parameter when the population achieves extremely low depletion (Brooks et al., 2010).

Nevertheless, informative priors for steepness for many species can be obtained using meta-analyses of global compilations of the results of stock assessments for data-rich species (e.g., Punt et al., 2005; Shertzer and Conn, 2012; Wiff et al., 2018). Thus, we recommend that the use of JABBA-Select in such cases. On the other hand, in data-poor situations, basing an assessment on the LPM with a prior for *r* based on a demographic analysis may be the only possibility. Analysts developing priors for *r* need to consider several potential pitfalls related to the information on which the prior for *r* is based:

- The number of eggs produced should not be considered as an estimate of the number of pups (i.e., assuming $p_0 = 1$) as this will overestimate *r* all things being equal.
- The may be value in reducing the mean of the prior (the 50%

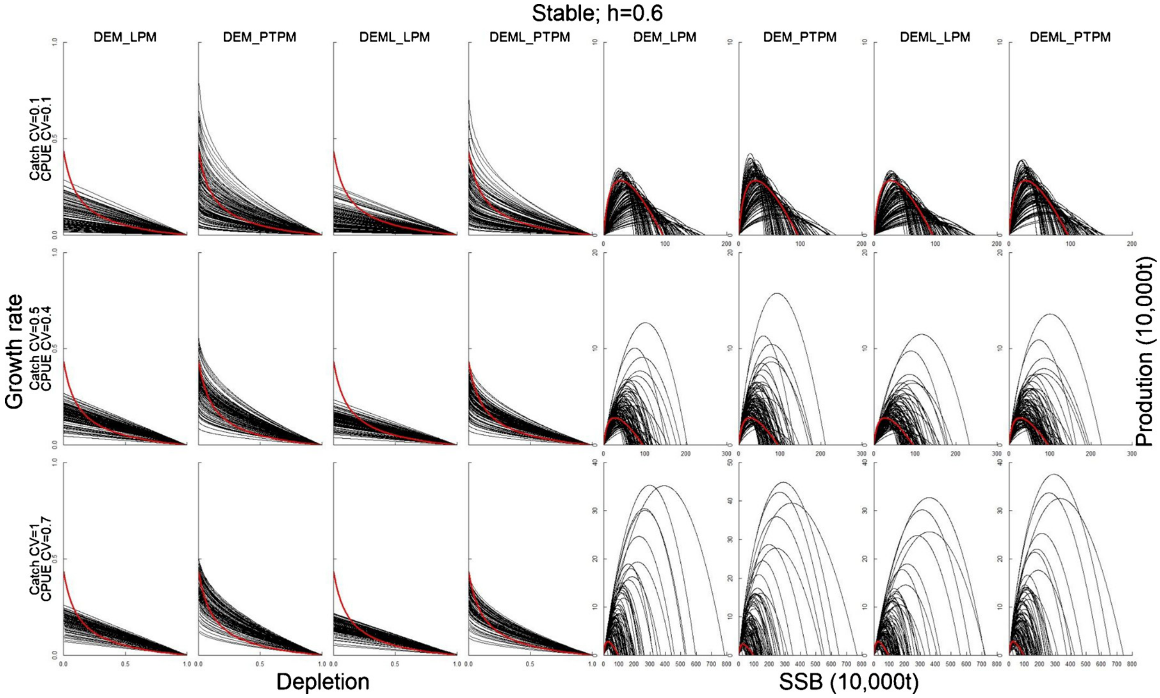

Fig. 6. As for Fig. 5, but for the stable catch series.

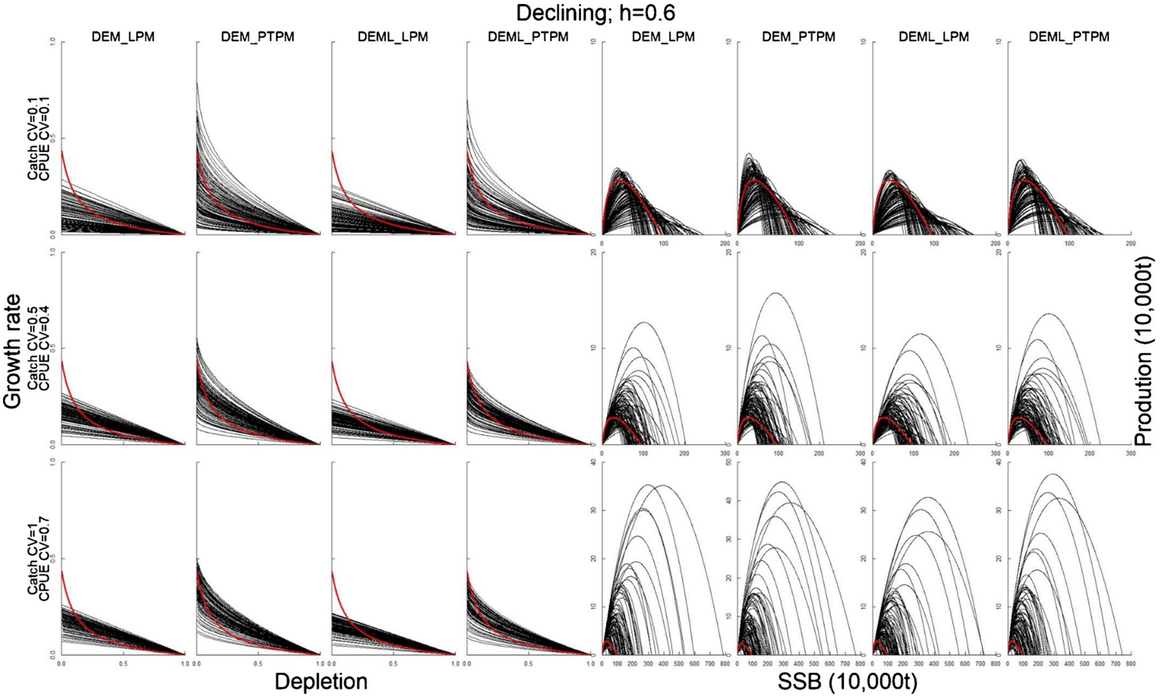

Fig. 7. As for Fig. 5, but for the declining catch series.

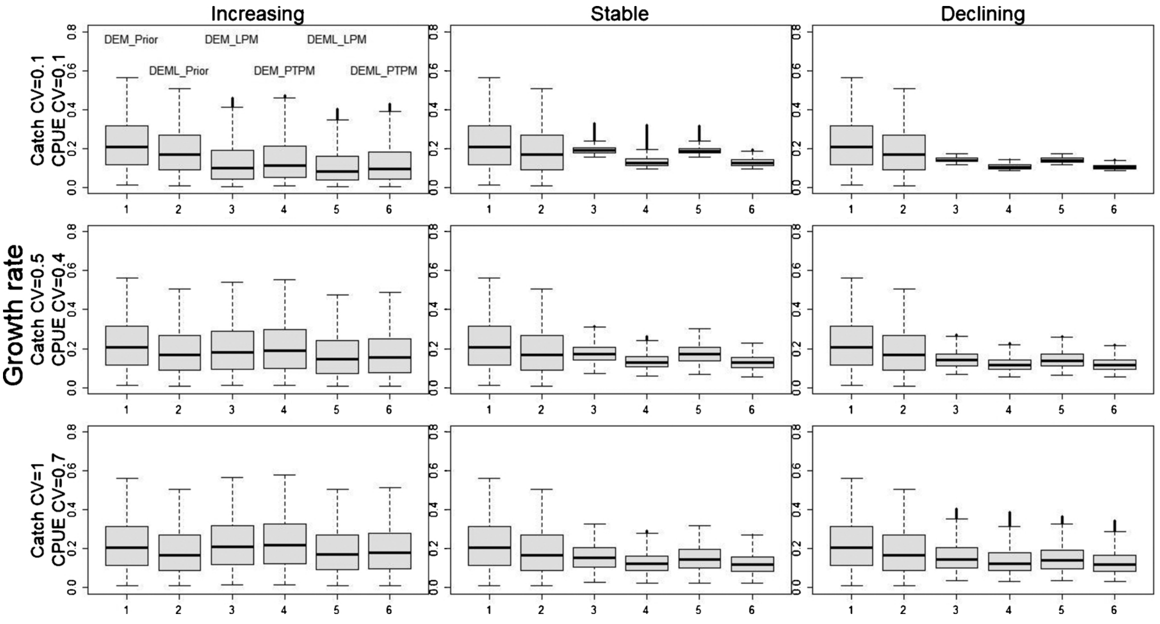

Fig. 8. Priors and posteriors for r from 100 stochastic simulations where h = 0.6.

considered here is arbitrary but was approximately correct for the scenarios considered) if the demographic analysis is based on demographic parameters collected from the population when it was at low stock size (i.e., representative of the true r) but with a stable age structure, and the assessment is based on the LPM. If the data are informative, the prior will be updated (Fig. 8) but if not, a reduced mean will lead to less bias for model outputs such as MSY.

- The value of r will be negatively biased if the demographic analysis is based on values of demographic parameters collected when the population is not at low stock size. Cortés (2016) and Gedamke et al. (2007) note that the relationship between the actual value of r (r_{true}) and r corresponding to the status of the population when the demographic data were collected (r_{current}) is approximately $r_{\text{true}} = r_{\text{curent}}(1 - \text{depletion})$ so basing the demographic analysis on parameter values obtained when the stock was half its unexploited level would be equivalent to the 50% adjustment referred to above.

This paper focused on the consequences of ignoring age-structured dynamics and observation error, and future work could examine the additional consequence of errors in assumptions regarding selectivity, and the nature of process error. In addition, the problems caused by discarding, under- or mis-reporting of catches, conflicting CPUE series, and misclassification of species may play a more important role for some data-poor cases, especially by-catch species. However, as noted by Rudd and Branch (2017), several of the quantities on which management advice is based will be robust to such errors. Nevertheless, the consequences of these uncertainties need further research. The OM (and the actual assessment for Indian Ocean blue shark) was based on the Beverton–Holt stock-recruitment relationship for which B_{MSY}/K is almost always below 0.5 (Punt et al., 2014). Further work could explore alternative stock-recruitment relationships (e.g., Punt and Cope, 2019) for which the production function is more representative of long-lived species such as sharks for which B_{MSY}/K is believed to be larger than 0.5, and hence for which the prior for Pella–Tomlinson p parameter would need to include values larger than 1.

Finally, the purpose of this paper is not to denigrate the use of SPMs, particularly in data-poor situations. SPMs, although unable to fully capture the true dynamics of an age-structured population (Punt and Szuwalski, 2012), at least make use of data on relative abundance, which will minimally improve estimation of depletion, as long as the model is able to mimic the trend in abundance index (Piner et al., 2011). Many other methods for estimating population status for use in management have been developed based on population dynamics models (e.g., CMSY; Froese et al., 2017). However, their performances are unlikely to be better than methods that make use of data on population trends, especially catch-only models that were not intended for long-term management (Wetzel and Punt, 2015). Methods that simply involve transforming priors for population parameters into estimates of biomass and MSY will be even more subject to the concerns outlined here, and should be avoided in favor of methods that update priors given available data.

CRediT authorship contribution statement

Zhe Geng: Conceptualization, Methodology, Software, Writing - original draft, Formal analysis, Investigation. **André E. Punt:** Writing - review & editing, Methodology, Supervision. **Yang Wang:** Data curation. **Jiangfeng Zhu:** Supervision, Funding acquisition, Conceptualization, Project administration, Resources, Validation. **Xiaojie Dai:** Project administration, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was conducted using data from the IOTC Working Party on Ecosystems and Bycatch (WPEB). We would like to thank Joel Rice and Sarah Martin for providing these data. We acknowledge suggestions from faculty and students in the MAPAM laboratory in the School of Aquatic and Fishery Sciences (SAFS) at the University of Washington. The editor and two anonymous reviewers are thanked for their comments on an earlier version of this manuscript. This paper was funded by Shanghai Ocean University with the project of National Natural Science Foundation of China (#41676120) and the China Scholarship Council.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fishres.2020.105636.

References

Aires-da-Silva, A.M., Gallucci, V.F., 2007. Demographic and risk analyses applied to management and conservation of the blue shark (*Prionace glauca*) in the North Atlantic Ocean. Mar. Freshw. Res. 58 (6), 570–580. https://doi.org/10.1071/MF06156.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control. 19, 716–723.

Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. 7 (4), 434–455.

Brooks, E.N., Powers, J.E., Cortés, E., 2010. Analytical reference points for age-structured models: application to data-poor fisheries. ICES J. Mar. Sci. 67 (1), 165–175. https://doi.org/10.1093/icesjms/fsp225.

Carruthers, T.R., Punt, A.E., Walters, C.J., MacCall, A., McAllister, M.K., Dick, E.J., Cope, J.M., 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fish. Res. 153, 48–68. https://doi.org/10.1016/j.fishres.2013.12.014.

Caswell, H., 2002. Matrix Population Models: Construction, Analysis and Interpretation. Sinauer Associates, Sunderland, Massechusetts.

Coelho, R., Mejuto, J., Domingo, A., Yokawa, K., Liu, K.M., Cortés, E., Romanov, E.V., Silva, C., Hazin, F., Arocha, F., Mwilima, A.M., Bach, P., Zárate, V.O., Roche, W., Lino, P.G., García-Corté, B., Ramos-Cartelle, A.M., Forselledo, R., Mas, F., Ohshimo, S., Courtney, D., Sabarros, P.S., Perez, B., Wogerbauer, C., Tsai, W.P., Carvalho, F., Santos, M.N., 2018. Distribution patterns and population structure of the blue shark (Prionace glauca) in the Atlantic and Indian Oceans. Fish Fish. Oxf. (Oxf) 19, 90–106. https://doi.org/10.1111/faf.12238.

Cope, J.M., 2013. Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. Fish. Res. 142, 3–14. https://doi.org/10.1016/j.fishres.2012.03.006.

Cope, J.M., Thorson, J.T., Wetzel, C.R., DeVore, J.L., 2015. Evaluating a prior on relative stock status using simplified age-structured models. Fish. Res. 171, 101–109. https://doi.org/10.1016/j.fishres.2014.07.018.

Cortés, E., 2008. Comparative life-history and demography of pelagic sharks. In: Camhi, M.D., Pikitch, E.K., Babcock, E.A. (Eds.), Sharks of the Open Ocean: Biology, Fisheries and Conservation. Blackwell, Oxford, UK, pp. 309–322.

Cortés, E., 2016. Perspectives on the intrinsic rate of population growth. Methods Ecol. Evol. 7, 1136–1145. https://doi.org/10.1111/2041-210X.12592.

Costello, C., Ovando, D., Hilborn, R., Gaines, S.D., Deschenes, O., Lester, S.E., 2012. Status and solutions for the world's unassessed fisheries. Sci. 338, 517–520. https://doi.org/10.1126/science.1223389.

Dichmont, C.M., Brown, I.W., 2010. A case study in successful management of a data-poor fishery using simple decision rules: the Queensland spanner crab fishery. Mar. Coast. Fish. 2, 1–13. https://doi.org/10.1577/C08-034.1.

Fournier, D.A., Sibert, J.R., Majkowski, J., Hampton, J., 1990. MULTIFAN: a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). Can. J. Fish. Aquat. 47, 301–317. https://doi.org/10.1139/f90-032.

Froese, R., Demirel, N., Coro, G., Kleisner, K.M., Winker, H., 2016. Estimating fisheries reference points from catch and resilience. Fish.Fish. 18 (3), 506–526. https://doi.org/10.1111/faf.12190.

Gedamke, T., Hoenig, J.M., Musick, J.A., DuPaul, W.D., Gruber, S.H., 2007. Using demographic models to determine intrinsic rate of increase and sustainable fishing for elasmobranchs: pitfalls, advances, and applications. N. Am. J. Fish. Manag. 27, 605–618. https://doi.org/10.1577/M05-157.1.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7 (4), 457–472.

Kelly, C.J., Codling, E.A., 2006. Cheap and dirty'fisheries science and management in the North Atlantic. Fish. Res. 79, 233–238. https://doi.org/10.1016/j.fishres.2006.03.007.

Lee, H.-H., Maunder, M.N., Piner, K.R., Methot, R.D., 2012. Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? Fish. Res. 125–126, 254–261. https://doi.org/10.1016/j.fishres.2012.03.001.

Lee, H.H., Piner, K.R., Maunder, M.N., Taylor, I.G., Methot Jr., R.D., 2017. Evaluation of alternative modelling approaches to account for spatial effects due to age-based movement. Can. J. Fish. Aquat. Sci. 74, 1832–1844. https://doi.org/10.1139/cjfas-2016-0294.

Maunder, M.N., Punt, A.E., 2013. A review of integrated analysis in fisheries stock assessment. Fish. Res. 142, 61–74. https://doi.org/10.1016/j.fishres.2012.07.025.

Maunder, M.N., Watters, G.M., 2003. A general framework for integrating environmental time series into stock assessment models: model description, simulation testing, and example. Fish. Bull. (Wash. D. C.) 2003 (101), 89–99. http://aquaticcommons.org/id/eprint/15108.

McAllister, M., Duplisea, D.E., 2011. Production model fitting and projection for Atlantic redfish (Sebastes fasciatus and Sebastes mentella) to assess recovery potential and allowable harm. DFO Can. Sci. Advis. Sec. Res. Doc 2011/057.

McAllister, M.K., Pikitch, E.K., 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 (9), 1871–1890. https://doi.org/10.1139/f01-114.

Methot Jr., R.D., Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99. https://doi.org/10.1016/j.fishres.2012.10.012.

Millar, R.B., Meyer, R., 2000. Non-linear state space modeling of fisheries biomass dynamics by using Metropolis-Hastings within-Gibbs sampling. J. R. Stat. Soc. Ser. C (Applied Stat.) 49, 327–342. https://doi.org/10.1111/1467-9876.00195.

Pella, J.J., Tomlinson, P.K., 1969. A generalized stock production model. Bull. Inter-Am. Trop. Tuna. Comm. 13, 416–497.

Piner, K., Lee, H., Maunder, M., Methot, R., 2011. A simulation-based method to determine model misspecificaton: examples using natural mortality and population dynamics models. Mar. Coast. Fish. 3, 336–343. https://doi.org/10.1080/19425120.2011.611005.

Polacheck, T., Hilborn, R., Punt, A.E., 1993. Fitting surplus production models: comparing methods and measuring uncertainty. Can. J. Fish. Aquat. Sci. 50, 2597–2607. https://doi.org/10.1139/f93-284.

Punt, A.E., Cope, J.M., 2017. Extending integrated stock assessments models to use non-depensatory three-parameter stock-recruitment relationships. Fish. Res. https://doi.org/10.1016/j.fishres.2017.07.007.

Punt, A.E., Cope, J.M., 2019. Extending integrated stock assessments models to use non-depensatory three-parameter stock-recruitment relationships. Fish. Res. 217, 46–57.

Punt, A.E., Hilborn, R., 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish Biol. Fish. 7, 35–63. https://doi.org/10.1023/A:1018419207494.

Punt, A.E., Szuwalski, C., 2012. How well can F_{MSY} and B_{MSY} be estimated using empirical measures of surplus production? Fish. Res. 134, 113–124. https://doi.org/10.1016/j.fishres.2012.08.014.

Punt, A.E., Smith, D.C., Koopman, M.T., 2005. Using Information for 'data-rich' Species to Inform Assessments of 'data-poor' Species Through Bayesian Stock Assessment Methods. Final Report to Fisheries Research and Development Corporation Project No. 2002/094. Queenscliff. Primary Industries Research Victoria.

Punt, A.E., Smith, A.D.M., Smith, D.C., Tuck, G.N., Klaer, N.L., 2014. Selecting relative abundance proxies for B_{MSY} and B_{MEY}. ICES J. Mar. Sci. 71, 469–483.

R Development Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rice, J., 2017. . Stock Assessment Blue Shark (*Prionace Glauca*) in the Indian Ocean Using Stock Synthesis. (accessed 16 August 2017. https://www.iotc.org/sites/default/files/documents/2017/08/IOTC-2017-WPEB13-33_0.pdf.

Richard, D., Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99. https://doi.org/10.1016/j.fishres.2012.10.012.

Rudd, M.B., Branch, T.A., 2017. Does unreported catch lead to overfishing? Fish Fish. Oxf. (Oxf) 18, 313–323. https://doi.org/10.1111/faf.12181.

Schaefer, M.B., 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bull. Inter-Am. Trop. Tuna. Comm. 1, 23–56.

Schnute, J., 1981. A versatile growth model with statistically stable parameters. Can. J. Fish. Aquat. Sci. 38, 1128–1140. https://doi.org/10.1139/f81-153.

Shertzer, K.W., Conn, P.B., 2012. Spawner-recruit relationships of demersal marine fishes: prior distribution of steepness. B. Mar. Sci. 88, 39–50. https://doi.org/10.5343/bms.2011.1019.

Simpfendorfer, C.A., Bonfil, R., Latour, R., 2005. Mortality estimation. In: Musick, J.A., Bonfil, R. (Eds.), Elasmobranch Fisheries Management Techniques. FAO, Rome..

Sissenwine, M.P., Shepherd, J.G., 1987. An alternative perspective on recruitment overfishing and biological reference points. Can. J. Fish. Aquat. Sci. 44, 913–918. https://doi.org/10.1139/f87-110.

Su, Y.S., Masanao, Y., 2015. R2jags: using r to run' JAGS'. https:/CRAN.R-project.org/package = R2jags.

Takeuchi, Y., Senba, Y., Nakano, H., 2004. Demographic analysis on Atlantic blue and shortfin mako sharks. Col. Vol. Sci. Pap. ICCAT. 58, 1157–1165.

Tsai, W.P., Sun, C.L., Punt, A.E., Liu, K.M., 2014. Demographic analysis of the shortfin mako shark, *Isurus oxyrinchus*, in the Northwest Pacific using a two-sex stage-based matrix model. ICES J. Mar. Sci. 71 (7), 1604–1618. https://doi.org/10.1093/icesjms/fsu056.

Walker, T.I., 1998. Can shark resources be harvested sustainably? A question revisited with a review of shark fisheries. Mar. Freshw. Res. 49, 553–572. https://doi.org/10.1071/MF98017.

Walters, C.J., Hilborn, R., Christensen, V., 2008. Surplus production dynamics in declining and recovering fish populations. Can. J. Fish. Aquat. Sci. 65, 2536–2551. https://doi.org/10.1139/F08-170.

Wetzel, C.R., Punt, A.E., 2015. Evaluating the performance of data-moderate and catch-only assessment methods for U.S. west coast groundfsh. Fish. Res. 171, 170–187. https://doi.org/10.1016/j.fishres.2015.06.005.

Wiff, R., Flores, A., Neira, S., Caneco, B., 2018. Estimating steepness of the stock-recruitment relationship in Chilean fish stocks using meta-analysis. Fish. Res. 200, 61–67. https://doi.org/10.1016/j.fishres.2017.12.004.

Winker, H., Carvalho, F., Kapur, M., 2018. JABBA: just another Bayesian biomass assessment. Fish. Res. 204, 275–288. https://doi.org/10.1016/j.fishres.2018.03.010.

Winker, H., Carvalho, F., Thorson, J.T., Kell, L.T., Parkera, D., Kapurf, M., Sharmag, R., Boothi, A.J., Kerwath, S.E., 2020. JABBA-Select: incorporating life-history and fisheries' selectivity into surplus production models. Fish. Res. 222 (10), 53–55. https://doi.org/10.1016/j.fishres.2019.105355.

Zhang, Z., 2013. Evaluation of logistic surplus production model through simulations. Fish. Res. 140, 36–45. https://doi.org/10.1016/j.fishres.2012.12.003.

Zhu, J., Chen, Y., Dai, X., Harley, S.J., Hoyle, S.D., Maunder, M.N., Aires-da-Silva, A.M., 2012. Implications of uncertainty in the spawner–recruitment relationship for fisheries management: an illustration using bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean. Fish. Res. 119–120, 89–93. https://doi.org/10.1016/j.fishres.2011.12.008.