Original Articles

Time-varying natural mortality in fisheries stock assessment models: identifying a default approach

Kelli F. Johnson[1]*, Cole C. Monnahan[2], Carey R. McGilliard[3,4], Katyana A. Vert-pre[1,5], Sean C. Anderson[6], Curry J. Cunningham[1], Felipe Hurtado-Ferro[1], Roberto R. Licandeo[7], Melissa L. Muradian[2], Kotaro Ono[1], Cody S. Szuwalski[1], Juan L. Valero[1,8], Athol R. Whitten[1], and A. E. Punt, [1]

[1]*School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA*
[2]*Quantitative Ecology and Resource Management, University of Washington, Box 352182, Seattle, WA 98195-2182, USA*
[3]*National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA*
[4]*Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Box 955020, Seattle, WA, USA*
[5]*School of Forest Resources and Conservation, University of Florida, Box 110410, Gainesville, FL 32611, USA*
[6]*Earth to Ocean Research Group, Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada V5A 1S6*
[7]*Fisheries Centre, Aquatic Ecosystems Research Laboratory, University of British Columbia, Vancouver, BC, Canada V6T 1Z4*
[8]*Center for the Advancement of Population Assessment Methodology, 8901 La Jolla Shores Drive, La Jolla, CA 92037, USA*

*Corresponding author: tel: +1 206 543 4270; fax: +1 206 685 7471; e-mail: kfjohns@uw.edu

A typical assumption used in most fishery stock assessments is that natural mortality (M) is constant across time and age. However, M is rarely constant in reality as a result of the combined impacts of exploitation history, predation, environmental factors, and physiological trade-offs. Misspecification or poor estimation of M can lead to bias in quantities estimated using stock assessment methods, potentially resulting in biased estimates of fishery reference points and catch limits, with the magnitude of bias being influenced by life history and trends in fishing mortality. Monte Carlo simulations were used to evaluate the ability of statistical age-structured population models to estimate spawning-stock biomass, fishing mortality, and total allowable catch when the true M was age-invariant, but time-varying. Configurations of the stock assessment method, implemented in Stock Synthesis, included a single age- and time-invariant M parameter, specified at one of the three levels (high, medium, and low) or an estimated M. The min – max (i.e. most robust) approach to specifying M when it is thought to vary across time was to estimate M. The least robust approach for most scenarios examined was to fix M at a high value, suggesting that the consequences of misspecifying M are asymmetric.

Keywords: model misspecification, natural mortality, population models, reference points, simulation, Stock Synthesis, time-varying.

Introduction

Regulations for many commercial fisheries are based on the results of stock assessments. In particular, the current status of the stock, spawning-stock biomass (SSB), and fishing mortality (F) are often estimated using age- and/or size-structured fishery stock assessments fit to available data. Stock assessment methods often assume certain population characteristics [e.g. natural mortality (M) and growth] and fishery characteristics [e.g. selectivity and catchability (q)] are time-invariant. However, there is

overwhelming evidence that many ecological processes vary substantially with time for various reasons, such as fluctuations in temperature (McCarty, 2001; Walther et al., 2002; Genner et al., 2004), regime shifts (Holbrook et al., 1997; Vert-pre et al., 2013), overfishing (Pondella and Allen, 2008), changes in fishing behaviour (Hilborn and Walters, 1992), physiological processes (Ziegler et al., 2003), changes in the abundance or area inhabited by the stock (Peterman and Steer, 1981; Winters and Wheeler, 1985; Harley et al., 2001), and changes in predator–

prey dynamics (Fu *et al.,* 2001; Tyrrell *et al.,* 2011; Neira and Arancibia, 2013).

M is related to an individual's age and size, and can be highly variable over time (Zheng *et al.,* 1995; Chu *et al.,* 2008; McCoy and Gillooly, 2008). Younger fish often experience higher predation rates (Lorenzen, 1996), older fish may undergo senescence or cumulative reproductive stress (Mangel, 2003; Moustahfid *et al.,* 2009), predator–prey dynamics can change over time impacting average values of *M* (Walters and Maguire, 1996; Tyrrell *et al.,* 2011; Neira and Arancibia, 2013), and changing ocean conditions or intense fishing pressure may lead to life history adaptations (Swain, 2011; Jørgensen and Holt, 2013). However, most population models for fisheries management assume *M* is constant for all ages and across time.

M is an influential parameter in population models, potentially because its magnitude relates directly to stock productivity (Pauly, 1980; Lee *et al.*, 2011; Kenchington, 2013). Recommendations for the optimal *F* will often scale with *M*, and are therefore highly dependent on the chosen or estimated value of *M* (Andrews and Mangel, 2012). Thus, misspecification in *M* may lead to over- or underestimates of critical management quantities, such as unfished biomass (B_0), depletion (current biomass relative to B_0), maximum sustainable yield (*MSY*), and stock–recruitment relationships (Thompson, 1994; Punt and Walker, 1998; Clark, 1999). *M* is often confounded with *q*, *F*, and recruitment (Schnute and Richards, 1995), and the estimation of even a single value for *M* is difficult (Pope, 1972; Mertz and Myers, 1997), except perhaps when data are available from the initial years of the fishery (Magnusson and Hilborn, 2007). While it is difficult to estimate a single constant *M* within a stock assessment, several methods exist to estimate *M* externally to an assessment model. Typically, these methods are based on life history invariants and require extensive prior biological knowledge (Kenchington, 2013). Furthermore, these methods can be imprecise, biased, and rarely provide estimates of uncertainty (Schnute and Richards, 1995; Maunder and Wong, 2011). Currently, no default approach exists for analysts confronted with multiple external estimates of *M*. Past studies have used a variety of techniques to incorporate such estimates into stock assessments: calculate an average (Brodziak *et al.*, 2011), fit the assessment model for each estimate (Zhang and Megrey, 2006), develop Bayesian priors based on the estimates (Jiao *et al.*, 2012), and use bootstrap techniques to determine the most robust estimate (Methot and Wetzel, 2013). A few studies estimated *M* successfully within an assessment model, but this approach required a considerable amount of data, and circumstances where *M* can be estimated within the model remain unclear and controversial (Lee *et al.*, 2011; Francis, 2012).

Several methods exist to account for time-varying parameters in stock assessment models (Fournier, 1983; Shepherd and Pope, 2002; Thorson, 2011), although no single method provides unbiased estimates of terminal biomass and *F* in all situations (Radomski *et al.*, 2005; Wilberg and Bence, 2006; Linton and Bence, 2011). Simulations show that not accounting for time-varying *M* when it is present, or incorrectly specifying time-varying *M* when it is in fact stationary will lead to biases in *SSB* and other parameters (Fu and Quinn, 2000; Punt *et al.*, 2013). Furthermore, ignoring age-specific *M* often leads to smaller biases than ignoring time-varying *M*, suggesting that it is more important to account for time-varying *M* (Deroba and Schueller, 2013). The lack of a reliable benefit from adding time-varying complexity to an assessment model has led most analysts to rely on the status quo of ignoring or averaging

time-varying aspects of parameters (Fox, 1975; Clark, 1999; Maunder and Wong, 2011). Further study is needed to understand the potential impacts of ignoring time-varying effects when they are suspected to occur, but are unknown.

Simulation modelling was used here to explore the implications of various ways of handling *M* in stock assessments and to develop guidelines for when it is unclear if or how *M* varies over time. The performance of several assessment model configurations was evaluated for scenarios where the true value of *M* was either constant or changed over time. The ability of each statistical age-structured population model configuration to estimate *SSB* and *F*, as well as implications for key management quantities [e.g. the total allowable catch (*TAC*) for the next year] were compared. The results were used to identify the stock assessment configurations that were most robust to uncertainty about the value and characteristics of *M*.

Material and methods
Overview of simulation framework
Monte Carlo simulations were used to evaluate the performance of four model configurations (estimation methods; EMs) across a range of life history types, historical time-trajectories of *F*, and time-trajectories in *M*. Each combination of simulated truth (operating model; OM) and EM, hereinafter referred to as a scenario, consisted of three steps: (i) simulate a 100-year population dynamics time-series with recruitment process error, (ii) apply the EM to data sampled from the OM with observation error, and (iii) compare estimates of relevant quantities with their "true" values as defined by the OM. These steps were repeated 100 times with different process errors (recruitment deviations) and observation errors for each scenario. Larger sample sizes (e.g. 500) produced similar results, thus a sample size of 100 was chosen to keep computational time requirements within reasonable limits (Supplementary Figure S1).

Analyses were conducted using the ss3sim simulation framework (Anderson *et al.*, 2014a, b), an open-source software package implemented in the R statistical software environment (R Core Team, 2013). The integrated stock assessment framework Stock Synthesis 3 (SS3, version V3.24O; Methot and Wetzel, 2013), a statistical age-structured population modelling framework frequently used to conduct assessments in many parts of the world (see Appendix A of Methot and Wetzel, 2013), was used as both the OM and EM.

The operating models
Three configurations of the OM were used to represent three life histories: cod-like, flatfish-like, and sardine-like, i.e. life histories that are slow-growing and long-lived; fast-growing and long-lived; and fast-growing and short-lived, respectively. The life history types differed in their biological parameters (e.g. growth, *M*, age-at-maturity, etc.; see Table 1 for details), based on the following stocks: North Sea cod (*Gadus morhua*; R. Methot, NMFS, NOAA, pers. comm.), yellowtail flounder (*Limanda ferruginea*; R. Methot, NMFS, NOAA, pers. comm.), and Pacific sardine (*Sardinops sagax caeruleus*; Hill *et al.*, 2012). The Ricker stock–recruit function used by Hill *et al.* (2012) was altered to a Beverton–Holt stock–recruit function with a steepness specific to sardine (Myers *et al.*, 1999) to enable comparisons among the three life history types. All configurations of the OM were age-structured, with sexes combined.

Seven trends in *M* (Figure 1a) were investigated, where *M* either increased or decreased in a knife-edged manner to a new value in a

Table 1. Life history, fishery, and modelling parameters used for each life history type (cod-like, flatfish-like, and sardine-like).

Parameter (units)	Symbol	Value (cod, flatfish, sardine)	Estimated	Lower bounds	Upper bounds
Natural mortality (year^{-1})	M	0.2, 0.2, 0.4	At times	0.5	500
Reference age (year)	a_1	0.5, 0.5, 0.5	No	–	–
Maximum age (year)	A_{max}	25, 25, 15	No	–	–
Mean length-at-age (cm)	L_a	$L_a = L_\infty + (L_l - L_\infty)e^{-K(a-a1)}$	–	–	–
Length at a_1 (cm)	L_1	20, 12.7, 10	Always	0.5	1000
Length at A_{max} (cm)	L_∞	132, 47.4, 25	Always	0.5	1000
Growth rate (year^{-1})	K	0.2, 0.35 0.4	Always	0.5	500
CV L_1 (−)	CV_1	0.1, 0.2, 0.14	Always	0.5	500
CV L_∞ (−)	CV_∞	0.1, 0.2, 0.05	Always	0.5	500
Length−weight scaling (kg cm$^{-\beta}$)	α	6.8e−6, 1.0e−5, 1.7e−5	No	–	–
Allometric factor (−)	β	3.1, 3, 2.9	No	–	–
Fraction mature (−)	φ_l	$\varphi_l = (1 + e_1^{\Omega}(L_1 - \Omega_2))^{-1}$	–	–	–
Maturity slope (cm^{-1})	Ω_1	−0.27, −0.42, −0.90	No	–	–
Length at 50% maturity (cm)	Ω_2	38.2, 28.9, 15.9	No	–	–
Log mean virgin recruitment (−)	$ln\,R_0$	18.7, 10.5, 16.0	Always	0.5	500
Steepness (−)	h	0.65, 0.76, 0.59	No	–	–
Recruitment variability (−)	σ_r	0.4, 0.7, 0.73	No	–	–
Mean fishery length-at-50% selectivity (cm)	S_1	38.2, 28.9, 15.9	Always	0.5	500
Fishery length selectivity slope (cm)	S_2	10.6, 7.0, 3.3	Always	0.5	500
Survey length-at-50% selectivity (cm)	S_3	30.5, 23.1, 12.7	Always	0.5	500
Survey length selectivity slope (cm)	S_4	10.6, 7.0, 3.3	Always	0.5	500
Log-catchability	$ln\,q$	0, 0, 0	Always	−20	+20
Survey observation error s.d.	σ_s	0.2, 0.2, 0.2	No	–	–

Lower and upper bounds are reported as percentages, which are multiplied by the true value to determine the bounds used in the estimation model for estimated parameters, except for catchability where the actual bounds were reported. M was only estimated in a subset of the EMs.
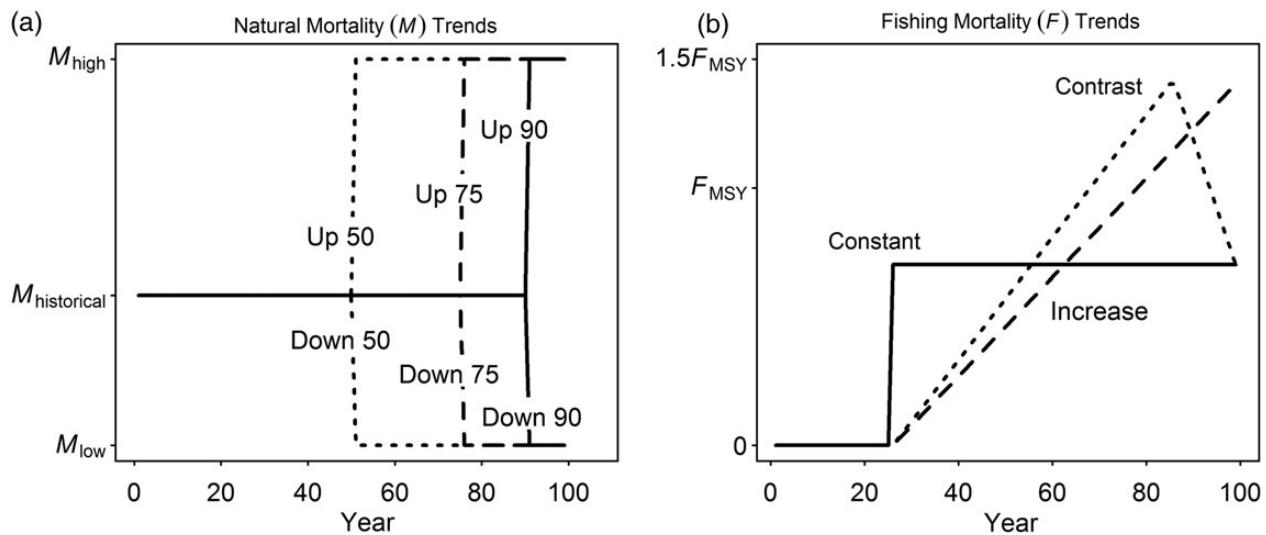

Figure 1. Natural mortality and fishing mortality trends used in the OM.

single year ("block"). Block changes were generated to represent abrupt changes due to predator–prey dynamics, such as for Chilean hake (*Merluccius gayi gayi*) where large decreases in hake are attributed to sudden influxes of jumbo squid (*Dosidicus gigas*; Arancibia and Neira, 2008) or abrupt changes due to sudden regime shifts, such as for several pandalid species in the Gulf of Alaska where abrupt increases in temperature led to large declines in abundance (Anderson, 2000).

Time-trends in M were defined relative to M in year 0 ($M_{historical}$), and M was age-invariant always. The changes in M were scaled relative to their impacts on *MSY* rather than in absolute terms to facilitate comparisons across life histories. The value to which M changed, M^*, corresponded to an *MSY* that was either 130% (M_{low}) or 70% (M_{high}) of the *MSY* corresponding to $M_{historical}$. M increased or decreased abruptly in one of the three years (year 50, 75, or 90) to M^* (Figure 1a).

Patterns in F were chosen to determine whether they changed the relative performance of models when M is time-varying, although the general influence of trajectories of F on the performance of stock assessment models has been well studied (Bence *et al.*, 1993,

Ono *et al.*, 2015). For all scenarios, years 1 through 25 had zero fishing and acted as a burn-in period. In subsequent years, fully selected *F* varied among years in one of the three ways: a constant *F* equal to the value that produced 0.95*MSY* on the left limb of the production curve ("constant"); a linear increase to the *F* corresponding to 0.95*MSY* on the right limb of the production curve ("increase"); and a 60 year linear increase to the *F* corresponding to 0.95*MSY* (right limb), followed by a 14 year linear decrease to the *F* corresponding to 0.95*MSY* (left limb; "contrast"; Figure 1b). For all scenarios, *F* in year 100 was the deterministic F_{MSY} calculated using the true value of *M* from the last year of the OM.

Fishery selectivity was length-based, mimicked the length-based maturity ogive, and was time-invariant for all life histories. It was assumed to be asymptotic to simplify the analysis and maximize the ability to estimate *M*, which is known to be difficult or impossible when selectivity is dome-shaped (He *et al.*, 2011; Lee *et al.*, 2011). For the survey, the length at which 50% of individuals were selected was specified as 80% of the length at which 50% of individuals were selected for the fishery to ensure that the survey sampled younger fish than were caught in the fishery.

Data generation

Catch was reported yearly without error from the start of the fishery to year 99 (terminal year). Fishery length- and age-composition data were generated every 10 years starting in year 25, every 5 years starting in year 45, and annually starting in year 70. Sample sizes for both the fishery length- and age-frequencies were 20 in years 25 and 35, increased by 10 from 40 to 80 between the years of 45 and 65, and were maintained at 100 beginning in year 70 to mimic an increasing emphasis on the collection of composition data. Scientific surveys occurred every other year starting in year 75, and provided an index of abundance, length-frequencies based on 100 samples, and age-frequencies based on 100 samples.

The survey composition data were generated using a multinomial distribution. This distribution assumes homogeneous capture probabilities across length or age bins and perfect mixing, which is not realistic for actual fishery composition samples (Maunder, 2011). Consequently, the Dirichlet distribution was used to generate fishery composition data with twice the standard deviation of the multinomial (i.e. over-dispersed) to more accurately reflect the information contained in actual fishery data (Aanes and Pennington, 2003; Hulson *et al.*, 2011).

Estimation methods

The EMs were based on the structure of the OM except in terms of the estimation of *M* (see Table 1 for a list of estimated parameters). The EM was configured with the correct effective sample size for both the multinomial and Dirichlet samples (e.g. for year 70, $N_{eff} = 100$ and $100/2^2$, respectively), assuring the correct statistical weight was given to the composition data (Hulson *et al.*, 2012). Bounds for all parameters were specified at extremely wide values, so that the full effects of model misspecification could be realized (Table 1). The following EMs were explored, leading to 252 scenarios (three life history types, three *F* trajectories, seven true trends in *M*, and four EM configurations):

(i) assumed to be constant and equal to $M_{historical}$ ("historical");

(ii) assumed to be constant and equal to M_{high} ("high");

(iii) assumed to be constant and equal to M_{low} ("low"); and

(iv) a single *M* was estimated ("estimated").

Sensitivity analyses

Simulations were also conducted for a series of fixed *M* values in the EM to create profiles of *M* vs. performance metrics and further evaluate the impact of misspecifying *M*. This sensitivity analysis was conducted for two cod-like scenarios: one where the true *M* was constant, as a reference case, and one where the true *M* underwent a block increase in year 75, representing a sudden regime shift or influx of predators during a period in which data on the change in *M* may be available. These scenarios were chosen because they were two of the more informative cases (i.e. produced a covariance matrix over the widest range of fixed values used for *M* in the EM). The series of assumed *M* values for both scenarios ranged from $M_{historical} - 0.07$ to $M_{historical} + 0.07$ in steps of 0.01.

Model performance

Estimation performance was evaluated using measures of relative error, $RE = (\hat{\theta} - \theta)/\theta$, and the median absolute relative error, $MARE = \text{median}(|\hat{\theta} - \theta|/\theta)$, where $\hat{\theta}$ are estimated values and θ are true values from the OM. The MARE was used to assess performance by combining bias and precision into a single metric. The median values were used because the mean values can be heavily influenced by occasional outlying estimates. REs were calculated for the time-series of *SSB* and *F* values, a forecasted *TAC* for year 100, as well as for virgin recruitment (R_0), *q*, and *M* (when it was estimated). The *TAC* for year 100 was based on the estimated F_{MSY}, where SS3 was conditioned to calculate F_{MSY} based on the biological parameters from the last year of the simulation. *SSB* and *F* from the terminal year were reported because these quantities are often important for forecasting and management.

Knowing precisely how a parameter varies across time is unrealistic. Therefore, the assessment configuration that minimized the maximum possible MARE (the min−max solution) in terminal *SSB* and *TAC* across the entire range of investigated true *M* scenarios was identified. The min−max solution identifies the stock assessment configuration that minimizes the possible losses given no information about *M*. In addition, the min−max solution was calculated for the situation where an analyst has data regarding $M_{historical}$ (e.g. from tagging data or a pre-exploitation estimate), but does not have information regarding $M^\star$.

Model fits were assessed for convergence, where convergence was defined as the production of a covariance matrix and no estimated parameters on bounds (Supplementary Table S1). A time-series was regenerated when a replicate failed to converge. It is likely that an analyst conducting assessments could have made adjustments to achieve convergence in many of these cases (e.g. changed initial values, bounds, and phases). Such manual tuning is not practical in the context of a simulation study and therefore new datasets were generated instead.

Results

Time-invariant *M* in the OM

The EMs performed well, given our chosen performance metrics (i.e. estimating terminal *SSB*, terminal *F*, *TAC*, and *M*, if estimated), when the true *M* was time-invariant and *M* was estimated or correctly specified in the EM (i.e. a self-test). The median absolute relative errors (MAREs) of terminal *SSB* and *TAC*, were close to zero when the true *M* was time-invariant and *M* was correctly specified in the EM (Tables 2 and 3, column "Constant"). The slope for survey selectivity (S_4) was poorly estimated (i.e. large MARE) when *M* was correctly specified or estimated (Table 4), particularly for cod (Table 5).

Table 2. MARE in terminal *SSB* for the cod-like life history.

		True trend in *M*						
Fishing scenario	**Estimation method**	**Constant**	**Up 50**	**Up 75**	**Up 90**	**Down 50**	**Down 75**	**Down 90**
Constant	Estimated *M*	0.30	–	–	–	–	–	–
	Fixed $M_{historical}$	0.12	–	–	–	–	–	–
Increase	Estimated *M*	0.15	0.36	0.40	0.15	0.44	*0.80*	0.15
	Fixed M_{low}	0.31	*0.47*	0.43	0.23	0.12	0.18	0.34
	Fixed $M_{historical}$	0.13	0.32	0.24	0.13	*0.44*	0.20	0.13
	Fixed M_{high}	1.04	0.18	0.36	1.23	*4.00*	2.13	1.03
Contrast	Estimated *M*	0.15	0.31	0.32	0.18	0.40	*0.45*	0.17
	Fixed M_{low}	0.27	*0.38*	0.36	0.18	0.13	0.19	0.32
	Fixed $M_{historical}$	0.11	0.26	0.19	0.13	*0.29*	0.14	0.14
	Fixed M_{high}	0.78	0.18	0.38	1.04	*1.98*	1.21	0.59

Values are reported for each of the three fishing patterns: constant, increase, and contrast. Italicized values are the scenario with the largest MAREs for each EM (largest MARE in a given row). The min–max approach compares two MAREs: (i) estimating *M* and (ii) the maximum MARE encountered when *M* is specified at M_{low}, $M_{historical}$, or M_{high}. Underlined values are the minimum MARE of the two options, indicating which EM leads to the min–max solution (the stock assessment method for which the analyst could go least wrong given no information about *M*).

Table 3. Same as Table 2, except for *TAC*.

		True trend in *M*						
Fishing scenario	**Estimation method**	**Constant**	**Up 50**	**Up 75**	**Up 90**	**Down 50**	**Down 75**	**Down 90**
Constant	Estimated *M*	0.36	–	–	–	–	–	–
	Fixed $M_{historical}$	0.13	–	–	–	–	–	–
Increase	Estimated *M*	0.20	0.48	0.55	0.24	0.58	*1.19*	0.21
	Fixed M_{low}	0.41	*0.63*	0.58	0.41	0.15	0.20	0.38
	Fixed $M_{historical}$	0.14	0.43	0.34	0.14	*0.62*	0.33	0.13
	Fixed M_{high}	1.52	0.21	0.47	1.47	*5.91*	3.22	1.73
Contrast	Estimated *M*	0.19	0.42	0.45	0.19	0.58	*0.72*	0.19
	Fixed M_{low}	0.34	*0.54*	0.51	0.33	0.13	0.21	0.34
	Fixed $M_{historical}$	0.11	0.35	0.28	0.11	*0.44*	0.22	0.11
	Fixed M_{high}	1.12	0.21	0.42	1.13	*3.00*	1.84	1.06

Table 4. MAREs in parameter estimates for the cod-like life history.

Fishing scenario	**Estimation method**	R_0	q	L_∞	K	CV_{young}	CV_{old}	S_1	S_2	S_3	S_4
Constant	Estimated *M*	0.27	0.33	0.02	0.03	0.05	0.11	0.03	0.14	0.02	0.26
	Fixed $M_{historical}$	0.06	0.10	0.02	0.04	0.05	0.11	0.03	0.14	0.01	0.28
Increase	Estimated *M*	0.13	0.12	0.02	0.03	0.05	0.10	0.02	0.10	0.02	0.29
	Fixed $M_{historical}$	0.05	0.08	0.02	0.03	0.05	0.10	0.02	0.10	0.01	0.29
Contrast	Estimated *M*	0.13	0.12	0.02	0.04	0.04	0.10	0.02	0.12	0.02	0.27
	Fixed $M_{historical}$	0.06	0.08	0.02	0.04	0.04	0.09	0.02	0.12	0.01	0.28

Values are reported for the self-test (i.e. time invariant *M* in the OM and a correctly specified EM) for each of the three fishing patterns: constant, increase, and contrast.

Table 5. MAREs in parameter estimates for all three life history types when *M* was time invariant or decreased to *M** year 50 in the OM and was misspecified at M_{high} in the EM.

Life history type	**Natural mortality**	L_∞	K	CV_{young}	CV_{old}	S_1	S_2	S_3	S_4
Cod	Constant	0.02	0.04	0.04	0.10	0.03	0.13	0.02	0.25
	Down 50	0.02	0.04	0.04	0.12	0.03	0.13	0.02	0.26
Flatfish	Constant	0.06	0.16	0.17	0.13	0.02	0.07	0.02	0.15
	Down 50	0.04	0.13	0.17	0.09	0.02	0.09	0.04	0.17
Sardine	Constant	0.02	0.05	0.05	0.16	0.01	0.04	0.03	0.07
	Down 50	0.01	0.04	0.05	0.11	0.02	0.04	0.04	0.08

Results are for the contrast fishing scenario.

The accuracy of parameter estimates for the remaining parameters depended on the fishing scenario. MAREs for growth, recruitment, and fishery selectivity parameters were close to zero for the increase and contrast fishing scenarios (Table 4). The model failed to accurately estimate *q* and R_0 on a consistent basis when *M* was estimated for the constant fishing scenario (Table 4). Specifically, 7% of the estimates of R_0 had a relative error > 500% when *M* was estimated and *F* was constant. Thus, because of the lack of information in the data under the constant fishing scenario, further results will focus on the increase and contrast fishing scenarios.

The EM estimated *SSB* and *F* more precisely when there was survey data. The median REs for *SSB* and *F* were close to zero for all years after year 74, which marked the start of the scientific

survey (Figures 2 and 3, respectively). The REs for *F* (Figure 3) were qualitatively similar, but opposite in sign compared with results for *SSB* (Figure 2). Therefore, only the results for *SSB* are discussed further below.

Trends in the results for the life histories were similar, differing only by the width of the inter-quartile ranges (Figures 2 and 3). In addition, results were similar for the increase and contrast *F* scenarios [e.g. fail to reject the null hypothesis of equal RE in *TAC* between the increase and contrast *F* scenarios when *M* was time-invariant in the OM and fixed at $M_{historical}$ in the EM; $P(0.45 > F_{1,594}) = 0.50$; Supplementary Table S2]. Therefore, only results for the contrast *F* scenario for the cod-like life history are presented when the true *M* is time-invariant.

Estimates of *M* were relatively unbiased when the true *M* was time-invariant (Figure 4, top box plot). As expected, the width of the distribution for the RE of *SSB* increased (Figure 2) and the

MARE for *TAC* increased (Table 3) when *M* was estimated rather than fixed at the true value. The distribution of the RE in *SSB* was positively skewed for these scenarios, although the median was close to zero.

Misspecifying time-invariant *M*

Misspecifying *M* when it was time-invariant in the OM led to biased estimates of *SSB* and *F*. The REs for *SSB* were negative when the true *M* was time-invariant and *M* was misspecified in the EM at a value lower than true *M* (Figure 5b) and were positive when *M* was misspecified higher than the true value (Figure 5c). The MARE for *SSB* increased with time for both of these scenarios. The REs in terminal *SSB* increased non-linearly with increasing misspecification (Figure 6a). The trend was more pronounced when *M* was fixed higher rather than lower than the true value in the EM (Figures 5c and 6a). For example, under the contrast fishing scenario, the
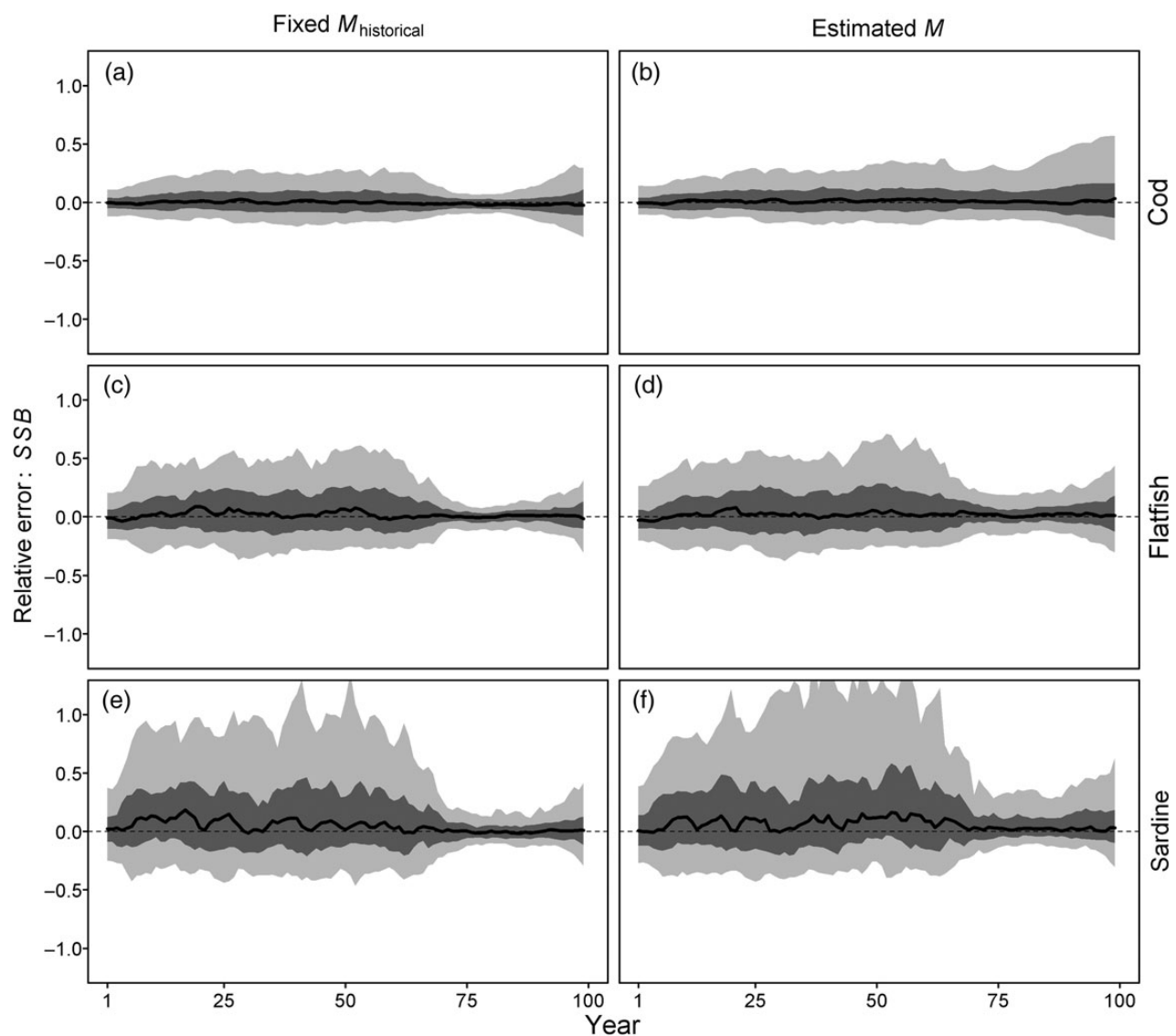

Figure 2. Time trajectories of relative error (RE) for *SSB* for the three life history types under the contrast fishing mortality scenario and when the true *M* was time-invariant. The left column depicts results when *M* was fixed at the true value and the right column depicts results when *M* was estimated. Values of zero (dashed line) indicate no difference between the OM value and the estimate from the EM. The shaded regions cover 50% and 90% of the relative errors, and the line denotes the median. Data from scientific surveys were available starting in year 75.
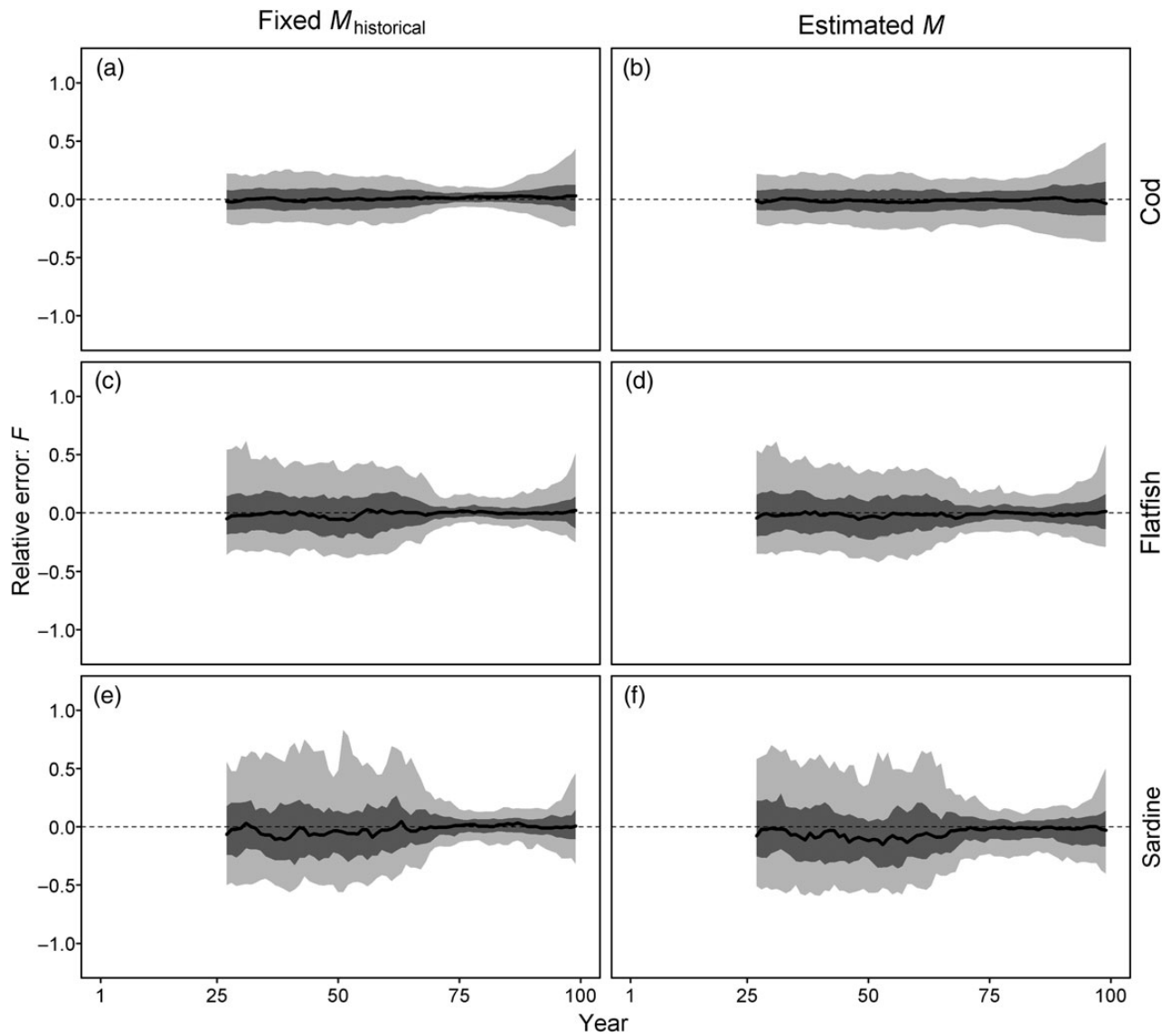
Figure 3. As for Figure 2, except results are shown for the relative error (RE) for *F*.
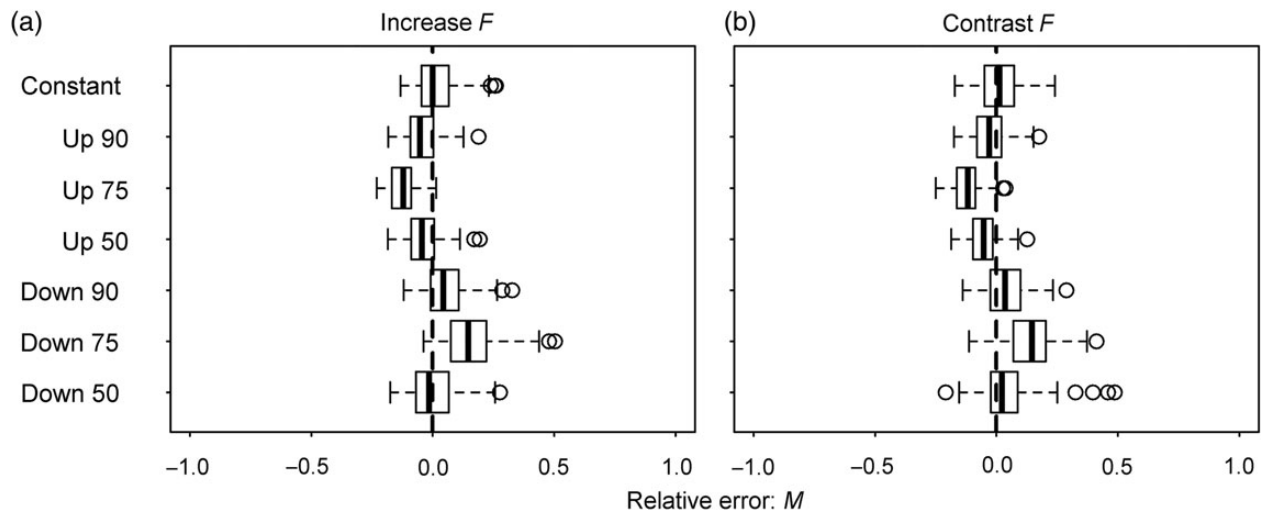

Figure 4. Boxplots for the relative errors (REs) in $M_{historical}$ for the cod-like life history.
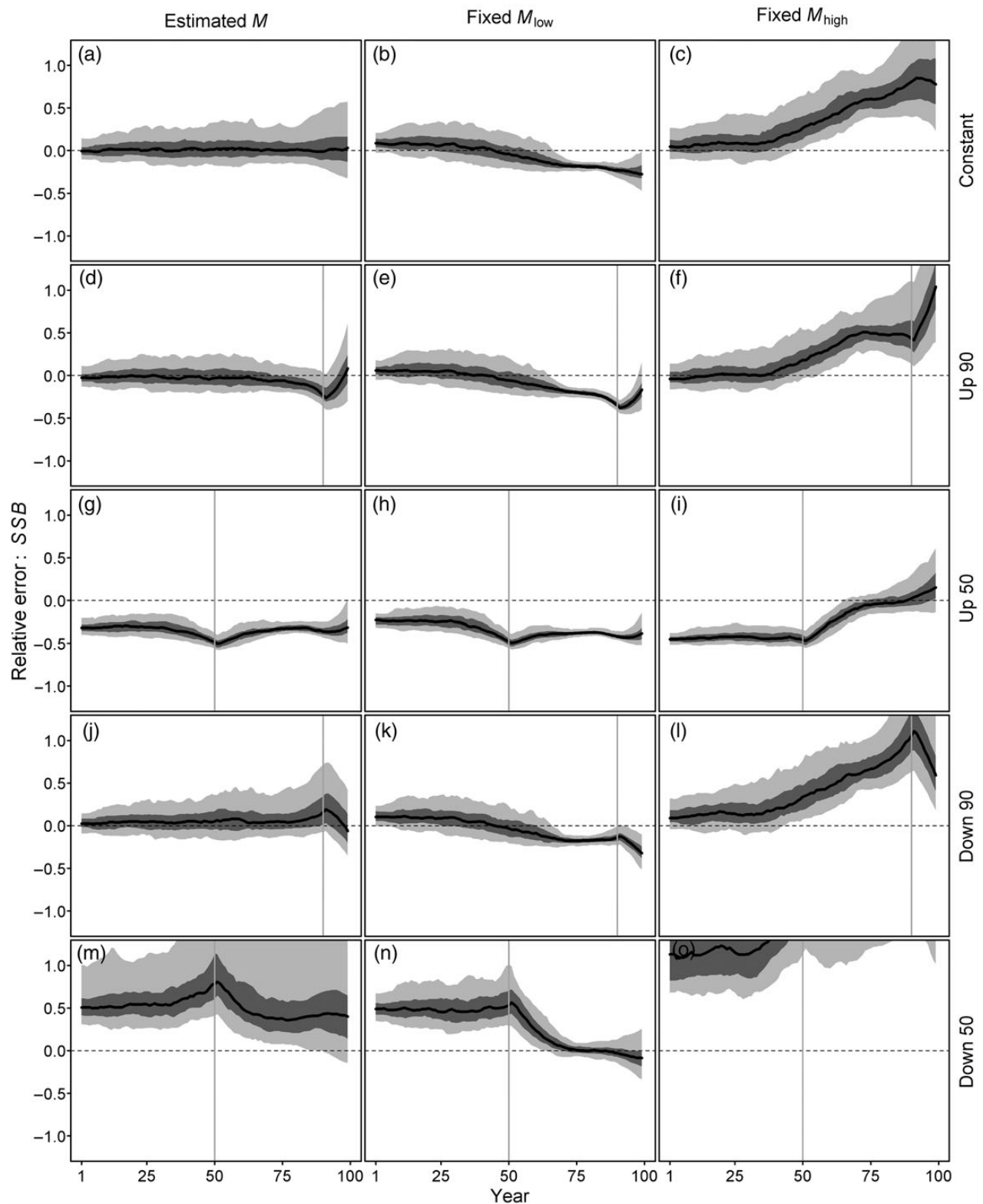
Figure 5. Time trajectories of relative error (RE) for *SSB* for the cod-like life-history type under a contrast fishing scenario when the true *M* changed in a block like fashion in years 50 or 90 or was constant across all years. *M* was either estimated, fixed at a high value, or fixed at a low value. Values of zero (dashed line) indicate no difference between the OM value and the estimate from the EM. The solid vertical line indicates the year that *M* changed to *M**. The shaded regions cover 50% and 90% of the relative errors, and the line denotes the median.
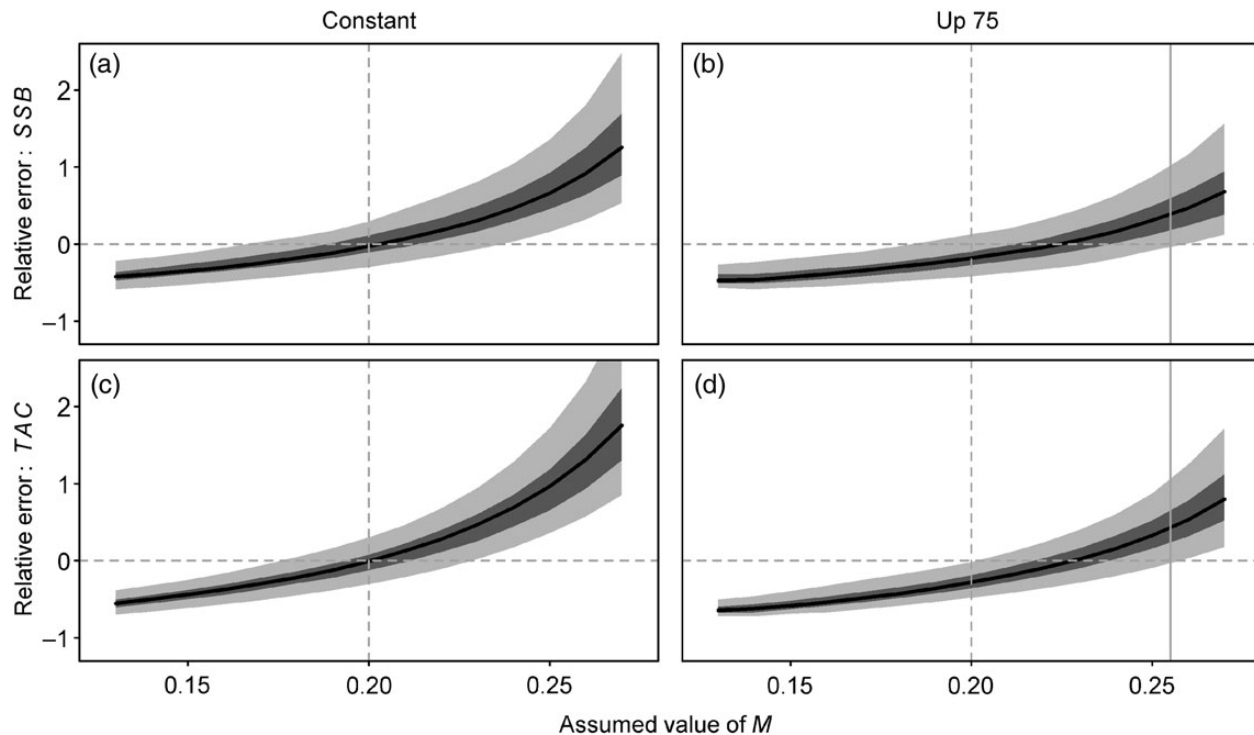
Figure 6. Relative error (RE) for terminal *SSB* and *TAC* as a function of the assumed value for *M*, when the true *M* is constant across years (left column) or increases to *M** in year 75 (right column) for the cod-like life history under the contrast *F* scenario. The value to which the true *M* changes, *M**, is denoted by a solid vertical line.

median RE in terminal *SSB* was ∼1 compared with −0.5 when *M* was specified 0.07 year^{-1} higher vs. 0.07 year^{-1} lower than the truth, respectively (Figure 6a). Patterns of RE for *TAC* (Figure 6c) were qualitatively similar to *SSB* (Figure 6a).

Misspecifying time-varying *M*

For scenarios where *M* increased or decreased to a new value (*M**) in which the EM converged, results were qualitatively similar between the life history types and thus only results for the cod-like life history are included in the main text. Plots for the flatfish- and sardine-like life histories, when the estimation model converged, are included in the Supplementary material. The EM failed to produce a covariance matrix for many scenarios involving the sardine-like life history when *M* increased early in the time-series.

Fixing *M* at a high value in the EM led to biased estimates. The 50% inter-quartile ranges for the RE for terminal *SSB* did not contain zero when *M* was fixed at a high value in the EM for any investigated trend in true *M* (Figure 5c, f, i, l, and o). The bias in terminal *SSB* (Figure 5f and i) and R_0 (Figure 7h and l) decreased when the increase to *M** occurred earlier in the time-series and *M* was fixed at a high value in the EM. In contrast, the bias in *SSB* increased as *M* decreased to *M** earlier in the time-series and *M* was fixed at a high value in the EM (Figure 5l and o). As expected, the REs for B_0 increased as the REs for terminal *SSB* decreased for these scenarios. Patterns of bias for *SSB* were similar, though opposite in sign and smaller in scale, when *M* was fixed at a low value in the EM (Figure 5).

The EM tended to underestimate $M_{historical}$ when *M** was above $M_{historical}$ and vice-versa (Figure 4). The inter-quartile ranges for the RE in *M* were generally similar for the increase (Figure 4a) and contrast fishing mortality scenarios (Figure 4b). The accuracy of

predictions for terminal *SSB* was similar, although always less accurate, when *M* was estimated compared with fixed at $M_{historical}$ in the EM for all investigated trends in *M* (Table 2). The REs in *SSB* displayed an abrupt peak or valley near the year in which the true *M* changed to *M** (Figure 5). The bias was similar before and after the abrupt change, with the largest bias in unfished and terminal *SSB* occurring when the model was misspecified for 50 years (Figure 5g and m).

Using the average of $M_{historical}$ and *M** led to the least amount of bias in terminal *SSB*. The rate of change of the median RE in terminal *SSB* with the fixed value of *M* was relatively small for values between $M_{historical}$ and *M** compared with values greater than *M** (Figure 6b). The general pattern was similar for *TAC*, although the bias was slightly more negative for values less than $M_{historical}$ than for terminal *SSB* (Figure 6d).

The EM compensated for misspecifications in *M* by adjusting *q*, a parameter that relates observed biomass to stock size, and R_0, which defines maximum recruitment. EMs for each life history type increased R_0 to account for more fish when *M* was misspecified at a high value (Figures 7 and 8, panels d, h, l, p, and t). The cod-like life history had larger REs for *q* compared with the short-lived (i.e. sardine-like) life history for the same scenario (Figures 7 and 8). Conversely, the sardine-like life history, where a larger proportion of the population is made up of new recruits compared with a long-lived fish, relied more on R_0 to account for the misspecification in *M* rather than *q*.

Min–max solution

Min–max solutions are useful to minimize the maximum loss given a range of scenarios. Here, the min–max solution specifies the stock
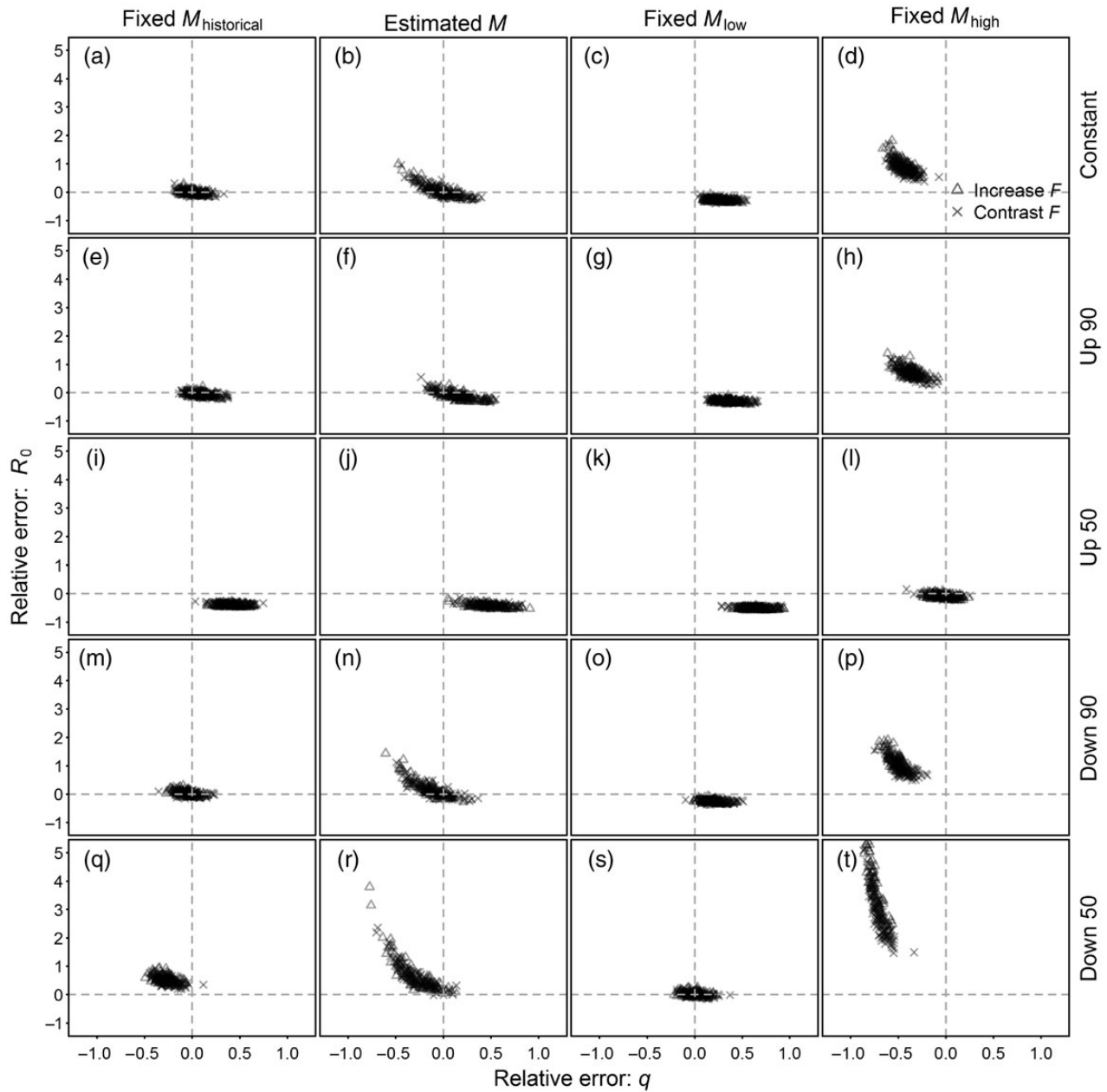
Figure 7. Correlation between the relative errors (REs) for R_0 and q for the cod-like life history for two fishing patterns. Positive errors in parameter estimates occur above and to the right of the dashed line, whereas negative errors occur below and to the left of the dashed line.

assessment configuration for which the analyst could go least wrong. Two scenarios are considered: the analyst has no prior information about M, and prior information about $M_{historical}$ is known.

For an analyst wishing to minimize the maximum RE in terminal *SSB*, with no information on how or when the true M is increasing or decreasing, the min–max solution was to estimate a single time-invariant M (Table 2). Estimating M was also the min–max solution with regard to *TAC*, although the scenario with the minimum MARE for a given EM was not necessarily the same between terminal *SSB* (Table 2) and *TAC* (Table 3).

If the analyst had information regarding the value of $M_{historical}$, the min–max approach was to fix M at $M_{historical}$, but this is only possible if an external estimate of $M_{historical}$ is available, which is

unlikely. The difference between best and worst MAREs for a given EM was smaller under the contrast fishing scenario (Tables 2 and 3).

Discussion

Estimating M, or fixing M at the long-term historical value if information about $M_{historical}$ is available, was the most robust EM across a wide array of time-trends in the true M. Fixing M higher than the true value led to the poorest performance for most trends in the true M (i.e. performance was unevenly biased when M was fixed at comparable high and low values). Although M_{high} and M_{low} were not equidistant from $M_{historical}$, results from the sensitivity analyses confirmed that fixing M at a low value can lead to less bias in
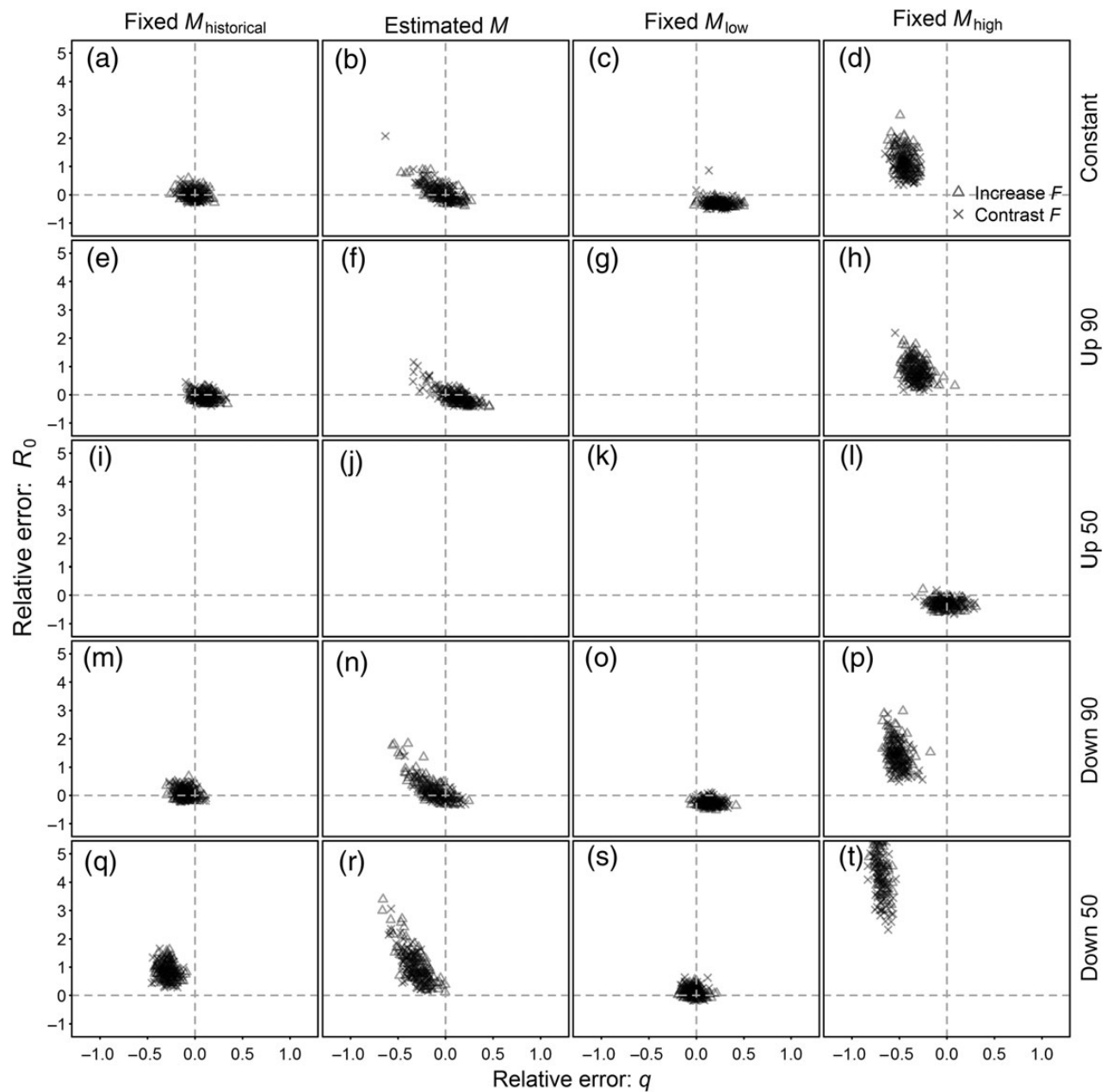
Figure 8. Correlation between the relative errors (REs) for R_0 and q for the sardine-like life history for two fishing patterns. Positive errors in parameter estimates occur above and to the right of the dashed line, whereas negative errors occur below and to the left of the dashed line. Empty plots indicate scenarios where the optimization algorithm failed to converge.

SSB than fixing *M* at a high value (Figure 6). In contrast, Deroba and Schueller (2013) found that the absolute values of the REs did not differ when *M* was fixed at symmetrically high and low values. Furthermore, they found that the REs were similar in scale when the true *M* was increasing or decreasing, although opposite in sign. Although Deroba and Schueller (2013) used a different modelling framework, the Age-Structured Assessment Program (Legault and Restrepo, 1999), the contradictory results are most likely not a result of differing frameworks. Rather, the difference in conclusions is likely a result of different quantities of data (e.g. length of time-series and sample sizes) being supplied to the EM, and differing assumptions regarding biological parameters (e.g. recruitment variability) and sampling schemes (e.g. selectivity) for each simulation.

It is well known that the estimation of *M* is heavily dependent on the quality and availability of length- or age-composition data (Beverton and Holt, 1957; Maunder and Wong, 2011). Simulations were limited to a sample size of 100 for each type of composition data in any given year, a sample size thought to be reflective of actual contemporary commercial fishery data (Haltuch *et al.,* 2011). Furthermore, simulations were provided data from just 12 scientific surveys. The choice to include a small quantity of data relative to previous studies [e.g 40 years of annual data with sample sizes of 500 per year (He *et al.,* 2011)] was made to provide realistic advice to analysts who are often provided even less data. Appropriate weighting of data sources can be challenging (Francis, 2011), and the results presented here may be subject to change if the models were tuned and

data were weighted differently (Maunder and Wong, 2011). For instance, estimates of M may be less biased if more weight is given to the age-composition data and less weight to the index of abundance from the scientific survey. Conversely, estimates of SSB, F, and TAC may be more biased if M is fixed in the EM and down-weighted indices of abundance lead to further bias in q, particularly when the true M is decreasing and M is fixed at a high value in the EM. Further research is needed to determine how weighting affects performance when M varies across time.

Quantifying the uncertainty in estimates of M whether they are estimated external to the assessment or internally can be challenging. Although the MARE was lower when M was set to $M_{\text{historical}}$, knowing $M_{\text{historical}}$ is rarely, if ever, known as precisely as was assumed here. Even the rate at which a stock is fished can change the uncertainty in estimates of M, with all else being equal. Simulations that estimated M, estimated it with varying degrees of uncertainty depending on the fishing scenario. When stocks were fished at a constant rate, some replicates estimated M and various performance metrics accurately but provided biologically implausible values of other parameters, such as q and R_0. Thus, these results provide further evidence for the lack of information in constant fishing scenarios and the value of both one- and two-way trip scenarios (Hilborn and Walters, 1992; Magnusson and Hilborn, 2007).

The EMs did not include estimating a time-varying M, although this has the potential to reduce bias in parameter estimates and better fit the data (Jiao *et al.*, 2012). Currently, the benefit of adding additional parameters to model time-varying effects is uncertain, and estimating time-varying M when the true M is constant across time can lead to biased estimates of SSB and recruitment (Fu and Quinn, 2000). In addition, multiple parameters have the potential to vary with time, and it becomes unclear which process should be accounted for and with what parameters. For example, Fu and Quinn (2000) found it is better to account for time-varying q by modelling time-varying M rather than q itself. Furthermore, it has been recommended that time-varying M only be estimated when a pronounced temporal trend is present (Fu and Quinn, 2000) and when survey data are uncorrelated (Myers and Cadigan, 1995a, b). In reality, an analyst will probably not know the true M and whether or not it varies with age or time, let alone whether or not the variation is pronounced. Nevertheless, simulations show a stationary M is estimable in SCAA models if the model is correctly specified and if the data lack patterns in the residuals (Lee *et al.*, 2011). Further research is needed to determine when it is best to include added model complexity compared with assuming M is time-invariant.

The results were not qualitatively robust to trends in the true fishing pattern. Fishing at a constant rate compared with the contrast scenario led to less information in the data and higher MAREs for all three model outputs (terminal SSB, terminal F, and TAC) for the self-test. It is well known that F is confounded with M, making it difficult to separate fishing from total mortality if catch-at-age data do not include both years of high and low fishing effort (Beverton and Holt, 1957; Hilborn and Walters, 1992). Lapointe *et al.* (1989) also note differences in estimates with varying levels of F when M is misspecified. Their results may be specific to virtual population analysis (VPA) where the errors are influenced by the ratio of F/M, with smaller ratios leading to larger errors (Vetter, 1988; Bradford and Peterman, 1989), which get increasingly larger as the VPA works backwards in time from the oldest to the youngest (Saila *et al.*, 1985). Schnute and Richards (1995) also found that fishing at a lower and more constant rate led

to higher REs for q, but no change in the estimates of M. Conversely, Deroba and Schueller (2013) found that trends in F across time had little effect on the REs of parameter estimates when the true M (age- or time-specific) trended across years.

The estimates of M and SSB were positively correlated, i.e. population size increased as the estimate of M increased to account for the greater mortality. Therefore, biases in SSB will not be consistent across a time-series when M is misspecified some of the time. Depletion, which is thought to be a robust indicator of population status (He *et al.*, 2011), will be biased in these scenarios because the biases of terminal and virgin SSB will be opposite in sign. Terminal SSB was least biased when M was fixed in the EM at the true $M_{\text{historical}}$ or estimated. There were two exceptions to this finding: (i) when M was fixed at $M^{\star}$ in the EM and the true M increased or decreased to $M^{\star}$ in year 50, or (ii) when fishing mortality increased with time, M decreased in year 75, and M was fixed at the true $M^{\star}$ value from the last 75 years of the time-series. Presumably, this is because the model obtains information about M from the initial age structure and can compensate for the misspecification in the later years by under- and overestimating other parameters such as q, F, and recruitment (Schnute and Richards, 1995).

The focus of this study was on the misspecification of M; therefore, all other parameters were correctly specified. In actuality, assessment models likely misspecify several parameters to various degrees. Assessments can produce unbiased estimates of management quantities if the misspecification acts in countering ways when more than one parameter is misspecified (Tyler *et al.*, 1989). Whether the misspecification works in a countering or parallel way largely depends on which parameters are estimated, the abundance and quality of available data, and the shape of the selectivity pattern (He *et al.*, 2011). Previous simulations show that when selectivity is dome-shaped, estimates of M are often unreliable, steepness is often estimated against its higher bound, and there can be convergence issues (He *et al.*, 2011; Lee *et al.*, 2011). Therefore, the results may not be applicable to scenarios where selectivity is dome-shaped. Consequently, results from the current study were interpreted to represent relative performance of the EMs rather than absolute performance. More work is needed to explore the effects of multiple parameter misspecification and its implications on the estimation of parameters and reference points.

Supplementary material

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Acknowledgements

Funding

References

Aanes, S., and Pennington, M. 2003. On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. ICES Journal of Marine Science, 60: 297–303.

Anderson, P. J. 2000. Pandalid shrimp as indicators of ecosystem regime shift. Journal of Northwest Atlantic Fishery Science, 27: 1–10.

Anderson, S. C., Monnahan, C. C., Johnson, K. F., Ono, K., and Valero, J. L. 2014a. ss3sim: an R package for fisheries stock assessment simulation with Stock Synthesis. PLOS ONE, doi: 10.1371/journal.pone.0092725.

Anderson, S. C., Monnahan, C. C., Johnson, K. F., Ono, K., Valero, J. L., Cunningham, C. J., Hurtado-Ferro, F., *et al.* 2014b. ss3sim: fisheries stock assessment simulation testing with Stock Synthesis. R package version 0.8.0. http://cran.r-project.org/package=ss3sim.

Andrews, K. I., and Mangel, M. 2012. Asymptotic size and natural mortality of long-lived fish for data poor stock assessments. Fisheries Research, 127–128: 45–48.

Arancibia, H., and Neira, S. 2008. Overview of the Chilean hake (*Merluccius gayi*) stock, a biomass forecast, and the jumbo squid (*Dosidicus gigas*) predator–prey relationship off central Chile (33°S–39°S). California Cooperative Oceanic Fisheries Investigations Report, 49: 104–111.

Bence, J. R., Gordoa, A., and Hightower, J. E. 1993. Influence of age-selective surveys on the reliability of Stock Synthesis assessments. Canadian Journal of Fisheries and Aquatic Sciences, 50: 827–840.

Beverton, R. J. H., and Holt, S. J. 1957. On the dynamics of exploited fish populations. Fisheries Investigations Series II, Vol. 19. Ministry of Agriculture, Fisheries, and Food, London, 553 pp.

Bradford, M. J., and Peterman, R. M. 1989. Incorrect parameter values used in virtual population analysis (VPA) generate spurious time trends in reconstructed abundances. *In* Effects of Ocean Variability on Recruitment and on Evaluation of Parameters Used in Stock Assessment, pp. 87–99. Ed. by R. J. Beamish, and G. A. McFarlane, Canadian Special Publication on Fisheries and Aquatic Science, 108. 379 pp.

Brodziak, J., Ianelli, J., Lorenzen, K., and Methot, R. D., Jr. 2011. Estimating natural mortality in stock assessment applications. U.S. Department Commerce, NOAA, Technical Memo, NMFS-F/SPO-199. 38 pp.

Chu, C. Y., Chien, H. K., and Lee, R. D. 2008. Explaining the optimality of U-shaped age-specific mortality. Theoretical Population Biology, 73: 171–180.

Clark, W. 1999. Effects of an erroneous natural mortality rate on a simple age-structured stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 56: 1721–1731.

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

Fournier, D. A. 1983. An analysis of the Hecate Strait Pacific cod fishery using an age-structured model incorporating density-dependent effects. Canadian Journal of Fisheries and Aquatic Sciences, 40: 1233–1243.

Fox, W. W. 1975. Fitting the generalized stock production model by least-squares and equilibrium approximation. Fishery Bulletin, 73: 23–37.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences, 68: 1124–1138.

Francis, R. I. C. C. 2012. The reliability of estimates of natural mortality from stock assessment models. Fisheries Research, 119: 133–134.

Fu, C., Mohn, R., and Fanning, L. P. 2001. Why the Atlantic cod (*Gadus morhua*) stock off eastern Nova Scotia has not recovered. Canadian Journal of Fisheries and Aquatic Sciences, 58: 1613–1623.

Fu, C., and Quinn, T. J., II. 2000. Estimability of natural mortality and other population parameters in a length-based model: *Pandalus borealis* in Kachemak Bay, Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 57: 2420–2432.

Genner, M. J., Sims, D. W., Wearmouth, V. J., Southall, E. J., Southward, A. J., Henderson, P. A., and Hawkins, S. J. 2004. Regional climatic warming drives long-term community changes of British marine fish. Proceedings of the Royal Society of London B: Biological Sciences, 271: 655–661.

Haltuch, M. A., Hicks, A., and See, K. 2011. Status of the U.S. petrale sole resource in 2010. Pacific Fishery Management Council, Portland, OR.

Harley, S. J., Myers, R. A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Canadian Journal of Fisheries and Aquatic Sciences, 58: 1760–1772.

He, X., Ralston, S., and MacCall, A. D. 2011. Interactions of age-dependent mortality and selectivity function in age-based stock assessment models. Fishery Bulletin, 109: 198–216.

Hilborn, R., and Walters, C. J. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. Chapman and Hall, Routledge. 570 pp.

Hill, K. T., Crone, P. R., Lo, N. C. H., Demer, D. A., Zwolinski, J. P., and Macewicz, B. J. 2012. Assessment of the Pacific sardine resource in 2012 for U.S. management in 2013. Pacific Fishery Management Council, Portland, OR.

Holbrook, S. J., Schmitt, R. J., and Stephens, J. S., Jr. 1997. Changes in an assemblage of temperate reef fishes associated with a climate shift. Ecological Applications, 7: 1299–1310.

Hulson, P-J. F., Hanselman, D. H., and Quinn, T. J., II. 2011. Effects of process and observation errors on effective sample size of fishery and survey age and length composition using variance ratio and likelihood methods. ICES Journal of Marine Science, 68: 1548–1557.

Hulson, P-J. F., Hanselman, D. H., and Quinn, T. J., II. 2012. Determining effective sample size in integrated age-structured assessment models. ICES Journal of Marine Science, 69: 281–292.

Jiao, Y., Smith, E. P., O'Reilly, R., and Orth, D. J. 2012. Modelling non-stationary natural mortality in catch-at-age models. ICES Journal of Marine Science, 69: 105–118.

Jørgensen, C., and Holt, R. E. 2013. Natural mortality: its ecology, how it shapes fish life histories, and why it may be increased by fishing. Journal of Sea Research, 75: 8–18.

Kenchington, T. J. 2013. Natural mortality estimators for information-limited fisheries. Fish and Fisheries, doi:10.1111/faf.12027.

Lapointe, M. F., Peterman, R. M., and MacCall, A. D. 1989. Trends in fishing mortality rate along with errors in natural mortality rate can cause spurious time trends in fish stock abundances estimated by virtual population analysis (VPA). Canadian Journal of Fisheries and Aquatic Sciences, 46: 2129–2139.

Lee, H-H., Maunder, M. N., Piner, K. R., and Methot, R. D., Jr. 2011. Estimating natural mortality within a fisheries stock assessment model: an evaluation using simulation analysis based on twelve stock assessments. Fisheries Research, 109: 89–94.

Legault, C. M., and Restrepo, V. R. 1999. A flexible forward age-structured assessment program. ICCAT Collective Volume of Scientific Papers, 49: 246–253.

Linton, B. C., and Bence, J. R. 2011. Catch-at-age assessment in the face of time-varying selectivity. ICES Journal of Marine Science, 68: 611–625.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology, 49: 627–642.

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

Mangel, M. 2003. Environment and longevity: the demography of the growth rate. Population and Development Review, 29: 57–70.

Maunder, M. N. 2011. Review and evaluation of likelihood functions for composition data in stock-assessment models: estimating the effective sample size. Fisheries Research, 109: 311–319.

Maunder, M. N., and Wong, R. A. 2011. Approaches for estimating natural mortality: application to summer flounder (*Paralichthys dentatus*) in the U.S. mid-Atlantic. Fisheries Research, 111: 92–99.

McCarty, J. P. 2001. Ecological consequences of recent climate change. Conservation Biology, 15: 320–331.

McCoy, M. W., and Gillooly, J. F. 2008. Predicting natural mortality rates of plants and animals. Ecology Letters, 11: 710–716.

Mertz, G., and Myers, R. A. 1997. Influence of errors in natural mortality estimates in cohort analysis. Canadian Journal of Fisheries and Aquatic Sciences, 54: 1608–1612.

Methot, R. D., Jr, and Wetzel, C. R. 2013. Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86–99.

Moustahfid, H., Tyrrell, M. C., and Link, J. S. 2009. Accounting explicitly for predation mortality in surplus production models: an application to longfin inshore squid. North American Journal of Fisheries Management, 29: 1555–1566.

Myers, R. A., Bowen, K. G., and Barrowman, N. J. 1999. Maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences, 56: 2404–2419.

Myers, R. A., and Cadigan, N. G. 1995a. Statistical analysis of catch-at-age data with correlated errors. Canadian Journal of Fisheries and Aquatic Sciences, 52: 1265–1273.

Myers, R. A., and Cadigan, N. G. 1995b. Was an increase in natural mortality responsible for the collapse of northern cod? Canadian Journal of Fisheries and Aquatic Sciences, 52: 1274–1285.

Neira, S., and Arancibia, H. 2013. Food web and fish stock changes in central Chile: comparing the roles of jumbo squid (*Dosidicus gigas*) predation, the environment, and fisheries. Deep Sea Research II: Topical Studies in Oceanography, 95: 103–112.

Ono, K., Licandeo, R., Muradian, M. L., Cunningham, C. J., Anderson, S. C., Hurtado-Ferro, F., Johnson, K. F., *et al.* 2015. The importance of length and age composition data in statistical age-structured models for marine species. ICES Journal of Marine Science, 72: 31–43.

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. Journal du Conseil International pour l'Exploration de la Mer, 39: 175–192.

Peterman, R. M., and Steer, G. J. 1981. Relation between sport-fishing catchability coefficients and salmon abundance. Transactions of the American Fisheries Society, 110: 585–593.

Pondella, D. J., and Allen, L. G. 2008. The decline and recovery of four predatory fishes from the Southern California Bight. Marine Biology, 154: 307–313.

Pope, J. G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. ICNAF Research Bulletin, 9: 65–74.

Punt, A. E., Trinnie, F., Walker, T. I., McGarvey, R., Feenstra, J., Linnane, A., and Hartmann, K. 2013. The performance of a management procedure for rock lobsters, *Jasus edwardsii*, off western Victoria, Australia in the face of non-stationary dynamics. Fisheries Research, 137: 116–128.

Punt, A. E., and Walker, T. I. 1998. Stock assessment and risk analysis for the school shark (*Galeorhinus galeus*) off southern Australia. Marine and Freshwater Research, 49: 719–731.

Radomski, P., Bence, J. R., and Quinn, T. J., II. 2005. Comparison of virtual population analysis and statistical kill-at-age analysis for a recreational, kill-dominated fishery. Canadian Journal of Fisheries and Aquatic Sciences, 62: 436–452.

R Core Team. 2013. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Saila, S. B., Lorda, E., and Walker, H. A. 1985. The analysis of parameter error propagation in simple fishery models. Marine Resource Economics, 1: 235–246.

Schnute, J. T., and Richards, L. J. 1995. The influence of error on population estimates from catch-age models. Canadian Journal of Fisheries and Aquatic Sciences, 52: 2063–2077.

Shepherd, J. G., and Pope, J. G. 2002. Dynamic pool models II: short-term and long-term forecasts of catch and biomass. *In* Handbook of Fish Biology and Fisheries, Volume 2: Fisheries, pp. 164–188. Ed. by P. J. B. Hart, and J. D. Reynolds. Blackwell Science, Oxford, 428 pp.

Swain, D. P. 2011. Life-history evolution and elevated natural mortality in a population of Atlantic cod (*Gadus morhua*). Evolutionary Applications, 4: 18–29.

Thompson, G. G. 1994. Confounding of gear selectivity and the natural mortality rate in cases where the former is a nonmonotone function of age. Canadian Journal of Fisheries and Aquatic Sciences, 51: 2654–2664.

Thorson, J. T. 2011. Auxiliary and focal assessment models: a proof-of-concept involving time-varying catchability and fishery stock-status evaluation. ICES Journal of Marine Science, 68: 2264–2276.

Tyler, A. V., Beamish, R. J., and McFarlane, G. A. 1989. Implications of age determination errors to yield estimates. *In* Effects of Ocean Variability on Recruitment and an Evaluation of Parameters Used in Stock Assessment Models. Ed. by R. J. Beamish, and G. A. McFarlane. Canadian Special Publication of Fisheries and Aquatic Sciences, 108: 27–35.

Tyrrell, M. C., Link, J. S., and Moustahfid, H. 2011. The importance of including predation in fish population models: implications for biological reference points. Fisheries Research, 108: 1–8.

Vert-pre, K. A., Amoroso, R. O., Jensen, O. P., and Hilborn, R. 2013. Frequency and intensity of productivity regime shifts in marine fish stocks. Proceedings of the National Academy of Sciences of the United States of America, 110: 1779–1784.

Vetter, E. F. 1988. Estimation of natural mortality in fish stocks: a review. Fishery Bulletin, 86: 25–43.

Walters, C., and Maguire, J. J. 1996. Lessons for stock assessment from the northern cod collapse. Reviews in Fish Biology and Fisheries, 6: 125–137.

Walther, G-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J-M., *et al.* 2002. Ecological responses to recent climate change. Nature, 416: 389–395.

Wilberg, M. J., and Bence, J. R. 2006. Performance of time-varying catchability estimators in statistical catch-at-age analysis. Canadian Journal of Fisheries and Aquatic Sciences, 63: 2275–2285.

Winters, G. H., and Wheeler, J. P. 1985. Interaction between stock area, stock abundance, and catchability coefficient. Canadian Journal of Fisheries and Aquatic Sciences, 42: 989–998.

Zhang, C-I., and Megrey, B. A. 2006. A revised Alverson and Carney model for estimating the instantaneous rate of natural mortality. Transactions of the American Fisheries Society, 135: 620–633.

Zheng, J., Murphy, M. C., and Kruse, G. H. 1995. A length-based population model and stock–recruitment relationships for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 52: 1229–1246.

Ziegler, P. E., Frusher, S. D., and Johnson, C. R. 2003. Space–time variation in catchability of southern rock lobster *Jasus edwardsii* in Tasmania explained by environmental, physiological, and density-dependent processes. Fisheries Research, 61: 107–123.

Handling editor: Mikko Heino