ICES Journal of
Marine Science

ICES Journal of Marine Science (2015), 72(1), 31–43. doi:10.1093/icesjms/fsu007

Original Article

The importance of length and age composition data in statistical age-structured models for marine species

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

[1]*School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA*
[2]*Fisheries Centre, Aquatic Ecosystems Research Laboratory (AERL), University of British Columbia, Vancouver, BC, Canada V6T 1Z4*
[3]*Quantitative Ecology and Resource Management, University of Washington, Box 352182, Seattle, WA 98195, USA*
[4]*Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada V5A 1S6*
[5]*Joint Institute of the Study of the Atmosphere and Ocean, University of Washington, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA, USA*
[6]*Center for the Advancement of Population Assessment Methodology (CAPAM), 8901 La Jolla Shores Drive, La Jolla, CA, USA*
[7]*School of Forest Resources and Conservation, University of Florida, PO Box 110240, Gainesville, FL 32611, USA*

**Corresponding author: tel: +1 206 685 3609; fax: +1 206 685 7471; e-mail: kotarono@uw.edu*

Ono, K., Licandeo, R., Muradian, M. L., Cunningham, C. J., Anderson, S. C., Hurtado-Ferro, F., Johnson, K. F., McGilliard, C. R., Monnahan, C. C., Szuwalski, C. S., Valero, J. L., Vert-pre, K. A., Whitten, A. R., and Punt, A. E. The importance of length and age composition data in statistical age-structured models for marine species. – ICES Journal of Marine Science, 72: 31–43.

Received 28 August 2013; accepted 12 January 2014; advance access publication 20 February 2014.

Management of marine resources depends on the assessment of stock status in relation to established reference points. However, many factors contribute to uncertainty in stock assessment outcomes, including data type and availability, life history, and exploitation history. A simulation–estimation framework was used to examine the level of bias and accuracy in assessment model estimates related to the quality and quantity of length and age composition data across three life-history types (cod-, flatfish-, and sardine-like species) and three fishing scenarios. All models were implemented in Stock Synthesis, a statistical age-structured stock assessment framework. In general, the value of age composition data in informing estimates of virgin recruitment (R_0), relative spawning-stock biomass (SSB_{100}/SSB_0), and terminal year fishing mortality rate (F_{100}), decreased as the coefficient of variation of the relationship between length and age became greater. For this reason, length data were more informative than age data for the cod and sardine life histories in this study, whereas both sources of information were important for the flatfish life history. Historical composition data were more important for short-lived, fast-growing species such as sardine. Infrequent survey sampling covering a longer period was more informative than frequent surveys covering a shorter period.

Keywords: age and length composition data, cod, fisheries modelling, flatfish, sardine, simulation testing, stock assessment, Stock Synthesis, survey information.

Introduction

Stock assessments produce estimates of population abundance to guide decisions regarding future regulation of harvest (Hilborn and Walters, 1992). However, stock assessments are subject to varying levels of uncertainty in current and historical abundance estimates, which must be quantified and conveyed to resource managers and policy-makers. Many factors contribute to uncertainty in stock assessment outcomes including data availability, fish life-history traits, characteristics of the fishery, the type of model used to fit available data, variability in model parameters over time, and assumptions made when conducting assessments, among others.

"Integrated" stock assessments use multiple data types to elucidate the population dynamics of focal species and estimate both model parameters and derived outputs (Maunder and Punt, 2013). Additionally, this type of model does not require complete time-series of all data sources, and can be applied to various life histories (Bence *et al.*, 1993; Punt *et al.*, 2002; Yin and Sampson, 2004; Klaer and Wayte, 2011). The data available for stock assessments often include catch (e.g. landings, discard, bycatch), effort, age/length composition data, length- or weight-at-age information, indices of abundance, maturity schedules, or tagging data (Hilborn and Walters, 1992; Quinn and Deriso, 1999). Catches, along with indices of abundance, generally provide information on the trend and scale of a population (Shepherd, 1984; Magnusson and Hilborn, 2007), while tagging and composition data can reveal changes in population structure, growth, natural mortality, and recruitment (Chen *et al.*, 2003; Magnusson and Hilborn, 2007). However, the quality of composition data may differ among data sources. Fishery-dependent information, while often relatively inexpensive to collect, is frequently less informative (has greater uncertainty) about the population of interest than directed scientific (i.e. fishery-independent) survey data, due to targeting behaviour by fishers and progressive changes in fishing gear, among other factors (Hilborn and Walters, 1992). Conversely, scientific surveys are designed to provide more accurate information about the population dynamics of the surveyed species because gear selectivity, the spatial distribution of sampling effort, and other confounding factors are controlled directly (Keller *et al.*, 2008). Thus, it must be emphasized to establish which data sources are most informative to efficiently and effectively allocate sampling effort (Ralston and Ianelli, 1998; Chen *et al.*, 2003; Rotherham *et al.*, 2007).

The quantity and quality of data required for an accurate stock assessment is a function of the species' life-history traits. Intuitively, short-lived species might require more frequent data collection than long-lived species, so cohorts can be tracked and fluctuations in recruitment strength through time estimated. On the other hand, long-lived species might require longer time-series of composition data to estimate growth for older cohorts. Most previous studies examining the importance of data quantity and quality on the efficacy of stock assessment methods have focused on a single species. For example, Chen *et al.* (2003) examined the effect of the quantity of abundance and length composition data on stock assessment of an abalone fishery, and stressed the importance of fishery-independent data in reducing bias in stock assessment outputs. Magnusson and Hilborn (2007) investigated under which conditions natural mortality rate and the steepness of the spawner–recruit relationship can be reliably estimated, and what type of catch history leads to greater estimation accuracy, focusing on Atlantic cod (*Gadus morhua*). Wetzel and Punt (2011) emphasized the importance of length composition data for fast-growing, early-maturing, and medium-lived species, in a data-limited situation. We extend these previous analyses by examining the influence of differences in data quantity and quality on the efficacy of assessment model estimation for three life-history types and three alternative catch histories with varying levels of contrast.

Stock Synthesis (SS), an integrated stock assessment framework, is currently used to assess a wide range of species around the world (see Appendix B in Methot and Wetzel, 2013, for a complete list of stocks assessed using SS). We use SS in this study as a simulation–estimation platform to address three questions: (i) What is the value of age and length composition data for various life histories? (ii) In addition to the length composition data for each life history, how much and how often should age data be collected? (iii) Which historical patterns in fishing mortality provide the most information to stock assessment methods across life histories and data quantity/quality scenarios?

Material and methods

General approach

The impacts of composition data quantity and quality, catch history, and life-history type were quantified by simulating abundance trends with appropriate process and observation error and applying SS to these generated data. This simulation–estimation approach was performed using SS as both an operating model (OM) and as an estimation model (EM). Employing SS as both OM and EM removed the potential for error due to structural differences between the OM and the EM, so that any effects of the factors under evaluation were more easily identifiable. All models presented here were implemented in Stock Synthesis 3 (SS3, version V3.24O, Methot and Wetzel, 2013) and using *ss3sim* (Anderson *et al.*, 2013, 2014), an open-source software package implemented in the R statistical software environment (R Core Team, 2013). Full details on the population dynamics model and equations can be found in Appendix A of Methot and Wetzel (2013). The simulation–estimation cycle consisted of three main steps: (i) simulate a 100-year population dynamics time-series with randomly drawn recruitment process error using the OM, (ii) apply the EM to data sampled from the OM with randomly drawn observation error in the abundance index and composition data, and (iii) repeat steps (i) and (ii) until 100 unique converged iterations were obtained. Then, estimates of relevant quantities of interest to management were compared with their "true" values as defined in the OM.

Operating models

The base model used in this study was a single-sex age-structured model without spatial structure. Three life-history types (cod-like, flatfish-like, and sardine-like) were simulated and parameterized based on the most recent assessment models for North Sea cod (*G. morhua*) (R. Methot, NMFS, NOAA, pers. comm.), yellowtail flounder (*Limanda ferruginea*) (R. Methot, pers. comm.), and Pacific sardine (*Sardinops sagax caeruleus*) (Hill *et al.*, 2013). The relationship between spawning-stock biomass (*SSB*) and recruitment for the sardine-like life history was altered from the Ricker stock–recruit function used by Hill *et al.* (2013) to a Beverton–Holt stock–recruit function with a steepness value specific to sardine (Myers *et al.*, 1999) to match the stock–recruitment relationships assumed for the other life-history types and facilitate comparison. Process error was included in the OM by adding independent, bias-corrected lognormal random deviates to the recruitment time-series with standard deviation specific to each life history (σ_R) (Anderson *et al.*, 2013).

Three patterns in fishing mortality rate, *F*, were simulated: (i) a constant *F* starting in year 25, equal to the value that produces an $F = F_{MSY}$ for each life history ("*F*1: flat *F*"), (ii) a linear increase from zero in year 25 to an $F_{high} > F_{MSY}$ that leads to a catch at equilibrium of 0.85MSY ("*F*2: fishing down"), and (iii) a linear increase from zero in year 25 to F_{high} in year 85, followed by a linear decrease for 15 years to F_{low} that is a value of *F* leading to a catch at equilibrium of 0.85MSY ("*F*3: two-way trip") (Figures 1 and 2). Selectivity was modelled as a logistic function of length in the OM. The selectivity curve for each fishery followed the maturity ogive for the corresponding life history. Survey selectivity was also assumed to be asymptotic, with the length at 50% selectivity set at

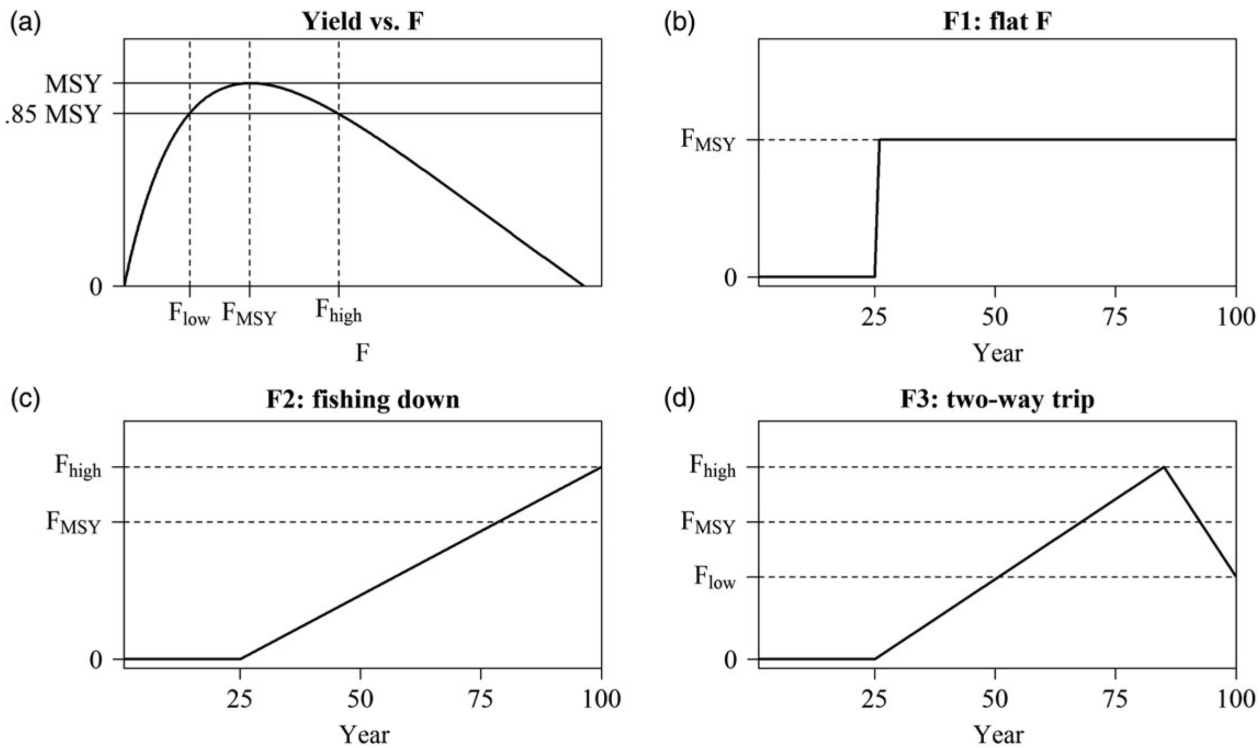

Figure 1. (a) Yield curve is for a hypothetical life history. The fishing mortality (*F*) values at MSY and 85% of MSY (F_{MSY}, F_{low}, and F_{high}; dashed vertical lines) were used to create three catch trajectories. (b–d) The three catch trajectories used in the OM.

80% of that for the fishery. A lower age at 50% selectivity allows the survey to capture smaller individuals, hence making it more informative by conferring additional information about recruitment (see selectivity curves in Haltuch *et al.*, 2013, for example).

The data provided to the EM were created by adding observation error to the "true" indices of abundance and composition data (both length and age) from the fishery and survey. The indices of abundance were generated by adding bias-corrected lognormal errors, with a log standard deviation of 0.2 or 0.4 for the fishery and survey, respectively, to the expected values of each index. Survey age and length composition data were sampled from OM-generated true values using a multinomial distribution. Fishery length and age data were sampled from a Dirichlet distribution to account for non-random, overdispersed catch-at-age or -length samples (Kitada *et al.*, 1994; Williams and Quinn, 1998; Aanes and Pennington, 2003; Hulson *et al.*, 2011). The overdispersion sample size multiplier, *c*, was set to 2 (Table 1). The effective sample sizes (N_{eff}) for the composition data were specified in the EM in terms of an equivalent number of multinomial samples to take into account the effect of the Dirichlet overdispersion,

$$N_{eff} = \frac{n}{c^2}, \quad (1)$$

where *n* is the input sample size.

Complete descriptions of the structural specifications for the cod, flatfish, and sardine OMs are outlined in Table 1.

Estimation models

Parameters that were estimated in the EM included: R_0 (virgin recruitment), the time-series of recruitment deviations, fishery and survey selectivity parameters, catchability coefficients (survey or fishery, depending on the scenario), and parameters describing somatic growth except for the coefficient of variation (*CV*) for the age–length relationship (Table 1). All other parameters (i.e. natural mortality, fecundity, length–weight relationship, recruitment variability, steepness) were fixed at their true values in the EM across data realizations.

Cases

Twelve cases that differed in frequency, quantity, and quality of index and composition data were created to examine the relative importance of composition data in informing stock assessments across life histories (Figure 3). These cases were designed to represent typical fishery situations as outlined below. The true catch time-series was available to the EM for all cases. For cases where both survey and fishery indices were available, only the survey index was supplied to the EM unless otherwise noted, mimicking current assessment practice for the US West Coast (PFMC, 2008).

Base case (Case 0)

The base case reflected a data-rich situation where the fishery began in year 25 of the 100-year simulation period and age and length composition data were collected twice between years 25 and 40, then every 5 years for the next 25 years, and then every year for the remainder of the modelled period (37 years with data in total). The survey index of abundance was available every other year starting in year 75 (13 observations). Length and age composition sample sizes for the fishery began at $n = 20$ in year 25 and gradually increased to $n = 100$ by the last 30 years (Figure 3). Sample sizes for the survey compositions were set at $n = 100$ for the entire period. This case was consistent with some US West Coast fisheries (e.g. petrale sole, Haltuch *et al.*, 2013). All subsequent cases had the same configuration as the Base case unless noted below.

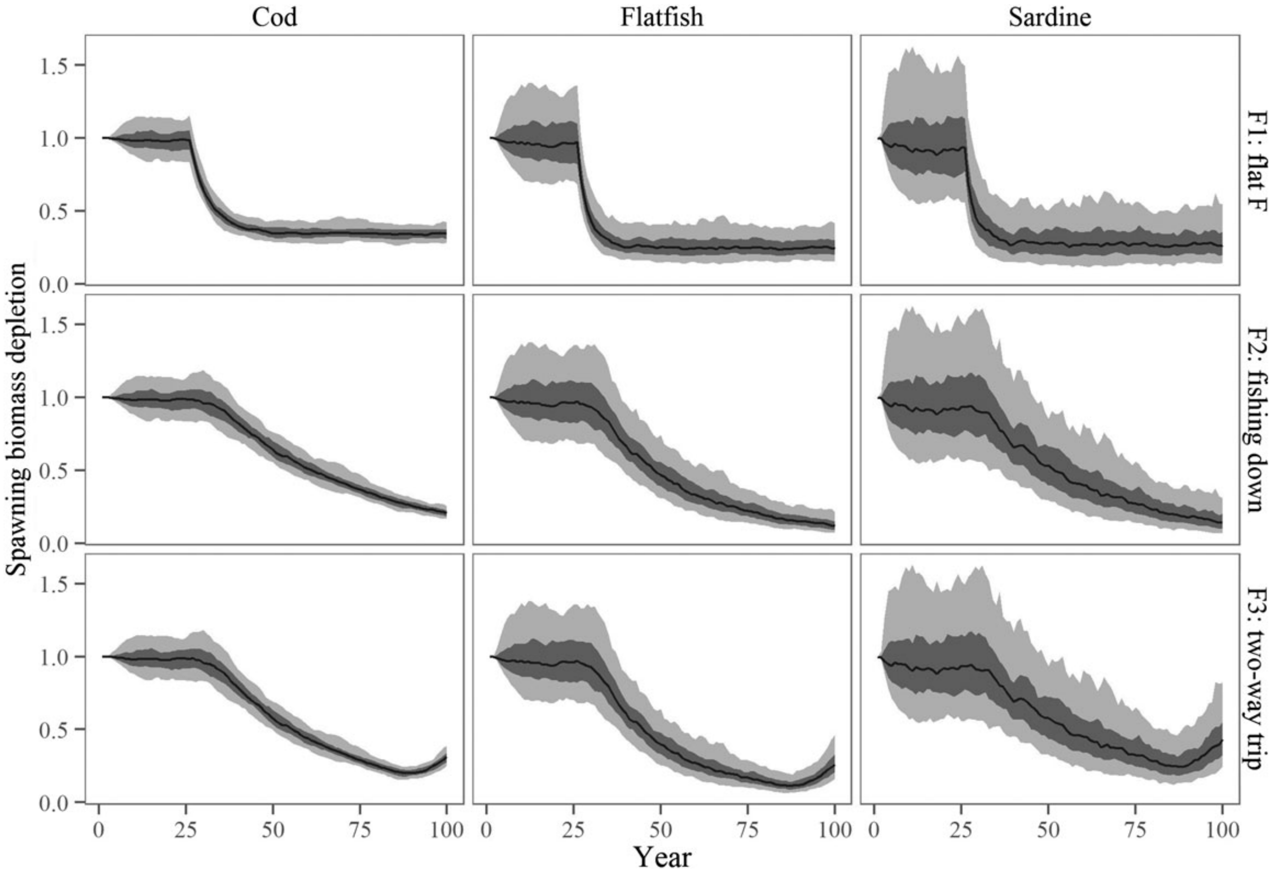

Figure 2. Time trajectories of *SSB* for the three life histories (cod, flatfish, and sardine) and for the three *F* patterns (*F*1, *F*2, and *F*3) for the OMs.

Cases with length data only

Case 1: Included only fishery index, catch, and length compositions, and mimicked a tuna-like fishery.

Case 2: Included a survey index and length compositions from both the fishery and survey, and mimicked some US West Coast groundfish fisheries.

Cases with length and age data

Case 3: Mimicked an Australian-like fishery that collected age and length composition information from the fishery only. Since this case only collected fishery data, it utilized the fishery index of abundance.

Case 4: Chosen to reflect a fishery that lacked historical age and length composition data. This case corresponded to the Base case, but with fishery and survey composition data collected every year starting in year 70.

Case 5: This case had a less frequent survey relative to the Base case. The survey was run only every fifth year starting in year 75 (five observations) to represent a situation where funding for a fishery-independent survey effort is limited.

Case 6: Similar to Case 5 with limited survey effort, however here the survey started in year 90 and ran every other year (five observations).

Case 7: This case examined the impact of sample size for composition data, and mimics a fishery where these data are difficult or expensive to collect. For the fishery, composition sample size reached a maximum of $n = 50$ in year 45 and remained at 50 for the remainder of the modelled period. Sample sizes for the survey were reduced to $n = 20$ for the entire period.

Case 8: This case added infrequently collected age composition data to Case 1. Fishery age compositions were collected every fifth year for the last 25 years of the modelled period (five observations). The fishery index was used in place of the survey index of abundance to make this scenario directly comparable with Case 1. This represented a situation where there was limited funding to age fish.

Case 9: This case added the infrequent collection of age data to Case 1. Fishery age compositions were collected every other year for the last 10 years of the modelled period (five observations). The fishery index was again used in place of the survey index of abundance.

Case 10: This case added infrequently collected age data to Case 2. Survey age compositions were collected every fifth year for the last 25 years of the modelled period (five observations).

Case 11: Similar to Case 10, but here survey age compositions were collected every other year for the last 10 years of the modelled period (five observations).

These 12 data cases were applied to the three fishing scenarios and three life-history types, leading to 108 separate scenarios.

Questions addressed by case comparisons

What is the value of age vs. length composition data?

The relative influence of age and length composition data was investigated using two comparisons. First, comparing the Base case with Case 2 demonstrated the impact on estimation performance of collecting age composition data in addition to length data from both fishery and survey. Second, comparing Case 1 with Case 3

Table 1. Values for the parameters of the OM for the three life-history types (see appendix A of Methot and Wetzel (2013) for the full description of symbols and equations, except for the last four symbols of the table).

Parameter	Symbols	Values			Estimated
		Cod	Flatfish	Sardine	
Reference age (year)	a_3	0.5	0.5	0.5	No
Plus group age (year)	A	25	25	15	No
Number of age bins	–	25	25	16	–
Number of length bins	–	45	42	45	–
Natural mortality in (year^{-1})	M	0.20	0.20	0.40	No
Steepness	h	0.65	0.76	0.59	No
Mean log virgin recruitment	$\ln R_0$	18.7	10.5	16.0	Yes
Recruitment deviation	σ_R	0.40	0.70	0.73	No
Length	l		Intermediate parameter		
Mean length-at-age a_3 (cm)	L_1	20.0	12.7	10.0	Yes
Coefficient of variation of L_1	CV_1	0.10	0.20	0.14	No
Mean asymptotic size (cm)	L_∞	132.0	47.4	25.0	Yes
Coefficient of variation of L_∞	CV_2	0.10	0.2	0.05	No
Growth coefficient (year^{-1})	k	0.20	0.35	0.40	Yes
Scaling constant for weight – length relationship	Ω_1	6.8e − 6	1.0e − 5	1.7e − 5	No
Allometric factor	Ω_2	3.1	3.0	2.9	No
Maturity slope	Ω_3	−0.27	−0.42	−0.90	No
Length-at-50% maturity (cm)	Ω_4	38.2	28.9	15.9	No
Age – length transition matrix	ϕ_a		Intermediate calculation		
Mean fishery length-at-50% selectivity (cm)	$\beta_{1,\text{fishery}}$	38.2	28.9	15.9	Yes
Fishery length selectivity slope (cm)	$\beta_{2,\text{fishery}}$	10.6	7.0	3.3	Yes
Mean survey length-at-50% selectivity (cm)	$\beta_{1,\text{survey}}$	30.5	23.5	12.7	Yes
Survey length selectivity slope (cm)	$\beta_{2,\text{survey}}$	10.6	7.0	3.3	Yes
Fishery catchability coefficient	Q_{fishery}	1.0	1.0	1.0	At times
Survey catchability coefficient	Q_{survey}	1.0	1.0	1.0	At times
Sample size for fishery composition data	n_{fishery}	20 – 100	20 – 100	20 – 100	No
Sample size for survey age/length composition data	n_{survey}	25/100	25/100	25/100	No
Coefficient of variation for the fishery index of abundance	CV_{fish}	0.4	0.4	0.4	No
Coefficient of variation for the survey index of abundance	CV_{surv}	0.2	0.2	0.2	No
Overdispersion parameter for the fishery composition data	c_{par}	2	2	2	No

highlighted the effect of collecting fishery age information in addition to fishery length composition data.

What is the influence of age composition collection frequency and duration?

Comparing Cases 8 and 9 with Case 1 examined the impact on assessment performance of collecting fishery age composition more frequently over a shorter period (Case 9) or less frequently over a longer time (Case 8), given that fishery length composition data are available. Similarly, comparison of Cases 10 and 11 with Case 2 allowed an evaluation of the impact of additional survey age composition data when length composition data were already collected from the survey and fishery.

What is the influence of frequency and duration of survey composition data?

Comparison of Cases 5 and 6 with the Base case broadened the scope of the previous question by examining whether survey composition data (both length and age) should generally be collected more frequently over a shorter period or less frequently over a longer period when fishery data were collected.

What is the impact of historical fishery composition data?

Comparison of the Base case with Case 4 allowed an evaluation of whether a historical time-series of fishery length and age compositions reduced error in estimating parameters of interest.

What is the impact of composition data sample size?

Comparison of the Base case with Case 7 addressed the question of how reducing the sample size of both fishery and survey composition (length and age) affected the performance of the EM.

Which fishing pattern produces the most informative data?

Comparing the estimation performance among fishing patterns $F1$, $F2$, and $F3$ across all scenarios evaluated which fishing pattern was most informative for estimating quantities of interest.

Performance measures

The relative spawning-stock biomass (relative SSB, which equals SSB_{100}/SSB_0), fishing mortality in the terminal year (F_{100}), and R_0 were calculated for each replicate and compared with the true values from the OM. The bias and accuracy of the EM were determined by calculating the median relative error (MRE) and the median absolute relative error (MARE) across iterations within a scenario.

$$\text{MRE} = \text{median}\left(\frac{E_{(1)} - T_{(1)}}{T_{(1)}}, \ldots, \frac{E_{(100)} - T_{(100)}}{T_{(100)}}\right), \quad (2)$$

$$\text{MARE} = \text{median}\left(\left|\frac{E_{(1)} - T_{(1)}}{T_{(1)}}\right|, \ldots, \left|\frac{E_{(100)} - T_{(100)}}{T_{(100)}}\right|\right), \quad (3)$$

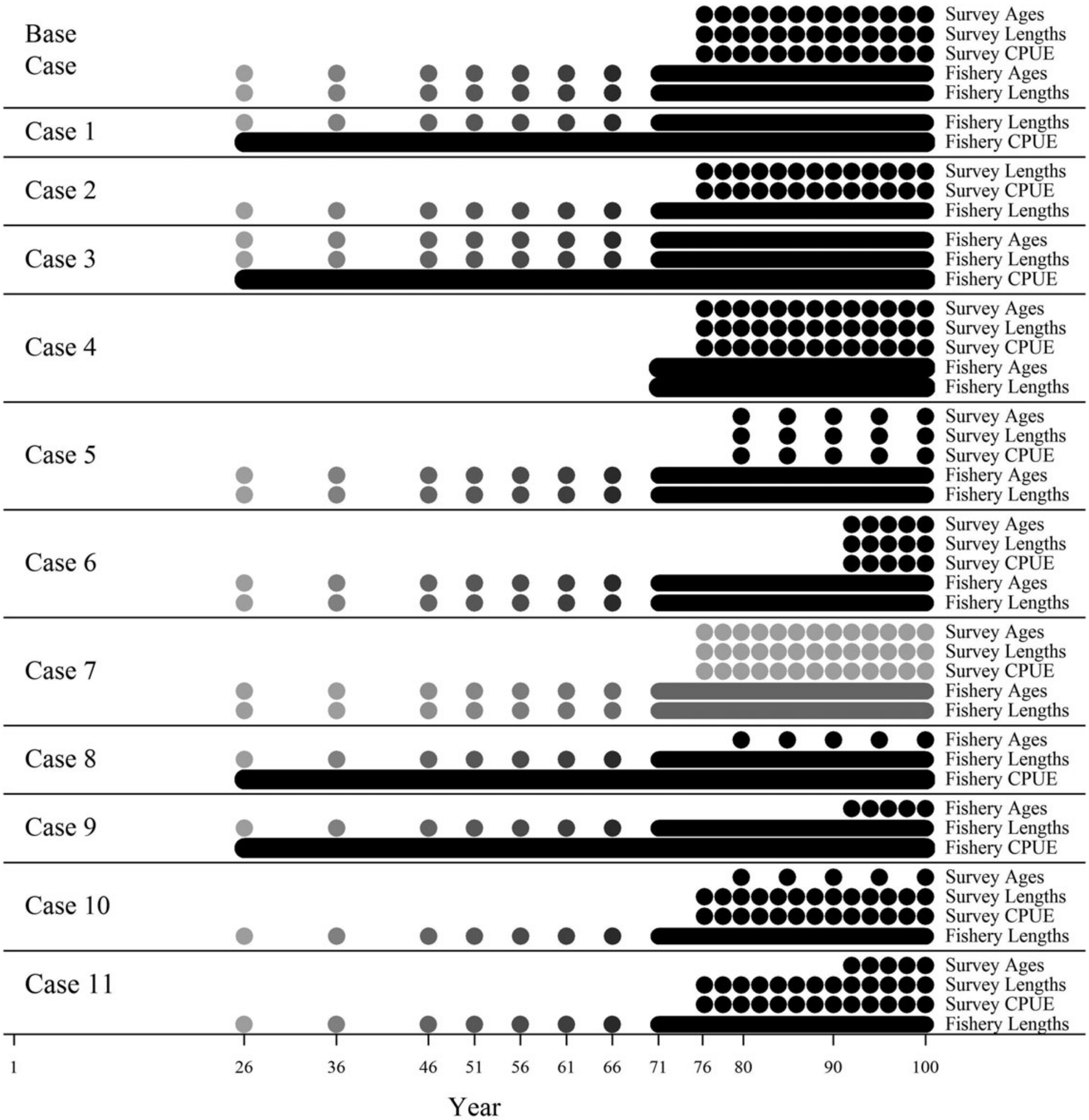

Figure 3. Presence/absence of data types and data quantity by scenario. Colour gradient corresponds to changes in sample size where lightest grey represents $n = 20$ and black represents $n = 100$. Catch data are available without error for the entire modelled period for all scenarios.

where E is the estimated quantity of interest, T the true value, and the subscript indicates the iteration number. Changes in model performance among scenarios can be evaluated by direct comparison of MRE and MARE values.

Results

Base case

Model parameters were accurately estimated (MARE at or below 16%) with good precision (low variability in relative error) and had low bias (MRE below 4%) for the Base case. R_0 was especially well estimated, with a lower MARE and higher precision (tighter 90% interval on the relative error) than for the relative *SSB* or F_{100}. Among the three species, management quantities for cod were estimated more accurately for all three parameters, and estimation performance was worst for sardine (Figure 4).

Benefit of age composition data relative to the length data

The collection of age composition data (fishery and survey) in addition to the length data generally improved estimation accuracy (i.e. MARE) and reduced bias (i.e. MRE) (Figure 4, Case 2 vs. Base). Among the three parameters of interest, the estimates of relative *SSB* and F_{100} generally improved when age composition data were available, whereas R_0 did not. Age information improved accuracy and reduced bias for the flatfish more than for the cod and

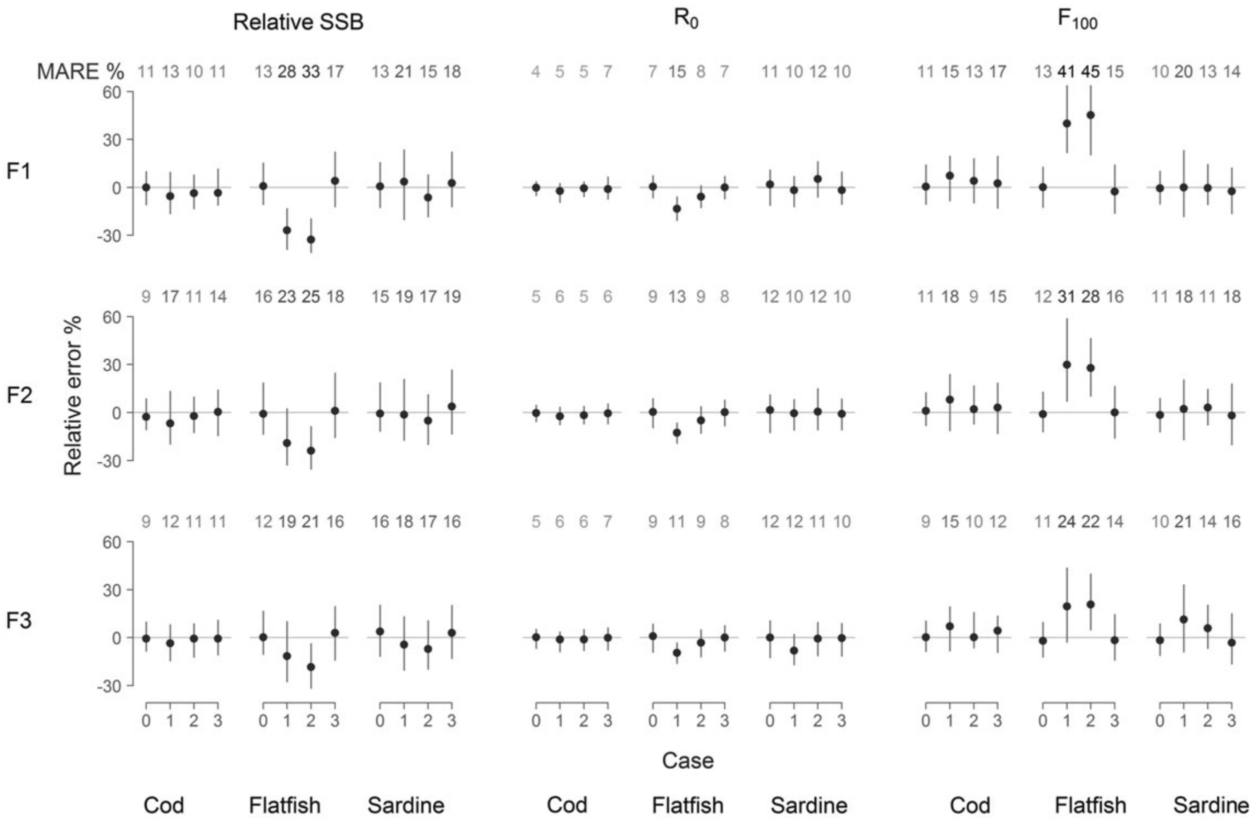

Figure 4. Estimation performance for relative *SSB*, R_0, and F_{100} by species (columns), fishing pattern (rows), and data case Base (0), 1, 2, and 3 (*x*-axes). Dots indicate the MRE and vertical lines depict the interquartile range. MARE values are printed above each case and their colour gradient indicates the level of estimation accuracy; darker grey represents higher MARE values which imply lower accuracy.

sardine. For example, the MARE of relative *SSB* and F_{100} for the flatfish decreased by 10–20 and 10–30%, respectively (across fishing scenarios). Similarly, there was, respectively, a 20–30% and 20–45% reduction in bias for these parameters. Estimation performance for the sardine and cod was much less affected by the availability of age composition data (Base case) when both fishery and survey length composition were present (Case 2); estimation accuracy increased (MARE decreased) only by 0–5% for all parameters (except R_0 for sardine, *F*2 scenario) compared with Case 2 while the effect on bias was minimal (Figure 4). However, estimation accuracy became much worse compared with the Base case, with increased bias for the cod and sardine when only fishery length data were available (Figure 4, Case 1 vs. Base). In contrast, the estimate of relative *SSB* for the flatfish life history became more accurate when only fishery length data were used without the survey length data (Figure 4, Cases 1 vs. 2).

Value of fishery age composition data: Cases 1 vs. 3

Compared with when only fishery length data were collected (Case 1), the addition of fishery age composition data (Case 3) generally improved the accuracy of relative *SSB*, R_0, and F_{100} for all species (Figure 4). Changes were largest for the flatfish model, where the accuracy of the relative *SSB*, R_0, and F_{100} improved by 5–10, 5–10, and 10–25%, respectively, and lowest for the cod where changes in accuracy were <5% (Figure 4, Cases 1 vs. 3). However, the MARE of R_0 and F_{100} for the cod increased slightly for some fishing scenarios (Figure 4). Parameter bias was also reduced when fishery age composition data were available (Figure 4). Here again, estimation performance for the flatfish improved the most, with a reduction in absolute bias of 10–20, 10–15, and 15–35% for relative *SSB*, R_0, and F_{100}, respectively. This improvement was also accompanied by a shift in the direction of bias. The estimates of relative *SSB* and R_0 were negatively biased when length were the only available composition data (Figure 4, Case 3), while the bias became positive when age data were added (Figure 4, Case 1).

Impact of frequency and duration of age composition data

Effect of frequency and duration of additional fishery age composition data: Cases 8 and 9 vs. 1

Accuracy generally improved for flatfish when some fishery age composition data in addition to fishery length composition data were available (Figure 5, Cases 8 and 9 vs. 1). However, there was not a strong signal of improvement for the cod and sardine life histories. Comparison of Case 8 with Case 9 showed that frequency and duration of age data did not markedly affect the performance metrics. Collecting age data less frequently over a longer period (Case 8) slightly improved estimation accuracy compared with more condensed sampling (Case 9) for flatfish, but did not consistently affect the estimation accuracy for cod and sardine (Figure 5). Overall, the amount of bias resulting from Cases 8 and 9 were similar, and lacked a discernible pattern across life histories (Figure 5, Cases 8 and 9 vs. 1).

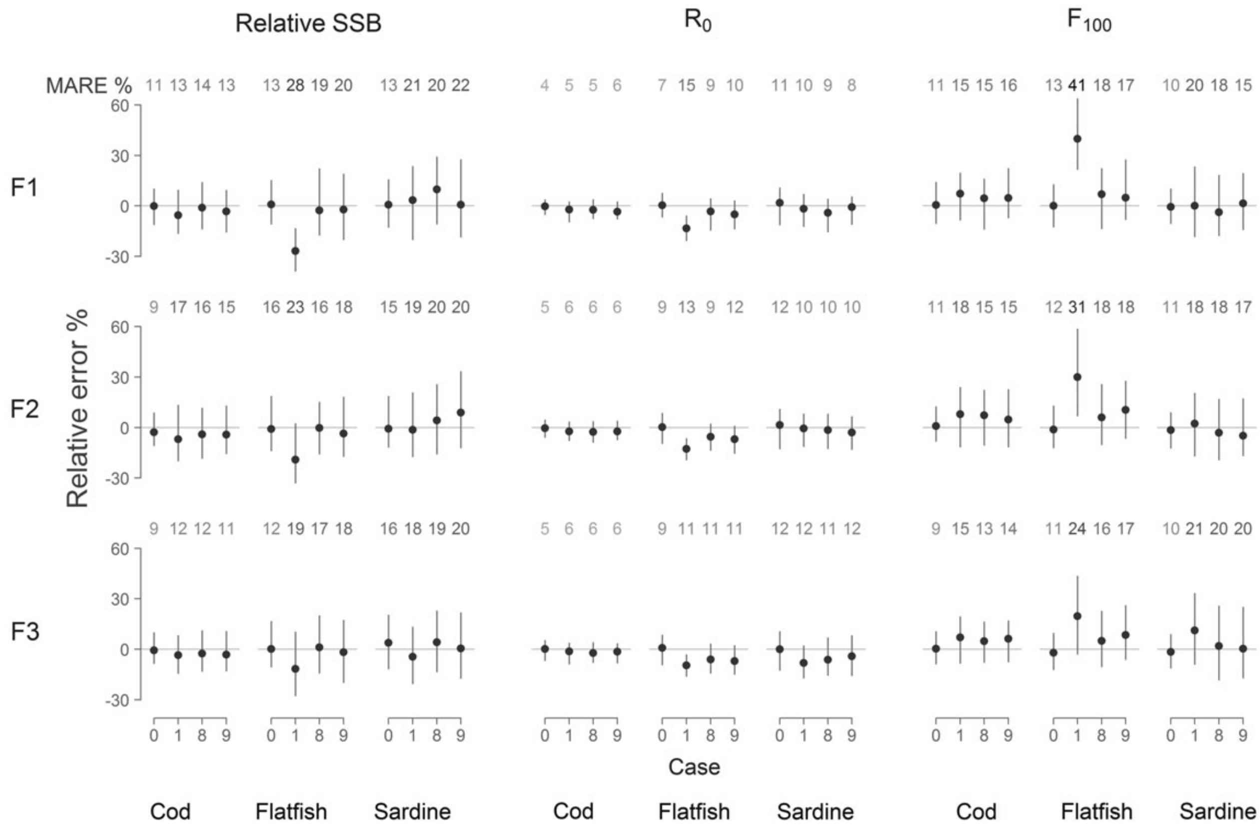

Figure 5. Estimation performance for relative *SSB*, R_0, and F_{100} by species (columns), fishing pattern (rows), and data case Base (0), 1, 8, and 9 (*x*-axes). Dots indicate the MRE and vertical lines depict the interquartile range. MARE values are printed above each case and their colour gradient indicates the level of estimation accuracy; darker grey represents higher MARE values which imply lower accuracy.

Effect of frequency and duration of additional survey age composition data: Cases 10 and 11 vs. 2

In general, accuracy was improved and bias was reduced for all life histories, particularly the flatfish, when some survey age data were provided in addition to fishery and survey length composition data (Figure 6, Cases 10 and 11 vs. 2). Some notable exceptions include estimation of R_0 for the sardine and flatfish under the *F*2 and *F*3 fishing patterns. However, the frequency and duration of age data collection did not markedly affect the performance metrics. Cases 10 and 11 displayed similar bias and accuracy across species (Figure 6), and again lacked a clear pattern for which case demonstrated better accuracy or lower bias.

Impact of frequency and duration of survey composition data: Cases 5 and 6 vs. Base

Less frequent collection of age and length survey composition data (Cases 5 and 6) compared with the Base case generally decreased accuracy in the relative *SSB* and F_{100} by 0–10% for all species (Figure 7). This decrease was particularly evident for relative *SSB* and F_{100} when survey samples were collected over the last few years (Case 6) and not spread out over a longer period (Case 5). On the other hand, accuracy for R_0 remained unchanged relative to the Base case under either survey schedule (Figure 7). Survey timing did not markedly bias any of our estimates across fishing patterns and life histories (Figure 7).

Impact of historical fishery composition data: Case 4 vs. Base

The sardine was the only life history that showed a consistent, although slight, decrease in accuracy when historical fishery composition data were not available (Case 4) compared with when they were (Base case) (Figure 8). Estimation performance (accuracy or bias) for the cod and flatfish was not strongly affected by missing historical data.

Impact of composition data sample size: Case 7 vs. Base

Estimation performance for all species did not qualitatively differ from the Base Case when the sample size of composition data was reduced throughout the entire time-series (Case 7) (Figure 8).

Information contained in the catch history

The accuracy of the parameter estimates varied between life histories and *F* patterns. Over the 12 simulated cases, for the cod and flatfish life histories, the accuracy of *SSB* and F_{100} was highest under catch scenario *F*3. In contrast, the accuracy of R_0 was highest for *F*1. Fishing pattern *F*1 led to the highest accuracy in estimating relative *SSB* and R_0 for the sardine life history and *F*2 led to the highest accuracy for F_{100}, but only by a small margin (Supplementary Table S1).

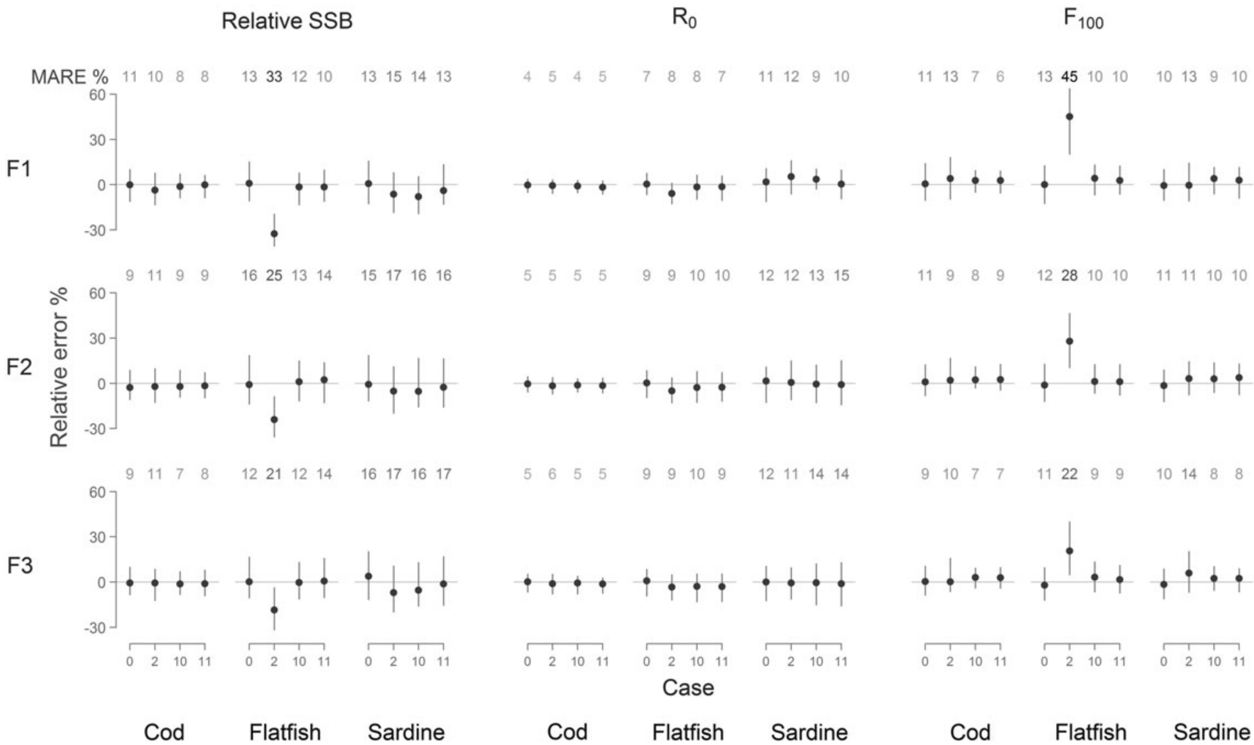

Figure 6. Estimation performance for relative *SSB*, R_0, and F_{100} by species (columns), fishing pattern (rows), and data case Base (0), 2, 10, and 11 (*x*-axes). Dots indicate the MRE and vertical lines depict the interquartile range. MARE values are printed above each case and their colour gradient indicates the level of estimation accuracy; darker grey represents higher MARE values which imply lower accuracy.

Discussion

What is the value of age and length composition data for different life-history types?

This study revealed that the value of length and age composition data differed among life histories. Age data tended to improve the accuracy of the EM for all species, but was particularly important for the flatfish model. Estimation performance for the cod and sardine was almost as good as for the Base case when only length data were available, especially if survey length data were also collected. This was mostly due to the combined effect of growth and selectivity. The cod and sardine life histories had low variation in length-at-age (Table 1), implying it was possible to track cohorts in the length composition data alone and therefore estimate growth and year-class strength. For example, a low *CV* of length-at-age for small individuals (CV_1 in Table 1) meant that the ages of small fish could be precisely determined based solely on length data, which allowed SS to estimate recruitment accurately. Additional simulations with higher *CV*s for length-at-age for the cod confirmed this and showed that the accuracy of relative *SSB* and F_{100} for the cod became similar to those for the flatfish (Supplementary Figure S1). Moreover, both fishery and survey selectivity retained smaller fish, which also helped inform estimates of recruitment, thus explaining the general increase in accuracy for R_0. However, performance for the cod and sardine models was not as accurate and unbiased as for the Base case when only fishery length composition data were available, because fishery composition data were assumed to be overdispersed.

How much and how often should these data be collected for different life-history types?

Results suggest that composition data should ideally be collected from the start of a fishery. However, this is often not possible because sampling programmes usually begin after exploitation is well underway, which could lead to a potential decrease in accuracy and an increase in bias in the estimates of relative *SSB* for short-lived, fast-growing species such as sardine. However, the impact of missing historical composition data was not considerable for the sardine and it is likely to be even less for longer-lived species, as individual cohorts are still present in the composition data from the more recent period.

A surprising result was that the sample size did not greatly affect the performance of the estimation method as long as information was available throughout the entire period that the fishery was active. Nevertheless, our results indicate that it is better to perform a survey infrequently over a longer period, than more frequently over a short period. Similar to findings by Chen *et al.* (2003) and Wetzel and Punt (2011), the relative *SSB* estimates tended to be positively biased when survey data were unavailable; adding even a small amount of survey data improved estimation performance. This is likely because (i) cohort strength is harder to identify over a short survey-sampling period and (ii) survey data are more informative for estimating recruitment than fishery data.

Collecting age data is generally expensive and time-consuming (Begg *et al.*, 2005). Therefore, it is of practical importance to determine (i) which life-history types require age compositions for

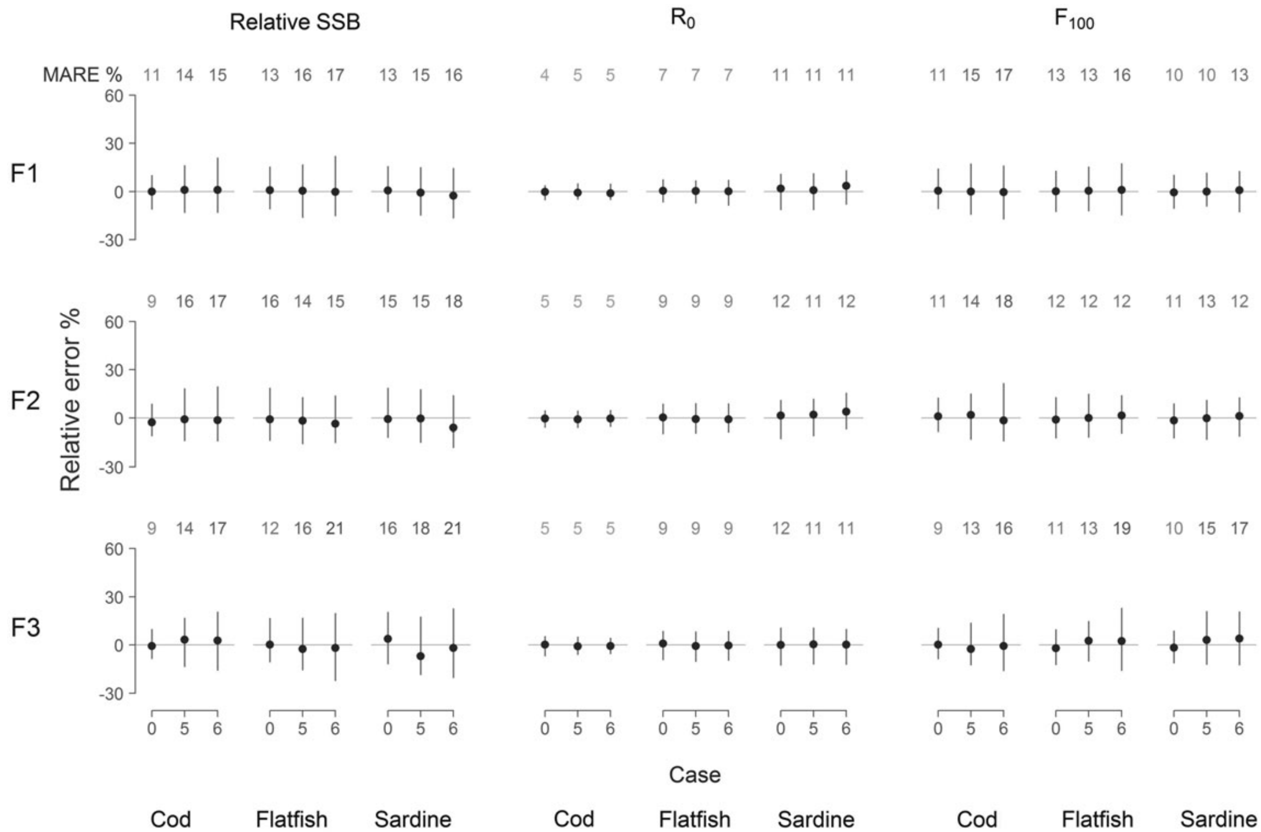

Figure 7. Estimation performance for relative *SSB*, R_0, and F_{100} by species (columns), fishing pattern (rows), and data case Base (0), 5, and 6 (*x*-axes). Dots indicate the MRE and vertical lines depict the interquartile range. MARE values are printed above each case and their colour gradient indicates the level of estimation accuracy; darker grey represents higher MARE values which imply lower accuracy.

acceptable levels of assessment accuracy and (ii) how frequently sampling should be conducted. As noted above, age data seem to be informative for the flatfish life history, and the inclusion of a few years of age compositions improved accuracy and reduced bias. This was especially true if survey age data were collected instead of the fishery age data because the latter were subject to overdispersion. However, the sampling frequency for the age data had a limited effect on both accuracy and bias. On the other hand, age composition data were not necessarily required to produce accurate estimates of management-related outputs for the cod and sardine life histories. This result likely reflects the effect of the *CV* of length-at-age data as noted previously, which highlights the importance of collecting mark-recapture data or conducting laboratory experiments that can provide information about variability in fish growth.

Information contained in the catch data

Magnusson and Hilborn (2007) found that data collected from a one-way trip (fishing down) were as informative as those collected from a two-way trip (fishing down then rebuilding) for statistical catch-at-age models. However, the current study showed that the *F* pattern which produced the most informative data depended on the life history and the quantities of interest. For example, the two-way trip scenario (*F*3) led to the most accurate estimates of relative *SSB* and F_{100} and the least accurate estimate of R_0 for flatfish and cod life histories.

Limits and future research

Several conclusions can be drawn regarding the importance of composition data for the three life histories. However, these findings depend on a number of assumptions. First, fishery composition data were generated with overdispersion to account for schooling behaviour, but without bias. In reality, fishery composition data could be subject to sources of bias. For instance, commercial vessels do not fish at random, but instead target fish of specific sizes or ages (Holland and Sutinen, 1999). Such additional sources of bias can potentially increase the importance of survey composition data compared with fishery data. Moreover, uncertainty in estimates of quantities of interest to management might become even larger if there are systematic changes in, for example, fish behaviour or environmental change (Linton and Bence, 2011; Johnson *et al.*, 2015). Additionally, fishery index data were assumed to be proportional to the true biomass. However, fishery index data can sometimes be subject to hyperstability or hyperdepletion (Hilborn and Walters, 1992; Harley *et al.*, 2001), which could increase the relative importance of survey data with respect to the fishery data. Lastly, many parameters in this study were fixed at their true values, which is an advantage many analysts lack in real assessment situations.

Second, age data were simulated without error in this study, while in reality, they are subject to reading error (Begg *et al.*, 2005). Future work could examine the levels of error at which age data become less valuable for certain species, such as flatfish, which require age data for accurate parameter estimation.

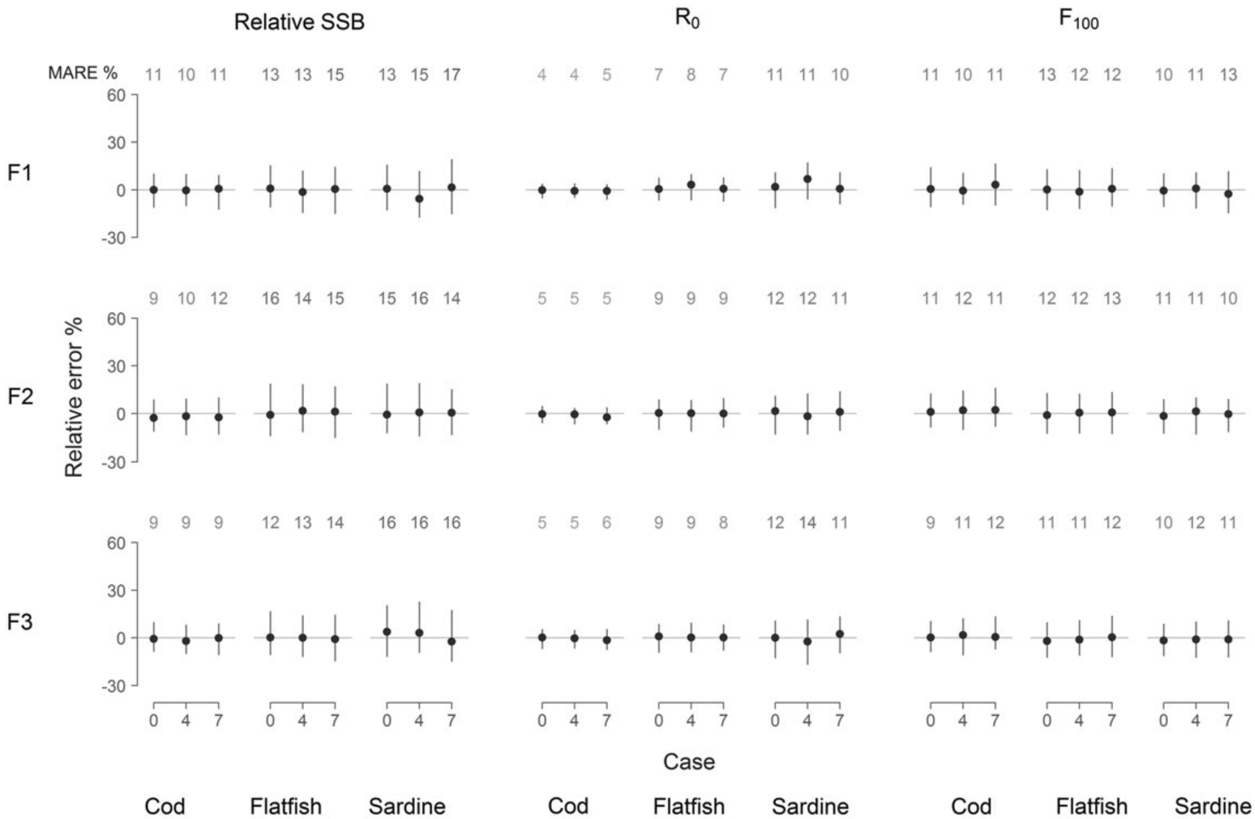

Figure 8. Estimation performance for relative *SSB*, R_0, and F_{100} by species (columns), fishing pattern (rows), and data case Base (0), 4 and 7 (x-axes). Dots indicate the MRE and vertical lines depict the interquartile range. MARE values are printed above each case and their color gradient indicates the level of estimation accuracy; darker grey represents higher MARE values which imply lower accuracy.

Third, in recent years, it has become commonly accepted that time-varying selectivity is likely occurring in most fisheries (Sampson and Scott, 2011, 2012). Additional simulations that incorporated time-varying fishery selectivity in the OM showed that the general findings of this study still hold under such circumstances (Supplementary Figures S2–S6). The lower effective sample size for the fishery composition data in these simulations, due to the overdispersion, reduced the impact of the fishery composition data to the total likelihood of the model and thus alleviates some of the potential bias introduced by the time-varying selectivity. A higher effective sample size for the fishery data confirmed these results and the Base case results became biased (Supplementary Figure S7). The latter highlights the importance of properly weighting the influence of different datasets (Francis, 2011). In addition to simulating time-varying data, future work could also consider estimating time-varying selectivity when assessing the quantity and quality of data required to accurately estimate model outputs of interest and the temporal trend or variability in selectivity parameters. Such future work could take advantage of the recent methodologies for performing model selection across alternative ways of modelling time-varying selectivity processes (Linton and Bence, 2011; Punt *et al.*, 2013).

Although the focus of this paper was on data quantity and quality, there are other ways to increase the accuracy of information provided by assessments. Intuitively, the addition of informative data should reduce bias and increase accuracy of parameter estimates, but there remain some parameters that are inherently difficult to estimate, e.g. natural mortality rate and steepness (Schnute and Richards, 1995; Magnusson and Hilborn, 2007; Conn *et al.*, 2010; Maunder and Wong, 2011; Maunder, 2012). The use of prior information for such parameters would be a useful way to incorporate additional information into assessments potentially provided by conducting mark recapture experiments (Whitlock *et al.*, 2012), applying meta-analytic methods (Myers *et al.*, 1995; Michielsens and McAllister, 2004) or using life-history theory (Forrest *et al.*, 2008).

Supplementary material

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

Acknowledgements

This work was partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative agreement No. NA10OAR4320148, Contribution No. 2193 and 2195. RL was supported by Conicyt. MLM was funded by Exxon Valdez Oil Spill Trustee Council, grant 13120111-Q. SCA was supported by Fulbright Canada and NSERC. We thank two anonymous reviewers and Shijie Zhou for their comments on earlier versions of the manuscript and Richard Methot, Jim Ianelli, and Ian Taylor (NOAA Fisheries) for discussions and advice. This research addresses the good practices in stock assessment modelling program of the Center for the Advancement of Population Assessment Methodology (CAPAM).

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 S. C., Monnahan C. C., Johnson K. F., Ono K., and Valero J. L. 2013. ss3sim: An R package for fisheries stock assessment simulation with Stock Synthesis. arXiv:1312.6450.

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

Begg G. A., Campana S. E., Fowler A. J., and Suthers I. M. 2005. Otolith research and application: current directions in innovation and implementation. Marine and Freshwater Research, 56: 477.

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

Chen Y., Chen L., and Stergiou K. I. 2003. Impacts of data quantity on fisheries stock assessment. Aquatic Sciences—Research Across Boundaries, 65: 92–98. http://link.springer.com/10.1007/s000270300008.

Conn P. B., Williams E. H., and Shertzer K. W. 2010. When can we reliably estimate the productivity of fish stocks? Canadian Journal of Fisheries and Aquatic Sciences, 67: 1–13.

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

Forrest R. E., Martell S. J. D., Melnychuk M. C., and Walters C. J. 2008. An age-structured model with leading management parameters, incorporating age-specific selectivity and maturity. Canadian Journal of Fisheries and Aquatic Sciences, 65: 286–296.

Haltuch M. A., Ono K., and Valero J. 2013. Status of the U.S. petrale sole resource in 2012. Pacific Fishery Management Council, Portland, OR. http://www.pcouncil.org/

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.

Hilborn R., and Walters C. J. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, New York.

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

Holland D. S., and Sutinen J. G. 1999. An empirical model of fleet dynamics in New England trawl fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 56: 253–264.

Hulson P. J. F., Hanselman D. H., and Quinn T. J. 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. http://icesjms.oxfordjournals.org/cgi/doi/10.1093/icesjms/fsr102.

Johnson K. F., Monnahan C. C., McGilliard C. R., Vert-pre K. A., Anderson S. C., Cunningham C. J., Hurtado-Ferro F., *et al.* 2015. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. ICES Journal of Marine Science, 72: 137–150.

Keller A. A., Horness B. H., Fruh E. L., Simon V. H., Tuttle V. J., Bosley K. L., Buchanan J. C., *et al.* 2008. The 2005 U.S. West coast bottom trawl survey of groundfish resources off Washington, Oregon, and California: Estimates of distribution, abundance, and length composition. Seattle.

Kitada S., Hiramatsu K., and Kishino H. 1994. Estimating mortality rates from tag recoveries: incorporating over-dispersion, correlation, and change points. ICES Journal of Marine Science, 51: 241–251.

Klaer N., and Wayte S. 2011. Demersal MSE for trawl fish in the Southern and Eastern Scalefish and Shark Fishery and other like-species. Hobart, Australia.

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. http://icesjms.oxfordjournals.org/cgi/doi/10.1093/icesjms/fsq173.

Magnusson A., and Hilborn R. 2007. What makes fisheries data informative? Fish and Fisheries, 8: 337–358. http://doi.wiley.com/10.1111/j.1467-2979.2007.00258.x.

Maunder M. N. 2012. Evaluating the stock–recruitment relationship and management reference points: application to summer flounder (Paralichthys dentatus) in the U.S. mid-Atlantic. Fisheries Research, 125–126: 20–26.

Maunder M. N., and Punt A. E. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research, 142: 61–74.

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.

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

Michielsens C. G. J., and McAllister M. K. 2004. A Bayesian hierarchical analysis of stock–recruit data: quantifying structural and parameter uncertainties. Fisheries Research, 1047: 1032–1047.

Myers R. A., Barrowman N. J., Hutchings J. A., and Rosenberg A. A. 1995. Population-dynamics of exploited fish stocks at low population-levels. Science, 269: 1106–1108.

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. http://www.nrcresearchpress.com/doi/abs/10.1139/f99-201.

PFMC. 2008. Pacific coast groundfish fishery assessment and fishery evaluation. Portland, OR. http://www.pcouncil.org/

Punt A. E., Hurtado-Ferro F., and Whitten A. R. 2013. Model selection for selectivity in fisheries stock assessments. Fisheries Research. http://dx.doi.org/10.1016/j.fishres.2013.06.003.

Punt A. E., Smith A. D. M., and Cui G. R. 2002. Evaluation of management tools for Australia's South East fishery 1. Modelling the south east fishery taking account of technical interactions. Marine and Freshwater Research, 53: 615–629.

Quinn T. J., and Deriso R. B. 1999. Quantitative Fish Dynamics. Oxford University Press, Oxford, UK.

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

Ralston S., and Ianelli J. N. 1998. When lengths are better than ages: the complex case of Bocaccio. Fishery Stock Assessment Models, AK-SG-98-01: 451–468.

Rotherham D., Underwood A. J., Chapman M. G., and Gray C. A. 2007. A strategy for developing scientific sampling tools for fishery-independent surveys of estuarine fish in New South Wales. Australia, 1512–1516.

Sampson D. B., and Scott R. D. 2011. A spatial model for fishery age-selection at the population level. Canadian Journal of Fisheries and Aquatic Sciences, 68: 1077–1086.

Sampson D. B., and Scott R. D. 2012. An exploration of the shapes and stability of population-selection curves. Fish and Fisheries, 13: 89–104.

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. 1984. The availability and information content of fisheries data. *In* Exploitation of marine communities SE-4, pp. 95–109.

Ed. by R. M. May. Springer, Berlin, Heidelberg. http://dx.doi.org/10.1007/978-3-642-70157-3_4.

Wetzel C. R., and Punt A. E. 2011. Performance of a fisheries catch-at-age model (Stock Synthesis) in data-limited situations. Marine Freshwater Research, 62: 927–936.

Whitlock R. E., McAllister M. K., and Block B. A. 2012. Estimating fishing and natural mortality rates for Pacific bluefin tuna (*Thunnus orientalis*) using electronic tagging data. Fisheries Research, 119–120: 115–127. http://linkinghub.elsevier.com/retrieve/pii/S0165783612000021 (last accessed 21 August 2013).

Williams E. H., and Quinn T. J. 1998. A parametric bootstrap of catch-age compositions using the Dirichlet distribution. Fishery Stock Assessment Models, AK-SG-98-01: 371–384.

Yin Y., and Sampson D. B. 2004. Bias and precision of estimates from an age-structured stock assessment program in relation to stock and data characteristics. North American Journal of Fisheries Management, 24: 865–879.

Handling editor: Shijie Zhou