NRC Research Press
ARTICLE

Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries

Merrill B. Rudd and James T. Thorson

Abstract: In fisheries with limited capacity for monitoring, it is often easier to collect length measurements from fishery catch than quantify total catch. Conventional stock assessment tools that rely on length measurements without total catch do not directly account for variable fishing mortality and recruitment over time. However, this equilibrium assumption is likely violated in almost every fishery, degrading estimation performance. We developed an extension of length-only approaches to account for time-varying recruitment and fishing mortality. This Length-based Integrated Mixed Effects (LIME) method at a minimum requires a single year of length data and basic biological information but can fit to multiple years of length data, catch, and an abundance index if available. We use simulation testing to demonstrate that LIME can estimate how much fishing has reduced spawning output in the most recent year across a variety of scenarios for recruitment and fishing mortality. LIME improves data-limited fisheries stock assessments by its flexibility to incorporate additional years or types of data if available and obviates the need for equilibrium assumptions.

Résumé : Dans les pêches caractérisées par une capacité de surveillance limitée, il est souvent plus facile de recueillir des mesures de la longueur sur les prises que de quantifier les prises totales. Les outils traditionnels d'évaluation des stocks qui reposent sur les mesures de la longueur sans données sur les prises totales ne tiennent pas directement compte de la variabilité dans le temps de la mortalité par pêche et du recrutement. Cette hypothèse d'équilibre n'est toutefois probablement pas respectée dans presque toutes les pêches, ce qui se traduit par une performance réduite des estimations. Nous avons mis au point une extension des approches reposant uniquement sur la longueur pour tenir compte de la variabilité temporelle du recrutement et de la mortalité par pêche. Cette méthode axée sur les effets mixtes intégrés basés sur la longueur (LIME) nécessite au minimum une seule année de données sur la longueur et de renseignements biologiques de base, mais peut être calée sur plusieurs années de données sur la longueur, les prises et un indice d'abondance, si ces données sont disponibles. Nous avons fait des essais de simulation pour démontrer que la méthode LIME peut estimer la réduction causée par la pêche de l'apport de la reproduction durant l'année la plus récente pour un éventail de scénarios de recrutement et de mortalité par pêche. L'approche LIME améliore les évaluations de stocks pour les pêches caractérisées par des données limitées parce qu'elle permet d'intégrer d'autres années ou types de données éventuellement disponibles et elle élimine la nécessité de faire appel à des hypothèses d'équilibre. [Traduit par la Rédaction]

Introduction

Many fisheries worldwide lack the quality and quantity of data used in classical stock assessments but must deal with limited information to make management decisions (Quinn et al. 2016). Stock assessments can provide a quantitative starting point for developing management strategies and monitoring the impacts of management. For example, the Magnuson–Stevens Fisheries Conservation and Management Act mandates fisheries managers in the United States to set catch limits based on the "best available science", involving advice from stock assessment to inform those limits (Darcy and Matlock 1999; Methot et al. 2014). The Marine Stewardship Council similarly requires fisheries seeking certification to go through a stock assessment process to determine their sustainability (Gulbrandsen 2009). In the case of small-scale fisheries in developing nations, it is possible to manage a fishery using only harvest control rules to meet management objectives, without formal stock assessment estimating status relative to reference points (Mahon 1997). However, beyond their use to evaluate management strategies for fishery resources (Carruthers et al. 2014), modeling tools can help with community engagement in the scientific process, such as conflict resolution (Butler et al. 2006) and integration of local knowledge to support cooperation between fishers and scientists (Neis 1992; Azzurro et al. 2011).

Many stocks worldwide remain unassessed, e.g., the UN Food and Agriculture Organization global fishery statistics database includes 19 624 unique combinations of country and taxa (FAO 2016), while the RAM Legacy Stock Assessment Database only includes 1268 stock assessments (Ricard et al. 2012). The regions of the world with the fewest stock assessments relative to the number of stocks perform worse across fishery management attributes (Melnychuk et al. 2017), indicating that the presence of fishery monitoring and assessment may have some relation with successful management. In the following, we define "data-limited" as any stock with uninformative data (no contrast to

---

Received 11 April 2017. Accepted 26 August 2017.

M.B. Rudd. School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195, USA.
J.T. Thorson.* Fisheries Resource Assessment and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, WA 98112, USA.
Corresponding author: Merrill Rudd (email: merrillrudd@gmail.com).

*James T. Thorson currently serves as an Associate Editor; peer review and editorial decisions regarding this manuscript were handled by Terrance Quinn II.

A correction was made to the e-First version of this paper on 20 March 2018 prior to the final issue publication. The current online and print versions are identical and both contain the correction.

Can. J. Fish. Aquat. Sci. **75**: 1019–1035 (2018) dx.doi.org/10.1139/cjfas-2017-0143

provide information on rates of change) or lacking data types typically used in statistical estimates of stock status (e.g., fishery-independent surveys, proportion at ages) (Costello et al. 2012; Dowling et al. 2016). Data-limited stock assessment methods are in great demand in both developed and developing nations due to the ubiquity of data-limited fisheries and the need for science-based management decisions (Wetzel and Punt 2011; Dowling et al. 2015; Chrysafi and Kuparinen 2016).

Length-based assessment methods are a vital component of the data-limited stock assessment toolbox because it is much easier to obtain reliable length measurements of a portion of the fishery catch than to measure total catch or record effort data for many small-scale or nontarget species (Harley et al. 2001; Kokkalis et al. 2015; Nadon et al. 2015; Prince et al. 2015*b*). Similarly, age information and a fishery-independent survey that are representative of total abundance are prohibitively expensive or impossible to collect for most fisheries in the world. Prominent length-based methods for estimating reference points in data-limited fisheries include length-based spawning potential ratio (LB-SPR) (Hordyk et al. 2015) and mean-length mortality estimation methods (Nadon et al. 2015). LB-SPR uses length-composition data and assumptions about biological parameters to make a rapid assessment of stock status relative to unfished levels assuming equilibrium conditions (Hordyk et al. 2015; Prince et al. 2015*b*). While LB-SPR can use multiple years of length data, status determination is based on one year of data at a time (i.e., estimates of status over multiple years are based on that year's length composition alone). Mean-length mortality estimators (e.g., Gedamke and Hoenig 2006), first developed by Beverton and Holt (1957), assume that fishing mortality directly influences mean length of the catch and have been used for assessments in the US South Atlantic, Pacific islands, and Caribbean (Ehrhardt and Ault 1992; Ault et al. 2005, 2008; Gedamke and Hoenig 2006; Nadon et al. 2015). As measures of stock status, these length-based methods derive the spawning potential ratio (SPR) reference point, defined as the proportion of unfished reproductive potential at a given level of fishing pressure (Goodyear 1993).

Length-based assessment methods estimating stock status assume that recruitment and fishing mortality arise from deterministic relationships or have not changed over a period significant for management and the life history of the species (termed "equilibrium assumptions"). Equilibrium assumptions are often violated (Gedamke and Hoenig 2006), as recruitment is quite variable for most species and fishing mortality changes with markets and other socioeconomic factors in the fishing community (Thorson et al. 2013, 2014*a*). Stochastic ocean conditions and productivity regime shifts may cause recruitment to vary erratically, gradually, or periodically at any given time (Vert-pre et al. 2013; Thorson et al. 2014*b*; Szuwalski et al. 2015). However, the violation of the equilibrium assumption may be difficult to detect. For example, the equilibrium assumption may appear valid when the mean length is constant over time (Gedamke and Hoenig 2006; Nadon et al. 2015). If recruitment is then constant over time, increasing fishing mortality will lead to decreasing mean length as the larger individuals are harvested and only smaller individuals remain in the population. However, constant fishing mortality and a recruitment pulse would also lead to a decrease in mean length, with more young individuals entering the population. Given both variable fishing mortality and recruitment processes occurring on the same population as well as errors when measuring mean length, a mean-length time series may appear constant when time-varying population processes are instead cancelling each other out.

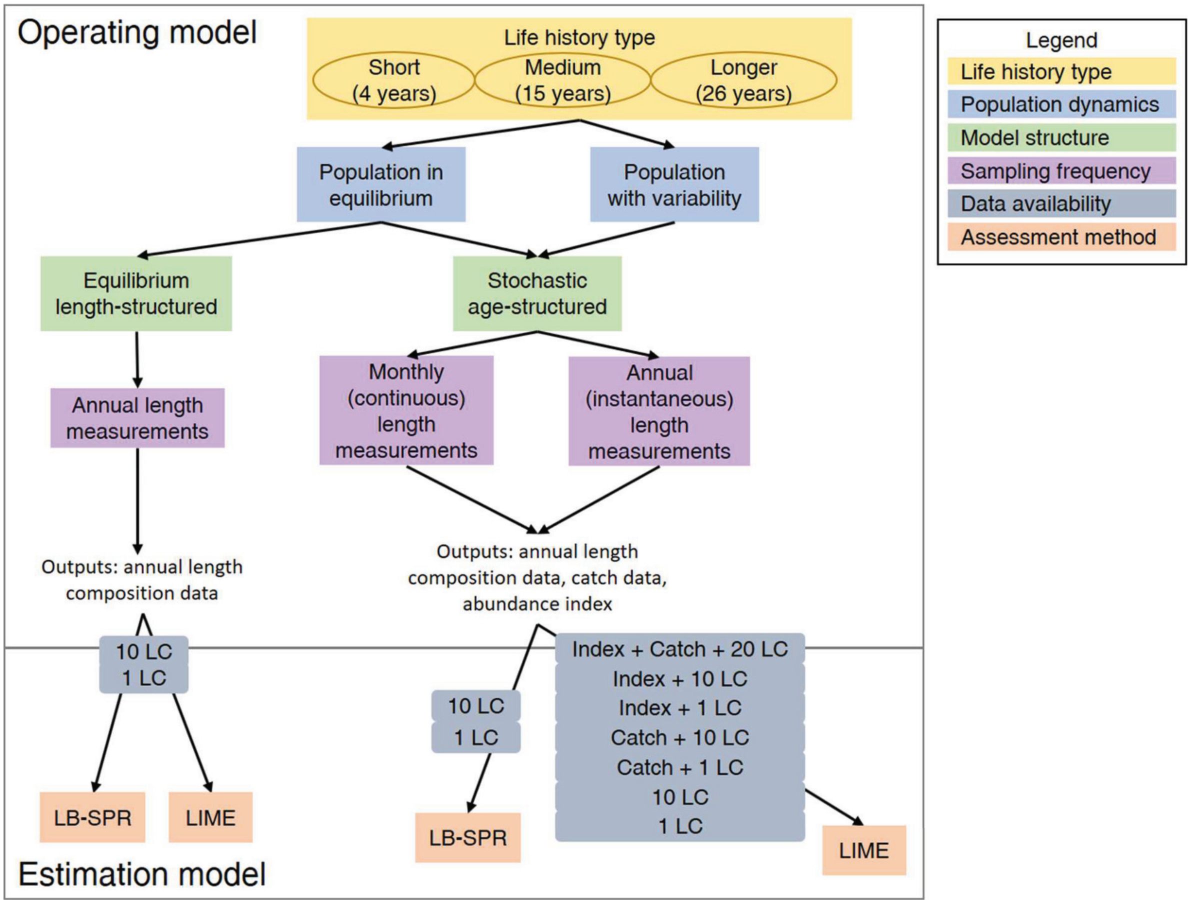
As an alternative to the equilibrium assumptions, a mixed-effects model can be used to deal with important demographic changes by estimating random variation in recruitment, fishing mortality, or other biological processes as well as the magnitude (variance) of random variation in each process (de Valpine and Hastings 2002; Buckland et al. 2004; Schnute and Haigh 2007; Thorson and Minto 2015). Mixed-effects models can directly account for variation arising from natural processes or measurement processes separately and therefore improve performance in nonlinear fisheries models (de Valpine and Hastings 2002; Ono et al. 2012; Thorson et al. 2015*b*). A main criticism of length-based methods with equilibrium assumptions is that with a single year of length-composition data and general understanding of biological parameters, it is impossible to determine whether a larger proportion of small fish in the catch is caused by strong cohort in recent years or by the removal of larger fish from the system. Accounting for random variation in recruitment, fishing mortality, and observation error arising from the process of sampling fish lengths from the population helps to tease apart each of these processes and better identify the true state of the fish population.

The aim of this study is to introduce a new length-based, integrated, mixed-effects (LIME) model and demonstrate its statistical performance when estimating reference points assuming that only length-composition and basic biological information are available. This method builds upon the catch curve stock-reduction analysis model (Thorson and Cope 2014), which includes an estimate of mortality from the age composition and at least one year of total fishery catch to estimate maximum sustainable yield (MSY) based reference points without assuming information about final biomass relative to unfished biomass (as generally used in stock reduction analysis). As an extension, LIME uses samples of length in place of the more resource-intensive samples of age and can estimate the SPR reference point if catch data are unavailable. To demonstrate the LIME model, we used simulation testing to (*i*) demonstrate that LIME is unbiased across several life-history types and patterns of fishing mortality and recruitment variability and able to include more years of length measurements and catch and (or) abundance index data, (*ii*) examine the sensitivity of the model to sample size of length measurements and error in input parameters, and (*iii*) compare LIME against LB-SPR to assess performance under various violations of model assumptions, including the timing of sampling within the year and modeling monthly time steps.

Methods

LIME is an age-structured population dynamics model with the ability to (1) account for variable fishing mortality and recruitment when only length data are available and (2) treat multiple years and types of data in an integrated manner to improve estimates of fishing mortality changes over time. The minimum inputs for the LIME assessment method are data on the length composition of the catch from a single year as well as assumed life-history information, including the length at age relationship, an assumed natural mortality rate, and length at 50% maturity. LIME estimates annual fishing mortality rates, lengths at 50% and 95% selectivity to the fishing gear, and the Dirichlet-multinomial parameter θ as fixed effects. The effective sample size of length data is linearly related to the input sample size with intercept $(1 + \theta)^{-1}$ and slope $\theta(1 + \theta)^{-1}$ (Thorson et al. 2017*a*). LIME can be differentiated from other age-structured models (e.g., Stock Synthesis) in that annual recruitments are treated as random effects, where mean and standard deviation of a distribution for recruitment are additionally estimated as fixed effects. Another key difference is that LIME does not require catch data: if no information on the scale of population size is available, recruitment will be estimated relative to average levels for an unfished population. As measures of stock status, we derived the SPR reference point to compare results with LB-SPR. We also derived the $F_{30\%}$ and $F_{40\%}$ reference points (the fishing mortality rates that would result in SPR of 30% and 40%, respectively; Clark 2002). We derived MSY by finding the fishing mortality rate that results in the highest yield per recruit. When total catch data were available (thereby providing information on scale of the population size), LIME would estimate equi-

Fig. 1. Diagram of the simulation study. "LC" stands for length composition and numbers indicate number of years in the 20 year time period modeled. In the case of the instantaneous annual length measurement scenario, we compared multiple sample sizes of length measurements annually (1000, 500, 200, 100, 50, and 20). For the other scenarios, we assumed the base case of 200 length measurements annually. [Colour online.]



librium recruitment, which would scale the MSY based on a per-recruit equation to a scale appropriate for the population size.

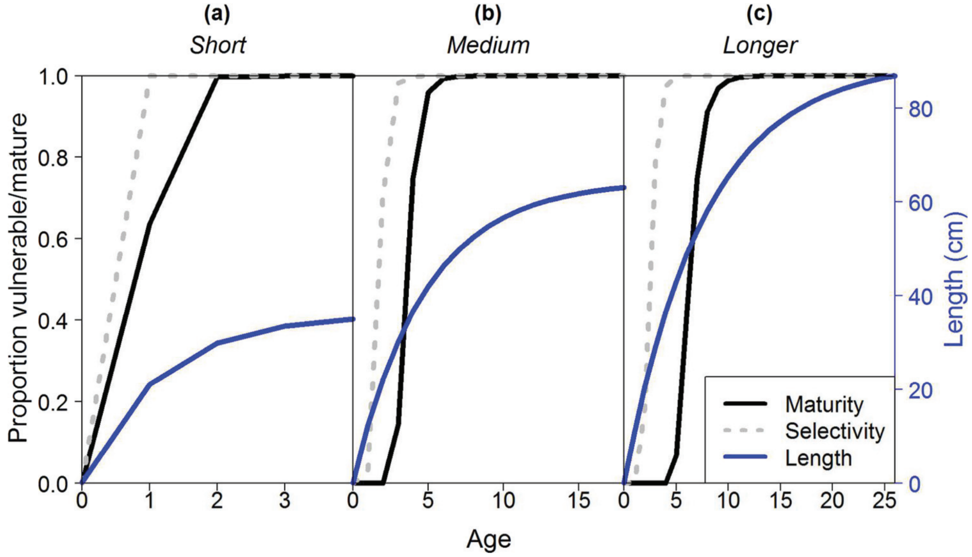
We developed an operating model to simulate true populations and generate data under a variety of fishing, recruitment, and life-history scenarios. We then used this operating model to explore the estimation performance of LIME for different scenarios regarding recruitment, fishing mortality, data availability, and life history (Fig. 1). We conducted all simulation modeling using the open-source statistical software R (R Core Team 2016) and all estimation in the R package *Template Model Builder* (TMB) (Kristensen et al. 2016) as implemented in our R package *LIME* (https://github.com/merrillrudd/LIME, doi:10.5281/zenodo.834404, version 1.0.0).

Operating model

We tested LIME for three different life-history types chosen to reflect the types of taxa for which length-based assessments are commonly demanded (Fig. 2). These life-history types were (*a*) a short-lived fish, mimicking rabbitfish (*Siganus sutor*): $L_\infty = 36.2$ cm, $k = 0.87$, $M = 1.49$, $L_m^{50} = 20.2$ cm, maximum age = 4; Hicks and McClanahan 2012), (*b*) a medium-lived fish, mimicking spotted rose snapper (*Lutjanus guttatus*): $L_\infty = 64.6$ cm, $k = 0.21$, $M = 0.43$, $L_m^{50} = 34.0$ cm, maximum age = 15; Bystrom 2015), and (*c*) a longer-lived fish, mimicking red grouper (*Epinephelus morio*): $L_\infty = 90.0$ cm, $k = 0.13$, $M = 0.18$, $L_m^{50} = 50.0$ cm, maximum age = 26; Heemstra and Randall 1993) (Table 1). Each simulated population began with biomass at a fraction of unfished biomass, drawn from a uniform distribution between 0.05 and 0.95.

For each life-history scenario, we tested LIME performance under three scenarios of fishing mortality and recruitment variability (Fig. 3). The first is the "equilibrium scenario", which matches the nonvariable fishing mortality and recruitment assumptions of LB-SPR. The equilibrium scenario involved fishing mortality and recruitment constant over a 20 year period, with a standard deviation for fishing mortality and recruitment set to a negligible 0.01. The second scenario, the "two-way base scenario", involved a linear change from the fishing mortality that would result in the randomly chosen initial depletion to the rate associated with 20% SPR ($F_{20\%}$) over the first 7 years of the 20 year time series. This change could be positive or negative depending on the F associated with randomly selected value for initial depletion and $F_{20\%}$. Over the next 7 years of the time series, F was constant at $F_{20\%}$ and then the fishing rate decreased linearly down to half of $F_{20\%}$ for the last 6 years of the time series. $F_{20\%}$ was calculated deterministically based on the biological information and selectivity associated with each life-history type. Thus, this value would not vary between different scenarios of variability within a life-history type. The fishing mortality time series in the "two-way base scenario" varied between simulation iterations in (*a*) the randomly chosen initial depletion and (*b*) the lognormally distributed deviations around this two-way trip (following the same equation for lognormal recruitment deviates in Table 2, eq. 1, except using the standard deviation for fishing mortality, σ_F, equal to 0.2). While we calculated $F_{20\%}$ for the medium-lived and longer-lived life-history types, we fixed $F_{20\%}$ to 3.0 for the short-lived life-history type. Based on the assumption of instantaneous annual fishing

Fig. 2. Selectivity, maturity, and length at age curves for the three life-history types tested in the simulation study mimicking (*a*) rabbitfish (*Siganus sutor*), a short-lived fish (asymptotic length L_∞ = 36.2 cm, von Bertalanffy k = 0.87, natural mortality M = 1.49, length at 50% maturity L_m^{50} = 1 year, maximum age A = 4 years; Hicks and McClanahan 2012), (*b*) spotted rose snapper (*Lutjanus guttatus*), a medium-lived fish (L_∞ = 64.6 cm, k = 0.21, M = 0.43, L_m^{50} = 4 years, A = 15 years; Bystrom 2015), and (*c*) red grouper (*Epinephelus morio*), a longer-lived fish (L_∞ = 90 cm, k = 0.13, M = 0.18,L_m^{50} = 7 years, A = 26 years; Heemstra and Randall 1993). [Colour online.]



mortality, $F_{20\%}$ for the short-lived fish was calculated to be much greater than could be supported even by very high fishing capacity. Recruitment was variable and autocorrelated over a 20 year period (Table 2, eq. 2). This scenario included a standard deviation of recruitment residuals σ_R equal to 0.737 and a first-order autoregressive coefficient equal to 0.426, the mean of the predictive distribution from a meta-analysis of recruitment variability in global fish orders (Thorson et al. 2014*b*). A third scenario, the "one-way base scenario", involved the same recruitment variability and autocorrelation as the "two-way scenario" but with fishing mortality changing linearly from the rate that would result in the randomly chosen initial depletion to $F_{20\%}$ over the 20 year period. Thus, this change could be positive or negative depending on the F associated with the randomly chosen initial depletion relative to $F_{20\%}$ for the life-history type. This combination of scenarios tests the ability to track how the population processes are changing over time with variability in the system.

In the operating model, we assumed that the natural mortality rate was constant, known, and independent of size or age. We modeled individual growth using a von Bertalanffy growth function (Table 3, eq. 1; Fig. 2). We assumed that maturity at length m_j was based on a one-parameter logistic function (Table 3, eq. 2; Fig. 2). We converted maturity at length to maturity at age using a normal distribution with standard deviation a function of the coefficient of variation of the age–length curve (Table 3, eq. 3). Selectivity at length follows a two-parameter logistic model (Table 3, eq. 4; Fig. 2), with estimated parameter length at 50% selectivity and a second parameter δ representing the difference between length at 95% and 50% selectivity. We modeled mass at age as an allometric function of individual length at age (Table 3, eq. 5). We calculated annual total biomass as a function of the abundance and mass at age (Table 3, eq. 6). Spawning biomass was a function of the total annual biomass and the proportion mature at age (Table 3, eq. 7). These processes contributed to an underlying age-structured model (Table 3, eq. 8). Parameter definitions and input values are listed in Table 1.

Data generation

We generated length data by simulating underlying age-structured dynamics and then sampling length composition from the vulnerable population instantaneously at the beginning of each year. First, we calculated the probability of being in a length bin for individuals of each age (Table 2, eq. 3). We then calculated the probability of harvest in each length bin each year as the proportion of the abundance at age vulnerable to the fishing gear each year multiplied by the probability of being in a length bin given age (Table 2, eq. 4). We used a multinomial probability distribution to generate the length frequency of samples of the catch in each length bin over time (Table 2, eq. 5).

We assumed that a sample size n of 200 individuals were measured annually and that this was the "true" sample size of the length data. We compared the base case of 200 length measurements annually to model performance under sample sizes of 1000, 500, 100, 50, and 20. We used the large sample size of 1000 to confirm that the model is unbiased and precise across data availability scenarios under ideal circumstances and then tested the alternate, lower sample sizes to assess performance under more realistic sample sizes.

We tested seven different scenarios of data availability. Two scenarios included only length data, assuming that 1 or 10 years of length data were available (the "one-length composition", "10-length compositions" scenarios). Data from these two scenarios were used to demonstrate the value of additional years of length data for each life-history type. We also explored five scenarios of additional data availability to demonstrate the integrated nature of LIME. The data availability scenarios included (*i*) a "data-rich" scenario with 20 years of total catch, 20 years of an abundance index, and 20 years of length data, used as proof-of-concept that LIME works when a high amount of informative data exists, (*ii*) "index plus 10" with 20 years of an abundance index and 10 years of length data, (*iii*) "index plus 1" with 20 years of an abundance index and 1 year of length data, (*iv*) "catch plus 10" with 20 years of total catch data and 10 years of length data, and (*v*) "catch plus 1" with 20 years of total catch data and 1 year of length data. We assume that catch and abundance indices are lognormally distributed with a log standard deviation of 0.2 (Table 1). The "catch plus 1" scenario is essentially a stock-reduction analysis while replacing the assumed information regarding final biomass (as used in the stock-reduction analysis)

Table 1. Parameter definitions, including parameter input values for the base scenario for each of the three life-history types, and whether the parameter is fixed ("Fixed") or estimated ("Est.") depending on the data availability scenario or if the parameter is used in data generation only ("Sim.").

		Input value for life-history type			Data scenario			
Symbol	Description	Short	Medium	Longer	Rich	Index + LC	Catch + LC	LC only
Biological								
L_∞	Asymptotic length	36.2 cm	64.6 cm	90.0 cm	Fixed	Fixed	Fixed	Fixed
k	Brody growth coefficient	0.87	0.21	0.13	Fixed	Fixed	Fixed	Fixed
t_0	Age at length = 0	–0.01	–0.01	–0.01	Fixed	Fixed	Fixed	Fixed
L_m^{50}	Length at 50% maturity	20.2 cm	34.0 cm	50.0 cm	Fixed	Fixed	Fixed	Fixed
α	Length–mass scalar	0.0597	0.0245	0.0264	Fixed	Fixed	Fixed	Fixed
β	Length–mass allometric	2.75	2.79	2.96	Fixed	Fixed	Fixed	Fixed
A	Maximum age	4	18	26	Fixed	Fixed	Fixed	Fixed
M	Natural mortality	1.49	0.43	0.18	Fixed	Fixed	Fixed	Fixed
h	Steepness parameter	1	1	1	Fixed	Fixed	Fixed	Fixed
CV_L	Coefficient of variation for the length–age curve	0.1	0.1	0.1	Fixed	Fixed	Fixed	Fixed
J	Maximum length bin	54 cm	97 cm	135 cm	Fixed	Fixed	Fixed	Fixed
R_0	Equilibrium recruitment	1	1	1	Est.	Fixed	Est.	Fixed
σ_R	Recruitment standard deviation	0.737	0.737	0.737	Est.	Est.	Est.	Est.
ρ	Recruitment autocorrelation	0.426	0.426	0.426	Sim.	Sim.	Sim.	Sim.
Nonbiological								
L_s^{50}	Length at 50% selectivity	11.3 cm	20.0 cm	25.0 cm	Est.	Est.	Est.	Est.
δ	Difference $L_s^{95} - L_s^{50}$ (expressed here as ratio the L_s^{95}/L_s^{50} to compare across life histories)	1.3	1.3	1.3	Est.	Est.	Est.	Est.
q	Catchability coefficient	1×10^{-5}	1×10^{-5}	1×10^{-5}	Est.	Est.	Fixed	Fixed
σ_F	Fishing mortality penalty standard deviation	0.2	0.2	0.2	Fixed	Fixed	Fixed	Fixed
σ_c	Standard deviation for observed catch	0.2	0.2	0.2	Fixed	Fixed	Fixed	Fixed
σ_I	Standard deviation for observed abundance index	0.2	0.2	0.2	Fixed	Fixed	Fixed	Fixed
θ	Dirichlet-multinomial parameter related to effective sample size	10	10	10	Est.	Est.	Est.	Est.
n	Sample size of length measurements	200	200	200	Fixed	Fixed	Fixed	Fixed
ϕ	Cumulative normal probability distribution							

Note: Values separated by commas indicate alternate values for sensitivity analysis. Note that scenarios including length composition (LC) are divided into two scenarios with either 1 or 10 years of length data. Parameter input values are based on short-lived *Siganus sutor* (Hicks and McClanahan 2012), medium-lived *Lutjanus guttatus* (Bystrom 2015), and longer-lived *Epinephelus morio* (Heemstra and Randall 1993).

with a length-based catch curve to estimate fishing mortality in the final year, which is analogous to Thorson and Cope (2014), except using length-composition samples. We calculated the expected catch at age using the Baranov catch equation (Table 2, eq. 6), summing to obtain total annual catch (Table 2, eq. 7), based on the true fishing mortality time series and selectivity specified in the operating model. The generated standardized abundance index I_t was proportional to spawning biomass (Table 2, eq. 8).

When not otherwise stated, we generated data on an annual time step, which assumes that length-composition data are collected instantaneously at the beginning of the year. In fisheries where only length data and biological information are available (i.e., no catch or abundance index), it is possible that length data would be available on less than an annual time step. We tested scenarios where length data were collected on a monthly time step and then either pooled into an annual time step or kept on a monthly time step. The monthly data collection scenario is more representative of fisheries occurring year-round, with no specific season. Furthermore, length data on a monthly time step would be more representative of short-lived fish growth to account for fish observed between the midpoints of each age class.

Estimation model

The structure of LIME follows the structure of the operating model using eqs. 1–8 in Table 3 and eqs. 3, 4 and 6–8 in Table 2 to derive the predicted catch, index, and length composition. LIME requires at least 1 year of length data but is flexible to include annual fishery catch and (or) an abundance index. During our simulation testing, we assumed that LIME had the correct values for the von Bertalanffy length at age relationship, including asymptotic length (L_∞), growth coefficient (k), and age at length = 0 (t_{0-}), length–mass parameters, natural mortality (M), and the one-parameter logistic maturity at length schedule. Future studies can obtain these values from local studies, FishBase (Froese 1990), or global meta-analyses of fish life-history parameters (Thorson et al. 2017*b*). We tested LIME performance under violations of these assumptions in sensitivity analyses. We also fixed the values for catch and abundance index observation error and coefficient of variation in the process error for the age–length curve. For all data availability scenarios, the model estimates as fixed effects the annual fishing mortality, lengths at 50% and 95% selectivity, the recruitment standard deviation, and the Dirichlet-multinomial parameter θ related to the effective sample size of length measurements in each year. In scenarios when catch data are unavailable, there is no information on the scale of the population. In these scenarios, we fixed mean recruitment to a relative value of 1.0 so that the model does not estimate the scale of the population but only the annual deviations in recruitment, and estimates of relative reference points (e.g., SPR) can be derived but reference

Fig. 3. Scenarios of fishing mortality and recruitment under equilibrium and variable conditions for the three life-history types with corresponding trajectories of relative spawning biomass. Shaded regions represent the area between the 5th and 95th percentiles of the generated data and the lines show three randomly chosen iterations out of 100 as examples. Scenarios labeled "equilibrium" demonstrate that the initial depletion may start between 0.05 and 0.95, but fishing mortality remains constant to produce that level of depletion and recruitment is constant at 1.0 over time. Scenarios labeled "two-way" demonstrate fishing mortality that increases to a fishing mortality rate $F_{20\%-}$ that results in 20% spawning potential ratio over the first 7 years, stays at $F_{20\%}$ for 7 years, and then drops to half of $F_{20\%}$ for the last 6 years. Scenarios labeled "one-way" demonstrate a change from equilibrium fishing mortality resulting in an initial depletion between 0.05 and 0.95 to the fishing mortality $F_{20\%}$ over the 20 year period. [Colour online.]

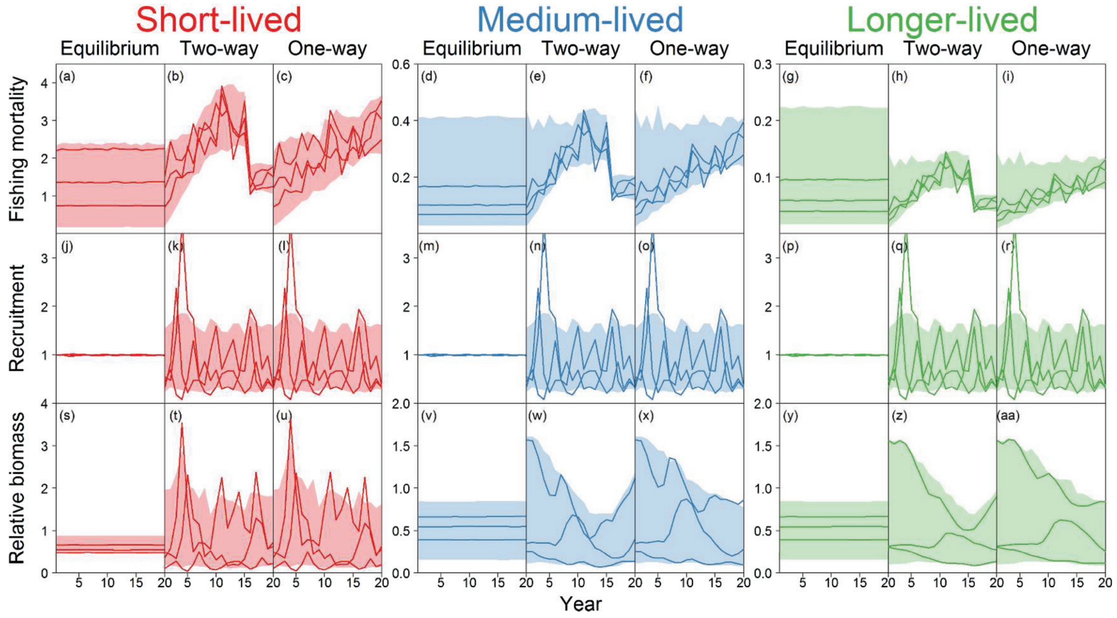


Table 2. Functions for generating data in the operating model.

	Equation	Description
1	$\varepsilon_t \sim \text{Lognormal}(0, \sigma_R)$	Nonautocorrelated recruitment deviations
2	$\tau_t = \varepsilon_t \quad t = 1$ $\tau_t = \varepsilon_t \tau_{t-1} \rho \sqrt{1-\rho^2} \quad t > 1$	Autocorrelated recruitment deviations
3	$p_{j,a} = \begin{cases} \phi\left(\dfrac{j - L_a}{L_a \text{CV}_L}\right), j = 1 \\ \phi\left(\dfrac{j - L_a}{L_a \text{CV}_L}\right) - \phi\left(\dfrac{j - 1 - L_a}{L_a \text{CV}_L}\right), 1 < j < J \\ 1 - \phi\left(\dfrac{j - 1 - L_a}{L_a \text{CV}_L}\right), j = J \end{cases}$	Probability of being in a length bin given age
4	$\pi_j = p_{j,a} \dfrac{\sum_{a=0}^{A} N_{a,t} S_a}{N_t}$	Predicted probability of harvest by length bin
5	$\tilde{\pi}_j \sim \text{Multinomial}(n, \pi_j)$	Generated probability of harvest by length bin
6	$C_{a,t} = [F_t S_a/(M + F_t S_a)] N_{a,t} [1 - \exp(M + F_t S_a)]$	Annual catch at age
7	$C_t = \sum_{a=0}^{A} C_{a,t}$	Annual catch
8	$I_t = qB_t$	Abundance index

Note: All except sqs. 2 and 5 are also used in the Length-based Integrated Mixed Effects (LIME) estimation model to predict values for the observed data.

points based on the scale of the population (e.g., MSY) are not meaningful. By contrast, when total catch data are available for at least one year, we can estimate mean recruitment and derive spawning biomass. When an index of abundance was available, we estimated the catchability coefficient for that index as an additional parameter q. A list of parameters estimated and fixed for each scenario is presented in Table 1.

We treated annual recruitment as a random effect in LIME, where recruitment each year is a function of an expected recruitment based on a Beverton–Holt stock–recruitment relationship (Table 4, eq. 1) and the estimated recruitment standard deviation (Table 4, eq. 2). For the simulation experiments in this paper, we fixed the steepness parameter h of the Beverton–Holt stock–recruitment function at 1.0, meaning that expected recruitment is

Table 3. Population dynamic equations used in the operating model and LIME estimation model.

	Equation	Description
1	$L_a = L_\infty\{1 - \exp[-k(a - t_0)]\}$	von Bertalanffy length at age a
2	$m_j = 1\Big/\left[1 + \exp\left(L_m^{50} - j\right)\right]$	Maturity at length
3	$m_a = \sum_{j=1}^{J} m_j(1/L_a\mathrm{CV}_L\sqrt{2\pi})\exp\left[-(j - L_a)^2/2L_a\mathrm{CV}_L^2\right]$	Maturity at age
4	$S_j = 1\Big/\left(1 + \exp\left\{-\left[\ln(19)\left(j - L_s^{50}\right)\right]\Big/\left(L_s^{95} - L_s^{50}\right)\right\}\right)$	Logistic selectivity at length
5	$w_a = \alpha L_a^{\beta}$	Mass at age
6	$B_t = \sum_{a=1}^{A} N_{a,t}w_a$	Annual total population biomass
7	$\mathrm{SB}_t = \sum_{a=0}^{A} N_{a,t}w_a m_a$	Annual spawning biomass
8	$N_{a,t} = \begin{cases} R_t, a = 0 \\ N_{a-1,t}\exp(-M - F_t S_{a-1}), 0 < a < A \text{ and } t = 1 \\ \dfrac{N_{a-1,t}\exp(-M - F_t S_{a-1})}{1 - \exp(-M - F_t S_{a-1})}, a = A \text{ and } t = 1 \\ N_{a-1,t-1}\exp(-M - F_{t-1}S_{a-1}), 0 < a < A \text{ and } t > 1 \\ (N_{a-1,t-1} + N_{a,t-1})\exp(-M - F_{t-1}S_{a-1}), a = A \text{ and } t > 1 \end{cases}$	Abundance at age over time
9	$\mathrm{SB}_0 = \sum_{a=0}^{A} R_0\exp(-aM)w_a m_a$	Unfished spawning biomass
10	$E_0 = \sum_{a=0}^{A}\exp(-aM)w_a m_a$	Expected lifetime egg production (unfished)
11	$E_f = \sum_{a=0}^{A}\exp[-a(M + FS_a)]w_a m_a$	Expected lifetime egg production (fished)
12	$\mathrm{SPR} = E_f/E_0$	Spawning potential ratio

Table 4. Components of the joint likelihood function in the LIME model as well as performance metrics across iterations of generated data.

	Equation	Description
1	$\mu_{Rt} = \dfrac{4hR_0\mathrm{SB}_{t-1}}{\mathrm{SB}_0(1 - h) + \mathrm{SB}_{t-1}(5h - 1)}$	Expected annual recruitment based on Beverton–Holt stock–recruitment relationship
2	$R_t \sim \mathrm{Lognormal}(\mu_{Rt}, \sigma_R^2)$	Annual recruitment R_t arising from a lognormal distribution
3	$F_t \sim \mathrm{Normal}(F_{t-1}, \sigma_F^2)$	Penalty on annual fishing mortality F_t
4	$\sigma_R \sim \mathrm{Lognormal}(0.7, 0.2^2)$	Recruitment standard deviation penalty
5	$\log L(\pi_j, \theta \mid \tilde{\pi}_j, n) = \log\Gamma(n + 1) - \sum[\log\Gamma(n\tilde{\pi}_j + 1)] + \log\Gamma(\theta n)$ $-\log\Gamma(n + \theta n) + \sum[\log\Gamma(n\tilde{\pi}_j + \theta n\pi_j) - \log\Gamma(\theta n\pi_j)]$	Dirichlet-multinomial log-likelihood
6	$n_{\mathrm{eff}} = (1 + \theta n)/(1 + \theta)$	Effective sample size
7	$C_t^{\mathrm{obs}} \sim \mathrm{Lognormal}\left(\sum_{a=0}^{A} C_t, \sigma_C^2\right)$	Lognormal likelihood (catch)
8	$I_t^{\mathrm{obs}} \sim \mathrm{Lognormal}\left(qB_t, \sigma_I^2\right)$	Lognormal likelihood (abundance index)
9	$\mathrm{MRE} = \mathrm{median}[(x_{\mathrm{estimated}} - x_{\mathrm{true}})/x_{\mathrm{true}}]$	Median relative error to quantify bias
10	$\mathrm{MARE} = \mathrm{median}\lvert[(x_{\mathrm{estimated}} - x_{\mathrm{true}})/x_{\mathrm{true}}]\rvert$	Median absolute relative error to quantify precision

constant among years and independent of the spawning stock biomass the previous year. It is possible to fix the steepness parameter at a value less than 1.0, where the mean of the lognormal distribution is the predicted number of recruits from the Beverton–Holt stock–recruitment relationship, as opposed to 1.0 or the equilibrium unfished recruitment. We chose to test LIME with 1.0 for the mean of the lognormal distribution to determine how well the model can estimate annual recruitment with all variation from recruitment deviates. Unfished spawning biomass is calculated using the same equation as fished spawning biomass (Table 3, eq. 7) but without the fishing mortality and selectivity terms (Table 3, eq. 9). Unlike the operating model, we did not account for autocorrelation in recruitment in the LIME estimation model.

We applied a random-walk penalty on annual estimates of fishing mortality, which shrinks the estimate of fishing mortality in year $t + 1$ towards its estimate in year t (Table 4, eq. 3). Drastic changes in fishing mortality between years are unlikely in the real world given costs of entering or leaving a fishery, but the random-walk process accommodates gradual changes in fishing mortality (Nielsen and Berg 2014). However, a fixed value of 0.2 for the standard deviation of the fishing mortality penalty does allow the model to estimate variability in fishing mortality if supported by the data (e.g., exit and entry from the fishery).

We also placed a lognormal penalty on the standard deviation of recruitment deviations, σ_R. Based on the meta-analysis conducted by Thorson et al. (2014), the log of the mean was set to the log of 0.737, with a log standard deviation of 0.353 (Table 4, eq. 4). The prior aids in the convergence of the σ_R parameter estimation, preventing the estimate from going to an unlikely value. There is also an upper bound on σ_R at 2.0.

The joint log-likelihood of the observed data is the sum of the log-likelihoods of the observed length data, log probability of fishing mortality and recruitment variation, and the log-likelihood of the catch and abundance index, if available. We include bias correction for recruitment deviations using the TMB bias correction feature (Thorson and Kristensen 2016). For scenarios that include abundance index and catch data, a lognormal probability distribution was assumed to describe error in both data types (Table 4, eq. 7, eq. 8). The respective observation errors σ_I and σ_C are fixed a priori (not estimated as parameters). We assumed that the length data arose from a Dirichlet-multinomial probability distribution with estimated parameters θ_c related to the effective sample size of length measurements each year. Many stock assessment methods use a multinomial distribution to fit age or length data, but the effective sample size must be calculated externally (Francis 2014). By contrast, the Dirichlet distribution can represent variability in the proportions in each length bin, but the parameters do not correspond to the easily interpretable effective sample size of length data for which model results are highly sensitive. The Dirichlet-multinomial is an alternative to these two distributions, estimating an additional parameter θ within the integrated model (Thorson et al. 2017*a*). The effective sample size is a nonlinear function of input sample size (Table 4, eq. 6) (Thorson et al. 2016). As θ approaches infinity, the effective sample size is equal to the observed sample size, and the multinomial distribution is a special case of the Dirichlet-multinomial distribution.

In model runs, we assessed that the model had converged if the final gradient for all parameters was less than 0.001. If the initial model run did not converge (resulting in NAs or a high final gradient), the model would be run up to 10 additional times with starting values equal to the estimates from the nonconverged model plus a random number drawn from a normal distribution with mean zero and standard deviation 0.2. For each combination of life-history type, data availability scenario, fishing mortality pattern, and recruitment dynamics, we obtained 100 iterations of generated data and ran the estimation model for each set.

Comparison to LB-SPR

We ran LB-SPR from the R package *LBSPR* (Hordyk et al. 2015) with 1 year (LBSPR1) and 10 years (LBSPR10) of length data using the operating model described above. LB-SPR requires as input the length data in each year and the ratio of natural mortality to the von Bertalanffy growth coefficient (M/k) as well as inputs similar to those required for LIME: the von Bertalanffy asymptotic length parameter, coefficient of variation of the asymptotic length, length at 50% and 95% maturity, length–mass parameters, and starting values for the length at 50% and 95% selectivity (Table 1). We assumed that the coefficient of variation of the asymptotic length was equivalent to the coefficient of variation of the entire age–length curve. For the base runs, we used the true value for the length at 95% maturity and selectivity from the two-parameter logistic curves used in the operating model. LB-SPR uses these inputs to calculate the abundance at relative age at equilibrium. LB-SPR estimates the ratio of fishing mortality to natural mortality and the lengths at 50% and 95% selectivity to best fit the predicted and observed length composition proportions and derives SPR, outputting estimates for these four values for each year with length data (Hordyk et al. 2015).

We note that this simulation experiment is comparing performance of LIME and LB-SPR using an operating model that is more similar to LIME. To further compare the strengths and limitations of each method, we used the *LBSPR* package to simulate length data under the LB-SPR equilibrium, length-based structure, so that we could also test LIME and LB-SPR using an operating model that is more similar to LB-SPR.

A major issue in the comparison of models was the time step used. To test model performance, we tested both the LIME and the LB-SPR methods using the monthly data generation feature in the operating model, where sampling is spread out over all 12 months instead of collected instantaneously at one point. We then compared LIME and LB-SPR performance pooling the monthly length data into an annual distribution and then running each model on an annual time step. These additional simulation tests were designed to compare these two methods under different data-generating models, with the goal of providing guidance for which methods may be appropriate under a variety of real-world conditions.

Reference points

We calculated SPR (Table 3, eqs. 10–12) as a biological reference point, used as a proxy for MSY when information on the scale of population size is not available, and for comparison with LB-SPR. A harvest strategy that targets a fishing mortality rate that is expected to result in 40% of unfished spawning output (termed "40% SPR") is considered risk averse for many species (Clark 2002). Therefore, we calculated these values as examples of possible fishing mortality reference points that could be used to compare with other length-based assessment methods.

Sensitivity tests

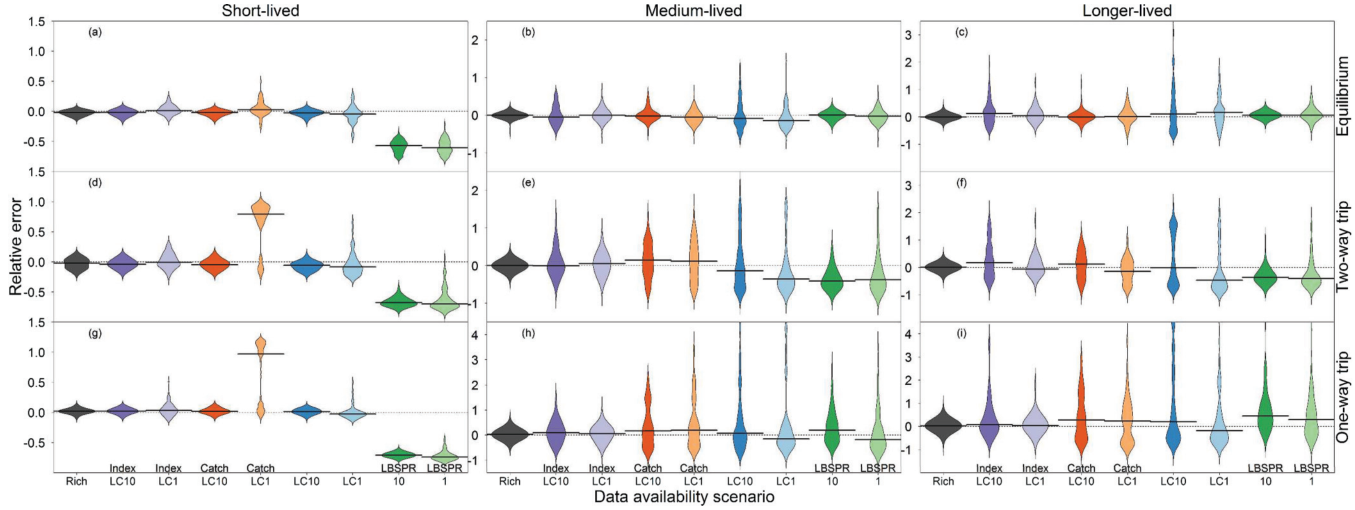
Preliminary exploration suggested that LIME performed poorly for short-lived species when using an annual time step; therefore, we explored LIME performance on a monthly time step. We generated monthly length data for the short-lived life-history type and ran LIME on a monthly time step by specifying the number of years as number of months and ages as fractions of a year and dividing the input natural mortality rate M by 12 so that mortality and growth occur in each month (Appendix, Fig. A1).

We also included sensitivity tests to LIME base models to understand biases associated with imperfect knowledge about species biology, fishery characteristics, and low sample sizes. We assessed the performance of LIME with (*a*) parameter misspecification of ±25% for each of the life-history inputs (M, CV_L, L_∞, k, and L_m^{50}) and (*b*) sample sizes of length data of 20, 50, 100, 500, and 1000 independent samples annually.

Model performance

To assess the ability of the model to accurately and precisely estimate quantities of management interest, we consider bias and precision (Table 4, eqs. 9 and 10) between estimated and true SPR in the last year of data across the 100 iterations of simulated data. We used median relative error (Table 4, eq. 9) to quantify bias and median absolute relative error (Table 4, eq. 10) to quantify precision. To understand the ability of the model to accurately capture uncertainty, we computed the "interval coverage", the proportion of iterations out of 100 where the true value of a population parameter in the terminal year is within the 50% confidence intervals. A well-performing model would have close to nominal coverage (i.e., a 50% coverage interval will contain the true value

Fig. 4. Distribution of relative error ((estimated – true)/true) for spawning potential ratio (SPR) in the current year for 100 iterations of simulated populations across the Length-based Integrated Mixed Effects (LIME) method and length-based SPR (LB-SPR) data availability scenarios for the three life-history types and scenarios of equilibrium and variable fishing mortality and recruitment with 200 samples of length measurements annually. The gray bean represents the "data-rich" scenario, which verifies that LIME is unbiased and most precise when an unrealistically high amount of data are available. Darker colours represent data availability scenarios with 10 years of length data and lighter colors represent the scenario with 1 year of length data available. Each life-history type has a different *y*-range. [Colour online.]



in 50% of simulation replicates). Coverage is presented to illustrate whether confidence intervals accurately capture model uncertainty.

Results

LIME performance across life-history, variability, and data scenarios

Based on the Monte Carlo simulations, LIME can estimate unbiased SPR when length data are available and biological characteristics are correctly specified across various scenarios of fishing mortality and recruitment patterns (Fig. 4). When tested using the same model structure and assumptions in the operating model, LIME performed best for short-lived fish, with increasing bias and decreasing precision for medium-lived fish and a further deterioration for longer-lived fish (Table 5). The bias in SPR in the terminal year was 0.002 for short-lived, −0.003 for medium-lived, and 0.016 for longer-lived fish across all data availability and population variability scenarios with 200 length measurements annually. Additional scenarios illustrate further LIME strengths and weaknesses (Table 5). With only 1 year of length data, LIME estimated SPR with bias of −0.038 for the short-lived life history, −0.186 for medium-lived, and −0.152 for longer-lived across variability scenarios. Integrating 10 years of length data improved accuracy in the estimation of SPR in the terminal year on average for each life history type (short-lived from −0.038 to −0.016, medium-lived from −0.186 to −0.048, and longer-lived from −0.152 to 0.102). The only exception within variability scenarios was decreased accuracy with 10 years of length data for the longer-lived life-history type under the one-way variability scenario (from −0.181 to 0.200). While 10 years of length data increased precision of SPR estimates for the short-lived life-history type over 1 year of length data (from 0.072 to 0.041), the additional length data decreased precision on average for the medium- and longer-lived life-history types (medium-lived from 0.334 to 0.391 and longer-lived from 0.529 to 0.626) (Table 5).

LIME provides the added advantage over other length-based assessment methods of including a catch time series and abundance index, if available, which further increases accuracy and precision under most cases. Across all scenarios, including an abundance index decreased bias (from −0.047 to 0.020) compared with when only length data were available. An abundance index particularly improved LIME performance for the short- and medium-lived life-history types. Breaking down by life-history types, bias decreases with an abundance index included for the short-lived life-history type (from −0.026 to 0.005) and medium-lived life-history type (from −0.126 to 0.025) but increases for the longer-lived life-history type (from −0.049 to 0.072) on average. Across all life-history types, including an abundance index increases precision (short-lived from 0.054 to 0.052, medium-lived from 0.346 to 0.190, and longer-lived from 0.594 to 0.257).

Adding a catch time series to length data decreased bias minimally on average across all scenarios (from −0.047 to 0.031). The improvement in accuracy is strongest under equilibrium conditions (from −0.036 to −0.007). Given variable fishing mortality and recruitment including catch data with length composition increased bias (from −0.055 to 0.118). This was mainly due to high bias in the "catch plus 1" scenario for short-lived fish (Table 5; Fig. 4). Excluding this anomalous scenario, including a catch time series decreased bias for a population with variability (from −0.055 to 0.047). The "catch plus 10" scenario removed the bias from the "catch plus 1" scenario for the short-lived life-history type under variability scenarios. However, the "catch plus 10" scenario did not necessarily improve bias or precision over the "catch plus 1" scenario for the medium- and longer-lived life-history types under variability scenarios (Table 5). This indicates that it may be advisable to include a catch time series only when more than 1 year of length data are available for a short-lived life-history type and consider the possibility of overestimating SPR when fishing mortality is changing over time for medium- and longer-lived life-history types when catch data are included.

LIME converged for 95% of iterations of generated data across life-history types, variability scenarios, data availability scenarios, and sample sizes of length data. LIME converged for 100% of iterations of generated data for the data-rich scenario (Fig. 5). With only length data, LIME converged for 95% of iterations across scenarios. The LIME convergence rate was 92% including catch data and 94% including an abundance index. Nonconvergence was always due to a high final gradient in parameter estimation, as opposed to any parameter being estimated at the upper or lower bound.

The interval coverage for a 50% confidence interval of LIME was 61% across all scenarios of life history, variability scenarios, data availability, and sample sizes of length data. This indicates that

Table 5. LIME performance in estimation of SPR in the terminal year across life-history types, scenarios of data availability, and patterns in fishing mortality and recruitment variability, compared with length-based SPR (LB-SPR).

	Short			Medium			Longer		
Data availability scenario	Equilibrium	Variability, two-way *F*	Variability, one-way *F*	Equilibrium	Variability, two-way *F*	Variability, one-way *F*	Equilibrium	Variability, two-way *F*	Variability, one-way *F*
Bias (MRE)									
Rich	–0.019[a]	–0.017[a]	0.023[a]	–0.004[a]	0.005[a]	0.034[a]	0.001[a]	0.012[a]	0.023[a]
Index + LC10	–0.014[a]	–0.039[a]	0.023[a]	–0.049[a]	–0.001[a]	0.094[b]	0.125[c]	0.173[c]	0.076[b]
Index + LC1	0.012[a]	–0.006[a]	0.033[a]	–0.003[a]	0.060[a]	0.065[b]	0.040[a]	–0.061[b]	0.029[a]
Catch + LC10	–0.018[a]	–0.045[a]	0.018[a]	–0.021[a]	0.150[c]	0.165[c]	–0.002[a]	0.124[c]	0.280[d]
Catch + LC1	0.029[a]	0.796[e]	0.972[e]	–0.056[b]	0.119[c]	0.193[c]	0.010[a]	–0.146[c]	0.236[d]
LC10	–0.022[a]	–0.056[a]	0.016[a]	–0.086[b]	–0.137[c]	0.072[b]	0.104[c]	–0.006[a]	0.200[d]
LC1	–0.045[a]	–0.082[b]	–0.025[a]	–0.140[c]	–0.355[e]	–0.154[c]	0.166[c]	–0.468[e]	–0.181[c]
LBSPR10	–0.570[e]	–0.676[e]	–0.708[e]	0.006[a]	–0.405[e]	0.196[c]	0.062[b]	–0.359[e]	0.452[e]
LBSPR1	–0.603[e]	–0.698[e]	–0.738[e]	–0.030[a]	–0.372[e]	–0.181[c]	0.059[b]	–0.403[e]	0.296[d]
Precision (MARE)									
Rich	0.027[a]	0.069[b]	0.028[a]	0.038[a]	0.044[a]	0.101[c]	0.038[a]	0.042[a]	0.089[b]
Index + LC10	0.039[a]	0.051[b]	0.038[a]	0.185[c]	0.176[c]	0.295[d]	0.244[d]	0.449[e]	0.352[e]
Index + LC1	0.046[a]	0.079[b]	0.062[b]	0.105[c]	0.207[d]	0.235[d]	0.180[c]	0.223[d]	0.223[d]
Catch + LC10	0.025[a]	0.055[b]	0.035[a]	0.080[b]	0.382[e]	0.476[e]	0.077[b]	0.423[e]	0.630[e]
Catch + LC1	0.060[b]	0.796[e]	0.972[e]	0.129[c]	0.429[e]	0.496[e]	0.161[c]	0.313[e]	0.714[e]
LC10	0.028[a]	0.063[b]	0.031[a]	0.246[d]	0.499[e]	0.470[e]	0.498[e]	0.743[e]	0.702[e]
LC1	0.072[b]	0.126[c]	0.039[a]	0.216[d]	0.437[e]	0.362[e]	0.301[e]	0.705[e]	0.549[e]
LBSPR10	0.570[e]	0.676[e]	0.708[e]	0.066[b]	0.412[e]	0.404[e]	0.078[b]	0.367[e]	0.486[e]
LBSPR1	0.603[e]	0.698[e]	0.738[e]	0.113[c]	0.478[e]	0.432[e]	0.120[c]	0.501[e]	0.536[e]

Note: Bias (top) is measured as MRE and precision (bottom) is measured as MARE from 100 iterations of generated data using the LIME age-structured operating model assuming instantaneous sampling with 200 length measurements annually.

[a] Bias/precision less than 5%.

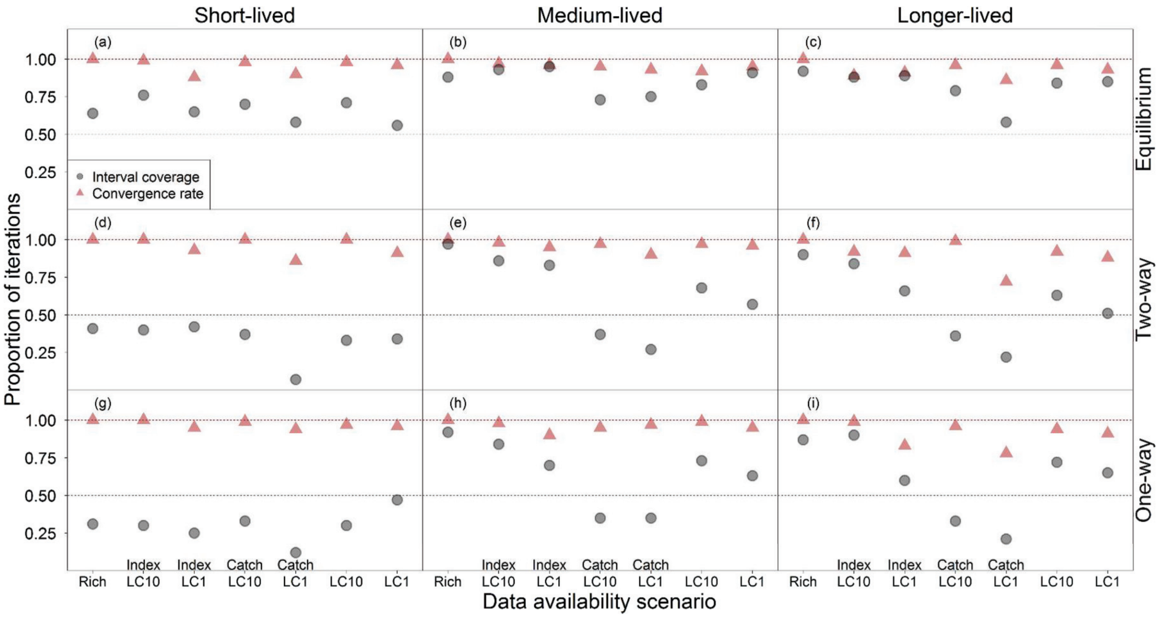
[b] Bias/precision less than 10%.

[c] Bias/precision less than 20%.

[d] Bias/precision less than 30%.

[e] Bias/precision greater than 30%.

Fig. 5. Proportion of iterations out of 100 where the true value of the SPR lies within the 50% confidence intervals for each life-history type across various patterns of fishing mortality and recruitment (gray circles) compared with the convergence rates for the same scenario (pink triangles). [Colour online.]



confidence interval estimates from LIME are generally informative about model uncertainty but may be wider (more conservative) on average. LIME estimated confidence intervals that were too narrow or wide for individual life history, variation, and data availability scenarios (Fig. 5). The "catch plus 1" scenario had the lowest coverage across life-history and variability scenarios (42%), likely because this data availability scenario had the most bias, and thus the true SPR would likely not fall within the confidence intervals at the expected rate. For the short-lived life-history type, LIME interval coverage was only 43%. While the method is expected to be unbiased for this life-history type on average, the confidence intervals are often too tight to include the true SPR within the 50% intervals at a rate of 50%. LIME estimated wider confidence intervals for the medium-lived life-history type, with

Table 6. LIME bias and precision in estimating SPR in the terminal year for a variety of sample sizes of independent length measurements (in all cases, 1 year of length data are provided to the model).

	Short			Medium			Longer		
Sample size	Equilibrium	Variability, two-way F	Variability, one-way F	Equilibrium	Variability, two-way F	Variability, one-way F	Equilibrium	Variability, two-way F	Variability, one-way F
Bias (MRE)									
1000	–0.033[a]	–0.070[b]	–0.029[a]	0.034[a]	–0.369[e]	–0.201[d]	–0.061[b]	–0.449[e]	–0.201[d]
500	–0.028[a]	–0.117[c]	–0.041[a]	–0.001[a]	–0.505[e]	–0.157[c]	–0.031[a]	–0.621[e]	–0.145[c]
200	–0.045[a]	–0.082[b]	–0.025[a]	–0.140[c]	–0.355[e]	–0.154[c]	0.166[c]	–0.468[e]	–0.181[c]
100	–0.027[a]	–0.070[b]	–0.028[a]	–0.055[b]	–0.370[e]	–0.201[d]	–0.127[c]	–0.480[e]	–0.206[d]
50	–0.082[b]	–0.124[c]	–0.041[a]	–0.182[c]	–0.499[e]	–0.163[c]	–0.141[c]	–0.621[e]	–0.107[c]
20	–0.171[c]	–0.153[c]	–0.043[a]	–0.295[d]	–0.431[e]	–0.214[d]	–0.465[e]	–0.751[e]	–0.451[e]
Precision (MARE)									
1000	0.070[b]	0.148[c]	0.048[a]	0.110[c]	0.443[e]	0.422[e]	0.240[d]	0.648[e]	0.474[e]
500	0.053[b]	0.172[c]	0.047[a]	0.125[c]	0.514[e]	0.423[e]	0.245[d]	0.663[e]	0.519[e]
200	0.072[b]	0.126[c]	0.039[a]	0.216[d]	0.437[e]	0.362[e]	0.301[e]	0.705[e]	0.549[e]
100	0.087[b]	0.147[c]	0.046[a]	0.217[d]	0.453[e]	0.429[e]	0.397[e]	0.665[e]	0.503[e]
50	0.108[c]	0.178[c]	0.046[a]	0.343[e]	0.510[e]	0.455[e]	0.421[e]	0.665[e]	0.506[e]
20	0.228[d]	0.168[c]	0.045[a]	0.374[e]	0.469[e]	0.610[e]	0.547[e]	0.754[e]	0.615[e]

Note: Bias (top) is measured as MRE and precision (bottom) is measured as MARE from 100 iterations of generated data using the LIME age-structured operating model assuming instantaneous sampling.

[a] Bias/precision less than 5%.

[b] Bias/precision less than 10%.

[c] Bias/precision less than 20%.

[d] Bias/precision less than 30%.

[e] Bias/precision greater than 30%.

interval coverage 72%, and longer-lived life-history types, with an interval coverage of 67%.

LIME sensitivity to sample size, input parameters, and selectivity estimation

With 1 year of length data, LIME estimated SPR with greater precision with more length measurements but not necessarily with greater accuracy (Table 6). Across all life history types, accuracy and precision improved greatly in the equilibrium scenario between 20 and 1000 length measurements collected annually. The short-lived life-history type experienced the strongest improvements in accuracy and precision (Table 6). This pattern was strongest with a population in equilibrium but was generally maintained for scenarios of variability as well. Accuracy generally reached a plateau with greater than 100 samples of length measurements for the short-lived life-history type. Alternatively, accuracy and precision improved for the medium- and longer-lived life-history types with at least 500 samples of length measurements for a population at equilibrium. With variability, there were no clear patterns in accuracy improvement for the medium- and longer-lived life-history types with increasing sample size of length measurements. This indicates that accuracy in SPR estimates is more likely to be improved with more data types than increasing sample size, particularly for medium- and longer-lived life histories (Tables 5 and 6).

Like all age- or length-based methods, LIME performance is sensitive to the correct specification of life-history information (Fig. 6). When all biological parameters were correctly specified, LIME estimated unbiased SPR on average with 10 years of length data (bias = –0.017). When L_∞ was misspecified as 25% greater than the truth, bias increased (to –0.554), meaning that on average, LIME estimated SPR to be lower than the truth. In this case, we would expect to see larger fish in the observed data. If those fish are not present in the length data, length-based models attribute the difference to a higher fishing mortality and thus a lower SPR than the truth. On the other hand, when L_∞ was misspecified as 25% lower than the truth, LIME estimated SPR higher than the truth on average (bias = 0.761). Assuming k was 25% lower than the truth, LIME estimated SPR higher than the truth on average (bias = 0.478). When k was assumed to be 25% higher than the truth, LIME estimated SPR lower than the truth on average but to a lower degree than the other biases due to life history misspecification (bias = –0.146). However, this lack of severe bias is simply due to some model runs resulting in estimates much greater than or lower than the truth, while none are unbiased. When M was fixed 25% higher than the truth, LIME estimated SPR to be higher than the truth (bias = 0.285). A fish that is faster to die can generally sustain a higher fishing pressure, and the interpretation of the higher SPR is that the population has more of its potential spawning biomass than it truly does. When M was assumed to be 25% lower than the truth, LIME estimated SPR lower than the truth (bias = –0.352).

Fixing length at 50% maturity at a value 25% higher than the truth resulted in negative bias in SPR (bias = –0.383). In this case, the LIME model attributes a lower proportion of the population as being mature, leading the estimated SPR to be lower than the truth. The opposite is true when length at maturity was 25% lower than the truth; estimates of SPR were biased higher than the truth (bias = 0.152). LIME was relatively insensitive to misspecification of CV$_L$ for the age–length curve, but a higher CV$_L$ resulted in a relatively lower estimate of SPR on average (bias = –0.044) and a higher CV$_L$ resulted in a relatively higher estimate of SPR on average (bias = 0.047).

Comparing LIME and LB-SPR

Simulation testing demonstrated that the LIME and LB-SPR methods perform well under different conditions. LB-SPR is biased when the equilibrium conditions are violated (Table 5). However, testing of LB-SPR against the age-structured LIME operating model with instantaneous sampling at the beginning of each year resulted in poor performance for the short-lived life-history type under equilibrium conditions (Tables 5 and 7; Fig. 7). We verified that LB-SPR performs well across life-history types when tested using an operating model that matches its model assumptions. LIME, on the other hand, overestimated SPR for the short-lived life-history type with 1 year of length data and, similar to the performance under its own operating model, had low precision for the medium- and longer-lived life-history types with 10 years of length data (Table 7). When length data were collected monthly and pooled into an annual length composition, LB-SPR performance improved over the annual model for the short-lived life-history type, but LIME overestimated SPR (Table 7; Fig. 7). Running

Fig. 6. Comparison of the distribution of estimation error for SPR in the terminal year with the life-history parameters fixed at ±25% of their true value from 100 iterations of generated data across the various life-history scenarios for the equilibrium (blue) and the two-way fishing mortality scenario with variable and autocorrelated recruitment (red), with other biological parameters fixed at their true values, including 10 years of length data with 200 length measurements annually. Life-history parameters include natural mortality, asymptotic length, the von Bertalanffy growth coefficient, length at 50% maturity, and the coefficient of variation for the age–length curve. [Colour online.]

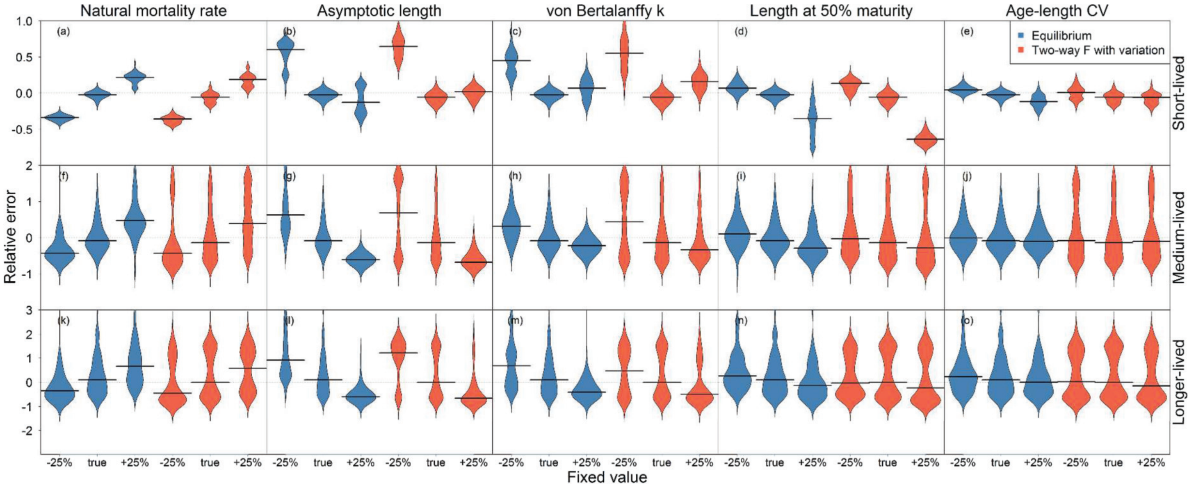


Table 7. Comparison of LIME and LB-SPR performance in estimation of SPR in the terminal year between equilibrium scenarios with length data collected annually and using an annual model, length data collected monthly but pooled annually for an annual model, and using the LB-SPR operating model.

	Short				Medium			Longer		
Data availability scenario	Annual data and model	Monthly data, annual model	Monthly data, monthly model	LB-SPR	Annual data and model	Monthly data, annual model	LB-SPR	Annual data and model	Monthly data, annual model	LB-SPR
Bias (MRE)										
LC10	−0.022^{a}	0.468^{e}	−0.040^{a}	0.068^{b}	−0.086^{b}	0.072^{b}	−0.133^{c}	0.104^{c}	0.184^{c}	−0.112^{c}
LC1	−0.045^{a}	0.484^{e}	0.034^{a}	0.195^{c}	−0.140^{c}	−0.104^{c}	−0.097^{b}	0.166^{c}	0.168^{c}	−0.113^{c}
LBSPR10	−0.570^{e}	0.136^{c}	NA	−0.005^{a}	0.006^{a}	0.090^{b}	0.002^{a}	0.062^{b}	0.117^{c}	0.010^{a}
LBSPR1	−0.603^{e}	0.111^{c}	NA	0.021^{a}	−0.030^{a}	0.056^{b}	−0.003^{a}	0.059^{b}	0.114^{c}	−0.030^{a}
Precision (MARE)										
LC10	0.028^{a}	0.468^{e}	0.102^{c}	0.140^{c}	0.246^{d}	0.272^{d}	0.260^{d}	0.498^{e}	0.442^{e}	0.429^{e}
LC1	0.072^{b}	0.484^{e}	0.130^{c}	0.233^{d}	0.216^{d}	0.209^{d}	0.133^{c}	0.301^{e}	0.311^{e}	0.167^{c}
LBSPR10	0.570^{e}	0.138^{c}	NA	0.050^{b}	0.066^{b}	0.094^{b}	0.048^{a}	0.078^{b}	0.117^{c}	0.050^{b}
LBSPR1	0.603^{e}	0.154^{c}	NA	0.152^{c}	0.113^{c}	0.139^{c}	0.138^{c}	0.120^{c}	0.157^{c}	0.149^{c}

Note: We also compared the annual model runs with monthly length data run on a monthly time step for the short-lived life-history type. Bias (top) is measured as MRE and precision (bottom) is measured as MARE from 100 iterations of generated data.

[a] Bias/precision less than 5%.

[b] Bias/precision less than 10%.

[c] Bias/precision less than 20%.

[d] Bias/precision less than 30%.

[e] Bias/precision greater than 30%.

LIME on a monthly time step with monthly length data decreased bias compared with running on an annual time step (from 0.468 to −0.040 with 10 years of monthly length data and from 0.484 to 0.034 with 1 year of monthly length data) (Table 7).

Discussion

This study demonstrates that LIME can be a valuable tool for fisheries stock assessment when at least 1 year of length data and basic biological information are available for the species (Table 8). LIME estimates reference points more accurately than LB-SPR under many common scenarios. The LIME assessment method has several characteristics that make it particularly useful in situations where an abundance index and (or) catch data cannot be collected reliably, as data collection programs improve to include more data types and when recruitment and fishing mortality are believed to be changing more rapidly than would match the equilibrium assumptions required for existing methods with the same data requirements.

The *length-based* aspect of the model setup allows the analyst to fit to length-, rather than age-, composition data from the catch. The current toolbox of data-limited stock assessment methods includes length-only methods with equilibrium assumptions (Hordyk et al. 2015; Kokkalis et al. 2015; Nadon et al. 2015), ad hoc assumptions about changes in mortality (Gedamke and Hoenig 2006), or age-structured models that cannot fit to length-composition data (Martell and Froese 2013; Thorson and Cope 2014). MULTIFAN (Fournier et al. 1990) and Stock Synthesis (Methot and Wetzel 2013) can use length instead of age data to gain information on cohort strength and total mortality but require a catch time series to estimate stock status. The flexibility of LIME to fit to length composition, as opposed to age composition, is more realistic for capacity-limited fisheries. LIME can be used in conjunction with

Fig. 7. Comparison of the LIME and LB-SPR methods in the equilibrium state via the distribution of relative error (estimated – true)/true) for SPR in the current year for 100 iterations of simulated populations from the LIME age-structured operating model using annual length composition collected instantaneously in the year (*a*–*c*), monthly length data collection pooled into annual length compositions (*d*–*f*), and the LB-SPR operating model based on relative ages (*g*–*i*). The methods are compared across life-history types. Darker colors represent data availability scenarios with 10 years of length data, and lighter colors represent the scenario with one year of length data available. Each scenario has a different *y*-range. [Colour online.]

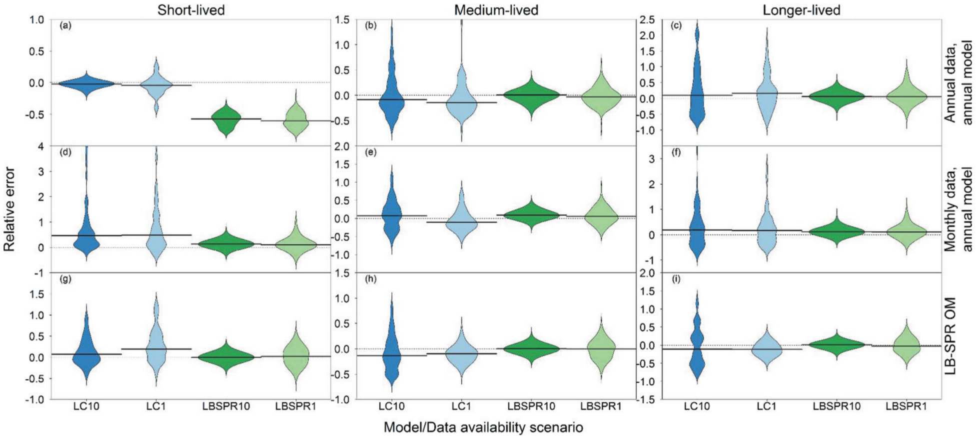


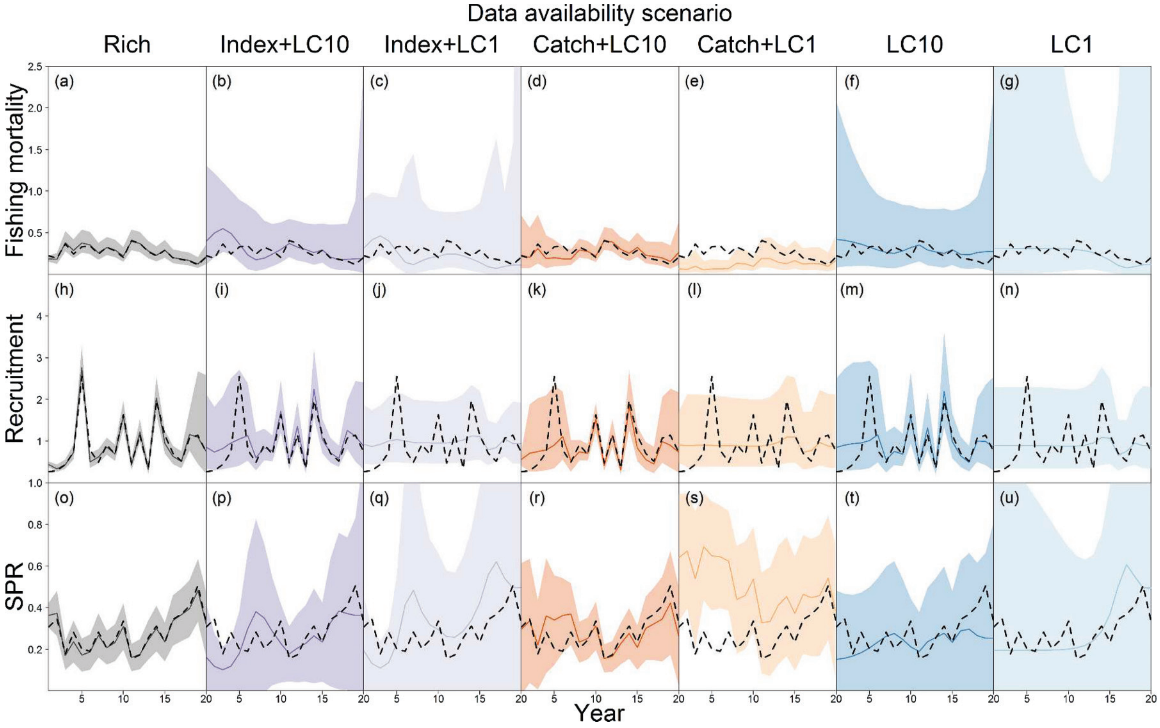
Table 8. Summary table of key questions and conclusions.

Objective	Conclusion
1. Is LIME unbiased:	
a. across life-history types?	Yes, if the assumption of instantaneous length sampling is met and growth parameters are known with no error. Performance is best for life histories with longevity less than 20 years under this assumption
b. with population variability?	Yes. It is recommended to use more than 1 year of length data for short- and medium-lived life-history types and include catch or an abundance index if longer-lived
c. with inclusion of catch and (or) an abundance index?	Yes, with more than 1 year of length data for short-lived species. May overestimate SPR for medium- and longer-lived species if the population is not in equilibrium
2. Is LIME sensitive to:	
a. sample size of length measurements?	Yes. Accuracy and precision generally improve with higher annual sample size. There is a performance plateau for 100 samples for short-lived fish and 500 samples for medium- or longer-lived fish
b. error in input parameters?	Relatively insensitive to error in the CV of the age–length curve. If asymptotic length, von Bertalanffy growth coefficient, or length at 50% maturity is input lower than the truth or natural mortality is input higher than the truth, SPR will be estimated higher than the truth (and vice versa)
3. Comparison of LIME and LB-SPR under equilibrium conditions	
a. Instantaneous annual sampling from age-structured model	LIME unbiased across life-history types; LB-SPR estimates lower SPR for short-lived life history
b. Continuous sampling (monthly) from age-structured model	LIME should use monthly time steps for short-lived species but can use annual time steps for medium- and longer-lived species
c. Length-structured operating model	LIME overestimates SPR for short-lived and underestimates SPR for medium- and longer-lived. LB-SPR unbiased across life-history types

other length-based methods as a diagnostic tool to see if variations in recruitment or fishing mortality is being predicted by the model based on the length data and how the results of the multiple assessment types would vary in light of those possibly violated assumptions.

The *integrated* nature of LIME is useful in situations where monitoring programs are continuing to be developed over time, incorporating more years of length data or other data types in the same assessment framework. LB-SPR is a quick way of conducting an assessment with only 1 year of length data, but as monitoring programs improve, the method must be applied independently over multiple years (Prince et al. 2015*a*, 2015*b*). There are also many catch-based methods that assume that a catch time series is available, but they are not thoroughly tested in fitting to length-, rather than age-, composition data (Martell and Froese 2013; Sabater and Kleiber 2014; Thorson and Cope 2014). LIME allows for these data to be included into the same assessment framework rather than requiring a switch between assessment models as monitoring programs develop and new data types become available (Maunder and Punt 2013). Also, few existing models can incorporate an abundance index and length data without catch data. LIME can be used in this scenario to estimate a relative SPR reference point with no measure of scale. This scenario is common for small-scale fisheries, bycatch species, or other fisheries where trends in abundance are available but total harvest is not.

Fig. 8. Example of model estimates (colour lines) with 95% confidence intervals (shaded areas) and the true simulated population dynamics (broken line) for each of the data availability scenarios for one iteration of a simulated population with a medium-lived life history, 200 length measurements annually in the length data, for the base variation scenario. [Colour online.]



Finally, the *mixed effects* aspect of LIME extends length-based methods by estimating changes in recruitment and fishing mortality over time. Estimating random effects has been integrated into existing stock assessment models, such as Stock Synthesis (Thorson et al. 2015*a*) and the state-space assessment model framework (Nielsen and Berg 2014). Many assessments for European stocks now use the state-space assessment model to separate process and observation errors as an objective method of weighting data in age-structured models (Berg et al. 2014). Computing the marginal likelihood for mixed-effects models was previously too computationally challenging, hindering wide application of mixed-effects models. Now, programs such as TMB can do these computations much faster than previous options (Nielsen and Berg 2014). To illustrate the benefits of LIME, a vignette is available on the repository site that walks the user through a simple example, with tips for model interpretation and convergence. An app using the R package *shiny* (Chang et al. 2017) is also available on the repository site (https://github.com/merrillrudd/LIME_shiny) that allows analysts to run the model in a graphical user interface.

Simulation testing in this study demonstrated the best LIME performance with only length data for the short-lived life-history type, with comparatively lower performance for medium- and longer-lived fish. With only length data, LIME performs well for the shorter-lived fish likely because the model is tracking cohorts through the length data to estimate recruitment deviations. This is likely difficult for the longer-lived life histories due to variation in the age–length relationship (especially for older fish) and because each individual cohort represents a small proportion of total abundance (hence requiring more samples to track each individual cohort). The increasing uncertainty in a longer-lived fish's age with increasing length blurs the cohorts as they age, making it difficult to track recruitment events without a much longer time series of length data. With 10 years of length data, accuracy in estimating SPR increases but precision decreases for the medium- and longer-lived life-history types, providing support for the increased ability to track recruitment events but propagating uncertainty in the age–length curve for older individuals. With short-lived fish that only live to 4 years old, each length more clearly matches up with an age group. Regardless of fish growth, 1 year of length data holds information on an entire generation for a short-lived fish. Even if recruitment is occurring monthly or seasonally and not annually, a 10 year time series of length data can inform variation in cohort strength across several generations of the short-lived fish.

By testing LIME and LB-SPR using multiple operating models, we identified scenarios in which each method excels. LB-SPR performs better than LIME for estimating SPR if the population is at equilibrium and only length data are available. If the population is not in equilibrium, then LIME performs better. If the data are collected continuously throughout the year (e.g., monthly), LIME is expected to estimate higher SPR than the truth for short-lived fish. In this case, analysts should run LIME using monthly (or shorter) time steps to account for fish growth during the year. This bias does not occur for medium- or longer-lived fish, where the growth during the year is less rapid. On the other hand, LB-SPR likely underestimates SPR for short-lived fish if data are collected instantaneously but is unbiased if data are collected continuously.

An important question for management is often how many years or how much data are enough, particularly for fisheries where funding for monitoring is limited. The simulation testing in this study demonstrated that we can get accurate but highly uncertain estimates of SPR, fishing mortality, and recruitment from only 1 year of length data on average (Fig. 8). As expected, collecting more data (10 years of length data instead of 1) resulted in greater accuracy and precision in LIME estimates of SPR (Fig. 4; Table 5). When using only length data, LIME performance is best with at least 500 independent length measurements per year for medium- and longer-lived life histories but is expected to perform well with 100 or more length measurements per year for short-lived life-history types (Table 6). However, including an abundance index or catch time series improves bias and precision to a greater extent than collecting more independent length measurements (Tables 5 and 6). Managers must determine how much uncertainty they are comfortable with to set monitoring goals. Collecting more years of length data, taking more independent length measurements during each year of length data collection, conducting surveys, and monitoring catch data hold varying levels of information used to estimate variable fishing mortality and recruitment. Their costs and benefits should be weighed to help managers prioritize data collection.

Including an abundance index or catch time series improves LIME performance in most scenarios. Catch data provide useful information on the scale of the population, which is not possible with length data alone. The abundance index provides useful information to help inform the trajectory of the population from which the length data arose. The abundance index could be more informative on the state of the population than length data alone, as long as the abundance index is proportional to abundance and any significant changes in fishing mortality and recruitment occurred during the surveyed time series. Our simulation study shows that the biases associated with including a time series with 1 year of length data are often overcome with more years of length data. The exception is for the longer-lived life-history type for which LIME performs worse when including an abundance index with 10 years of length composition data than if the abundance index was excluded.

However, the simulation study also identified some potential issues with data conflict when an abundance index or catch time series is included with length data. Data conflict can be diagnostic of poor data quality, such as catch misreporting, indices from spatial areas with ontogenetic differences, or length data not representative of the fishery. Data quality is an equally important issue as limited data types and should be considered in LIME applications. In the case of this study, however, poor data quality was not the culprit behind any data conflict because we used the true catch data, abundance index from a single area, and representative length data to test LIME. To avoid data conflict, Maunder and Piner (2017) recommend modeling process error explicitly, most commonly via time-varying recruitment, as a better alternative to down-weighting or eliminating data conflicts. LIME takes this approach, but data conflict may still occur, particularly because recruitment deviations are treated as a random effect even if this is not the correct model process for which the data hold conflicting information (Maunder and Piner 2017). Other options to avoid data conflict are to estimate the variance parameter for observed data outside of the stock assessment model (Lee et al. 2014; Maunder and Piner 2017). Alternatively, an analyst could identify whether data conflict is occurring in the LIME model by likelihood profiling individual data components (Ichinokawa et al. 2014) or retrospective analysis, which quantifies the impact of additional years of data on the stock assessment output (Hurtado-Ferro et al. 2014).

Like all length-based assessment methods, fixing the biological parameters at their true values is an important first step for estimating unbiased reference points. LB-SPR circumvents this strong assumption by using Beverton–Holt life-history invariants to estimate the expected age or length structure, as opposed to assuming known values of natural mortality and growth (Prince et al. 2015*a*). Sensitivity tests in this study demonstrating the impact of misspecifying biological parameters provide support for local studies of species growth rates for stocks assessed using LIME. Analysts should be aware of how SPR is expected to be biased given their assumptions on fixed values for input parameters or model structure. A next step for LIME is to use Bayesian priors on biological parameters to more thoroughly represent the uncertainty in population parameter estimates relevant to management (e.g., from FishLife; Thorson et al. 2017*b*). Sensitivity tests and likelihood profiles should be conducted on different levels of dome-shaped selectivity to understand how SPR is expected to be biased if the model structure is misspecified.

Length-based stock assessments are good starting points for making management decisions with limited data and monitoring capacity. The shrinkage of poorly estimated parameters towards an estimated distribution (as implemented within mixed-effects models) has been shown to increase accuracy and precision in stock assessments (Thorson et al. 2013; Nielsen and Berg 2014), but high uncertainty in estimates of stock status will always result if there are only 1 or 2 years of length data. Any stock assessment deals with data limitation and uncertainty and it is vital to appropriately represent and communicate this uncertainty to managers.

Ideally, conclusions on stock status should be drawn using an ensemble of assessment models with varying structure and assumptions (Stewart and Martell 2015; Anderson et al. 2017). In the data-limited context, existing models represent a relatively small range of alternative model assumptions. We have shown that LIME presents a way to represent both process and observation uncertainty. LIME can complement other length-based (Gedamke and Hoenig 2006; Hordyk et al. 2015; Nadon et al. 2015) and catch-only methods (Carruthers et al. 2014; Rosenberg et al. 2014) with the strength of estimating recruitment variability. More accurate and precise estimates of recruitment variability can help decipher whether decreased mean length is due to fishing pressure or recruitment variability and understanding the range of possible levels of recruitment into the future. We therefore conclude that LIME is a step forward in dealing with uncertainty in decision-making for fisheries where length data are collected.

Acknowledgements

M.B.R. was funded by the National Science Foundation IGERT Program on Ocean Change and a University of Washington School of Aquatic and Fishery Sciences fellowship. We thank J. Cope, T. Branch, R. Hilborn, J. Hastie, M. McClure, and T.R. McClanahan for their early review of this manuscript and A. Hordyk for discussions and feedback throughout the method development process.

References

Anderson, S.C., Cooper, A.B., Jensen, O.P., Minto, C., Thorson, J.T., Walsh, J.C., Afflerbach, J., Dickey-Collas, M., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Mosqueira, I., Rosenberg, A.A., and Selig, E.R. 2017. Improving estimates of population status and trend with superensemble models. Fish Fish. **18**(4): 732–741.

Ault, J.S., Smith, S.G., and Bohnsack, J.A. 2005. Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. ICES J. Mar. Sci. **62**(3): 417–423. doi:10.1016/j.icesjms.2004.12.001.

Ault, J.S., Smith, S.G., Luo, J., Monaco, M.E., and Appeldoorn, R.S. 2008. Length-based assessment of sustainability benchmarks for coral reef fishes in Puerto Rico. Environ. Conserv. **35**(3): 221–231. doi:10.1017/S0376892908005043.

Azzurro, E., Moschella, P., and Maynou, F. 2011. Tracking signals of change in Mediterranean fish diversity based on local ecological knowledge. PLoS One, **6**(9): e24885. doi:10.1371/journal.pone.0024885. PMID:21966376.

Berg, C.W., Nielsen, A., and Kristensen, K. 2014. Evaluation of alternative age-based methods for estimating relative abundance from survey data in relation to assessment models. Fish. Res. **151**: 91–99. doi:10.1016/j.fishres.2013.10.005.

Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Ministry of Agriculture, Lowestoft, UK.

Buckland, S.T., Newman, K.B., Thomas, L., and Koesters, N.B. 2004. State-space models for the dynamics of wild animal populations. Ecol. Modell. **171**: 157–175. doi:10.1016/j.ecolmodel.2003.08.002.

Butler, J.R.A., Middlemas, S.J., Graham, I.M., Thompson, P.M., and Armstrong, J.D. 2006. Modelling the impacts of removing seal predation from Atlantic salmon, *Salmo salar*, rivers in Scotland: a tool for targeting conflict resolution. Fish. Manag. Ecol. **13**(5): 285–291. doi:10.1111/j.1365-2400.2006.00504.x.

Bystrom, A.B. 2015. Análisis de características biológico-pesqueras del pargo manchado (*Lutjanus guttatus* (Steindachner, 1869)) y tendencias socio-ecológicas de la pesca artesanal con líneas de fondo en el distrito de Bejuco, Pacífico de Costa Rica. M.Sc. thesis, Manejo de Recursos Naturales, Universidad Estatal a Distancia, San Jose, Costa Rica.

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

Chang, W., Cheng, J., Allaire, J.J., Xie, Y., and McPherson, J. 2017. Shiny: web application framework for R. R package version 1.0.0. https://CRAN.R-project.org/package=shiny.

Chrysafi, A., and Kuparinen, A. 2016. Assessing abundance of populations with limited data: lessons learned from data-poor fisheries stock assessment. Environ. Rev. **24**: 25–38. doi:10.1139/er-2015-0044.

Clark, W.G. 2002. $F_{35\%}$ revisited ten years later. North Am. J. Fish. Manage. **22**(1): 251–257. doi:10.1577/1548-8675(2002)022<0251:FRTYL>2.0.CO;2.

Costello, C., Ovando, D., Hilborn, R., Gaines, S.D., Deschenes, O., and Lester, S.E. 2012. Status and solutions for the world's unassessed fisheries. Science, **338**: 517–520. doi:10.1126/science.1223389. PMID:23019613.

Darcy, G.H., and Matlock, G.C. 1999. Application of the precautionary approach in the national standard guidelines for conservation and management of fisheries in the United States. ICES J. Mar. Sci. **56**(6): 853–859. doi:10.1006/jmsc.1999.0533.

de Valpine, P., and Hastings, A. 2002. Fitting population models incorporating process noise and observation error. Ecol. Monogr. **72**(1): 57–76. doi:10.1890/0012-9615(2002)072[0057:FPMIPN]2.0.CO;2.

Dowling, N.A., Dichmont, C.M., Haddon, M., Smith, D.C., Smith, A.D.M., and Sainsbury, K. 2015. Empirical harvest strategies for data-poor fisheries: a review of the literature. Fish. Res. **171**: 141–153. doi:10.1016/j.fishres.2014.11.005.

Dowling, N.A., Wilson, J.R., Rudd, M.B., Babcock, E.A., Dougherty, D., and Gleason, M. 2016. FishPath: a decision support system for assessing and managing data- and capacity-limited fisheries. *In* Assessing and managing data-limited fish stocks. *Edited by* T.J.I. Quinn, J.L. Armstrong, M.R. Baker, J. Heifetz, and D. Witherell. Alaska Sea Grant, University of Alaska Fairbanks. pp. 59–96. doi:10.4027/amdlfs.2016.03.

Ehrhardt, N.M., and Ault, J.S. 1992. Analysis of two length-based mortality models applied to bounded catch length frequencies. Trans. Am. Fish. Soc. **121**: 115–122. doi:10.1577/1548-8659(1992)121<0115:AOTLMM>2.3.CO;2.

FAO. 2016. The state of world fisheries and aquaculture 2016. Contributing to food security and nutrition for all. FAO, Rome.

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

Francis, R.I.C.C. 2014. Replacing the multinomial in stock assessment models: a first step. Fish. Res. **151**: 70–84. doi:10.1016/j.fishres.2013.12.015.

Froese, R. 1990. FishBase: an information system to support fisheries and aquaculture research. ICLARM Fishbyte, **8**(3): 21–24.

Gedamke, T., and Hoenig, J.M. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. Trans. Am. Fish. Soc. **135**(2): 476–487. doi:10.1577/T05-153.1.

Goodyear, C.P. 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. *In* Risk evaluation and biological reference points for fisheries management. *Edited by* S.J. Smith, J.J. Hunt, and D. Rivard. Can. Spec. Publ. Fish. Aquat. Sci. No. 120.

Gulbrandsen, L.H. 2009. The emergence and effectiveness of the Marine Stewardship Council. Mar. Policy, **33**(4): 654–660. doi:10.1016/j.marpol.2009.01.002.

Harley, S.J., Myers, R.A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Can. J. Fish. Aquat. Sci. **58**(9): 1760–1772. doi:10.1139/f01-112.

Heemstra, P.C., and Randall, J.E. 1993. FAO Species Catalogue. Vol. 16. Groupers of the world (family Serranidae, subfamily Epinephelinae). An annoted and illustrated catalogue of the grouper, rockcod, hind, coral grouper, and lyretail species known to date. Fish. Synop. **125**(16). FAO, Rome.

Hicks, C.C., and McClanahan, T.R. 2012. Assessing gear modifications needed to optimize yields in a heavily exploited, multi-species, seagrass and coral reef fishery. PLoS ONE, **7**(5): e36022. doi:10.1371/journal.pone.0036022. PMID:22574133.

Hordyk, A., Ono, K., Valencia, S., Loneragan, N., and Prince, J. 2015. A novel length-based empirical estimation method of spawning potential ratio (SPR) and tests of its performance, for small-scale, data-poor fisheries. ICES J. Mar. Sci. **72**(1): 217–231. doi:10.1093/icesjms/fsu004.

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., and Punt, A.E. 2014. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES J. Mar. Sci. **72**(1): 99–110. doi:10.1093/icesjms/fsu198.

Ichinokawa, M., Okamura, H., and Takeuchi, Y. 2014. Data conflict caused by model mis-specification of selectivity in an integrated stock assessment model and its potential effects on stock status estimation. Fish. Res. **158**: 147–157. doi:10.1016/j.fishres.2014.02.003.

Kokkalis, A., Thygesen, U.H., Nielsen, A., and Andersen, K.H. 2015. Limits to the reliability of size-based fishing status estimation for data-poor stocks. Fish. Res. **171**: 4–11. doi:10.1016/j.fishres.2014.10.007.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B. 2016. TMB: automatic differentiation and Laplace approximation. J. Stat. Softw. **70**: 1–21. doi:10.18637/jss.v070.i05.

Lee, H.H., Piner, K.R., Methot, R.D., and Maunder, M.N. 2014. Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: an example using blue marlin in the Pacific Ocean. Fish. Res. **158**: 138–146. doi:10.1016/j.fishres.2013.12.017.

Mahon, R. 1997. Does fisheries science serve the needs of managers of small stocks in developing countries. Can. J. Fish. Aquat. Sci. **54**(9): 2207–2213. doi:10.1139/f97-112.

Martell, S., and Froese, R. 2013. A simple method for estimating MSY from catch and resilience. Fish Fish. **14**(4): 504–514. doi:10.1111/j.1467-2979.2012.00485.x.

Maunder, M.N., and Piner, K.R. 2017. Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. Fish. Res. **192**: 16–27. doi:10.1016/j.fishres.2016.04.022.

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

Melnychuk, M.C., Peterson, E., Elliott, M., and Hilborn, R. 2017. Fisheries management impacts on target species status. Proc. Natl. Acad. Sci. **114**: 178–183. doi:10.1073/pnas.1609915114.

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

Methot, R.D., Tromble, G.R., Lambert, D.M., and Greene, K.E. 2014. Implementing a science-based system for preventing overfishing and guiding sustainable fisheries in the United States. ICES J. Mar. Sci. **71**(2): 183–194. doi:10.1093/icesjms/fst119.

Nadon, M.O., Ault, J.S., Williams, I.D., Smith, S.G., and DiNardo, G.T. 2015. Length-based assessment of coral reef fish populations in the Main and Northwestern Hawaiian Islands. PLoS One, **10**(8): e0133960. doi:10.1371/journal.pone.0133960. PMID:26267473.

Neis, B. 1992. Fishers' ecological knowledge and stock assessment in Newfoundland. Newfoundland Studies, **8**(2): 155–178.

Nielsen, A., and Berg, C.W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fish. Res. **158**: 96–101. doi:10.1016/j.fishres.2014.01.014.

Ono, K., Punt, A.E., and Rivot, E. 2012. Model performance analysis for Bayesian biomass dynamics models using bias, precision and reliability metrics. Fish. Res. **125–126**: 173–183.

Prince, J., Hordyk, A., Valencia, S.R., Loneragan, N., and Sainsbury, K. 2015*a*. Revisiting the concept of Beverton–Holt life-history invariants with the aim of informing data-poor fisheries assessment. ICES J. Mar. Sci. **72**(1): 194–203. doi:10.1093/icesjms/fsu011.

Prince, J., Victor, S., Kloulchad, V., and Hordyk, A. 2015*b*. Length based SPR assessment of eleven Indo-Pacific coral reef fish populations in Palau. Fish. Res. **171**: 42–58. doi:10.1016/j.fishres.2015.06.008.

Quinn, T.J.I., Armstrong, J.L., Baker, M., Heifetz, J., and Witherell, D. 2016. Assessing and managing data-limited fish stocks. Alaska Sea Grant, University of Alaska Fairbanks.

R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.r-project.org/.

Ricard, D., Minto, C., Jensen, O.P., and Baum, J.K. 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish Fish. **13**(4): 380–398. doi:10.1111/j.1467-2979.2011.00435.x.

Rosenberg, A.A., Fogarty, M.J., Cooper, A.B., Dickey-Collas, M., Fulton, E.A., Gutierrez, N.L., Hyde, K.J.W., Kleisner, K.M., Kristiansen, T., Longo, C., Minte-Vera, C., Minto, C., Mosqueira, I., Chato Osio, G., Ovando, D., Selig, E.R., Thorson, J.T., and Ye, Y. 2014. Developing new approaches to global stock status assessment and fishery production potential of the seas. *In* FAO Fisheries and Aquaculture Circular No. 1086. FAO, Rome.

Sabater, M., and Kleiber, P. 2014. Augmented catch-MSY approach to fishery management in coral-associated fisheries. *In* Interrelationships between corals and fisheries. *Edited by* S.A. Bortone. CRC Press. pp. 199–218. doi:19.1201/b17159-12.

Schnute, J.T., and Haigh, R. 2007. Compositional analysis of catch curve data, with an application to Sebastes maliger. ICES J. Mar. Sci. **64**(2): 218–233. doi:10.1093/icesjms/fsl024.

Stewart, I.J., and Martell, S.J.D. 2015. Reconciling stock assessment paradigms to better inform fisheries management. ICES J. Mar. Sci. **72**(8): 2187–2196.

Szuwalski, C.S., Vert-Pre, K.A., Punt, A.E., Branch, T.A., and Hilborn, R. 2015. Examining common assumptions about recruitment: a meta-analysis of recruitment dynamics for worldwide marine fisheries. Fish Fish. **16**(4): 633–648. doi:10.1111/faf.12083.

Thorson, J.T., and Cope, J.M. 2014. Catch curve stock-reduction analysis: an alternative solution to the catch equations. Fish. Res. **171**: 33–41. doi:10.1016/j.fishres.2014.03.024.

Thorson, J.T., and Kristensen, K. 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. Fish. Res. **175**: 66–74. doi:10.1016/j.fishres.2015.11.016.

Thorson, J.T., and Minto, C. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES J. Mar. Sci. **72**(5): 1245–1256. doi:10.1093/icesjms/fsu213.

Thorson, J.T., Minto, C., Minte-Vera, C.V., Kleisner, K.M., and Longo, C. 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. Can. J. Fish. Aquat. Sci. **70**(12): 1829–1844. doi:10.1139/cjfas-2013-0280.

Thorson, J.T., Cope, J.M., and Patrick, W.S. 2014*a*. Assessing the quality of life history information in publicly available databases. Ecol. Appl. **24**: 217–226. doi:10.1890/12-1855.1. PMID:24640546.

Thorson, J.T., Jensen, O.P., and Zipkin, E.F. 2014*b*. How variable is recruitment for exploited marine fishes? A hierarchical model for testing life history theory. Can. J. Fish. Aquat. Sci. **71**(7): 973–983. doi:10.1139/cjfas-2013-0645.

Thorson, J.T., Skaug, H., Kristensen, K., Shelton, A., Ward, E., Harms, J., and Benante, J. 2015*a*. The importance of spatial models for estimating the strength of density dependence. Ecology, **96**(5): 1202–1212. doi:10.1890/14-0739.1. PMID:26236835.

Thorson, J.T., Hicks, A.C., and Methot, R.D. 2015*b*. Random effect estimation of time-varying factors in Stock Synthesis. ICES J. Mar. Sci. **72**: 178–185. doi:10.1093/icesjms/fst211.

Thorson, J.T., Johnson, K.F., Methot, R.D., and Taylor, I.G. 2017*a*. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fish. Res. **192**: 84–93. doi:10.1016/j.fishres.2016.06.005.

Thorson, J.T., Munch, S.B., Cope, J.M., and Gao, J. 2017*b*. Predicting life history parameters for all fishes worldwide. Ecol. Appl. **27**(8): 2262–2276. doi:10.1002/eap.1606.

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. Proc. Natl. Acad. Sci. U.S.A. **110**(5): 1779–1784. doi:10.1073/pnas.1214879110. PMID:23322735.

Wetzel, C.R., and Punt, A.E. 2011. Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. Fish. Res. **110**(2): 342–355. doi:10.1016/j.fishres.2011.04.024.

Appendix

Fig. A1. Size at age for the yearly (blue) versus monthly (white–orange) time steps for the short-lived life-history type. White lines (monthly) overlapping blue lines (yearly) demonstrate that size at age is identical at the start of each year and orange lines increasing in density indicate the size at age distributions during the following months within each year. [Colour online.]