

## Original Article

# Performance evaluation of data-limited, length-based stock assessment methods

Lisa Chong <sup>1,2\*</sup>, Tobias K. Mildenerger <sup>3</sup>, Merrill B. Rudd<sup>4,5</sup>, Marc H. Taylor <sup>6</sup>, Jason M. Cope<sup>7</sup>, Trevor A. Branch<sup>4</sup>, Matthias Wolff<sup>2</sup>, and Moritz Stäbler<sup>2</sup>

<sup>1</sup>Fisheries and Aquatic Sciences Program, School of Forest Resources and Conservation University of Florida, Gainesville, FL 32603, USA

<sup>2</sup>Leibniz Centre for Tropical Marine Research (ZMT), Fahrenheitstraße 6, 28359 Bremen, Germany

<sup>3</sup>National Institute of Aquatic Resources, Technical University of Denmark, 2800 Kgs. Lyngby, Kemitorvet, Denmark

<sup>4</sup>School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA

<sup>5</sup>Scaleability LLC, 4425 Letitia Avenue South, Seattle, WA 98118, USA

<sup>6</sup>Marine Living Resources, Thünen Institute of Sea Fisheries, Herwigstr. 31, 27572 Bremerhaven, Germany

<sup>7</sup>Fishery Resource Analysis and Monitoring Division, National Oceanic and Atmospheric Administration, National Marine Fisheries Service Northwest Fisheries Science Center, 2725 Montlake Boulevard East, Seattle, WA 98112-2097, USA

\*Corresponding author: tel: +1 201 221 6117; e-mail: [lisa.chong8594@gmail.com](mailto:lisa.chong8594@gmail.com).

Chong, L., Mildenerger, T. K., Rudd, M. B., Taylor, M. H., Cope, J. M., Branch, T. A., Wolff, M., and Stäbler, M. Performance evaluation of data-limited, length-based stock assessment methods. – ICES Journal of Marine Science, 77: 97–108.

Received 24 July 2019; revised 10 October 2019; accepted 11 October 2019; advance access publication 11 November 2019.

Performance evaluation of data-limited, length-based methods is instrumental in determining and quantifying their accuracy under various scenarios and in providing guidance about model applicability and limitations. We conducted a simulation–estimation analysis to compare the performance of four length-based stock assessment methods: length-based Thompson and Bell (TB), length-based spawning potential ratio (LBSPR), length-based integrated mixed effects (LIME), and length-based risk analysis (LBRA), under varying life history, exploitation status, and recruitment error scenarios. Across all scenarios, TB and LBSPR were the most consistent and accurate assessment methods. LBRA is highly biased, but precautionary, and LIME is more suitable for assessments with time-series longer than a year. All methods have difficulties when assessing short-lived species. The methods are less accurate in estimating the degree of recruitment overfishing when the stocks are severely overexploited, and inconsistent in determining growth overfishing when the stocks are underexploited. Increased recruitment error reduces precision but can decrease bias in estimations. This study highlights the importance of quantifying the accuracy of stock assessment methods and testing methods under different scenarios to determine their strengths and weaknesses and provides guidance on which methods to employ in various situations.

**Keywords:** data-limited fishery, individual-based model, length-based assessment, MSY, simulation–estimation analysis, spawning potential ratio

## Introduction

Fisheries are considered data-limited if there are insufficient data to conduct a comprehensive quantitative, model-based stock assessment to estimate time-series of biomass and fishing mortality relative to their reference points (Dowling *et al.*, 2019). Nevertheless, even with limited data, some aspects of stock status can be inferred. Data-limited assessment methods are increasingly used for management purposes to report on the regional status of

fisheries across many stocks and to assess the status of individual data-limited stocks as inputs to management decisions (Dowling *et al.*, 2015, 2019; Chrysafi and Kuparinen, 2016). In data-limited fisheries, length-frequency data from commercial catches are often the primary data type collected because they are relatively economical and easy to collect (Pilling *et al.*, 2008; Hordyk *et al.*, 2015a; Mildenerger *et al.*, 2017). As a result, numerous length-based methodologies have been developed. Prominent methods

include the length-based Thompson and Bell (TB) model (Thompson and Bell, 1934), length-based spawning potential ratio (LBSPR; Hordyk et al., 2015b), length-based integrated mixed effects (LIME; Rudd and Thorson, 2018), and length-based risk analysis (LBRA; Ault et al., 1998, 2008, 2019).

TB is a yield-per-recruit (YPR) model that evaluates stock status relative to fishing and selectivity reference levels using length composition data (Thompson and Bell, 1934; Sparre and Venema, 1998; Mildenerger et al., 2017). LBSPR is a well-known length-based model that assesses stock status by comparing the spawning potential as measured through the length composition data to that expected in an unfisher stock (Hordyk et al., 2015b). LIME relaxes the equilibrium assumptions of other methods, accounting for time-varying recruitment and fishing mortality (though assumes constant selectivity), and derives population parameters associated with an age-structured model and length compositions (Rudd and Thorson, 2018). LBRA uses the mean length of the catch to calculate reference points that address sustainability risks (Ault et al., 1998, 2008, 2019). These approaches can be used to estimate spawning potential ratio (SPR) and  $F/F_{MSY}$ , which are commonly used as indicators for recruitment overfishing and growth overfishing, respectively. SPR is the proportion of the unfisher reproductive potential left at any given level of fishing pressure (Hordyk et al., 2015b). SPR is 100% in an unexploited stock, and 0% in a stock with no spawning (e.g. all mature fish have been removed or all female fish have been caught).

The aim of this study is to quantify accuracy and precision for four length-based data-limited methods (TB, LBSPR, LIME, and LBRA) under various life history, exploitation, and recruitment scenarios given only a single year of length-frequency data, which is a limited field period common in very data-limited fisheries (Tsfaye et al., 2016; Herrón et al., 2018; Tuda, 2018; Abobi et al., 2019). This comparison allows for the identification of the method most suitable in different data-limited assessment scenarios by revealing the strengths and weaknesses of each method with reference to how well it captures true stock status, estimates key reference points, and characterizes uncertainty. It further helps to expose discrepancies in the performance of the methods (Cadrin and Dickey-Collas, 2015) and thus provides guidance about model applicability, expected bias, and advantages or disadvantages in uncertainty characterization. Although several studies have tested the performance of these methods through simulation testing (Hordyk et al., 2015b; Rudd and Thorson, 2018), this study differs in using an individual-based modelling (IBM) framework to track individuals in populations rather than using an approximation of lengths distributed in a population by age, thus offering a new and important test of these length-based approaches. The IBM framework also provides an alternative way of conducting a simulation–estimation analysis to ensure that the operating model is distinct from the estimation model.

## Methods

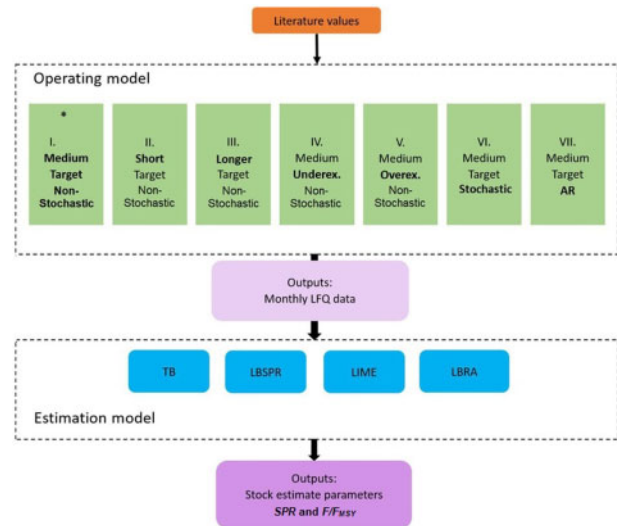
We conducted a simulation–estimation analysis using an individual-based population model as the operating model, which simulated population dynamics and generated length composition catch data. The assessment models were given the “true” (i.e. unbiased) input parameter values for the mean somatic growth and mortality rates and were then used to estimate various reference points and derived quantities relating to stock status and exploitation rate. This simulation loop allows us to

compare how far the outputs of the assessment models are from the “true” stock status estimates and investigate the sensitivities of the models to violated model assumptions. An overview of the experimental design is depicted in Figure 1.

## Operating model

The stock dynamics were simulated using the “fishdynr” R package (Taylor, 2017), which contains several models for simulating stock or population dynamics and fisheries management. The package’s function “virtualPop” creates an IBM of a fish stock with certain life history traits subjected to a fishing fleet with specific selectivity characteristics and fishing pressure. Information about the modelling approach for growth, mortality, selectivity, and recruitment is outlined by Taylor and Mildenerger (2017). Functions and equations for the population dynamics used in the operating model are listed in Table 1.

We simulated seven scenarios to test for the effects of variations in life history, fishing exploitation level, and recruitment types on assessment model performance. The base model was comprised of a medium-lived species (maximum age  $\approx 18$  years), an exploitation rate at the target level of SPR ( $\approx 40\%$ ), and constant recruitment with no recruitment variability. From this base model, we varied one of the three characteristics (life history, current exploitation status, or recruitment) according to the parameter values in Table 2. For each scenario, we generated 300 IBM replicates, which differed from each other due to random



**Figure 1.** Simulation study methodology diagram. There are seven operating model setups. Scenarios differ in life history with (I) medium-lived (base model\*), (II) short-lived, (III) and longer-lived stocks. All fish longevity simulations were run with constant recruitment and exploitation at target level ( $SPR \sim 0.4$ ). From the base model, the exploitation scenario, (IV) underexploited and (V) overexploited, or the recruitment scenario, (VI) stochastic, and (VII) autocorrelation (AR), changes. For each operating model, 1 year of monthly length-frequency data and the “true” stock estimate parameters were simulated and extracted. The “true” life history values (asymptotic length  $L_{\infty}$ , growth coefficient  $K$ , natural mortality  $M$ , and length at 50% maturity  $L_{50}^m$ ) from the operating models and the length-frequency data were then used as input for the simulated assessments with the four length-based estimation models: (i) length-based TB, (ii) LBSPR, (iii) LIME, and (iv) LBRA.

**Table 1.** Functions and population dynamic equations used for generating stocks and length-frequency data in the operating models.

Number	Description	Function/equation
1.1	von Bertalanffy growth function	$L_t = L_\infty [1 - \exp(-K(t - t_0))]$
1.2	Variability in $L_\infty$ of von Bertalanffy growth function	$L_{\infty,i} = L_{\infty,P} \times \text{Lognorm}(0, CV_{L_\infty})$
1.3	Variability in $K$ of von Bertalanffy growth function	$K_i = K_P \times \text{Lognorm}(0, CV_K)$
1.4	Length–weight relationship	$W = \alpha L^\beta$
1.5	Selectivity/maturity probability	$\text{Sel or Mat} = \frac{1}{1 + \exp\left[\frac{L_t - L_{50}^m}{\ln\left(\frac{0.75}{1-0.75}\right) - \ln\left(\frac{0.25}{1-0.25}\right)}\right]}$
1.6	Non-autocorrelated recruitment deviation	$R_t = \text{Lognorm}(0, \sigma_R^2)$
1.7	Autocorrelated recruitment deviation	$\tau_t = \epsilon_t; t = 1$
1.8		$\tau_t = \epsilon_t \tau_{t-1} \sqrt{1 - \rho^2}; t > 1$
1.9	Spawning-stock biomass	$SSB_t = \sum_{a=0}^A N_{a,t} w_a m_a$
1.10	Beverton–Holt relationship	$N_{recruits} = \frac{r_{max} BH * SSB_t}{beta + SSB_t}$ beta = 1 (constant recruitment)
1.11	Total mortality	$Z = M + F$
1.12	Fishing mortality	$F = q \times E$
1.13	Probability of death	$p_{death} = 1 - \exp(-Z \times t_{incr})$
1.14	$R_{and}$ = random number generated	If $R_{and} = p_{death}$ , individual dies
1.15	Probability of death due to either natural or fishing mortality ( $0 = M, 1 = F$ )	$p_{M \text{ or } F} = \text{sample}[c(0, 1), \text{prob} = c(\frac{M}{Z}, \frac{F}{Z})]$
1.16	Unfished expected lifetime egg production	$E_0 = \sum_{a=0}^A SSB_0 w_a m_a$
1.17	Fished expected lifetime egg production	$E_f = \sum_{a=0}^A SSB_f w_a m_a$
1.18	SPR	$SPR = \frac{E_f}{E_0}$
1.19	Yield	$Y_{F, L, t} = F_t B_{L, t}$

**Table 2.** Input values used to generate length-frequency data with the operating models.

Parameter	Changes	Asymptotic length ( $L_\infty$ ; cm)	Growth coefficient ( $K$ )	Age at length = 0 ( $t_0$ )	Natural mortality ( $M$ )	Fishing mortality ( $F$ )	Theoretical maximum age ( $A_{MAX}$ )
Life history	Base	64.6	0.21	-0.01	0.32	0.13	18
	Short	36.2	0.87	-0.01	0.9	0.45	4
	Longer	90	0.13	-0.01	0.18	0.08	26
Exploitation level	Underexploited	64.6	0.21	-0.01	0.32	0.06	18
	Overexploited	64.6	0.21	-0.01	0.32	0.28	18
Recruitment	Stochastic	64.6	0.21	-0.01	0.32	0.13	18
	AR	64.6	0.21	-0.01	0.32	0.13	18

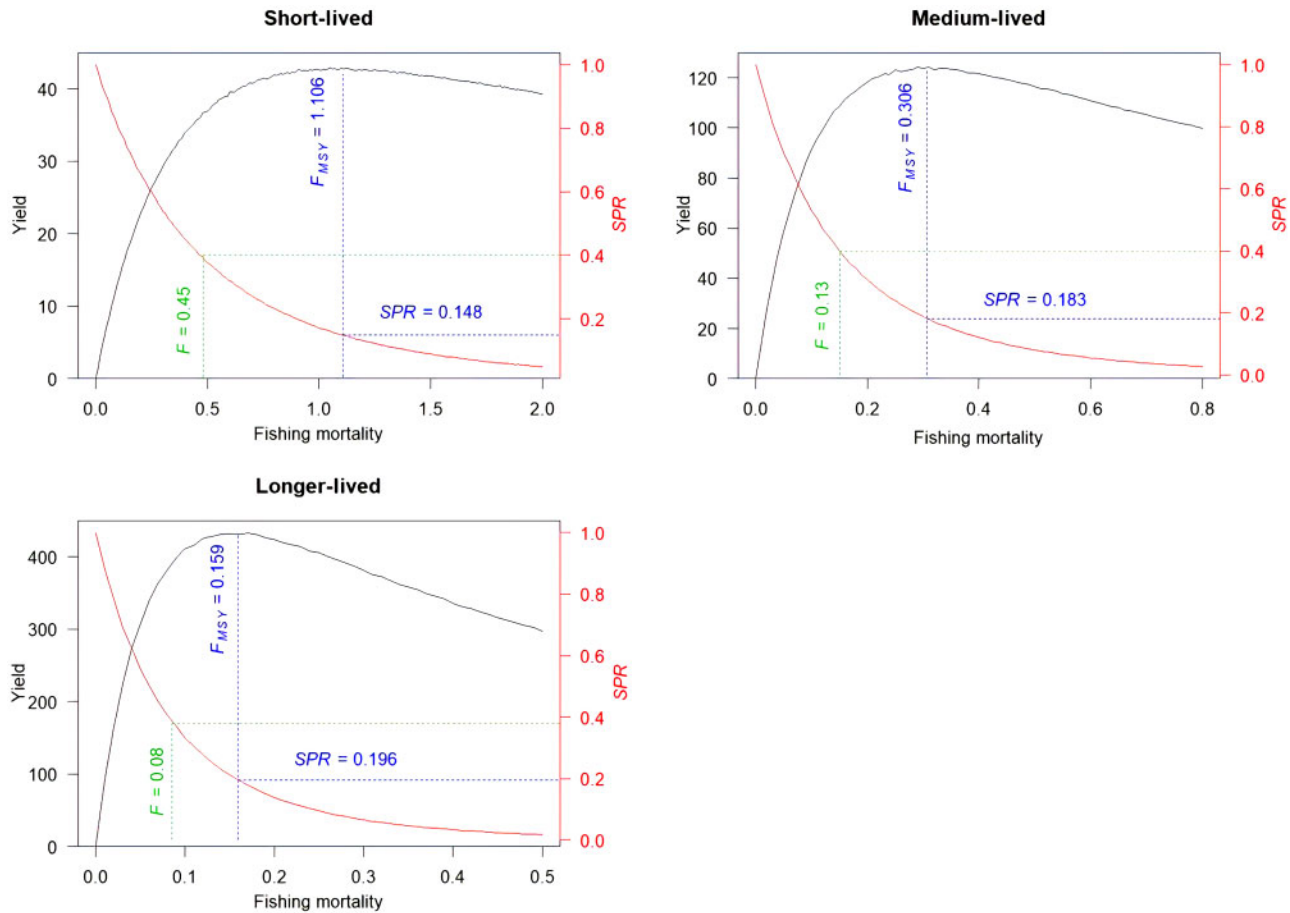
  

Parameter	Changes	Length at 50% maturity ( $L_{50}^m$ )	Width of maturity ogive (cm)	Recruitment standard deviation ( $\sigma_R$ )	Length at 50% selectivity ( $L_{50}^s$ )	Width of selectivity ogive (wqs; cm)	Bin size (cm)
Life history	Base	34	6.8	0.01	11	2.2	2
	Short	20.2	4.04	0.01	9	1.8	1
	Longer	50	10	0.01	20	4	3
Exploitation level	Underexploited	34	6.8	0.01	11	2.2	2
	Overexploited	34	6.8	0.01	11	2.2	2
Recruitment	Stochastic	34	6.8	0.4537	11	2.2	2
	AR	34	6.8	0.737	11	2.2	2

The base model is comprised of the medium-lived species at the target exploitation level with constant, non-stochastic recruitment. From the base model, changes in life history (short- or longer-lived), exploitation level (under- or overexploited), and recruitment type [stochastic (Thorson, in press) or autocorrelated (AR)] was made, respectively, to create the other six scenarios.

sampling. Each simulation replicate consisted of a 35-year simulation period, of which the first 10 years had no fishing activity and the remaining 25 years were fished at the desired exploitation

rate at a monthly time-step. This provided sufficient time for the model to reach equilibrium. We then only extracted the final year of each simulation replication (year 35) to provide monthly



**Figure 2.** Visual representation of the calculation and YPR of the “true” SPR and  $F_{MSY}$  based on three life histories. The “true” SPR and  $F_{MSY}$  were calculated by pushing the IBM forward 100 years and producing a YPR curve. The fishing mortality ( $F$ ) for each scenario was simulated based on this calculation as well.

length-frequency data from catches (200 individuals per month). For each simulation replicate, we calculated the “true” SPR and  $F/F_{MSY}$  based on life history (Figure 2; detailed description is provided by in the Supplementary Figure S1).

All operating models assumed von Bertalanffy growth and logistic-type selectivity and maturity. The three life histories were based on those used by Rudd and Thorson (2018) to test the LIME approach: *Siganus sutor* (Kenyan rabbitfish) for the short-lived (Hicks and McClanahan, 2012), *Lutjanus guttatus* (spotted rose snapper) for the medium-lived (Bystrom, 2016), and *Epinephelus morio* (red grouper) for the longer-lived (Heemstra and Randall, 1993) life history types. Examples of each life history are shown in Figure 3 as length-frequency graphs for one simulation replicate. All simulations and analyses were conducted using the statistical programming language R (R Core Team, 2018).

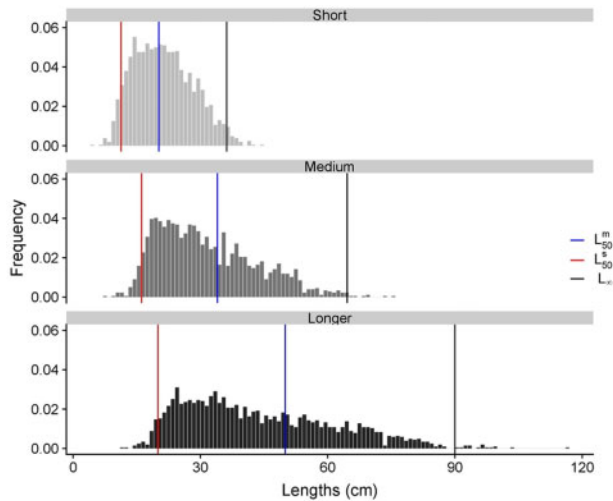
### Estimation models

The estimation models refer to the length-based assessment methods that derive estimates of stock status from sampled length-frequency catch data. All estimation models compared in this study assume the true values for life history parameters are known (e.g. growth, maturity, and natural mortality parameters; see Inputs in Table 3) and, therefore, the same values that were used in the operating models were supplied to each of the

methods as inputs. As a result, the estimation capabilities of each method are better than would be expected if these parameter values were estimated with uncertainty or incorrect, as is normally the case. Life history parameters, fishing mortality rate, and recruitment are often confounding factors affecting the shape of the length-frequency distribution; thus, the focus of this article is on sensitivities of methods estimating stock status based on an IBM approach alone rather than confounding it with biased life history parameters, because previous simulation studies have already shown the effects of incorrect life history parameters (Punt, 2003; Deroba and Schueller, 2013; Hordyk et al., 2015b; Rudd and Thorson, 2018).

The selectivity values  $L_{50}^s$  and  $L_{95}^s$ , lengths at 50% and 95% selectivity, respectively, were calculated using the length-converted catch curve (LCCC) from the R package TropFishR (Mildenberger et al., 2017) and used as the selectivity inputs for all estimation models. The LCCC method plots the natural log of catch vs. age, estimates total mortality ( $Z$ ) from the negative slope of the regression of the curve and derives a selection ogive. For TB,  $Z$  was calculated from the LCCC and used as an input value, similar to how TB is applied to data in reality. The four length-based methods assessed, (i) TB, (ii) LBSR, (iii) LIME, and (iv) LBRA, are contained within the R packages TropFishR, LBSR (Hordyk et al., 2015b), LIME (Rudd and Thorson, 2018), and





**Figure 3.** Length-frequency distribution graphs. For each life history scenario (short, medium, and longer),  $L_{50}^m$ ,  $L_{50}^s$ , and  $L_{\infty}$  are visualized (in blue, red, and black, respectively) over the length-frequency graphs for one iteration.

fishmethods (Nelson, 2017), respectively. The inputs, assumptions, and outputs of each of the methods, including the LCCC method, are listed in Table 3.

### Performance metrics

Many studies have explored the levels of *SPR* to be used as target and limit reference points, often applying *SPR* of 30% as a limit and 40% as a target reference point (Mace and Sissenwine, 1993; Clark, 2002; Hordyk et al., 2015c).  $F/F_{MSY}$  is a reference point relating the current fishing mortality ( $F$ ) to the level that would sustain maximum yield ( $F_{MAX}$ ). There is a common practice of linking  $F_{MAX}$  to  $F_{MSY}$  based on the assumption that recruitment is independent of spawning-stock size for fishing mortalities between 0 and  $F_{MAX}$  (Reynolds et al., 2001) and that selectivity is asymptotic and around the maturity ogive. These assumptions are often invalid for most species and many fisheries. However, in this study, density-dependent effects on recruitment were negligible, resulting in equivalency between  $F_{MAX}$  and  $F_{MSY}$ . Although reference points based on *MSY* are based on a measure of magnitude (i.e. catch),  $F_{MSY}$  can be calculated with length only instead of catch using relative YPR (Larkin, 1977; Holt and Talbot, 1978; Caddy and McGarvey, 1996; Punt and Smith, 2001).

For each of the seven scenarios above, we compared *SPR* and  $F/F_{MSY}$  in the terminal year as the performance metrics, which represent stock status and fishing intensity of the current state relative to a reference point. We then quantified bias and precision of *SPR* and  $F/F_{MSY}$  and derived values used to calculate  $F/F_{MSY}$ ,  $F$ , and  $F_{MSY}$  (results of  $F$  and  $F_{MSY}$  are in the Supplementary Tables S1 and S2, Figures S2 and S3). Relative errors between the estimated and true value were calculated as performance metrics for each derived output in order to quantify bias using the median relative error (*MRE*) and precision using the median absolute relative error (*MARE*):

$$MRE = \text{median}\left(\frac{x_{est} - x_{true}}{x_{true}}\right) \quad (1)$$

$$MARE = \text{abs}\left[\text{median}\left(\frac{x_{est} - x_{true}}{x_{true}}\right)\right], \quad (2)$$

where  $x_{est}$  is the estimated value (calculated from the respective estimation model) and  $x_{true}$  is the true value (calculated from the operating models). Values closer to zero represent the least biased (*MRE*) and most precise (*MARE*) results.

## Results

### Performance of the length-based methods

TB and LBSPR were <30% biased and imprecise across scenarios when estimating the reference points (Figures 4 and 5, Tables 4 and 5). Of the four scenarios, TB was the least biased and most precise in estimating *SPR* for the short-lived and stochastic scenarios with <10% bias and imprecision. When estimating  $F/F_{MSY}$ , TB performed the best in the base and overexploited scenarios, each being <10% biased and precise. The LBSPR estimate of *SPR* was the closest to the truth in the overexploited (<5% biased and imprecise) and autocorrelated scenarios, and was the least biased in estimating  $F/F_{MSY}$  in the underexploited, stochastic, and autocorrelated scenarios. LBSPR was relatively robust as natural mortality was fixed ( $M/K$  input; Table 5), which increases precision in the calculation of  $F/M$  and thus *SPR* and  $F/F_{MSY}$ . Rudd and Thorson (2018) found that LBSPR performed better when the stocks were in equilibrium and when the operating model matched LBSPR's assumptions; this result is supported given that the base IBM assumptions included constant recruitment (except in the stochastic and autocorrelated scenarios), and fishing. LBSPR was found to be accurate when the underlying selectivity is asymptotic (Hordyk et al., 2015b; Rudd and Thorson, 2018; Pons et al., 2019).

LIME was the least biased and most precise when estimating *SPR* for the base and longer-lived scenarios (Table 4) but usually performed the worst in estimating *SPR* for the over- and underexploited scenarios (Table 4) and  $F/F_{MSY}$  for short-lived and autocorrelated recruitment scenarios (Table 5). The relatively high bias and low precision for these estimates indicate that the complexity of this method requires longer time-series; a single year of monthly length-frequency data is insufficient for this method to untangle the effects of fishing mortality and recruitment on the length-frequency data. Lastly, LBRA only performed well in the short-lived scenario. Due to the method's constraint in using a "truncated model" of Ehrhardt and Ault (1992), meaning that  $L_{MAX}$  is the cut-off for the upper length class, some of the lengths above  $L_{MAX}$  were removed. Truncations lead to overestimation of  $Z$  (Then et al., 2015), which results in underestimation of *SPR* and overestimation of  $F/F_{MSY}$ . This was especially evident in the medium- and longer-lived scenarios (Figures 4 and 5). Although LBRA was more biased than the other methods, it gave low values of *SPR* and high values of  $F/F_{MSY}$ , meaning that it is more precautionary, a more desired feature in data-limited approaches if error is unavoidable.

### Across life history, exploitation level, and recruitment scenarios

Life histories have a clear impact on the performance of these four length-based assessment methods (Figures 4 and 5). Notably, each method had difficulties in assessing short-lived stocks, especially when estimating  $F$  and  $F_{MSY}$ , and consequently

**Table 3.** Summary of methods.

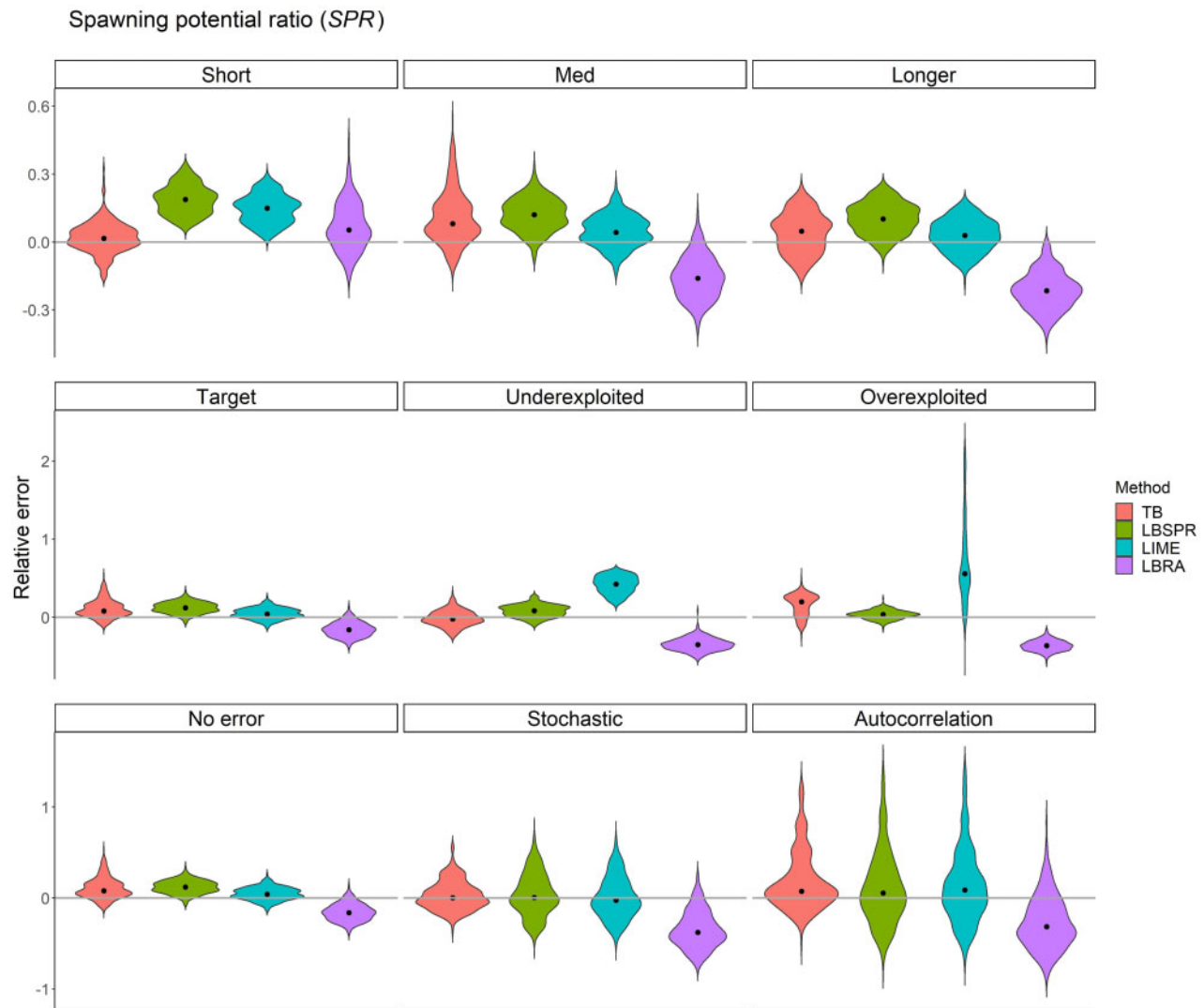
Method	Inputs	Assumptions	Outputs
LCCC (not part of the comparison)	<ul style="list-style-type: none"> <li>Length-frequency data (yearly catch vector)</li> <li>von Bertalanffy growth function (<math>L_\infty, K, t_0</math>)</li> <li>Natural mortality (<math>M</math>)</li> </ul>	<ul style="list-style-type: none"> <li>(i) Total mortality is constant for all length classes</li> <li>(ii) Selectivity follows logistic curve (width of curve calculated from <math>L_{50}^s</math> and <math>L_{95}^s</math>)</li> </ul>	<ul style="list-style-type: none"> <li><b>Length at 50% and 95% selectivity (<math>L_{50}^s</math> and <math>L_{95}^s</math>)</b></li> <li><b>Total mortality (<math>Z</math>) (used to calculate <math>F</math>)</b></li> </ul>
Length-based TB	<ul style="list-style-type: none"> <li>Length-frequency data</li> <li>von Bertalanffy growth function (<math>L_\infty, K, t_0</math>)</li> <li>Length-weight relationship (<math>\alpha</math> and <math>\beta</math>)</li> <li>F-at-length-array (fishing mortality for each length class; calculated based on selectivity)</li> <li>Natural mortality (<math>M</math>)</li> <li>Total mortality (<math>Z</math>)</li> <li>Length at 50% selectivity and maturity (<math>L_{50}^s</math> and <math>L_{50}^m</math>)</li> <li>Width of selectivity and maturity logistic curve</li> </ul>	<ul style="list-style-type: none"> <li>(i) Stock is in equilibrium</li> <li>(ii) Natural mortality is constant</li> <li>(iii) Selectivity and maturity follow logistic curve</li> </ul>	<ul style="list-style-type: none"> <li>Precautionary reference levels (<math>F_{0.1}, F_{0.5}, E_{0.5}</math>)</li> <li>Exploitation, yield, abundance and catch across vector of fishing mortalities</li> <li>Current exploitation, yield, abundance and catch</li> <li><b>Current <math>F</math></b></li> <li><b>SPR</b></li> <li><b><math>F/F_{MSY}</math> or <math>F/F_{MSY}</math></b></li> <li><b><math>SPR_{MSY}</math></b></li> </ul>
LBSPR	<ul style="list-style-type: none"> <li>Length-frequency data</li> <li>Asymptotic length (<math>L_\infty</math>)</li> <li>Coefficient of variation of <math>L_\infty</math> (<math>CV_{L_\infty}</math>)</li> <li><math>M/K</math> (calculated from <math>M</math> and <math>K</math>)</li> <li>Length-weight relationship (<math>\alpha</math> and <math>\beta</math>)</li> <li>Length at 50% and 95% selectivity (<math>L_{50}^s</math> and <math>L_{95}^s</math>)</li> <li>Length at 50% and 95% maturity (<math>L_{50}^m</math> and <math>L_{95}^m</math>)</li> </ul>	<ul style="list-style-type: none"> <li>(i) Stock is in equilibrium</li> <li>(ii) Natural mortality and growth rates are constant</li> <li>(iii) Selectivity and maturity follow a logistic curve</li> <li>(iv) Both sexes have the same growth curve and the sex ratio is equal</li> <li>(v) The lengths at each age are normally distributed around a mean length-at-age value.</li> </ul>	<ul style="list-style-type: none"> <li><math>F/M</math> ratio</li> <li>Length at 50% and 95% selectivity (<math>L_{50}^s</math> and <math>L_{95}^s</math>)</li> <li><b><math>F/M</math> ratio (used to calculate <math>F</math>)</b></li> <li><b>SPR</b></li> <li><b><math>F/F_{MSY}</math></b></li> <li><b><math>SPR_{MSY}</math></b></li> </ul>
LIME	<ul style="list-style-type: none"> <li>Length composition data</li> <li>von Bertalanffy growth function (<math>L_\infty, K, t_0</math>)</li> <li>Length-weight relationship (<math>\alpha</math> and <math>\beta</math>)</li> <li>Natural mortality (<math>M</math>)</li> <li>Length at 50% and 95% selectivity (<math>L_{50}^s</math> and <math>L_{95}^s</math>)</li> <li>Length at 50% (<math>L_{50}^m</math>)</li> </ul>	<ul style="list-style-type: none"> <li>(i) Natural mortality is constant</li> <li>(ii) Selectivity and maturity follow a logistic curve</li> </ul>	<ul style="list-style-type: none"> <li>(Length data only)</li> <li>Recruitment</li> <li>Spawning biomass</li> <li>Mean length</li> <li>Length at 50% and 95% selectivity (<math>L_{50}^s</math> and <math>L_{95}^s</math>)</li> <li><b>Current <math>F</math></b></li> <li><b>SPR</b></li> <li><b><math>F/F_{MSY}</math></b></li> <li><b><math>SPR_{MSY}</math></b></li> </ul>
LBRA	<ul style="list-style-type: none"> <li>Length composition data</li> <li>von Bertalanffy growth function (<math>L_\infty, K, t_0</math>)</li> <li>Coefficient of variation of length at age (<math>CV_{L_\infty}</math>)</li> <li>Length-weight relationship (<math>\alpha</math> and <math>\beta</math>)</li> <li>Natural mortality (<math>M</math>)</li> <li>Theoretical maximum age (<math>\hat{a}_i</math>)</li> </ul>	<ul style="list-style-type: none"> <li>(i) Average annual constant recruitment</li> <li>(ii) Selectivity and maturity follow a logistic curve</li> <li>(iii) The lengths at each age are normally distributed around the mean length</li> <li>(iv) The observed maximum age (<math>\hat{a}_i</math>) deviates are described by the exponential probability density function (used to calculate <math>M</math>)</li> </ul>	<ul style="list-style-type: none"> <li><math>B/B_{MSY}</math></li> <li><b>Total mortality (<math>Z</math>) [used to calculate fishing mortality (<math>F</math>)]</b></li> <li><b>SPR</b></li> <li><b><math>F/F_{MSY}</math></b></li> <li><b><math>SPR_{MSY}</math></b></li> </ul>

The data inputs, assumptions, and expected outputs are listed for each method including the LCCC. In the outputs, the estimates in bold are those used this study uses for the comparison.

$F/F_{MSY}$ . The increased bias and decreased precision of  $F/F_{MSY}$  stem from the decreased accuracy of the calculation of fishing mortality and  $F_{MSY}$ . All methods overestimated  $SPR$ ; LBSPR and LIME performed the worst and TB and LBRA performed the best across the life histories. For medium- and longer-lived species, TB, LBSPR, and LIME overestimated  $SPR$ , whereas LBRA on average underestimated  $SPR$ . In the medium-lived scenario, TB,

LBSPR, and LIME were negatively biased in calculating  $F/F_{MSY}$ , whereas LBRA was positively biased. For the longer-lived scenario, TB and LBSPR underestimated  $F/F_{MSY}$ , whereas LIME and LBRA overestimated  $F/F_{MSY}$ .

Although the performance of the methods under different exploitation scenarios varied among reference points, it is evident that stocks that are either under- or overexploited are more



**Figure 4.** Violin plots of relative error for  $SPR$  for 300 iterations per scenario with 200 individuals per month sampled in a single year across three life histories (short-, medium-, and longer-lived), three exploitation levels (target, under, and overexploited), and three recruitments (no error, stochastic error, and autocorrelation pattern and error). Four methods (length-based TB [in orange], LBSPR [in green], LIME [in blue], and LBRA [in violet]) were analysed. The grey horizontal line is the zero relative error line, and the black dot is the median relative error indicating bias. Each plot has a different  $y$ -axis with a smoother tail.

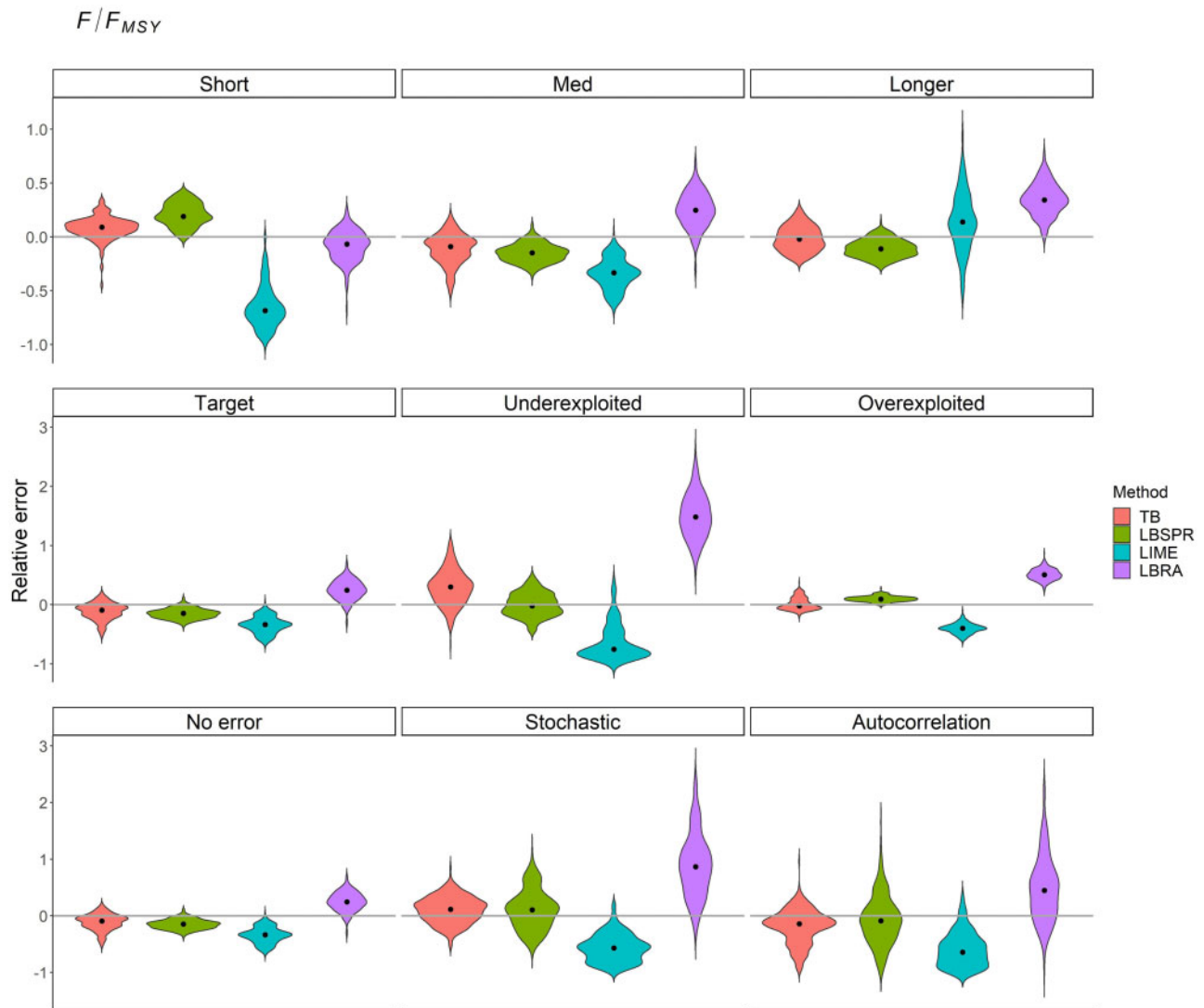
difficult to assess than those that are fished near the target exploitation level ( $SPR \approx 40\%$ ). When stocks are severely overexploited, the methods are less accurate in estimating  $SPR$ . When the stocks are severely underexploited, the methods present inconsistencies in estimating  $F/F_{MSY}$ . The error stems from the calculation of fishing mortality as seen in [Supplementary Figure S1](#) and [Table S1](#). For both exploitation levels, the inaccuracy of  $F/F_{MSY}$  is due to the large bias and imprecision in calculating the fishing mortality.

The impacts of increasing recruitment error were also evident in each of the methods, as precision decreased in the stochastic and autocorrelated scenarios. Although precision decreased, the bias in  $SPR$  was also lower in most of the methods because no recruitment error was implemented in the base model. In the stochastic scenario, each of the methods on average underestimated  $SPR$ , with three methods (TB, LBSPR, and LIME) also being  $<5\%$  biased, whereas in the autocorrelated scenario,

the three methods on average overestimated  $SPR$ , but remained  $<10\%$  biased. When estimating  $F/F_{MSY}$  in general, the bias generally decreased with the addition of recruitment error compared with when there was no error. The high bias in estimating  $F/F_{MSY}$  stems from  $F_{MSY}$ , and the low precision stems from fishing mortality.

## Discussion

This study is a first in simulation testing length-based models using an IBM framework, thus adding a level of independence in population dynamics not seen in other studies. In similar studies ([Ault et al., 1998](#); [Hordyk et al., 2015b](#); [Rudd and Thorson, 2018](#)), the operating models used are identical to the assessment models and assume that all dynamic processes are fully understood. An alternatively structured operating model can help avoid this problem and identify misleading assumptions that may be implicit in the design of an assessment model ([Cao et al., 2016](#)).



**Figure 5.** Violin plots of relative error for  $F/F_{MSY}$  for 300 iterations per scenario with 200 individuals per month sampled in a single year across three life histories (short-, medium-, and longer-lived), three exploitation levels (target, under-, and overexploited), and three recruitments (no error, stochastic error, and autocorrelation pattern and error). Four methods (length-based TB [in orange], LBSPR [in green], LIME [in blue], and LBRA [in violet]) were analysed. The grey horizontal line is the zero relative error line, and the black dot is the median relative error indicating bias. Each plot has a different y-axis with a smoother tail.

In addition, population-based methods (age or length) make assumptions regarding error in length-at-age to create length distribution within each age class (Cao *et al.*, 2016). Our operating model considers the Rosa Lee phenomenon, which is usually not seen in many assessment models, where slow-growing fish may have lower mortality rates because they become susceptible to fishing selection (i.e. reach  $L_{50}^s$ ) at a larger age than faster-growing individuals (Lee, 1912; Kraak *et al.*, 2019).

### Performance of methods

Stock assessment methods often perform poorly with short-lived species, as the annual time-step does not provide enough information about their dynamics (Thiaw *et al.*, 2011; Alemany *et al.*, 2015; Maunder and Piner, 2015). In addition, the biomass of short-lived species is more sensitive to environmental variability because of their fast growth rates and short generation times

(Winemiller, 2005; Dichmont *et al.*, 2006; Pinsky *et al.*, 2011). Hordyk *et al.* (2015b) also found that length-based methods tend to be biased for short-lived species, as these methods often rely on detecting the signal of fishing mortality on the upper tail of the length composition. Rudd and Thorson (2018) noted that with increasing length and age, the cohorts tend to be indistinct. In general, longer-lived species may have lower *SPR* levels, as there is a relationship between longevity and sensitivity of *SPR* to exploitation pressures (Nadon *et al.*, 2015). Thus, for longer-lived species, spawning biomass is represented by older individuals, and their numbers can be reduced even with low fishing rate. Medium-lived species have new recruits at an early age, which allows better detection of information about the population, and are not as vulnerable to fishing due to a low reproductive rate. Varying life histories have an impact on the quality of length-frequency data as they affect the ideal sample size, sparseness in small or large length classes, and effects of selectivity and fishing



**Table 4.** Bias (MRE) and precision (MARE) table of SPR from length-based TB, LBSPR, LIME, and LBRA performance across life histories, exploitation levels, and recruitment types.

Model	Life history			Exploitation level			Recruitment		
	Short	Medium <sup>a</sup>	Longer	Target <sup>a</sup>	Under-exploited	Over-exploited	No error <sup>a</sup>	Stochastic	Auto-correlation
Bias (MRE)									
TB	0.016	0.081	0.048	0.081	-0.023	0.197	0.081	-0.003	0.076
LBSPR	0.188	0.121	0.102	0.121	0.081	0.035	0.121	-0.007	0.056
LIME	0.149	0.042	0.029	0.042	0.424	0.560	0.042	-0.023	0.089
LBRA	0.054	-0.159	-0.215	-0.159	-0.350	-0.364	-0.159	-0.377	-0.315
Precision (MARE)									
TB	0.043	0.084	0.073	0.084	0.066	0.197	0.084	0.099	0.146
LBSPR	0.187	0.121	0.102	0.121	0.082	0.044	0.121	0.150	0.218
LIME	0.149	0.062	0.057	0.062	0.424	0.560	0.062	0.163	0.226
LBRA	0.074	0.159	0.215	0.159	0.350	0.364	0.159	0.377	0.337

<sup>a</sup>These are components of the base model, and thus are of a single scenario.

**Table 5.** Bias (MRE) and precision (MARE) table of  $F/F_{MSY}$  from length-based TB, LBSPR, LIME, and LBRA performance across life histories, exploitation levels, and recruitment types.

Model	Life history			Exploitation level			Recruitment		
	Short	Medium <sup>a</sup>	Longer	Target <sup>a</sup>	Under-exploited	Over-exploited	No error <sup>a</sup>	Stochastic	Auto-correlation
Bias (MRE)									
TB	0.091	-0.091	-0.022	-0.091	0.298	-0.018	-0.091	0.112	-0.142
LBSPR	0.189	-0.148	-0.111	-0.148	-0.022	0.096	-0.148	0.105	-0.089
LIME	-0.684	-0.334	0.138	-0.334	-0.752	-0.400	-0.334	-0.568	-0.644
LBRA	-0.066	0.247	0.343	0.247	1.484	0.505	0.247	0.867	0.448
Precision (MARE)									
TB	0.096	0.114	0.089	0.114	0.306	0.075	0.114	0.197	0.190
LBSPR	0.189	0.148	0.111	0.148	0.145	0.096	0.148	0.223	0.282
LIME	0.684	0.334	0.188	0.334	0.752	0.400	0.334	0.568	0.644
LBRA	0.098	0.253	0.343	0.253	1.484	0.505	0.253	0.867	0.469

Lighter colours indicate better models: the lightest red colour indicates bias/precision <5%, and the darkest red colour indicates bias/precision >30%.

<sup>a</sup>These are components of the base model, and thus are of a single scenario.

pressure. Although this study does not address common issues associated with the quality and bias of length-frequency data, we can, however, conclude that medium-lived species seem the easiest to assess as their cohorts can be tracked, and it is easier to sufficiently sample length classes.

We also show that fishing mortality in overexploited and underexploited stocks is harder to assess with the four tested methods. In the majority of the scenarios, fishing mortality was underestimated, which highlights the challenge of accurately estimating mortality rates and emphasizes the need for conservative interpretation of assessment outcomes, particularly in assessments that can only estimate stock status based on mortality rates (e.g. exploitation). In a future study, the influence of historical

fishing patterns (e.g. a history of increasing and decreasing fishing mortality) would be of interest to investigate as fishing mortality is usually not constant, and a consistently high fishing mortality results in a different length distribution than that seen in a recently developed fishery.

Many stock assessment models, including length-based methods, assume equilibrium conditions. However, this assumption is typically incorrect as recruitment variation changes the age structure of a population with time (Haddon, 2001). Recruitment patterns can vary greatly among stocks (e.g. pulsed, autocorrelated), seasonal variations and modes (uni- vs. bimodal; Isaac, 1990) and should not be overlooked. Despite the underlying uncertainties about recruitment error and type and whether a fishery is in

non-equilibrium conditions, one could still apply TB or LBSPR as they were less biased with a single year of data. Although this study lightly addressed the effects of including recruitment variability, and many management strategy evaluation studies investigate different levels of recruitment error, further studies should investigate how this may affect data-limited fisheries assessments.

### Caveats of length data

Typically, length data representing a single year of sampling are frequently the case in tropical, data-limited areas. However, some methods have limited abilities with only a single year of data as there is no information to tease apart recruitment from fishing mortality leading to high uncertainty. Time-series data are not always guaranteed to give better estimates (Carruthers *et al.*, 2014; Rudd and Thorson, 2018; Dowling *et al.*, 2019); rather, the validity of the assumptions of the method and the attributes of the stock life history are more important to assessment outcomes (Carruthers *et al.*, 2014). Although stock assessments usually use multiyear data, the quality of the data and the performance of the methods are often more important in determining the reliability of the assessment. In addition, the ideal sample size and sampling period are dependent on life history and should be considered in the assessment.

Length-frequency data obtained from small-scale, data-limited fisheries are often strongly biased due to mixed gear selectivity not being properly accounted for or because not all gear types are sampled. The size composition of the catch usually reflects a mix of sizes due to the use of a variety of fishing gears. This issue was addressed for tropical fisheries by Wolff *et al.* (2015), who found that different gear selectivities impact YPR and spawning biomass. The data generation of our simulation study assumed asymptotic selectivity and  $L_{50}^s = 0.25 L_{\infty}$ . Ideally, the selection characteristics of the gear(s) should be known prior to any length-frequency analysis, and (if possible) catch length-frequency data should first be reconstructed based on the selectivity features of the gear. A study by Pons *et al.* (2019) investigated the influence of different gear selectivity from multiple fleets on length composition data of scombrids in the Atlantic Ocean and found that accounting for multiple selectivity curves reduces bias in estimating *SPR*. Further consideration and investigation of gear selectivity influences on length data are warranted.

### Guidance to practitioners and conclusions

TB and LBSPR are methods based on YPR or spawning per recruit, meaning that they do not assume any stock–recruitment relationship. These are important assumptions of the two models, and when violated, their performances will be biased. A common practice in using TB is to perform the length-based Jones' cohort analysis beforehand (Mildenberger *et al.*, 2017), which relies on catch in numbers/biomass at age and can estimate fishing mortality per length class. However, it should be noted that this approach adds error from the cohort analysis to TB. We, therefore, recommend calculating fishing mortality per length class using the LCCC with the estimation of selectivity pattern.

When there is conflicting information and results among methods, we recommend using the information provided in this study to advise which type of model is less biased for a given combination of life history and other factors. This study guides which type of model, out of the four tested, can be considered most reliable given a stock's life history, recruitment, and

exploitation. If those properties are unknown (and they most often will be), TB and LBSPR should be run in parallel. Minor differences in outcomes would then indicate uncertainties due to model structure, whereas major discrepancies point to model assumptions potentially being overlooked. Although we highlight which methods performed well in a given scenario, we also encourage using a combination of length-based methods to compare their performances and define a range of possible stock estimates. In the absence or uncertainty of such knowledge on life history, recruitment, and exploitation, LBSPR and TB can be expected to perform more consistently than LIME and LBRA in the rapid assessment of limited data of a single year.

These four length-based methods can be used for rapid assessments in data-limited fisheries to provide a cost-effective starting point for management. As the assumptions and sensitivities of each method were analysed in this study, scientists and managers can use this information to quickly assess data-limited stocks for an indication of stock status, thus providing guidance on which methods to employ given a situation. For the development of new data-limited assessment approaches, understanding how different life histories, exploitation levels, and recruitment errors affect existing methods is essential as it highlights the weaknesses of the current methods.

### Supplementary data

[Supplementary material](#) is available at the *ICESJMS* online version of the manuscript.

### Funding

TKM was funded in part by the EMFF project “ManDaLiS—Improving the management basis for Danish data-limited stocks” (33113-B-16-085), which is funded by the European Maritime and Fisheries Fund and the Danish Fisheries Agency. TAB was funded in part by the Richard C. and Lois M. Worthington Endowed Professor in Fisheries Management. MBR was funded by the Joint Institute for the Study of the Atmosphere and Ocean, NOAA Cooperative Agreement No. NA150AR4320063 (2017–2019), Contribution No. 2018-0161.

### Acknowledgements

We thank James Thorson and Adrian Hordyk for their technical support in applying FishLife and LBSPR tools, respectively. We acknowledge Andi Stephens, the editor, and three anonymous reviewers for comments that significantly improved the paper.

### References

- Abobi, S. M., Mildenberger, T. K., Kolding, J., and Wolff, M. 2019. Assessing the exploitation status of main fisheries resources in Ghana's reservoirs based on reconstructed catches and a length-based bootstrapping stock assessment method. *Lake and Reservoir Management*, 1: 1–20.
- Alemay, J., Foucher, E., Vigneau, J., and Robin, J. P. 2015. Stock assessment models for short-lived species in data-limited situations: case study of the English Channel stock of cuttlefish (*Sepia officinalis*). *Tools and Strategies for Assessment and Management of Data-Limited Fish Stocks*, 30th Lowell Wakefield Fisheries Symposium. May 12–15, 2015, Anchorage, Alaska.
- Ault, J., Bohnsack, J., and Meester, G. 1998. A retrospective (1979–1996) multispecies assessment of coral reef fish stocks in the Florida Keys. *Fishery Bulletin*, 96: 395–414.
- Ault, J. S., Smith, S. G., Bohnsack, J. A., Luo, J., Stevens, M. H., DiNardo, G. T., Johnson, M. W., *et al.* 2019. Length-based risk

- analysis for assessing sustainability of data-limited tropical reef fisheries. *ICES Journal of Marine Science*, 76: 165–180.
- 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. *Environmental Conservation*, 35: 221–231.
- Bystrom, A. 2016. 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.
- Caddy, J. F., and McGarvey, R. 1996. Targets or limits for management of fisheries? *North American Journal of Fisheries Management*, 16: 479–487.
- Cadrin, S. X., and Dickey-Collas, M. 2015. Stock assessment methods for sustainable fisheries. *ICES Journal of Marine Science*, 72: 1–6.
- Cao, J., Guan, W., Truesdell, S., Chen, Y., and Tian, S. 2016. An individual-based probabilistic model for simulating fisheries population dynamics. *Aquaculture and Fisheries*, 1: 34–40.
- 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. *Fisheries Research*, 153: 48–68.
- Chrysafi, A., and Kuparinen, A. 2016. Assessing abundance of populations with limited data: lessons learned from data-poor fisheries stock assessment. *Environmental Reviews*, 24: 25–38.
- Clark, W. G. 2002. F35% revisited ten years later. *North American Journal of Fisheries Management*, 22: 251–257.
- Deroba, J. J., and Schueller, A. M. 2013. Performance of stock assessments with misspecified age- and time-varying natural mortality. *Fisheries Research*, 146: 27–40.
- Dichmont, C. M., (Roy) Deng, A., Punt, A. E., Venables, W., and Haddon, M. 2006. Management strategies for short-lived species: the case of Australia's Northern prawn fishery. *Fisheries Research*, 82: 204–220.
- Dowling, N. A., Dichmont, C. M., Haddon, M., Smith, D. C., Smith, A. D. M., and Sainsbury, K. 2015. Guidelines for developing formal harvest strategies for data-poor species and fisheries. *Fisheries Research*, 171: 130–140.
- Dowling, N. A., Smith, A. D. M., Smith, D. C., Parma, A. M., Dichmont, C. M., Sainsbury, K., Wilson, J. R., *et al.* 2019. Generic solutions for data-limited fishery assessments are not so simple. *Fish and Fisheries*, 20: 174–188.
- Ehrhardt, N. M., and Ault, J. S. 1992. Analysis of two length-based mortality models applied to bounded catch length frequencies. *Transactions of the American Fisheries Society*, 121: 115–122.
- Haddon, M. 2001. *Modelling and Quantitative Methods in Fisheries*. Chapman & Hall/CRC, Boca Raton, FL. 406 pp.
- Heemstra, P. C., and Randall, J. E. 1993. *FAO Species Catalogue, Volume 16, Groupers of the world (Family Serranidae, Subfamily Epinephelinae)*. An annotated and illustrated catalogue of the grouper, rockcod, hind, coral grouper and lyretail species known to date. *FAO Fisheries Synopsis*, No. 125, Vol. 16. Rome, FAO. 382 pp.
- Herrón, P., Mildenerberger, T. K., Díaz, J. M., and Wolff, M. 2018. Assessment of the stock status of small-scale and multi-gear fisheries resources in the tropical Eastern Pacific region. *Regional Studies in Marine Science*, 24: 311–323.
- 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: e36022.
- Holt, S. J., and Talbot, L. M. 1978. New principles for the conservation of wild living resources. *Wildlife Monographs*, 59: 3–33.
- Hordyk, A., Ono, K., Sainsbury, K., Loneragan, N., and Prince, J. 2015a. Some explorations of the life history ratios to describe length composition, spawning-per-recruit, and the spawning potential ratio. *ICES Journal of Marine Science*, 72: 204–216.
- Hordyk, A., Ono, K., Valencia, S., Loneragan, N., and Prince, J. 2015b. A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. *ICES Journal of Marine Science*, 72: 217–231.
- Hordyk, A. R., Loneragan, N. R., and Prince, J. D. 2015c. An evaluation of an iterative harvest strategy for data-poor fisheries using the length-based spawning potential ratio assessment methodology. *Fisheries Research*, 171: 20–32.
- Isaac, V. J. 1990. *The Accuracy of Some Length-Based Methods for Fish Population Studies (ICLARM Contribution)*. International Center for Living Aquatic Resources Management, Manila, Philippines. 81 pp.
- Kraak, S. B. M., Haase, S., Minto, C., and Santos, J. 2019. The Rosa Lee phenomenon and its consequences for fisheries advice on changes in fishing mortality or gear selectivity. *ICES Journal of Marine Science*, doi.org/10.1093/icesjms/fsz107.
- Larkin, P. A. 1977. An epitaph for the concept of maximum sustained yield. *Transactions of the American Fisheries Society*, 106: 1–11.
- Lee, R. M. 1912. An investigation into the methods by growth determination in fishes by means of scales. *Publications de Circonstance Conseil Permanent International Pour L'Exploration de la Mer*, 63: 34 pp.
- Mace, P. M., and Sissenwine, M. P. 1993. How much spawning per recruit is enough? *Canadian Special Publication of Fisheries and Aquatic Sciences*, 120: 101–118.
- Maunder, M. N., and Piner, K. R. 2015. Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science*, 72: 7–18.
- Mildenerberger, T. K., Taylor, M. H., and Wolff, M. 2017. TropFishR: an R package for fisheries analysis with length-frequency data. *Methods in Ecology and Evolution*, 8: 1520–1527.
- 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: e0133960.
- Nelson, G. A. 2017. *Fishmethods: Fishery Science Methods and Models in R*. <https://CRAN.R-project.org/package=fishmethods> (last accessed 1 March 2019).
- Pilling, G. M., Apostolaki, P., Failler, P., Floros, C., Large, P. A., Morales-Nin, B., Reglero, P., *et al.* 2008. Assessment and management of data-poor fisheries. *In Advances in Fisheries Science: 50 Years on from Beverton and Holt*, pp. 280–305. Ed. by A. Payne, J. Cotter, and T. Potter. Blackwell Publishing Ltd., Oxford, UK. 568 pp.
- Pinsky, M. L., Jensen, O. P., Ricard, D., and Palumbi, S. R. 2011. Unexpected patterns of fisheries collapse in the world's oceans. *Proceedings of the National Academy of Sciences*, 108: 8317–8322.
- Pons, M., Kell, L. T., Rudd, M. B., Cope, J. M., and Lucena Frédo, F. 2019. Performance of length-based data-limited methods in a multifleet context: application to small tunas, mackerels, and bonitos in the Atlantic Ocean. *ICES Journal of Marine Science*, 76: 960.
- Punt, A. E. 2003. The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. *Fisheries Research*, 65: 391–409.
- Punt, A. E., and Smith, A. D. M. 2001. The gospel of maximum sustainable yield in fisheries management: birth, crucifixion and reincarnation. *In Conservation of Exploited Species*, pp. 41–65. Ed. by J. D. Reynolds, G. M. Mace, K. H. Redford, and J. G. Robinson. Cambridge University Press, New York, NY. 546 pp.
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (last accessed 1 March 2019).

- Reynolds, J. D., Mace, G. M., Redford, K. H., and Robinson, J. G. (Eds). 2001. Conservation of Exploited Species. Cambridge University Press, New York, NY. 546 pp.
- Rudd, M. B., and Thorson, J. T. 2018. Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, 75: 1019–1035.
- Sparre, P., and Venema, S. C. 1998. Introduction to tropical fish stock assessment. Part. 1: Manual. FAO Fisheries Technical Paper 306/1, Rev. 2. FAO, Rome. 407 pp.
- Taylor, M. H. 2017. fishdynr: fisheries science related population dynamics models. <https://github.com/marchtaylor/fishdynr> (last accessed 1 March 2019).
- Taylor, M. H., and Mildenberger, T. K. 2017. Extending electronic length frequency analysis in R. *Fisheries Management and Ecology*, 24: 330–338.
- Tesfaye, G., Wolff, M., and Taylor, M. 2016. Gear selectivity of fishery target resources in Lake Koka, Ethiopia: evaluation and management implications. *Hydrobiologia*, 765: 277–295.
- Then, A. Y., Hoenig, J. M., Hall, N. G., and Hewitt, D. A. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. *ICES Journal of Marine Science*, 72: 82–92.
- Thiaw, M., Gascuel, D., Thiao, D., Thiaw, O. T., and Jouffre, D. 2011. Analysing environmental and fishing effects on a short-lived species stock: the dynamics of the octopus *Octopus vulgaris* population in Senegalese waters. *African Journal of Marine Science*, 33: 209–222.
- Thompson, W. F., and Bell, F. H. 1934. Biological statistics of the Pacific halibut fishery. (2) Effect of changes in intensity upon total yield and yield per unit of gear. International Fisheries Commission, Report No. 8. 49 pp.
- Thorson, J. T. in press. Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. *Fish and Fisheries*.
- Tuda, P. M. 2018. Assessing the state and impacts of the artisanal reef fisheries and their management implications in Kenyan south coast. PhD thesis, Universität Bremen: Biologie/Chemie, Bremen, Germany. 198 pp.
- Winemiller, K. O. 2005. Life history strategies, population regulation, and implications for fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences*, 62: 872–885.
- Wolff, M., Taylor, M. H., and Tesfaye, G. 2015. Implications of using small meshed gillnets for the sustainability of fish populations: a theoretical exploration based on three case studies. *Fisheries Management and Ecology*, 22: 379–387.

Handling editor: Emory Anderson