Accounting for Variability and Biases in Data-limited Fisheries Stock Assessment

Merrill B. Rudd

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Trevor A. Branch, Co-chair

Ray Hilborn, Co-chair

James T. Thorson

Program Authorized to Offer Degree:

School of Aquatic and Fishery Sciences

University of Washington

Abstract

Accounting for Variability and Biases in Data-limited Fisheries Stock Assessment

Merrill B. Rudd

Joint Chairs of the Supervisory Committee:
Associate Professor Trevor Branch
Professor Ray Hilborn
School of Aquatic and Fishery Sciences

Many regions of the world have very few stocks assessed, often due to limited data quality or quantity or lack of trained scientists to apply and interpret stock assessments. These same areas with fewer assessments perform worse across fishery management attributes, including research capacity, management, enforcement, and socioeconomics. Some studies have used the limited data available to approximate the status of these "unassessed" stocks and find them to be declining compared to assessed stocks. Global assessments of "unassessed" stocks are informed by many attributes of the stock, one of which is the trend in reported catch to the U.N. Food and Agriculture Organization. These assessments assume that the catch data are accurate, at least in trend if not in magnitude. However, quantifying catch is a major challenge due to commonly misreported sectors such as discards, small-scale, recreational, and illegal, unreported, unregulated fisheries. One approach is to improve catch data by improving monitoring programs or conducting catch reconstructions. Another approach is to collect alternative data types to conduct stock assessments. Stocks that are not assessed because they have the lowest information or research capacity appear more likely to be unsustainable, and further research and management action is required to improve the status of the data- and research-capacity-limited fisheries.

This dissertation addresses issues of data quality in the monitoring process and data limitation in the assessment process. Chapter 1 addresses an issue of data quality in asking, “does unreported catch lead to overfishing?” We used simulation analysis to demonstrate that if catch is misreported at a constant rate, surplus production models can still estimate unbiased stock status and would avoid overfishing with effective management. If catch misreporting is increasing over time, stock assessments would recommend harvest limits that would lead to under-exploitation, while decreasing reporting rates lead to over-exploitation. This question is relevant to fisheries around the world dealing with uncertainties in catch reporting in their stock assessment, as well as current debates over the use of catch reconstructions.

In Chapter 2 stock assessments are examined for fisheries that only have length composition data and biological information, since it is often easier to collect length measurements than to quantify total catch. Length measurements from samples of the catch, referred to as length composition data, can be contrasted with expected length composition in an unfished state to reveal information about fishing mortality, recruitment, and selectivity. Most data-poor length-based stock assessment methods assume the population is in equilibrium, i.e. that fishing mortality and recruitment have not changed within one generation of the species. In this chapter we present a Length-based Integrated Mixed Effects (LIME) model that relaxes this equilibrium assumption and directly estimates variable fishing mortality and recruitment using the same data inputs as other length-based methods. Using simulation testing we demonstrated LIME performs best for life history types with a maximum age of less than 20 years, and is unbiased across a range of recruitment and fishing mortality patterns, provided individual growth parameters are known. LIME also has the capability of including multiple years of length data,

abundance indices, and catch time series when available. LIME is a flexible new tool for stock assessments of fish usually caught as bycatch and other small-scale fisheries.

Chapter 3 applies LIME and the equilibrium-based Length-Based Spawning Potential Ratio (LB-SPR) methods to assess a medium-lived Costa Rican spotted rose snapper, *Lutjanus guttatus*, and short-lived Kenyan rabbitfish, *Siganus sutor*. LIME estimated the Costa Rican snapper fishery to be overfished in the most recent year of data after a period of full exploitation, whereas LB-SPR estimated more variability in stock status throughout the time series but the fishery was above the target reference point in the most recent year of data. LIME estimated the rabbitfish fishery to have undergone a period of overexploitation in the late 1990s and early 2000s, but fishing mortality has continually decreased over time resulted in a recovered stock in the most recent year of data. Alternatively, LB-SPR estimated the stock slightly less than the SPR target. Chapter 3 presented the first empirical LIME assessment and comparison with a commonly-used alternative method, and presented guidelines for future LIME applications.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am very thankful for a large network of people who have encouraged and supported me through my PhD. Foremost thanks to the IGERT Program on Ocean Change and Terrie Klinger for the eight quarters of funding that opened my door to SAFS, leading me down an unbelievable road of incredible people, research, teaching, traveling, and general happiness in Seattle that I could never have predicted. Beyond the funding, IGERT served as an enriching, supportive community that strongly shaped my PhD. I thank the School of Aquatic and Fishery Sciences for supporting me via eight additional quarters of fellowship funding. This allowed my research to continue naturally towards contributions I personally found most interesting and useful, as opposed to dictation from other grants.

I thank my supervisors, Trevor Branch and Ray Hilborn, and committee members Jim Thorson, Eddie Allison, and Jennifer Ruesink. Your feedback and brainstorming connected me with other subjects within fisheries and statistics so that my work could benefit from a greater context and build on new tools to make my mark on the field. Your influence in expanding my network has been equally important, providing me with such a wide range of collaborators and opportunities throughout my PhD to make it a unique and enriching experience that I hope to build on as I enter my career.

A modeler’s impact is greatly reduced without collaboration with field scientists who work to collect data. I acknowledge the Wildlife Conservation Society and Tim McClanahan for access to the Kenyan coral reef fishery landings dataset, with special thanks to Caroline Abunge for taking me snorkeling around a community-managed marine park during my visit to

Mombasa. I thank Conservation International and PRETOMA, specifically Ana Gloria Guzman and Andy Bystrom, for the Costa Rican spotted rose snapper data, as well as additional funding to conduct the stock assessment and helping me better understand the nuances of the dataset and fishery. I am very grateful for the opportunity to work with the NCEAS SNAPP data-limited fisheries working group, particularly Jono Wilson, Natalie Dowling, Jason Cope, Adrian Hordyk, Jeremy Prince, and George Maina. This group put me on a fast track for learning about data-limited fisheries stock assessment and management, a breadth of knowledge I am so thankful to have been able to experience at this point in my young career.

I am grateful for the SAFS community, of which I am so honored to have been a part for the past four years. Gracias Maite and Ricky, for the countless hours drinking mate or wine, teaching courses in fisheries or how to be Latina. I thank the Branch lab and the rest of the Hilborn lab, past and present, for seeing me through my rollercoasters. I am thankful for my IGERT clamily, for being a unit of support and adventure while learning the trials and tribulations of interdisciplinary science together.

I thank my dad for always optimistically reminding me that I've "always figured out how to make it happen before!" when I wasn't sure if I could and my mom for being an incredible role model of balancing family and a career, while also being transparent that she made it work due to her decisions and sacrifice, not by random chance. I thank my brother Nathan for being my moral compass and true mad scientist, my friends far and wide for reminding me not to be a weenie and always accepting the great bucket of love I have for you all, and Ripley for always pushing me forward and encouraging me to take a great leap.

DEDICATION

To my family.

INTRODUCTION

Stock assessments are meant to provide scientific, quantitative evaluations to inform fisheries management (Hilborn and Walters 1992). However, due to high data requirements and training in stock assessment, only about 50% of global catch have assessments (R. Hilborn, RAM Legacy Stock Assessment Database, *personal communication*). Many of these data-limited fisheries are small-scale, located in developing nations, and support communities on many levels, providing income (Cunningham et al. 2009), diversifying livelihoods (Allison and Ellis 2001), supplying food security (McClanahan et al. 2013) and nutrition (Hall et al. 2013), and supporting culture and well-being (Armitage et al. 2012, Kittinger et al. 2012). Generally, data-limited fisheries in data-poor regions are estimated to be in worse condition than stocks in data-rich regions with management based on full stock assessments (Costello et al. 2012, Hilborn and Ovando 2014). While data-limited fisheries have been sustainable for generations, changing oceans, markets, and increasing human population densities are threatening the sustainability of these fisheries and the value they represent (Andrew et al. 2007, Cinner et al. 2013a, Ruckelshaus et al. 2013). There are many approaches in the toolbox of data-limited assessment methods, each with varying objectives: to estimate stock status relative to biological reference points, estimate fishing pressure, explore population processes, and approximate uncertainty. Each approach comes with major assumptions impacting their ability to provide scientific advice to inform management.

Data-limited stock assessment is a large umbrella with several underlying families of methods (Chrysafi and Kuparinen 2015, Dowling et al. 2016). These methods can be broadly categorized as i) risk or vulnerability assessment, ii) indicator-based, iii) life history-based, iv) catch-only, v) length-based, and vi) model-based. These groupings are based on the type of inputs required, and may vary within family in whether they are static versus dynamic, baseline versus non-baseline, and whether model outputs apply to effort or catch guidelines.

On the very data-limited end of the spectrum are risk and vulnerability assessment, indicator-based methods, and life history-based methods. Risk and vulnerability assessments, such as productivity-susceptibility analysis (Patrick et al. 2009), are often used to prioritize species management, providing a general indication of which species are at greatest risk to overfishing, climate change, or other threats based on their life history and other qualitative information (Fujita et al. 2014, Cope et al. 2015, Zhou et al. 2016). Indicator-based methods are based on proxy reference points, using one or multiple indicators (e.g. mean length or length composition compared to length at maturity, changes in species composition, changes in catch per unit effort, etc.) to determine when management action is required (Die and Caddy 1997, Froese 2004, Shin et al. 2005, Ault et al. 2014). Life history-based approaches use empirical relationships or simple population models to estimate a snapshot of stock status, including per recruit analysis and matrix models estimating population growth rate at different levels of fishing mortality (Gabriel et al. 1989, Goodyear 1993, Morris et al. 2010).

In the intermediate realm of data-limited assessments are catch- and length-based methods, placed in this category because they typically assume the biological information from the very data-limited methods are also available. There are a wide variety of catch-only methods, which require a time series of total catch data, usually since initial exploitation, and information on population growth rate and scale to estimate maximum sustainable yield-based reference points (MacCall 2009, Berkson et al. 2011, Dick and MacCall 2011, Martell and Froese 2013, Rosenberg et al. 2017). Length-based methods are those that rely on mean length or length composition data, using lengths as a proxy for ages. This family of methods makes assumptions about life history attributes, either growth, natural mortality, and maturity explicitly or life history invariants, to estimate stock status (Thorson and Cope 2014, Hordyk et al. 2015, Then et al. 2015a).

At the upper end of the range of data-limited methods are model-based approaches, such as surplus production models and some simple types of statistical catch-at-age models, which allow for accounting of population dynamic processes (Hilborn 2001, Cope 2013). Population models are often included in specific approaches under the other data-limited method families (Dick and MacCall 2011, Nadon et al. 2015). While they require more assumptions and more data, they provide greater ability to explore the population processes upon which the data arose. Although surplus production and statistical-catch-at-age models are not data-limited to the same degree as some of the previous approaches discussed here, their performance will decline if data quality is low, which can place them at the upper end of the range of data-limited assessments.

This thesis focuses on questions related to data-quality issues in model-based methods and data-limitation issues in length-based methods. Chapter 1 asks the question: “Does unreported catch lead to overfishing?” Using simulation analysis, we quantified the direct impact of different scenarios of misreported catch on stock assessment results using surplus production models and proceeding management advice. While a surplus production model is in the upper range of data-limitation, many assessments worldwide suffer from catch misreporting, an issue of data quality. Using a surplus production model, which includes an abundance index, as opposed to a catch-only method, allowed for direct accounting of biases associated with catch misreporting (Rudd and Branch 2017).

The latter two chapters focus on length-based methods, moving away from assuming catch data are available or reliable. Chapter 2 presents the Length-based Integrated Mixed Effects (LIME) method, an extension of length-based assessment methods to directly account for variable fishing mortality and recruitment (Rudd and Thorson, in review). LIME, like other equilibrium-based length-based methods, requires a minimum of one year of length composition data with assumptions about growth, natural mortality, and maturity to estimate the spawning potential ratio (SPR) biological reference point. The major difference between LIME and other length-based methods is its treatment of recruitment as a random effect, assuming annual recruitment arises from a distribution with a mean and standard deviation. LIME estimates the standard deviation, selectivity, and annual fishing mortality, as fixed effects. LIME differs from other integrated methods, such as Stock Synthesis, in that it can estimate stock status without catch data. In Chapter 2 we used simulation analysis to estimate bias and precision in LIME across various scenarios of fishing mortality and recruitment variability, life history types, data availability, and sampling frequency, and compared LIME with an equilibrium-based method currently used in practice (Hordyk et al. 2015). Chapter 3 applies LIME to assess a medium-lived (13 years) spotted rose snapper (*Lutjanus guttatus*) stock from Costa Rica and a short-lived (4 years) rabbitfish (*Siganus sutor*) stock from Kenya. We developed guidelines for applying LIME, compared results to an equilibrium-based alternative, and provided stock assessments for the species to guide management in their respective locales.

The main value of this body of work is the challenge to strong assumptions about data and system processes in data-limited assessments, particularly the development of a novel assessment method to match the data availability of many fisheries worldwide that better accounts for variability in the natural and fisheries systems. LIME presents an alternate approach to estimate stock status for data-limited fisheries, relaxing the strong equilibrium assumption.

Chapter 1. DOES UNREPORTED CATCH LEAD TO OVERFISHING?

Abstract

Catches are commonly misreported in many fisheries worldwide, resulting in inaccurate data that hinder our ability to assess population status and manage fisheries sustainably. Under-reported catch is generally perceived to lead to overfishing, and hence, catch reconstructions are increasingly used to account for sectors unreliably reported, including illegal harvest, recreational and subsistence fisheries, and discards. However, improved monitoring and/or catch reconstructions only aid in the first step of a fisheries management plan: collecting data to make inferences on stock status. Misreported catch impacts estimates of population parameters, which in turn influences management decisions, but the pattern and degree of these impacts are not necessarily intuitive. We conducted a simulation study to test the effect of different patterns of catch misreporting on estimated fishery status and recommended catches. If, for example, 50% of all fishery catches are consistently unreported, estimates of population size and sustainable yield will be 50% lower, but estimates of current exploitation rate and fishery status will be unbiased. As a result, constant under- or over-reporting of catches results in recommended catches that are sustainable. However, when there are trends in catch reporting over time, the estimates of important parameters are inaccurate, generally leading to underutilization when reporting rates improve, and overfishing when reporting rates degrade. Thus, while quantifying total catch is necessary for understanding the impact of fisheries on businesses, communities and ecosystems, detecting trends in reporting rates is more important for estimating fishery status and setting sustainable catches into the future.

1.1 INTRODUCTION

Scientific predictions to inform fisheries management are usually based on stock assessments that rely on catches, survey time series, and biological information (Hilborn and Walters 1992). Fishery-independent surveys are the most informative for changes in population size over time (Francis 2011), but catches are also required to assess fishing pressure in relation to population size (Branch et al. 2011). The credibility of these fisheries stock assessment models hinges on the quality of data inputs and viability of model assumptions.

Catch data are the most common source of information provided by a fishery, but not always the easiest to obtain. Catches alone cannot be used to assess the status of a fishery, as they are a product of complex human behaviors and not a direct feedback from the ecosystem (Branch et al. 2011). In cases where catch data are unreliable or non-existent (as is the case for protected species), it is possible to assess the status of fish populations without catch data, for example using survey data only (Cook 1995), length composition of a sample of the catch (Cope and Punt 2009), or mark-recapture studies (Coggins et al. 2006). However, the amount harvested is useful for estimating the productivity of the population, its ability to withstand increased fishing pressure (Hilborn and Sibert 1988), and its importance in providing food, employment, and other resources to society (McClanahan et al. 2013). When catch is misreported, estimates of total productivity are altered, hindering the ability of managers to balance conservation and food security goals. The most common type of misreported catch is under-reporting, due to the decentralized nature of many fishing sectors, and poor monitoring and enforcement in much of the world (Pitcher et al. 2002; Agnew et al. 2009; Pauly and Zeller 2016). Illegal harvest, discards, subsistence, and other non-commercial fishing sectors are often unreported to management bodies (Pitcher et al. 2002), leading to overall under-reported catch for fish populations, which averages 53% globally (Pauly and Zeller 2016). Since 1950, the United Nations Food and Agriculture Organization (FAO) has compiled a fisheries landings database from their member countries (Food and Agriculture Organization 2014). Due to its ease of access and wide geographic scale, the FAO database has been instrumental in global fisheries research, for example in testing data-limited assessment methods (Thorson et al. 2012), exploring the status of unassessed fish populations (Costello et al. 2012), and examining the impacts of biodiversity loss on ecosystem services (Worm et al. 2006). However, the FAO database has been criticized since it often does not include commonly unreported fishery sectors (Pauly and Froese 2012). This is largely due to operational problems with many monitoring programs, including a lack of quality control, supervision for data collection, and insufficient resources (Weyl et al. 1999). Due to these logistical issues, the majority of funding for fisheries management in these regions is spent on catch monitoring, without adequate support for assessment of the monitoring data, management of the resource, and enforcement of fisheries regulations (Darwall and Allison 2002).

One option to deal with catch misreporting has been the estimation of the total catch to account for the full impact of fishing. Fisheries science and management bodies in developed nations (e.g. the National Marine Fisheries Service in the United States and the International Council on the Exploration of the Seas in Europe) often devote considerable effort to account for catch misreporting issues for commercially or ecologically important fisheries. This accounting is performed on a stock-by-stock basis, and cannot be done easily for a whole country. To fill this gap, there have been extensive efforts by the Sea Around Us Project to quantify the world's misreported catch via reconstruction methods to produce a corrected global database of fishery catches (Pauly and Zeller 2016, Watson 2017). Catch reconstruction methods involve a search for local clues to missing data, interpolation between data points, and an analysis of the sale and consumption of fish products (Zeller et al. 2007; Plagányi et al. 2011). The general motivation behind these catch reconstructions is that estimating the degree of misreported catch is better than assuming it is zero (Pauly 1998; Pitcher et al. 2002; Cressey 2015).

However, monitoring of fisheries ecological, social, and economic performance is only one aspect of a management strategy. The data collected from these monitoring programs or catch reconstructions is then fed into the stock assessment process, and in regions with effective fisheries management, becomes the basis for managing fisheries via decision rules. Although much work at the national and global level has aimed to reconstruct catches, this is not possible for many of the world's decentralized, small-scale fisheries due to the sheer number of species harvested. Another issue is that reconstructions en masse may be biased up or down due to assumptions that simplify the complex nature of many fleets (Pauly et al. 2014). The resulting incorrect catch records could in turn bias estimates of population status and sustainable yield coming from fisheries stock assessments, although the direction and magnitude of these biases are not obvious. One assumption is that under-reported catch leads scientists and managers to think the fish population is doing better than it truly is, in turn leading to higher recommended harvest limits than would be sustainable (Metuzals et al. 2008). Individual assessments on stocks with under-reported catch have estimated both lower total population sizes and lower sustainable harvest limits than assessments using catch reconstructions, although these studies have not teased out the effect of the under-reported catch from other simultaneous changes in data and model structure (Patterson 1998; Groeneveld 2003; Hammond and Trenkel 2005; Zeller et al. 2008; Plagányi et al. 2011). Thus there is a need to explore how changes in catch reporting rates over time impact fisheries stock assessment outputs and the resulting catch limits set by management.

The objective of this study is to quantify the direct impact of different scenarios of misreported catch on stock assessment results and proceeding management recommendations.

1.2 MATERIALS AND METHODS

Rationale for different catch scenarios

We used simulations to compare the true and estimated values of total biomass, maximum sustainable yield (MSY), exploitation rate (u), carrying capacity (K), intrinsic growth rate (r), and population status (biomass B and exploitation rate u compared to that which would produce MSY; B_{MSY} and u_{MSY}, respectively). These simulations were conducted with the following catch misreporting scenarios: 100% reporting, constant under-reporting, constant over-reporting, increasing reporting rate over time, and decreasing reporting rate over time. Constant under-reporting occurs when distinct sectors of the fishery are not reported over time and the size of that sector remains constant, or when catch is under-estimated in catch reconstructions. Constant over-reporting occurs when promotions or bonuses are based on output, as in China (Watson and Pauly 2001), or due to assumptions made in catch reconstructions such as over-estimates of discard rates (Chaboud et al. 2015). Increasing reporting rates in major commercial fisheries in developed countries have resulted from increased observer coverage, better estimates of discards, and reduction in illegal harvest for some of the world’s most commercially valuable species, such as tuna (Agnew et al. 2009), and lobster (Hilborn et al. 2005). Declining reporting rates could result from changes in management, such as the adoption of a quota system, or increases in the number of artisanal and recreational fisheries over time as coastal populations boom and technology improves, as seen in the Gulf of Mexico (Coleman et al. 2004). These patterns are demonstrated in the catch reconstructions; most of the FAO areas have constant reporting rates over time, but some increase and some decrease (Pauly and Zeller 2016).

Modelling overview

To test the effect of misreported catches on estimates of stock status, we simulated the stock assessment process (Figure 1.1), generating time series of true population biomass and true catch, and running stock assessments on observed catch (based on different catch reporting scenarios) and survey indices sampled from the true biomass. We then estimated carrying capacity and maximum sustainable yield (MSY) using the Pella-Tomlinson surplus production model (Figure 1.1). We compared the estimated population trajectory, population parameters, and management reference points to their true values to better understand the impacts of magnitude and trend of catch reporting over time. We then projected the population forward based on management at estimated MSY reference points to demonstrate the impact of the different catch misreporting scenarios on meeting management targets.

Simulations

We used a single true catch time series for all catch misreporting scenarios, which increased to a peak within the first half of the time series, then decreased to very low levels compared to the peak, and then increased and stabilized during the second half of the time series. This two-way scenario has been shown to provide information about population status and growth rates compared to other patterns of catch time series (Magnusson and Hilborn 2007). We also tested a true catch time series that continuously increased over time (a one-way trip) in a sensitivity analysis.

The simulated "true" population follows a Pella-Tomlinson (discrete logistic) function over 35 years:

$$\mathrm{B}_{t}^{true} = \left(\mathrm{B}_{t-1}^{true} + \frac{z^{z/(z-1)}}{z-1} MSY^{true} \left(\frac{\mathrm{B}_{t-1}^{true}}{K^{true}} - \left(\frac{\mathrm{B}_{t-1}^{true}}{K^{true}}\right)^{z}\right) - C_{t-1}^{true}\right) e^{\tau_2 - \frac{\sigma_B^2}{2}} \quad (1.1)$$

where B_t^{true} is the true population biomass in year *t*, *z* is the Pella-Tomlinson scaling parameter that determines what biomass level produces maximum sustainable yield (MSY), *K* is the carrying capacity, C_t^{true} is the true catch in year t, and ε_t is the multiplicative process error in time *t*. The process error is normally distributed around zero:

$$\varepsilon_t \sim N(0, \sigma) \quad (1.2)$$

where σ_B^2 is the process error standard deviation (set to 0.1 in all scenarios to represent a reasonable level of stochasticity in the population dynamics). In this parameterization of the Pella-Tomlinson model, MSY is a function of the intrinsic rate of growth (*r*), *K*, and *z*:

$$\mathrm{MSY}^{true} = \frac{r^{true} K^{true} z}{(z+1)^{\frac{1}{z}+1}} \quad (1.3)$$

We chose the Pella-Tomlinson over a more complex age-structured model because we wanted to focus on the effect of catch misreporting in the targeted population, rather than including the interaction with age-specific selectivity, fecundity, and other dynamics. Understanding how catch misreporting impacts biases in age-structured models would be a logical next step of this analysis (for example, the discard or harvest of juvenile fish), but is beyond the scope of this study.

We assumed true values of *K*=1,000, *r*=0.2, and *z*=1.188. The value of *z* is fixed such that the biomass at maximum sustainable yield, B_{MSY}, is equal to 40% of *K*, which is the average for exploited marine fisheries based on a global meta-analysis of stock assessments (Thorson et al. 2012). All population simulations started with a biomass at 75% of the carrying capacity in the first year, thus explicitly assuming that exploitation started before the first year in the model. Starting at carrying capacity did not influence the overall findings. Other simulation studies have varied the population dynamics parameters within the simulation model (Hammond and Trenkel 2005; Ono et al. 2014), but we kept these variations to a minimum to demonstrate the impacts of misreported catch on stock assessment output.

From the true population generated by the operating model, including a true value of survey catchability, q, of 0.01, we generated a survey index I_t subject to multiplicative lognormal observation error δ_t with standard deviation σ_I (set to 0.1):

$$\begin{gathered} I_t^{obs} = q^{true} B_t^{true} \exp(\delta_t - \frac{\sigma^2}{2}) \\ \delta_t \sim N(0, \sigma_I) \end{gathered} \tag{1.4}$$

We assumed that survey data were available for every year for which catch data were available. By assuming a good quality survey index, we were able to tease out biases in stock assessment output associated with misreported catch, without the confounding effect of biases and uncertainty resulting from a poor survey time series. The generated survey index and the "reported" catch time series were then used as data inputs in the estimation model. The reported catch time series $C_t^{reported}$ was a function of the true catch time series C_t^{true} and the annual reporting rate R_t:

$$C_t^{reported} = C_t^{true} R_t \tag{1.5}$$

We considered five scenarios of catch misreporting: 100% reported in all years, constant under-reporting of R_t = 50%, constant over-reporting of R_t = 150%, linearly increasing R_t from 40% to 90%, and linearly decreasing R_t from 90% to 40%. The only changing variables across these scenarios were the magnitude and trend in catch reporting (Figure 1.1).

To estimate the expected bias from the catch misreporting scenarios on stock assessment output of interest, we ran a deterministic version of the operating model, with process and observation error set negligibly low (0.001), and we also ran stochastic simulations, where the operating model was run with standard deviation values for both process and observation error of 0.1. To visualize the variation expected in estimated parameter values given process and observation error, we iterated each catch misreporting scenario 1,000 times, generating different true populations (from process error) and survey time series (from observation error) (Figure 1.1).

Estimation

The estimation model takes the reported catch time series and survey index as data inputs, and is also based on the Pella-Tomlinson surplus production model with *z* fixed at 1.188.

$$\begin{array}{ll} \mathrm{B}_t = KP_0 & \text{if t=1} \\ B_t = \left(B_{t-1} + \frac{z^{z/(z-1)}}{z-1} MSY \left(\frac{B_{t-1}}{K} - \left(\frac{B_{t-1}}{K}\right)^z\right) - C_{t-1}^{reported}\right) & \text{if t<1} \end{array} \qquad (1.6)$$

The estimated parameters include *MSY*, *K*, and P_0 (the proportion of carrying capacity in the first year). The survey catchability constant *q* can be derived analytically from the ratio of the survey index data and predicted biomass:

$$\hat{q} = \exp\left(\frac{1}{n}\sum \ln\left(\frac{I_t^{obs}}{B_t}\right)\right) \qquad (1.7)$$

where *n* is the number of years of data. The derived value of *q* is used to derive a predicted value of the survey index:

$$I_t = \hat{q}B_t \qquad (1.8)$$

The observation error, $\hat{\sigma}$, can also be derived analytically from the deviations between the predicted and observed survey index and the number of years of survey index data:

$$\hat{\sigma} = \sqrt{\frac{\sum\left(ln(I_t^{obs}) - ln(I_t)\right)^2}{n-1}} \tag{1.9}$$

The *K* parameter was bounded between the maximum catch and five times the true *K*, i.e. 150–5000 with true *K* at 1000. *MSY* was given an upper bound of five times the true value of MSY, i.e. 1–280 with true *MSY* at 56. P_0 was bounded between 0.01 and 1. The parameters *K* and *MSY* were estimated in log space. The estimation of these parameters involved minimizing a negative log likelihood function that assumes the indices are log-normally distributed, after removing constant terms:

$$NLL = n\ln(\hat{\sigma}) + \sum \frac{\left(\ln(I_t^{obs}) - \ln(I_t)\right)^2}{2\sigma^2} \tag{1.10}$$

The estimation model and minimization were performed in AD Model Builder (Fournier et al. 2012). The posfun function was used to ensure that biomass could not fall below zero to ensure differentiability. To guarantee that the model converged on the maximum likelihood estimate, for each simulation (iteration of population given catch time series, process error of true population, and catch misreporting scenario), we ran the model ten times with different randomly chosen starting values within 25% of the true parameter values, retaining the estimated parameter values that resulted in the lowest negative log likelihood.

We then compared the outputs of the estimation model, estimates of *MSY*, *K*, *r*, population biomass, and B_{MSY} to the true values from the operating model. We were particularly interested in the estimates of these key population and management parameters in the terminal year of data, which is the key information about population status that is available to a manager blind to catch misreporting after collecting 35 years of data.

Projections

Typically, stock assessments would be used by managers to recommend sustainable harvest rates, as exemplified by fishing at u_{MSY} into the future. This would be enacted by multiplying the estimated u_{MSY} by the estimated current population biomass to set a catch limit. We simulate the effects of this kind of control rule for five years into the future, assuming the same conditions of catch misreporting as in the terminal year of data. We then compared the expected equilibrium population size and overfished status (B/B_{MSY}) based on this harvest strategy for each scenario of catch misreporting.

1.3 RESULTS

We found that constant under-reported catch leads to a more conservative view of the productivity of the resource, with total population size, MSY, and exploitation rate under-estimated on average compared to the simulated truth (Figure 1.2l). The degree to which these parameters were under-estimated is proportional to the reporting rate: a 50% reporting rate results in estimated population size, MSY, and exploitation rate that is 50% of the truth. Conversely, constant over-reported catch by 50% leads to estimates of the population size, MSY, and exploitation rate that are 50% higher than the actual values (Figure 1.2m). Thus catch data are important in scaling the estimated population size and MSY: when reported catches are lower than the truth, the population is perceived to be less productive.

However, when reporting rates were constant over time, neither under-reporting nor over-reporting biased estimates of population status (Figure 1.3l,m,q,r). For these two scenarios, biomass and exploitation rate relative to those that would produce MSY (i.e. B/B_{MSY} and $\underline{u}/u_{\text{MSY}}$) were correctly estimated (Figure 1.2l,m). The reason these are correctly estimated is that the identical, proportional biases in biomass estimates and in MSY cancel out when calculating the ratios B/B_{MSY} and u/u_{MSY}. In other words, given a reliable index of abundance, only a proportional index of catch (which need not be absolute) is needed to obtain reliable estimates of population status relative to reference points.

The biases in population parameter estimates were not as straightforward when catch reporting rates changed over time. When the reporting rate improved over time, estimates of population size was 20% of the truth, MSY was 60% of the truth, and carrying capacity was 75% of the truth (Figure 1.2n). Population status relative to reference points was also poorly estimated, with B/B_{MSY} being far below, and u/u_{MSY} being far above, the true values (Figure 1.2n). Since our perception under this scenario is that the population is overfished, this results in precautionary management. Thus, under the scenario of increasing reporting rates over time, applying an estimated exploitation rate into the future that is thought to lead to biomass at B_{MSY} and catches at MSY will in reality result in catches below MSY, biomass above B_{MSY} (Figure 1.3n), and exploitation rates well below u_{MSY} (Figure 1.3s).

The opposite result occurs when reporting rate decreases over time. Here exploitation rate is estimated to be 20% of the true value (Figure 1.2o), and the population size is estimated to be 90% of the truth. Similar to when the reporting rate improves over time, these non-proportional biases prevent unbiased estimates of population status. Now, though, we are in the danger zone where we have an overly-optimistic perception of the population, analysts would erroneously perceive the population to be lightly fished, and future catches would be set too high, resulting in overfishing. Thus future projections that involve deteriorating reporting rates demonstrate that if exploitation rate is estimated to be far below u_{MSY}, managers aiming to fish at the u_{MSY} target will rapidly increase catches and cause swift population collapse (Figure 1.3e,j,o,t). Thus, from a conservation perspective, it is far worse to have declining reporting rates over time than increasing reporting rates over time.

Sensitivity tests examining the influence of different values of K and r revealed that these variables did not influence the overall patterns found. The only exception in this sensitivity analysis were estimates of status relative to MSY reference points when r was low (e.g., r=0.05) and there were trends in misreported catch over time. These differences were likely due to the lack of response in the population to the changing population due to slow growth rates, and the decreased contrast in the catch time series with trends in misreporting. Use of the one-way catch time series decreased the ability of the simulation model to estimate population parameters reliably even with no observation or process error, thus adding uncertainty in parameter estimates and direction of biases given catch misreporting. Therefore, the biases reported in this study are less applicable to very slow-growing fish, or with less informative catch time series, but the overall message remains consistent: exploitation rate will be over-estimated with increasing reporting, under-estimated with decreasing reporting, but unbiased with constant misreporting, and unintuitive biases in estimates of stock status relative to reference points will occur with trends in misreporting, but not with constant misreporting.

1.4 DISCUSSION

Our results demonstrate how misreported catch at a constant rate can still lead to unbiased estimates of stock status and sustainable management. The biases in this case are simply in the estimates of the scale of population size, but do not affect the estimates of relative values such as exploitation rate and estimated biomass to B_{MSY} ratios. Trends in catch reporting can, however, lead to unintuitive biases in estimates of stock status. The biases in estimates of population parameters are not proportional to the reporting rates, nor are they identical. When trends in catch misreporting have occurred or are occurring, improved catch data can only help to estimate sustainable reference points and stock status.

Thus, our results demonstrate that monitoring programs and catch reconstructions should be geared towards understanding the trend in catch reporting over time when data collection and research capacities are limited in management strategy development. These findings have far-reaching implications for assessments with misreported catch, fisheries for which assessments could be conducted based on new catch reconstruction data, and the development of future monitoring and assessment programs. The trend in reporting rates is likely cheaper to estimate than the accurate magnitude of catches. These trends in misreporting can then be accounted for in stock assessments, leading to more sustainable catch-based harvest control rules. While catch reporting rates are likely increasing for many fisheries with either improved accounting or decreasing of discard rates (Kelleher 2005; Davies et al. 2009), it is likely that reporting rates for other fisheries may be decreasing as demand for seafood increases and more people participate in commonly unreported artisanal fleets (Salas et al. 2007). However, many of the catch reconstructions from FAO areas appear to demonstrate fairly constant misreporting rates over time (Pauly and Zeller 2016). Thus, the catch misreporting scenarios of the world’s fisheries are likely as varied as the world’s fisheries themselves. Case-specific simulation studies would help managers understand the potential biases in assessment results due to the misreporting of catch and effort for the fishery (Omori et al. 2016).

While understanding the total catch is critical in understanding overall ecosystem impacts and the economic value of fisheries, total catch alone is insufficient to conduct stock assessments and choose sustainable harvest control rules. When survey data are not available or precise, catch data are increasingly used alongside biological information in catch-based, data-limited assessment methods (Carruthers et al. 2014). With minimal information to account for age structure, these methods often use biomass dynamic models similar to the one used in this study (Dick and MacCall 2011). With constant misreporting, we would expect these data-limited assessment methods to estimate population size and MSY scaled to the level of misreporting, without biased trends. However, without well-informed estimates of current population depletion, these methods would likely perform poorly in estimating population status when catch reporting rate changes over time. Approximately 78% of global reported fishery catch has undergone some form of stock assessment (FAO 2016) with 53% of global reported catch accounted for within the RAM Legacy Stock Assessment Database (Ricard et al. 2012; Costello et al. 2016; RAM Legacy Stock Assessment Database 2017). However, the remaining 22% of global reported catch remains unassessed, and nestled in this category are the small-scale and non-commercial fishery sectors that are often unreported. These same fisheries provide crucial nutrition to densely populated coastal peoples with otherwise low protein diets (Food and Agriculture Organization 2014), adding even more weight to the critical need to estimate trends in catch and improve catch reporting.

Catch reconstructions show great potential for finding efficient outcomes in the tradeoff between resource utilization and conservation when paired with the stock assessment and decision rule aspects of management strategy development. For example, when the Hawaiian bottomfish resource was assessed with a catch time series that included both reported commercial and reconstructed recreational catch, the estimated MSY increased compared to that which was estimated when only reported commercial catch was considered (Zeller et al. 2008). Accounting for the unreported sector provides more confidence to scientists and managers that the assessment has improved and the advantage of a potentially increased harvest limit for fishermen due to the higher estimated MSY. However, the benefit of this reconstruction is its resolution to catch at the population level, as opposed to the country level.

As has been previously experienced by FAO, catch monitoring and reconstructions are expensive to conduct on a stock-by-stock basis, and unrealistic to keep up-to-date at a time scale relevant for management (Darwall and Allison 2002). Many published reconstructions, such as that for small-scale fisheries in the U.S. flag-associated island nations, are at the country level (Zeller et al. 2007). Our results show that in terms of setting sustainable harvest limits, understanding this general trend at the population level is more important than the overall catch trend in the country. This is increasingly important as catch reconstructions replace previously misreported catch time series in fisheries stock assessments. Since catch reconstructions involve making assumptions based on limited data on unreported catches (Belhabib et al. 2014, 2015; Chaboud et al. 2015), reconstructed catch time series could conceivably be lower, higher, or have different trends than true catches. Given these scenarios, our results provide substantive advice for making management decisions in the face of uncertainty.

Our study shows that improved catch reconstructions and catch monitoring programs are valuable in terms of assessing the scale of harvest, since they result in better estimates of total food produced and income earned from fishing. Accurate catches are also necessary to estimate true MSY, biomass, and carrying capacity of a particular fishery, and hence the relative importance of that population as predator or prey within food webs. However, catch reconstructions and monitoring would need to be conducted stock by stock, which is time-consuming and expensive. While the time and money spent to undergo catch reconstructions or improve monitoring programs have their advantages, moving forward in the fisheries management process will likely require management strategies robust to limited data and capacity for most of the world’s fisheries. Some of the key components to catch reconstructions, such as information from local experts, could help to identify trends in catch reporting over time. Catch time series could then be adjusted accordingly, or this expert information would be useful for various other assessment methods and harvest control rules. While expert information may not account for all harvest from the system, it could still be used to help set sustainable harvest rates into the future. Under-reported catch itself does not lead to overfishing unless reporting rates deteriorate over time, suggesting a key area of focus for catch reconstruction efforts, resource management strategies, and data collection.

Acknowledgements

This work was conducted while 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. This work has benefitted from feedback from R. Hilborn and A.E. Punt for comments on the manuscript. M.B.R. conducted the analyses. Both M.B.R. and co-author T.A. Branch designed the study, discussed the results and jointly wrote the manuscript.

1.5 FIGURES

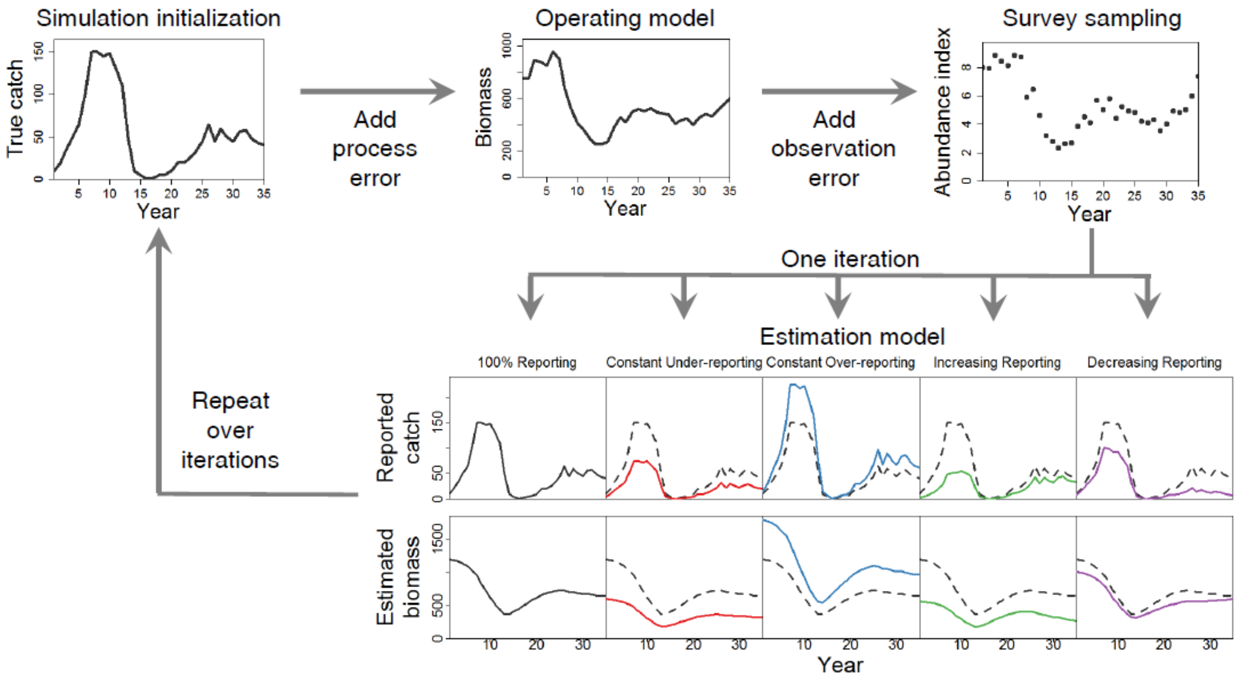

Figure 1.1. Flow diagram of the main steps in the simulation study. The Simulation Initialization consists of a deterministic, true catch time series, which is used in the Operating Model, along with process error, to generate a true population biomass. The Survey Sampling step then samples an abundance index, which is used as a data input along with reported catch time series in the Estimation Model. The Estimation Model estimates the parameters of a biomass-dynamic model to derive population biomass and management reference points (dashed lines indicate the true trajectory, solid lines indicate the reported catch under the respective catch misreporting scenario).

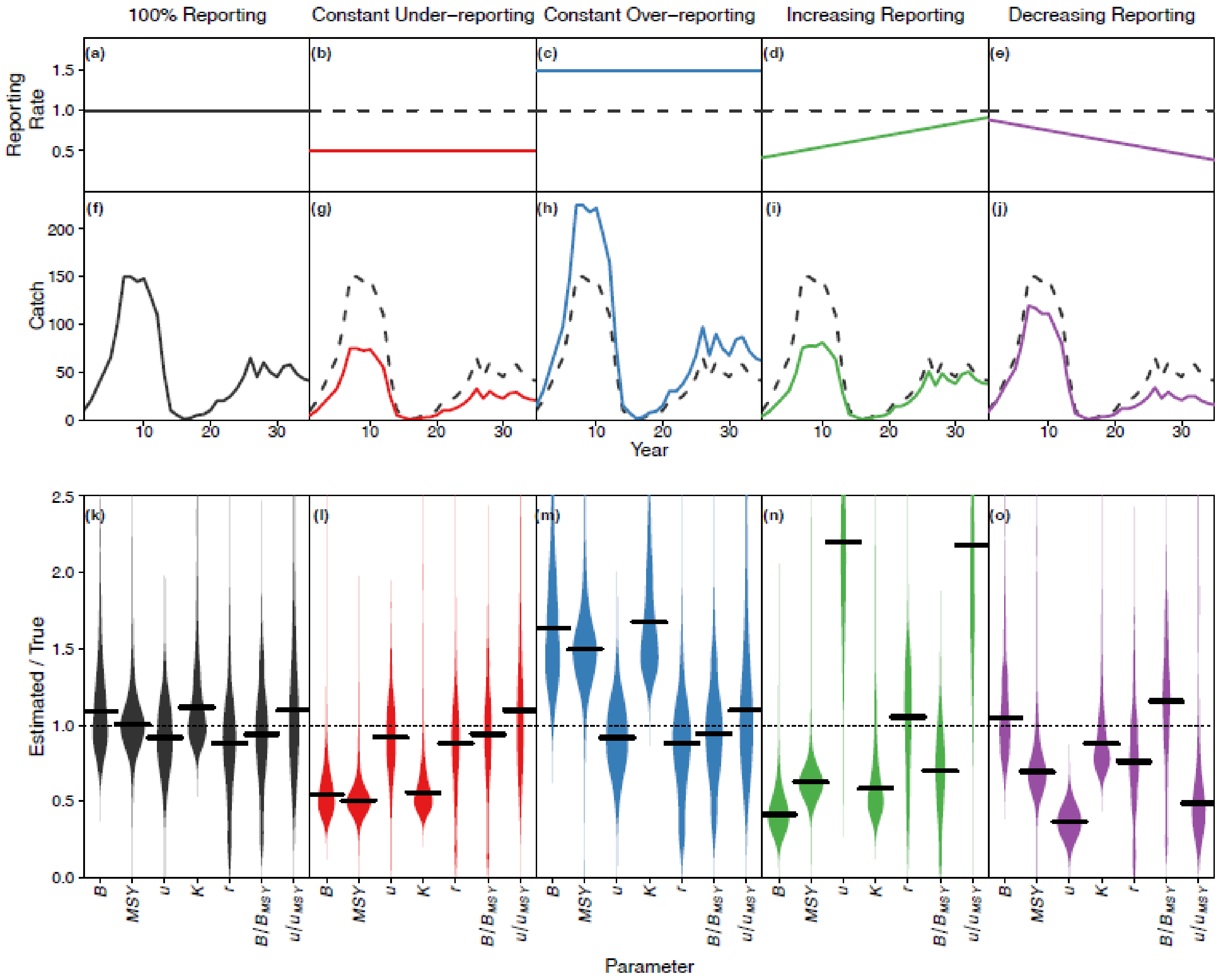

Figure 1.2. Effect of different patterns of catch reporting on estimates of population parameters. The reporting rate (a-e), time series of catches (f-j), and the distribution of estimated:true parameter values for 1000 simulation runs (k-o) are displayed in each row, respectively. The catch scenarios by column are (a) 100% reporting, (b) constant under-reporting, (c) constant over-reporting, (d) increasing trend in reporting, (e) decreasing trend in reporting. The ratio of estimated:true is given for the following parameters: biomass at the end of the time series (B), maximum sustainable yield (MSY), exploitation rate (u), carrying capacity (K), intrinsic rate of growth (r), biomass relative to that yielding MSY (B/B_{MSY}), and exploitation rate relative to that yielding MSY (u/u_{MSY}). Dashed gray traces in panels b-e and g-j represent the 100% reporting rate and catch, respectively. The thinner, dashed line in panels k-o represents where the estimated and true values of each parameter are equal, and the short black lines represent the median of the distribution for the 1000 simulation runs.

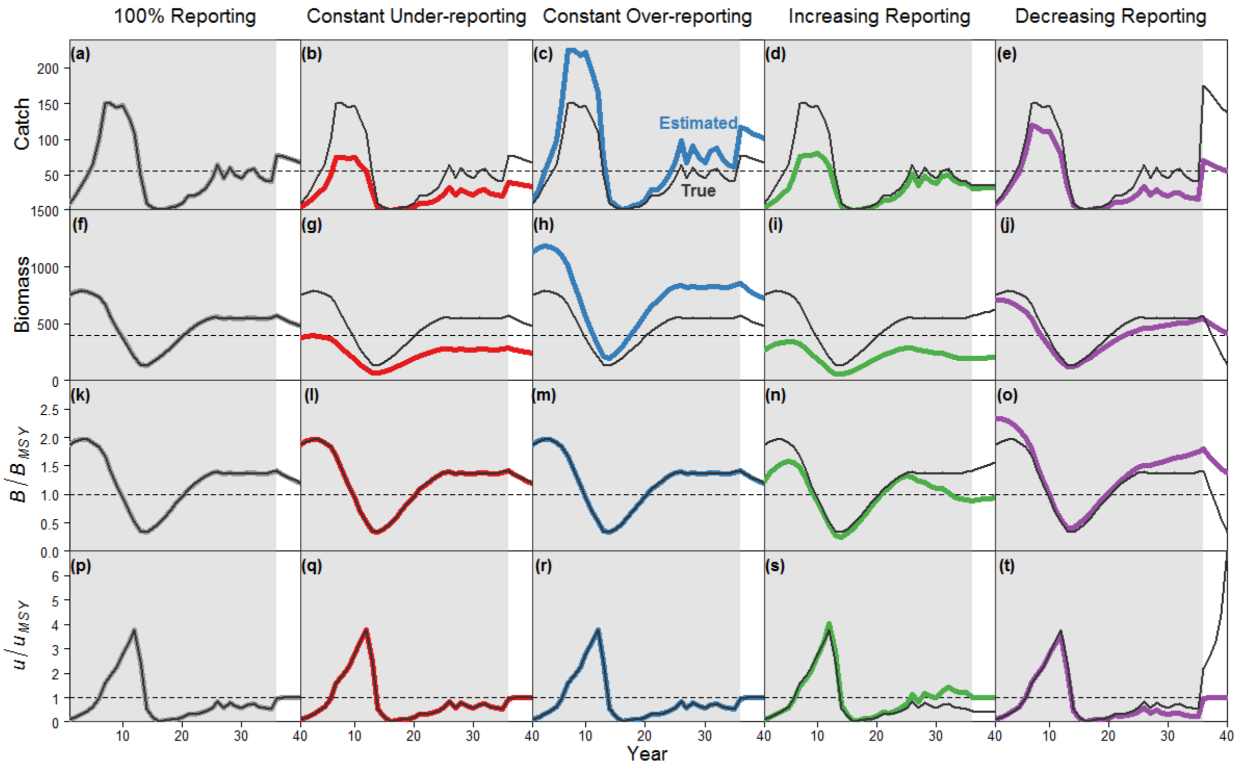

Figure 1.3. Time series of estimated (thick lines) and true (thin lines) status, and future projections. These are deterministic trajectories of (a-e) catch, (f-j) biomass, (k-o) B/B_{MSY}, and (p-t) u/u_{MSY} under each scenario for catch reporting (thick lines). Shaded background areas designate the time period with catch and survey data available, and white areas indicate the five-year projections where harvest rate is set at u_{MSY}. All three scenarios with constant reporting rates lead to sustainable management, but trends in catch misreporting lead to biomass that is higher than MSY levels for increasing reporting rates (i), and lower than MSY levels for declining reporting rates (j). The dashed line indicates true MSY (a-e), true B_{MSY} (f-j), target B/B_{MSY} (k-o), and target u/u_{MSY} (p-t).

Chapter 2. ACCOUNTING FOR VARIABLE RECRUITMENT AND FISHING MORTALITY IN LENGTH-BASED STOCK ASSESSMENTS FOR DATA-LIMITED FISHERIES

Abstract

In fisheries with limited capacity for monitoring, it is often easier to collect length measurements from fishery catch than quantify total catch. Stock assessment tools that rely on length measurements without total catch cannot 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 requires a single year of length composition data and basic biological information at a minimum, but can fit to multiple years of length composition 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 moves data-limited fisheries stock assessments forward in its flexibility to adapt as additional years of data or more data types are available, and obviates the need for equilibrium assumptions.

2.1 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 (MSC) 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, modeling tools can help with conflict resolution (Butler et al. 2006), supplement and support local knowledge (Neis 1992, Azzurro et al. 2011), and evaluate management strategies for fishery resources (Carruthers et al. 2014). Many stocks worldwide remain unassessed, e.g., the U.N. 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 1,268 stock assessments (Ricard et al. 2012). In the following, we define “data-limited” as any stock with uninformative data (no contrast to 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, Chrysafi and Kuparinen 2015, Dowling et al. 2015).

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 non-target species (Harley et al. 2001, Nadon et al. 2015, Prince et al. 2015b). 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 quick assessment of stock status relative to unfished levels assuming equilibrium conditions (Hordyk et al. 2015, Prince et al. 2015b). 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 fishing mortality directly influences mean length of the catch, and have been used for assessments in the U.S. 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). 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. 2014b, 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. With both variable fishing mortality and recruitment processes occurring on the same population as well as errors when measuring mean length, it is possible 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. 2015b). 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 lots of small fish in the catch is caused by a high recruitment year class, 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 only length composition and basic biological information are available. This method builds upon the catch-curve stock reduction analysis (CCSRA) 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 MSY-based reference points without the assumption of final biomass relative to unfished biomass needed to conduct 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. LIME estimates reference points more precisely than LB-SPR under many common scenarios, with the added benefit of accounting for recruitment variation and potentially including catch or an abundance index that may be available in some data-limited stock assessment methods.

2.2 MATERIALS AND METHODS

LIME is an age-structured population dynamics model with the ability to (1) account for variable fishing mortality and recruitment when only length composition data are available, and (2) to 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 a parameter representing the ratio of input and effective sample size (Thorson et al. in press) as fixed effects. 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, the mean recruitment will be fixed to a relative value of 1.0. 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 providing information on scale of the population size, LIME would estimate equilibrium 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 (Figure 2.1). 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. 2015) as implemented in our R package *LIME* (https://github.com/merrillrudd/LIME).

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 (Figure 2.2). These life history types were: (a) a short-lived fish, mimicking rabbitfish (*Siganus sutor*, L_∞=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_∞=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_∞=90.0 cm, k=0.13, M=0.18, L_m^{50}=50.0 cm, maximum age=26, Heemstra and Randall 1993) (Table 2.1). 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 (Figure 2.3). The first is the “equilibrium scenario”, which matches the non-variable 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 seven years of the 20-year time series, then constant fishing at $F_{20\%}$ for seven years, then the fishing rate decreasing linearly down to half of $F_{20\%}$ for the last six years of the time series. $F_{20\%}$ was fixed to 3.0 for the short-lived life history type since it increased to a level above reasonable capacity. Recruitment is variable and autocorrelated over a 20-year period. This scenario included a standard deviation of recruitment residuals σ_R equal to 0.737, the mean of the predictive distribution from a meta-analysis of recruitment variability in global fish orders (Thorson et al. 2014b). We assumed a first-order autoregressive coefficient equal to 0.426, the mean of the predictive distribution from the same meta-analysis of recruitment variability for global fish orders (Thorson et al. 2014b). A third scenario, the "one-way base scenario", involved the same recruitment variability and autocorrelation as the base 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. 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 the natural mortality rate was constant, known, and independent of size or age. We modeled individual growth using a von Bertalanffy growth function (Table 2.2, eq. 1; Figure 2.2). We assumed maturity at length m_j was based on a one-parameter logistic function (Table 2.2, eq. 2; Figure 2.2). 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 2.2, eq. 3). Selectivity-at-length follows a two-parameter logistic model (Table 2.2, eq. 4; Figure 2.2), with estimated parameter length at 50% selectivity and a second parameter δ representing the difference between length at 95% and 50% selectivity. We modeled weight at age as an allometric function of individual length at age

(Table 2.2, eq. 5). We calculated annual total biomass as a function of the abundance and weight-at-age (Table 2.2, eq. 6). Spawning biomass was a function of the total annual biomass and the proportion mature at age (Table 2.2, eq. 7). These processes contributed to an underlying age-structured model (Table 2.2, eq. 8). Parameter definitions and input values are listed in Table 2.1.

Data generation

We generated length composition data by simulating underlying age-structured dynamics and then instantaneously sampling length composition from the vulnerable population during each year. First we calculated the probability of being in a length bin for individuals of each age (Table 2.3, eq. 1). We then calculated the probability of harvest in each length bin each year based on the size of the vulnerable population abundance at age (Table 2.3, eq. 2). We used a multinomial probability distribution to generate the length frequency of samples of the catch in each length bin over time (Table 2.3, eq. 3).

We assumed a sample size n of 200 individuals were measured annually, and that this was the effective sample size of the length composition data. We compared the base case of 200 length measurements annually to model performance under sample sizes of 1,000, 500, 50, and 20. We used the large sample size of 1,000 to confirm 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 composition data, assuming one and ten years of length composition data were available (the “one length composition”, “ten length compositions” scenarios). Data from these two scenarios were used to demonstrate the value of additional years of length composition 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 composition 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 composition data; (iii) “index plus 1” with 20 years of an abundance index and 1 year of length composition data; (iv) “catch plus 10” with 20 years of total catch data and 10 years of length composition data; and (v) “catch plus 1” with 20 years of total catch data and 1 year of length composition data. We assume that catch and abundance indices are lognormally distributed with a log-standard deviation of 0.2 (Table 2.1). The “catch plus 1” scenario is essentially a stock-reduction analysis using 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 2.3, eq. 4), summing to obtain total annual catch (Table 2.3, eq. 5), 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 2.3, eq. 6).

When not otherwise stated, we generated data on an annual time step, which assumes length composition data are collected instantaneously within the year. In fisheries where only length composition data and biological information are available (i.e. no catch or abundance index) it is possible length composition data would be available on less than an annual time step. We tested scenarios where length data were collected on a monthly time step, 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 composition 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 (Chapter 3).

Estimation model

LIME requires at least one year of length composition 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-weight parameters, natural mortality (M), and the one-parameter logistic maturity-at-length schedule. 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. Estimates of relative reference points (e.g. SPR) can be derived, but reference 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 2.1.

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 2.4, eq. 1), and the estimated recruitment standard deviation (Table 2.4, 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 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 2.2, eq. 7) but without the fishing mortality and selectivity terms (Table 2.2, eq. 9).

The structure of LIME follows the structure of the operating model using equations 1-8 in Table 2.2 and equation 1 in Table 2.3 to derive the predicted catch, index, and length composition.

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 2.4, eq. 3). Drastic changes in fishing mortality between years is 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 2.4, 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 composition data, log-probability of fishing mortality and recruitment variation, and the log-likelihood of the catch and abundance index, if available. For scenarios that include abundance index and catch data, a lognormal probability distribution was assumed to describe error in both data types (Table 2.4, eq. 7, eq. 8). The respective observation errors σ_I and σ_C are fixed *a priori* (not estimated as parameters). We assumed the length composition 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 composition 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 as it was designed for continuous composition data, but the parameters do not correspond to the easily interpretable effective sample size of length composition 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. 2016). The effective sample size is a nonlinear function of input sample size (Table 2.4, 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 0.01. If the initial model run did not converge (resulting in NAs or a high final gradient), the model would be run up to ten additional times with starting values equal to the estimates from the non-converged 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 one year (LBSPR 1) and ten years (LBSPR 10) of length composition data using the operating model mimicking the age structure of LIME. LB-SPR requires as input the length composition 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-weight parameters, and starting values for the length at 50% and 95% selectivity (Table 2.1). We assumed 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 LIME 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 composition data (Hordyk et al. 2015).

It is important to note, however, that the above methods are cross-testing LB-SPR with the LIME operating model, while LIME is being self-tested under its own model structure with instantaneous sampling within each time step (Needle 2014). To further compare the strengths and limitations of each method, we used the *LBSPR* package to simulate length composition data under the LB-SPR equilibrium, length-based structure, self-testing LB-SPR and cross-testing LIME under equilibrium conditions for all three life history types.

A major issue in the comparison of models was the time step used. To test model performance, we cross-tested both the LIME and LB-SPR methods using the monthly data generation feature in the operating model, where sampling is spread out over all twelve months instead of collected instantaneously at one point. We compared LIME and LB-SPR performance when we pooled the monthly length data into an annual distribution and then ran each assessment on an annual time step. These additional simulation tests addressed the objective of comparing the two methods, 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 2.2, eq. 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 to other length-based assessment methods.

Sensitivity tests

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 composition 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, ages in months instead of years, and dividing the input natural mortality rate M by twelve so that mortality occurs in each month.

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 composition data of 20, 50, 100, 500, and 1,000 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 2.4, eq. 9-10) between estimated and true SPR in the last year of data across the 100 iterations of simulated data. We used median relative error (MRE; Table 2.4, eq. 9) to quantify bias, and median absolute relative error (MARE; Table 2.4, eq. 10) to quantify precision (Ono et al. 2012). 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 in 50% of simulation replicates). Coverage is contrasted with bias and precision to interpret results for management purposes.

2.3 RESULTS

LIME performance across life history, variability, and data scenarios

Based on the Monte Carlo simulations, LIME can estimate unbiased SPR when length composition data are available and biological characteristics are correctly specified across various scenarios of fishing mortality and recruitment patterns (Figure 2.4). When self-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 2.5). The bias in SPR in the terminal year was -0.001 for short-lived, -0.005 for medium-lived, and -0.02 for longer-lived fish across all data availability and population variability scenarios with 200 length measurements annually. Additional data availability scenarios and variability conditions pinpointed further LIME strengths and weaknesses. With only one year of length composition data, LIME estimated bias in SPR of -0.02 for the short-lived life history, -0.04 for medium-lived, and 0.12 for longer-lived. Integrating ten years of length composition data improves accuracy in the estimation of SPR in the terminal year for LIME when applied to the short-lived and medium-lived life history types, but not necessarily for the longer-lived life history type (Table 2.5, Figure 2.4), decreasing bias by 45% for the short-lived life history and 72% for the medium-lived life history, but increasing bias by 25% for the longer-lived life history. Ten years of length data improved LIME accuracy for the long-lived life history type in the two-way fishing mortality scenario, but bias increased by 20% in the equilibrium scenario and 158% for the one-way fishing mortality scenario.

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 life history types, adding an abundance index to the length composition data improved accuracy by 46% and precision by 52%. An abundance index particularly improved LIME performance for the short- and medium-lived life history types. With ten years of length composition data and an abundance index, accuracy improved for the short-lived life history type by 79% and medium-lived life history type by 74%. Under this same data scenario bias increased for the longer-lived life history by 86% (Figure 2.4). The increase in bias for a longer-lived fish with an abundance index is likely due to data conflict with the length composition data. An abundance index with one year of length composition data is expected to decrease bias by 61%, but with ten years of length composition data the abundance index leads to bias of 9%.

Adding a catch time series to length composition data improved accuracy by 14% and increased precision by 25%. The improvement in accuracy is strongest under equilibrium conditions, with decreased bias of 75% and increased precision by 62% on average. Under conditions of variability including catch data with length composition increased bias by 57% and increased precision by 18% (Table 2.5, Figure 2.4). This was mainly due to high bias in the “catch plus one” scenario for short-lived fish (Figure 2.4). Excluding this anomalous scenario, including a catch time series with length composition data improved accuracy by 36% and precision by 29% under scenarios of variability. We also observed this anomaly to a lesser degree for one-way fishing mortality scenarios for the other life history types (Table 2.5, Figure

2.4). While the “catch plus ten” scenario removed the bias from the “catch plus one” scenario for the short-lived life history type under variability scenarios, it did not necessarily improve bias for the medium- and longer-lived life history types (Table 2.5). This indicates that it may be advisable to include a catch time series only when more than one year of length composition data are available for a short-lived life history type, and consider the possibility of over-estimating SPR when fishing mortality is changing over time for medium- and longer-lived life history types.

LIME converged for 86% of iterations of generated data across life history types, variability scenarios, data availability scenarios, and sample sizes of length composition data. LIME converged for 90% of iterations of generated data for the data-rich scenario (Figure 2.5). With only length composition data, LIME converged for 85% of iterations across scenarios. Including a catch time series or abundance index along with length composition data, the convergence rates were 86% for each. Non-convergence 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 55% across all scenarios of life history, variability scenarios, data availability, and sample sizes of length composition data. This indicates LIME performs well, with close to nominal coverage in that the true value of SPR should fall within the 50% confidence intervals for 50% of iterations by the definition of a confidence interval. However, LIME is more conservative (estimating a smaller confidence interval) or permissive (estimating a wider confidence interval) when we break down the interval coverage by life history, variation, and data availability scenario (Figure 2.5). The “catch plus one” scenario usually had the lowest coverage, 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. The longer-lived the fish, the more permissive LIME is in estimates of uncertainty. For the short-lived life history type, LIME interval coverage was only 37%. 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 was more permissive in its estimates of uncertainty for the population parameters for the medium-lived life history type, with interval coverage 66%, and longer-lived life history types, with an interval coverage of 62%.

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

Bias and precision in SPR estimates generally improved with more independent length measurements collected annually (Table 2.6). The short-lived life history types experienced the strongest improvements in accuracy between 20 and 1,000 length measurements collected annually (Table 2.6). While some scenarios within the life history and variability cases demonstrated the ability to estimate unbiased SPR with fewer than 100 samples per year, there was high variability in performance across cases, and low precision even in cases estimated to be unbiased. For the short-lived life history type, bias was less than 10% across variability scenarios with 100 or more samples per year. LIME was unable to estimate unbiased SPR for any sample size with two-way fishing mortality for the medium or longer-lived life history type with only one year of length composition data (Table 2.6). For other fishing mortality scenarios, performance generally hit a plateau with 100 samples or greater per year.

Like all age or length-based methods, LIME performance is sensitive to the correct specification of life history information (Figure 2.6). When all biological parameters are correctly specified, LIME estimates unbiased SPR on average with ten years of length composition data (MRE = -0.019). When L_∞ was misspecified as 25% greater than the truth, bias increased to -0.424, 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 composition 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, with bias of 0.847. When k was assumed to be 25% lower than the truth, LIME estimated SPR higher than the truth on average, with bias of 0.481. Alternatively, assuming k was 25% higher than the truth did not result in biased estimates of SPR on average, with bias -0.028. However, this apparent lack of bias is simply due to the median being close to zero between some biases greater than zero and some less than zero, while none are unbiased. As expected, LIME estimated biased SPR lower than the truth on average when k was misspecified 25% higher than the truth, with bias of -0.220 for the medium-lived life history type and -0.419 for the longer-lived life history type. When k was misspecified 25% higher than the truth for the shorter-lived life history type, LIME estimated SPR higher than the truth on average, with bias 0.201. This pattern is unexpected, and could have been due to unrealistically high values of k tested for the short-lived fish, where k is already high. When M was fixed 25% higher than the truth, we estimated SPR to be higher than the truth, with bias of 0.259 (Figure 2.6). 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. LIME estimated SPR biased lower than the truth when M was assumed to be 25% lower than the truth, with bias of -0.341.

Assuming the length at 50% maturity 25% higher than the truth resulted in biased SPR lower than the truth, with bias -0.414 on average. 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, with bias of 0.158 on average. 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, with bias of -0.035, and a higher CV_L resulted in a relatively higher estimate of SPR on average, with bias of 0.053.

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 2.5). However, testing of LB-SPR against the age-structured LIME operating model with instantaneous sampling resulted in poor performance for the short-lived life history type under equilibrium conditions (Table 2.5, Table 2.7, Figure 2.7). When self-testing LB-SPR under its own operating model with relative ages based on the *M/k* ratio, we verified that LB-SPR performs well across life history types. LIME, on the other hand, under-estimated SPR for the medium- and longer-lived life history types, and over-estimated SPR for the short-lived life history type with the LB-SPR operating model (Table 2.7, The cross-test for both methods, when length data were collected monthly and pooled into an annual length composition, demonstrated decreased LIME performance and improved LB-SPR performance for the short-lived life history type (Table 2.7, Figure 2.7).

LIME's over-estimation of SPR when length data were collected on a monthly time step for the short-lived life history type was mitigated when we ran LIME on a monthly time step. When length data were collected on a monthly time step for short-lived fish, running LIME on a monthly time step improved accuracy by 95% with ten years of length composition data and 88% with one year of data. Precision improved by 86% with ten years of length composition data and 27% with one year of data (Table 2.7). Chapter 3 goes into more detail on the importance of running LIME on shorter time steps for shorter-lived life history types.

2.4 DISCUSSION

This study demonstrates that the length-based integrated mixed-effects (LIME) method can be a valuable tool for fisheries stock assessment when at least one year of length composition data and basic biological information are available for the species (Table 2.8). 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 and when recruitment and fishing mortality are believed to be changing more regularly 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, Nadon et al. 2015) while catch-only methods programmed with underlying age dynamics as opposed to length (Martell and Froese 2013, Thorson and Cope 2014) fall in a separate category. The MULTIFAN (Fournier et al. 1990) and Stock Synthesis (Methot and Wetzel 2013) programs can use length data instead of age to gain information on cohort strength and total mortality, but require a catch time series to estimate stock status. The flexible ability 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 other length-based methods as a diagnostic tool to see if variation in recruitment or fishing mortality are being predicted by the model based on the length composition 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 composition data or other data types in the same assessment framework. LB-SPR is a quick way of conducting an assessment with only one year of length composition data, but as monitoring programs improve, the method must be applied independently over multiple years (Prince et al. 2015a, 2015b). There are also many catch-based methods that assume a catch time series is available, but they are not thoroughly tested in fitting to length, rather than age, composition data (Sabater and Kleiber 2013, Martell and Froese 2013, Thorson and Cope 2014). LIME allows for these data to be included into the same assessment framework, rather than requiring the capacity to switch assessment programs as monitoring programs develop and new data types become available (Maunder and Punt 2013). There are also not many tools available when there is an abundance index and length data, but no 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 scenarios where trends in abundance are available but total harvest is not.

Finally, the *mixed effects* aspect of LIME extends the current length-based methods by using the data to inform the magnitude of variability of time-varying processes, notably recruitment. Estimation of random effects has been integrated into existing stock assessment models, such as Stock Synthesis (Thorson et al. 2015a) and the state-space assessment model (SAM) framework (Nielsen and Berg 2014). Many assessments for European stocks now use SAM 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 communicate the virtues 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 can allow people to run the model in a user-friendly interface.

Simulation testing in this study demonstrated much better LIME performance for short- and medium-lived life history types with only length composition data given the simulation tests. With only length composition data, LIME performs well for the shorter-lived fish likely because the model is tracking cohorts through the length composition data to estimate recruitment deviations. This is likely difficult for the longer-lived life history 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 composition data. With short-lived fish that only live to four years old, each length more clearly matches up with an age group. Regardless of fish growth, one year of length composition data holds information on an entire generation for a short-lived fish. Even if recruitment is occurring monthly or seasonally and not annually, a ten-year time series of length composition data hold several generations' worth of information to inform whether there have been higher or lower recruitment pulses in that period.

By conducting a combination of self- and cross-testing for the LIME and LB-SPR methods, we identified scenarios in which each method excels. If the population is at equilibrium, data collection occurs continuously throughout the year, and only length composition data are available, LB-SPR performs better than LIME for estimating SPR. If the population is not in equilibrium, then LIME performs better.

There are some potential biases in LIME. 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. This bias is mitigated for medium- or longer-lived fish, where the growth during the year is less rapid. In this case, LIME should run using monthly time steps or shorter, to account for fish growth during the year. The main takeaway points for using the LIME and LB-SPR methods are that LB-SPR is expected to be biased if equilibrium conditions are violated, LB-SPR likely under-estimates SPR for short-lived fish if data are collected instantaneously, and LIME is expected to over-estimate SPR for short-lived fish if data are collected continuously but evaluated on an annual time step.

As expected, collecting more data (ten years of length data instead of one) resulted in greater accuracy and more precision in LIME estimates of SPR. An important question for management is often how many years of data are enough, particularly for fisheries where monitoring funding or capacity are 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 one year of length composition data on average (Figure 2.8). Managers will have to determine how much uncertainty they are comfortable with to set monitoring goals. The longer the time series, the better information on the processes occurring within the system that can help to identify the underlying explanations for the observed data.

In many cases, including an abundance index or catch time series is expected to improve accuracy and precision. However, the simulation study also identified some potential issues with data conflict when an abundance index or catch time series are included with length composition data. Catch data provides 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 one year of length composition data are often overcome with more years of length composition data. The exception is for the longer-lived life history type, for which we demonstrated LIME performs worse. To avoid data conflict, Maunder and Piner (2015) 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 2015). 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 2015). 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 a vital precondition 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. 2015a). 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. 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.

While length-based stock assessments are good starting points for making management decisions when limited data and capacity are available, there is no “silver bullet” assessment method that can consistently increase certainty in stock status with limited information. 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 one or two years of length composition 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, the toolbox of available methods is often restricted in the range of assumptions that can be tested. We have shown that LIME presents a way to represent uncertainty in both process and observation processes. LIME can complement other length-based (LB-SPR, mean length mortality estimators) and catch-only methods (Rosenberg et al. 2014, Carruthers 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. I thank J. Cope, R. Hilborn, T. Branch, 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.

2.5 TABLES

Table 2.1. 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."). Values separated by commas indicate alternate values for sensitivity analysis. Note that scenarios including "LC" are each divided into two scenarios with either one or ten years of length composition 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).

	Description	Input value			Data Scenario			
Symbol		Life history type				Index + LC	Catch + LC	LC only
Biological		Short	Medium	Longer	Rich			
L_∞	Asymptotic length	36.2cm	64.6cm	90.0cm	Fixed	Fixed	Fixed	Fixed
k	Brody growth coefficient	0.87	0.21	0.13	Fixed	Fixed	Fixed	Fixed
t_o	Age at length=0	-0.01	-0.01	-0.01	Fixed	Fixed	Fixed	Fixed
L_m^{50}	Length at 50% maturity	20.2cm	34.0cm	50.0cm	Fixed	Fixed	Fixed	Fixed
α	Length-weight scalar	0.0597	0.0245	0.0264	Fixed	Fixed	Fixed	Fixed
β	Length-weight 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	54cm	97cm	135cm	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.
Non-biological								
L_s^{50}	Length at 50% selectivity	11.3cm	20.0cm	25.0cm	Est.	Est.	Est.	Est.
δ	Difference L_s^{95} - L_s^{50} (expressed here as ratio L_s^{95}/L_s^{50} to compare across life histories)	1.3	1.3	1.3	Est.	Est.	Est.	Est.
σ_d	Standard deviation for right-side selectivity	20, 40	2, 5	15, 30	Sim.	Sim.	Sim.	Sim.
q	Catchability coefficient	1e-5	1e-5	1e-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							

Table 2.2. 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	$m_j = 1/(1 + \exp(L_m^{50} - j))$	Maturity-at-length
3	$m_a = \sum_{j=1}^{J} m_j \left(\frac{1}{L_a CV_L \sqrt{2\pi}}\right) \exp(-\frac{(j - L_a)^2}{2 L_a {CV_L}^2})$	Maturity-at-age
4	$S_j = 1/(1 + \exp\left(-\frac{\ln(19)\,(j - L_s^{50})}{L_s^{95} - L_s^{50}}\right))$	Logistic selectivity-at-length
5	$w_a = \alpha L_a^{\beta}$	Weight-at-age
6	$B_t = \sum_{a=1}^{A} N_{a,t} w_a$	Annual total population biomass
7	$SB_t = \sum_{a=0}^{A} N_{a,t} w_a m_a$	Annual spawning biomass
8	$N_{a,t} = \begin{cases} R_t, & a = 0 \\ N_{a-1,t}\exp(-M - F_t S_{a-1}), & 0 < a < A \text{ and t=1} \\ \frac{N_{a-1,t}\exp(-M - F_t S_{a-1})}{1 - \exp(-M - F_t S_{a-1})}, & a = A \text{ and t=1} \\ N_{a-1,t-1}\exp(-M - F_{t-1} S_{a-1}), & 0 < a < A \text{ and t > 1} \\ (N_{a-1,t-1} + N_{a,t-1})\exp(-M - F_{t-1} S_{a-1}), & a = A \text{ and t > 1} \end{cases}$	Abundance at age over time
9	$SB_0 = \sum_{a=0}^{A} R_0 \exp(-aM) w_a m_a$	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	$SPR = \frac{E_f}{E_0}$	Spawning potential ratio

Table 2.3. Functions for generating data in the operating model. All except Equation 3 are also used in the LIME estimation model to predict values for the observed data.

	Equation	Description
1	$p_{j,a} = \begin{cases} \emptyset\left(\frac{j - L_a}{L_a CV_L}\right), & j = 1 \\ \emptyset\left(\frac{j - L_a}{L_a CV_L}\right) - \emptyset\left(\frac{j - 1 - L_a}{L_a CV_L}\right), & 1 < j < J \\ 1 - \emptyset\left(\frac{j - 1 - L_a}{L_a CV_L}\right), & j = J \end{cases}$	Probability of being in a length bin given age
2	$\pi_j = p_{j,a} \frac{\sum_{a=0}^{A} N_{a,t} S_a}{N_t}$	Predicted probability of harvest by length bin
3	$\tilde{\pi}_j \sim Multinomial(n, \pi_j)$	Generated probability of harvest by length bin
4	$C_{a,t} = \frac{F_t S_a}{M + F_t S_a} N_{a,t}(1 - \exp(-M - F_t S_a))$	Annual catch at age
5	$C_t = \sum_{a=0}^{A} C_{a,t}$	Annual catch
6	$I_t = qB_t$	Abundance index

Table 2.4. Components of the joint likelihood function in the LIME model, as well as performance metrics across iterations of generated data.

	Equation	Description
1	$\mu_{Rt} = \frac{4hR_0 SB_{t-1}}{SB_0(1-h) + SB_{t-1}(5h-1)}$	Expected annual recruitment based on Beverton-Holt stock-recruit relationship
2	$R_t \sim Lognormal(\mu_{Rt}, \sigma_R^2)$	Annual recruitment R_t arising from a lognormal distribution
3	$F_t \sim Normal(F_{t-1}, \sigma_F^2)$	Penalty on annual fishing mortality F_t
4	$\sigma_R \sim Lognormal(0.7, 0.2^2)$	Recruitment standard deviation penalty
5	$logL(\pi_j, \theta \mid \tilde{\pi}_j, n) = log\Gamma(n+1) - \sum (log\Gamma(n\tilde{\pi}_j + 1)) + log\Gamma(\theta n) - log\Gamma(n + \theta n) + \sum (log\Gamma(n\tilde{\pi}_j + \theta n \pi_j) - log\Gamma(\theta n \pi_j))$	Dirichlet-multinomial log-likelihood
6	$n_{eff} = (1 + \theta n)/(1 + \theta)$	Effective sample size
7	$C_t^{obs} \sim Lognormal(\sum_{a=0}^{A} C_t, \sigma_C^2)$	Lognormal likelihood (catch)
8	$I_t^{obs} \sim Lognormal(qB_t, \sigma_I^2)$	Lognormal likelihood (abundance index)
9	$MRE = median\left(\frac{x_{estimated} - x_{true}}{x_{true}}\right)$	Median relative error to quantify bias
10	$MARE = median\left(\left\lvert\frac{x_{estimated} - x_{true}}{x_{true}}\right\rvert\right)$	Median absolute relative error to quantify precision

Table 2.5. LIME performance in estimation of SPR in the terminal year across scenarios of data availability and patterns in fishing mortality and recruitment variability, compared with the cross-tested LB-SPR. Bias (top) is measured as median relative error and precision (bottom) is measured as median absolute relative error from 100 iterations of generated data using the LIME age-structured operating model assuming instantaneous sampling with 200 length measurements annually. Dark green indicates bias/precision less than 5%, light green indicates bias/precision less than 10%, yellow indicates bias/precision less than 20%, light blue indicates bias/precision less than 30%, and dark blue indicates bias/precision greater than 30%.

Data availability scenario	Short			Medium			Longer		
Bias (MRE)	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
Rich	-0.016	-0.015	0.014	-0.005	0.009	0.018	-0.006	0.017	0.046
Index+LC10	-0.016	-0.024	0.018	-0.044	0.014	0.085	0.158	0.277	0.029
Index+LC1	0.021	0.006	0.041	0.002	0.079	0.046	0.050	0.024	0.045
Catch+LC10	-0.016	-0.028	0.015	-0.022	0.140	0.206	-0.009	0.101	0.279
Catch+LC1	0.027	0.821	0.972	-0.047	0.067	0.179	0.024	-0.147	0.222
LC10	-0.026	-0.051	0.025	-0.041	-0.056	-0.018	0.129	-0.008	0.281
LC1	-0.056	-0.082	-0.015	-0.115	-0.317	-0.126	0.107	-0.443	-0.109
LBSPR10	-0.570	-0.676	-0.708	0.006	-0.405	0.196	0.062	-0.359	0.406
LBSPR1	-0.603	-0.698	-0.738	-0.030	-0.369	-0.165	0.059	-0.403	0.271
Precision (MARE)									
Rich	0.023	0.047	0.031	0.039	0.055	0.102	0.033	0.042	0.098
Index+LC10	0.039	0.047	0.034	0.172	0.206	0.314	0.251	0.447	0.348
Index+LC1	0.058	0.082	0.068	0.097	0.208	0.247	0.162	0.210	0.250
Catch+LC10	0.032	0.054	0.035	0.081	0.423	0.409	0.075	0.426	0.598
Catch+LC1	0.066	0.821	0.972	0.127	0.374	0.514	0.134	0.382	0.646
LC10	0.038	0.067	0.037	0.245	0.510	0.486	0.380	0.788	0.757
LC1	0.098	0.128	0.033	0.215	0.441	0.374	0.304	0.720	0.625
LBSPR10	0.570	0.676	0.708	0.066	0.412	0.404	0.077	0.366	0.456
LBSPR1	0.603	0.698	0.738	0.113	0.471	0.432	0.120	0.501	0.5084

Table 2.6. LIME bias and precision in estimating SPR in the terminal year for a variety of sample size of independent length measurements. In all cases one year one year of length composition data are provided to the model. Bias (top) is measured as median relative error and precision (bottom) is measured as median absolute relative error from 100 iterations of generated data using the LIME age-structured operating model assuming instantaneous sampling. Dark green indicates bias/precision less than 5%, light green indicates bias/precision less than 10%, yellow indicates bias/precision less than 20%, light blue indicates bias/precision less than 30%, and dark blue indicates bias/precision greater than 30%.

Sample size	Short			Medium			Longer		
Bias (MRE)	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
1000	-0.033	-0.029	-0.012	0.059	-0.250	-0.108	-0.021	-0.311	-0.081
500	-0.029	-0.055	-0.019	0.011	-0.348	-0.116	0.038	-0.410	-0.253
200	-0.056	-0.082	-0.015	-0.115	-0.317	-0.126	0.107	-0.443	-0.109
100	-0.072	-0.083	-0.025	-0.038	-0.370	-0.185	-0.113	-0.430	-0.173
50	-0.087	-0.103	-0.038	-0.117	-0.444	-0.166	-0.061	-0.560	-0.084
20	-0.145	-0.150	-0.042	-0.294	-0.426	-0.201	-0.461	-0.721	-0.440
1000	0.077	0.109	0.031	0.112	0.478	0.450	0.249	0.743	0.629
500	0.062	0.119	0.032	0.126	0.442	0.325	0.258	0.748	0.661
200	0.098	0.128	0.033	0.215	0.441	0.374	0.304	0.720	0.625
100	0.141	0.149	0.047	0.206	0.475	0.434	0.376	0.700	0.552
50	0.116	0.156	0.049	0.323	0.503	0.469	0.426	0.671	0.615
20	0.227	0.169	0.048	0.401	0.462	0.567	0.540	0.742	0.617

Table 2.7. 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. 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 median relative error and precision (bottom) is measured as median absolute relative error from 100 iterations of generated data. Dark green indicates bias/precision less than 5%, light green indicates bias/precision less than 10%, yellow indicates bias/precision less than 20%, light blue indicates bias/precision less than 30%, and dark blue indicates bias/precision greater than 30%.

Data availability scenario	Short				Medium			Longer		
Bias (MRE)	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
LC10	-0.026	0.724	0.011	0.366	-0.041	0.037	-0.102	0.129	0.065	-0.106
LC1	-0.056	0.547	-0.064	0.366	-0.115	-0.103	-0.102	0.107	0.099	-0.105
LBSPR10	-0.570	0.142	NA	-0.033	0.006	0.086	-0.011	0.062	0.117	-0.007
LBSPR1	-0.603	0.125	NA	-0.033	-0.030	0.056	-0.011	0.059	0.114	-0.007
LC10	0.038	0.724	0.102	0.366	0.245	0.275	0.105	0.129	0.372	0.106
LC1	0.098	0.547	0.401	0.366	0.215	0.215	0.105	0.107	0.374	0.106
LBSPR10	0.570	0.142	NA	0.059	0.066	0.093	0.061	0.062	0.117	0.070
LBSPR1	0.603	0.159	NA	0.059	0.113	0.139	0.061	0.059	0.161	0.070

Table 2.8. 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 one year of length composition 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 one year of length composition data for short-lived species. May over-estimate SPR for medium and longer-lived species.
2.	*Is LIME sensitive to:*	
a.	Sample size of length measurements?	Yes, accuracy and precision generally improve with higher annual sample size. Scenarios with 100 samples and greater generally see a performance plateau.
b.	Error in input parameters?	Relatively insensitive to error in the CV of the age-length curve. Asymptotic length, von Bertalanffy growth coefficient, If asymptotic length, von Bertalanffy growth coefficient, or length at 50% maturity are input higher than the truth, or natural mortality input lower than the truth, SPR will be estimated lower 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 self-test, LB-SPR cross-test)	LIME unbiased across life history types; LB-SPR estimates lower SPR for short-lived life history.
b.	Sampling monthly from age-structured model (LIME and LB-SPR cross-test)	LIME should use monthly time steps for short-lived species, but can use annual time steps for medium and longer-lived species. LB-SPR slightly over-estimates SPR.
c.	Length-structured operating model (LIME cross-test, LB-SPR self-test)	LIME over-estimates SPR for short-lived and under-estimates SPR for medium- and longer-lived. LB-SPR unbiased across life history types.

2.6 FIGURES

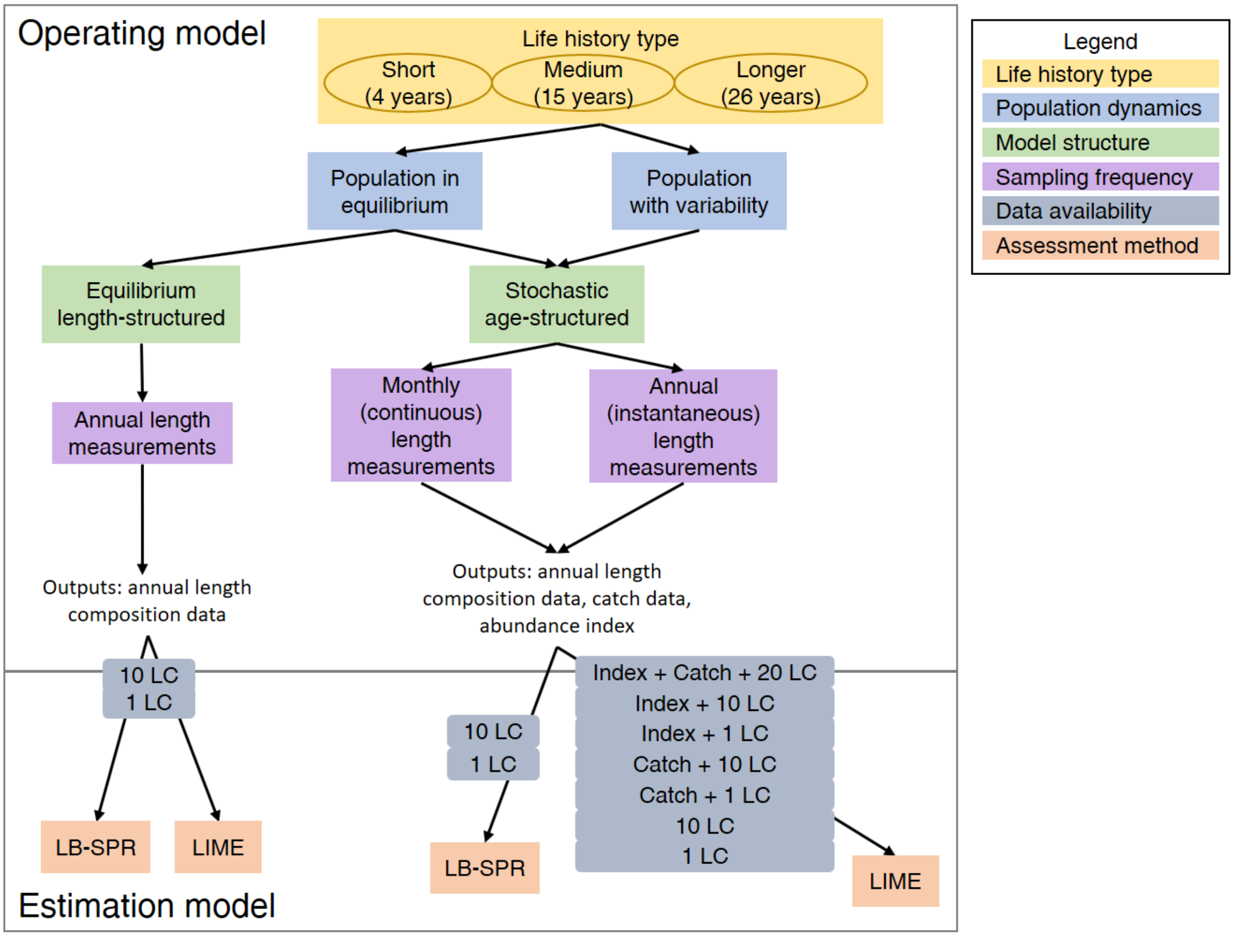

Figure 2.1. Diagram of the simulation study. "LC" stands for length composition data, numbers indicate number of years in the 20-year time-period modeled. In the case of the instantaneous, annual length measurement scenario, we compared multiple sample sizes of length measurements annually (1,000, 500, 200, 100, 50, and 20). For the other scenarios, we assumed the base case of 200 length measurements annually.

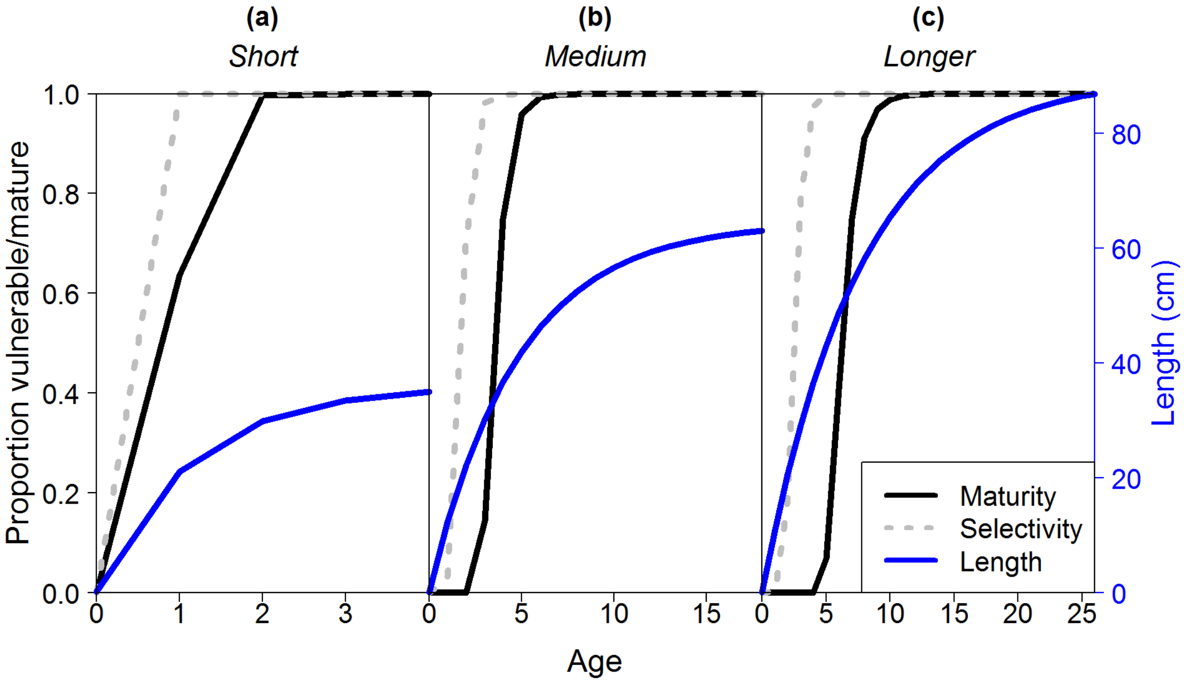

Figure 2.2. Selectivity, maturity, and length at age curves for the three life history types tested in the simulation study, mimicking: (a) rabbitfish (*Siganus sutor*), a short-lived fish (L_∞=36.2 cm, k=0.87, M=1.49, a_m^{50}=1 year, maximum age A=4 years, Hicks and McClanahan 2012), (b) spotted rose snapper (*Lutjanus guttatus*), a medium-lived fish (asymptotic length L_∞=64.6 cm, von Bertalanffy k=0.21, natural mortality M=0.43, length at 50% maturity a_m^{50}=4 years, maximum age A=15 years; Bystrom 2015), and (c) red grouper (*Epinephelus morio*), a longer-lived fish (L_∞=90 cm, k=0.13, M=0.18, a_m^{50}=7 years, maximum age A=26 years; Heemstra and Randall 1993).

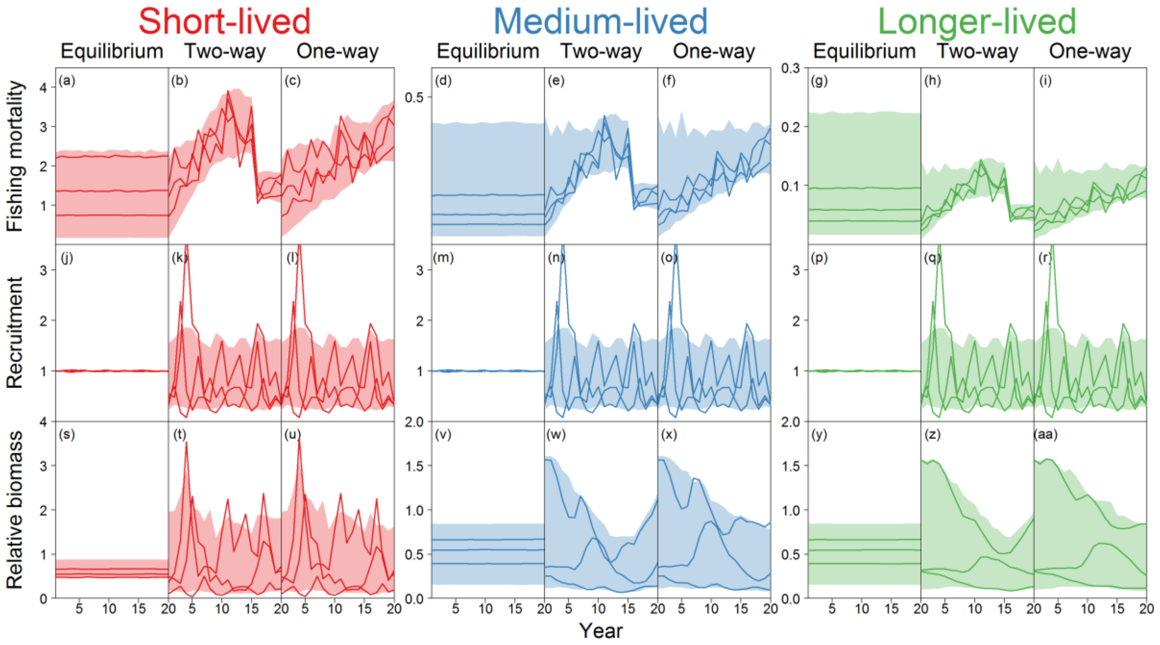

Figure 2.3. Scenarios of fishing mortality and recruitment under equilibrium and variable conditions for the three life history types, with corresponding trajectories of relative spawning biomass. Scenarios labeled "equilibrium" demonstrate that the initial depletion may start between 0.05 and 0.95, 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% SPR over the first seven years, stays at $F_{20\%}$ for seven years, and then drops to half of $F_{20\%}$ for the last six years. Scenarios labeled "one-way" demonstrate a change from equilibrium fishing mortality resulting in an initial depletion between 0.05 to 0.95 to the fishing mortality $F_{20\%}$ over the 20-year period.

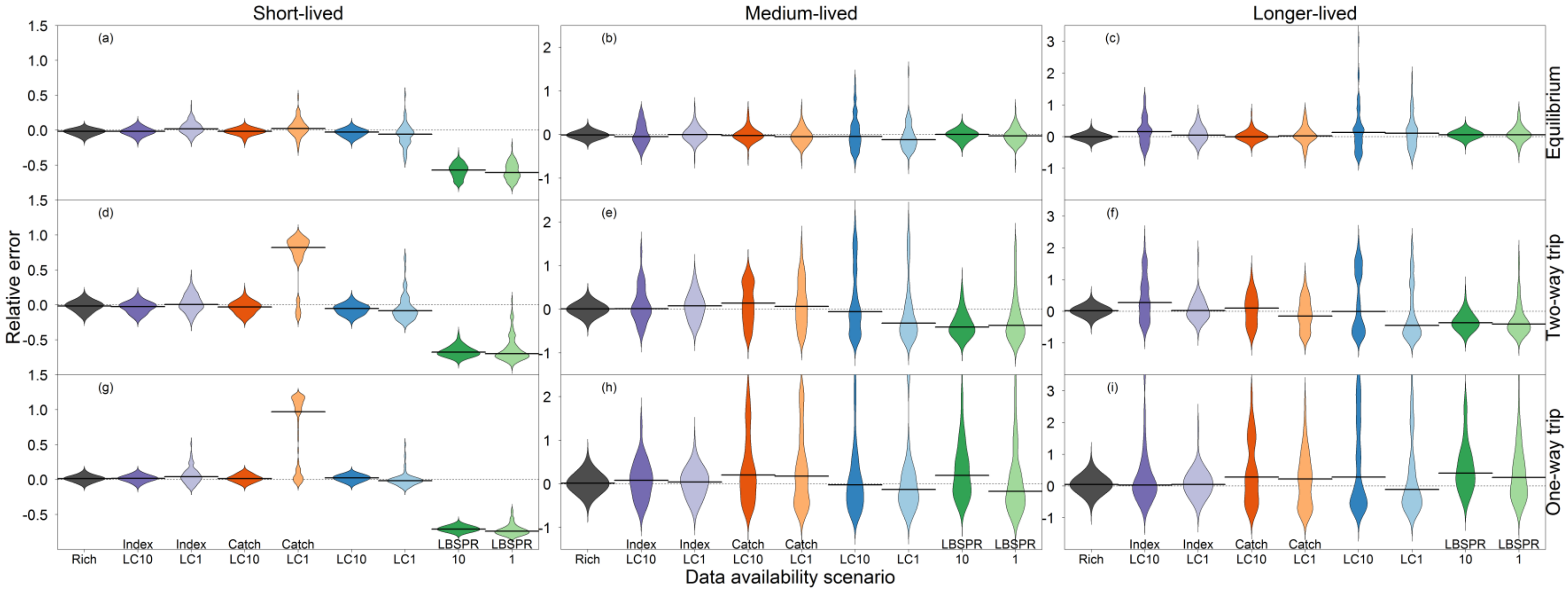

Figure 2.4. Distribution of relative error (*(estimated-true)/true*) for spawning potential ratio (SPR) in the current year for 100 iterations of simulated populations across the LIME and LB-SPR data availability scenarios for the three life history types and scenarios of equilibrium and variable fishing mortality and recruitment with 200 samples of length measurements annually. The gray bean represents the “data-rich” scenario, which verifies LIME is unbiased and most precise when an unrealistically high amount of data are available. Darker colors represent data availability scenarios with 10 years of length composition data, and lighter colors represent the scenario with one year of length composition data available. Each life history type has a different y-range.

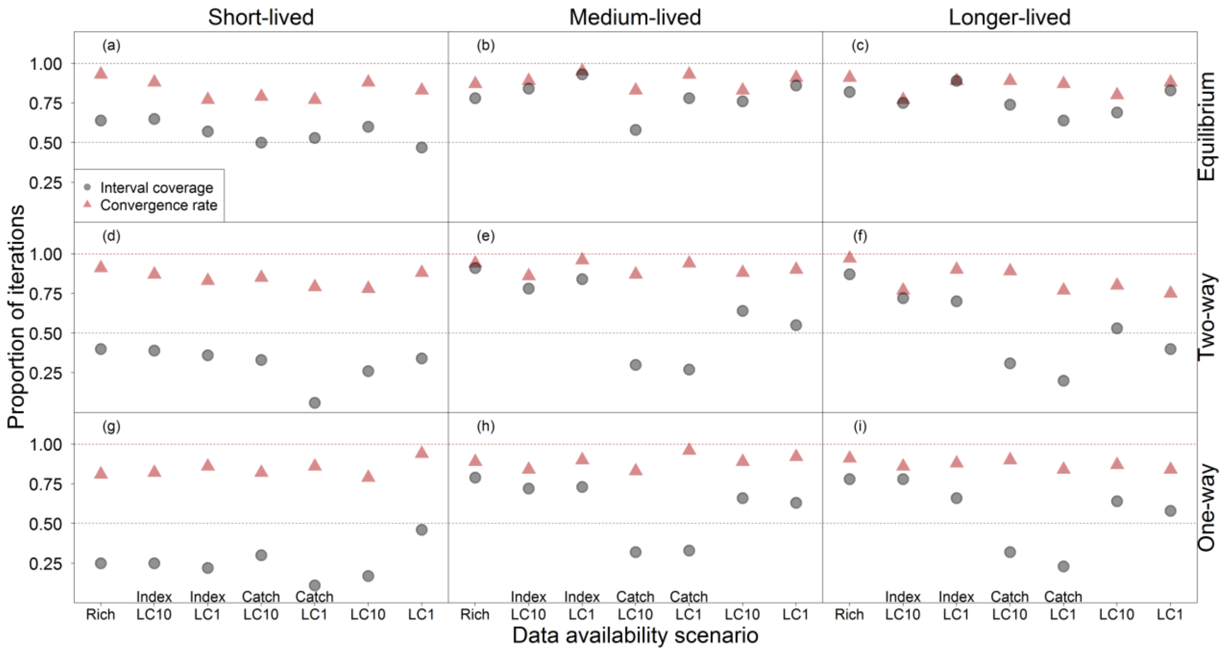

Figure 2.5. Proportion of iterations out of 100 where the true value of the spawning potential ratio (SPR) lies within the 50% confidence intervals for each life history type across various patterns of fishing mortality and recruitment (gray circles) compared to the convergence rates for the same scenario (pink triangles).

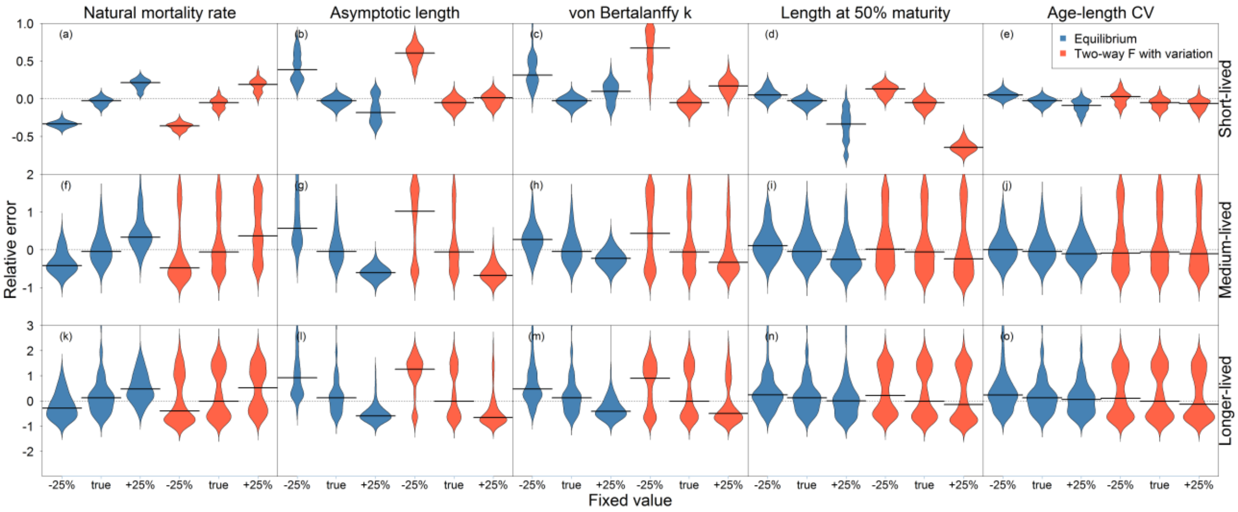

Figure 2.6. Comparison of the distribution of estimation error for spawning potential ratio (SPR) in the terminal year with the life history parameters fixed at +/- 25% of their true value from 100 iterations of generated data across the various life history scenarios for the equilibrium (blue) and the 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 composition data with 200 length measurements annually. Life history parameters include natural mortality, asymptotic length, von Bertalanffy growth coefficient, length at 50% maturity, and the coefficient of variation for the age-length curve.

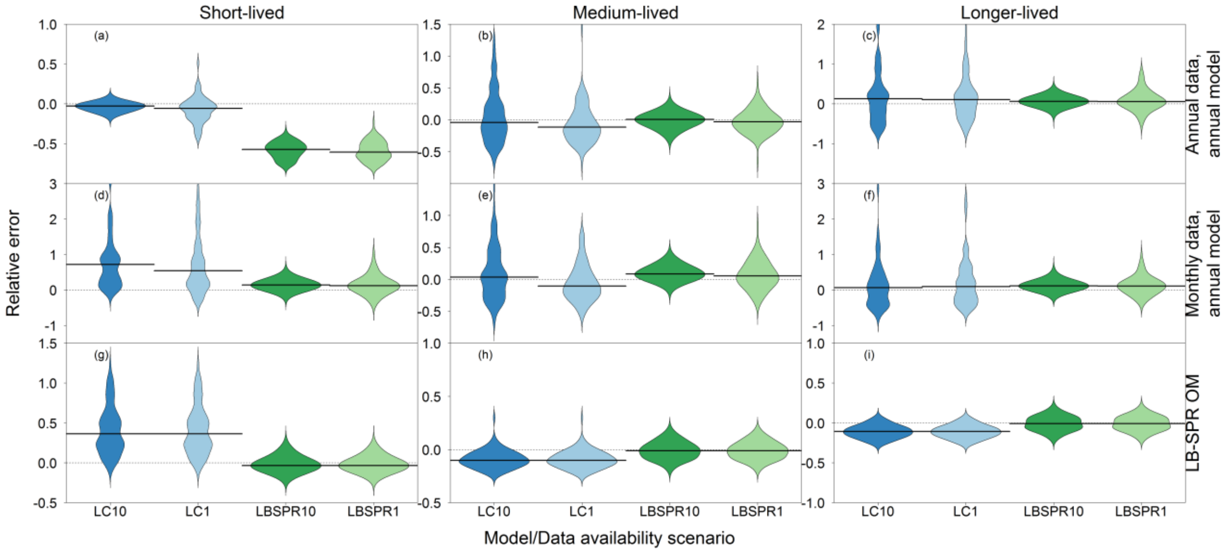

Figure 2.7. Comparison of the LIME and LB-SPR methods in the equilibrium state via the distribution of relative error *(estimated-true)/true* for spawning potential ratio (SPR) in the current year for 100 iterations of simulated populations from the LIME age-structured operating model using annual length 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. The methods are compared across life history types. Darker colors represent data availability scenarios with 10 years of length composition data, and lighter colors represent the scenario with one year of length composition data available. Each scenario has a different y-range.

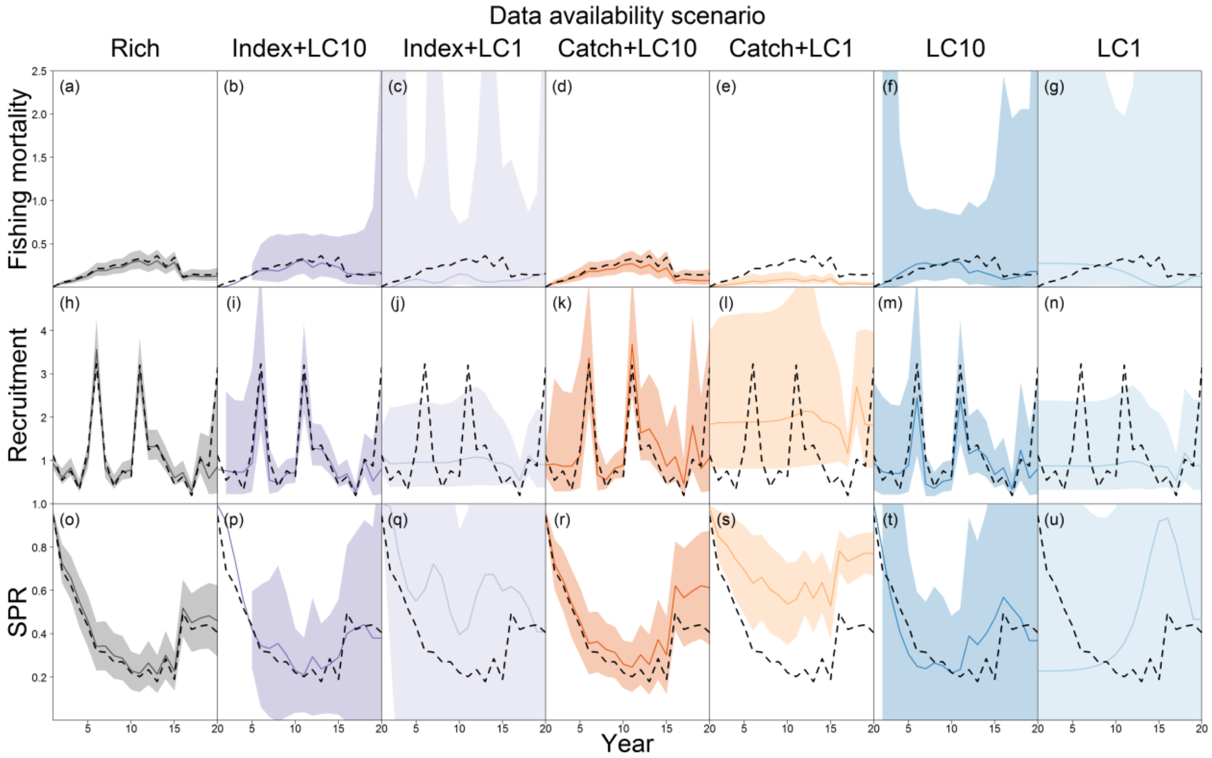

Figure 2.8. Example of model estimates (color lines) with 95% confidence intervals (shaded areas) and the true simulated population dynamics (black dotted line) for each of the data availability scenarios for one iteration of a simulated population with medium-lived life history, 200 length measurements annually in the length composition data, for the base variation scenario.

Chapter 3. APPLICATION OF LENGTH-BASED METHODS TO DATA-LIMITED FISHERIES

Abstract

Many data-limited stock assessments are oriented towards estimating fishing pressure or stock status relative to a biological reference point, but are often limited estimating population processes or quantifying uncertainty. Using length composition data information on growth, maturity, and natural mortality, we conducted stock assessments of two previously unassessed, small-scale fishery stocks: Costa Rican snapper, *Lutjanus guttatus*, a medium-lived fish with a maximum age of 13 years, and Kenyan rabbitfish, *Siganus sutor*, a short-lived fish with a maximum age of 4 years. We compared stock assessment results, including the spawning potential ratio (SPR) reference point, from two length-based methods that required approximately the same inputs, (i) the Length-based Integrated Mixed Effects (LIME) method that can account for variability in fishing mortality and recruitment and (ii) Length-Based Spawning Potential Ratio (LB-SPR) method, which assumes equilibrium conditions. LIME estimated the Costa Rican spotted rose snapper stock to be at the target SPR of 0.40 from 2009-2013, but dropping below the overfished level of 0.30 in 2014 and 2015. LB-SPR estimates of stock status were more variable across the time series, but on average indicated sustainable levels at the target SPR with the estimate in 2015 above the target. LIME estimated the rabbitfish stock to have undergone a period of overfishing in the late 1990s and early 2000s, but fishing mortality continuously declined since 2000 leading to a recovered SPR above the target 0.40 by 2014. LB-SPR estimates also indicated a period of overfished status followed by higher SPR on average, but estimated a fully exploited stock in 2014, not quite recovered due to the high interannual

variability in SPR estimates. These represent the first reference-point based stock assessments in each case, and we provide guidelines for length-based assessment application for future stocks in which only biological information and length composition data are available.

3.1 INTRODUCTION

Large-scale commercial fisheries in developed nations are commonly assessed due to economic incentives to keep those resources sustainable (Ricard et al. 2012), but assessment and management of small-scale fisheries are difficult due to the decentralized nature of the resource users, low research capacity, ecosystem uncertainties, and lack of management enforcement (Hicks and McClanahan 2012, Cinner et al. 2013, McClanahan et al. 2014). For these reasons, focusing limited funding and capacity on management in developing nations is more important than conducting stock assessments for small-scale fisheries with limited data (Mahon 1997). In turn, this has led to a greater emphasis on empirical harvest control rules for unassessed data-limited fisheries (Dowling et al. 2015). The benefits of small-scale fisheries are often more than just economic, including diversified livelihoods (Allison and Ellis 2001), food security (Bene et al. 2010, McClanahan et al. 2013), nutrition (Hall et al. 2013), cultural values, and well-being (Armitage et al. 2012). Lack of data and low research capacity leads to a greater emphasis on empirical harvest control rules based on proxy reference points for unassessed, small-scale fisheries (Mahon 1997, Dowling et al. 2015). However, single-species assessments are still valid goals for species that are in higher demand or are caught at a greater frequency. Stock assessments can estimate stock status relative to biological reference points, explore hypotheses on the processes that have led to the current population status, and estimate uncertainty to provide guidance for management decisions.

The most common biological reference points in fisheries stock assessments are the biomass and fishing mortality associated with maximum sustainable yield (MSY). However, some measure of population scale (typically either total catch or catchability for a survey index), is required to estimate MSY-based reference points, and it is difficult to record total annual catches for many small-scale or bycatch fisheries. Length measurements are much easier to reliably collect than total catch. Given assumptions about biological information, comparing length distribution in fishery catch with that expected in an unfished state yields information on fishing mortality, recruitment, and selectivity.

Spawning potential ratio (SPR) is a relative measure of stock status that can serve as a proxy for FMSY, and is defined as the ratio of potential lifetime egg production between the fished and unfished state (Goodyear 1993). SPR is increasingly used as a reference point in data-limited stock assessments where length composition data are available but total catch is unavailable or unreliable (Nadon et al. 2015, Prince et al. 2015b). SPR is useful as a metric of fishing intensity independent of stock steepness, where an SPR of x% arises from a population under constant virgin recruitment fished at a rate of F_x% (Cordue 2012). SPR has been used in length-based assessments in the U.S. South Atlantic (Ault et al. 2005), Gulf of Mexico (Cummings et al. 2016), Caribbean (Ault et al. 2008), and the Indo-Pacific (Prince et al. 2015b), as well as in assessments on data-rich age-structured assessments in conducted in Stock Synthesis (e.g. Cope and Key 2009, Methot and Wetzel 2013, Thorson and Wetzel 2015). For many life histories, an SPR of 40% is often used as target, while an SPR of 30% is a threshold under which a population would be considered overfished (Clark 2002, Nadon et al. 2015, Johnson et al. 2016, Stawitz et al. 2016). However, targets are determined somewhat arbitrarily by species (Mace and Sissenwine 1993), with long-lived rockfish typically having an SPR target of 50% (Mayo et al. 2007, Thorson and Wetzel 2015) and flatfish an SPR target of 30% (Hicks and Wetzel 2011, Stawitz et al. 2016). In place of estimating the fishing mortality rate associated with MSY (F_{MSY}), it is possible to calculate the fishing mortality rate that would result in an SPR of 40% ($F_{40\%}$) or 30% ($F_{30\%}$).

The length-based spawning potential ratio (LB-SPR) model has been used in data-limited stock assessments of smaller-scale fisheries in developing nations (Prince et al. 2015b). LB-SPR requires one year of length composition data and estimates of the maturity ogive, asymptotic length, and the Beverton-Holt life history invariant M/k, which is the ratio of natural mortality and the von Bertalanffy growth parameter (Beverton and Holt 1957, Jensen 1996, 1997, Hordyk et al. 2015, Prince et al. 2015a). Instead of assuming M and k are known without error, LB-SPR assumes that the M/k ratio is known without error. LB-SPR has been used to quickly conduct assessments of previously unassessed, data-limited stocks (Prince et al. 2015b). However, LB-SPR assumes the population is in equilibrium, i.e. that fishing mortality and recruitment are not changing on a time scale relevant for fish life history or fisheries management. When more than one year of length composition data are available, LB-SPR is run independently each year, resulting in separate annual estimates of SPR, selectivity parameters, and the ratio of fishing mortality to natural mortality (F/M). While these predictions usually fit the observed data well, estimates of selectivity curves, fishing mortality, and SPR may vary unrealistically between years. Thus the equilibrium assumptions underlying LB-SPR may result in time series of estimates that do not tell a cohesive story of fishery status over time. There is an option to smooth estimates of F/M, selectivity, and SPR using a Kalman filter so that these parameters do not vary unrealistically between years. This in turn would affect the model fits to the length composition data, causing a mismatch between observed data and predicted population processes.

An alternative to LB-SPR is the length-based integrated mixed effects (LIME) method. LIME relaxes the assumption of equilibrium conditions and estimates recruitment variability and annual fishing mortality in an integrated framework (Chapter 2; Rudd and Thorson, in review). The minimum requirements for LIME are the same as for LB-SPR, but LIME uses a full age-structured model requiring assumptions about both M and k instead of just the ratio between the two. While these assumptions require more information than LB-SPR, LIME can allow for greater flexibility in modeling age- or length-structured processes. LIME estimates recruitment variation, length at 50% selectivity, the effective sample size of length composition data, and annual fishing mortality, and derives the recruitment deviations by treating annual recruitment as a random effect. If length composition data are available for more than one year, they are all included in the same model to estimate a single selectivity curve for all years and fishing mortality that can vary among years (with a penalty to restrict inter-year variability). LIME can also include time series of total catch and abundance indices, if available.

Simulation testing comparing the performance of the LIME and LB-SPR methods in estimating SPR found that LIME performs better for short-lived species across a range of life histories, fishing mortality, and recruitment scenarios (Chapter 2). The simulations also suggested that both LIME and LB-SPR will estimate SPR without bias for medium-lived species (maximum age 15) on an annual time-step, and that LIME will better track changes in fishing mortality and recruitment than LB-SPR when the population is not in equilibrium. However, when data are collected continuously throughout the year, and the species has a shorter lifespan, LIME will perform worse (Chapter 2). Notably, when the length data are collected on a monthly time step but pooled for an annual time-step in the model, LIME will over-estimate SPR while LB-SPR will be unbiased. To estimate unbiased SPR when monthly length composition data are available, LIME should be run on a monthly time-step (Chapter 2).

The objective of this study is to estimate stock status relative to the SPR reference point for medium-lived spotted rose snapper (*Lutjanus guttatus*, maximum age 13 years) from the Nicoya Peninsula in Costa Rica, and short-lived rabbitfish (*Siganus sutor*, maximum age 4 years) from the southern coast of Kenya. We used the length composition of the catch and assumptions about growth, natural mortality, and maturity to compare estimates of SPR coming from LIME and LB-SPR. This work is the first formal, stock-status-based assessment of both species, and the first comparison of results from LB-SPR and LIME using real-world stock assessment data.

3.2 MATERIALS AND METHODS

3.2.1 *LIME overview*

Inputs to LIME include biological parameters: natural mortality rate, von Bertalanffy growth parameters, allometric length-weight parameters, and maturity-at-age ogive; and length composition data. Values for biological parameters should come from local studies. If local studies are not available, distributions of plausible values in FishBase (Froese and Pauly 2017) can be used from the same species in a different region, or other species in the same genus. For natural mortality, the Barefoot Ecologist’s Toolbox hosts a tool (Cope 2017) which calculates a distribution of natural mortality estimates from a variety of methods given the information available, such as von Bertalanffy growth parameters, temperature, and the gonadosomatic index (e.g. Jensen 1996, Gislason et al. 2010, Then et al. 2015).

LIME uses a dynamic age-structured model and assumes the biological input parameters are known without error, length-at-age is normally distributed, natural mortality rates are constant over time, and growth rates are constant across cohorts. LIME requires at least one year of length composition data, which is assumed to be collected instantaneously within the time step, but does not require length composition data in every time step. LIME estimates annual fishing mortality, length at 50% selectivity, and the standard deviation of recruitment, all in log space (Chapter 2). In addition, LIME estimates a Dirichlet-multinomial θ parameter related to effective sample size to fit the predicted length composition to the observed using a Dirichlet-multinomial likelihood function (Thorson et al. 2016). When using a monthly time step, LIME estimates annual fishing mortality and recruitment deviations, but distributes both evenly throughout each month by dividing the estimated annual value by the number of seasons (e.g. 12 months). As an alternative to the built-in two-parameter logistic selectivity model, it is possible to fix the selectivity curve directly. This allows for the exploration of an alternate structure for selectivity than the two-parameter logistic curve included in LIME (e.g. a longer-sloping logistic curve or dome-shaped selectivity). It is also possible to specify the length-at-age curve directly. A case where this would be useful is sequential hermaphrodites, which have different growth curves before and after they change sexes but may be subject to one fishery.

LIME assumes annual recruitment is drawn from a distribution with an expected equilibrium value (μ) and standard deviation (σ_R). Because neither dataset contains information on scale to estimate total population size, we ran the models assuming μ fixed at 1.0. The resulting recruitment deviations represent relative recruitment compared to an equilibrium, unfished state. We estimate σ_R penalized by a prior with mean 0.737 and standard deviation of

0.353, with these values obtained from a meta-analysis on the variability of recruitment for exploited marine fish (Thorson et al. 2014b).

Populations are modeled over the inclusive years or shorter time steps where length composition data are available. LIME calculates the probability of being in a length bin given age from the length-at-age curve and assumed coefficient of variation (CV) around the curve (Chapter 2). We assumed a CV of 0.1, as this value is typically assumed to be about 0.1 in data-rich assessments (Johnson et al. 2016), and the analogous CV for asymptotic length is assumed to be 0.1 in LB-SPR assessments (Prince et al. 2015b). To obtain the predicted length composition, LIME multiplies the probability of being in a length bin given age by the probability of being harvested at an age from the derived abundance at age and estimated selectivity-at-age. LIME assumes length composition is sampled instantaneously at a single point in the time step, and does not account for individual growth during the time step (Figure 3.1; Chapter 2).

The joint likelihood of the observed data is the product of a Dirichlet-multinomial distribution fit to length composition data, a random walk penalty on fishing mortality, and a penalty on the recruitment deviations (Chapter 2). Key outputs from LIME include: the estimated annual fishing mortality rates, estimated selectivity curve, recruitment standard deviation, and derived recruitment deviations, SPR, the fishing mortality rate that would produce 30% SPR ($F_{30\%}$; the overfishing limit), and $F_{40\%}$ (the target fishing mortality).

LIME uses the R package Template Model Builder (TMB), which minimizes the negative log-likelihood, calculates exact derivatives with respect to all parameters, and fits mixed effects models using the Laplace approximation to approximate the marginal likelihood of fixed effects while integrating over random effects (Kristensen et al. 2016). We determined model convergence if the absolute value of the final gradient for each parameter is less than 0.01. After minimizing the negative log likelihood, LIME has the option to run Newton's method as an extra assurance of model convergence at the global minimum of the likelihood space. While we did not run the extra Newton steps for simulation testing, we ran three extra Newton steps during LIME application to real stocks after experiencing some cases with complex likelihood surfaces. If the convergence requirements are not met after the first run and the specified number of Newton steps, LIME runs up to ten more times with starting values equal to the previous estimates from the non-converged model plus a random number drawn from a normal distribution with mean zero and standard deviation 0.2. TMB uses the delta method to calculate standard errors for all estimated parameters and derived values of interest (e.g. SPR). Like AD Model Builder (Fournier et al. 2012), automatic differentiation in TMB evaluates the Hessian matrix to approximate the variance-covariance matrix for the estimated parameters (Magnusson et al. 2013). The delta method estimates the variance of each derived quantity assuming both estimation bias and the higher-than-quadratic terms of the Taylor series are negligible (Seber 1973).

3.2.2 *LB-SPR overview*

Inputs to LB-SPR include the ratio of natural mortality to the von Bertalanffy growth coefficient (*M/k),* asymptotic length in the von Bertalanffy growth equation, the coefficient of variability in asymptotic length, the size at maturity schedule, and length composition data. LB-SPR uses maximum likelihood methods to estimate the length at 50% and 95% selectivity and the relative fishing mortality (*F/M*), using a multinomial likelihood to fit the predicted and observed length composition data. These values are then used to calculate SPR. LB-SPR uses a static, equilibrium-based relative age-structured model, where “age” is in arbitrary units ranging from 0, associated with length 0, to 1, associated with asymptotic length. All processes typically included in an age-structured model, such as selectivity, growth, and predicted length composition, are standardized to asymptotic size. The *M/k* ratio can be determined from meta-analysis, life history theory, expert judgment, or biological studies (Hordyk et al. 2015). LB-SPR estimates parameters for each year independently. We obtained estimates of fishing mortality *F* by multiplying the estimated *F/M* ratio by the assumed natural mortality rate used in the base model for each species. Variances are calculated for each parameter using the Hessian matrix and delta method, as calculated using the *optim* function in *R.*

3.2.3 *Assessment of Costa Rican snapper*

Data collection

Consultants and interns from the Sea Turtle Restoration Program (PRETOMA) in Costa Rica collected fishery-dependent data on local fishing boats during nightly trips to fishing sites from 2007-2015, for a range of species including snapper. Bottom longline is the most commonly used fishing method for snapper (Bystrom 2015). Gillnet is the second most common, most often used when catches are low (A. Bystrom, personal communication). Experimental longline collected a small percentage of length samples in the dataset. Observations for each longline set included: duration, number of hooks, number of fish caught by species, and length and weight measurements. PRETOMA scientists also collected length and weight measurements from dockside landings. These observations were made year-round, although there were several months when data were not collected or fishing activity did not occur. The median number of sampling days per month was 5, with a maximum of 27 days. Sampling is representative of the

Gulf of Nicoya, although it is possible snapper extends beyond the range of the sampled fishery, and thus an implicit assumption of this assessment is that the whole population is experiencing similar population dynamics and fishing pressure.

Life history

Bystrom (2015) estimated length-weight and von Bertalanffy parameters from the PRETOMA database for Costa Rican snapper. Bystrom (2015) estimated von Bertalanffy growth parameters from length frequency data on a monthly time-step from July 2007 through June 2008 using ELEFAN I (Pauly 1987) in FiSAT II (Fisheries and aquaculture software 2017). We used estimates of length at 50% maturity from Rojas (1996) study, as in Bystrom (2015) (Table 3.1).

To estimate mortality rates, Bystrom (2015) used a length-converted-to-age catch curve based on the estimated von Bertalanffy growth parameters and the FiSAT II program to estimate that $Z = 0.77$. Bystrom (2015) calculated natural mortality M of 0.43 using the Pauly (1980) empirical relationship relating M to mean temperature 19°C, from temperature measurements on the fishing grounds. By definition, $F = Z - M$, therefore $F = 0.34$. However, we used a base value for M of 0.38, which is the median of estimates from multiple methods in the natural mortality tool (range 0.16 to 0.48, Cope 2017). While lower than the value of 0.43 used in Bystrom (2015), we believe this value better accounts for the other hypotheses for determining natural mortality.

We ran LIME on a quarterly time step for Costa Rican snapper instead of an annual time step. Simulation testing demonstrated that with the medium-lived nature of snapper, LIME can adequately model growth during the year even though it assumes the length composition data are collected instantaneously during the annual time step (Chapter 2), because for all lengths fully selected to the gear, there is reasonable probability they could be assigned to an age given the CV of the growth curve (Figure 3.1a). For example, if a fish of length 10 cm is harvested and in the length composition data, LIME on an annual time step assumes a very low likelihood of observing a 10-cm fish in the length composition due to the gap between lengths of age-1 fish and age-2 fish in the growth curve during one year. This would lead to LIME fitting the model poorly, unable to account for that 10-cm fish. However, on a quarterly time step, the growth curve predicts a 10-cm fish to be between about 5 and 8 "quarters" in age, which is somewhere between ages 1 and 2 (Figure 3.1b). The shorter the time step, however, the more uncertainty there is in the estimated age of a specific length, resulting in greater difficulty in teasing apart recruitment, fishing mortality, and selectivity (Figure 3.1c).

Length composition data

Length was recorded for 58% of all captured fish. To obtain the quarterly length composition data to use in the base model, we weighted the proportion of catch in each 1 cm length bin for each of the two main gears (bottom longline, 88%, and gillnet, 11%) by total observed catch numbers each quarter by gear (Figure 3.2). We obtained the proportion of samples in each length bin *j* each quarter *y* by gear g ($p_{g,j,y}$) from the annual length bin proportions ($p_{j,y}$), and number of fish caught in that quarter for each gear ($n_{g,y}$) as follows:

$$w_{j,y} = \sum_{g=1}^{G} p_{j,y} n_{g,y} \qquad (3.1)$$

Weighting the length composition data by gear accounts for the differing probability of being harvested in each length bin for each gear given selectivity differences, while allowing LIME to estimate only a single selectivity curve to represent the fisheries across gears and years (Figure 3.2). Another possible approach would be to consider separate fleets for each fishing gear, each with a separate selectivity curve. This approach would require estimation of length at 50% and 95% selectivity for each fishing gear. Total selectivity on the stock would be an additive process. For example, if 20% of 30-cm fish are vulnerable to a gillnet gear and 40% are vulnerable to bottom longline, the 60% of 30-cm fish would be vulnerable to the fishing mortality on the stock. This approach would allow the bottom longline selectivity to apply to the years where gillnet fishing was minimal, but adjust the total selectivity acting on the population in the years when both bottom longline and gillnet fishing occurred.

We assumed fishing selectivity is asymptotic, rather than dome-shaped, as assumed in previous. Recent analysis of biological characteristics of the fishery (Bystrom 2015) reported the length or age at first selectivity with no discussion on the shape of the selectivity curve. Dome-shaped selectivity could occur if larger fish move off rocky outcroppings targeted by fishing and out to sandy bottom areas as they age, but there is no evidence for this ontogenetic shift and no information on stock structure (A. Bystrom, *personal communication*). It could also occur if there is spatial variation in fishing pressure. For example, high population connectivity to an area not heavily fished would also result in dome-shaped selectivity. Bottom longline is preferred by fishermen because it captures a wider range of fish lengths, from smaller individuals to the largest known individuals, while gillnets are used when fishing times are more difficult, and catch only larger individuals. Generally, most snapper are caught by longlines, except in 2014 when more were caught by gillnets (Figure 3.2). Future information or data collection may result in a change in base model from asymptotic to dome-shaped selectivity.

3.2.4 *Assessment of Kenyan rabbitfish*

Data collection

Data used were collected by the Wildlife Conservation Society from 1996 to 2014 as part of their sampling of coral reef fish landings along the Kenyan coast. The dataset includes length measurements sampled from fishery landings for 262 species in 66 families, harvested by 24 different gears across 23 sites under 10 different types of management. Each observation is tied to a specific boat and the number of fishermen on the boat. The maximum number of days sampled in a year was 38 in 2007, while sampling was conducted on just one or two days in 1996, 1997, 2010, and 2011. The median number of days sampled in a year was 15. The 19-year database spans 228 months, with at least one day of sampling conducted in 76 of those months (33%) (Figure 3.3). The top priority species for assessment are the top three species in the dataset: rabbitfish (*Siganus sutor*), representing 20.7% of the recorded length measurements in the database, marbled parrotfish (*Leptoscarus vaigiensis*), 15.2% of length measurements, and emporers (*Lethrinus lentjan*), 6.7% of length measurements.

Life history

Information for LIME on growth, maturity, and starting values for selectivity parameters came from local studies (Ntiba and Jaccarini 1988, 1990), and were used for the previous equilibrium-assuming catch curve analysis (Hicks and McClanahan 2012). However, the natural mortality rate used in Hicks and McClanahan (2012) came from a meta-analysis from FishBase and not local studies, so for consistency with the Costa Rican snapper assessment we used the median of the distribution from the natural mortality tool (Cope 2017) for an M of 1.36 (range 0.53 to 1.88).

Due to the short-lived nature of rabbitfish (maximum age 4 years), LIME is unlikely to model the length composition accurately on an annual time step, and therefore a monthly time step was used. The coefficient of variation in the age-length curve allows for some uncertainty in the age-length distribution, but is not enough to overcome the instantaneous assumption for short-lived fish that grow very quickly over the course of one year (Chapter 2). Sampling length measurements during the year results in collecting fish that have grown from the average length at age-1, but are not quite at the average length for age-2 (Figure 3.1).

Length composition

We pooled the daily length measurements into 1-cm bins for each month. Stock assessment results are sensitive to the selectivity-at-length relationship, and each gear has a different selectivity curve. However, insufficient data exist to estimate this many parameters, and therefore data-limited, length-based stock assessment methods all assume one selectivity curve acting on the population. To correct for catches from different gear, for the base model we therefore weighted the length composition data by the number of individuals sampled from each gear (Figure 3.3), as for Costa Rican snapper (eq. 3.1).

3.2.5 *Base models for both stocks*

For the base LIME models for Costa Rican snapper and Kenyan rabbitfish, we input the growth parameters, length-weight parameters, natural mortality rate, maturity ogive (Table 3.1) and the weighted length composition data in counts, and estimated length at 50% selectivity, annual fishing mortality, recruitment standard deviation, and the θ Dirichlet-multinomial parameter. Derived outputs were mean length in the catch and relative depletion from equilibrium spawning biomass. For LIME, length composition data should be in counts, rather than proportions, as this provides information to estimate the θ parameter that is related to the effective sample size of the length data. As in the previous example, we compared the LIME estimates of SPR and selectivity with maximum likelihood estimates from LB-SPR using the R package *LBSPR* (https://github.com/adrianhordyk/LBSPR).

3.2.6 *Sensitivities*

We conducted several sensitivity tests to determine how results change with altered input data. In the first sensitivity, instead of using weighted length composition data, we ran the model using only data from the most commonly used gear: bottom longline for Costa Rican snapper (88% of length samples), and basket trap for Kenyan rabbitfish (47% of length samples). We also ran the models using data from the months with sampling above the median number of days per month, which included quarters with more than 14 sampling days for Costa Rica, and months with greater than two days of sampling for Kenya.

We ran the base model with alternate assumptions about biological input parameters. We explicitly compared LIME and LB-SPR model estimates with L_∞, k, and M fixed at 25% lower and higher than the base model value. We ran negative log likelihood (NLL) profiles for LIME over a distribution of possible values for the natural mortality rate M, asymptotic length L_∞, von Bertalanffy growth parameter k, and recruitment standard deviation σ_R. The range of values for the natural mortality rate spanned the minimum and maximum of estimates from the natural mortality tool, the range of values for the growth parameters spanned the minimum and maximum of estimates for the respective genus from FishBase, and σ_R was profiled between 0.05 and 0.95. The goal of the NLL profile was to determine how sensitive results are to misspecifications of these values in the base model, and to highlight the importance of integrating across uncertainty in these parameters for future LIME applications. In addition, we ran sensitivity tests comparing LIME and LB-SPR model estimates with L_∞, *k*, and *M* fixed at 25% lower and higher values than in the base models.

An additional sensitivity test in LIME examined the impact of changing logistic selectivity curve to dome-shaped selectivity. We parameterized the left side of the dome-shaped selectivity curve as the two-parameter logistic curve from the base model, with length at 50% and 95% selectivity fixed to the estimates from the base model. We added descending right-hand selectivity as a normal distribution from the first length at 100% selectivity from the base model down to where 10% at maximum length.

3.3 RESULTS

3.3.1 *Costa Rican snapper*

Base model

The base LIME model estimated that SPR in 2015 was 0.24 (95% interval 0.14-0.34), and fishing mortality was 0.86 (Table 3.2, Figure 3.4). 87% of the confidence interval for SPR was below the 0.30 threshold that would categorize the stock as overfished and greater than 99% of the confidence interval was below the SPR target of 0.40. The model estimated the SPR was stable above the overfished threshold, between 0.31 and 0.42 before 2011, but decreased to the most recent estimate of 0.24 as fishing mortality increased from 0.39 in 2012, past the reference point $F_{40\%}$ of 0.42 between 2012 and 2013, to 0.86 in 2015 (Table 3.2, Figure 3.4). Estimated recruitment variability in the base LIME model was 0.64. LIME estimated a low recruitment year in 2010, approximately 16% of equilibrium recruitment, before increasing back to nearequilibrium recruitment the rest of the time series (Figure 3.4). The estimated length at 50% selectivity was 36.3 cm and length at 95% selectivity was 43.6 cm. This implies that the estimated length at 50% selectivity is greater than the assumed length at 50% maturity of 34.0 cm (Rojas 1996). LIME estimated constant mean length throughout the time series until 2013. We would expect the estimated decline in fishing mortality in in 2009 and low recruitment year in 2010 to lead to an increase in mean length, which LIME estimated from 2011 to 2013, a time lag associated with the length at selectivity. LIME estimated a decrease in mean length from 2013 through 2015 as the fishing mortality was estimated to increase (Figure 3.4). Similar patterns occur in the spawning biomass relative to the first year in the time series ("first-year relative spawning biomass"). LIME estimated increasing first-year relative spawning biomass after the increases in recruitment in 2011 and 2012 after the period of low recruitment in 2010, but then decreased first-year relative spawning biomass as fishing mortality ramped up through 2015 (Figure 3.4). Patterns in first-year relative spawning biomass and SPR are different because SPR is based on per-recruit estimates and fishing mortality each year, whereas relative spawning biomass also accounts for the recruitment deviations.

In contrast to LIME, LB-SPR estimated that Costa Rican snapper was not overfished in 2015, with SPR of 0.49 in 2015 (95% interval 0.46-0.52; Table 3.2, Figure 3.4), and less than 1% of the confidence interval for SPR was below the 0.40 SPR target in 2015. LB-SPR estimated fluctuation in SPR between an SPR just above the overfished threshold and above the target threshold between 2012 and 2015. LB-SPR estimated length at 50% selectivity to vary between 27.1 cm in 2010 and 57 cm in 2014 (Table 3.2). The outlying length at 50% selectivity in 2014 (Figure 3.4) is associated with a marked change in the shape of the length composition data (Figure 3.2). The slope of the selectivity curve represented by the ratio of length at 95%

selectivity to 50% selectivity varies from a low of 1.13 in 2010 to a high of 1.27 in 2012. In 2014, when the shape of the length composition data was much different than other years, LB-SPR attributed the change to a shift in selectivity (Table 3.2). The associated *F/M* ratio was much higher than other estimates, leading to an unrealistically high estimate of *F* when multiplied by the base input *M*. When using the smoother option to prevent unrealistic variation between years of data, estimates of SPR were stable between 0.45 and 0.48 throughout the nine-year period, but estimates of *F/M* were much higher than the independently estimated values due to the influence of the high *F/M* estimate in 2012 (Figure 3.4).

Sensitivities

Across sensitivity tests, LIME generally estimated SPR for the Costa Rican snapper stock less than the 0.30 SPR overfished threshold (Figure 3.5), while LB-SPR generally estimated the snapper stock above the 0.40 SPR target. There were exceptions to these overall results which demonstrate the importance of conducting sensitivity analyses for making management decisions. LIME estimated the population was not overfished when we assumed *M* was 25% greater than the base model input value or L_∞ was 25% less than the base model. LB-SPR estimated the snapper stock above the 0.40 SPR target for all sensitivities except when considering only the bottom longline gear, or when assuming a lower *M/k* ratio where LB-SPR estimated the stock at or below the 0.30 overfished threshold (Table 3.3, Figure 3.5).

When assuming *M* was 25% higher than the base model, LIME estimated SPR to be 0.34, still less than the target SPR of 0.40 but above the SPR threshold 0.30 (Table 3.3). The likelihood profile across estimates of *M* from the natural mortality tool is flat (Figure 3.6e), meaning different values of *M* do not affect the fit to the length composition data, but the estimated level of fishing mortality would be higher if natural mortality was lower to account for the total mortality expected from the length composition data. Attributing more of the mortality to fishing mortality leads to lower estimates of SPR. Adjusting M to be lower or higher than the value used in the base model had more profound impacts on the LB-SPR assessment. If M was assumed to be 25% lower than the base model, this changed the M/k ratio from 1.79 to 1.34, resulting in an estimated SPR of 0.27, down from 0.49 in the base model (Table 3.3). The likelihood profiles over the von Bertalanffy asymptotic length L_∞ based on the distribution for *Lutjanus* species on FishBase had more of a parabolic shape than natural mortality, implying the model would fit the data better dependent on the assumed value for L_∞ (Figure 3.6f). When L_∞ was 25% less than the base model, LIME estimated the snapper stock to be close to be in an unfished condition (Table 3.3). However, this sensitivity is unlikely given we know the fishery is exploiting the snapper stock, with the high negative log likelihood relative to other models as statistical evidence. The likelihood profile over the von Bertalanffy growth coefficient k had a bimodal pattern (Figure 3.6g). The higher value of k associated with the likelihood minimum that was not used in the base case was associated with higher estimates of fishing mortality and lower estimates of SPR, indicating a more depleted population than the base model suggests.

When considering length composition data from the bottom longline gear only, LIME and LB-SPR both estimated the Costa Rican snapper stock to be in an overfished state. LIME estimated a greater increase in fishing mortality from 2013 to 2015 than estimated in the base model (Figure 3.5d), resulting in estimates of a more overfished population (Figure 3.5b,Table 3.3). LIME estimated the population to be below target levels with SPR at 0.18 (95% interval 0.10-0.26), with 99% of the interval being below 0.30 SPR and 100% below 0.40 SPR (Table 3.2). Using only data from the bottom longline gear, LB-SPR estimated the population to be in an overfished state in 2015, as opposed to above target levels using the base model (Table 3.3, Figure 3.5c). LB-SPR estimated SPR at 0.23 in 2015 (95% CI 0.21 to 0.25). Selectivity estimates did not vary as widely between years as they did with the weighted length composition data between bottom longline and gillnet gears (Figure 3.5g). This lack of strong changes in selectivity led to all estimates of *F* falling in a reasonable range, less than four times the natural mortality rate (Figure 3.5e).

Considering high sampling periods only, assuming they are more representative of the fishery, LIME estimated slightly lower fishing mortality and recruitment and higher SPR than the base model in 2015, although still estimated the stock in an overfished state. Using only high sampling periods does not affect LB-SPR estimates, since LB-SPR runs separately for each year of data.

Dome-shaped selectivity did not make much of a difference from logistic selectivity for the Costa Rican snapper. Dome-shaped selectivity resulted in lower estimates of fishing mortality and higher estimates of SPR. Fishing mortality estimates for 2015 decreased from 0.86 in the base model to 0.83 with dome-shaped selectivity, and SPR estimates increased from 0.24 in the base model to 0.25 with dome-shaped selectivity (Table 3.3).

3.3.2 *Kenyan rabbitfish*

Base model

The base model for LIME estimated SPR in 2014 to have been 0.59, but with high uncertainty (95% CI 0.21-0.98). LIME estimated 7% of the SPR confidence interval below the overfished threshold SPR 0.30 for the rabbitfish stock, and 16% of the confidence interval below the SPR target 0.40 (Table 3.4, Figure 3.7). LIME estimated the rabbitfish stock was overfished in the late 1990s and early 2000s, with fishing mortality greater than the $F_{30\%}$ reference point of 1.06 (Figure 3.7). Fishing mortality reached a peak in 2000 with a rate of 1.87, but since 2000 was estimated to be declining down to the estimate of 0.41 in 2014. SPR increased above the target SPR 0.40 by 2008 reached was recovered at 0.59 SPR estimated in 2014. LIME estimated the standard deviation for recruitment to be 0.72 (Table 3.4). This estimate is close to the median recruitment standard deviation for perciforms of 0.74 (Thorson et al. 2014b). LIME estimated a period of relatively high recruitment between 1998-2000, in the middle of the overfished period. Recruitment estimates were relatively normally distributed around the equilibrium throughout the rest of the time series. LIME estimated length at 50% selectivity of 12.7 cm and length at 95% selectivity at 17.6 cm, for a ratio of length at 95% to 50% selectivity of 1.4. Based on this selectivity curve, rabbitfish are estimated to be selected to the fishing gears before their length at 50% maturity, assumed to be 20.2 cm (Ntiba and Jaccarini 1990). LIME predicted mean length to be increasing on average, with increases and decreases throughout the year, leading to questions about serial depletion during the fishing year. First-year relative spawning biomass increases drastically throughout the time series, up to almost eight times relative spawning biomass in 1996 likely due to the decline in fishing mortality. Relative spawning biomass is more variable than SPR because it also reflects changes in annual recruitment (Figure 3.7).

LB-SPR estimated SPR to be 0.39 in 2014 (95% interval 0.29-0.48), close to the SPR target of 0.40 (Table 3.4, Figure 3.7). LB-SPR estimated 4% of the confidence interval for rabbitfish SPR below the threshold SPR 0.30 in 2014 and a 60% of the confidence interval below the SPR target 0.40. LB-SPR estimated the rabbitfish stock to be overfished in 1998, 2000, 2005, and 2006, which aligns with the period LIME estimated the stock to be overfished (Figure 3.7). The only exception was in 2004, when LB-SPR estimated the stock to be above the SPR target when LIME estimated SPR to be below the 0.30 threshold. LB-SPR estimated a slower increase in SPR than LIME, estimating it was still overfished in 2007 when LIME estimated SPR closer to the SPR target. From 2008 through 2014, LB-SPR estimates of SPR varied widely, but mostly above the overfished threshold except in 2013.The smoothed results estimated SPR to be just below the SPR target for the late 1990s and early 2000s, and just over the SPR target from 2008 through 2013 until 2014 when it dips just below the SPR target (Table 3.5). LB-SPR estimated a wide range of selectivity curves throughout the time series. LB-SPR estimated length at 50% selectivity to be 26.2 cm in 1996, 29.2 cm in 2010, and 25.4 cm in 2011, while estimates ranged from 10.1 cm to 18.5 cm in the other years (Figure 3.7). The years 1996, 2010, and 2011 had relatively sparse data collection (Figure 3.3), resulting in SPR estimates of 0.99 for 1996 and 2011 and very high uncertainty in 2010, although the maximum likelihood estimate was 0.41. Data was relatively sparse and evenly distributed across many length bins in 2004, which was associated with estimated selectivity much lower than other years, with length at 50% selectivity 10.1 and high uncertainty in SPR.

Sensitivities

Across all sensitivity tests, generally LIME estimated SPR in 2014 greater than the SPR target 0.40, while LB-SPR estimated SPR to be lower than the target SPR (Table 3.5, Figure 3.8). Both models had exceptions to these results in some cases of alternative biological parameter inputs. Assuming M was 25% lower than the base model resulted in an SPR estimate of 0.34, below the SPR target 0.40 (Table 3.5). Like the Cost Rican snapper assessment, the likelihood profile for LIME over estimates of M from the natural mortality tool was very flat (Figure 3.9e), so different values of M only affected estimates of fishing mortality, not the model fits. LB-SPR results were also sensitive to the assumed rate for *M*, with estimates of SPR decreasing as *M* decreases (Table 3.5). The likelihood profiles for the von Bertalanffy growth parameters had clear minima, unlike the profile for *M* (Figure 3.9f, g), indicating that the model fit was sensitive to the assumed values for the growth parameters. Assuming L_∞ was 25% greater than the base model resulted in estimated SPR 0.17, an overfished stock. LB-SPR also estimated an overfished stock in 2014 when L_∞ was assumed to be 25% greater than the base model, with SPR of 0.14 compared to 0.39 in the base model (Table 3.5). The patterns in SPR estimates for LIME and LB-SPR are the same if von Bertalanffy *k* was fixed higher or lower than the base model; lower leads to higher estimates of SPR, higher leads to lower estimates of SPR (Table 3.5, Figure 3.9).

Using only length composition data collected from the basket trap, LIME estimated very similar trends in fishing mortality and SPR as the base model, estimating SPR of 0.63 in 2014 (95% interval 0.13-1.0), with 10% of the confidence interval below the 0.30 SPR threshold and 18% of the confidence interval below the SPR target (Table 3.5, Figure 3.8b, d). Selectivity was estimated to have a slightly higher slope than the base model (Figure 3.8f). Using only data from the majority basket trap gear, LB-SPR estimated essentially the same trends as the base model but one year lagged (Figure 3.8c, e) and much higher uncertainty than the base model (SPR 0.34, 95% interval 0-0.71). The high uncertainty in 2014 is associated with the very high selectivity parameters estimated that year, where length at 50% selectivity was estimated to be 31.6 cm, almost as much as the assumed asymptotic length of 36.2 cm (Figure 3.8g). In this case LB-SPR estimated very hard fishing pressure on only the largest individuals.

When only considering data from months with more than 2 sampling days, LIME estimated fishing mortality to have been much higher than the base model before 2000, estimating a similar steep decline in fishing mortality between 2000 and 2006 as in the base model (Figure 3.8d). After 2006, however, LIME using only high sampling months estimated fishing mortality to stabilize between 0.64 and 0.67, just below $F_{30\%}$ of 0.76 (Figure 3.8d). This resulted in SPR estimates below the target SPR 0.40 until 2007, with SPR remaining just over the target, between 0.41 and 0.45 from 2007 to 2014 (Figure 3.8b). The difference in estimates of *F* during the first half of the time series could be attributed to lack of estimation of a recruitment peak between 1998-2000 (Figure 3.8a). LB-SPR estimated much less variable selectivity when considering only high sampling years (Figure 3.8g), due to the high selectivity estimated only in the lower sampling periods.

Dome-shaped selectivity results in lower estimates of fishing mortality and higher estimates of SPR. Fishing mortality estimates for 2014 decreased from 0.41 in the base model to 0.31 with dome-shaped selectivity, and SPR estimates increased from 0.59 in the base model to 0.68 with dome-shaped selectivity.

3.4 DISCUSSION

This chapter documents the first applications of LIME to empirical data for two data-poor stocks, developed as an extension to equilibrium-based length-based methods that can account for variation in fishing mortality and recruitment over time. We compared LIME and LB-SPR assessments for two species with different life history types, representative of common data-limited, small-scale fisheries. Combined with simulation testing (Chapter 2), the empirical examples developed a course of action for application of these methods to more stocks, and suggest diagnostic criteria and ideas for better uncertainty estimates in future applications.

The base LIME model estimated the Costa Rican snapper stock to be overfished in the terminal year, while the Kenyan rabbitfish stock has recovered from a previous period of overfishing. The LB-SPR model estimated the Costa Rican snapper to be above the target SPR and the Kenyan rabbitfish to be at approximately the target SPR. As length-based models, both are sensitive to the biological input values, particularly the assumed values for M and L_∞. LIME attributed changes in the length composition data to variability in recruitment, keeping selectivity constant over time. LB-SPR allowed selectivity to change between years, but assumes recruitment is constant over time. Differences in the value for recruitment standard deviation is not likely to change annual population estimates enough to alter management decisions (Figure 3.6, Figure 3.9). However, the ability to use this parameter to account for variability in the annual recruitment process helps attribute changes in observed data to biological processes, that may help to improve management decisions.

Costa Rican snapper

The LIME base assessment estimated lower SPR than the LB-SPR base assessment, estimating the snapper stock was overfished in 2014 while LB-SPR estimated the stock was above the target SPR level. LB-SPR fits the data very well, but generally changes the predicted length composition between years by allowing parameters to vary widely between years. This ability was particularly important due to the drastic change in the shape of the length composition data in 2014, which has relatively more larger fish than smaller fish as opposed to the other years. LB-SPR attributed this change as a jump in selectivity and the F/M ratio, estimating that only the largest individuals are selected to the gear in that year, but that those individuals are subject to very high fishing pressure. LIME, on the other hand, assumes selectivity was fixed throughout the time series. LIME estimated the length at 50% selectivity to be 36.3 cm, which corresponds to an age between 3 and 4 years old. Thus, when there were relatively few small individuals in the length composition in 2014, LIME attributed this change to a recruitment failure in 2010. When the relative number of large individuals in the catch was much smaller compared to small individuals in 2015, LIME estimated higher fishing mortality in 2014 and 2015.

The drastic change in length composition shape in 2014 was due to the weighting of length composition data between bottom longline and gillnet gear (Figure 3.2). The unweighted length composition data was much more uniformly distributed between small and large fish, with smaller fish that year caught using gillnets and larger fish caught using bottom longline. However, there were many more fish caught in the bottom longline gear that did not have associated length data. Therefore, the weighted length composition represented many more large fish caught in 2014. When considering only bottom longline data, the length composition data in 2014 was not as skewed towards large individuals as in the weighted length composition data. LIME still predicted relatively fewer small individuals in 2014, leading to the same patterns estimated in fishing mortality, recruitment, and SPR. When considering only bottom longline data, LB-SPR also estimated increased fishing mortality and decreased SPR in 2014 and 2015.

To moderate the jumps in estimated selectivity and *F/M* ratio, LB-SPR uses a Kalman filter to smooth estimates so that parameters do not change unrealistically between years. In this case, the LB-SPR base model estimated SPR to remain just over the SPR target 0.40 throughout the time series, with selectivity each year estimated similar to LIME, with approximately the same slope meaning the estimates of 50% and 95% selectivity each year varied around the LIME estimates. However, the Kalman filter does not consider fits to the length composition data, and pulls the *F/M* parameter far above the base estimates if there was an outlier. Therefore, while the SPR and selectivity estimates may be more realistic for the fishery system, the corresponding *F/M* ratio implies that *F* was orders of magnitude higher than *M*, which is not realistic. Applying the Kalman filter to the fishing mortality in log-space with the estimated imprecision as the "weight" associated with the value in each year would prevent outlier estimates from having higher standard errors, providing less leverage in the smoothed estimate.

This assessment estimates the Costa Rican snapper stock was more likely to be overfished than could have been inferred from previous analyses. Bystrom (2015) assumed *M* was equal to 0.43 based on the Pauly (1980) method of estimating natural mortality from local estimates of water temperature. Catch curves using length frequency data estimated total mortality *Z* to be 0.77, resulting in a derived fishing mortality *F* of 0.34. This is much lower than LIME estimated *F* of 0.86 in 2015, although higher than the LB-SPR derived *F* of 0.21 in 2015 (by multiplying the estimated *F/M* of 0.58 by the assumed natural mortality rate of 0.38). If the catch curve analysis assumed $M = 0.38$, the median of estimates from many methods using the natural mortality tool, *F* would have been 0.39, higher than the initial estimate but not as high as estimated by LIME.

The Costa Rican snapper was estimated to be below the overfished SPR threshold of 0.30 for all data cases of LIME (base model, bottom longline data only, and high sampling years only) and when only the bottom longline data was considered using LB-SPR. The NLL from the biological input sensitivities can be compared with the NLL from the base model because they use the same data. None of the LIME sensitivities resulted in NLL significantly different than the base model. The LB-SPR sensitivity tests with NLL significantly lower than the base model were with lower *k* and higher *M*. Both sensitivities would lead to a higher *M/k* ratio. If indeed the

M/k ratio was higher, leading to using either a higher *M* or lower *k* in the LIME assessment, LIME and LB-SPR would be estimate a higher SPR.

Further research into the estimates and uncertainty of growth parameters is recommended. The NLL profile on *k* identified two minima. While the value of *k* used in the base model was close to the minimum associated with a lower value of *k*, the second minimum at a higher value of *k* is associated with lower estimates of SPR and higher estimates of fishing mortality (Figure 3.6). Unlike the flat shape of the *M* likelihood profile, the bimodal shape of the *k* profile is more problematic due to specific values of *k* leading to better fits of the length composition data, with uncertainty in which value of *k* should be used. This is possible due to the anti-correlation between *k* and L_∞. requiring further research into the most appropriate values to assume for the growth parameters (Pilling et al. 2002). One important next step would be to compare multiple methods of estimating von Bertalanffy growth parameters from the length data (Pilling et al. 2002) to identify which value of *k* from the bimodal NLL profile may be more appropriate for use in the assessment model and in turn, management decisions. Bayesian prior distributions would help integrate across the uncertainty in each of the biological parameters, helping to rule out combinations that do not make biological sense.

Kenyan rabbitfish

The LB-SPR assessment for rabbitfish was more conservative than the LIME assessment, estimating SPR at approximately the target SPR of 0.40, whereas LIME estimated the stock has recovered from a previous period of overfishing.

LB-SPR applied to Kenyan rabbitfish suffers from the same issues described for the Costa Rican snapper: selectivity, SPR, and *F/M* estimates vary widely and perhaps unrealistically between years. In the case of Kenyan rabbitfish, the hypothesis of varying selectivity is possible due to the large number of gears harvesting rabbitfish and variability in gear use between years. It is also possible for a short-lived fish like rabbitfish that spawning biomass could vary widely between years due to environmental variation, potentially represented in recruitment variation. It is unlikely that a short-lived fish in a heavily exploited population like Kenyan rabbitfish would be in equilibrium. LB-SPR assumes selectivity varies by year, while LIME assumes recruitment does. In truth both do, and additional information is needed to tease out which model would be more appropriate for a particular fishery. Unlike the Costa Rican snapper, estimates of *F/M* using the smoothing Kalman filter were more realistic. With the smoother, LB-SPR estimated *F/M* as 1.01, and 2.44 without the smoother, which projected forward at equilibrium would result in a recovered population with the smoother or overfished population without the smoother, leading to high uncertainty using LB-SPR in management for this case.

While LB-SPR estimated lower SPR than LIME for the Kenyan rabbitfish stock, LIME estimates provide some evidence of success of coral reef fishery management in Kenya. Active management of area closures occurred throughout southern Kenya beginning in 1968 at the oldest marine park in Malindi, and at other parks in the region in 1972, 1978, and 1991 (McClanahan et al. 2007). McClanahan et al. (2007) estimated exponential increases in biomass of *Siganidae* species (such as the rabbitfish assessed here) due to an increase in fish between 10-20 cm between 1992 and 2004. During this period, LIME estimated clear trends in increasing SPR, decreasing fishing mortality, and increasing first-year relative spawning biomass for *Siganus sutor*. LIME also estimated increased recruitment during this period, attributing a relative increase in small individuals in the length composition in 2000 to a recruitment peak between 1998-2000. This recruitment event could have been reflected in the McClanahan et al. (2007) study identifying increased *Siganidae* biomass.

LIME also addresses concerns in the Hicks and McClanahan (2012) equilibrium-based assessment of coral reef fisheries that the populations are not in fact in equilibrium. While Kenyan fisheries have seen reasonable success from area-based management, Hicks and McClanahan (2012) concluded that gear restrictions have not been as successful. In their assessment of gear restrictions, Hicks and McClanahan estimated F of 1.66 using the equilibrium-based catch curve method, estimating total mortality of 3.15 for *Siganus sutor* assuming M of 1.49 from meta-analysis of methods included in FishBase. This F estimate falls within the 95% confidence intervals estimated by LIME during the overfished period before 2006, and is in the range of the equilibrium-based estimates from LB-SPR throughout the time series. By relaxing the equilibrium assumptions, LIME estimates further reductions in fishing pressure that are not evident in the catch-curve method.

Guidelines for LIME application

As the first empirical application of LIME, these two case studies suggest some guidelines for applying LIME in future assessments (Table 3.4). The first step is to gather life history information. Ideally, estimates of growth and length at maturity would come from recent studies on the same stock. If these are not available, older estimates should be used, and if these are not available, then recent estimates from the same species in other areas should be used. Failing this, estimates of von Bertalanffy growth parameters for related species around the world are often available from FishBase and are generally consistent with regionally compiled values from expert opinion (Thorson et al. 2014a). A similar process should be used for natural mortality, although except in a few cases where natural mortality can be directly estimated from mark-recapture studies (Rudd et al. 2014), natural mortality estimates will often come from imprecise relations among life history invariants (Kenchington 2014). To account for variability in natural mortality estimates, analysts can use the distribution from the natural mortality tool (Cope 2017) based on estimates from many different methods. Beyond sensitivity analyses, future applications should better characterize the full distribution of uncertainty using Monte Carlo simulations to draw from each input distribution simultaneously (Walters et al. 2006, Dick and MacCall 2011), transition from frequentist to Bayesian inference to integrate across the uncertainty from the biological parameter prior distributions (Newman et al. 2009, Martell and Froese 2013, Prince et al. 2015b), or use likelihood-based techniques to assess uncertainty from fixed parameters (MacCall 2013). However, the range of uncertainty in biological parameters is much greater than the values 25% higher or lower tested in simulation analyses found to produce substantial bias in LIME. Biological parameters, either misspecified or with high uncertainty in prior distributions, could be too imprecise to result in meaningful management advice from LIME (Minte-Vera et al. 2017).

After determining the biological inputs for LIME (based on stock-specific studies or descriptive statistics from the distributions), the next step is to determine an appropriate time step for the model by examining the probability of being a given age and length based on the assumed biological inputs (Figure 3.1). Given the age-structured nature of LIME in its current form, it is important that it is possible for a fish to be any length across ages. For example, if the Kenyan *Siganus sutor* were modeled on an annual time step, there would be a near-zero probability for an individual to be between 2 cm (large age 0 fish) and 15 cm (small age-1 fish) in length would approach zero. However, those lengths appear in the observed length composition data. Alternatively, when modeled on a monthly time step, no length bin has a near-zero probability of occurring across ages in months. In the case of the medium-lived Costa Rican snapper, an annual time step would lead to a model with near-zero probability of a fish being between 1 and 5 cm in length, and near-zero probability of being 10 cm during the jump from age-1 to age-2. A quarterly time-step adequately models the snapper's growth, while a monthly time step is exceedingly fine-scale. This fine-scale age-length relationship would lead to greater uncertainty in age given length, leading to difficulty in teasing apart fishing mortality from recruitment given the length composition data (see the longer-lived fish in Chapter 2).

The time-step of the model is limited by the resolution of length composition data. If the length composition data are only available on an annual time step, while analysis of the age-length probabilities would recommend a shorter time step, LIME would require an annual time-step. In this case, it is possible LIME would not be able to account for the growth between annual age groups, resulting in biased estimates of SPR or non-convergence. We recommend using 1 cm length bins, if possible, as pooling the length data into larger bins loses information on fishing pressure or recruitment pulses and does not provide any gain in terms of model convergence or run time (Monnahan et al. 2016). Pooling data into wider length bins could mediate issues in modeling the population on a longer time scale than can be represented in the growth model (Figure 3.1), but modeling on a shorter time scale will fix the problem better than wider length bins. If modeling on a shorter time scale is not an option, we recommend scenario-specific simulation testing with the stock's life history attributes to determine whether the issue modeling growth can be accounted for with larger length bins.

Generally, LIME is not very sensitive to starting values for recruitment standard deviation, selectivity, or fishing mortality. Best practices would be to set the starting values for length at 50% selectivity at approximately halfway between the first observed length in the length composition data and the modal length, and recruitment standard deviation at the taxa median from meta-analysis (Thorson et al. 2014b). LIME could take several minutes to run, but if it takes up to an hour it is most likely due to non-convergence. In this case, a text file would be saved in the output directory indicating a high final gradient, NAs in the standard error report, or indication that the model could not run. If the LIME model does not converge, the most likely reason in our experience is that parameter starting values are far from the truth or length composition data are not informative (are too similar in each year).

In some cases, we identified potential non-convergence even when the final gradient was low enough to meet convergence criteria. This occurred in the Costa Rican snapper assessment over the NLL profile for the von Bertalanffy *k* parameter. To overcome this issue, we used Newton's method after minimizing the negative log-likelihood and used the model with the lowest NLL from many different starting values. In this case, TMB had difficulty finding the global minimum over the likelihood surface when there was likely some combination of growth parameter misspecification and uninformative length composition data. We recommend running Newton's method after the optimization step, which is now written as an option in LIME, and testing many different starting values to ensure TMB has converged on the global minimum.

In the case of non-convergence indicated by a high final gradient, an option is to fix the recruitment standard deviation or selectivity curves. Fishing mortality, recruitment, and selectivity all covary conditionally on a set of length composition data. By making further, transparent assumptions about selectivity, it may still be possible to estimate levels and trends in fishing mortality and recruitment deviations. When possible, however, analysts should estimate both the length at 50% and 95% selectivity, as the shape of the selectivity curve could have significant impacts on model estimates. As an example using LIME, when we fixed the ratio of length at 95% to 50% selectivity at different values between 1.1 (very steep) and 1.7 (widely sloping), these changes greatly impacted estimates of recruitment and fishing mortality. When we ran this sensitivity test for the Costa Rican snapper, widely sloping selectivity curves led to very high estimates of recruitment in the early years, with those recruits essentially populating the fishery (Figure 3.10). Fishing mortality increased over time, leading to estimates of an overfished stock as the population was not getting repopulated by recruits (Figure 3.10). LB-SPR estimated this widely sloping selectivity curve during years with low sampling, likely leading to estimates of fishing mortality biased high. The Kenyan rabbitfish stock was most sensitive to selectivity curves that were very steep (Figure 3.11). In this case, LIME estimated the population to be essentially unfished in 2014, an unlikely hypothesis.

Once LIME has converged and appropriate sensitivity analyses have been explored, an important final step is the communication of results. Length-based methods in general are sensitive to biological inputs and often estimate high uncertainty. For this reason, LIME and LB-SPR are more likely to be viewed useful tools to use as a starting point for estimates of stock status rather than highly precise estimates of stock status. Results from these models are likely to spur discussion between scientists, managers, and stakeholders to consider which inputs are reliable and which need to be improved.

Future directions

This chapter presents a first application of the LIME stock assessment method, but there is certainly room for improvement. While we modeled both stocks using time steps less than one year, we assumed fishing mortality and recruitment were distributed evenly throughout the year, during each time step. Previous studies for both stocks indicated that recruitment likely occurs during specific seasons during the year (Ntiba and Jaccarini 1990, Rojas 1996). Future LIME applications could better include this information by specifying the proportion of recruits entering the population during each time step within a year. The model could also be extended so that users could directly specify the length at each age, rather than assuming von Bertalanffy growth. This would extend LIME applications to sequentially hermaphroditic fish, very common in tropical regions' small-scale fisheries, which often have one pattern of growth at the beginning of their lives and another after they change sexes (Munday et al. 2004).

There has also been interest in applying LIME to unexploited populations, to account for variation in natural mortality and recruitment associated with environmental disturbances and climate change. Because length data are relatively inexpensive to collect, it would be possible to use length composition data of the total population to estimate changes in total mortality instead of fishing mortality.

A next major development would be transitioning to a Bayesian method of inference. A Bayesian algorithm, such as Markov Chain Monte Carlo (Buckland et al. 2007, Newman et al. 2009) or Hamiltonian Monte Carlo that has recently been implemented in TMB (Monnahan et al. 2017), would allow for direct propagation of uncertainty in biological input parameters in specifying prior distributions. The method would then sample from posterior distributions of parameters of interest. While uncertainty in estimates of SPR and fishing mortality are already high in the current approach, explicitly quantifying uncertainty, and due to each input specifically, would help pinpoint information that needs to be improved to decrease uncertainty.

Management strategy evaluation is also needed to compare the performance of LIME relative to other data-poor methods, especially since estimates of stock status tend to have high uncertainty. If LIME is unbiased on average and decisions are made based on the median expected value after integrating uncertainty in input values, perhaps the high uncertainty is an indication that data collection should improve but management does not need to wait for uncertainty to decrease. LIME was developed in the context of data-limited, small-scale, previously unassessed fisheries, the same context that experiences issues with lack of management enforcement. Finally, stock assessment does not exist in a vacuum, and must occur alongside continued data collection, science- and social context-based management decisions, and management enforcement. Only when all of these conditions are met, is it reasonable to expect sustainable exploitation of fisheries.

Acknowledgements

M.B.R. recognizes and acknowledges Conservation International for the Costa Rican spotted rose snapper dataset and support for this project. I thank A. Bystrom for sharing his part of the spotted rose snapper data and analysis. I thank T. McClanahan, C. Abunge, and the team of scientists at the Wildlife Conservation Society in Mombasa for the Kenyan coral reef fishery dataset and for hands-on learning about the fishery and management system during my visit. Many thanks to J. Cope for helpful tips and insights during the data-limited assessment process.

3.5 TABLES

Table 3.1. Parameter symbols, definitions, fixed values, and starting values for estimated parameters for the species assessed in this study. Normally distributed prior distributions are designated with *N(*mean,standard deviation).

Parameter	Description	Costa Rican snapper	Kenyan rabbitfish
Fixed			
M/k	Ratio used in LB-SPR	1.79	1.60
M	Natural mortality (yr^{-1})	0.38	1.39
L_∞	Asymptotic length (cm)	64.6	36.2
k	Growth coefficient (yr^{-1})	0.21	0.87
t_0	Age at length=0	-0.01	-0.24
α	Weight-length scalar (kg)	0.0245	0.0600
β	Weight-length allometric	2.79	2.76
L_m^{50}	Length at 50% maturity (cm)	34.0	20.2
CV_L	Growth curve coefficient of variation	0.10	0.10
A	Maximum age (years)	13	4
N_s	Number of time steps per year modeled	4	12
Estimated			
L_s^{50}	Length at 50% selectivity (cm)	20.0	11.0
δ	Ratio between length at 95% and 50% selectivity	1.3	1.3
σ_R	Standard deviation for recruitment	$N(0.737,0.353)$	$N(0.737,0.353)$
F_1	Fishing mortality in the first time step	1.0	1.0
F_t	Annual fishing mortality	$N(F_{t-1}, 0.2)$	$N(F_{t-1}, 0.2)$
θ	Dirichlet-multinomial parameter	10.0	10.0

Table 3.2. Estimates of SPR for Costa Rican *Lutjanus guttatus*, compared between LIME and LB-SPR base models.

	SPR		L_{50} (cm)		L_{95} (cm)		F/F$_{\text{reference}}$	
Year	LIME	LB-SPR	LIME	LB-SPR	LIME	LB-SPR	LIME (F/F_{40})	LB-SPR (F/M)
2007	0.31	0.63	36.3	31.4	43.6	38.3	1.45	0.39
2008	0.31	0.22	36.3	34.3	43.6	41.1	1.44	2.23
2009	0.38	0.33	36.3	32.2	43.6	37.1	1.07	1.19
2010	0.40	0.66	36.3	27.1	43.6	30.7	1.01	0.29
2011	0.42	0.61	36.3	30.9	43.6	37.7	0.91	0.41
2012	0.42	0.35	36.3	35.7	43.6	45.3	0.92	1.36
2013	0.39	0.58	36.3	35.7	43.6	43.4	1.05	0.59
2014	0.25	0.33	36.3	57.0	43.6	68.7	1.94	37.4
2015	0.24	0.49	36.3	29.2	43.6	33.8	2.05	0.58

Table 3.3. Parameter estimates and derived quantities for Costa Rican *Lutjanus guttatus* in 2015 comparing LIME and LB-SPR base models and model runs with the majority gear only and high sampling periods, +/- 25% asymptotic length, von Bertalanffy *k*, and natural mortality, along with dome-shaped selectivity for LIME only. Asterisk (*) denotes a fixed value.

Model	Parameter/Derived value								
LIME	*SPR*	L^{50} *(cm)*	L^{95} *(cm)*	$F/F_{40\%}$	*F*	σ_R	*NLL*		
Base	0.24	36.3	43.6	2.05	0.86	0.64	3976		
MG	0.18	34.4	39.9	2.74	1.00	0.56	3085		
High	0.29	35.7	42.5	1.64	0.65	0.89	2101		
Linf -25%	1.00	44.5	54.6	0.00	0.00	0.72	4337		
Linf +25%	0.13	38.7	47.3	4.80	1.81	0.59	3997		
k -25%	0.29	35.8	42.4	1.64	0.72	0.78	3986		
k +25%	0.20	36.7	44.3	2.54	1.04	0.60	3979		
M -25%	0.16	36.1	43.5	3.28	0.95	0.76	3976		
M +25%	0.34	36.4	43.8	1.31	0.75	0.76	3976		
Dome	0.25	36.3*	43.6*	1.97	0.83	0.74*	3977		
LB-SPR	*SPR*	L^{50} *(cm)*	L^{95} *(cm)*	*F/M*	*Smooth SPR*	*Smooth* L^{50} *(cm)*	*Smooth* L^{95}*(cm)*	*Smooth F/M*	*NLL*
Base	0.49	29.2	33.8	0.58	0.46	36.8	44.2	8.08	1286
MG	0.23	29.4	33.9	1.46	0.40	31.6	37.9	1.04	635
High	0.49	29.2	33.8	0.58	0.44	39.0	47.1	10.3	821
Linf -25%	1.00	31.9	38.9	0.00	1.00	39.7	47.3	0.00	29625
Linf +25%	0.75	47.4	56.2	32.8	0.91	46.7	54.9	12.6	198398
k -25%	0.97	29.0	33.4	0.03	0.81	36.6	53.6	5.59	1208
k +25%	0.31	29.2	33.9	1.05	0.30	37.0	44.7	10.1	1307
M -25%	0.27	29.2	34.0	1.21	0.27	37.0	44.7	10.5	1310
M +25%	0.82	29.0	33.5	0.15	0.74	36.5	43.4	6.01	1180

Table 3.4. Estimates of SPR for Kenyan *Siganus sutor*, compared between LIME and LB-SPR base models.

	SPR		L^{50} (cm)		L^{95} (cm)		$F/F_{reference}$	
Year	LIME	LB-SPR	LIME	LB-SPR	LIME	LB-SPR	LIME (F/F_{40})	LB-SPR (F/M)
1996	0.72	1.00	12.7	26.2	17.7	35.3	0.33	0.00
1997	0.32	0.51	12.7	16.5	17.7	18.4	1.30	0.53
1998	0.20	0.11	12.7	12.3	17.7	17.4	2.02	1.78
1999	0.16		12.7		17.7		2.33	
2000	0.15	0.08	12.7	12.4	17.7	14.7	2.44	2.12
2001	0.17		12.7		17.7		2.29	
2002	0.18		12.7		17.7		2.14	
2003	0.20		12.7		17.7		1.98	
2004	0.23	0.58	12.7	10.1	17.7	12.5	1.79	0.30
2005	0.27	0.12	12.7	16.5	17.7	22.8	1.51	2.47
2006	0.35	0.12	12.7	11.5	17.7	16.6	1.19	1.57
2007	0.39	0.25	12.7	15.5	17.7	23.3	1.04	1.19
2008	0.43	0.67	12.7	13.4	17.7	19.4	0.90	0.26
2009	0.48	0.49	12.7	13.0	17.7	16.5	0.78	0.47
2010	0.52	0.41	12.7	29.2	17.7	39.3	0.69	3.54
2011	0.53	1.00	12.7	25.4	17.7	39.9	0.66	0.00
2012	0.55	0.35	12.7	18.5	17.7	26.7	0.61	1.07
2013	0.59	0.07	12.7	17.5	17.7	24.4	0.54	4.28
2014	0.59	0.39	12.7	14.0	17.7	18.4	0.53	0.68

Table 3.5. Parameter estimates and derived quantities for Kenyan *Siganus sutor* in 2014 comparing LIME and LB-SPR base models and model runs with the majority gear only and high sampling periods, along with high dome-shaped selectivity for LIME only. Asterisk (*) denotes a fixed value.

Model	Parameter/Derived value								
LIME	*SPR*	*L^{50} (cm)*	*L^{95} (cm)*	*$F/F_{40\%}$*	*F*	*σ_R*	*NLL*		
Base	0.59	12.7	17.6	0.53	0.41	0.72	3653		
MG	0.63	13.7	17.3	0.45	0.37	1.02	2188		
High	0.45	12.5	17.6	0.84	0.64	0.45	2310		
Linf -25%	1.00	11.5	16.5	2.0e-6	1.3e-6	1.87	3925		
Linf +25%	0.17	13.5	19.1	2.36	1.94	0.59	3683		
k -25%	1.00	13.1	18.2	6.4e-8	4.6e-8	0.56	3654		
k +25%	0.33	12.6	17.6	1.28	1.05	0.76	3674		
M -25%	0.34	12.6	17.6	1.20	0.75	0.72	3653		
M +25%	0.93	12.8	17.7	0.07	0.06	0.69	3653		
Dome	0.68	12.7*	17.6*	0.49	0.31	0.74*	3653		
LB-SPR	*SPR*	*L^{50} (cm)*	*L^{95} (cm)*	*F/M*	*Smooth SPR*	*Smooth L^{50} (cm)*	*Smooth L^{95}(cm)*	*Smooth F/M*	*NLL*
Base	0.39	14.0	18.4	0.68	0.39	17.6	24.7	1.64	1028
MG	0.35	31.6	40.7	9.61	0.38	19.8	24.7	3.59	334
High	0.39	14.0	18.4	0.68	0.33	13.8	19.2	1.14	681
Linf -25%	1.00	14.0	18.6	0.00	0.82	17.6	24.3	0.37	2234
Linf +25%	0.14	14.6	19.3	1.73	0.13	19.3	26.4	5.05	858
k -25%	0.79	13.9	18.1	0.15	0.64	18.0	24.7	0.86	1138
k +25%	0.24	14.0	18.4	1.15	0.28	17.4	24.6	2.34	1025
M -25%	0.21	14.0	18.4	1.30	0.24	17.5	24.8	2.55	1090
M +25%	0.66	13.9	18.2	0.27	0.57	17.9	24.6	1.03	883

3.6 FIGURES

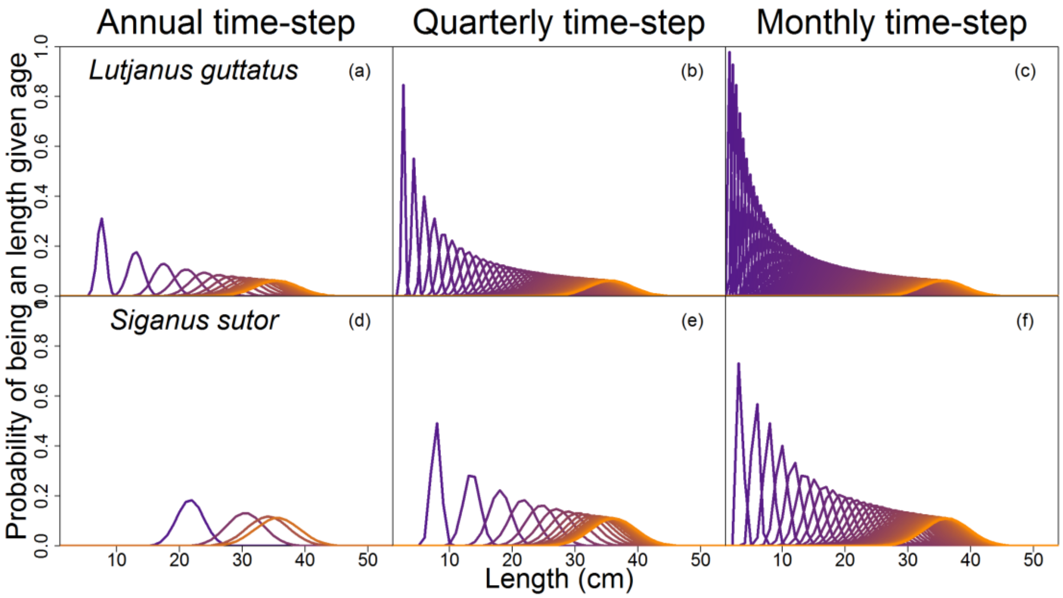

Figure 3.1. Probability of being a length given age, where each line represents a different age in years (a, d), quarter-years (b, e), or months (c, f) for the medium-lived *Lutjanus guttatus* (a-c) and short-lived *Siganus sutor* (d-f). Overlapping distributions can account for growth of the fish during the year without requiring an unrealistic jump from one distribution to the next between time steps.

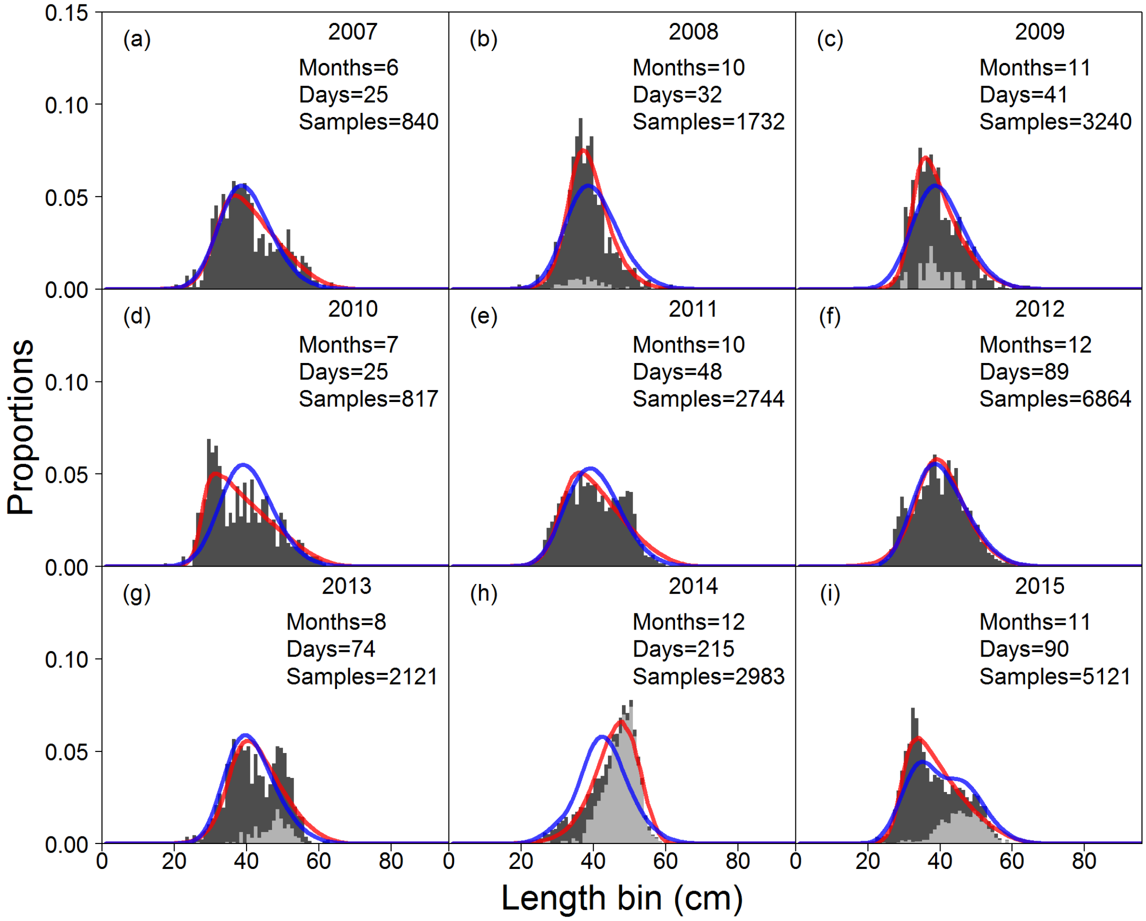

Figure 3.2. Length composition data with majority gear (dark gray) and other gears (light gray) with LIME (blue) and LB-SPR (red) model fits for Costa Rican snapper annual length composition data. Quarterly LIME estimates were added together for each year and weighted by number of monthly samples.

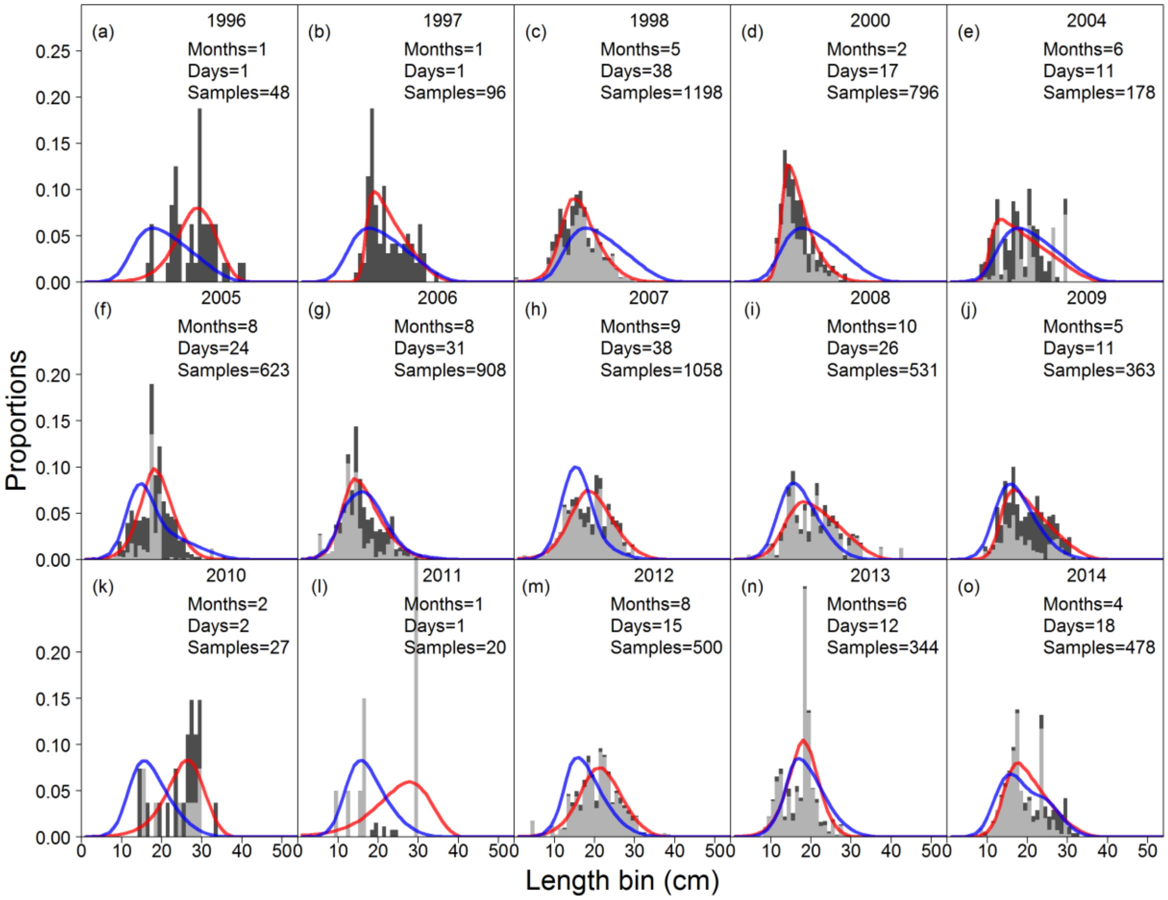

Figure 3.3. Length composition data with majority gear (dark gray) and other gears (light gray) with LIME (blue) and LB-SPR (red) model fits for Kenyan rabbitfish annual length composition data. Monthly LIME estimates were added together for each year and weighted by number of monthly samples. 65% of fish sampled in 2011 were in the 30-cm length bin (i).

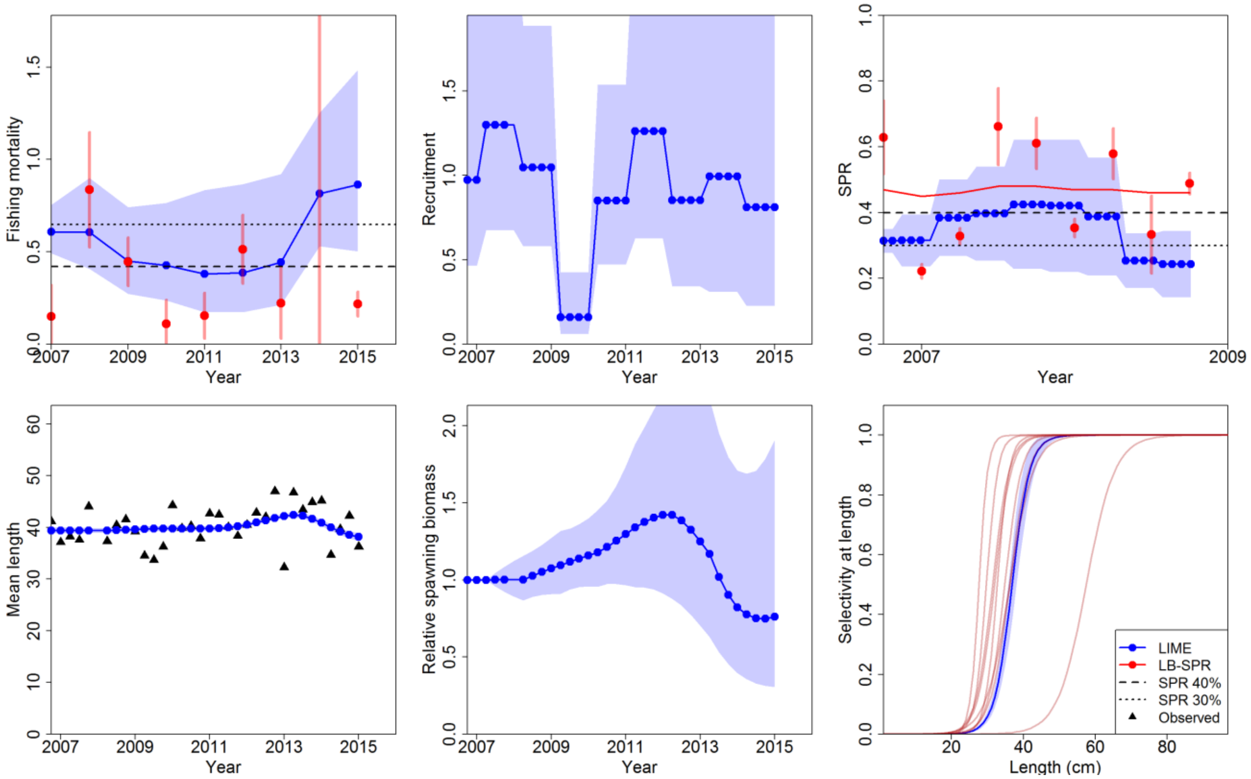

Figure 3.4. Model estimates of Costa Rican snapper fishing mortality, recruitment, SPR, mean length, first-year relative spawning biomass, and selectivity from LIME (blue) and estimates of SPR and selectivity for LB-SPR (red) using the weighted annual length composition data including all gears. Points designate months or years with data, lines (LIME) or points (LB-SPR) indicate maximum likelihood estimates, and shading (LIME) or error bars (LB-SPR) indicate 95% confidence intervals. Multiple red lines for selectivity represent the separate curves estimated for each year of LB-SPR. Black triangles represent observed mean length in each quarter.

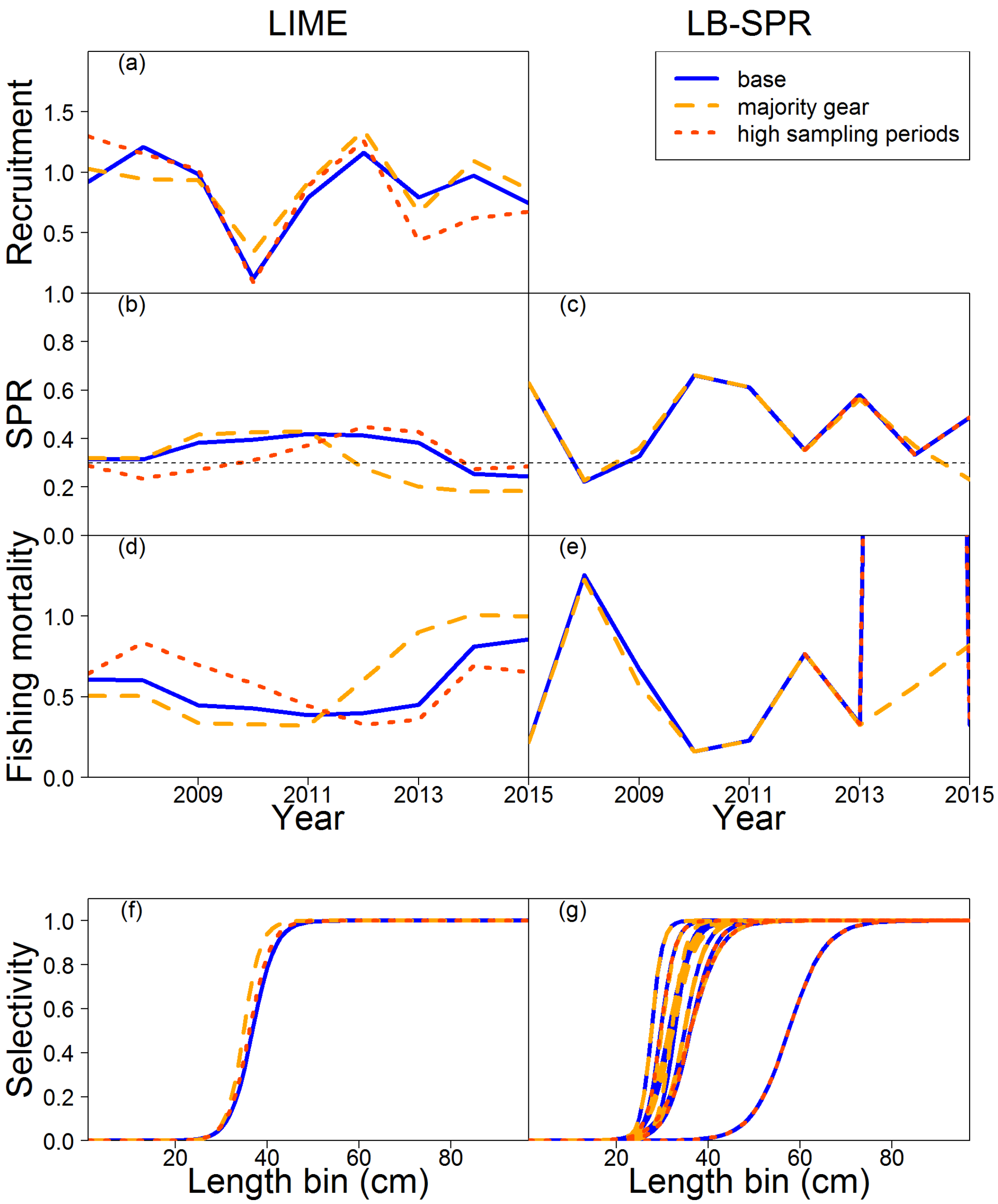

Figure 3.5. Maximum likelihood estimates for the base model, majority gear only, and high sampling periods only from LIME and LB-SPR, comparing recruitment (LIME only), SPR, fishing mortality, and selectivity for the Costa Rican snapper fishery. Multiple lines for each data availability scenario for LB-SPR selectivity represent the different estimated curves for each year. Fishing mortality rates for the base model and high sampling years extend up to F=14 in (e).

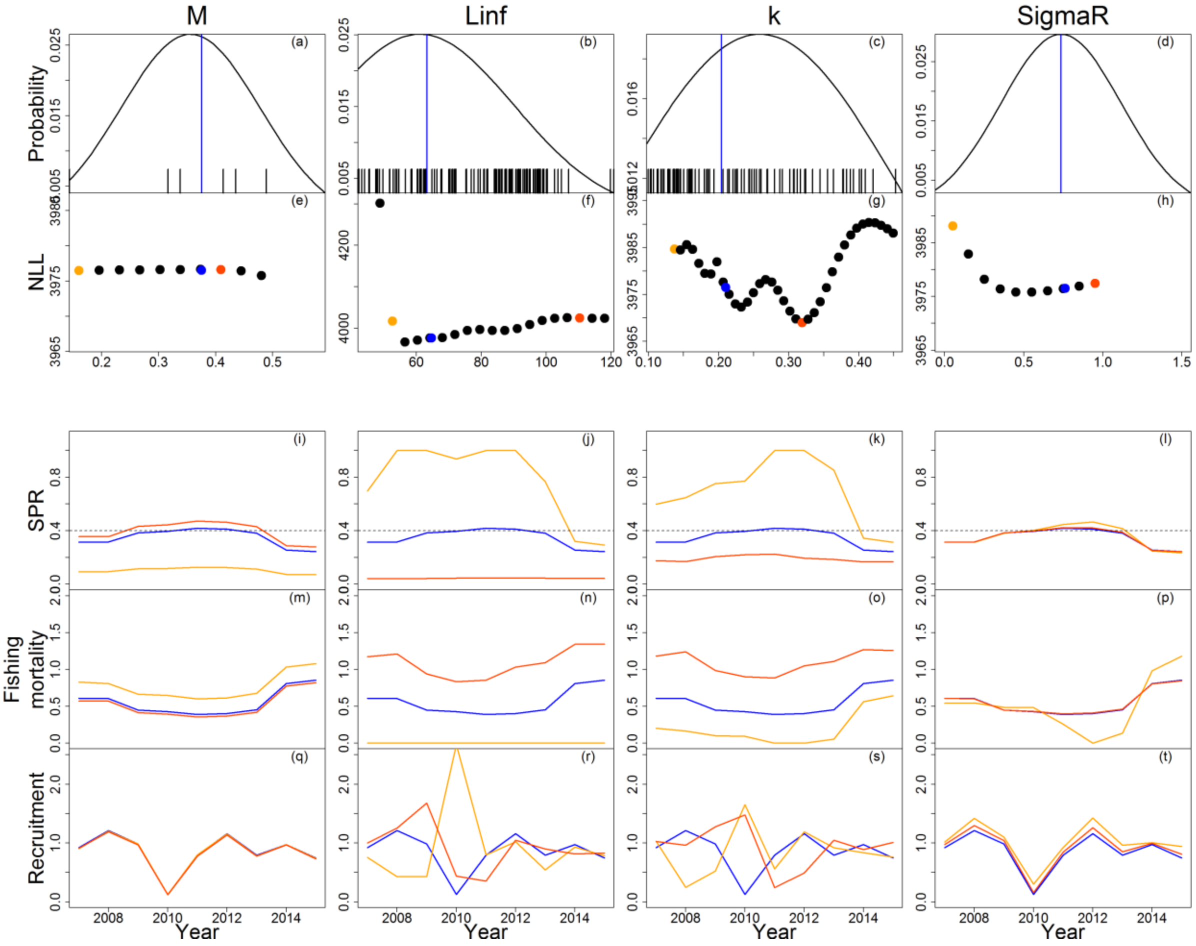

Figure 3.6. Costa Rican snapper distributions of natural mortality rate (*M*) from the natural mortality tool (a), asymptotic length (L_∞) and von Bertalanffy growth (*k*) from the *Lutjanus* genus from FishBase (b, c), and recruitment standard deviation (SigmaR) for perciforms (d, Thorson et al. 2014). Negative log likelihood profiles (e-h) and associated maximum likelihood estimates of spawning potential ratio for the LIME base model (blue) and alternate values (yellow and orange) for each parameter, with resulting spawning potential ratio (i-l). Black points indicate converged runs without examples of their SPR trajectories.

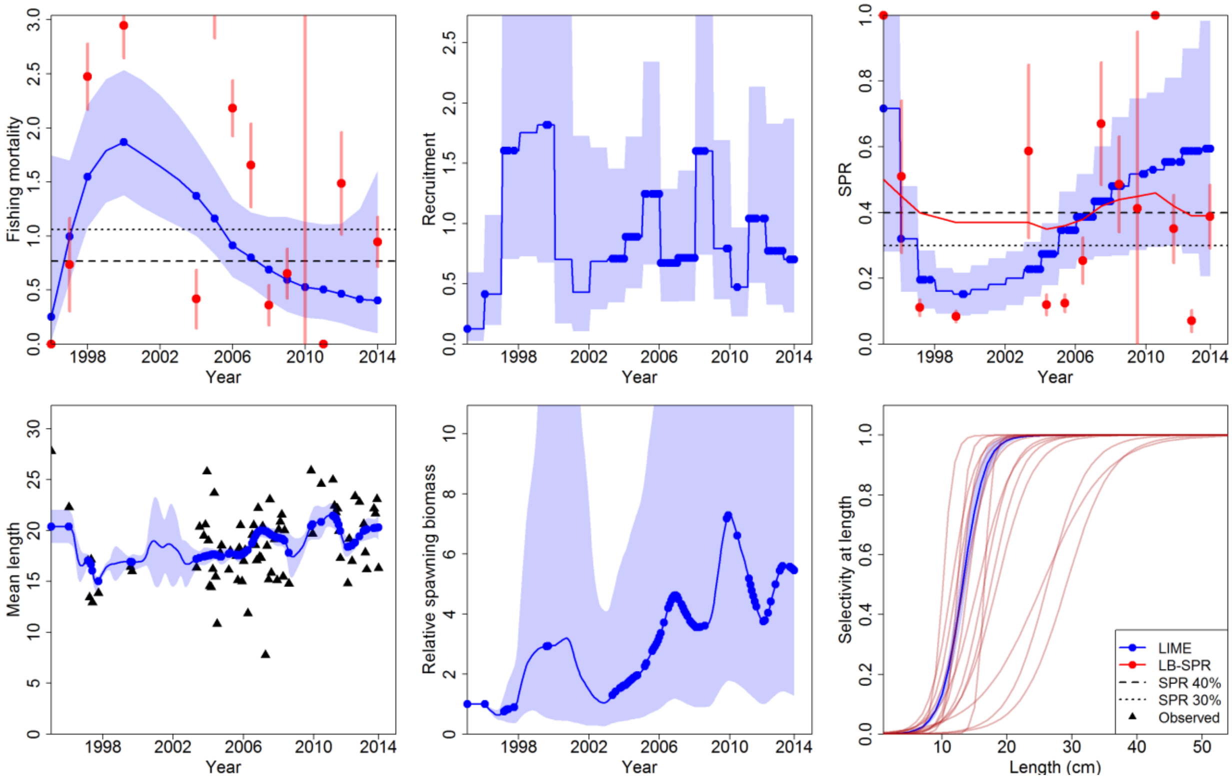

Figure 3.7. Model estimates of Kenyan rabbitfish fishing mortality, recruitment, SPR, mean length, first-year relative spawning biomass, and selectivity from LIME (blue) and estimates of SPR and selectivity for LBSPR (red) using the weighted annual length composition data including all gears. Points designate months or years with data, lines (LIME) or points (LB-SPR) indicate maximum likelihood estimates, and shading (LIME) or error bars (LB-SPR) indicate 95% confidence intervals. Multiple red lines for selectivity represent the separate curves estimated for each year of LB-SPR. Black triangles represent observed mean length in each month.

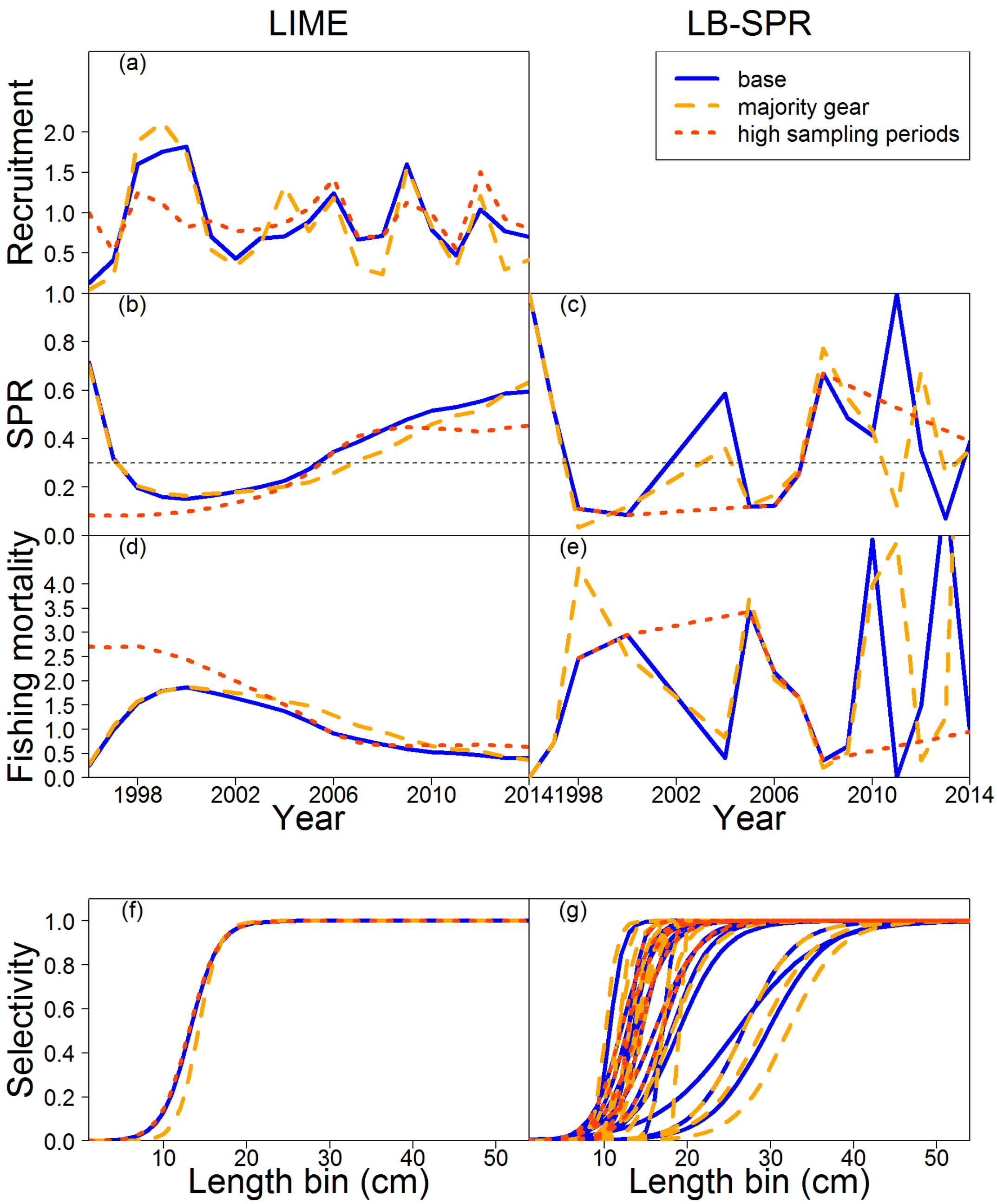

Figure 3.8. Maximum likelihood estimates for the base model, majority gear only, and high sampling periods only from LIME and LB-SPR, comparing recruitment (LIME only), SPR, fishing mortality, and selectivity for the Kenyan rabbitfish fishery. Multiple lines for each data availability scenario for LB-SPR selectivity represent the different estimated curves for each year.

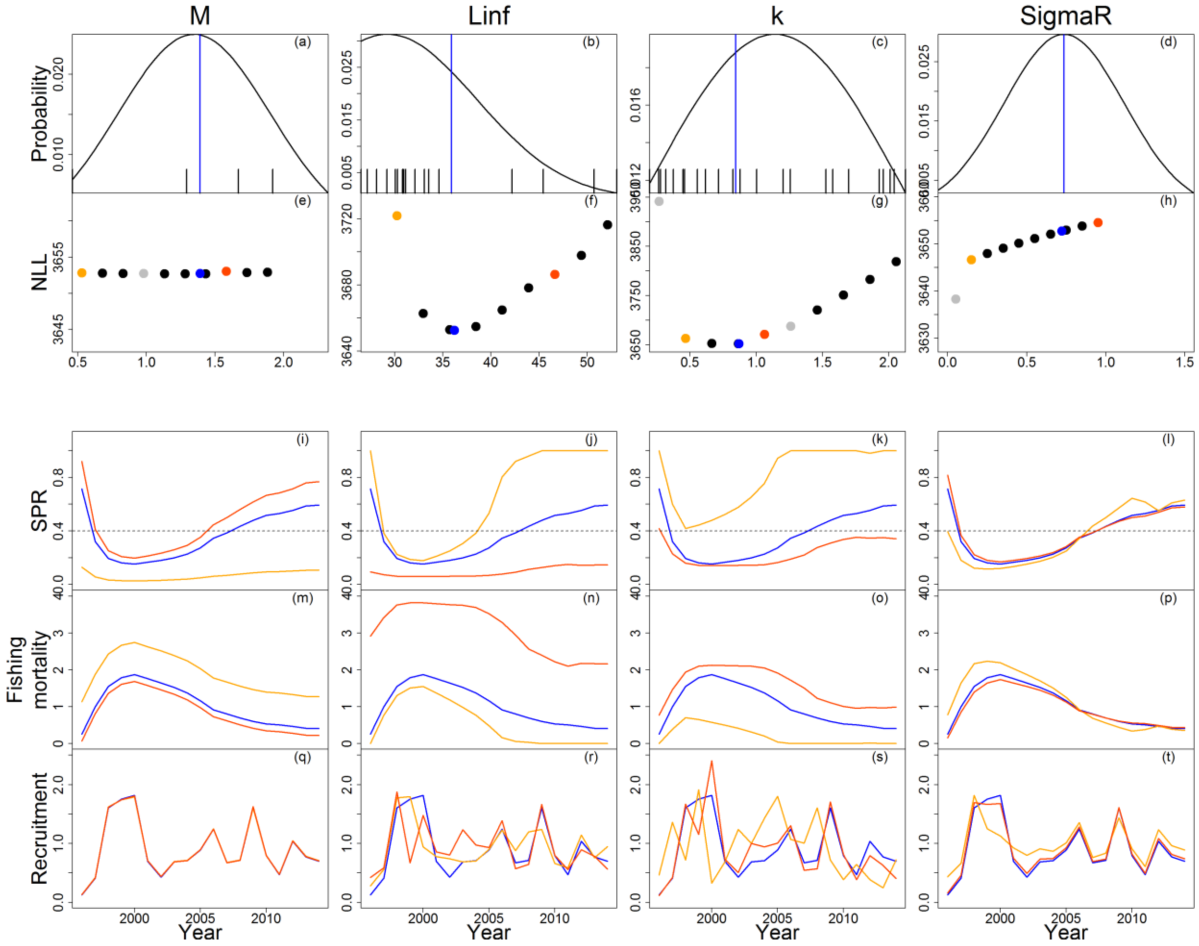

Figure 3.9. Kenyan rabbitfish distributions of natural mortality rate (*M*) from the natural mortality tool (a), asymptotic length (Linf) and von Bertalanffy growth (*k*) from the *Siganus* genus from FishBase (b, c), and recruitment standard deviation (SigmaR) for perciforms (d, Thorson et al. 2014). Negative log likelihood profiles (e-h) and associated maximum likelihood estimates of spawning potential ratio for the LIME base model (blue) and alternate values (yellow and orange) for each parameter, with resulting spawning potential ratio (i-l). Gray points indicate non-converged runs, and black points indicate converged runs without examples of their SPR trajectories.

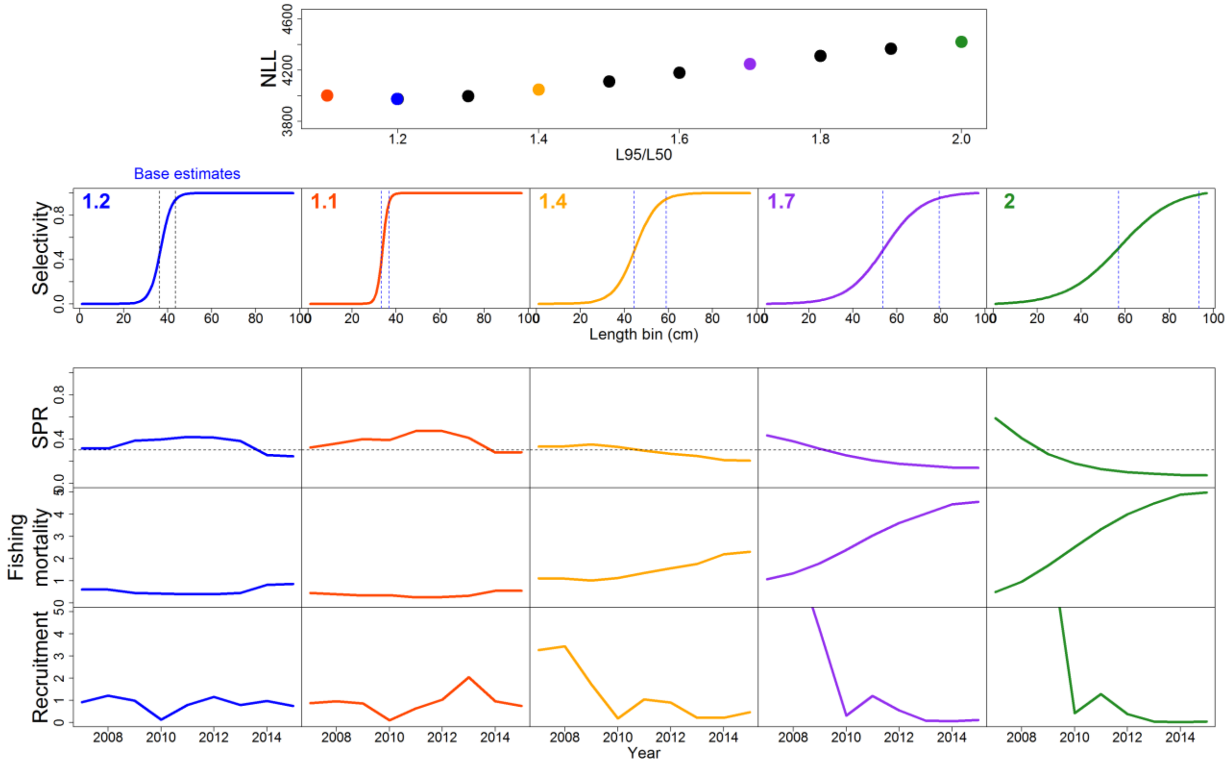

Figure 3.10. Negative log likelihood profile over values of δ (L_s^{95}/L_s^{50}) for the Costa Rican snapper, followed by base model estimated selectivity (column 1) compared to results when the δ parameter was fixed to 1.1, 1.4, 1.7, and 2.0, while length at 50% selectivity was estimated. This demonstrates the importance of estimating δ rather than fixing the parameter. If the δ parameter is fixed too high, this could result in misattributing changes to the length composition with unrealistically high recruitment that allows the population to persist while no new recruits enter the population but fishing mortality increases dramatically.

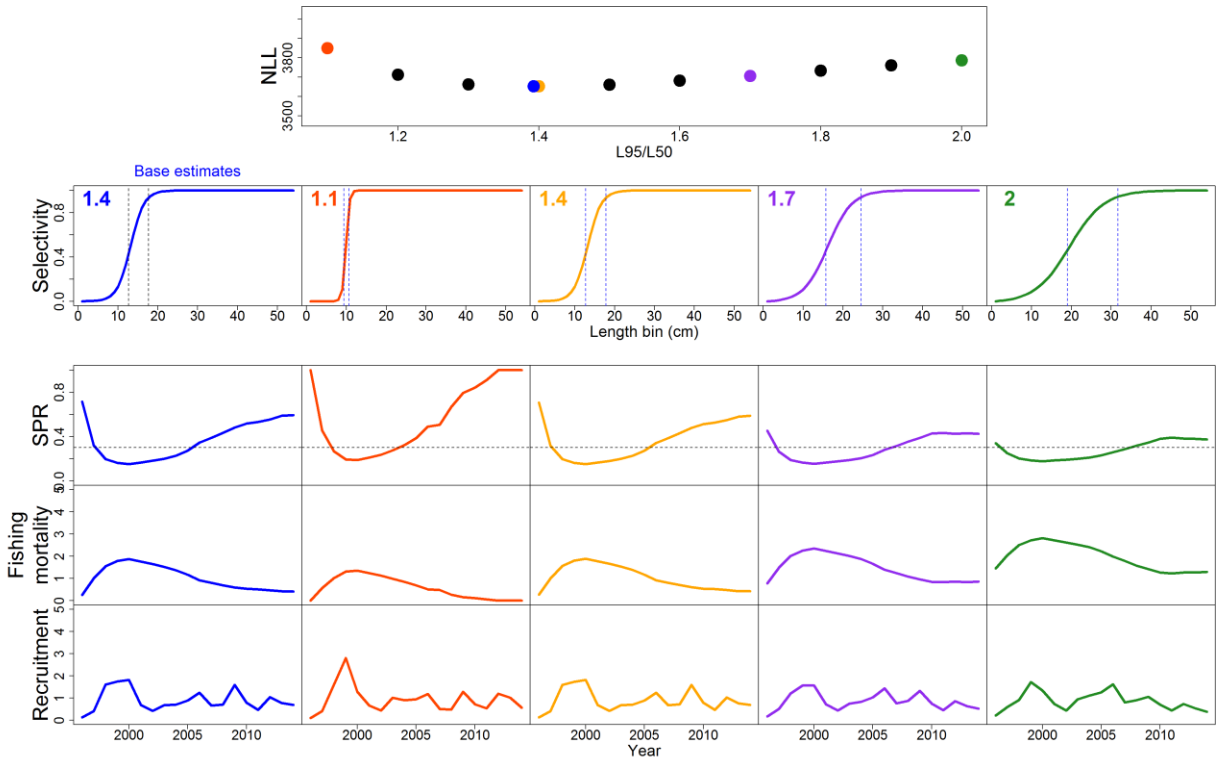

Figure 3.11. Negative log likelihood profile over values of δ (L_S^{95}/L_S^{50}) for Kenyan rabbitfish, followed by base model estimated selectivity (column 1) compared to results when the δ parameter was fixed to 1.1, 1.4, 1.7, and 2.0, while length at 50% selectivity was estimated. This demonstrates the importance of estimating δ rather than fixing the parameter, in that the level of stock recovery is highly dependent on the slope of the selectivity curve.

CONCLUSIONS

The unifying theme of this dissertation is the attempt to confront and address limited data quality and quantity. Data limitation is not an excuse to avoid management, but the information held in each data type must be addressed. While this dissertation focuses on catch misreporting in the context of surplus production models and length composition data, no assessment of any kind should run without thoroughly questioning the methods used to collect the data, how the data may be biased, how the assessment inputs may be biased, and how data collection can improve.

Chapter 1 posed the question, “does unreported catch lead to overfishing?” and we found that the answer is, “not necessarily.” Where catch has been misreported at a constant rate, estimates of maximum sustainable yield and population biomass will be biased proportionally to the level of misreporting, so the target exploitation rate would be set appropriately given an effective fisheries management system. If catch misreporting has changed over time, stock assessment would most likely lead to lower-than-necessary catch limits if reporting rates have improved or overfishing is reporting rates have decreased. These patterns are relevant for fisheries with suspected catch misreporting, trying to conduct management while knowing misreporting of any sort is ongoing. In these cases, it will be more useful to delve into patterns of misreporting than quantify total catch itself.

The findings in Chapter 1 could also be used in the assessment and management process surrounding the discussions on use of catch reconstructions. The United Nation’s Food and Agriculture Organization global fishery landings database, summarized in biennial *State of the World Fisheries and Aquaculture* reports (SOFIA; e.g. FAO 2016), is known to contain under-reported global landings (Garibaldi 2012) due to difficulties quantifying discards, artisanal, recreational, or other non-industrial catch. The Sea Around Us Project has worked to reconstruct catch country-by-country around the world, in an attempt to better account for those commonly unreported sectors (Pauly and Zeller 2016). However, the FAO does not cite catch reconstructions in the SOFIA reports, igniting criticism from the Sea Around Us Project (Pauly and Zeller 2017). Ye et al. (2017) respond to these criticisms that the methods of interpolation used in the catch reconstructions are not statistically sound. The response from FAO scientists emphasizes the same take-away message as the first chapter of this thesis: that the assumptions in catch reconstructions, or any catch dataset, must be thoroughly examined and understood before used in stock assessment and consequent management. Both FAO and the Sea Around Us Project can agree that it is inherently difficult to measure illegal, unreported, and unregulated catch, as well as other commonly unreported sectors. Yet, many management systems worldwide call for the setting of catch limits, despite uncertainties in the data. The results from Chapter 1 remind managers to consider the biases in the catch data at hand, and how they propagate through to management.

Chapter 2 presents LIME as an additional tool for the length-based data-limited methods toolbox, while Chapter 3 demonstrates how LIME can be used for medium- and short-lived species. One of the most important findings via simulation for future LIME applications is the need to consider the time-step in which to run the model that can adequately account for individual growth within the year. LIME and LB-SPR perform very differently when self- and cross-tested with operating models based on their own structure or an alternate structure. When the fish was short-lived, and the data were collected throughout the year and pooled together on an annual time step (as is the scenario for many coral reef fish species worldwide), age-structured models on an annual time step cannot adequately model individual growth during the year, leading to very poor model fits and poor performance. This is not just an issue for LIME, but for all age-structured models. For example, this remains true for the DLMtool (Carruthers and Hordyk 2015), an R package increasingly used for data-limited fisheries worldwide (Cummings et al. 2016). The operating model used to run the management strategy evaluation in DLMtool is age-structured, resulting in better performance for age-structured data-limited methods and poor performance for methods cross-tested. At this point, the DLMtool is not recommended for short-lived species due to the inability to model individual growth within an annual time step. This is a major drawback, since many of the world's data-limited fisheries are also short-lived.

While LIME is expected to be biased if it does not adequately model individual growth during the time step, LIME estimates have high uncertainty for longer-lived fish (e.g. greater than 20 years). This is due to determinate growth: at shorter lengths, fish have more predictable ages, but at longer lengths, fish could be a wide range of ages. This reduces the ability to track cohorts over time, and the ability to tease apart recruitment pulses from fishing mortality in the length composition data. Thus, we recommend using LIME for fish that live less than about 20 years, but the shorter the life history, the more care must be taken to model using an appropriate time step for individual growth.

The development of LIME moves length-based data-limited methods forward in its integrated nature and treatment of variability. Annual catch limits (ACLs) for U.S. fisheries with only length data and biological information are currently set using mean length methods or the DLMtool. Most mean length models use length as an empirical indicator outside of the statistical framework. The equilibrium assumption of mean-length methods and LB-SPR is limiting, and we cannot hope to find many fisheries that meet these assumptions. ACLs in the Gulf of Mexico and Caribbean have been set using the DLMtool. The DLMtool can rapidly compare different data-poor methods. A next step is to add LIME to the many methods in DLMtool. As stated earlier, however, the DLMtool is limited in its functionality for short-lived species. Future research is necessary to examine the effects of modeling size composition data with age structured models. There has also been interest in applying LIME to unexploited species that are impacted by environmental changes and climate change. Instead of modeling fishing mortality, LIME could try to estimate changes in natural mortality over time, assuming fishing mortality is zero, accounting for survey gear selectivity and recruitment variation. We expect LIME to be useful as a tool to estimate stock status for the many fisheries worldwide that have local biological information, biological information shared from similar stocks, and at least one year of length composition data. This could include small-scale fisheries, bycatch fisheries, and perhaps in the future, unexploited populations impacted by climate change.

BIBLIOGRAPHY

Agnew, D. J., J. Pearce, G. Pramod, T. Peatman, R. Watson, J. R. Beddington, and T. J. Pitcher. 2009. Estimating the worldwide extent of illegal fishing. PloS ONE 4:e4570.

Allison, E. H., and F. Ellis. 2001. The livelihoods approach and management of small-scale fisheries. Marine Policy 25:377–388.

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

Andrew, N. L., C. Béné, S. J. Hall, E. H. Allison, S. Heck, and B. D. Ratner. 2007. Diagnosis and management of small-scale fisheries in developing countries. Fish and Fisheries 8:227–240.

Armitage, D., C. Béné, A. T. Charles, D. Johnson, and E. H. Allison. 2012. The interplay of well-being and resilience in applying a social-ecological perspective. Ecology and Society 17.

Ault, J. S., J. A. Bohnsack, S. G. Smith, and J. Luo. 2005a. Towards sustainable multispecies fisheries in the Florida, USA, coral reef ecosystem. Bulletin of Marine Science 76:595–622.

Ault, J. S., S. G. Smith, and J. A. Bohnsack. 2005b. Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. ICES Journal of Marine Science 62:417–423.

Ault, J. S., S. G. Smith, J. A. Browder, W. Nuttle, E. C. Franklin, J. Luo, G. T. Dinardo, and J. A. Bohnsack. 2014. Indicators for assessing the ecological dynamics and sustainability of southern Florida’s coral reef and coastal fisheries. Ecological Indicators 44:164–172.

Ault, J. S., S. G. Smith, J. Luo, M. E. Monaco, and R. S. Appeldoorn. 2008. Length-based assessment of sustainability benchmarks for coral reef fishes in Puerto Rico. Environmental Conservation 35:221–231.

Azzurro, E., P. Moschella, and F. Maynou. 2011. Tracking signals of change in mediterranean fish diversity based on local ecological knowledge. PLoS ONE 6:e24885.

Belhabib, D., V. Koutob, A. Sall, V. W. Y. Lam, and D. Pauly. 2014. Fisheries catch misreporting and its implications: The case of Senegal. Fisheries Research 151:1–11.

Bene, C., B. Hersoug, and E. H. Allison. 2010. Not by rent alone: analyzing the pro-poor functions of small-scale fisheries in developing countries. Development Policy Review 28:325–358.

Berg, C. W., A. Nielsen, and K. Kristensen. 2014. Evaluation of alternative age-based methods for estimating relative abundance from survey data in relation to assessment models. Fisheries Research 151:91–99.

Berkson, J., L. Barbieri, S. Cadrin, S. Cass-Calay, P. Crone, M. Dorn, C. Friess, D. Kobayashi,

T. J. Miller, W. S. Patrick, S. Pautzke, S. Ralston, and M. Trianni. 2011. Calculating acceptable biological catch for stocks that have reliable catch data only (Only Reliable Catch Stocks - ORCS). NOAA Technical Memorandum NMFS-SEFSC-616.

Beverton, R. J. H., and S. J. Holt. 1957. On the Dynamics of Exploited Fish Populations. Ministry of Agriculture, Lowestoft, UK.

Branch, T. A., O. P. Jensen, D. Ricard, Y. Ye, and R. Hilborn. 2011. Contrasting global trends in marine fishery status obtained from catches and from stock assessments. Conservation Biology 25:777–786.

Buckland, S. T., K. B. Newman, C. Fernández, L. Thomas, and J. Harwood. 2007. Embedding Population Dynamics Models in Inference. Statistical Science 22:44–58.

Buckland, S. T., K. B. Newman, L. Thomas, and N. B. Koesters. 2004. State-space models for the dynamics of wild animal populations. Ecological Modelling 171:157–175.

Butler, J. R. A., S. J. Middlemas, I. M. Graham, P. M. Thompson, and J. D. Armstrong. 2006. Modelling the impacts of removing seal predation from Atlantic salmon, Salmo salar, rivers in Scotland: A tool for targeting conflict resolution. Fisheries Management and Ecology 13:285–291.

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.

Carruthers, T., and A. Hordyk. 2015. DLMtool: Data-Limited Methods Toolkit.

Carruthers, T. R., A. E. Punt, C. J. Walters, A. MacCall, M. K. McAllister, E. J. Dick, and J. Cope. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research 153:48–68.

Chaboud, C., M. Fall, J. Ferraris, A. Fontana, A. Fonteneau, F. Laloë, A. Samba, and D. Thiao. 2015. Comment on “Fisheries catch misreporting and its implications: The case of Senegal.” Fisheries Research 164:322–324.

Chang, W., J. Cheng, J. J. Allaire, Y. Xie, and J. McPherson. 2017. Shiny: Web Application Framework for R.

Chrysafi, A., and A. Kuparinen. 2015. Assessing abundance of populations with limited data: Lessons learned from data-poor fisheries stock assessment. Environmental Reviews 1:1–44.

Cinner, J. E., N. A. J. Graham, C. Huchery, and M. A. Macneil. 2013a. Global effects of local human population density and distance to markets on the condition of coral reef fisheries. Conservation biology 27:453–8.

Cinner, J. E., C. Huchery, E. S. Darling, A. T. Humphries, N. A. J. Graham, C. C. Hicks, N. Marshall, and T. R. McClanahan. 2013b. Evaluating social and ecological vulnerability of coral reef fisheries to climate change. PloS ONE 8:e74321.

Clark, W. G. 2002. F 35% revisited ten years later. North American Journal of Fisheries Management 22:251–257.

Coggins, L. G., W. E. Pine, C. J. Walters, and S. J. D. Martell. 2006. Age-structured mark-recapture analysis: A virtual-population-analysis-based model for analyzing age-structured sapture-recapture data. North American Journal of Fisheries Management 26:201–205.

Coleman, F. C., W. F. Figueira, J. S. Ueland, and L. B. Crowder. 2004. The Impact of United States Recreational Fisheries on Marine Fish Populations. Science 305:2002–2004.

Cook, R. M. 1995. A simple model for the analysis of research vessel data to determine stock trends. ICES C.M. D:12.

Cope, J. M. 2013. Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. Fisheries Research 142:3–14.

Cope, J. M. 2017. Natural Mortality Tool. GitHub Repository. https://github.com/shcaba/Natural-Mortality-Tool.

Cope, J. M., and M. Key. 2009. Status of cabezon (Scorpaenichthys marmoratus) in California waters as assessed in 2005. Page Pacific Fishery Management Council.

Cope, J. M., and A. E. Punt. 2009. Length-based reference points for data-limited situations: Applications and restrictions. Marine and Coastal Fisheries 1:169–186.

Cope, J. M., J. T. Thorson, C. R. Wetzel, and J. DeVore. 2015. Evaluating a prior on relative stock status using simplified age-structured models. Fisheries Research 171:101–109.

Cordue, P.L. 2012. Fishing intensity metrics for use in overfishing determination. ICES Journal of Marine Science 69:615-623.

Costello, C., D. Ovando, T. Clavelle, C.K. Strauss, R. Hilborn, M.C. Melnychuk, T.A. Branch, S.D. Gaines, C.S. Szuwalski, R.B. Cabral, D.N. Rader, and A. Leland. 2016. Global fishery prospects under contrasting management regimes. Proceedings of the National Academy of Sciences 113:5125–5129.

Costello, C., D. Ovando, R. Hilborn, S. D. Gaines, O. Deschenes, and S. E. Lester. 2012. Status and solutions for the world’s unassessed fisheries. Science 338:517–520.

Cressey, D. 2015. Eyes on the ocean. Nature 519:280–282.

Cummings, N., S. Sagarese, and Q. C. Huynh. 2016. An alternative approach to setting annual catch limits for data- limited fisheries: Use of the DLMtool and mean length estimator for six US Caribbean stocks. SEDAR46-RW-03. North Charleston, SC.

Cunningham, S., A. E. Neiland, M. Arbuckle, and T. Bostock. 2009. Wealth-based fisheries management: using fisheries wealth to orchestrate sound fisheries policy in practice. Marine Resource Economics 24:271–287.

Darcy, G. H., and G. C. Matlock. 1999. Application of the precautionary approach in the national standard guidelines for conservation and management of fisheries in the United States. ICES Journal of Marine Science 56:853–859.

Darwall, W. R. T., and E. H. Allison. 2002. Monitoring, assessing, and managing fish stocks in Lake Malawi/Nyassa: Current approaches and future possibilities. Aquatic Ecosystem

Health and Management 5:293–305.

Davies, R. W. D., S. J. Cripps, a. Nickson, and G. Porter. 2009. Defining and estimating global marine fisheries bycatch. Marine Policy 33:661–672.

Dick, E. J., and A. D. MacCall. 2011. Depletion-Based Stock Reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research 110:331–341.

Die, D. J., and J. F. Caddy. 1997. Sustainable yield indicators from biomass: Are there appropriate reference points for use in tropical fisheries? Fisheries Research 32:69–79.

Dowling, N. A., C. M. Dichmont, M. Haddon, D. C. Smith, A. D. M. Smith, and K. Sainsbury. 2015. Empirical harvest strategies for data-poor fisheries: A review of the literature. Fisheries Research 171:141–153.

Dowling, N. A., J. R. Wilson, M. B. Rudd, E. A. Babcock, D. Dougherty, and M. Gleason. 2016. FishPath: A Decision Support System for Assessing and Managing Data- and Capacity-Limited Fisheries. Pages 59–96*in* T. J. I. Quinn, J. L. Armstrong, M. R. Baker, J. Heifetz, and D. Witherell, editors.Assessing and managing data-limited fish stocks. Alaska Sea Grant, University of Alaska Fairbanks.

Ehrhardt, N. M., and J. S. Ault. 1992. Analysis of Two Length-Based Mortality Models Applied to Bounded Catch Length Frequencies. Transactions of the American Fisheries Society 121:115–122.

Fisheries and aquaculture software. 2017. FISAT II - FAO-ICLARM Stock Assessment Tool. FAO Fisheries and Aquaculture Department, Rome.

Food and Agriculture Organization. 2014. The State of the World Fisheries and Aquaculture: opportunities and challenges. Rome.

Food and Agriculture Organization. 2016. The State of World Fisheries and Aquaculture 2016. Contributing to food security and nutrition for all. Rome.

Fournier, D. A., J. R. Sibert, J. Majkowski, and J. Hampton. 1990. MULTIFAN a Likelihood-Based Method for Estimating Growth Parameters and Age Composition from Multiple Length Frequency Data Sets Illustrated using Data for Southern Bluefin Tuna (Thunnus maccoyii). Canadian Journal of Fisheries and Aquatic Sciences 47:301–317.

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software 27:233–249.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68:2228–2228.

Francis, R. I. C. C. 2014. Replacing the multinomial in stock assessment models: A first step. Fisheries Research 151:70–84.

Froese, R. 2004. Keep it simple : three indicators to deal with overfishing:86–91.

Froese, R., and D. Pauly. 2017. FishBase. www.fishbase.org.

Fujita, R., D. J. Thornhill, K. Karr, C. H. Cooper, and L. E. Dee. 2014. Assessing and managing data-limited ornamental fisheries in coral reefs. Fish and Fisheries 15:661–675.

Gabriel, W. L., M. P. Sissenwine, and W. J. Overholtz. 1989. Analysis of Spawning Stock Biomass per Recruit: An Example for Georges Bank Haddock. North American Journal of Fisheries Management 9:383–391.

Garibaldi, L. 2012. The FAO global capture production database: A six-decade effort to catch the trend. Marine Policy 36:760–768.

Gedamke, T., and J. M. Hoenig. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. Transactions of the American Fisheries Society 135:476–487.

Gislason, H., N. Daan, J. C. Rice, and J. G. Pope. 2010. Size, growth, temperature and the natural mortality of marine fish. Fish and Fisheries 11:149–158.

Goodyear, C. P. 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. Pages 67–81*in* S. J. Smith, J. J. Hunt, and D. Rivard, editors.Risk Evaluation and Biological Reference Points for Fisheries Management. Canadian Special Publication on Fisheries and Aquatic Science.

Groeneveld, J. C. 2003. Under-reporting of catches of South Coast rock lobster (Palinurus gilchristi), with implications for the assessment and management of the fishery. African Journal of Marine Science 25:407–411.

Gulbrandsen, L. H. 2009. The emergence and effectiveness of the Marine Stewardship Council. Marine Policy 33:654–660.

Hall, S. J., R. Hilborn, N. L. Andrew, and E. H. Allison. 2013. Innovations in capture fisheries are an imperative for nutrition security in the developing world. Proceedings of the National Academy of Sciences of the United States of America 110:8393–8.

Hammond, T. R., and V. M. Trenkel. 2005. Censored catch data in fisheries stock assessment. ICES Journal of Marine Science 62:1118–1130.

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

Heemstra, P. C., and J. E. Randall. 1993. An annotated and illustrated calalogue of the grouper, rockcod, hind, coral grouper, and lyretail species known to date. Page 382 pp. FAO Species Catalogue. Vol. 16. Rome.

Hicks, A. C., and C. Wetzel. 2011. The status of Dover sole (Microstomus pacificus) along the US West Coast in 2011. Page Pacific Fishery Management Council. Portland, OR.

Hicks, C. C., and T. R. McClanahan. 2012. Assessing gear modifications needed to optimize yields in a heavily exploited, multi-species, seagrass and coral reef fishery. PloS ONE 7:e36022.

Hilborn, R. 2001. Calculation of biomass trend, exploitation rate, and surplus production from survey and catch data. Canadian Journal of Fisheries and Aquatic Sciences 58:579–584.

Hilborn, R., J. M. L. Orensanz, and A. M. Parma. 2005. Institutions, incentives and the future of fisheries. Philosophical Transactions of the Royal Society of London. Series B, Biological sciences 360:47–57.

Hilborn, R, and D. Ovando. 2014. Reflections on the success of traditional fisheries management. ICES Journal of Marine Science 71:1040-1046.

Hilborn, R., and J. Sibert. 1988. Adaptive management of developing fisheries. Marine Policy 12:112–121.

Hilborn, R., and C. J. Walters. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. Springer.

Hordyk, A., K. Ono, S. Valencia, N. Loneragan, and J. Prince. 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 Journal of Marine Science 72:217–231.

Hurtado-Ferro, F., C. S. Szuwalski, J. L. Valero, S. C. Anderson, C. J. Cunningham, K. F. Johnson, R. Licandeo, C. R. McGilliard, C. C. Monnahan, M. L. Muradian, K. Ono, K. A. Vert-Pre, A. R. Whitten, and A. E. Punt. 2014. Looking in the rear-view mirror: Bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science 72:99–110.

Ichinokawa, M., H. Okamura, and Y. Takeuchi. 2014. Data conflict caused by model mis-specification of selectivity in an integrated stock assessment model and its potential effects on stock status estimation. Fisheries Research 158:147–157.

Jensen, A. L. 1996. Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. Canadian Journal of Fisheries and Aquatic Sciences 53:820–822.

Jensen, A. L. 1997. Origin of the relation between K and Linf and the synthesis of relations among life history parameters. Canadian Journal of Fisheries and Aquatic Sciences 54:987–989.

Johnson, K. F., M. B. Rudd, M. Pons, C. A. Akselrud, Q. Lee, M. A. Haltuch, and O. S. Hamel. 2016. Status of the U.S. sablefish resource in 2015.

Kelleher, K. 2005. Discards in the world’s marine fisheries: An update. FAO Fisheries Technical Paper No. 470:131.

Kenchington, T. J. 2014. Natural mortality estimators for information-limited fisheries. Fish and Fisheries 15:533–562.

Kittinger, J. N., E. M. Finkbeiner, E. W. Glazier, and L. B. Crowder. 2012. Human Dimensions of Coral Reef Social-Ecological Systems. Ecology and Society 17:17.

Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. Bell. 2016. TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software 70:1–21.

Lee, H. H., K. R. Piner, R. D. Methot, and M. N. Maunder. 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. Fisheries Research 158:138–146.

MacCall, A.D. 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES Journal of Marine Science 66:2267–2271.

MacCall, A.D. 2013. Use of the delta method to evaluate the precision of assessments that fix parameter values. Fisheries Research 142:56-60.

Mace, P. M., and M. P. Sissenwine. 1993. How much spawning per recruit is enough? Pages 101–118Risk Evaluation and Biological Reference Points for Fisheries Management. Canadian Special Publication on Fisheries and Aquatic Science.

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

Magnusson, A., A. E. Punt, and R. Hilborn. 2013. Measuring uncertainty in fisheries stock assessment: The delta method, bootstrap, and MCMC. Fish and Fisheries 14:325–342.

Mahon, R. 1997. Does fisheries science serve the needs of managers of small stocks in developing countries. Canadian Journal of Fisheries and Aquatic Sciences 54:2207–2213.

Martell, S., and R. Froese. 2013. A simple method for estimating MSY from catch and resilience. Fish and Fisheries 14:504–514.

Maunder, M. N., and K. R. Piner. 2017. Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. Fisheries Research.

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

Mayo, R.K., J.K.T. Brodziak, J.M. Burnett, M.L. Traver, and L.A. Col. 2007. The 2005 assessment of Acadian redfish, *Sebastes fasciatus* Storer, in the Gulf of Maine/Georges Bank Region. U.S. Department of Commerce, *Northeast Fisheries Science Center Reference Document* 07-06, 32 p.

McClanahan, T., E. H. Allison, and J. E. Cinner. 2013. Managing fisheries for human and food security. Fish and Fisheries 16:78–103.

McClanahan, T. R., N. A. Graham, M. A. MacNeil, and J. E. Cinner. 2015. Biomass-based targets and the management of multispecies coral reef fisheries. Conservation Biology 29:409–417.

Methot, R. D., G. R. Tromble, D. M. Lambert, and K. E. Greene. 2014. Implementing a science-based system for preventing overfishing and guiding sustainable fisheries in the United States. ICES Journal of Marine Sciences 71:183–194.

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

Metuzals, K., C. Wernerheim, R. Haedrich, P. Copes, and A. Murrin. 2008. Data Fouling in Newfoundland's Marine Fisheries. Pages 121–137Making and Moving Knowledge: Interdisciplinary and Community-based Research in a World on the Edge. McGill-Queens University Press, Montreal.

Minte-Vera, C.V., M.N. Maunder, A.M. Aires-de-Silva, K. Satoh, and K. Uosaki. 2017. Get the biology right, or use size-composition data at your own risk. Fisheries Research 192:114-125.

Monnahan, C. C., K. Ono, S. C. Anderson, M. B. Rudd, A. C. Hicks, F. Hurtado-Ferro, K. F. Johnson, P. T. Kuriyama, R. R. Licandeo, C. C. Stawitz, I. G. Taylor, and J. L. Valero. 2016. The effect of length bin width on growth estimation in integrated age-structured stock assessments. Fisheries Research 180:103–112.

Monnahan, C.C., J.T. Thorson, and T.A. Branch. 2017. Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. Methods in Ecology and Evolution 8:339-348

Morris, J. A., K. W. Shertzer, and J. A. Rice. 2010. A stage-based matrix population model of invasive lionfish with implications for control. Biological Invasions 13:7–12.

Munday, P. L., A. L. Hodges, J. H. Choat, and N. Gust. 2004. Sex-specific growth effects in protogynous hermaphrodites. Canadian Journal of Fisheries and Aquatic Sciences 61:323–327.

Nadon, M. O., J. S. Ault, I. D. Williams, S. G. Smith, and G. T. DiNardo. 2015. Length-Based Assessment of Coral Reef Fish Populations in the Main and Northwestern Hawaiian Islands. PloS ONE 10:e0133960.

Needle, C. L. 2014. Using self-testing to validate the SURBAR survey-based assessment model. Fisheries Research 171:78–86.

Neis, B. 1992. Fishers' Ecological Knowledge and Stock Assessment in Newfoundland.

Newman, K. B., C. Fernández, L. Thomas, and S. T. Buckland. 2009. Monte carlo inference for state-space models of wild animal populations. Biometrics 65:572–83.

Nielsen, A., and C. W. Berg. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research 158:96–101.

Ntiba, M. J., and V. Jaccarini. 1988. Age and growth parameters of *Siganus sutor* in Kenyan marine inshore water, derived from numbers of otolith microbands and fish lengths. Journal of Fish Biology 33:465–470.

Ntiba, M. J., and V. Jaccarini. 1990. Gonad maturation and spawning times of *Siganus sutor* off the Kenya coast: evidence for definite spawning seasons in a tropical fish. Journal of Fish Biology 37:315–325.

Omori, K.L., J.M. Hoenig, M.A. Luehring, K. Baier-Lockhart. 2016. Effects of underestimating catch and effort on surplus production models. Fisheries Research 183:138-145.

Ono, K., R. Licandeo, M. L. Muradian, C. J. Cunningham, S. C. Anderson, F. Hurtado-ferro, K. F. Johnson, C. R. Mcgilliard, C. C. Monnahan, C. S. Szuwalski, J. L. Valero, K. A. Vert-pre, and A. R. Whitten. 2015. The importance of length and age composition data in statistical age-structured models for marine species. ICES Journal of Marine Science 72:31–43.

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

Patrick, W. S., P. Spencer, J. Link, J. Cope, J. Field, D. Kobayashi, P. Lawson, T. Gedamke, E. Cortes, O. Ormseth, K. Bigelow, and W. Overholtz. 2009. Using productivity and susceptibility indices to assess the vulnerability of United States fish stocks to overfishing. Fishery Bulletin 108:305–322.

Patterson, K. 1998. Assessing fish stocks when catches are misreported: model, simulation tests, and application to cod, haddock, and whiting in the ICES area. ICES Journal of Marine Science 55:878–891.

Pauly, D. 1987. A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic invertebrates. Pages 7–34*in* D. Pauly and G. R. Morgan, editors.Length-based methods in fisheries research. International Center for Living Aquatic Resources Management, Makati, Metro Manila, Philippines.

Pauly, D. 1998. Rationale for reconstructing catch time series. European Community Fisheries Cooperation Bulletin 11:4–7.

Pauly, D., D. Belhabib, R. Blomeyer, W. W. W. L. Cheung, A. M. Cisneros-Montemayor, D. Copeland, S. Harper, V. W. Y. Lam, Y. Mai, F. Le Manach, H. Österblom, K. M. Mok, L. van der Meer, A. Sanz, S. Shon, U. R. Sumaila, W. Swartz, R. Watson, Y. Zhai, and D. Zeller. 2014. China’s distant-water fisheries in the 21st century. Fish and Fisheries 15:474–488.

Pauly, D., and R. Froese. 2012. Comments on FAO’s state of fisheries and aquaculture, or “SOFIA 2010.” Marine Policy 36:746–752.

Pauly, D., and D. Zeller. 2016. Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. Nature Communications 7:10244.

Pauly, D., and D. Zeller. 2017. Comments on FAOs State of World Fisheries and Aquaculture (SOFIA 2016). Marine Policy 77:176–181.

Pilling, G.M., G.P. Kirkwood, and S.G. Walker. 2002. An improved method for estimating individual growth variablity in fish, and the correlation between von Bertalanffy growth parameters. Canadian Journal of Fisheries and Aquatic Sciences 59:424-432.

Pitcher, T. J., R. Watson, R. Forrest, H. P. Valtysson, and S. Guenette. 2002. Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish and Fisheries 3:317–339.

Plagányi, É., D. Butterworth, and M. Burgener. 2011. Illegal and unreported fishing on abalone—Quantifying the extent using a fully integrated assessment model. Fisheries Research 107:221–232.

Prince, J., A. Hordyk, S. R. Valencia, N. Loneragan, and K. Sainsbury. 2015a. Revisiting the concept of Beverton-Holt life-history invariants with the aim of informing data-poor fisheries assessment. ICES Journal of Marine Science 72:194–203.

Prince, J., S. Victor, V. Kloulchad, and A. Hordyk. 2015b. Length based SPR assessment of eleven Indo-Pacific coral reef fish populations in Palau. Fisheries Research 171:42–58.

Quinn, T. J. I., J. L. Armstrong, M. Baker, J. Heifetz, and D. Witherell. 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.

RAM Legacy Stock Assessment Database. 2017. www.ramlegacy.org. Accessed 15 June 2017.

Ricard, D., C. Minto, O. P. Jensen, and J. K. Baum. 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish and Fisheries 13:380–398.

Rojas, M. J. R. 1996. Fecundidad y épocas de reproducción del “pargo mancha” Lutjanus guttatus (Pisces: Lutjanidae) en el Golfo de Nicoya, Costa Rica. Revista de Biología Tropical 44/45:477–487.

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

Rosenberg, A. A., K. M. Kleisner, J. Afflerbach, S. C. Anderson, M. Dickey-Collas, A. B. Cooper, M. J. Fogarty, E. A. Fulton, N. L. Gutiérrez, K. J. W. Hyde, E. Jardim, O. P. Jensen, T. Kristiansen, C. Longo, C. V. Minte-Vera, C. Minto, I. Mosqueira, G. C. Osio, D. Ovando, E. R. Selig, J. T. Thorson, J. C. Walsh, and Y. Ye. 2017. Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries Around the World. Conservation Letters.

Ruckelshaus, M., S. C. Doney, H. M. Galindo, J. P. Barry, F. Chan, J. E. Duffy, C. a. English, S. D. Gaines, J. M. Grebmeier, a. B. Hollowed, N. Knowlton, J. Polovina, N. N. Rabalais, W. J. Sydeman, and L. D. Talley. 2013. Securing ocean benefits for society in the face of climate change. Marine Policy 40:154–159.

Rudd, M. B., R. N. M. Ahrens, W. E. I. Pine, and S. K. Bolden. 2014. Empirical, spatially explicit natural mortality and movement rate estimates for the threatened Gulf sturgeon (Acipenser oxyrinchus desotoi). Canadian Journal of Fisheries and Aquatic Sciences 71:1407–1417.

Rudd, M. B., and T. A. Branch. 2017. Does unreported catch lead to overfishing? Fish and Fisheries 18:313–323.

Sabater, M., and P. Kleiber. 2014. Augmented catch-MSY approach to fishery management in coral-associated fisheries. Pages 199–218*in* S. A. Bortone, editor.Interrelationships

Between Corals and Fisheries. CRC Press.

Salas, S., R. Chuenpagdee, J. C. Seijo, and A. Charles. 2007. Challenges in the assessment and management of small-scale fisheries in Latin America and the Caribbean. Fisheries Research 87:5–16.

Schnute, J. T., and R. Haigh. 2007. Compositional analysis of catch curve data, with an application to Sebastes maliger. ICES Journal of Marine Science 64:218–233.

Seber, G. A. F. 1973. The Estimation of Animal Abundance and Related Parameters. Griffin, London.

Shin, Y., M. Rochet, S. Jennings, J. Field, and H. Gislason. 2005. Using size-based indicators to evaluate the ecosystem effects of fishing. ICES Journal of Marine Science 62:384–396.

Stawitz, C. C., F. Hurtado-Ferro, P. Kuriyama, J. T. Trochta, K. F. Johnson, M. A. Haltuch, and O. S. Hamel. 2016. Stock Assessment Update: Status of the U. S. petrale sole resource in 2014. Page Pacific Fishery Management Council.

Stewart, I.J. and S.J.D. Martell. 2015. Reconciling stock assessment paradigms to better inform fisheries management. ICES Journal of Marine Science 72:2187-2196.

Szuwalski, C. S., K. A. Vert-Pre, A. E. Punt, T. A. Branch, and R. Hilborn. 2015. Examining common assumptions about recruitment: A meta-analysis of recruitment dynamics for worldwide marine fisheries. Fish and Fisheries 16:633–648.

Then, A. Y., J. M. Hoenig, T. Gedamke, and J. S. Ault. 2015a. Comparison of Two Length-Based Estimators of Total Mortality: A Simulation Approach. Transactions of the American Fisheries Society 144:1206–1219.

Then, A. Y., J. M. Hoenig, N. G. Hall, and D. A. Hewitt. 2015b. 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.

Thorson, J. T., T. A. Branch, and O. P. Jensen. 2012. Using model-based inference to evaluate global fisheries status from landings, location, and life history data. Canadian Journal of Fisheries and Aquatic Sciences 69:645–655.

Thorson, J. T., and J. M. Cope. 2014. Catch curve stock-reduction analysis: An alternative solution to the catch equations. Fisheries Research 171:33–41.

Thorson, J.T., J.M. Cope, and W.S. Patrick. 2014a. Assessing the quality of life history information in publicly available databases. Ecological Applications 24:217–226.

Thorson, J. T., A. C. Hicks, and R. D. Methot. 2015a. Random effect estimation of time-varying factors in Stock Synthesis. ICES Journal of Marine Science 72:178–185.

Thorson, J. T., O. P. Jensen, and E. F. Zipkin. 2014b. How variable is recruitment for exploited marine fishes? A hierarchical model for testing life history theory. Canadian Journal of Fisheries and Aquatic Sciences 71:973–983.

Thorson, J. T., K. F. Johnson, R. D. Methot, and I. G. Taylor. 2016. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fisheries Research.

Thorson, J. T., and C. Minto. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES Journal of Marine Science 72:1245–1256.

Thorson, J. T., C. Minto, C. V Minte-Vera, K. M. Kleisner, and C. Longo. 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. Canadian Journal of Fisheries and Aquatic Sciences 70:1829–1844.

Thorson, J.T., H. Skaug, K. Kristensen, A. Shelton, E. Ward, J. Harms, and J. Benante. 2015b. The importance of spatial models for estimating the strength of density dependence. Ecology 96:1202–1212.

Thorson, J. T., and C. Wetzel. 2015. The status of canary rockfish (Sebastes pinniger) in the California Current in 2015. Page Pacific Fishery Management Council.

de Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and obervation error. Ecological Monographs 72:57–76.

Vert-pre, K. A., R. O. Amoroso, O. P. Jensen, and R. Hilborn. 2013. Frequency and intensity of productivity regime shifts in marine fish stocks. Proceedings of the National Academy of Sciences of the United States of America 110:1779–84.

Walters, C. J., S. J. D. Martell, and J. Korman. 2006. A stochastic approach to stock reduction analysis. Canadian Journal of Fisheries and Aquatic Sciences 63:212–223.

Watson, R.A. 2017. A database of global marine commercial, small-scale, illegal and unreported fisheries catch 1950–2014. Scientific Data 4:170039

Watson, R., and D. Pauly. 2001. Systematic distortions in world fisheries catch trends. Nature 414:534–536.

Wetzel, C. R., and A. E. Punt. 2011. Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. Fisheries Research 110:342–355.

Weyl, O. L. F., M. M. Manase, W. Namoto, and P. L. Banda. 1999. Assessment of the artisanal fishery catch- and effort- data collection and analysis procedure in Mangochi District, Malawi. NARMAP, Malawi.

Worm, B., E. B. Barbier, N. Beaumont, J. E. Duffy, C. Folke, B. S. Halpern, J. B. C. Jackson, H. K. Lotze, F. Micheli, S. R. Palumbi, E. Sala, K. a Selkoe, J. J. Stachowicz, and R. Watson. 2006. Impacts of biodiversity loss on ocean ecosystem services. Science 314:787–790.

Worm, B., R. Hilborn, J. K. Baum, T. A. Branch, J. S. Collie, C. Costello, M. J. Fogarty, E. A. Fulton, J. A. Hutchings, S. Jennings, O. P. Jensen, H. K. Lotze, P. M. Mace, T. R. McClanahan, C. Minto, S. R. Palumbi, A. M. Parma, D. Ricard, A. A. Rosenberg, R. Watson, and D. Zeller. 2009. Rebuilding global fisheries. Science 325:578–585.

Ye, Y., M. Barange, M. Beveridge, L. Garibaldi, N. Gutierrez, A. Anganuzzi, and M. Taconet. 2017. FAO's statistical databases and the sustainability of fisheries and aquaculture: Comments on Pauly and Zeller 2017. Marine Policy:1–5.

Zeller, D., S. Booth, G. Davis, and D. Pauly. 2007. Re-estimation of small-scale fishery catches for U.S. flag-associated island areas in the western Pacific: the last 50 years. Fisheries Builletin 105:266–277.

Zeller, D., M. Darcy, S. Booth, M. K. Lowe, and S. Martell. 2008. What about recreational catch? Potential impact on stock assessment for Hawaii’s bottomfish fisheries. Fisheries Research 91:88–97.

Zhou, S., A. J. Hobday, C. M. Dichmont, and A. D. M. Smith. 2016. Ecological risk assessments for the effects of fishing: A comparison and validation of PSA and SAFE. Fisheries Research 183:518–529.

VITA

Merrill Rudd was born and grew up in Short Hills, New Jersey. She obtained her B.A. in Environmental Studies – Ecology track at Washington University in St. Louis, Missouri. After participating in the NOAA Research-Training-Recruitment marine resource population dynamics undergraduate workshop during her last semester at Washington University, she decided to pursue fisheries science and stock assessment via an M.S. degree at the University of Florida, studying Gulf sturgeon population dynamics with Drs. Rob Ahrens and Bill Pine. After earning her M.S. Merrill moved to Seattle to obtain her doctorate at the University of Washington's School of Aquatic and Fishery Sciences under the supervision of Ray Hilborn and Trevor Branch, funded by the NSF IGERT Program on Ocean Change. Merrill's next step will continue applying the work from her PhD towards the assessment of U.S. data-limited fisheries, and branch out into different types of stock assessments for fisheries around the world.