The impact of data, modeling approaches, and control rules on the performance of management strategies: applications to West Coast groundfish fishery

Chantell R. Wetzel

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

André Punt, Chair

Tim Essington

Ian Stewart

Program Authorized to Offer Degree:
School of Aquatic and Fishery Sciences

University of Washington

Abstract

The impact of data, model approaches, and control rules on the performance of management strategies: applications to West Coast groundfish fishery

Chantell R. Wetzel

Chair of the Supervisory Committee:
Professor André Punt
School of Aquatic and Fishery Sciences

The foundation of sustainable fisheries management is based on setting catch limits that will maintain stock biomasses at or near levels that produce maximum sustainable yield. Determining appropriate harvest rates and the resulting catches depends on the ability of available data and estimation methods to accurately estimate productivity, unfished biomass, and current biomass. However, uncertainty surrounds the ability to estimate these quantities accurately. Incorporating uncertainty into management advice can improve the ability of fishery managers to achieve their objectives, decreasing the risk of overfishing and increasing the probability of sustainably managing fish stocks.

This dissertation examines the impact of data availability to accurately estimate biological parameters and hence stock biomass, alternative estimation methods to set harvest limits that prevent overfishing, and the impact of differing management actions to rebuild and maintain stocks at target biomass levels. Simulation was used to address questions concerning data, modeling, and management decisions by creating a 'true' population to measure and evaluate outcomes. Chapter 1 highlights the importance of continued data collection when fish stocks are rebuilding from low biomasses. Retaining data collection at historical levels during rebuilding allowed for improved parameter estimation, which resulted in reduced variability in estimated stock size with larger average catches during rebuilding.

In contrast, when data collection intensity was reduced during rebuilding, the estimates of relative stock size become more variable between assessments, resulting in stocks being prematurely declared rebuilt to the target biomass. Estimates of stock productivity at the time of the first assessment were poor, but continued data collection during rebuilding allowed for improved estimates of steepness. However, when data were limited estimates of productivity were highly variable during rebuilding. Chapter 2 evaluated the impact of data on assessment performance by examining estimation methods designed for application when data limitations prevent the use of complex estimation approaches. Overall, the application of the catch-only (data-limited) and data-moderate estimation methods coupled with management buffers to reduce harvest had mixed results, failing to consistently protect stocks from experiencing overfishing. However, two methods were identified that performed well for rockfish stocks that prevented overfishing when the population had experienced limited historical exploitation. Specific recommendations are made for management of US west coast groundfish and when it might or might not be appropriate to apply such methods.

The final two chapters each performed a management strategy evaluation that examined alternative actions to meet the goals of managers and to explicitly identify the trade-offs among different approaches. Chapter 3 simulated fish stocks that were in an overfished state, below the management-defined minimum stock size threshold, and implemented alternative actions designed to allow rebuilding to the target biomass. Rebuilding approaches that applied precaution when determining rebuilding harvest rates, buffering against future uncertainty, were less responsive to noise in assessment estimates, had fewer changes in harvest rates during rebuilding, and successfully rebuilt stocks on time. The final chapter examined the trade-offs achieved by alternative harvest control rules, the ability of each strategy to maintain stocks at or near target biomass, and prevent stocks from becoming overfished for US west coast flatfish. All harvest control rules examined were effective at maintaining stocks at or near target biomasses and preventing stocks from declining below minimum stock size thresholds. However, trade-offs are associated with each harvest control rule. The harvest control rules that involved the highest proxy biomass targets resulted in higher probabilities of the relative spawning biomass being within 10% of the target biomass, with lower average annual variation in catch, but also resulted in the lowest average catches over the last twenty-five years of the management period. The more aggressive harvest control rules that involved lower proxy biomasses, resulted in higher average catches, but had higher average annual variation in catch and lower probabilities of the stock biomass being within 10% of the proxy biomass.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I have been fortunate to have many people that have encouraged and supported me throughout my graduate work and without them none of this would have been possible. I would like to start by thanking my advisor, André Punt. I was incredibly lucky to have him as my major advisor. Through his direction I have grown immensely and will forever be thankful to him.

The members of my committee have been instrumental in the creation of this thesis; Ray Hilborn, Tim Essington, Trevor Branch, and Ian Stewart. My committee has helped my remember the bigger picture, step beyond the math and models in order to create meaningful work by providing valuable feedback and guidance that improved this body of work.

Thank you to the Northwest Fisheries Science Center Fisheries Resource Analysis and Monitoring Division that generously supported me throughout my graduate work. I am so grateful to be a member of this team, which has allowed me to tackle important research questions and given me the opportunity to have invaluable hands-on experience in the field of fishery stock assessment.

My fellow lab-mates in the Punt lab, past and present, have truly enriched my graduate experience I am so thankful to each of them for their friendship, meaningful discussions, and for the laughs, which at times, were sorely needed.

I am grateful for my family and friends, to whom I am forever indebted to. A special thank you to my parents for their love and steadfast support that has made me the person I am. Without them, I would be lost. I would like to thank my little sister for being my friend, despite the horrors she endured at my hands growing up.

Finally, I would like to say thank you to my husband, Ryan, for not allowing me to take life too seriously and being my rock. Despite the jokes concerning my "life-time" student status, you always believed in me. I lack the words to express my gratitude and love.

DEDICATION

to

my family

INTRODUCTION

The concept of maximum sustainable yield (MSY) from fish stocks was first introduced in the 1930s (Russell, 1931; Hjort et al., 1933; Graham, 1935). The theory behind MSY within fishery science states as stocks are fished down, a surplus yield, an amount beyond the replacement biomass, is available for exploitation, and there exists a biomass at which the productivity of a stock would be maximized (B_{MSY}) and a fishing rate that would result in maximum yield (F_{MSY}). Managing stocks to obtain MSY would maximize the harvest obtained from fisheries, while preventing stocks from becoming depleted to biomass sizes well below B_{MSY}. These tenants have become the basis for legislation that governs fisheries management.

US federal fisheries are governed by the Magnuson Stevens Fishery Conservation and Management Act (MSA) and the Sustainable Fisheries Act (SFA). National Standard 1, included in the SFA, requires management measures be implemented to prevent overfishing and to achieve optimum yield (Sustainable Fisheries Act, 1996). The MSA created Regional Fishery Management Councils to govern federal fisheries within the US to meet these goals. The Fishery Management Councils have employed management measures defined within Fishery Management Plans (FMP). An FMP outlines the fish stocks under management, how decisions for management are developed, management targets, and measures to prevent overfishing. The Councils set biomass targets, target harvest rates, and biomass limits based on the concepts of B_{MSY} and MSY. However, fishery managers have often employed proxy values for B_{MSY} and F_{MSY} given the difficulty of estimating these quantities with certainty for individual stocks. Proxies have been developed to simplify management due to the logistic challenges of estimating and managing to B_{MSY} and F_{MSY} for each fish stock individually (especially for stocks where limited data prevents reliable estimation) while still allowing fisheries to obtain close to maximum benefit while preventing overfishing. US fisheries managed using proxies apply a B_{MSY} proxy defined in terms of biomass relative to unfished biomass termed B_{PROXY} and an F_{MSY} proxy harvest rate termed F_{PROXY}. Harvest is a function of biomass and the F_{PROXY}, termed the Overfishing Limit (OFL), a level of harvest that if exceeded would constitute overfishing and defines the maximum harvest level. Additionally, Fishery Management Councils define a minimum stock size threshold (MSST), a biomass level less than B_{PROXY}, which is the threshold level that if a stock biomass fell below would result in an overfished declaration, requiring a formal plan with a timeline to rebuild to B_{PROXY}.

The Pacific Fishery Management Council (PFMC) was established in 1982 to manage federal fisheries off the US west coast. The PFMC groundfish FMP defines three goals for fisheries management; 1) conservation (preventing overfishing and rebuilding overfished stocks through appropriate harvest levels), 2) economics (maximizing the value of the resource as a whole to users and communities), and 3) utilization (attaining the maximum biological yield to the overall groundfish fishery while meeting the constraints of the first goal and providing year-round availability of seafood) (Pacific Fishery Management Council, 2016). The complexity of the stocks off the US west coast and the fishery has prevented the PFMC from being completely successful in meeting each of these goals in tandem. The groundfish FMP contains 80+ species of groundfish that are highly diverse in life history and productivity; of which only approximately a third have been formally assessed (Pacific Fishery Management Council, 2016). Specifically, rockfish stocks (*Sebastes* spp.), which comprise a large portion of the stocks within the groundfish FMP, have presented complex challenges for sustainable management. Rockfish species have slow population dynamics, often with low intrinsic rates of increase, high longevities, and with relatively large unexploited standing stock biomasses that supported the development of target fisheries, but their sustainability given the known biology was questionable (Francis, 1986). By 2002, groundfish catches had been dramatically reduced relative to historical levels, the fishery had been deemed a federal economic disaster for coastal communities (Conway and Shaw, 2008), nine stocks had been declared overfished, and managers had introduced large closed areas to protect overfished rockfish species. Although rockfish species make up a large proportion of the stocks within the groundfish FMP, there are other species with alternative life histories (e.g. flatfish, shark and skates) that need to be considered in tandem when setting management actions. The PFMC grouped species within the FMP based on family groups or life history similarity; rockfish, flatfish, roundfish, and sharks and skates. Each group are managed using specific proxies for target biomass and target harvest rate (based on spawning biomass-per-recruit, F_{SPR}), determining the OFL given the stock size.

Given the complicated history of the US west coast groundfish fishery, the PFMC has attempted to apply a precautionary approach in recent years when setting harvest limits to prevent overfishing and to allow overfished stocks to rebuild as fast as possible, accounting for the societal and economic needs of the fishery. However, the unobservable nature of fisheries results in an inherent uncertainty about the true state of the population, and has led to the development of computer simulation to evaluate management actions, assessment estimation methods, and differing assumptions concerning population dynamics. Simulation is a useful tool for evaluating the performance of complex models. Simulation studies utilize a process whereby data are generated using a specific population dynamics model, termed an operating model, to represent a theoretical fish population (Peck, 2004). The main advantage to this type of work is that the "truth" is known and the experimenter can ask complex questions of the system, allowing for an increased understanding for managers and scientists related to the potential implications of policy actions and assessment assumptions (e.g. Smith et al., 1993; Butterworth and Punt, 1999). Specifically, simulation has proved to be a useful way to evaluate research questions regarding US west coast groundfish, providing fishery managers guidance on a variety of topics; e.g. modeling selective fisheries mortality (Taylor and Methot, 2013), the performance of alternative reference points for determining target stock sizes and harvest (Haltuch et al., 2008), and the influence of environmental conditions to maintain stocks at or near target biomass (Haltuch and Punt, 2011).

Extending the application of simulation to evaluate long-term implications, the potential trade-offs, and the likelihood of meeting management goals with current management actions led to the development of Management Strategy Evaluation (MSE; Smith, 1994). MSE was designed as a way to formally evaluate management actions and understand their potential biological and fishery implications. MSE uses simulation testing based on modeling with feedback to characterize the potential outcomes from management decisions. The International Pacific Halibut Commission and the International Whaling Commission were among the first management bodies to implement this approach to determine management strategies that would meet their Commission's goals (Southward, 1968; International Whaling Commission, 1992b; Punt and Donovan, 2007). Since then, MSEs have been widely used in fisheries management; e.g., South Africa (Punt and Butterworth, 1995), New Zealand (Starr et al., 1997), Australia (Punt and Smith, 1999), and the United States (Punt and Donovan, 2007; A'mar et al., 2009; International Joint Technical Committee for Pacific Hake, 2014).

The core framework of an MSE involves an operating model to represent the biological component (or system) of interest. The operating model can vary in complexity to account for multiple sources of error and uncertainty (process, observation, implementation, and model uncertainty) (Butterworth and Punt, 1999). Each version of the operating model is used to explore the impact of each of these potential factors and their combined effects. A management strategy, which typically involves some form of decision rule (pre-definied rule that governs management actions, e.g. adjustments to the harvest level based on stock size), based upon the perceived state of the stock. The action determined by the "decision rule" is then applied to the simulated population, which is then projected for an additional year (or specified period between management actions). The cycle is repeated for a pre-specified duration, following which performance statistics are calculated to evaluate the implications of the management actions. Performance statistics should be chosen that are specific to the questions at hand and that are easy for managers and stakeholders to understand (Francis and Shotton, 1997). Performance statistics often measure one of several quantities: predicted changes in catch (Starr et al., 1997), interannual variability in catch (International Whaling Commission, 1992a), average biomass levels, and the probability of biomass falling below a specified threshold (Francis, 1992; Punt and Smith, 1999). An MSE lays bare the risks and potential trade-offs among actions, which allow fishery managers to make more informed decisions based upon the current base of knowledge.

This work employs both simulation and MSE to inform the management of US west coast groundfish. The overfished declarations of multiple rockfish stocks in 1999 and 2000 (Pacific Fishery Management Council, 2016), resulted in large harvest reductions relative to historical catch to allow stocks to rebuild to target biomass levels (Hilborn et al., 2012). Harvest restrictions have resulted in avoidance behavior by fisherman for many overfished stocks (Kuriyama et al., 2016), reducing opportunities for the collection of fishery compositional data. Limited fishery data can pose a challenge for assessment when the fishery provides the primary source of information on a stock. Several of the overfished rockfish stocks fell into this category, where they are not well sampled by the main scientific survey off the US west coast, either due to the inability of the to sample rocky habitat using trawl gear or other restrictions on sampling locations (e.g. rockfish conservation areas or near-shore habitat). The combination of limited scientific observations and reduced fishery data has created a situation when the bulk of data for assessment informs the historical period with only limited information available to monitor the progress of rebuilding efforts. The objective of Chapter 1 was to evaluate the impact of reduced data during rebuilding in an assessment's ability to estimate spawning biomass, biological parameters, and identify when a stock had successfully rebuilt.

Preventing stocks from becoming overfished due to overfishing requires appropriate harvest limits that maintain stocks at or above the target biomass. Approximately two thirds of the 80+ stocks within the US west coast groundfish FMP have not been formally assessed due to a lack of available data as a consequence of limited fishery targeting, and insufficient time or money. In order to set harvest limits for previously unassessed stocks the PFMC adopted new data-moderate and catch-only (data-limited) estimation methods. However, with limited data, uncertainty about current biomass increases and additional precaution should be taken when setting harvest limits to prevent overfishing. Chapter 2 used management strategy evaluation to determine the ability of alternative estimation methods for data-moderate and data-limited stocks to prevent overfishing and maintain stocks at or above target biomasses.

Chapter 3 performed an MSE to evaluate alternative methods to rebuild overfished groundfish stocks to target biomass levels. Stocks that fall below MSST are declared overfished, and the SFA mandates that a formal rebuilding plan be implemented by management, outlining a timeline for rebuilding by a maximum year. However, Fishery Management Councils have the flexibility to determine what harvest rates to apply that will allow the stocks to rebuild to target biomass levels at or before the legally mandated deadline. Managers must consider the trade-offs among applying a conservative harvest policy that will allow for earlier rebuilding and the impact of restricted catches on fishery. The MSE in Chapter 3 explored risk-averse to risk-neutral approaches for rebuilding overfished groundfish stocks and provided a formal evaluation of the trade-offs for managers to consider when setting rebuilding harvest policies.

One of the primary goals of management is to maintain stocks at or near target biomass levels and to prevent stocks from becoming overfished, through the application of appropriate harvest limits. The PFMC employs the use of proxy targets and harvest rates, with a harvest control rule dictating how catches should be adjusted downward if the stock falls below the target biomass. Determining if harvest rates and target biomasses are consistent and that the harvest control rule behaves as designed is key for successful fisheries management. The PFMC has applied life history specific proxies and harvest control rules. Flatfish species (i.e. Pleuronectiformes) are considered highly productive, allowing for increased harvest relative to other less productive stocks (e.g. rockfish). However, the benefits or potential risks associated with alternative proxies and harvest control rules are not always apparent. Chapter 4 conducted an MSE to evaluate the trade-offs and suitability of alternative harvest control rules to maintain US west coast flatfish stocks at target biomass levels. Trade-offs between catches, annual variation in catches, and the probability of a stock falling below the MSST were detailed for managers.

Chapter 1

ARE WE THERE YET? THE IMPACT OF REDUCED DATA ON THE ABILITY TO MONITOR REBUILDING FOR OVERFISHED FISH STOCKS

Abstract

Select rockfish stocks off the US west coast are below target biomasses and are managed under rebuilding plans that limit the allowable harvest. However, limited harvest reduces the opportunity to collect fishery-dependent data, which is the primary source of information on changes in abundance for many rockfish stocks. A simulation study was conducted using operating models that involved time-invariant or time-varying parameters to evaluate the impact of reduced data to estimate spawning biomass and biological parameters during rebuilding. Decreased data during rebuilding resulted in increased among-simulation variation in estimates of spawning biomass in absolute terms and relative to unfished spawning biomass. Additionally, decreased data resulted in reduced average catches and increased inter-annual variation in catches during rebuilding compared to when data collection was maintained. The presence of time-varying parameters in the operating model that were not accounted for within the estimation method resulted in increased among-simulation variability about spawning biomass and relative spawning biomass compared to the time-invariant case, with the largest increase in variability occurring during rebuilding when data were reduced or eliminated. Retaining data collections at historical levels allowed for improved parameter estimation during rebuilding which resulted in reduced variability in estimated stock size and increased average catches during rebuilding.

1.1 Introduction

In the absence of an unexpected run of good recruitment, rebuilding overfished stocks requires a reduction in fishing mortality to a level that allows stock biomass to increase. In the US, federally-managed stocks that fall below a minimum stock size threshold (MSST) are declared overfished and are mandated to be rebuilt to target biomass levels in the shortest amount of time, accounting for biological and environmental conditions (Sustainable Fisheries Act, 1996). This can lead to substantial reductions in fishing effort relative to historical levels. The severity of restrictions during rebuilding can, for some stocks, lead to a situation where the ability to collect data becomes limited over the period when managers are likely most concerned about stock size (i.e. when the stock is under the rebuilding plan).

Data are necessary to determine the extent to which has rebuilt. The ability to measure the rate of recovery is crucial to management, and increased uncertainty due to limited data can impede monitoring to determine if a stock is on target to rebuild in a specified timeframe. Additionally, biological data are critical to improve estimates of key parameters within stock assessments (e.g. natural mortality, growth, recruitment compensation termed steepness) and can indicate incoming poor or strong year-classes (recruitment), which can impact estimates of relative stock biomass (the ratio of current biomass to unfished biomass) and rebuilding rates. Potential improvements in parameter estimates, and the ability to detect incoming fluctuations in recruitment during rebuilding could be restricted when collection of new biological data is limited due to harvest restrictions.

Overfished rockfish species off the US west coast have experienced large reductions in harvest during rebuilding. One example is yelloweye rockfish (*Sebastes ruberrimus*), which was declared overfished in 2002 (Methot and Piner, 2002). Similar to other rockfish species off the US west coast, catches of yelloweye rockfish were unsustainable during the 1980s and early 1990s. Catches of yelloweye rockfish were reduced dramatically relative to historical catches following the overfished declaration, where the allowable catch during the first year of rebuilding was reduced to approximately 10% of the catch four years earlier (Stewart et al., 2009).

The reduction of fishery catch, and resulting fishery data during rebuilding, presents a challenge for assessment and management. Many species of rockfish, such as yelloweye rockfish, are not reliably sampled by the main fishery-independent survey off the US west coast, either due to the inability of the survey to sample rocky habitat using trawl gear or other restrictions on sampling locations (e.g. rockfish conservation areas or near-shore habitat). Because these species are not well sampled, the majority of historical information (e.g. length, and age data) available for assessment derives primarily from recreational and commercial fishery samples. Yet, because of retention restrictions triggered by the rebuilding plan, recreational and commercial fishery behavior has been profoundly altered (Stewart et al., 2009). The most recent yelloweye rockfish assessment cited limited fishery data during rebuilding as a challenge to "produce conclusive information about the stock for the foreseeable future" (Stewart et al., 2009).

Understanding the long-term impact of reduced data on the ability to monitor a stock during rebuilding would provide insight and guidance for management. There have been numerous simulation studies evaluating the impact of data quality and quantity on the performance of stock assessment methods (e.g. Hilborn, 1977; Chen et al., 2003; Yin and Sampson, 2004; Magnusson and Hilborn, 2007; Wetzel and Punt, 2011a; Lee et al., 2012). However, studies often focus on the ability to estimate either management quantities or biological parameters. The simulation performed here evaluates the ability to accurately monitor rebuilding of an overfished long-lived rockfish stock where harvest and the collection of fishery data are restricted during rebuilding. The simulation study addresses three main questions; 1) does limited data result in increased uncertainty impacting the ability to detect when an overfished stock has recovered to the management target stock size (i.e. it is rebuilt), 2) are limited data from the fishery able to detect a shift in fishery selectivity resulting from changing fishing behavior during rebuilding, and 3) how are model estimates of stock size and biological parameters affected during periods of limited data?

1.2 Materials and Methods

1.2.1 General approach

A rockfish life history type common to the US west coast was simulated (Table 1.1), based on the life history for yelloweye rockfish. Yelloweye rockfish exhibit very low natural mortality and recruitment compensation (termed 'steepness'), even relative to other US west coast rockfish species and are therefore assumed to have slow population dynamics. The operating model was parameterized using higher natural mortality and steepness values to be more similar to other US west coast rockfish species, and to allow for shorter recovery periods (<100 years) for computational efficiency while still maintaining the characteristics of a rockfish life history.

Two alternative cases were simulated using the operating model to account for the potential impacts of time-varying natural mortality and fishery selectivity. The first case, referred to as "time-invariant", involved a single fixed natural mortality rate over the entire time period. The fishery selectivity was assumed (and fixed) to be asymptotic during historical period, dome-shaped during the overfished period, and then again asymptotic after the stock was rebuilt (Fig. 1.1). The simulated stocks were reduced to an overfished state (below MSST) at the time of the first assessment in year 50.

The second case, referred to as "time-varying", involved annual deviations in natural mortality and in the parameters on which the fishery selectivity pattern was based during the historical, overfished, and rebuilt periods (Fig. 1.1c and Fig. 1.1d). Annual deviations in fishery selectivity were applied to two selectivity parameters: 1) the length (in cm) at which the ascending limb of selectivity curve reached maximum selectivity (termed 'size at maximum selectivity', Fig. 1.1c), and 2) the width of the plateau for maximum selectivity (defined as a logistic function between peak and the maximum length) resulting in dome-shaped selectivity curve (termed 'width at maximum selectivity', Fig. 1.1d) during the years the stock was overfished. A standard error of 0.05 was applied annually about the size at maximum selectivity parameter for all years and a standard error of 0.20 was applied for the width at maximum selectivity parameter during the years the stock was estimated to be overfished. The level of variation about each parameter was selected to ensure that the ascending limb of the selectivity curve was greater than the 50% length at maturity (37 cm) within the operating model, and the width of maximum selectivity (creating dome-shaped curve) was small enough to allow potential detection by the estimation method (a detectable portion of the population with reduced selectivity due to dome-shaped curve). Annual deviations in natural mortality were autocorrelated.

The operating model was based on a single-sex age-structured model. An annual fishery catch-per-unit effort (CPUE) index was observed with error, length- and age-composition data were collected for selected years, and used by the estimation method to estimate population size and a catch level. The catches were then removed without error from the simulated stock. Data generation, catch estimation and stock updating was conducted in an iterative fashion for 100 years (termed the management period), a length of time that would allow for the stock to recover to the target biomass.

1.2.2 The operating model

The numbers-at-age at the start of the year are computed using the equation:

$$N_{t+1,a} = \begin{cases} R_t & \text{if a} = 0 \\ N_{t,a-1}e^{-(M_t+S_{t,a-1}F_t)} & \text{if } 1 \leq \text{a} < \text{A-1} \\ N_{t,A-1}e^{-(M_t+S_{t,A-1}F_t)} + N_{t,A}e^{-(M_t+S_{t,A}F_t)} & \text{if a} = \text{A} \end{cases} \tag{1.1}$$

where $N_{t,a}$ is the number of fish of age a at the start of the year t, R_t is the number of age-0 fish at the start of year t, $S_{t,a}$ is the selectivity during year t for fish of age a, A is the plus group (set equal to age 70), F_t is the instantaneous fishing mortality rate during year t, and M_t is the instantaneous rate of natural mortality during year t.

Natural mortality for year t is defined as:

$$M_t = Me^{-0.5\sigma_M^2 + \epsilon_t^M} \tag{1.2}$$

where M is the mean value of natural mortality, σ_M is the standard error of the annual deviations in natural mortality, and ϵ_t^M is the autocorrelated lognormal deviation in natural mortality for year t:

$$\epsilon_t^M = \rho\epsilon_{t-1}^M + \sqrt{1-\rho^2}\phi_t \quad \phi_t \sim \ N(0;\sigma_M^2) \tag{1.3}$$

where ρ is the level of autocorrelation associated with natural mortality and ϕ_t is the deviation in natural mortality for year t. The time-invariant natural mortality case assumed $\sigma_M = 0$ and hence $\epsilon_t^M = 0$.

The number of age-0 fish is related to spawning biomass according to the Beverton-Holt stock recruitment relationship:

$$R_t = \frac{4hR_0SB_t}{SB_0(1-h) + SB_t(5h-1)} e^{-0.5\sigma_R^2 + \epsilon_t^R} \quad \epsilon_t^R \sim N(0;\sigma_R^2) \tag{1.4}$$

where SB_0 is the unfished spawning biomass, SB_t is the spawning biomass at the start of the spawning season in year t, σ_R is the standard deviation of recruitment in log space, and h is steepness.

A non-equilibrium starting condition was created by applying equations (1.1) and (1.2) for the number of years equal to the maximum age prior to the start of fishing, with variation in recruitment. Historical catches for years 1-50 were generated so that the populations were at $0.15SB_0$ in year 50. This ratio of spawning biomass to unfished spawning biomass (relative spawning biomass) was selected to allow for correct detection by the estimation method that the stocks were in an overfished state, and that it would require an extended number of years for the stock to rebuild to the target biomass (where the loss of data could impact the long-term performance of the estimation method). The catch of fish of age a during year t

in numbers is given by:

$$C_{t,a} = \frac{S_{t,a}F_t}{M_t + S_{t,a}F_t} N_{t,a} \left(1 - e^{-M_t - S_{t,a}F_t}\right) \tag{1.5}$$

The observation model was used to generate a fishery CPUE index for each year t:

$$I_t = Q\widetilde{B}_t e^{-0.5\sigma_f^2 + \epsilon_t^f} \quad \epsilon_t^f \sim N(0; \sigma_f^2) \tag{1.6}$$

where Q is the catchability coefficient, σ_f is the standard deviation of catchability in log space, and the $\widetilde{B}_t$ is the selected biomass in the middle of year t:

$$\widetilde{B}_t = \sum_{a=1}^{A} w_a S_{t,a} N_{t,a} e^{-0.5(M_t + S_{t,a}F_t)} \tag{1.7}$$

where w_a is the weight of a fish of age a. The length- and age-composition data for the fishery were assumed to be multinomially distributed (see *Data Scenarios* section for details). Ageing error was assumed to be normally distributed with ages subject to a 5% standard deviation by age.

The fishery selectivity during the historical period (years 1-50) were assumed to be asymptotic (Fig. 1.1a and 1.1c). Fishery selectivity shifted to a dome-shaped (compared to the historical asymptotic) form (Fig. 1.1b and 1.1d) within the operating model during the period that the stock was estimated to be below the target biomass ($0.40SB_0$). Once the population was estimated to have recovered to above the target biomass, fishery selectivity reverted to the asymptotic form. The shift in selectivity was designed as a way to mimic a change in fisher behavior resulting from an overfished declaration (e.g. creation of rockfish conservation areas protecting portions of the stock or fisher avoidance of known specific habitat or areas associated with high abundance of the overfished stock). The change in shape of the selectivity curve depended on the estimated stock status rather than the true operating model status, i.e. changes in fisher behavior modeled by a change in selectivity were assumed to be driven by management restrictions based on the estimation method's perception of the stock rather than the true unobservable state of the simulated stock.

1.2.3 The estimation model

Stock synthesis (SS), an integrated statistical catch-at-age model (Methot and Wetzel, 2013), was the estimation method used to assess the simulated stocks. SS was applied for the first time in year 50 and then every 6th year thereafter. Assessment frequency for US west coast groundfish varies as a consequence of commercial importance (an indicator of exploitation), the time since last assessment, and dynamics of the stock (Methot, 2015). Long-lived rockfish species generally have slow dynamics, resulting in minimal fluctuations in biomass from year to year (assuming non-extreme harvesting). To mimic the likely cycle of assessments for this stock in real life the assessment was conducted every 6^{th} year.

Parameters determining unfished recruitment (R_0), steepness, growth, annual recruitment deviations, initial age-structure deviations, and the size and width at maximum selectivity for the fishery were estimated. Steepness was estimated using a diffuse beta prior within the estimation method. All other parameters were estimated without priors. Natural mortality, the variation of length-at-age, weight-at-length, the fecundity relationship, and the variation of recruitment (σ_R) were assumed known. The relative spawning biomass in the assessment year was estimated and the forecasted catches were determined using the harvest control rule adopted by the Pacific Fishery Management Council (PFMC) for rockfish (see below). The catches were removed from the operating population without error, fishery CPUE index, length- and age-composition data were then generated for the subsequent six years.

The harvest control rule adopted by the PMFC for rockfish involves a linear reduction in catch when a stock falls below $0.40SB_0$, with no fishing when the stock falls below $0.10SB_0$. The maximum catch, termed the overfishing level catch, (the catch corresponding to the proxy for the fishing mortality at which maximum sustainable yield is achieved and if surpassed would constitute overfishing) was set equal to the target relative spawning biomass-per-recruit harvest rate ($F_{0.50}$) multiplied by SB_t. The overfished level catch was reduced by a management buffer (0.956) that accounts for the uncertainty about current biomass for well-assessed stocks to determine the acceptable biological catch level (i.e. acceptable biological catch = 0.956 overfishing level catch, Ralston et al., 2011). The annual catch limit was set equal to the acceptable biological catch when the stock was above the target biomass, $0.40SB_0$, or reduced from the acceptable biological catch according to the harvest control rule when the stock fell below $0.40SB_0$.

One major simplification in this simulation design and actual management practice of US west coast groundfish was the omission of the rebuilding plans that are implemented when a stock is assessed to have fallen below the MSST (defined as $0.25SB_0$ for US west coast rockfish). In reality, harvest for stocks below the MSST is not based on the standard harvest control rule, but rather a rebuilding plan that determines catches until the stock is rebuilt to the target biomass (see Wetzel and Punt (2016) for additional details on PFMC rebuilding plans).

1.2.4 Data scenarios

Three data scenarios were created to explore the impact of data availability on the ability to monitor rebuilding of an overfished stock (Fig. 1.2). The data scenarios were designed to emulate a stock, like yelloweye rockfish, that is infrequently encountered by a fishery-independent survey (e.g. due to depth or habitat) and only fishery data were available. The historical length and age sample sizes were generally based on the effective sample sizes observed for yelloweye rockfish. Following the first assessment in year 50, the three scenarios have different data availability based on estimated stock status (e.g. overfished vs. rebuilt) in the assessment year.

The "full data" scenario maintained the fishery CPUE index and length- and age-composition data at the historical levels (prior to the stock being declared overfished in year 50) during rebuilding (Fig. 1.2). The "reduced data" scenario decreased the amount of data available from the fishery during rebuilding (Fig. 1.2). The length and age-composition data were reduced to 20% of the historical sample sizes during rebuilding and the fishery CPUE index was eliminated during the rebuilding period. The CPUE index resumes and composition sample sizes reverted to historical levels when the stock was estimated to have rebuilt to the target biomass. The "eliminated data" scenario had no fishery data during rebuilding (Fig. 1.2). The fishery CPUE index and composition data resumed at historical sample sizes when the stock was projected to be rebuilt.

The estimation method in the full and reduced data scenarios were allowed to estimate a change in selectivity from asymptotic to dome-shaped during the rebuilding period through the application of a time block on selectivity. However, the eliminated data scenario was forced to assume constant asymptotic selectivity in the assessment in all years because no fishery composition data were available to detect a potential shift in selectivity.

1.2.5 Performance Measures

The outcomes of the simulations for each case and data scenario were summarized using five metrics that were selected to evaluate the impact of data on estimation of indicators of stock status (e.g. relative spawning biomass) and management quantities (e.g. rebuilding catch):

1. The relative errors (REs) for estimated parameters, calculated as:

$$RE = \frac{E - T}{T} \tag{1.8}$$

   where E is the estimated quantity of interest and T is the true value from the operating model.

2. The percent root mean square error (RMSE), a measure of precision and bias, was calculated to assess the overall level of error given the amount of data available:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{(E_i - T_i)^2}{T_i^2}} \tag{1.9}$$

   where N is the number of simulations ($N = 100$).

3. The average (over simulations) of the total catch while the stock was recovering to the target biomass.

4. The annual average variability of the catches (abbreviation AAV), defined as:

$$AAV = 100\frac{\sum_t |C_t - C_{t+1}|}{\sum_t C_t} \tag{1.10}$$

   where C_t is the catch during year t.

5. The percentage of stocks that rebuilt to the target biomass and percent of stocks that remain overfished at the end of management period.

1.3 Results

1.3.1 Assessment performance with time-invariant parameters

The trends of the relative error about spawning biomass and relative spawning biomass were generally consistent among the full and reduced data scenarios (Figs. 1.3a-b and 1.4a-b). The median estimates of spawning biomass and relative spawning biomass were less than the true values during rebuilding for both scenarios (Figs. 1.3a-b and 1.4a-b). As expected, there was less among-simulation variability in the difference between operating model and estimated spawning biomass and relative spawning biomass for the full data scenario during the rebuilding period compared to the reduced and eliminated data scenarios (Figs. 1.3a-c and 1.4a-c). However, by the end of the management period, the among-simulation variability of errors in biomass metrics were similar between the full and reduced data scenarios. The eliminated data scenario resulted in median (across simulations) estimates of spawning biomass and relative spawning biomass errors that were similar to the true values, but were highly imprecise at the start of the management period (years 50-74) (Figs. 1.3c and 1.4c). The eliminated data scenario, in the absence of new data during rebuilding, projected stocks based on the historical data and new catches until rebuilt, at which time data collection resumed allowing the estimation method to estimate population status. The median estimates for the eliminated data scenario were less than the true values, with high among-simulation variability in error as stocks began to be projected to be rebuilt and data collection resumed. In contrast to the full and reduced data scenarios, the estimates of spawning biomass and the relative spawning biomass for the eliminated data scenario showed little improvement in the among-simulation variability in error estimates by the end of the management period (Figs. 1.3c and 1.4c).

The RMSE for the estimated relative spawning biomass for each assessment year shows the increased precision of the full data scenario during the rebuilding period compared to the reduced and eliminated data scenarios (Fig. 1.5a). The eliminated data scenario resulted in the highest RMSE over the entire management period (Fig. 1.5a). However, the RMSE for the reduced data scenario showed improvement over the management period as stocks began to be assessed rebuilt to the target biomass, and sample sizes returned to historical levels. The limited improvement in the RMSE for the eliminated data scenario was driven by the simulations that were never projected to have rebuilt to the target biomass (35 out of 100 simulations).

Examining the eliminated data scenario closer revealed a pattern in the performance of the estimation method based on the estimation of steepness in the first assessment year. The eliminated data scenario simulations were divided and plotted based on whether the estimation method projected the simulation rebuilt (65 simulations) or failed to rebuild (35 simulations) by the end of the management period. To allow comparison, the estimates from the full data scenario were also divided into the same two groups and plotted. The estimated spawning biomasses were considerably less that the true values in the first assessment year (Figs. 1.6b [white]) for the 35 simulations that were estimated not to be rebuilt by the end of the management period. The underestimates of spawning biomass (Fig. 1.6b [white]) were driven by estimates of steepness that were much less than the true value in the first assessment (Fig. 1.6d [white]). In the absence of new data, the underestimates of steepness resulted in the estimation method perceiving a less productive stock requiring an extended period to rebuild to the target biomass. However, with full data present, estimated quantities (spawning biomass and steepness) improved for this subset of simulations and were median unbiased by the end of the management period (Fig. 1.6a and 1.6c [white]). Note that the term "median unbiased" will be used to define cases in which the median of the relative errors equals zero.

The estimates of steepness varied across data scenarios. The full data scenario resulted in median unbiased estimates by the end of the management period (Fig. 1.7a). In contrast, the median of the estimates of steepness for the reduced data scenario were greater than the true steepness during the management period (Fig. 1.7b). The eliminated data scenario had the highest among-simulation variability among estimates of steepness during the management period (Fig. 1.7c) due to the mixture of rebuilt and not rebuilt stocks.

The median number of years estimated for the stocks to recover to the target biomass for the full data scenario was longer than the median recovery year within the operating model (Table 1.2). In contrast, both the reduced and eliminated data scenarios had shorter median recovery times compared to the operating model (Table 1.2). The contrast in estimated recovery times across the data scenarios was related to the average catch obtained during rebuilding along with the bias and variability of estimates. The median error associated with relative spawning biomass for the full data scenario was less than zero, with low among-simulation variability (compared to the other data scenarios) for all assessment years, resulting in estimates that predicted constant rebuilding but at a slower rate than the true stock (Fig. 1.4a). In contrast, the reduced data scenario had higher variability over time (i.e. within-simulation) across the estimates of error associated with relative spawning biomass (Fig. 1.4b). The variability of estimates between assessments resulted in stocks being estimated recovered earlier than was the case in the operating model due to estimation error driven by the limited composition samples during rebuilding.

The reduced data scenario had the lowest median average catch during rebuilding (Table 1.3), with the median rebuilding time estimated shorter than the true time to recovery within the operating model (Table 1.2). The eliminated data scenario which was entirely dependent upon historical data until the stocks were projected to be rebuilt, essentially projected the population forward with each assessment based on the initial parameter estimates from the historical data, resulting in high median average catches during rebuilding, and the lowest median AAV during rebuilding and across the entire management period (Table 1.3).

Reduction or elimination of data during rebuilding increased the among-simulation variability about estimates of the size at maximum fishery selectivity which defined the ascending limb of the selectivity curve (see Fig. 1.1a), with the median estimates generally equal to the true value for all data scenarios (Fig. 1.8a-c). The among-simulation variability of the estimates for the reduced and eliminated scenarios improved when the majority of the stocks were estimated to be rebuilt, and fishery composition sample sizes returned to historical levels. The full and reduced data scenarios were allowed to estimate dome-shaped selectivity during the rebuilding period (the eliminated data scenario did not allow for estimation of dome-shaped selectivity due to the absence of fishery composition data), and resulted in median estimates that exceeded the true values and were highly variable among simulations at start of the management period (Fig. 1.9a-b). The estimates that exceeded the true values for this parameter indicated that the data available were not sufficient to inform the estimation method about the severity of the dome-shape in selectivity during rebuilding (a higher estimated value implies the dome in selectivity occurs at larger sizes with a higher proportion of the population relative to the operating model at full selectivity). The full data scenario resulted in markedly improved estimates of the shape of the dome over the management period, compared to the reduced data scenario (Fig. 1.9a-b).

1.3.2 The impact of time-varying parameters

The case that assumed time-varying annual deviations in natural mortality and fishery selectivity generally resulted in increased among-simulation variation in estimation errors compared to the time-invariant case. The median error of estimates of spawning biomass at the time of the first assessment exceeded the true values and were highly variable among simulations (Fig. 1.3d-f). The among-simulation variance in errors of estimates of spawning biomass decreased markedly for the full data scenario after the first assessment (Fig. 1.3d).

However, this variability remained high for approximately the first twenty-five years of the management period for both the reduced and eliminated data scenarios, until approximately 50% of the simulated stocks were estimated recovered and the fishery sample sizes increased to historical levels (Fig. 1.3e-f). The full and reduced data scenarios resulted in median spawning biomass estimates that were generally smaller than the operating model values (Fig. 1.3d-e). However, the medians of the errors for relative spawning biomasses were variable over the management period (Fig. 1.4d-e). The medians of the estimates of relative spawning biomass for eliminated data scenario were larger than operating model values at the start of the management period, but became smaller than these values as stocks rebuilt to target biomass and data collection resumed (Fig. 1.4f).

Compared to the case with time-invariant parameters the RMSE was higher for all data scenarios when time-varying parameters were present within the operating model (Fig. 1.5). The RMSE about the estimated relative spawning biomass for the full data scenario was lower relative to the other scenarios for the entire management period (Fig. 1.5b). Similar to the time-invariant results, the RMSE of relative spawning biomass for the eliminated data scenario was the highest among the scenarios across the entire management period, peaking in assessment year 68 at 221% (a single simulation for the eliminated data scenario, with extreme outliers for two assessment years, was removed for a more informative summary of the RMSE).

The time-varying results for the eliminated data scenario were qualitatively similar to time-invariant case, where a large number of simulations failed to be projected rebuilt by the estimation method (32 simulations). As was observed in the time-invariant case, the simulations that failed to be projected rebuilt had median estimates of spawning biomass and relative spawning biomass below the operating model values at the time of the time of the first assessment, which were driven by estimates of steepness that were considerably lower than the true value (not shown).

The inclusion of time-varying parameters in the operating model resulted in shorter median estimated recovery times relative to the time-invariant case for the full and reduced data scenarios (Table 1.2). However, the median number of years to rebuild for the operating model stocks were similar between the time-varying and time-invariant cases. The estimation method estimated earlier recovery times for the time-varying case due to the increased variability in the estimates of relative spawning biomass, resulting in the estimation method having an increased frequency of erroneously estimating the biomass to be above the target stock size (Fig. 1.4).

The median average catch during the recovery period was the highest for the eliminated data scenario due to estimates of the relative spawning biomass that were higher than the true values at the start of the management period (Fig. 1.4 and Table 1.3). Additionally, the eliminated data scenario had the lowest median AAV during the rebuilding period (Table 1.3). In contrast, the eliminated data scenario, resulted in the highest number of stocks that never reached the target biomass (Table 1.3).

Inclusion of time-varying selectivity resulted in the median estimates of maximum selectivity across all data scenarios exceeding the mean of the operating model values (Fig. 1.8d-f), although the full data scenario resulted in the lowest among-simulation variation. The full and reduced data scenarios, which were allowed to estimate dome-shaped selectivity (width at maximum selectivity) during the recovery period, resulted in highly variable among-simulation estimates at the start of the management, period with the variability for the estimates decreasing earlier for the full data scenario (Fig. 1.9c-d).

1.4 Discussion

Maintaining fishery data at historical levels during rebuilding reduced the variation in estimates between assessments (i.e. over time within a simulation). While the full data scenario had less variation, the median (over simulations) estimates of spawning biomass and relative spawning biomass were consistently below the operating model values for much of the management period. This result is contrary to what might be expected when additional data are available. Explorations with simulations where there was a fishery independent survey that provided an index of abundance and composition data (length and age), determined that this underestimation of the true spawning biomass was eliminated if survey composition data were available along with fishery composition data (see *Appendix A*). The underestimation was driven by two key factors; the shape of fishery selectivity curve and data quantity. The fishery selectivity curve was specified to be greater than the maturity-at-length curve, resulting in only mature larger fish being selected by the fishery. The fishery data were informative about recruitment, but with a lag between recruitment to population and recruitment to the fishery. However, a fishery-independent survey selecting fish at smaller sizes yields information about recruitment to population earlier. Additionally, an increase in the length- and age-composition samples from multiple data sources can improve estimates of recruitment, spawning biomass, and relative spawning biomass (Yin and Sampson, 2004; Wetzel and Punt, 2011a).

Median relative errors for relative spawning biomass below zero for the full data scenario resulted in the estimation method failing to determine that the operating model population was at or above the target biomass (median number of rebuilding years greater than the operating model, Table 1.2), an outcome that would lead to extended harvest restrictions that were not warranted given the true state of the population, a situation fishery management would like to avoid. However, the reduced estimation variability (within- and among-simulations) offered by the full data scenario resulted in an improvement in the consistency of estimates by subsequent assessments, which offers a level of stability for fisheries managers and stakeholders. In contrast, the higher between-assessment variation in estimates of spawning biomass for the reduced data scenario resulted in stocks being estimated rebuilt when the true population was still below the target biomass which could have undesirable outcomes for fisheries management. Overly optimistic estimates of relative spawning biomass can result in overfishing when catch limits are set too high, leading to further reductions in biomass, potentially requiring an overfished declaration by a future assessment.

Loss of data during rebuilding resulted in a high number of simulations that failed to rebuild due to poor initial estimates of steepness, a key parameter, controlling how quickly a stock can rebuild from low biomass levels. In the absence of new data, the first and subsequent assessments were entirely dependent on the quality of the historical data to inform parameter estimates. The simulations that failed to correctly detect rebuilt stocks were driven by erroneously low estimates of steepness at the time of the first assessment. Estimating a stock to be less productive than the true population resulted in lower estimates of spawning biomass and relative spawning biomass, with the assessment setting harvest at levels well below the true acceptable biological catch. The reduced harvest allowed the population in the operating model to rebuild to or above the target biomass. However, in the absence of new (and informative) data, the estimation method was unable to detect the correct stock size. The operating model population represented a two-way trend of abundance (decline and increase in biomass) with fishery data available during the fishing down and recovery period, which previous studies have found informative in estimating steepness (Magnusson and Hilborn, 2007; Conn et al., 2010). This work showed that a one-way trip scenario in stock size with limited data may not be adequate to correctly estimate steepness, but the inclusion of even limited data can, with contrast in stock size, improve the estimation of steepness even if the initial assessment produced a poor estimate (Figs. 1.6c and 1.7).

The general trend in results when the operating model included time-varying natural mortality and fishery selectivity were similar to the time-invariant case, although the among-simulation estimates were more variable across all data scenarios. The estimation method assumed a single natural mortality across all years equal to the mean value that was used to generate the autocorrelated annual deviations in the operating model. This setup was a strategic choice that allowed variation in the composition data that the estimation method would not be able to account for, but would not be anticipated to result in strongly biased estimates due to model misspecification. Time-varying natural mortality is likely more complex with extended periods of high or low mortality based upon external factors (e.g. predator abundance, climate conditions), which could result in large biases in estimated quantities if not accounted for in an assessment (Johnson et al., 2015).

Shifts in the form of selectivity over time and the impact of annual deviations in selectivity led to mixed results. The estimation method consistently overestimated the mean size at maximum selectivity for all data scenarios with time-varying selectivity. The operating model selectivity involved normally distributed deviations to generate the annual shifts in selectivity and one would not a priori predict the estimation method to have a consistent bias in estimates. However, the estimation method was able to identify the change in the selectivity form (asymptotic to dome-shaped through a reduction in the width at peak selectivity) during the rebuilding years with similar error to the time-invariant case. Each case led to estimates that overestimated the width at maximum selectivity, defining the dome in selectivity (dome-shaped selectivity occurring at larger sizes with increased sizes subject to full selectivity compared to the operating model). This evaluation applied time blocks defined by the status of the stock to allow for shifts in selectivity, ignoring the annual deviations in the selectivity curve. Studies have evaluated other ways of estimating time-varying selectivity using state-space models (Nielsen and Berg, 2014), or the implications of applying time blocks vs. allowing a random walk component in selectivity parameters or catchability (Wilberg and Bence, 2006; Martell and Stewart, 2014). Further exploration should be conducted to evaluate if allowing a random walk or applying an alternative estimation method eliminates the bias observed in the estimated selectivity observed here and how data quantity and quality impacts these estimates. Additionally, if shifts in fishery selectivity are anticipated due to management actions, increased data collections may be required in order to achieve a similar level of precision in estimates during rebuilding.

This work highlights the benefits of continued data collection during rebuilding on the precision of estimates, but there are many additional reasons why retaining data streams is important. The first benefit of continued data collection is the potential ability to identify misspecification in the model assumptions. The estimation method and operating models applied here, generally made similar structural assumptions. However, the true state of nature is never known with confidence in real-world assessments. Continued data collection may allow for the identification of model misspecification in the structural assumptions (e.g. growth, recruitment), which will allow for better approximations to reality through models.

Specifically, there could be long-term impacts to a population that is depleted, which could negatively impact the ability of the stock to rebuild (e.g. Hixon et al., 2014) that additional data would be required to detect. The second benefit of continued data collections is that ability to identify potential long-term changes in stock dynamics (e.g. changes in climate-driven forces). There has been much work identifying potential links between climate and recruitment (Hollowed et al., 2011; Ianelli et al., 2011; Mueter et al., 2011; Stachura et al., 2014). Additionally, long-term varying climate conditions may result in changes in biological parameters (Swain and Benoît, 2015), which will impact the productivity of stocks (Legault and Palmer, 2015) that needs to be accounted for when setting harvest limits.

Acknowledgments

This work has benefitted from feedback provided by Ian Taylor (Northwest Fisheries Science Center), Vladlena Gertseva (Northwest Fisheries Science Center), Ian Stewart (International Pacific Halibut Commission), and Allan Hicks (International Pacific Halibut Commission).

1.5 Tables

Table 1.1: Life history and observation parameters used in the operating model and their treatment within the estimation model.

Parameter	Time-invariant	Time-varying	Treatment in estimation method
Natural mortality (M) (yr^{-1})	0.08	0.08	Fixed
Natural mortality standard error (σ_M)	0	0.10	
Natural mortality autocorrelation (ρ)	0	0.707	
Steepness (h)	0.65		Estimated
Maximum length (L_∞) (cm)	64		Estimated
Growth coefficient (k_t) (yr^{-1})	0.05		Estimated
Weight at length $w_l = \alpha L^\beta$ (kg)	α=1.50x10^{-5} β= 3		Fixed
Length at 50% maturity (cm)	37		Fixed
Recruitment variation (σ_R)	0.50		Fixed
Fishery CPUE standard error (σ_f)	0.30		Fixed
Fishery CPUE catchability coefficient (Q_f)	0.01		Analytically estimated

Table 1.2: The median and 90% simulation interval for the estimated number of years needed to rebuild to the target biomass, the operating model number of years needed to rebuild to target biomass, and the number of stocks that failed to rebuild to the target biomass determined by the estimation method (EM) and the operating model (OM) for each case and data scenario.

Selectivity/ data scenario	Estimated num. of rebuilding years		Operating model num. of rebuilding years		Num. of stocks that failed to rebuild	
	Median	90% SI	Median	90% SI	EM	OM
Time-invariant						
full data	43	(13-87)	34	(16-73)	7	4
reduced data	31	(19-61)	34	(14-83)	1	5
eliminated data	25	(14-72)	37	(14-87)	35	4
Time-varying						
full data	31	(13-91)	35	(13-85)	13	4
reduced data	25	(13-79)	32	(12-74)	8	2
eliminated data	25	(13-77)	36	(12-79)	32	5

Table 1.3: The median and 90% simulation intervals for the average catch during rebuilding, the AAV during rebuilding, and the AAV over all years for each case and data scenario.

Selectivity/ data scenario	Average catch during rebuilding		AAV during rebuilding		AAV all years	
	Median	90% SI	Median	90% SI	EM	OM
Time-invariant						
full data	44.0	(15.3-78.9)	6.0	(3.7-11.5)	3.2	(2.1-4.7)
reduced data	28.1	(14.6-57.9)	7.7	(4.0-14.5)	3.5	(2.3-5.3)
eliminated data	41.3	(19.9-83.8)	2.6	(1.3-4.4)	2.2	(1.3-3.9)
Time-varying						
full data	31.7	(11.0-75.4)	7.3	(4.4-17.5)	4.2	(2.7-5.9)
reduced data	25.1	(15.6-68.0)	8.9	(4.5-20.7)	4.5	(2.6-9.8)
eliminated data	36.3	(15.7-79.4)	2.3	(1.2-4.8)	2.8	(1.3-5.3)

1.6 Figures

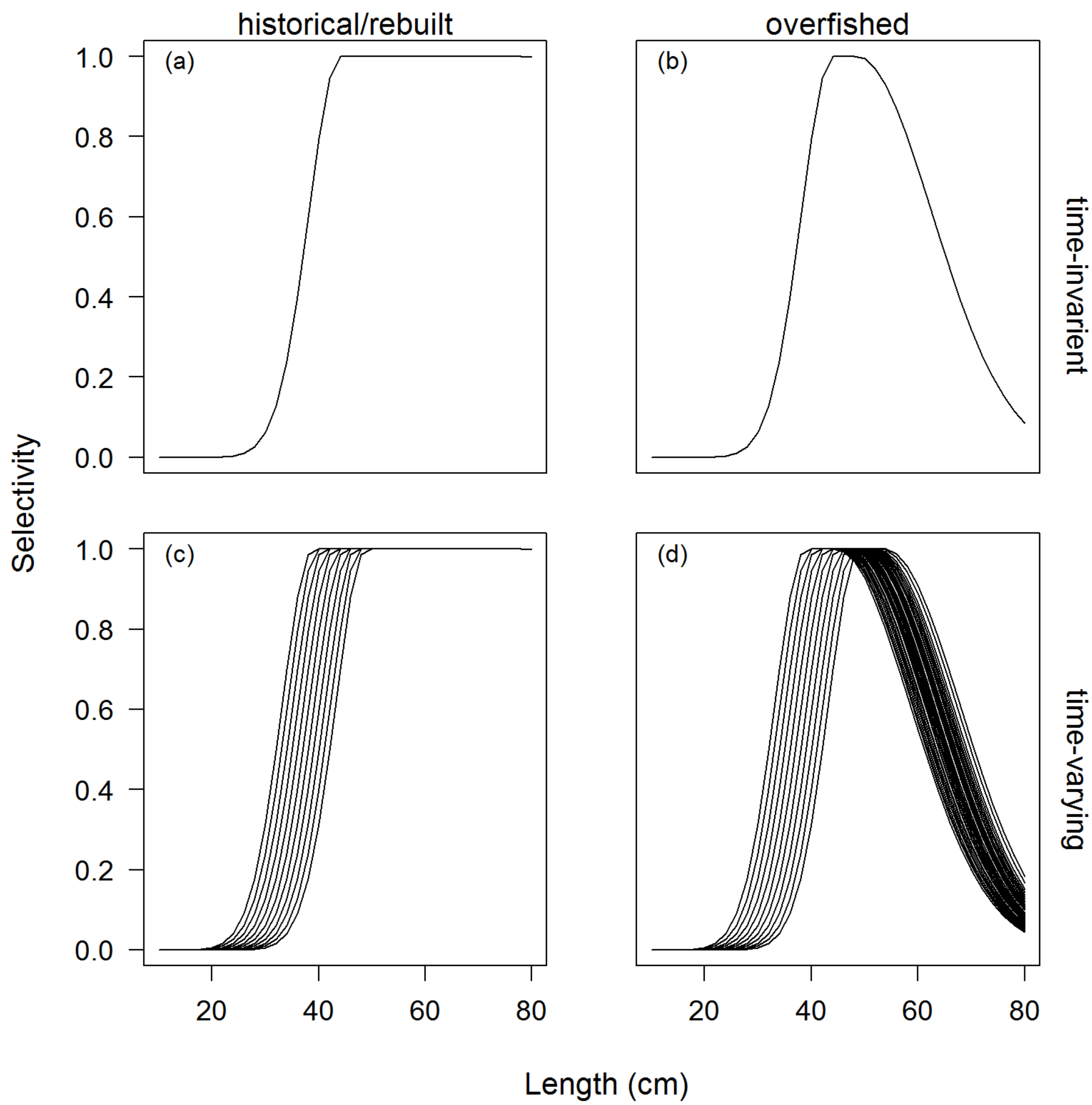

Figure 1.1: Fishery selectivity for either time-invariant or time-varying selectivity for the historical/rebuilt (a and c) and overfished (b and d) periods. A standard error of 0.05 was applied annually about the size at maximum selectivity, which defined the variability among the ascending limb of the selectivity curve (c and d), and a standard error of 0.20 was applied for the width at maximum selectivity that defined the length at which the dome in selectivity began while the stock was estimated overfished (d) (see Methot and Wetzel (2013), for additional details on double normal selectivity).

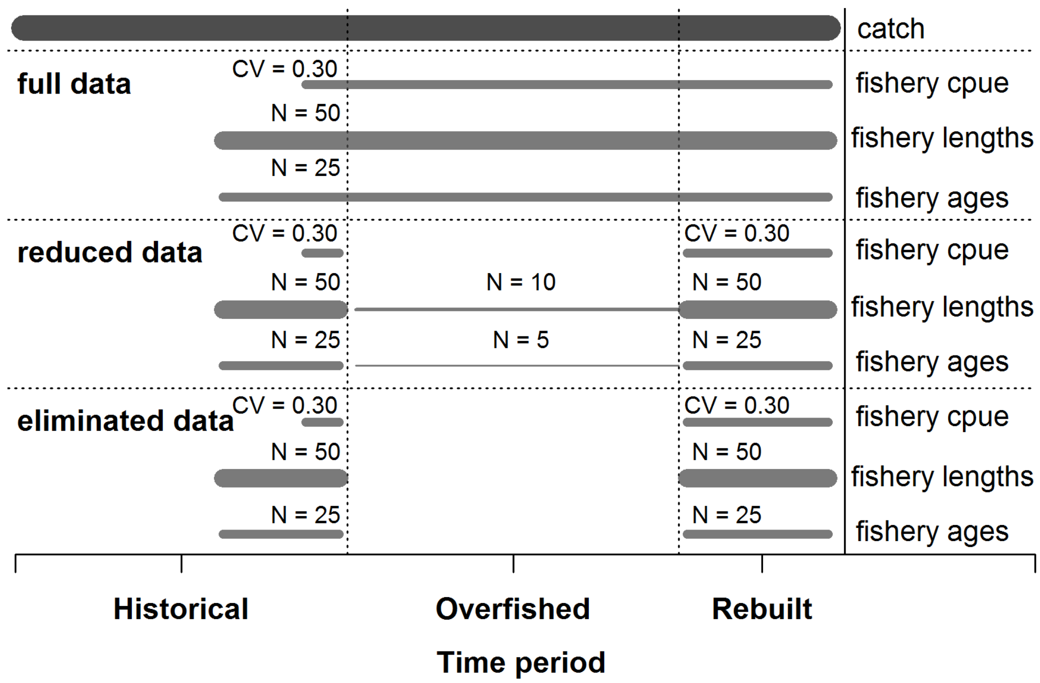

Figure 1.2: Summary of the data available for each of the data scenarios. Catches were known without error and were available for all data scenarios.

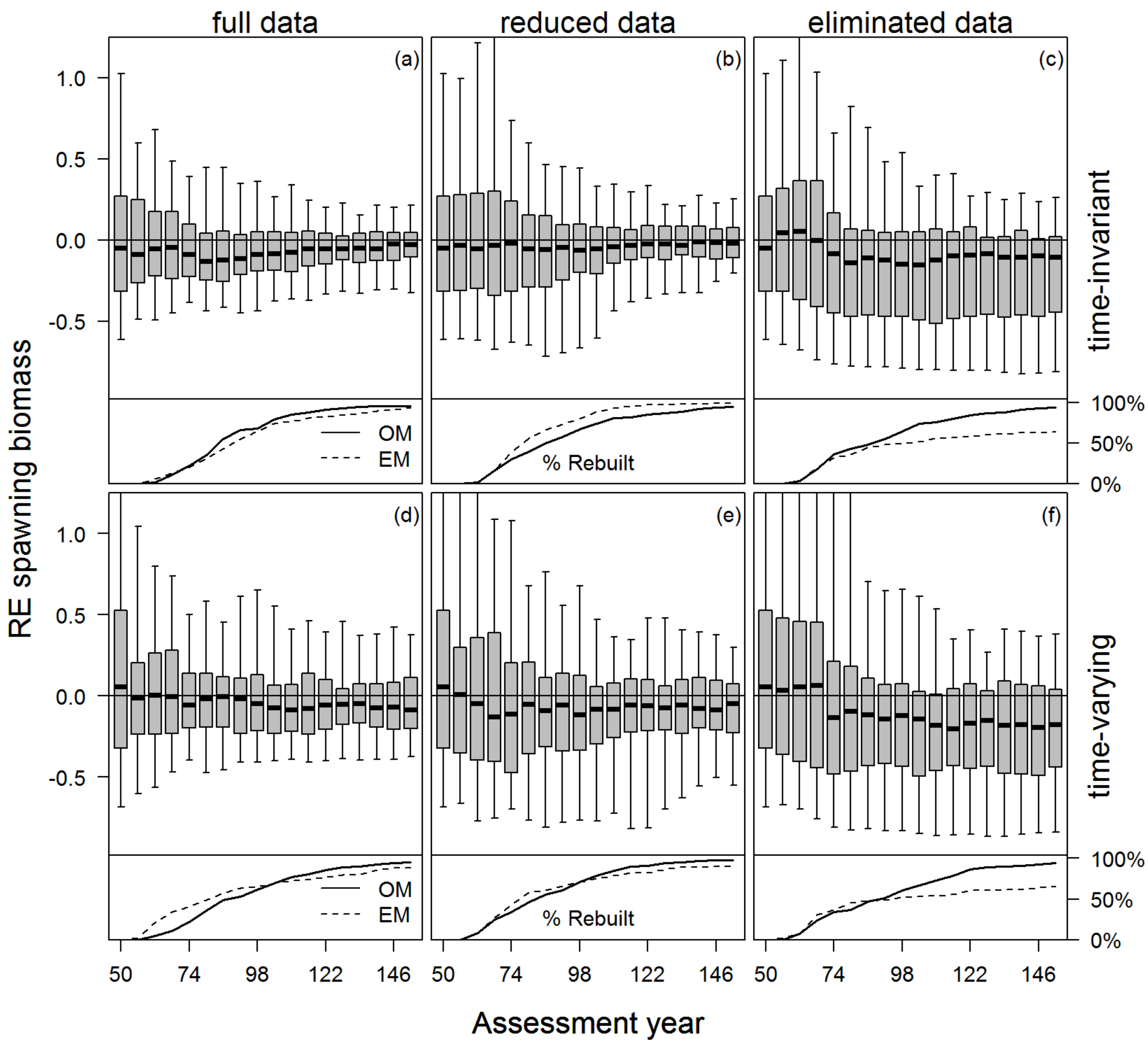

Figure 1.3: Relative error of estimated spawning biomass in each assessment year for each case and data scenario for all simulations (top panel) and the percentage of stocks that had rebuilt to the target biomass during the management period (bottom panel) within the operating model (OM; black line) and the estimation model (EM, dashed black line), with data collection consequently returning to historical levels when the EM determined the stock was rebuilt. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

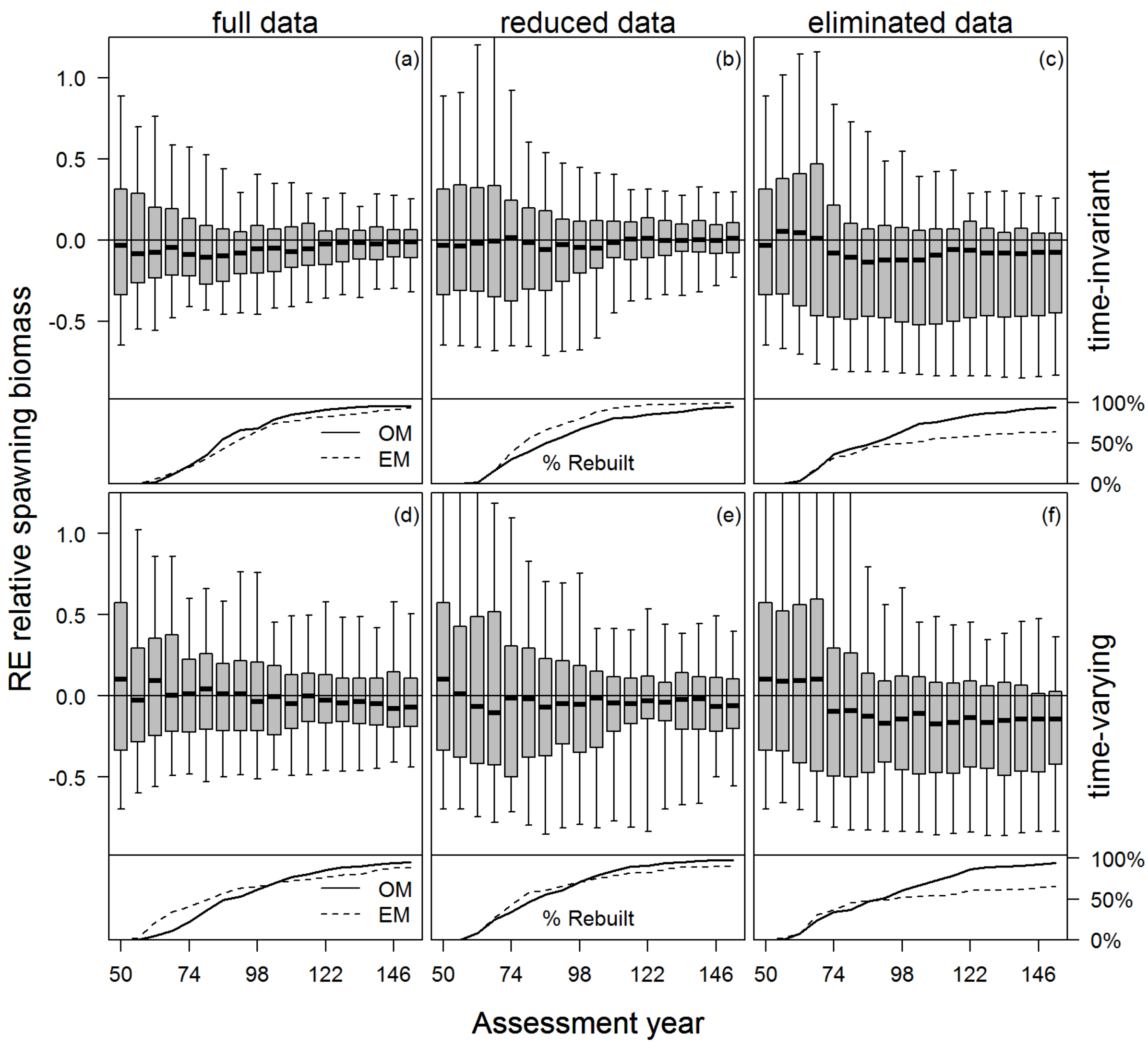

Figure 1.4: Relative error of estimated relative spawning biomass in each assessment year for each case and data scenario for all simulations (top panel) and the percentage of stocks that had rebuilt to the target biomass during the management period (bottom panel) within the operating model (OM; black line) and the estimation model (EM, dashed black line), with data collection consequently returning to historical levels when the EM determined the stock was rebuilt. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

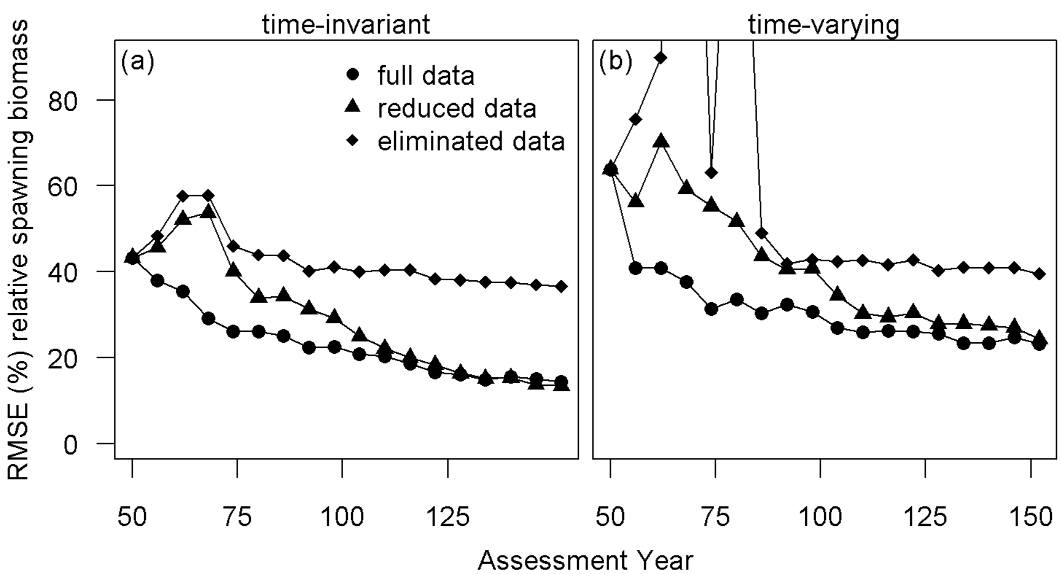

Figure 1.5: The root mean square error (RMSE) about relative spawning biomass by assessment year for each case and data scenario. The scale of the y-axis is the same for comparability of results between the time-invariant and the time-varying simulations. The time-varying eliminated data scenario peaked in year 68 at 221% RMSE.

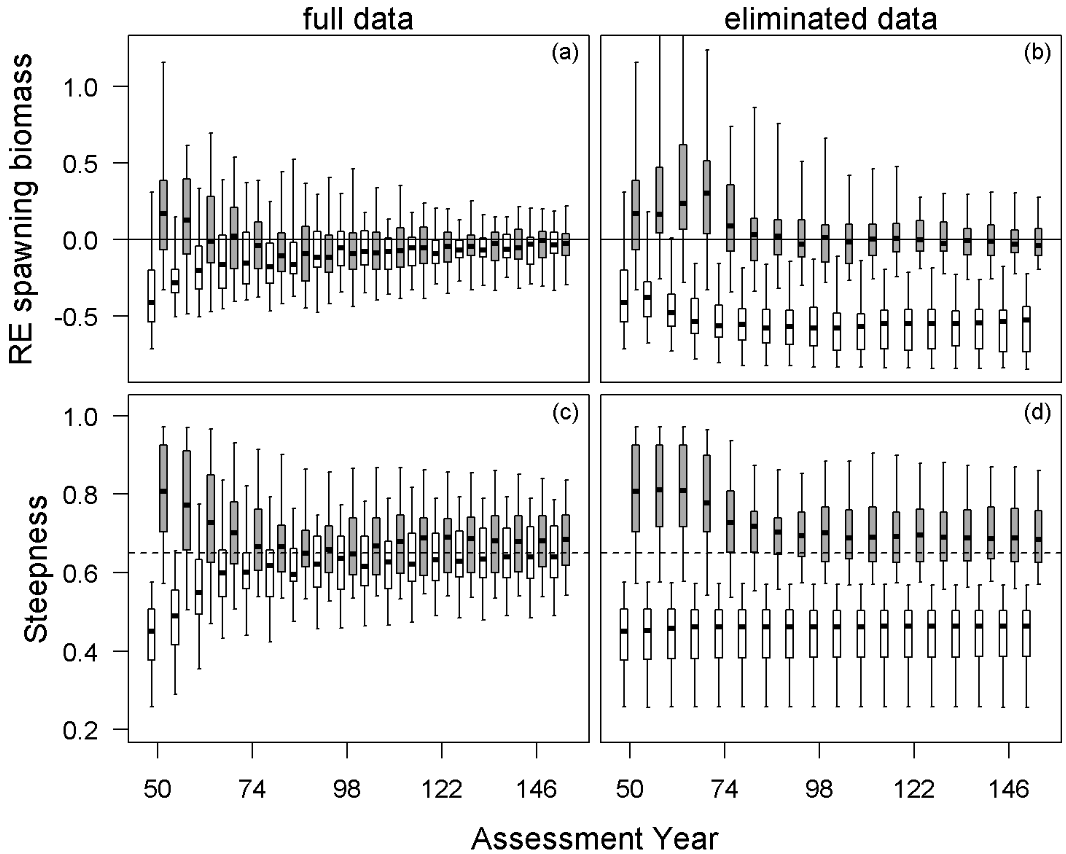

Figure 1.6: Relative error of spawning biomass and the estimates of steepness for the full and eliminated data scenarios for the time-invariant case, with the results divided by whether the simulated stock was estimated to be rebuilt (35 simulations [white]) or not (65 simulations [grey]) for the eliminated data scenario. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

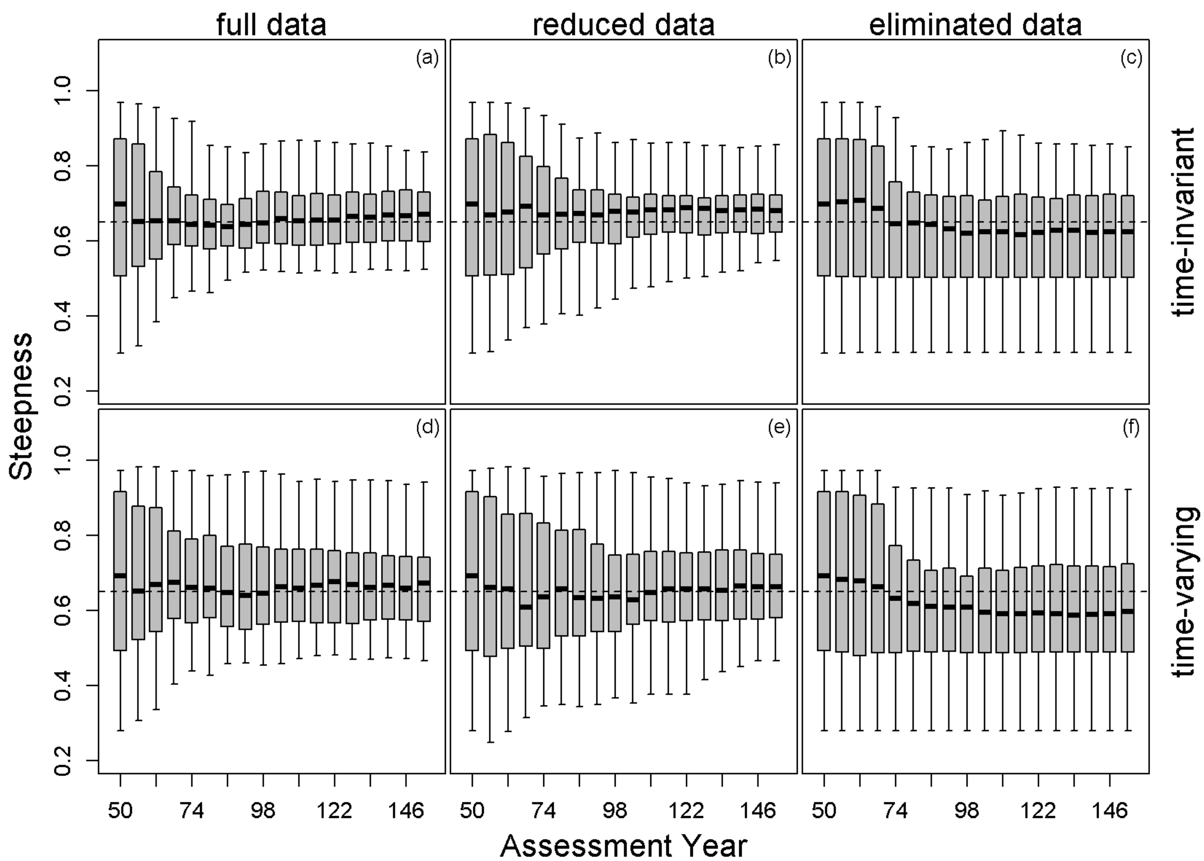

Figure 1.7: The estimates of steepness in each assessment year for each case and data scenario. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

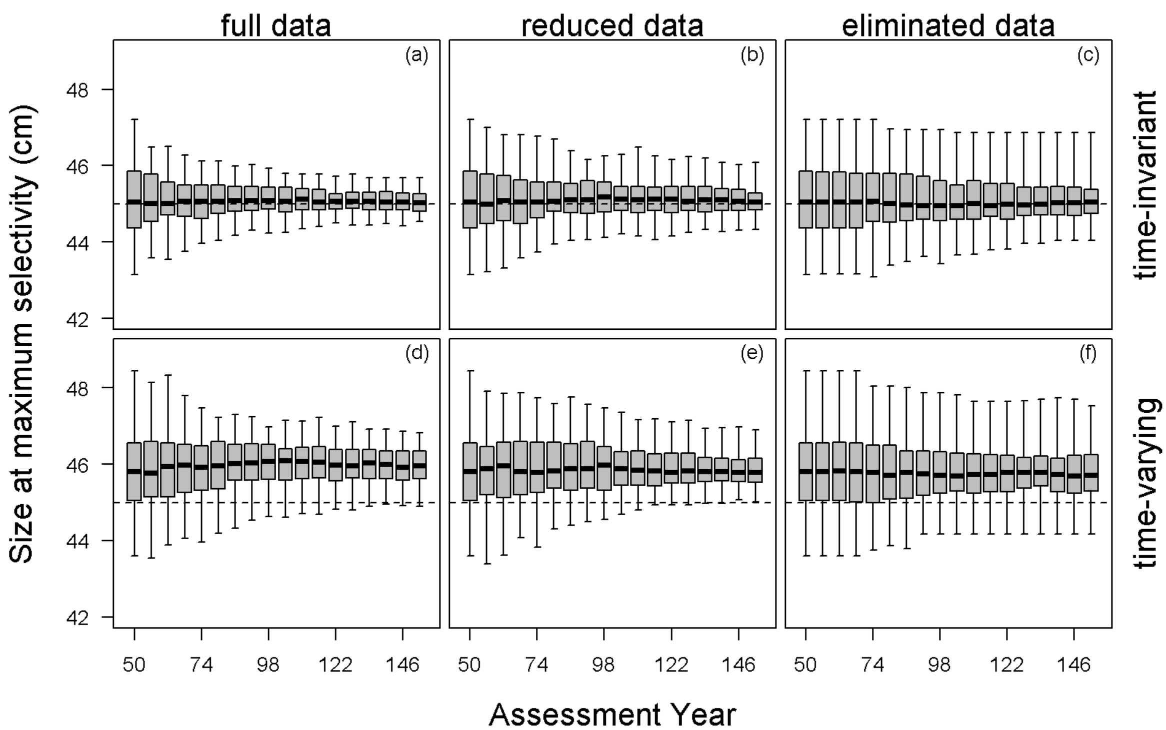

Figure 1.8: The estimated size at maximum fishery selectivity for each data scenario and case. The black dashed line indicates the operating model parameter value. The estimates from the data scenarios with time-varying selectivity were compared against the mean of the distribution from the operating model. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

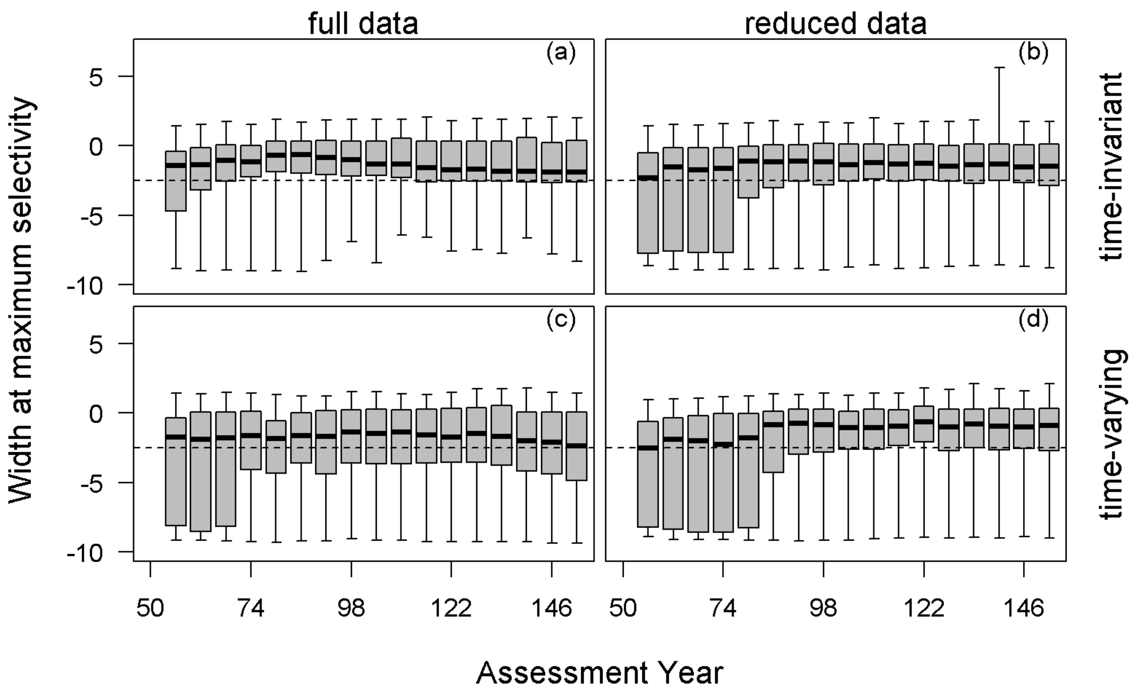

Figure 1.9: The estimated width at maximum selectivity (i.e. defines the extent of dome-shaped selectivity) for the fishery that occurred while the stock was overfished for each data scenario and case. The black dashed line indicates the operating model parameter value. The estimates from the data scenarios with time-varying were compared against the mean of the distribution from the operating model. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

Chapter 2

EVALUATING THE PERFORMANCE OF DATA-MODERATE AND CATCH-ONLY ASSESSMENT METHODS FOR US WEST COAST GROUNDFISH

Abstract

The estimation of sustainable harvest limits for stocks that have never been assessed and have limited data available can be challenging. Harvest limits for previously un-assessed US west coast groundfish have been set using catch-only methods (Depletion Corrected Average Catch [DCAC] and Depletion-Based Stock Reduction Analysis [DB-SRA]) for data-limited stocks, as well as catch and index based methods (Extended Depletion-Based Stock Reduction Analysis [XDB-SRA] and Extended Simple Stock Synthesis [XSSS]) for data-moderate stocks. To account for uncertainty and to prevent overfishing the harvest levels for US west coast groundfish are reduced based upon the estimation method and the amount of data used. A management strategy evaluation was applied to evaluate the performance of each estimation method to provide benchmark harvest levels for two life history types (US west coast flatfish and rockfish) under varying misspecifications of the parameter distributions used by these methods. Both data-moderate and catch-only (data-limited) methods resulted in overfishing > 0.50 (except XSSS in select scenarios) when the simulated stock was below the relative biomass target for both life histories. Each of the data-moderate methods (XDB-SRA and XSSS) that applied biomass index data failed to estimate the correct stock status in the first assessment when an overly optimistic prior distribution about the stock status was assumed. However, during the projection period as the biomass index lengthened, the estimates of current stock size improved for both of these estimation methods, reducing the probability of overfishing to < 0.50 (except XSSS for one scenario). The ability to incorpo-

rate index data by the data-moderate methods resulted in improved estimates, as the data became more informative, for stock status and the subsequent harvest limits, that when reduced to account for uncertainty resulted in population stock sizes that either remain stable or rebuild toward the biomass targets for both life histories. A notable exception was the performance of XDB-SRA for the flatfish life history when the stock was at the target level and the prior was assumed correctly. In this instance the index data were non-informative, resulting in overfishing due to overly optimistic estimates of relative population size.

2.1 Introduction

The Pacific Fishery Management Council (PFMC) manages federal fisheries off the US west coast. Currently, the groundfish fishery management plan includes 90+ species, of which only approximately a third have been formally assessed (Pacific Fishery Management Council, 2014b). The Council classifies groundfish into three categories based upon the type of assessment, and thus, harvest specification uncertainty: 1) category 1 stocks, where complex age-structured stock assessments that incorporate biomass indices and composition data (length and or age data) were applied and are assumed to have the lowest level uncertainty regarding status; 2) category 2 stocks, where some biological indicators are present (e.g. index data) and estimation methods that incorporate some limited data were applied and have a moderate level of uncertainty about stock status; and 3) category 3 stocks, which have limited data (e.g. catch) and harvest estimates are calculated using catch-based statistics and have a high degree of uncertainty surrounding stock status. The level of uncertainty surrounding stock status is accounted for when setting harvest limits Determining a harvest limit is typically done by two steps: 1) estimating an Overfishing Limit (OFL), a level of harvest that if exceeded would constitute overfishing and defines the maximum harvest level; and 2) setting an Acceptable Biological Catch (ABC), a level of harvest less or equal to the OFL where reductions account for scientific uncertainty in the OFL. The PFMC has termed the reduction between the OFL and the ABC as the "buffer", where the size of that buffer should be directly related to the level of uncertainty about the status which the council bases on the category of the stock (Pacific Fishery Management Council, 2010).

In 2010, the PFMC took the first step to establish harvest limits for the previously unassessed stocks included within their fishery management plan by applying catch-only harvest estimation methods, Depletion-Based Stock Reduction Analysis (DB-SRA) (Dick and MacCall, 2011) and Depletion Corrected Average Catch (DCAC) (MacCall, 2009). These methods are applicable in "data-free" situations because they require only limited information to estimate harvest levels. DB-SRA uses a full catch time series, a pre-specified distribution of relative biomass, and biological parameter distributions to determine the unfished biomass level required to produce the pre-specified ending relative biomass level resulting in estimates of the OFL, while DCAC calculates a one-time estimate of the distribution for a yield that would likely be sustainable given the sum of the known catch history and assumed biological parameter distributions. Stocks where harvest levels have been estimated by either DB-SRA or DCAC are considered to have the highest uncertainty about their status (classified as category 3) and subject to the largest reduction between the OFL and the ABC. While this was an important first step in calculating ABCs for West Coast groundfish, ideally the estimation of OFLs would be based on methods that incorporate data where available. The benefit of applying data based estimation methods would be an improved estimate of stock status with a reduced uncertainty, resulting in less of a reduction between the OFL and ABC, minimizing "lost harvest". Currently, stocks with high uncertainty (category 3) result in ABC values that are reduced from the OFL by a buffer value of 0.69 (ABC = 0.69 x OFL), while stocks with a moderate level of uncertainty (category 2) apply a buffer value of 0.83 to reduce the OFLs[1].

In 2013, two estimation methods that incorporate trend information, Extended Depletion-Based Stock Reduction Analysis (XDB-SRA) and Extended Simple Stock Synthesis (XSSS), were developed and applied to West Coast groundfish to estimate stock status and determine

[1]The buffer values here are the default values applied by the PFMC based upon stock category. However, the PFMC has applied alternative buffer values that resulted in larger reductions between the ABC and OFL for stocks within each category when a higher level of precaution was deemed warranted.

OFLs (Cope et al., 2015). Each method, similar to DCAC and DB-SRA, applies user-defined biological parameter and relative biomass prior distributions, but XDB-SRA and XSSS incorporate biomass indices to update these prior distributions. The incorporation of indices classifies these estimation methods as data-moderate (category 2) stock assessments, defined by the PFMC for West Coast groundfish. The underlying structure of XDB-SRA mimics DB-SRA by applying a delay-difference model that assumes equal growth between the sexes with age-based knife-edge maturity and selectivity, but XDB-SRA allows for updating of the prior distributions based upon model fits to observed index data (Pacific Fishery Management Council, 2014c). Extended Simple Stock Synthesis (XSSS) involves a simplified implementation of Stock Synthesis, an integrated statistical catch-at-age model (Methot and Wetzel, 2013). XSSS is an age-structured model that assumes length-based maturity (although it can be parameterized to be age-based) with selectivity equal to maturity (Pacific Fishery Management Council, 2014c). Similar to XDB-SRA, XSSS calculates model fits to observed index data, which allows for updating of the prior distributions to estimate stock status and harvest levels.

One of the prior distributions that must be specified by the user for each of these data-moderate and catch-only (data-limited) approaches is the relative biomass level in a specific year, a quantity that is often an estimated result from an assessment. Several simulation studies have evaluated the performance of the catch-only methods (DCAC and DB-SRA) in regards to the assumed prior distributions at estimating sustainable levels of harvests (Wetzel and Punt, 2011a; Carruthers et al., 2014). Wetzel and Punt (2011a) determined that the OFL can be overestimated when the prior distribution for relative biomass is specified higher than the true value. However, catch-only assessments (DCAC and DB-SRA) are subject to the greatest reduction in the OFL by a buffer value of 0.69 to account for uncertainty. Therefore, if this reduction is large enough, prior misspecification does not necessarily lead to overfishing when the estimated OFL is too great.

Similar to the catch-only methods, the data-moderate approaches, XDB-SRA and XSSS, also apply an assumed prior distribution for relative biomass. However, XDB-SRA and

XSSS differ from the catch-only methods by fitting to biomass indices to update the prior distribution based on the data. If the indices are informative, this process should reduce the influence of the user-specified priors, even when misspecified, resulting in posterior distributions based more strongly on the data and improving the OFL estimates. Data-moderate estimated OFLs are reduced by 0.83 to set ABCs, resulting in potentially higher harvest limits relative to the catch-only methods due to the reduced uncertainty about stock status and therefore potentially result in increased harvest limits relative to catch-only methods.

This paper applies a management strategy evaluation approach to evaluate the performance of data-moderate (XDB-SRA and XSSS) and catch-only estimation methods (DCAC and DB-SRA) and the subsequent ABC calculations for the management of West Coast groundfish stocks. The estimation methods are used to set ABCs for two simulated stocks, a flatfish and a rockfish, for a 25-year projection period. The status of the operating model population and error in parameter estimates at the end of the projection period are evaluated, and compared among methods to address three questions: 1) are the reductions applied to the OFLs sufficient to prevent overfishing for the data-moderate and catch-only methods; 2) what are the relative risks when parameters are misspecified for both data-moderate and catch-only methods; and, finally, 3) what are the potential trade-off for performing either a data-moderate or catch-only assessment?

2.2 Materials and Methods

2.2.1 General approach

The population was modeled using an age-structured operating model. An annual biomass index was observed with error for selected years, and was used by each index-based estimation method (XDB-SRA and XSSS) to estimate population size and OFLs. The estimated OFLs were then adjusted according to a pre-determined buffer (data-moderate 0.83 and catch-only [data-limited] 0.69) to determine ABCs. US federally managed fisheries are required to set an Annual Catch Limit (ACL) value which can be set equal to the ABC or reduced further to account for additional uncertainty (e.g. management). In this simulation the ACL was set equal to the ABC and then applied to the simulated stock. For clarity, the level of harvest removed from the stock will be referred to as the ABC rather than an ACL. The data generation, OFL estimation and stock updating were conducted in an iterative fashion for twenty-five years (Fig. 2.1).

Two life history types that are common to US west coast groundfish were simulated; a fast growing, short-lived flatfish and a slow growing, long-lived rockfish (Table 2.1).

2.2.2 The operating model

An age-structure population was simulated with stochastic recruitment subject to annual fishery removals where an observed biomass index with observation error was available for select years (See *Appendix B* for operating model details). The population in the operating model at the start of year 51 (immediately prior to the start of management based on OFLs and ABCs) was defined to be a specified proportion of the virgin biomass, with relative biomass level depending on the life history and the simulation scenario (see *Scenarios* section below).

The population was assessed using each estimation method at the start of year 51 based on an annual survey biomass index for years 31-50. The index length was a strategic selection because it covered a period long enough to offer contrast, if present, in the trajectory of the population. The catch-only estimation methods calculate either a single OFL (DCAC) or an OFL for each projection year (DB-SRA) and were not reapplied in the projection period because they do not use index data and do not update the prior distributions. In contrast, the data-moderate estimation methods (XDB-SRA and XSSS) estimated biomass and OFLs for the subsequent four years based on the available index data. A buffer, as applied by the PFMC (approximately 0.83 for category 2 data-moderate stock estimation methods [XDB-SRA and XSSS], and approximately 0.69 for category 3 catch-only stock estimation methods [DCAC and DB-SRA]), was applied to each OFL to calculate the ABCs. The ABCs were removed without error and the population projected forward for four years.

An additional four years of index data were then generated and provided to each data-moderate estimation method, which then re-assessed the population and calculated OFLs. This iterative process continued over twenty-five years, at which point the performance of the estimation methods were evaluated using performance measures (see *Performance measures* section below). DCAC, DB-SRA, and XDB-SRA were applied separately to each of the 100 simulated operating model populations. XSSS was applied to 60 simulated operating model populations. The number of XSSS simulations conducted was limited because it was more time intensive (due to it being age-structured) compared to the other estimation methods.

2.2.3 Estimation methods

DCAC

DCAC (MacCall, 2009) allows for the estimation of a likely sustainable yield for data-limited stocks based upon average observed catches, distributions for three biologically-based life history parameters, and assumed distribution for relative stock status. DCAC calculates a yield as:

$$\text{Sustainable Yield} = \frac{\sum_{t=1}^{N} C_t}{n + (1-\delta_{50})\left(\frac{B_{MSY}}{B_0}\frac{F_{MSY}}{M}M\right)^{-1}} \tag{2.1}$$

where C_t is the catch during year t, n is the length of the catch history, δ_{50} is the relative biomass in year 50, B_{MSY}/B_0 is the biomass that corresponds to maximum sustainable yield relative to carrying capacity, M is the instantaneous rate of natural mortality, and F_{MSY}/M is the ratio of the fishing mortality rate that corresponds to and M. Equation 2.1 has been re-parameterized from the version in MacCall (2009) where the original version applied a parameter (defined as the difference between the relative biomass at a previous point in time and the current status) which is 1- δ_{50}, assumed to be the relative biomass from the unfished state. This re-parameterization has been applied to DCAC, DB-SRA, and XDB-SRA to create consistency among all methods and how the PFMC defines their target relative biomass levels.

A Monte Carlo approach was applied to account for the parameter uncertainty. For each life history type and simulation scenario a total of 10,000 random draws were conducted based on distributions for each parameter, which generated a distribution of sustainable yield values.

DCAC was applied at the start of year 51 to calculate a yield based on specified distributions for each life history type and scenario (Tables 2.2 and 2.3). The median value of the resulting distribution was then applied as the OFL. The OFL was adjusted to an ABC value by applying the buffer (0.69*OFL) that was then removed from the stock for each of the 25 future years. DCAC was not reapplied during the projection period because this method defines a one-time only calculation of yield, and is not an update-able calculation.

DB-SRA

DB-SRA is based on stock reduction analysis (SRA) (Kimura et al., 1984; Walters et al., 2006). Dick and MacCall (2011) adapted the concepts with the addition of a relative biomass parameter. Similar to DCAC, DB-SRA uses Monte Carlo draws from four parameter distributions (for M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) to create probability distributions for current relative biomass and OFLs. DB-SRA is based on the following delay-difference production model that includes a time lag for recruitment and mortality:

$$B_t = B_{t-1} + P_t(B_{t-amat}) + (1 - e^{-M})(B_{t-amat} - B_{t-1}) - C_{t-1} \tag{2.2}$$

where B_t is the biomass at the start of year t, M is the instantaneous rate of natural mortality, and $P(B_t - amat)$ is the latent annual production based on a function of biomass in year $t - amat$, where $amat$ is the age at maturity. The latent annual production is determined by a hybrid between the Pella-Tomlinson-Fletcher model and the Schaefer surplus production model (see Dick and MacCall (2011) for additional detail).

Unfished biomass, B_0, is calculated separately for each parameter draw from the distributions by solving the equation $B_{t=50}/B_0 = \delta_{50}$ for B_0. The estimated OFL value for each year t is calculated as:

$$OFL_t = \left(1 - e^{-(M+F_{MSY})}\right)\left(\frac{F_{MSY}}{M + F_{MSY}}\right)B_t \tag{2.3}$$

This estimation method produces estimates of annual biomass and the corresponding OFLs for the entire projection period (years 51-75). The vector of serial OFLs was reduced by the PFMC category 3 data-limited buffer factor, 0.69, to produce ABCs that were annually removed from the simulated population.

XDB-SRA

XDB-SRA (Pacific Fishery Management Council, 2014c) builds upon the basic structure of DB-SRA, using the same four parameters and underlying population dynamics model. However, unlike DB-SRA, it updates the prior distributions for the parameters by fitting the model to a biomass index using adaptive importance sampling (see *Adaptive Importance Sampling* section). The application of adaptive importance sampling was based on 1,000 initial population trajectories created from draws for each of the four prior distributions and 1,000 draws at each step (except the last draw which was based on 2,000 draws). The parameter related to the relative biomass (δ_{50}) was always assumed to pertain to year 50 for consistency with how DCAC and DB-SRA were applied, even though XDB-SRA was applied in future years, unlike the catch-only estimation methods. The OFLs for the next four years were set equal to the median of the distribution derived from the estimated population size from each trajectory.

The ABCs for XDB-SRA were determined using the PFMC harvest control rule policy. The OFLs were reduced based upon the buffer value (0.83) to determine the ABCs for category 2 data-moderate assessments. Additionally, the ABCs were adjusted downward based on a life history specific harvest control rule which applies a linear reduction in harvest when a stock is below a pre-specified target relative biomass level. Harvest was reduced when the flatfish stock was below 0.25 of virgin biomass (B_0) or the rockfish stock was below $0.40B_0$, with no fishing when the relative biomass level was below the life history specific threshold value (flatfish: $0.05B_0$, rockfish: $0.10B_0$). The appropriate life history harvest control rule for West Coast groundfish was applied based on the posterior median relative biomass value. The ABCs were removed without error from the simulated population.

XSSS

XSSS simplifies Stock Synthesis for application in data-moderate situations (Pacific Fishery Management Council, 2014c). The population model underlying XSSS is sex- and age-structured, with a Beverton-Holt stock recruitment relationship, length-based maturity, with fishery and survey selectivity assumed equal to maturity. The Bayesian analyses were based on equivalent biological parameters to those used in XDB-SRA: M the instantaneous rate of natural mortality, h (also known as steepness), and the relative biomass in year 50 (δ_{50}). This parameterization has been assumed when applying XSSS for the West Coast data-moderate assessments (Cope et al., 2015) and in simpler extensions (Cope et al., 2013). The application of adaptive importance sampling for XSSS was based on 1,200 initial population trajectories created from draws for each of the four prior distributions and 700 draws at each step (except the last draw which was 1,200). The fewer draws for XSSS than XDB-SRA reduced model estimation time while preserving convergence in the posterior distributions. The value for $\log(R_0)$ was calculated so that the generated and model-predicted relative biomass levels for year 50 were the same for each parameter draw. The estimated OFLs for the next four years were calculated from the medians of the posterior distributions for the biomass trajectory as for XDB-SRA. The harvest control rule and the calculation of ABCs for category 2 data-moderate stocks were applied the same as for XDB-SRA.

2.2.4 Adaptive importance sampling

Adaptive importance sampling was applied to both XDB-SRA and XSSS to update prior parameter distributions based upon the fit to the index data. Adaptive importance sampling grew out of the foundation of sampling importance resampling (SIR) (Rubin, 1987, 1988). SIR samples parameter vectors from a prior distribution taken from a sampling envelope and has been applied in fishery stock assessment for parameter estimation (e.g. McAllister et al., 1994). However, the adaptive importance sampling approach updates the sampling envelope based upon iterative SIR draws and can be beneficial when the best sampling envelope is unknown or not well understood *a priori* due to correlation among parameters (Givens and Raftery, 1996; Kinas, 1996). See *Appendix B* for technical details.

2.2.5 Scenarios

Multiple scenarios were created to explore the performance of each estimation method and the sensitivity to parameter misspecification (Tables 2.2 and 2.3):

- Scenario T_{at}: All prior distributions were centered about the true values (T) and the simulated stock was at the target relative biomass (at) as defined by the PFMC based on the life history (flatfish 25%, rockfish 40% of virgin biomass) in year 50.

- Scenario $\mathrm{T}_{\mathrm{below}}$: All prior distributions were centered about the true values (T) and the true stock was below the target relative biomass (below) as defined by the PFMC based on the life history (flatfish 10%, rockfish 20% of virgin biomass) in year 50.

- Scenario $\mathrm{D}_{\mathrm{below}}$: The prior distribution for the relative population size in year 50 (δ_{50}) was centered about an optimistic value relative to the true value (D: relative biomass misspecified) and all other prior distributions were centered about the true values. The true stock was below the target relative biomass level (below) defined by the PFMC based on the life history (flatfish 10%, rockfish 20% of virgin biomass) in year 50.

- Scenario $\mathrm{DP}_{\mathrm{below}}$: The prior distributions for the productivity parameter (F_{MSY}/M or h) and the relative population size in year 50 (δ_{50}) were each centered about optimistic values relative to the true values (DP: relative biomass and productivity misspecified).

The true stock was below the target relative biomass level (below) defined by the PFMC based on the life history (flatfish 10%, rockfish 20% of virgin biomass) in year 50.

A challenge for any estimation method is determining which parameters need to be fixed or estimated, due to the lack of information in the data or correlations among the parameters. XDB-SRA and XSSS allow for the estimation of key biological parameters through the application of adaptive importance sampling, but it is uncertain what information the data (a biomass index) may contain to inform each estimation method to correctly estimate these parameters. Specific scenarios explore the performance of the estimation methods when the distribution for the relative stock status in the target year was misspecified and when this parameter was misspecified along with the productivity parameter. The final scenario, that evaluates the misspecification of the relative stock status and productivity parameter, was designed to investigate the impact of multiple misspecifications on estimation performance and whether precaution in setting the productivity parameter would offset the bias in the specification of relative biomass level. Similar model runs could have been conducted to explore the misspecification of the priors for M and B_{MSY}/B_0 (DCAC, DB-SRA, XDB-SRA), but they were beyond the scope of the current analysis.

2.2.6 Performance measures

The mean of the estimated parameter distribution for each of the model's input parameters was evaluated (DCAC, DB-SRA, and XDB-SRA: M, F_{MSY}/M, B_{MSY}/B_0 and δ_{50}; XSSS: M, h, and δ_{50}). The choice to summarize by the mean rather than the median for the parameter inputs was based upon the prior parameter distributions, which were specified by a mean value and were either assumed to be lognormal or beta distributed. If the data provided no information and the prior was not updated, resulting in a posterior equal to the prior, summarizing the distributions by the median would result in estimates that would differ from the input distribution values that generated the prior. The model-derived estimates (spawning biomass, current relative biomass, ABCs) were summarized by the medians of their distributions. The relative error of the estimates was calculated as:

$$RE = \frac{E - T}{T} \tag{2.4}$$

where E is the estimated value and T is the corresponding true value from the operating model.

The performance of each estimation method was evaluated using the following criteria:

1. Relative biomass at the end of the projection period relative to management targets.

2. The probability of overfishing (ABC $>$ true OFL).

3. The probability of being overfished (flatfish: $< 0.125B_0$, rockfish: $< 0.25B_0$) at the end of the projection period.

4. The percent of the estimated catch realized over the projection period relative to the operating model OFL by each estimation. A perfect performance for the catch-only estimation methods would be 69% (ABC = 69% OFL) and 83% (ABC = 83% OFL) for the data-moderate estimation methods.

5. The average annual variability of ABC catches (abbreviation AAV) over the projection period defined as:

   $$AAV = 100\frac{\sum_{t=1}^{N} |C_t - C_{t-1}|}{\sum_{t=1}^{N} C_t} \tag{2.5}$$

   where C_t is the catch during year t.

To evaluate estimation performance, simulations were conducted where the populations were managed with perfect knowledge and independent of the estimation methods. The "perfect knowledge" results provide information on the best possible performance given the current management guidelines. The true OFL and the resulting ABC, based on the operating model population status were determined and removed without error.

2.3 Results

2.3.1 Overview

All estimation methods resulted in overfishing when the δ_{50} prior was misspecified for both life history types. This was most notable for DB-SRA. The catch-only methods (DCAC and DB-SRA) generally avoided overfishing when the stocks were correctly judged to be at the relative biomass target at the time of the first assessment in year 51, estimating OFLs that resulted in ABC values at or below the true OFLs. The data-moderate estimation methods (XDB-SRA and XSSS) resulted in very different results when the stock was at the target biomass level with XDB-SRA estimating OFLs and resulting ABCs that were well above the true OFLs. All methods, except DCAC, resulted in overfishing when the stock was below the target biomass level at the time of the first assessment when the δ_{50} prior was misspecified at an overly optimistic value. However, within 10-15 years, the index data used by XDB-SRA and XSSS were sufficient to result in updated posterior distributions for the δ_{50} prior that estimated OFLs that produced ABC values which allowed the stocks to rebuild.

2.3.2 Failed simulations

A small number of simulations 'failed' for XDB-SRA and XSSS because a sample was drawn which resulted in a singular covariance matrix. The failed simulations only occurred for the flatfish life history. This behavior was commonly observed later in the projection period and was attributed to the parameters entering a narrowly defined parameter space which led to very few unique parameter combinations being supported by the data, resulting in repeated draws by adaptive importance sampling. The percent of simulations on which the results are based for XDB-SRA by scenario were; T_{at}: 99%, T_{below}: 90%, D_{below}: 94%, and DP_{below}: 100% and for XSSS by scenario were; T_{at}: 95%, T_{below}: 85%, D_{below}: 72%, and DP_{below}: 85%.

2.3.3 Flatfish life history

Scenario T_{at}

The status of the simulated stocks at the end of the projection period were variable within and among estimation methods for the $\mathrm{T_{at}}$ scenario, where all priors were assumed correctly and the simulated relative biomass was at the management target at the start of the projection period (Table 2.2). A high percentage of the stocks where XDB-SRA was applied for estimation were below the relative biomass target value ($0.25B_0$) in year 75, with a large subset of those below the overfished level ($0.125B_0$) (Table 2.4). In contrast, at the end of the projection period, DCAC, DB-SRA, and XSSS resulted in either none or a low percentage of simulated stocks below the target level (Table 2.4). The median operating model relative biomasses at the end of the projection period for these three methods were above the target levels (Fig. 2.2a). The OFLs and the resulting ABCs set by DCAC, DB-SRA, and XSSS were well below the true OFLs, while XDB-SRA substantially overestimated the true OFLs resulting in a probability of overfishing that was > 0.50 for each year the stock was assessed (Fig. 2.3a and Table 2.4). The median AAV for XDB-SRA was highest in the first five years of the projection period (Fig. 2.4a). However, the average catch increased over the projection period (Fig. 2.4a), resulting in the observed median population decline (Fig. 2.2b).

XDB-SRA and XSSS resulted in dissimilar median posterior estimates for several of the parameters. The medians of the posterior means for the δ_{50} parameter for XDB-SRA were consistently greater than the true value throughout the projection period (Fig. 2.5c), although this tendency declined over time. There was little contrast in the biomass index for the flatfish life history for years 31-50, thus the index data were insufficient to inform stock status, causing a bias for the relative biomass at the end of year 50. The medians of the posterior means for M and F_{MSY}/M from XDB-SRA were unbiased for the first two assessments, but for later assessments the among-simulation variance increased and most of the posterior means were biased low (Fig. 2.5a-b). Overall, the median of the posterior medians for the estimated assessment year relative biomass for XDB-SRA were greater than the true values and the among-simulation variance increased over the projection period despite the increased amount of data (Fig. 2.5e). The posterior distribution means for M and h from XSSS were generally centered about the true values for all years (Fig. 2.6a-b). The posterior distributions for relative stock status at the end of year 50 from XSSS emphasized values larger than the true value at the start of the projection period but the posterior medians were less than the true value by the time of the last assessment in year 74 (Fig. 2.6c). The distribution of posterior medians for the assessment year estimated relative biomass was unbiased for the first two assessments, but distributions became biased low relative to true stock status over the projection period (Fig. 2.6d).

Scenario T_{below}

The median relative biomass was below the target level at the end of year 50, and all prior distributions were unbiased for the T_{below} scenario (Table 2.2). This scenario resulted in a high percentage of stocks that were still below the management target level at the end of the projection period, with a portion of them below the overfished threshold or even extinct depending upon the estimation method (Table 2.4). The median time-trajectory of relative biomass varied among estimation methods for this scenario, with the estimated OFLs and the resulting ABCs set using DCAC, XDB-SRA, and XSSS leading to increases in relative biomass, while the median relative biomass of the populations where DB-SRA was applied declined towards zero by year 75 (Fig. 2.2c-d). DCAC and XDB-SRA set ABCs that were greater than the true OFLs during the early part of the projection period, resulting in a high probability of overfishing (approximately 1 and 0.75) (Fig. 2.3b). However, the probability of overfishing for DCAC declined to < 0.50 by approximately year 60 (Fig. 2.3b). DB-SRA estimated OFLs that resulted in ABCs that were greater than two times the true value OFLs resulting in a declining population (Fig. 2.2c and Table 2.5), and in the highest probability of overfishing among the estimation methods (Fig. 3b). XDB-SRA and XSSS resulted in similar median AAVs, but XSSS had less among-simulation variation in both average catch and AAV compared to XDB-SRA (Fig. 2.4b).

The medians of the posterior means were more variable and further from the true values for XDB-SRA compared to XSSS. The medians of the posterior means for M, F_{MSY}/M, and B_{MSY}/B_0 from XDB-SRA were less than the true values and error generally increased over time (Fig. 2.5f, 2.5g, and 2.5i). The first assessment resulted in the largest among-simulation variation and medians of the posterior means that were well below the true value for the δ_{50} parameter from XDB-SRA (Fig. 2.5h). The posterior distributions for estimated assessment year stock status from XDB-SRA were greater than the true value at the start of the projection period, but approached the true value with each subsequent assessment (Fig. 2.5j). The XSSS posterior distributions for M were less than the true value for all assessments (Fig. 2.6e). The posterior distributions for h had most of their mass less than the true value for the first assessment with increasing inter-simulation variation in the posterior medians over time (Fig. 2.6f). The medians of posterior distributions for the relative biomass in year 50 varied among assessments (Fig. 2.6g). The distribution of the posterior medians for the estimated assessment year stock status were roughly centered about the true value until the final two assessments (year 70 and 74) at which point the posterior medians were less than the true value (Fig. 2.6h).

Scenario D_{below}

The population was below the relative biomass target, and the prior for this parameter was misspecified at an overly optimistic value for scenario $\mathrm{D}_{\mathrm{below}}$ (Table 2.2). The results for this scenario were qualitatively similar to scenario $\mathrm{T}_{\mathrm{below}}$, with the following noteworthy exceptions. Specifically, DB-SRA resulted in high percentage of extinct stocks by the end of the projection period due to a high rate of overfishing (Fig. 2.2c and Tables 2.4-2.5) while XSSS led to overfishing with > 0.5 probability until year 65 (Fig. 2.3c) resulting in almost half of stocks being below the target level at the end of the projection period (Table 2.4). The posterior means from XSSS are notably more variable among simulations for scenario $\mathrm{D}_{\mathrm{below}}$ than for scenario $\mathrm{T}_{\mathrm{below}}$ (Fig. 2.6i-l vs. 2.6e-h), and in contrast to scenario $\mathrm{T}_{\mathrm{below}}$

the posterior means for the relative biomass in year 50 were estimated greater than the true values for each assessment year (Fig. 2.6k vs. 2.6g). The variability observed in the distributions for XSSS is due in part to the smaller simulation size for this scenario (43) relative to the other scenarios (See *Failed Simulations* section for details).

Scenario DP_{below}

The prior distributions for the productivity parameter and for the relative stock status in year 50 are centered on incorrect values, and the stock was below the relative biomass target for scenario $\mathrm{DP}_{\mathrm{below}}$ (Table 2.2). The results for this scenario in terms of relative stock status and probability of overfishing for DCAC, DB-SRA, and XDB-SRA were generally similar those for scenario $\mathrm{T}_{\mathrm{below}}$ (Figs. 2.2c, 2.2d, and 2.3b). In contrast, XSSS performed poorer compared to scenario $\mathrm{T}_{\mathrm{below}}$ (Fig. 2.2h vs. 2.2d) with estimated ABCs exceeding the true OFLs over the projection period (Table 2.5) resulting in 49% of the simulated populations below the relative biomass target at the end of the projection period (Table 2.4).

XDB-SRA and XSSS resulted in varying updating in the misspecified prior distributions. The medians of the posterior means for F_{MSY}/M from XDB-SRA were less than the true value for all years, showing little evidence for updating (Fig. 2.5q). The median of the posterior means for the δ_{50} parameter, which was also misspecified, was furthest from the true value for the first assessment and improved with each subsequent assessment (Fig. 2.5r). Although there was little updating for each of the misspecified priors, the medians of the posterior distributions for estimated assessment year relative stock status were similar to those observed in scenario $\mathrm{T}_{\mathrm{below}}$ (Fig. 2.5j vs. 2.5t) which was achieved through shifting the distributions of the other estimated model parameters. The posterior distribution for h, one of the misspecified parameters, from XSSS showed no evidence of updating towards the correct value (Fig. 2.6n). The median of the posterior means for the δ_{50} parameter in year 50, the prior which was misspecified, was furthest from the true value in the first assessment, but declined with subsequent assessments towards the true value (Fig. 2.6o). In contrast to XDB-SRA, the medians of the posterior distributions for the estimated assessment year relative stock status were much more variable among simulations and shifted to values greater than the true value compared to that observed in scenario T_{below} (Fig. 2.6p vs. 2.6h).

2.3.4 Rockfish life history

Scenario T_{at}

The median final relative biomass was above the target level for DCAC and DB-SRA, when the assumed prior distributions were centered on the true values and the stock was at the target level at the start of the projection period (Scenario T_{at}) (Fig. 2.7a and Table 2.3). The median final relative biomass was below the target level when management was based on XDB-SRA and XSSS (Fig. 2.7b), with a majority of the simulations below the target level at the end of the projection period (Table 2.6). The probability of overfishing occurring over the projection period was low for all estimation methods except for XDB-SRA, for which this probability increased to more than 0.50 by the end of the projection period (Fig. 2.8a and Table 2.7). The median AAV was comparable between XDB-SRA and XSSS, and was consistent over time (Fig. 2.9a). XDB-SRA had the highest average catch, which resulted in overfishing and a decline in the median population trajectory, while DCAC had the lowest average catch (Fig. 2.9a).

The medians of the posterior distributions for parameters were similar for both XDB-SRA and XSSS. The medians of the posteriors means for M, F_{MSY}/M, B_{MSY}/B_0 and were generally centered about the true value (Fig. 2.10a, 2.10b, and 2.10d). The among-simulation variance of the parameters increased over time, a counter-intuitive result given that the estimation method has more data as time goes on. This could, however, be attributed to the inconsistencies between the model assumptions and the way the data are generated. The medians of the posterior means for the δ_{50} parameter for XDB-SRA were greater than the true value (Fig. 2.10c) so the estimation method often assumed a more optimistic relative biomass level for the population when setting OFLs. The variability in the posterior means also increased over time for XSSS, with the posterior means being closest to the true values for the first assessment (Fig. 2.11a-c). Similar to XDB-SRA, XSSS estimates an overly optimistic relative biomass level in year 50, and the outlook of year 50 relative biomass becomes more optimistic as the projection continues (Fig. 2.11c). The result of this is an increase in the among-simulation variance for the estimates of estimated assessment year stock status with the median of the posterior medians being greater than the true value (Fig. 2.11d).

Scenario T_{below}

The median final relative biomass was below the target level at the start of the projection period and all prior distributions were centered about the correct values for scenario $\mathrm{T}_{\mathrm{below}}$ (Table 2.3). The time-trajectory of median biomass recovered slowly towards the target level for all estimation methods in this scenario (Fig. 2.7c-d). However, all of the simulated stocks were still below the target level at the end of the projection period due to the slow dynamics of the long-lived rockfish with a low steepness value governing population growth, where many simulations led to populations below the overfished threshold by the end of the projection period (Table 2.6). DCAC, DB-SRA, and XDB-SRA resulted in high probabilities of overfishing at the start of the projection period (Fig. 2.8b). The input assumption for DCAC and DB-SRA was that the stock was at or near a target value and not in need of rebuilding, which was not the case here and consequently resulted in an overly optimistic OFLs and hence ABCs (Table 2.7). DCAC and DB-SRA resulted in slightly higher average catches over the projection period, while XDB-SRA and XSSS had slightly lower but more variable catches (Fig. 2.9b and Table 2.7).

The posterior means for the estimated parameters for XDB-SRA were generally less than the true values and the magnitude of the errors increased over time, except for δ_{50} (Fig. 2.10f-i) The posterior means for δ_{50} were above the true value at the time of the first assessment but decreased towards the true values over the projection period (Fig. 2.10h). The posterior medians for the estimated assessment year stock status from XDB-SRA were above the true value (Fig. 2.10j). However, the mass of the distribution shifted to encompass the true value by the third assessment in year 58. The posterior means for the productivity parameter, h, from XSSS were consistently less than the true value following the first assessment (Fig. 2.11f). A similar pattern was observed for the distribution of the posterior means for M, which were near the true value in year 50, but shifted to lower values in all subsequent years (Fig. 2.11e). However, the median of the posterior medians for the estimated assessment year stock status from XSSS were only slightly above the true values, with the bulk of the distribution encompassing the true value for all years (Fig. 2.11h).

Scenario D_{below}

The population was initially below the target level and the mean of the prior distribution for the stock status parameter in year 50 was assumed incorrectly in scenario $\mathrm{D_{below}}$ (Table 2.3). Qualitatively, the results for XDB-SRA and XSSS were similar to those for scenario $\mathrm{T_{below}}$, with median relative biomass trajectories that were slowing increasing (Fig. 2.7f), although each method estimated OFLs that resulted in ABCs above the true values for the first few assessments, resulting in probabilities of overfishing > 0.50 for the corresponding period (Fig. 2.8c). In contrast, DCAC and DB-SRA performed poorer than for scenario $\mathrm{T_{below}}$, with a high probability of overfishing ($>$0.60) (Fig. 2.8c and Table 2.7), and a majority of the simulated populations below the overfished threshold by the end of the projection period (Table 2.6).

XDB-SRA resulted in overfishing and a high proportion of simulation populations below the overfished threshold at the end of the projection period because the posterior means of the δ_{50} parameter were too high during the early part of the projection period, likely due to the misspecification of this parameter and the data not being informative enough for the model to update it to the correct value (Fig. 2.10m). The posterior means for this parameter did move closer to the true value over time, although the majority of the posterior means were above the true value. Similar to XDB-SRA, the median of the posterior means for δ_{50} from XSSS were above the true value by the greatest amount in the first assessment and updated closer to the true value over time, although most of the posterior means remained above the true values (Fig. 2.11k).

Scenario DP_{below}

The population was below the target level, and the prior distributions for the productivity and the relative biomass priors in year 50 were assumed incorrectly for scenario $\mathrm{D_{Pbelow}}$ (Table 2.3). Both DB-SRA and XSSS resulted in high percentages of the simulated stocks below the overfishing threshold at the end of the projection period (Table 2.6) due to a high probability of overfishing (Fig. 2.8d), and large average catches (Fig. 2.9d and Table 2.7).

There was again little evidence of strong updating in either of the misspecified parameters, F_{MSY}/M and δ_{50} (Fig. 2.10q and 2.10r) for XDB-SRA. The estimates of h, the misspecified parameter, from XSSS updated away from the true value to favor smaller values, similar to scenario $\mathrm{T_{below}}$, although to a greater extent (Fig. 2.11n vs. Fig. 2.11f). The distribution of the posterior medians for δ_{50} in year 50 was above the true value, and moved closer to the true value in years 54 and 58, but then drifted away in the final assessments (Fig. 2.11o).

2.4 Discussion

The performance of XDB-SRA was poor for the flatfish life history, most markedly when the simulated stock was initially at the relative biomass target and all parameter distributions were specified correctly. The fast flatfish dynamics, along with the historical exploitation pattern, generally resulted in a biomass index with little contrast when the first assessment was conducted. The non-informative index data led to an overestimation of the spawning biomass, which resulted in both an overly optimistic estimate of relative biomass and OFLs. The reduction between the OFLs and the ABCs by the buffer was not large enough to prevent overfishing. This behavior continued for the subsequent assessments, with estimation performance often not improving until the true stock was sufficiently depleted to offer informative index data. The pattern of overly optimistic estimates of relative biomass for the first assessment was also evident in the scenarios where the stock was initially below the target level regardless if the prior for relative biomass was specified correctly or not, although the subsequent estimates improved fairly rapidly as the index data became more informative.

In contrast, overall XSSS resulted in median population trajectories that either maintained above or rebuilt the stock to the target level over the projection period for all scenarios for the flatfish life history. XDB-SRA and XSSS apply varying assumptions regarding the productivity of the stock and age-structure. XDB-SRA is a delay-difference model while XSSS has the advantage of full age-structured dynamics similar to the operating model. Also, while the operating model stock-recruitment relationship did not match the form assumed by either of the estimation methods, the rigid form of Shepherd stock-recruitment curve applied in the operating model is more akin to the Beverton-Holt form assumed by XSSS compared to the hybrid Pella-Tomlinson-Fletcher and Schaefer surplus production model assumed by XDB-SRA. The combined impact of an age-structured population and the stock-recruitment form for a highly productive stock (flatfish high steepness) is the likely cause for the varying performance between these estimation methods. However, XSSS also resulted in poor estimates of assessment year stock status for the first assessment in the scenarios where the relative biomass prior was misspecified (D_{below} and DP_{below}), although the posterior means were closer to the true value compared to XDB-SRA. Each of these methods failed to estimate posterior distributions that matched the true value about relative biomass when misspecified. The largest overestimation occurred at the time of the first assessment resulting in overfishing despite the reductions applied to the estimated OFLs by the category 2 data-moderate buffer.

The data-moderate estimation methods that used biomass index data, XDB-SRA and XSSS, performed similarly for the rockfish life history. There was little between-simulation variance for all methods for this life history, especially when compared to the flatfish, due to the slow rockfish dynamics and the inherent inertia of a long-lived species. The median population trajectories for the two data-moderate methods were similar to those when the stock was managed with perfect information (dotted gray lines in Fig. 2.7). The average catch attained by each of these methods was often close to that under "perfect information" for all scenarios. However, harvest estimates by XDB-SRA and XSSS each resulted in overfishing in specific scenarios. The estimated OFLs and the resulting ABCs for XSSS resulted in overfishing when the productivity and relative biomass were both misspecified (DP_{below}). Estimates for XDB-SRA resulted in overfishing when the stock was at the relative biomass target and all distributions were specified correctly (T_{at}), which resulted in an increasing probability of overfishing over the projection period (Fig. 2.8a). In this scenario for the rockfish life history the ABCs (OFLs reduced by the buffers) exceed true OFLs based upon the PFMC harvest control rule by only a small fraction resulting in very slight stock decline over the projection period. However, exceeding the OFL by any amount constitutes overfishing, a practice that fishery managers are mandated to prevent.

The catch-only methods, DCAC and DB-SRA, resulted in precautionary estimates for both life histories when they were at the target relative biomass level. These methods generally estimated lower OFL values compared to the data-moderate estimation methods, and were subject to further reductions by the category 3 buffer to determine the ABC values. However, the catch-only methods, most notably DB-SRA, generally did not perform well when the prior for relative biomass was misspecified when the stock was below the management target (D_{below}), a result consistent with previous findings (Wetzel and Punt, 2011a). In this scenario the probability of overfishing for the catch-only methods was often $>$ 0.50 despite the large reduction between the OFL and ABC. Comparing the two life histories, DCAC and DB-SRA performed somewhat better for the rockfish scenarios. Carruthers et al. (2014) drew a similar conclusion, noting that these estimation methods tended to better estimate the "windfall" biomass of the older ages classes associated with long-lived species. Although, for both life histories, DCAC generally resulted in median population trajectories that either grew or remained flat, overfishing still occurred for portions of the projection period when harvest estimates were derived based upon misspecified parameters. This method is considered a way to estimate a yield that will likely be sustainable. However, this was not always realized when the relative biomass prior was misspecified, even when the yield was reduced to account for uncertainty.

The catch-only methods were initially created as a means to improve short-term management over yield only (average catch) based harvest estimates and were not intended for long-term management. Given the general conclusion that both catch-only methods were sensitive to misspecification, especially DB-SRA, applying either for long-term management, could have undesirable properties, especially compared to the data-moderate options which have some ability to update priors based upon data. However, on the short-term (e.g. 10 years or less) each of the methods could be very useful for management. DB-SRA resulted in OFLs that when reduced to account for uncertainty avoided overfishing for each life history when the stocks were at the higher relative biomass levels. Based upon these results, this catch-only estimation method could be applied to determine short-term harvest estimates for stocks that have low historical exploitation. Also, DCAC generally resulted in OFLs that once reduced by the category 3 buffer (0.69) avoided overfishing over the long-term even when prior parameters were misspecified. Unfortunately, DCAC can only be applied for a one-time harvest calculation for a likely sustainable yield and alternative estimation methods would need to be applied in the future for updated harvest estimates, but once again this method could be applied for short-term estimates until more data can be incorporated for assessment. However, over the long-term, this work shows that it is beneficial to apply a method that can incorporate index data if available. The one notable exception to this general conclusion is the use of XDB-SRA for a life history with fast dynamics and for which there is little contrast in the index data.

The posterior distributions of the model parameters for both XDB-SRA and XSSS displayed some unexpected behaviors. Specifically, there were times when the posterior distributions for one, some, or all of the parameters shifted away from the true values over time. However, the resulting posterior distribution of the estimated assessment year stock status was approximately centered about the true value (e.g. rockfish $\mathrm{T}_{\mathrm{below}}$). An advantage of XDB-SRA and XSSS is that they each have only a few parameters that have biological significance for which the user may have some *a priori* general idea of acceptable priors. However, the results of this analysis suggest that great care should be taken when attempting to draw conclusions from the individual posterior distributions from each of the methods for all parameters. All model misspecification is incorporated into the parameter estimates and interpreting them individually may be misleading. For example, often in data-rich assessments that contain index and composition data (e.g. length and age), the steepness parameter is often fixed due to there being little information in the data (Lee et al., 2012). There should be no expectation that these data-moderate methods applying only a biomass index would be able to offer more information compared to data-rich methods. A proposed diagnostic for these estimation methods is to evaluate the posterior distributions for M, h or F_{MSY}/M, and B_{MSY}/B_0 for updating (Cope et al., 2015). Movement in these parameters indicates that there is information in the index data, which is unlikely, and could be an indication of parameter misspecification.

Generally, the highest probabilities of overfishing for both life histories occurred when the simulated stock was initially below the target relative biomass level, even when all parameter priors were specified correctly resulting in probabilities > 0.50 during the early part of the projection period (except for XSSS). Each of these methods tended to estimate OFLs that were too high, and the buffer values were insufficient to prevent overfishing for at least the first assessment. The performance of the data-moderate methods (XDB-SRA and XSSS) slowly improved as the amount of index data increased and became informative about a continuing population decline. Although the probability of overfishing decreased for the catch-only methods (DCAC and DB-SRA), this was not due to updating of the OFLs during the projection period. The estimated harvest exceeded the true OFL, which under the harvest control rule is designed to rebuild the stock, but the set ABCs were less than the corresponding productivity at the stock size allowing for population growth. If a stock is below the target level, specifying this information through the parameter distributions may not result in harvest estimates that avoid overfishing.

The results presented here highlight situations where specific estimation methods may be better suited given the data and the life history, or when alternative methods could perform equally well, and how to interpret estimates. However, these results are dependent upon the modeling assumptions applied and additional work should be conducted to explore how robust these results are to alternative operating model structures. A strategic choice was made to incorporate only limited model misspecification between the operating model and estimation methods (e.g. stock recruitment relationship, selectivity) with process error through autocorrelated recruitment deviations and observation error about the biomass index in the operating model. This work is the first simulation test of these data-moderate methods and it was deemed critical to determine their general performance under ideal conditions to allow interpretability of the results. If a method performed poorly it was important to be able identify if that was a property of the estimation method or a particular misspecification that was assumed.

Future research should extend this work by adding more complexity to the operating model and should explore additional assumptions implemented by data-moderate and catch-only estimation methods. Additionally, the data-moderate index based estimation methods evaluated here are very time intensive to simulation test, which limited the number of scenarios that could be explored. Future work should be conducted to evaluate ways to improve the speed of these estimation methods and alternative methods for simulation testing.

The estimation methods examined here were selected because they have been applied for management of West Coast groundfish. However, there are many additional methods that have been developed to address data-limited and data-moderate stocks that range from simpler methods that apply life history invariants to estimate maximum sustainable yield (e.g. Beddington and Kirkwood, 2005), or mean lengths to estimate spawning potential ratios (e.g. Ault et al., 2008), or methods that can incorporate catch and recent year fishing effort estimates (e.g. Walters et al., 2006) or apply a catch curve and limited recent year composition data to estimate fishing mortality (e.g. Thorson and Cope, 2015). Evaluating a collection of data-moderate and catch-only methods under a wide range of simulation scenarios will help identify when methods perform adequately and when they may lead to precautionary or risky advice for decision makers.

Acknowledgments

We would like to thank Alec MacCall (Southwest Fisheries Science Center), E.J. Dick (Southwest Fisheries Science Center), and Jason Cope (Northwest Fisheries Science Center) for useful discussions and feedback that assisted in the formation of this work. Also, thank you to the multiple reviewers of early versions of this paper for their thoughtful comments.

2.5 Tables

Table 2.1: Life history parameters used in the operating model for the flatfish and rockfish life history.

Parameter	Equation form	Rockfish life history Values		Flatfish life history Values	
		Males	Females	Males	Females
Natural mortality (M) (yr^{-1})	M = constant	0.05		0.20	
Steepness (h)		0.511		0.875	
Compensation (c)		1.5		1.75	
Fishing rate at B_{MSY} (F_{MSY}) (yr^{-1})		0.045		0.27	
Mean length-at-age in cm ($L_{\gamma,a}$)	$L_{\gamma,a} = L_{\infty,\gamma} + (L_{1,\gamma} - L_{\infty,\gamma})e^{-k_\gamma(a-a_3)}$				
Mean asymptotic size in cm ($L_{\infty,\gamma}$)	$L_{\infty,\gamma} = L_{1,\gamma} + \frac{L_{2,\gamma}-L_{1,\gamma}}{1-e^{-k_\gamma(a_4-a_3)}}$				
Reference age (a_3, a_4) (yr)		1		2.833	
		8		17.8	
Mean length-at a_3 ($L_{1,\gamma}$) (cm)		6.6		24.6	
Growth coefficient (k_γ) (yr^{-1})		0.26	0.13	0.30	0.14
Coefficient of variation of length-at-age ($\sigma_{0,\gamma,a}$)		0.08		0.08	
Body weight ($w_{l,\gamma}$)	$w_{l,\gamma} = \Omega_1 L_l^{\Omega_2}$	$1.6e^{-6}$		$7.2e^{-6}$	$3.4e^{-6}$

Table 2.2: Scenarios (assumed prior distributions) for each assessment method for the flatfish life history. The mean values of the prior distributions by scenario are given along with the standard deviation, and assumed distribution (where LN is the for lognormal distribution) with the misspecified parameters in italics by scenario. The scenarios are denoted by abbreviation to indicate the prior assumption (T= true values, P= the productivity parameter [F_{MSY}/M or h], and δ_{50} = relative stock status in year 50) and subscript that refers to whether the stock was at or below the management target biomass level in year 50. The true values are given in bold below the corresponding scenario.

	Parameters														
	All Models						XDB-SRA DB-SRA DCAC						XSSS		
Scenarios	M	SD	Dist.	δ_{50}	SD	Dist	$\frac{F_{MSY}}{M}$	SD	Dist.	$\frac{B_{MSY}}{B_0}$	SD	Dist.	h	SD	Dist.
T_{at}	0.20	0.25	LN	0.25	0.20	Beta	1.35	0.30	LN	0.26	0.10	Beta	0.88	0.10	Beta
True Values	**0.20**			**0.25**			**1.35**			**0.26**			**0.88**		
	M	SD	Dist.	δ_{50}	SD	Dist	$\frac{F_{MSY}}{M}$	SD	Dist.	$\frac{B_{MSY}}{B_0}$	SD	Dist.	h	SD	Dist.
$\mathrm{T}_{\mathrm{below}}$	0.20	0.25	LN	0.10	0.20	Beta	1.35	0.30	LN	0.26	0.10	Beta	0.88	0.10	Beta
$\mathrm{D}_{\mathrm{below}}$	0.20	0.25	LN	*0.25*	0.20	Beta	1.35	0.30	LN	0.26	0.10	Beta	0.88	0.10	Beta
$\mathrm{DP}_{\mathrm{below}}$	0.20	0.25	LN	*0.25*	0.20	Beta	*0.80*	0.30	LN	0.26	0.10	Beta	*0.75*	0.10	Beta
True Values	**0.20**			**0.10**			**1.35**			**0.26**			**0.8**		

Table 2.3: Scenarios (assumed prior distributions) for each assessment method for the rockfish life history. The mean values of the prior distributions by scenario are given along with the standard deviation, and assumed distribution (where LN is the for lognormal distribution) with the misspecified parameters in italics by scenario. The scenarios are denoted by abbreviation to indicate the prior assumption (T= true values, P= the productivity parameter [F_{MSY}/M or h], and δ_{50}= relative stock status in year 50) and subscript that refers to whether the stock was at or below the management target biomass level in year 50. The true values are given in bold below the corresponding scenarios.

	Parameters														
	All Models						XDB-SRA DB-SRA DCAC						XSSS		
Scenarios	M	SD	Dist.	δ_{50}	SD	Dist	$\frac{F_{MSY}}{M}$	SD	Dist.	$\frac{B_{MSY}}{B_0}$	SD	Dist.	h	SD	Dist.
T_{at}	0.05	0.25	LN	0.40	0.20	Beta	0.89	0.30	LN	0.45	0.10	Beta	0.50	0.10	Beta
True Values	**0.05**			**0.40**			**0.89**			**0.45**			**0.50**		
	M	SD	Dist.	δ_{50}	SD	Dist	$\frac{F_{MSY}}{M}$	SD	Dist.	$\frac{B_{MSY}}{B_0}$	SD	Dist.	h	SD	Dist.
T_{below}	0.05	0.25	LN	0.20	0.20	Beta	0.89	0.30	LN	0.45	0.10	Beta	0.50	0.10	Beta
D_{below}	0.05	0.25	LN	*0.40*	0.20	Beta	0.89	0.30	LN	0.45	0.10	Beta	0.50	0.10	Beta
DP_{below}	0.05	0.25	LN	*0.40*	0.20	Beta	*0.60*	0.30	LN	0.45	0.10	Beta	*0.40*	0.10	Beta
True Values	**0.05**			**0.20**			**0.89**			**0.45**			**0.50**		

Table 2.4: The percent of simulated flatfish stocks that are below the relative biomass target (25%), overfished (<12.5%), or extinct at the end of the projection period (year 75).

	Below Target %				Overfished %				Extinct %			
	Estimation Method				Estimation Method				Estimation Method			
Scenario	DCAC	DB-SRA	XDB-SRA	XSSS	DCAC	DB-SRA	XDB-SRA	XSSS	DCAC	DB-SRA	XDB-SRA	XSSS
T_{at}	0	17	81	21	0	9	59	4	0	4	15	0
T_{below}	41	60	51	10	37	55	30	2	31	48	14	0
D_{below}	44	88	55	49	40	86	32	33	36	83	14	14
DP_{below}	29	57	32	49	23	51	14	18	19	46	4	2

Table 2.5: The simulated flatfish stock median percent and 95% simulation interval of catch realized by each method relative to the operating model estimated OFL over the projection period. A perfect performance for the catch-only estimation methods would be 69% (ABC = 69% OFL) and 83% (ABC = 83% OFL) for the data-moderate estimation methods

	DCAC		DB-SRA		XDB-SRA		XSSS	
Scenario	Median %	95% SI	Median %	95% SI	Median %	95% SI	Median %	95% SI
T_{at}	23.3	(17.7-45.3)	46.6	(36.0-61.8)	167.7	(48.2-329.2)	52.3	(28.1-151.8)
T_{below}	70.6	(33.6-186.2)	133.1	(75.1-194.0)	107.1	(39.4-1347.2)	60.1	(38.0-143.1)
D_{below}	80.2	(37.4-194.9)	391.6	(248.3-558.4)	122.0	(40.3-1024.1)	100.6	(46.7-752.4)
DP_{below}	53.9	(26.2-167.7)	115.8	(58.5-204.3)	79.2	(33.1-644.4)	108.5	(50.0-353.2)

Table 2.6: The percent of simulated rockfish stocks that are below the relative biomass target (40%), overfished (<25%), or extinct at the end of the projection period.

	Below Target %				Overfished %				Extinct %			
	Estimation Method				Estimation Method				Estimation Method			
Scenario	DCAC	DB-SRA	XDB-SRA	XSS	DCAC	DB-SRA	XDB-SRA	XSS	DCAC	DB-SRA	XDB--SRA	XSS
T_{at}	0	31	93	78	0	1	11	3	0	0	1	0
T_{below}	100	100	100	100	56	58	54	28	0	0	1	0
D_{below}	100	100	100	100	78	100	76	67	0	25	0	0
DP_{below}	100	100	100	100	50	93	41	83	0	1	0	2

Table 2.7: The simulated rockfish stock median percent and 95% simulation interval of catch realized by each method relative to the operating model estimated OFL over the projection period. A perfect performance for the catch-only estimation methods would be 69% (ABC = 69% OFL) and 83% (ABC = 83% OFL) for the data-moderate estimation methods.

	DCAC		DB-SRA		XDB-SRA		XSSS	
Scenario	Median %	95% SI	Median %	95% SI	Median %	95% SI	Median %	95% SI
T_{at}	30.7	(26.4-37.8)	57.5	(52.1-64.8)	97.7	(60.0-165.7)	80.1	(52.4-147.1)
T_{below}	97.5	(44.2-129.0)	99.3	(86.7-117.0)	89.9	(53.9-149.6)	78.1	(46.9-153.6)
D_{below}	131.6	(112.2-154.8)	517.3	(241.4-2325.3)	116.6	(65.7-197.5)	99.4	(60.3-208.6)
DP_{below}	87.4	(66.7-120.1)	201.1	(174.3-274.5)	85.7	(47.3-142.7)	124.1	(76.9-251.8)

2.6 Figures

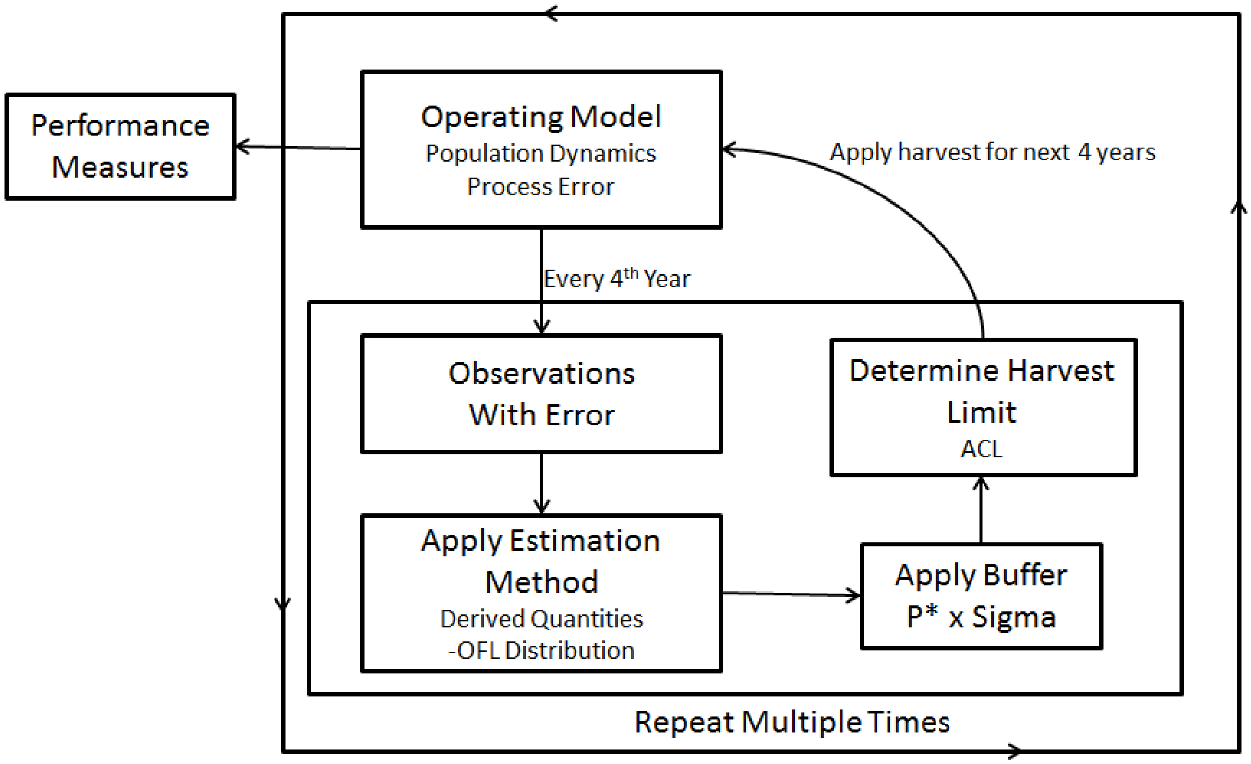

Figure 2.1: The process and order of operations for the simulation.

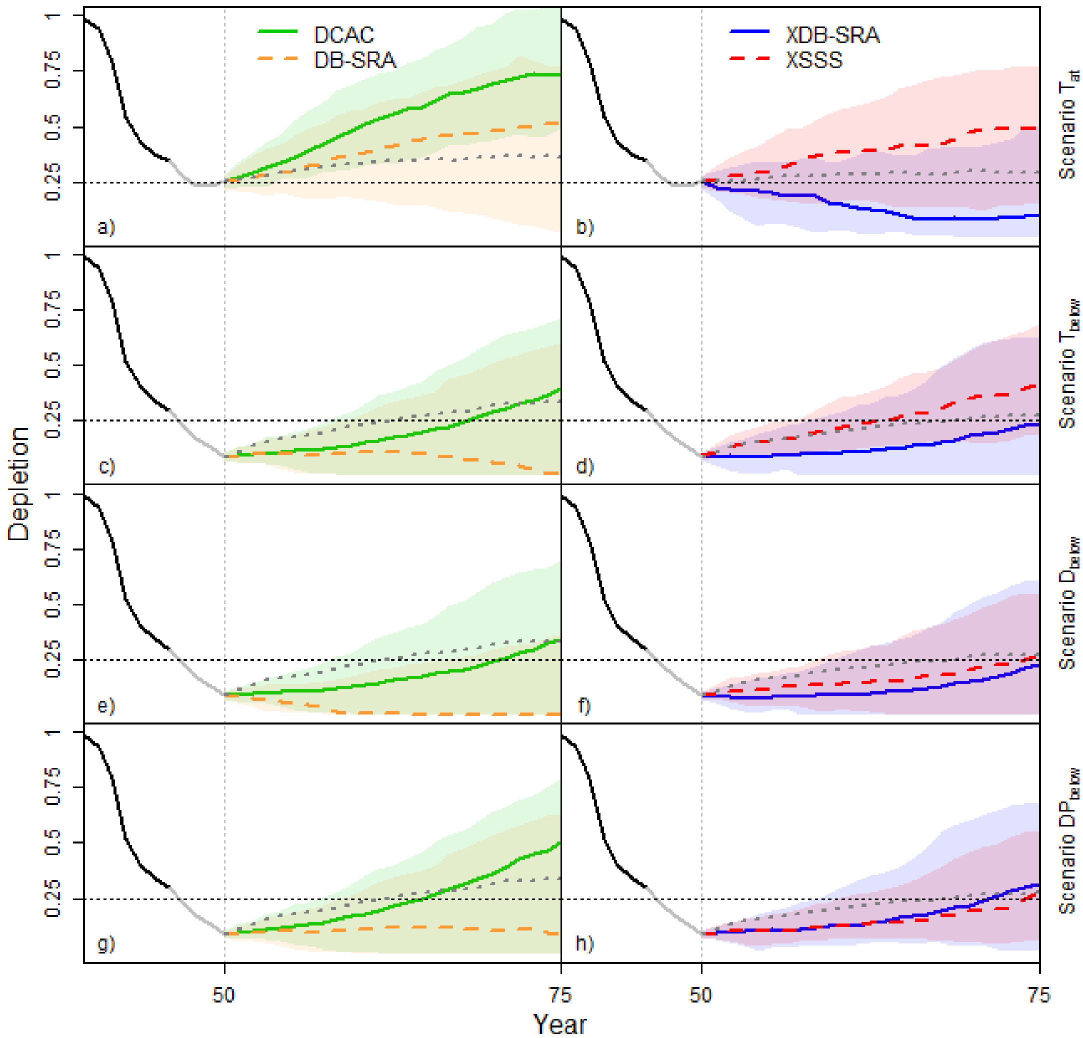

Figure 2.2: Time-trajectories of relative biomass for the flatfish population with 90% simulation intervals when the OFLs and ABCs are provided by: DCAC (green line and interval) and DB-SRA (orange dashed line and interval), shown in the left panels, and XDB-SRA (blue line and interval) and XSSS (red dashed line and interval), shown in the right panels for each of the four scenarios. The median relative biomass over the simulations if the stock was managed with perfect information from the operating model with the OFL adjusted by the appropriate buffer being removed without error is shown in each panel (dotted grey line). The years for which a biomass index was available for the first assessment in year 50 is shown by the light grey line of the time-trajectories of the simulated stocks prior to start of the projection period. The vertical dotted line indicates the start of the projection period and the horizontal dotted line indicates the target value for flatfish stocks set by the PFMC.

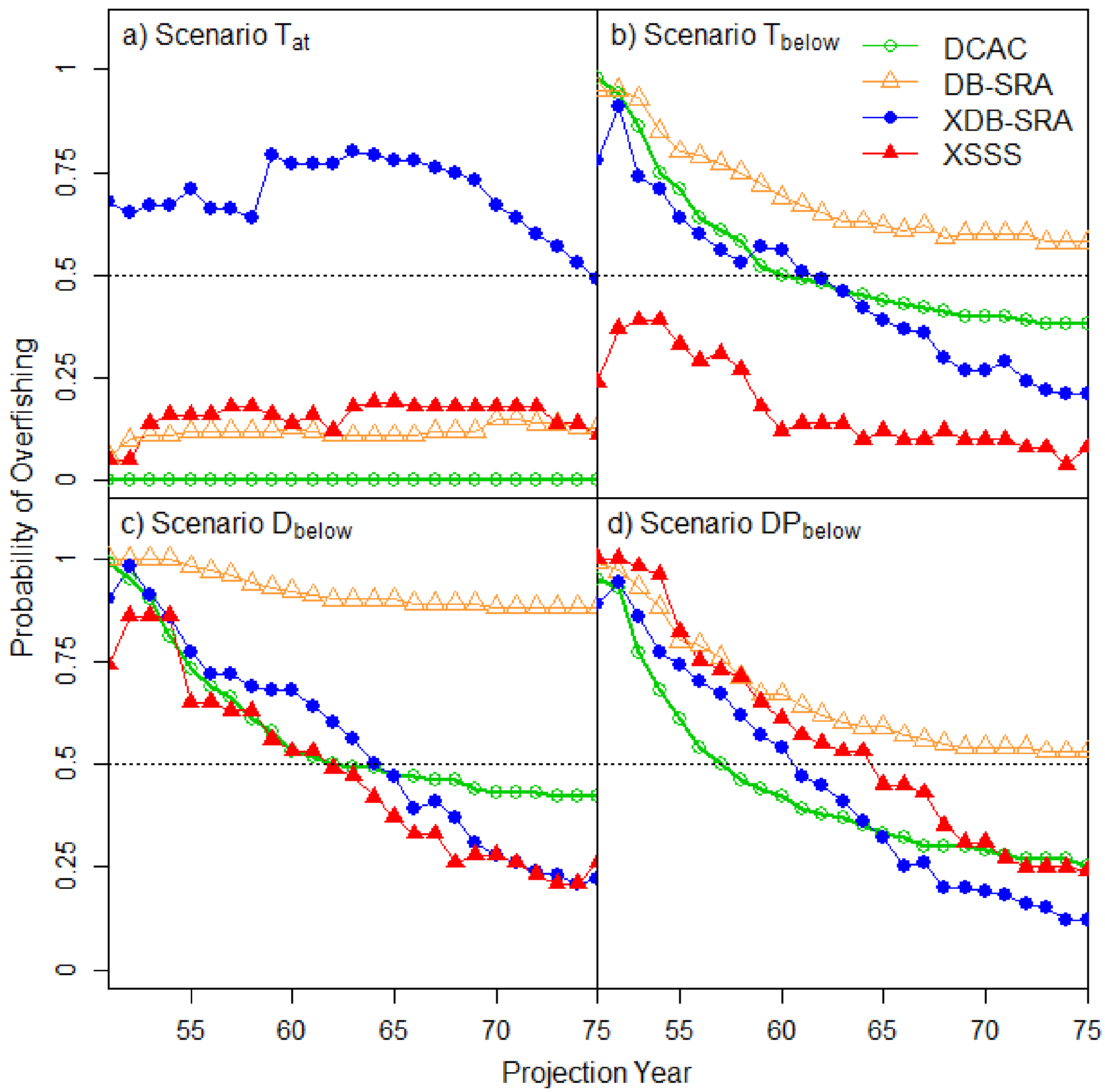

Figure 2.3: The probability of overfishing (ABC > true OFL) for the flatfish life history during the projection period for each assessment method (DCAC. DB-SRA, XDB-SRA, and XSSS) and scenario.

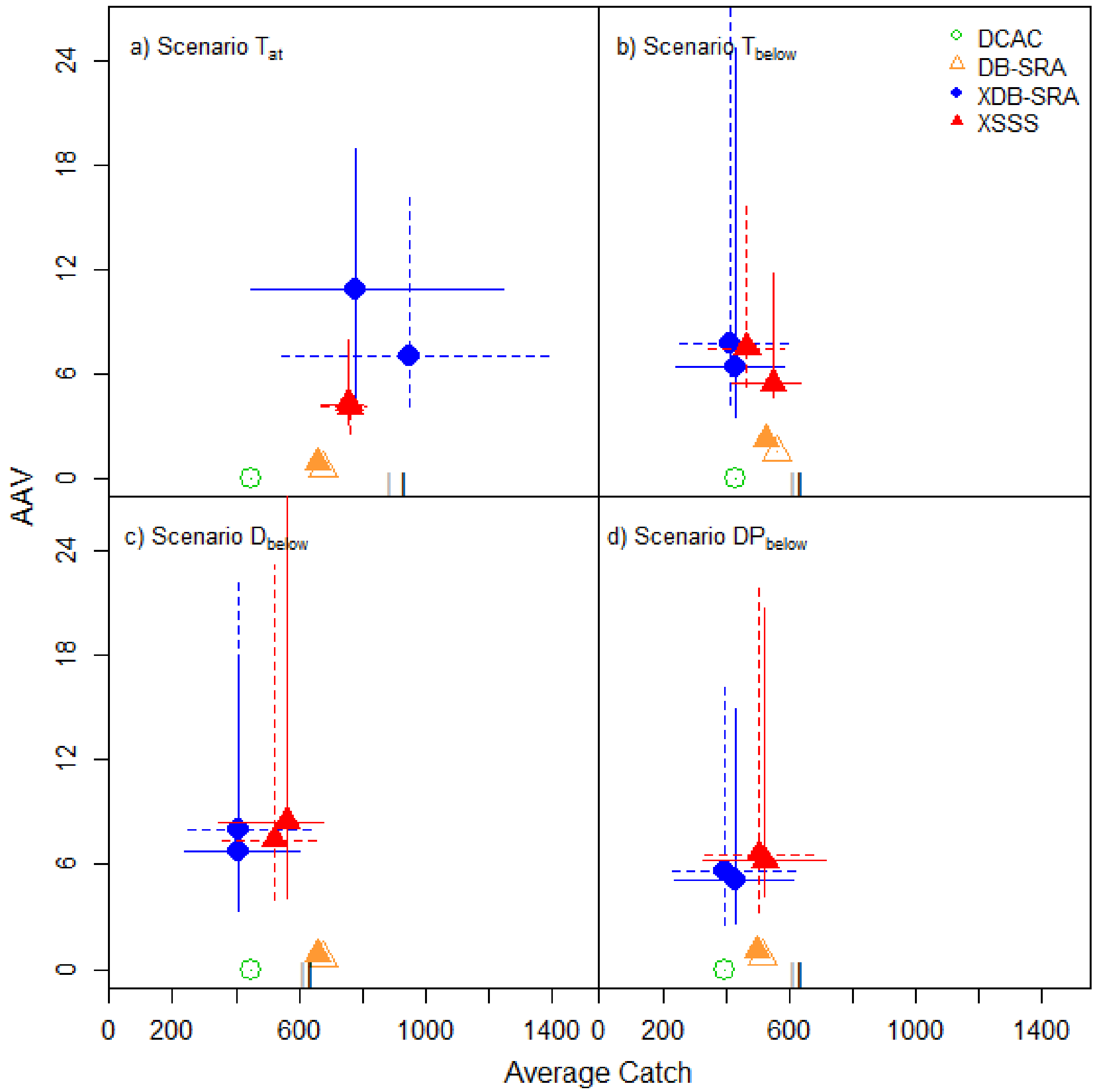

Figure 2.4: The average annual variation in catch vs the average catch (with 90% simulation intervals) after five years (solid line), and 25 years (dotted line) by assessment method; DCAC, DB-SRA (5 yrs: open diamond, 25 yrs: filled diamond), XDB-SRA, and XSSS, and each scenario. There is no inter-simulation variation in the *AAV* and average catch for DCAC and DB-SRA since each method estimates a time-series of harvest levels independent of data. The average catch over 25 years when the true maximum catch from the operating model was removed without error from the stock is shown on the x-axis for each buffer level (data-moderate category two [XDB-SRA and XSSS]: black and data-limited category three [DCAC and DB-SRA]: grey).

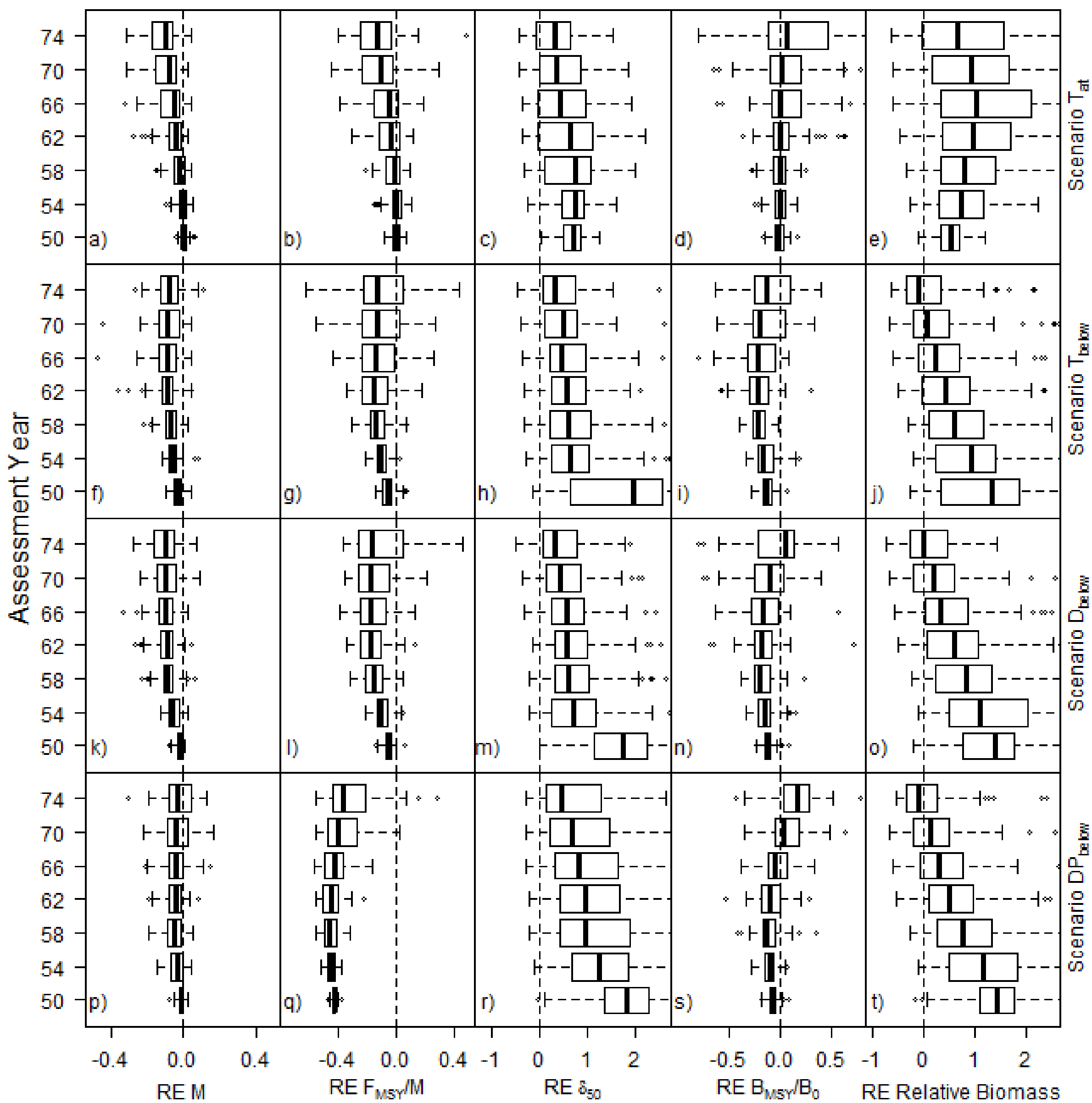

Figure 2.5: The distribution of the relative error between the true values and the means of the posterior distributions for the four leading parameters (M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) and relative error between the median of the posterior distribution for the assessment year stock status for XDB-SRA for the flatfish life history. The thick black line indicates the median of the posterior distribution means (M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) or medians (assessment year stock status). Results are shown for when each assessment is undertaken for each scenario.

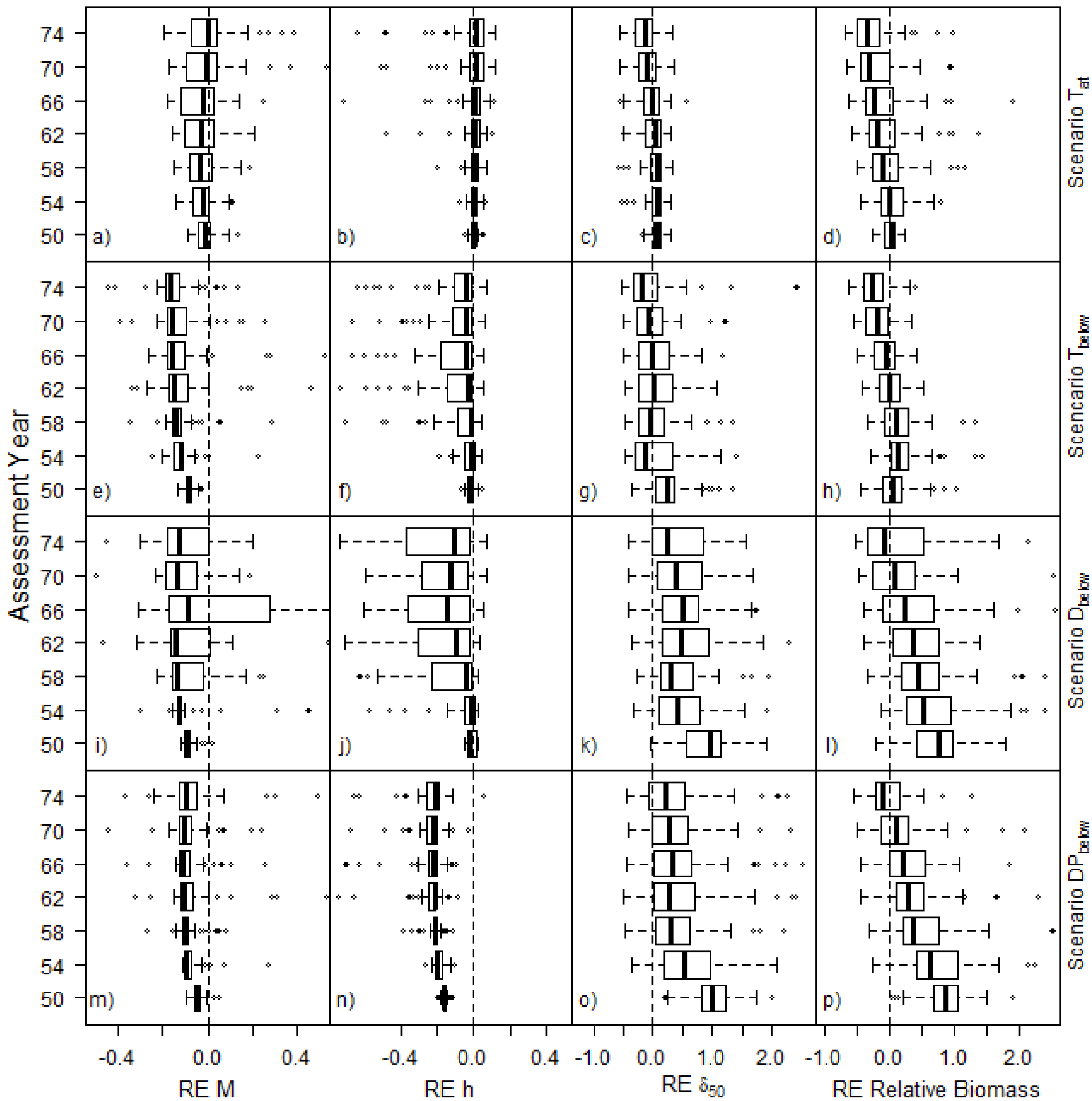

Figure 2.6: Relative error between the means of the posterior distributions for the three leading parameters (M, h, δ_{50}) and the relative error between the median of the posterior distributions for assessment year stock status and their true values for XSSS for the flatfish life history, expressed as relative to the true values. The thick black line indicates the median of the posterior distribution means (M, h, δ_{50}) or medians (assessment year stock status). Results are shown for when each assessment is undertaken for each scenario.

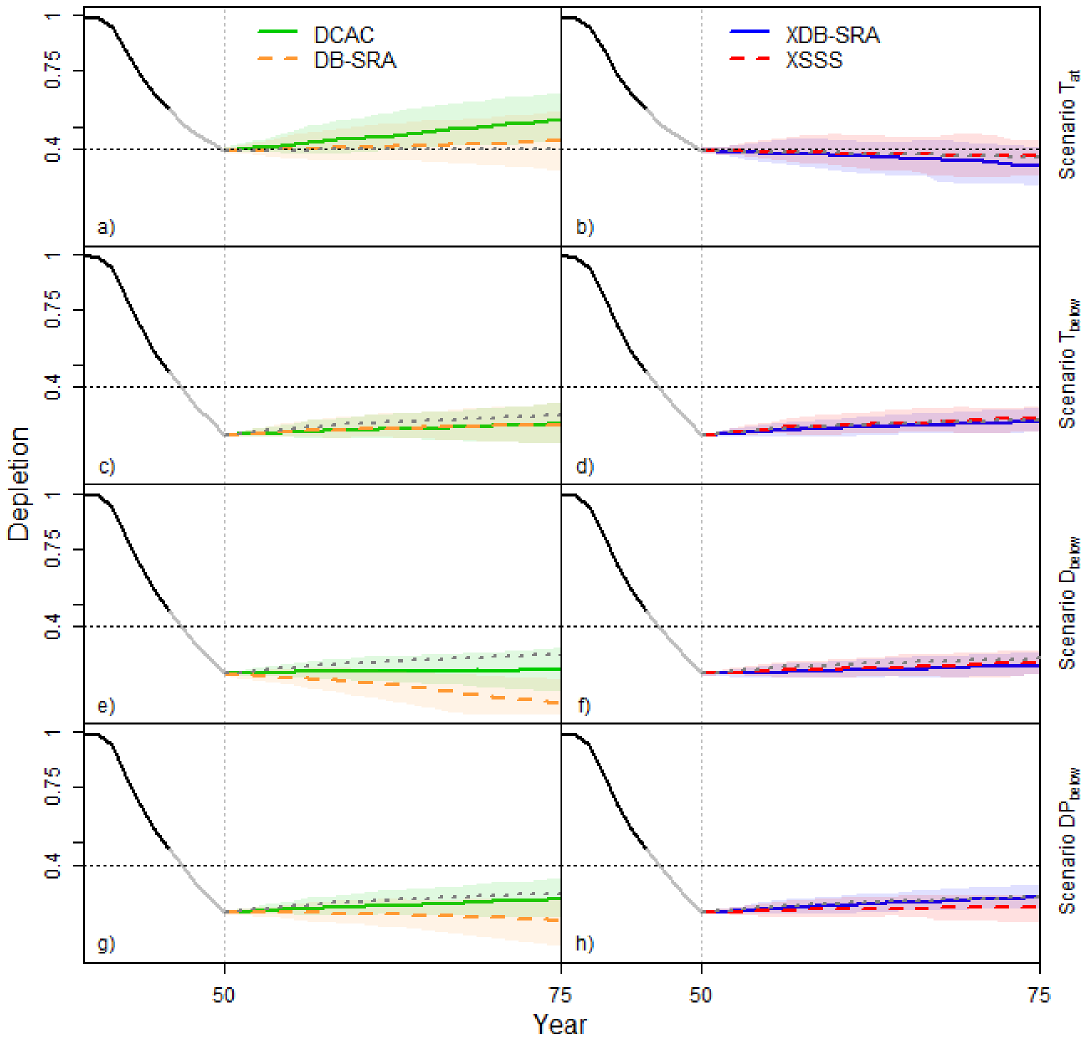

Figure 2.7: Time-trajectories of relative biomass for the rockfish population with 90% simulation intervals when the OFLs are provided by: DCAC (green line and interval) and DB-SRA (orange dashed line and interval), shown in the left panels, and XDB-SRA (blue line and interval) and XSSS (red dashed line and interval), shown in the right panels for each of the four scenarios. The median relative biomass over the simulations if the stock was managed with perfect information from the operating model with the OFL adjusted by the appropriate buffer being removed without error is shown in each panel (dotted grey line). The years for which a biomass index was available for the first assessment in year 50 is shown by the light grey line of the time-trajectories of the simulated stocks prior to start of the projection period. The vertical dotted line indicates the start of the projection period and the horizontal dotted line indicates the target value for rockfish stocks set by the PFMC.

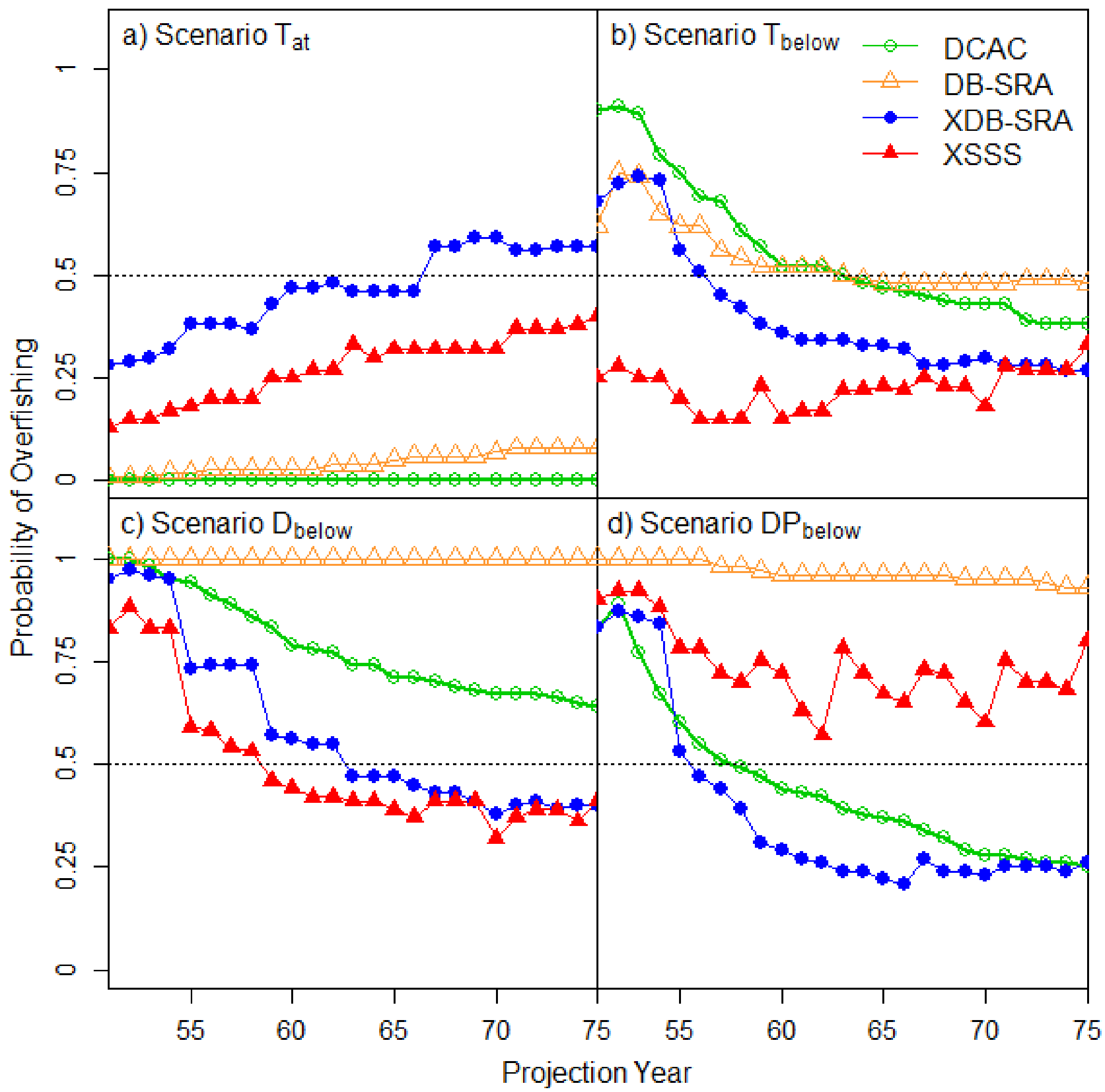

Figure 2.8: The probability of overfishing (ABC > true OFL) for the rockfish life history during the projection period for each assessment method (DCAC. DB-SRA, XDB-SRA, and XSSS) and scenario.

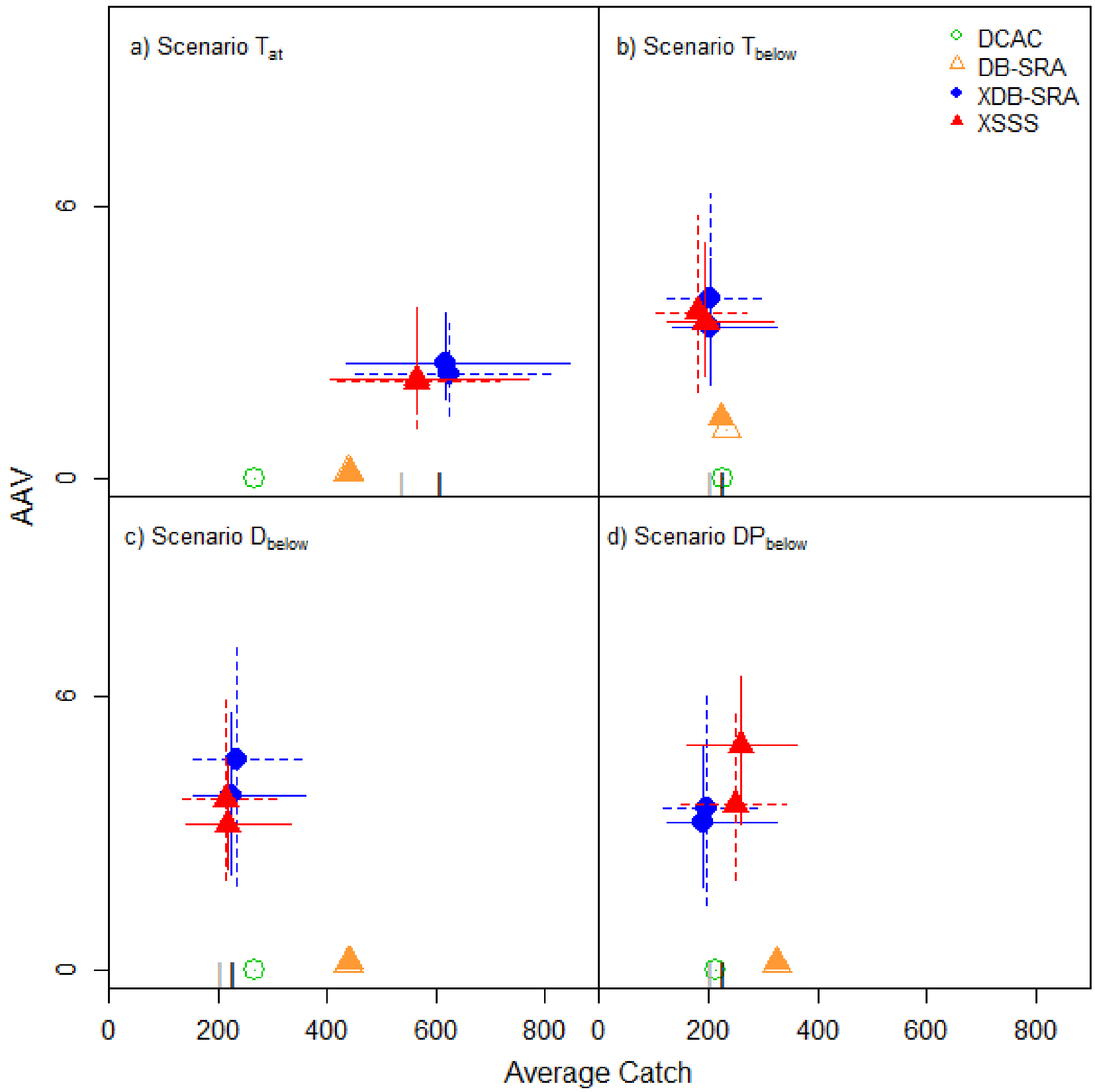

Figure 2.9: The average annual variation in catch vs the average catch (with 90% simulation intervals) after five years (solid line), and 25 years (dotted line) by assessment method; DCAC, DB-SRA (5 yrs: open diamond, 25 yrs: filled diamond), XDB-SRA, and XSSS, and each scenario. There is no inter-simulation variation in the *AAV* and average catch for DCAC and DB-SRA since each method estimates a time-series of harvest levels independent of data. The average catch over 25 years when the true maximum catch from the operating model was removed without error from the stock is shown on the x-axis for each buffer level (data-moderate category two [XDB-SRA and XSSS]: black and data-limited category three [DCAC and DB-SRA]: grey).

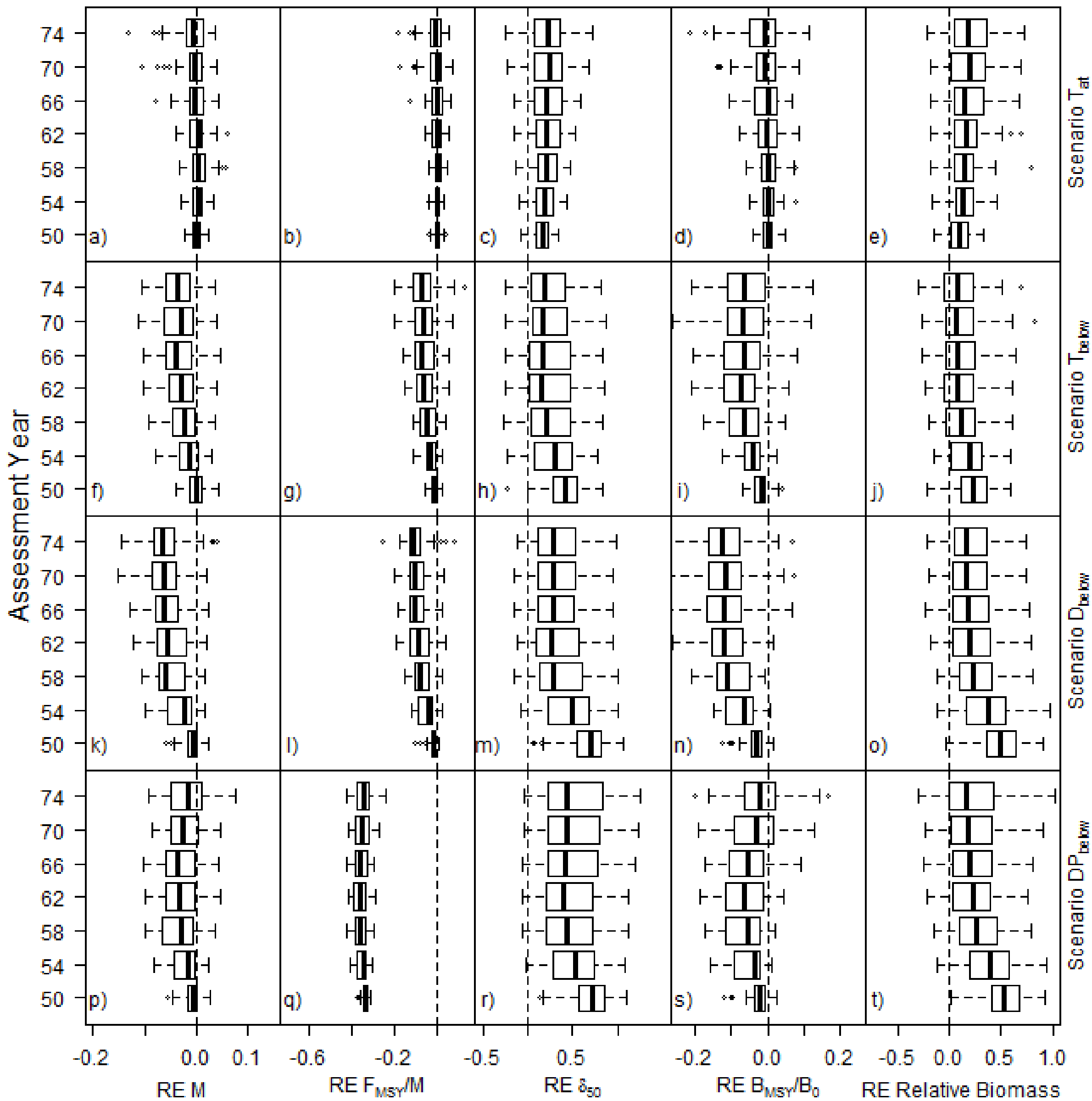

Figure 2.10: The distribution of the relative error between the true values and the means of the posterior distributions for the four leading parameters (M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) and relative error between the median of the posterior distribution for the assessment year stock status for XDB-SRA for the flatfish life history. The thick black line indicates the median of the posterior distribution means (M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) or medians (assessment year stock status). Results are shown for when each assessment is undertaken for each scenario.

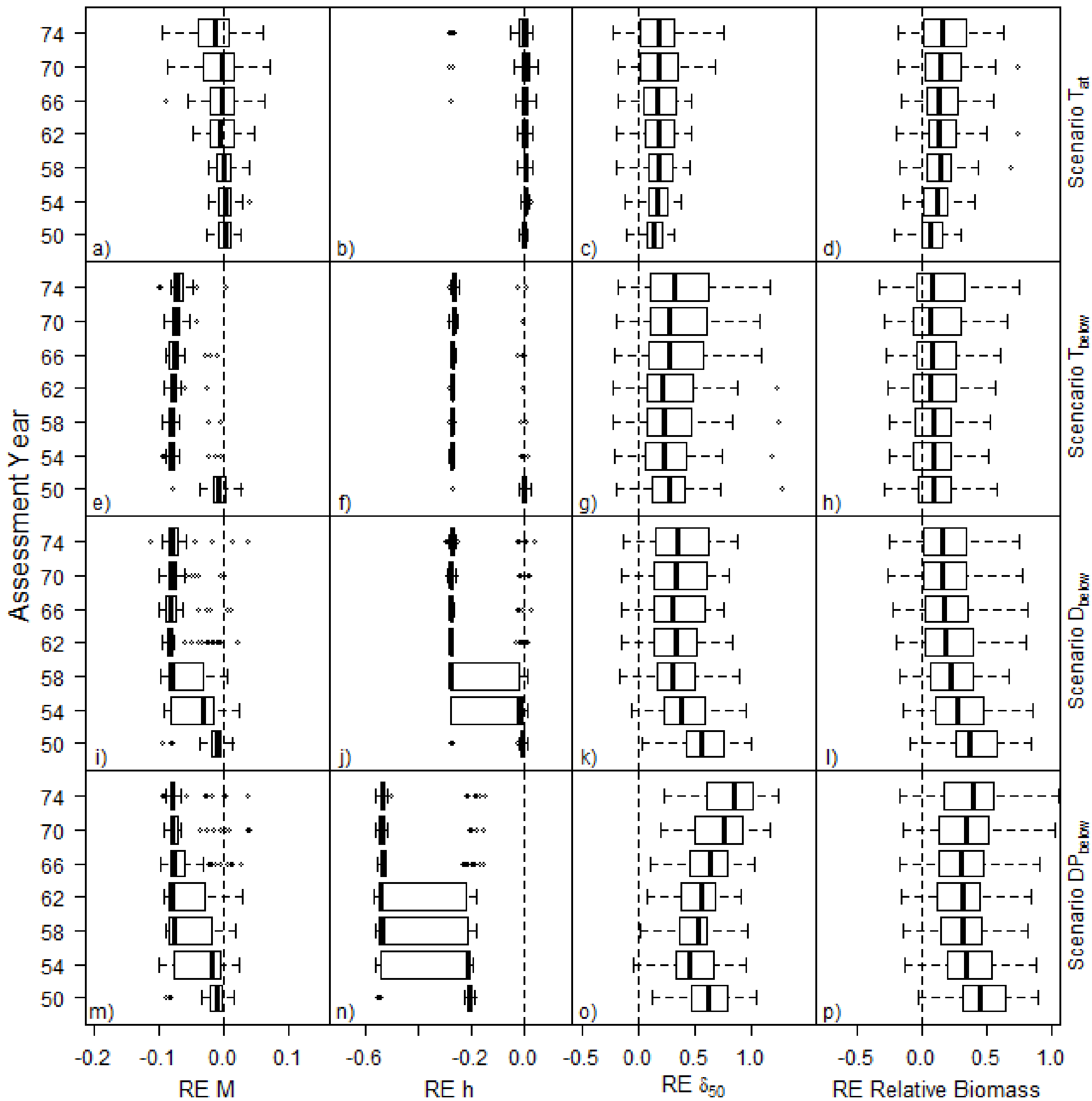

Figure 2.11: Relative error between the means of the posterior distributions for the three leading parameters (M, h, δ_{50}) and the relative error between the median of the posterior distributions for assessment year stock status and their true values for XSSS for the flatfish life history, expressed as relative to the true values. The thick black line indicates the median of the posterior distribution means (M, h, δ_{50}) or medians (assessment year stock status). Results are shown for when each assessment is undertaken for each scenario.

Chapter 3

THE IMPACT OF ALTERNATIVE REBUILDING STRATEGIES TO REBUILD OVERFISHED STOCKS

Abstract

Ending overfishing and rebuilding fish stocks to levels that provide for optimum sustainable yield is a concern for fisheries management worldwide. In the US, fisheries managers are legally mandated to end overfishing and to implement rebuilding plans for fish stocks that fall below minimum stock size thresholds. Rebuilding plans should lead to recovery to target stock sizes within ten years, except in situations where the life history of the stock or environmental conditions dictate otherwise. Federally managed groundfish species along the US west coast have diverse life histories where some are able to rebuild quickly from overfished status, while others, specifically rockfish (*Sebastes* spp.), may require decades for rebuilding. A management strategy evaluation which assumed limited estimation error was conducted to evaluate the performance of alternative strategies for rebuilding overfished stocks for these alternative US west coast life histories. Generally, the results highlight the trade-off between the reduction of catches during rebuilding vs. the length of rebuilding. The most precautionary rebuilding plans requiring the greatest harvest reduction resulted in higher average catches over the entire projection period compared to strategies that required a longer rebuilding period with less of a reduction in rebuilding catch. Attempting to maintain a 50% probability of rebuilding was the poorest performing rebuilding strategy for all life histories, resulting in a high number of changes to the rebuilding plan, increased frequency of failing to meet rebuilding targets, and higher variation in catch. The rebuilding plans that implemented a higher initial rebuilding probability ($\leq$ 60%) for determining rebuilding fishing mortality and targets generally resulted in fewer changes to the rebuilding plans and

rebuilt by the target rebuilding year, particularly for stocks with the longer rebuilding plans (e.g. rockfishes).

3.1 Introduction

Eliminating overfishing and achieving sustainable fisheries has been, and continues to be, a challenge worldwide. Up to 63% of global stocks have been estimated to be below biomass target reference points for maximum sustainable yield (Worm et al., 2009). In the US, 23% of federally-managed stocks were estimated to be below their biomass limit reference points (and thus meeting the overfished definition) in 2012, with 85 stocks being declared overfished during 1997–2011 (National Reserach Council, 2013). The US has made commitments to end overfishing and to rebuild overfished stocks. Reducing fishing mortality is the first critical step to end overfishing. Beyond reducing harvest, successful rebuilding of overfished stocks is greatly facilitated by implementation of rebuilding plans that have clearly defined objectives and strategies and that have stakeholder and management support (Mora et al., 2009). Additionally, successful rebuilding plans should consider a precautionary approach in the face of management and scientific uncertainty (Cadrin and Pastoors, 2008).

For US federally-managed fish stocks, rebuilding plans are required when fish stocks are declared overfished, i.e. when they are estimated to be below their minimum stock size threshold (Sustainable Fisheries Act, 1996). Such plans have been shown to successfully rebuild overfished stocks to target levels in many cases when fishing mortality was reduced to a rate that would allow population growth in the absence of unexpected changes in productivity (Milazzo, 2012; National Reserach Council, 2013). The development of US rebuilding plans involves three key factors: (i) the Magnuson–Stevens Act, (ii) the National Standard Guidelines, and (iii) court cases. The National Marine Fisheries Service (NMFS) has provided guidelines for the features of a rebuilding plan (Sustainable Fisheries Act, 1996; Federal Register, 1998). US rebuilding plans are required to define the following components (Table 3.1): (i) the target year for rebuilding (T_{TARGET}), (ii) the minimum amount of time that would allow rebuilding in the absence of fishing with at least a 50% probability (T_{MIN}), and

(iii) the maximum amount of time targeted for rebuilding the stock (T_{MAX}). Guidelines from NMFS dictate that a stock must be rebuilt within 10 years (i.e. $T_{MAX} = 10$ years, $T_{TARGET} \leq 10$ years) if T_{MIN} is less than 10 years, but the upper limit for rebuilding (T_{MAX}) may be set as high as T_{MIN} plus one mean generation time if the stock is unable to rebuild within 10 years ($T_{MIN} > 10$ years). The target year for rebuilding (T_{TARGET}) must fall between T_{MIN} and T_{MAX}.

From 1999 to the present, ten US west coast groundfish stocks have been declared overfished and have required rebuilding plans (some of which have since been declared rebuilt; Pacific Fishery Management Council, 2014c). During this period, the PFMC, which makes management recommendations for federally-managed West Coast fish stocks, has been subject to lawsuits filed directly in opposition to rebuilding plans for overfished groundfish stocks (e.g., Natural Resources Defense Council (NRDC) v. NMFS, 421 F.3d 872 (9th Cir. 2005)) which have had lasting implications on the development of management plans for West Coast groundfish stocks. The stocks that have been declared overfished are highly diverse, ranging from stocks deemed able to rebuild within 10 years (e.g., Pacific hake (*Merluccius productus*) and petrale sole (*Eopsetta jordani*)) to stocks that may require very long rebuilding periods (e.g., 70+ years for yelloweye rockfish (*Sebastes ruberrimus*) (Pacific Fishery Management Council, 2014c). The PFMC faces the challenge of implementing rebuilding plans that will successfully rebuild stocks across this range of circumstances while meeting the mandate set by the Magnuson–Steven Act requiring a stock to be rebuilt "in as short as possible, taking into account the needs of the fishing communities".

US west coast federal fisheries management controls fishing mortality rates by setting harvest rates based on spawning potential ratios (SPRs). SPR is a measure of the impact of fishing mortality on the projected average contribution of each recruit to the spawning output (thus, the smaller the SPR value, the higher fishery exploitation). Current practice for establishing rebuilding plans on the US west coast includes projections that apply a range of fishing mortality rates, expressed in terms of SPR, to determine the minimum year for rebuilding in the absence of fishing (T_{MIN}) which, combined with the mean generation time, determines the maximum year for rebuilding (T_{MAX}) to occur (Table 3.1). The regional management council then selects a target year for rebuilding (T_{TARGET}: must fall between T_{MIN} and T_{MAX}) and the associated SPR that reflects a desired level of probability to rebuild the stock (must be ≥50%). The results from the projections across this range of SPR rates and a range of realized stock dynamics (with process error modeled using recruitment deviations, although other sources of uncertainty are often considered, such as assessment uncertainty) represent a "rebuilding analysis".

Stocks that are managed under a rebuilding plan are monitored during rebuilding, and subsequent rebuilding analyses are conducted to ensure that the stock remains on course to rebuild by the target year according a prespecified probability (P_{TARGET}). Adjustments are made to the SPR as needed to meet rebuilding targets. Additionally, changes in the understanding of population scales during rebuilding could require adjustments to the rebuilding SPR and rebuilding timelines. The rebuilding analysis provides the scientific guidance for determining the rebuilding targets and the harvest rates for the rebuilding plan.

During rebuilding, managers generally prefer minimal revisions to the rebuilding plan to minimize the impact of harvest reductions to the stakeholders, while still meeting rebuilding targets (and for ease of application). Continuity during rebuilding also provides a measure of predictability for fishery stakeholders and allows them to plan accordingly. Rebuilding strategies that are overly sensitive to assessment noise can result in needless changes to the rebuilding plan, increasing the variability in catches during rebuilding.

Punt and Ralston (2007) conducted a management strategy evaluation (MSE) for rockfish stocks that evaluated the performance of several alternative rebuilding strategies, the method for assessing rebuilding progress, and the guidelines for adjusting rebuilding plans based on changes in perceived stock status. This paper provides an updated MSE for rebuilding US west coast groundfish stocks that has been developed iteratively based on discussions and feedback received from stakeholders, groundfish management advisory bodies, and the PFMC. Specifically, this paper evaluates the performance of six rebuilding strategies across various West Coast life histories that apply alternative approaches to set initial rebuilding harvest rates and when to update harvest rates during rebuilding. A variety of sensitivity analyses are also undertaken that explore the sensitivity to model misspecification, the frequency of assessment, or alternative thresholds for updating harvest rates during rebuilding.

3.2 Materials and Methods

3.2.1 General approach

The majority of life history strategies of fishes that are federally managed along the US west coast fall into the categories of either periodic or intermediate strategy (King and McFarlane, 2003). Periodic strategies are defined as slow-growing, long-lived demersal species with low variability in recruitment, and intermediate strategies, as defined by King and McFarlane (2003), have mid-range longevity (10–20 years) that can have dramatic changes in biomass. Two intermediate and two periodic life history strategies were simulated: (i) flatfish with a moderately high natural mortality rate and a high recruitment compensation rate (e.g. petrale sole (*Eopsetta jordani*) and Dover sole (*Microstomus pacificus*)), (ii) roundfish with an intermediate natural mortality and recruitment compensation rate (e.g. Pacific hake (*Merluccius productus*) and lingcod (*Ophiodon elongatus*)), (iii) medium-lived rockfish with a moderately low natural mortality rate and moderate recruitment compensation rate (e.g. greenstriped rockfish (*Sebastes elongates*) and widow rockfish (*Sebastes entomelas*)), and (iv) long-lived rockfish with a low natural mortality rate and a low recruitment compensation rate (e.g. canary rockfish (*Sebastes pinniger*) and yelloweye rockfish (*Sebastes rubberimus*)) (Table 3.2). For ease of presentation, the intermediate and periodic life history strategies will be referred to as either flatfish, roundfish, medium-, or long-lived rockfish life histories. Additionally, these life histories generally correspond to the categorization of stocks as applied by federal US west coast management.

The simulation study involves three separate submodels: (i) an operating model which simulates the population, (ii) an estimation model that conducts assessments and rebuilding analyses, and (iii) a management decision model that determines the management actions following alternative strategies. The simulated population was age-structured, where an annual index of abundance was observed with error, and age composition data were collected for selected years. These data were used by the stock estimation method to estimate population size and project the catch. When a stock was estimated to be below the minimum stock size threshold, as defined given its life history for the first time (i.e. the stock was not currently under a rebuilding plan), the assessment estimated catch was modified based on a rebuilding plan that calculated a SPR that would result in a given estimated probability of recovery at a specific future point in time. The rebuilding strategy was applied, and the stock assessment was updated iteratively for a specified number of years based on life history that generally allowed for recovery to target biomass levels under a variety of conditions (flatfish and roundfish, 50 years; medium-lived rockfish, 75 years; long-lived rockfish, 125 years). Results for alternative rebuilding strategies and sensitivities for each life history were based on 100 simulated stocks.

3.2.2 Operating model

The numbers-at-age at the start of the year are computed as:

$$N_{t+1,\gamma,a} = \begin{cases} 0.5R_t & \text{if a = 0} \\ N_{t,\gamma,a-1}e^{-(M_\gamma+S_{\gamma,a-1}F_t)} & \text{if } 1 \leq \text{a} < \text{A-1} \\ N_{t,\gamma,A-1}e^{-(M_\gamma+S_{\gamma,A-1}F_t)} + N_{t,\gamma,A}e^{-(M_\gamma+S_{\gamma,A}F_t)} & \text{if a = A} \end{cases} \tag{3.1}$$

where $N_{t+1,\gamma,a}$ is the number of fish of sex γ and age a at the start of year t, is the number of age-0 animals at the start of year t, $S_{\gamma,a}$ is the selectivity by sex and age, A is the plus group, F_t is the instantaneous fishing mortality rate during year t, and M_γ is the instantaneous rate of natural mortality for sex γ.

The number of age-0 fish is related to spawning biomass according to the Beverton and

Holt (1957) stock–recruitment relationship:

$$R_t = \frac{4hR_0SB_t}{SB_0(1-h)+SB_t(5h-1)}e^{-0.5\sigma_R^2+\epsilon_t^R} \quad \epsilon_t^R \sim N(0;\sigma_R^2) \tag{3.2}$$

where SB_0 is the unfished spawning biomass, SB_t is the spawning biomass at the start of the spawning season in year t, R_0 is the unfished recruitment, σ_R is the standard deviation of recruitment in log space, and h is the recruitment compensation (also known as steepness).

A non-equilibrium starting condition was created by applying equations (3.1) and (3.2) for the number of years equal to the maximum age for each life history, with variation in recruitment and no fishing. Following this, an initial fishery was simulated (along with the population) over 50 years, with the catch of fish of sex γ and age a during year t in numbers determined by:

$$C_{t,\gamma,a} = \frac{S_{\gamma,a}F_t}{M_\gamma + S_{\gamma,a}F_t}N_{t,\gamma,a}\left(1-e^{-M_\gamma - S_{\gamma,a}F_t}\right) \tag{3.3}$$

This simulated historical fishing mortality increased linearly over 50 years such that, in all cases, the populations were in an overfished state (flatfish $0.05SB_0$, roundfish and rockfish $0.10SB_0$; Table 3.2) at the time of the first assessment in year 50, based on the PFMC minimum biomass threshold levels for each life history type (flatfish $0.125SB_0$, roundfish and rockfish $0.25SB_0$). The fishery and the survey both assumed an age-based logistic selectivity (Table 3.2).

An annual survey index of abundance ($CV = 0.20$) and age composition data (n = 100) from the survey and the fishery were available for 20 years prior to the first assessment, and catches were known without error for all years. Index and age composition data were generated annually following the first assessment. The start and frequency of the survey were selected to mimic the data available for West Coast groundfish stocks.

The observation model was used to generate an index of abundance for each year t:

$$I_t = Q\widetilde{B}_t e^{-0.5\sigma_s^2+\epsilon_t^s} \quad \epsilon_t^s \sim N(0;\sigma_s^2) \tag{3.4}$$

where Q is the catchability coefficient for the survey (arbitrarily set equal to 1, since the scale does not matter here, given how this index is included in the assessment), and σ_s is the standard deviation of survey catchability in log space (see Table 3.2). The expected biomass index is given by:

$$\widetilde{B}_t = \sum_{\gamma} \sum_{a=1}^{A} w_{\gamma,a} S_{s,\gamma,a} N_{t,\gamma,a} e^{-0.5(M_\gamma + S_{f,\gamma,a} F_t)} \tag{3.5}$$

where $w_{\gamma,a}$ is the average weight by sex at age, $S_{s,\gamma,a}$ is the selectivity for the survey by sex and age, and $S_{f,\gamma,a}$ is the selectivity for the fishery by sex and age. The observed age composition data for the fishery and survey catch were assumed to be multinomially distributed.

3.2.3 Estimation method and rebuilding analysis

The simulated stocks were assessed using Stock Synthesis (Methot and Wetzel, 2013), an integrated statistical catch-at-age model. Growth, natural mortality, and the steepness of the Beverton-Holt stock–recruitment relationship were assumed to be known without error. The unfished recruitment (R_0), annual recruitment deviations, and the selectivity parameters for the survey and the fishery were estimated. The ratio of the current spawning biomass to the unfished spawning biomass (relative stock status) was estimated and, based on the estimated stock status, one of three actions was performed:

1. If the relative stock status was estimated to be below the minimum stock size threshold, as defined by the PFMC by life history type (flatfish $0.125SB_0$; rockfish and roundfish $0.25SB_0$) for the first time, the stock was declared overfished and a rebuilding analysis was performed which defined the initial rebuilding plan for the stock, setting a rebuilding harvest level (SPR) associated with the predefined probability of rebuilding by a maximum year ($\mathrm{P_{INIT}}$).

2. If the stock was already under a rebuilding plan and estimated to still be below the target biomass level, a rebuilding analysis was conducted to evaluate the current probability of rebuilding by the target year. The rebuilding SPR and rebuilding targets

were adjusted, if necessary, so that rebuilding could occur within the allowable time.

3. If the stock size was found to be above the target biomass level, the stock was declared rebuilt, and catches were estimated using the default harvest control rule (Fig. 3.1). The PFMC harvest control rule reduces the catch linearly when stock is below the target stock size (flatfish: $0.25SB_0$, roundfish and rockfish $0.40SB_0$), to zero when the stock is at or below the management lower threshold (flatfish: $0.05SB_0$, roundfish and rockfish $0.10SB_0$, although this is never applied here).

The approach to rebuilding plans and subsequent analyses during rebuilding (i.e. updated rebuilding analyses to evaluate the probability of meeting rebuilding targets given the current harvest rate) can vary by region within the US (National Reserach Council, 2013). The process implemented for developing a rebuilding plan here was based on the current practice for the US west coast groundfish:

1. The unfished biomass, SB_0, was calculated by multiplying the spawning output-per-recruit in the absence of exploitation by the arithmetic average recruitment (R_0) for the first 10 years of the assessment period.

2. Future recruitment was generated from a Beverton-Holt stock–recruitment relationship with process error variation around that median relationship.

3. The minimum time to rebuild ($\mathrm{T_{MIN}}$) was defined as the median year in which spawning biomass exceeded the management target (flatfish: $0.25SB_0$, rockfish and roundfish: $0.40SB_0$) in the absence of fishing.

4. The maximum time to rebuild ($\mathrm{T_{MAX}}$) was defined relative to minimum time required. If a stock could rebuild in less than 10 years in the absence of fishing, the maximum time allowed for rebuilding equaled 10 years (i.e. a current requirement for US rebuilding plans). However, if the minimum time required to rebuild was greater than 10 years, the $\mathrm{T_{MAX}}$ was defined as $\mathrm{T_{MIN}}$ plus one mean generation.

5. The initial rebuilding SPR was defined as the value that would result in recovery of the stock by $\mathrm{T_{MAX}}$ equal to a prespecified initial rebuilding probability ($\mathrm{P_{INIT}} \geq 0.50$).

6. The target year to rebuild ($\mathrm{T_{TARGET}}$) was set equal to the first year that the stock was projected to recover to the management target with a $\geq 50\%$ probability based on the specified rebuilding SPR.

The initial rebuilding analysis determined the parameters for the rebuilding plan ($\mathrm{T_{MIN}}$, $\mathrm{T_{MAX}}$, $\mathrm{T_{TARGET}}$, and SPR) (Table 3.1). The ensuing year's catches were determined by the rebuilding SPR. Subsequent rebuilding analyses evaluated four questions (Fig. 3.1): (i) will the stock rebuild by the target year with a probability greater than a prespecified minimum probability ($\mathrm{P_{TARGET}}$) by applying the current rebuilding plan SPR, (ii) if no, is there an SPR that would result in rebuilding by the target year, (iii) if no, is there an SPR for which the stock would be projected to rebuild if the target year was set to the maximum rebuilding year ($\mathrm{T_{MAX}}$), and (iv) if there is an SPR that met one of the above conditions, would the resulting catch be $> 50\%$ of the previous year's catch? When none of the first three criteria could be met, or when the fourth criterion was not met, the rebuilding plan was determined to be a failure and a new rebuilding plan was implemented that updated the rebuilding parameters. The SPR set by the new rebuilding plan was constrained so that it did not result in a lower (i.e. more aggressive) SPR compared to the SPR in the failed rebuilding plan.

Stocks that successfully rebuilt to the target biomass level were subsequently managed based on the PFMC harvest control rules for non-overfished stocks, where catch was calculated based on the life history SPR proxy value (flatfish: $SPR_{30\%}$, roundfish: $SPR_{45\%}$, rockfish: $SPR_{50\%}$).

3.2.4 Alternative management actions: rebuilding strategies

This work evaluated the performance of alternative initial probability of recovery ($\mathrm{P_{INIT}}$) determining a rebuilding strategy and target probability ($\mathrm{P_{TARGET}}$) of recovery threshold values while rebuilding, as applied to rebuild West Coast groundfish stocks. In practice, the value for the probability of recovery by the maximum year allowed for rebuilding (T_{MAX}) is selected by the PFMC. The current guideline from the Council is that the initial rebuilding plan will select an SPR corresponding to a probability of recovery by target year with $\geq$ 50% probability (P_{INIT}; although it has often been set much higher than 50%; Pacific Fishery Management Council, 2014a). The subsequent rebuilding analyses conducted during rebuilding evaluate whether the current SPR was predicted to result in at least a 50% probability (P_{TARGET}) of rebuilding by the target year. If the probability of recovery to the target year with the current SPR falls below 50%, the current practice of the Council is to adjust the SPR to a value that corresponds to a 50% probability of recovery.

The following alternative rebuilding strategies were simulated and their performance evaluated (Table 3.3):

1. "*Status quo*" – The "*status quo*" strategy attempted to mimic as best as possible the species-specific rebuilding strategies used by the PFMC for rebuilding West Coast groundfish stocks. The SPR in the initial rebuilding plan was determined based on a rebuilding probability of 60% (i.e. $P_{INIT} = 60\%$) by T_{MAX}. The stock and fishery were simulated for four more years, assessed, and if the stock was estimated still below the biomass target, a new rebuilding analysis was performed to determine if the stock was on target to rebuild by the target year, based on the rebuilding SPR. The SPR was adjusted upwards (i.e. reducing fishing mortality) during rebuilding to maintain at least a 50% probability (P_{TARGET}) of rebuilding by the target year (T_{TARGET}). If no SPR was found that predicted rebuilding by the target year with at least a 50% probability, the target year was revised and set equal to the current value for T_{MAX}. An updated SPR was selected that would rebuild the stock by the new target year ($T_{TARGET} = T_{MAX}$) with a 50% probability. However, the rebuilding plan was declared a failure if the stock was predicted to be unable to rebuild by the T_{MAX} under any SPR. If a rebuilding plan failed, a new rebuilding plan was conducted (calculating

new values for SPR, $\mathrm{T}_{\mathrm{TARGET}}$, and $\mathrm{T}_{\mathrm{MAX}}$) and implemented in the current year (Fig. 3.1).

2. "Flexible" – The SPR in the initial rebuilding plan was determined based on rebuilding by $\mathrm{T}_{\mathrm{MAX}}$ with a $\mathrm{P}_{\mathrm{INIT}} = 60\%$. The SPR was adjusted upwards if the predicted probability ($\mathrm{P}_{\mathrm{TARGET}}$) of rebuilding by target year under the current SPR fell below 40% to a SPR that was estimated to rebuild by the target year given a 50% probability. Other specifications are as for the "*status quo*" strategy.

3. "Risk averse" – The SPR in the initial rebuilding plan was determined based on rebuilding by $\mathrm{T}_{\mathrm{MAX}}$ with a probability of 75%. The SPR was adjusted upwards if the predicted probability of rebuilding by the target year under the current SPR fell below 60% to a SPR that was estimated would rebuild with a 60% probability by $\mathrm{T}_{\mathrm{MAX}}$.

4. "Risk neutral" – The SPR in the initial rebuilding plan was determined based on rebuilding by the $\mathrm{T}_{\mathrm{MAX}}$ with a 50% probability ($\mathrm{T}_{\mathrm{TARGET}} = \mathrm{T}_{\mathrm{MAX}}$). The SPR was adjusted upwards if the predicted probability of rebuilding by the target year under the current SPR fell below 50% to a SPR that was estimated would rebuild with a 50% probability by $\mathrm{T}_{\mathrm{MAX}}$.

5. "Fixed" – The SPR in the initial rebuilding plan was determined based on rebuilding by $\mathrm{T}_{\mathrm{MAX}}$ with a $\mathrm{P}_{\mathrm{INIT}} = 60\%$. During rebuilding, the SPR was not updated until the rebuilding target year. If the stock was estimated not to have rebuilt by the target year, the SPR was set equal to either 125% of the PFMC SPR maximum sustainable yield proxy value by life history (SPR_{PROXY}: flatfish $SPR_{30\%}$, roundfish $SPR_{45\%}$, and rockfish $SPR_{50\%}$) or remained at the rebuilding SPR, whichever value was higher, until the stock was estimated to be rebuilt (i.e. this constraint prevented catch from increasing when the stock failed to rebuild if the rebuilding SPR was more conservative relative to 125% of the SPR_{PROXY}).

6. "Constant harvest rate" – The "constant harvest rate" rebuilding strategy deviates from all other strategies. The "constant harvest rate" strategy did not apply a rebuilding plan, but allowed for rebuilding by reducing harvest by setting the SPR rate to 125% of the PFMC SPR_{PROXY}. Since a rebuilding plan was not performed, an estimated minimum year, target year, and maximum year for recovery, including the rebuilding probability, were not estimated. While this rebuilding strategy would not currently be allowed under US law, it does represent a rebuilding alternative that may be applied outside the US.

The alternative rebuilding plans and sensitivities, except the constant harvest rate and fixed strategies including the fixed sensitivity test (see *Sensitivities* section below), applied some general rules to govern the amount catch could change between rebuilding analyses. The lower and upper limits of the multiplicative change to catches were 50 and 120%, respectively, of the previous catch. If the new estimated catch exceeded the upper bound, the catch was lowered. If the new estimated catch was below the lower bound, the target year was changed to maximum year for rebuilding and a new catch was estimated based on the updated target year. If the estimated catch was still below one-half of the previous catch, the current rebuilding plan was deemed a failure and a new plan was put in place (calculating new values for T_{TARGET} and T_{MAX}). In this case, the new rebuilding SPR was constrained to not be lower (result in a higher harvest rate) than the previous plan's SPR. In addition, these conditions for limiting the degree of changing in catch levels were not applied when the stock was first declared overfished and the initial rebuilding plan put in place, which is consistent with actual practice when a West Coast groundfish stock is initially placed under a rebuilding plan (although in practice, management has two years to implement a rebuilding plan and may proactively reduce catches prior to the rebuilding plan implementation). Similarly, when a stock was declared rebuilt, no conditions on the degree of change in catch were applied. The limits on the change in catch were arbitrarily selected, but were designed to capture the PFMC behavior when altering catch during and between rebuilding plans. Historically, management has been reactive to reduce catches during rebuilding based on more pessimistic assessments, but has been more apt to take a precautionary approach when the perception of stock biomass becomes more optimistic restricting large increases in catch during rebuilding.

3.2.5 Sensitivities

A number of sensitivity analyses were conducted to evaluate the performance of the rebuilding strategies given specific assumptions (Table 3.3):

1. "*Status quo* – natural mortality" – The natural mortality rate was biased high by 10% in the estimation model relative to the true (operating model) value in assessments through the first half of the initially estimated rebuilding period. Assuming a positively biased natural mortality value in the assessment will result in an estimate of stock status that is less pessimistic regarding the true state of the stock. A misspecification of 10% was applied because it was a level of error that still resulted in the estimation method estimating the stock to be overfished in the first assessment year. The natural mortality rate in the estimation method was updated to the true value halfway through the initially estimated rebuilding period ($T_{TARGET}/2$). This sensitivity explored the impact of overly optimistic assessment estimates for an extended period during rebuilding on the likelihood of meeting the rebuilding targets.

2. "*Status quo* – steepness" – The steepness parameter was biased high by 10% in the estimation model relative to the true (operating model) value in assessments through the first half of the initially estimated rebuilding period. A similar logic was applied in selecting a positive bias of 10%, as was considered for natural mortality, where this value resulted in the assessment estimating a less depleted stock relative to its true operating model status, but the stock was still estimated to be overfished in the first year. Steepness in the estimation method was updated to the true value halfway through the initially estimated rebuilding period ($T_{TARGET}/2$). Similar to the natural mortality sensitivity, this sensitivity explored the impact of overly optimistic assessment

estimates for an extended period during rebuilding on the likelihood of meeting the rebuilding targets.

3. "*Status quo* – assessment frequency" – The assessment frequency during rebuilding was either increased or decreased based on the life history type. The frequency of assessment increased for the shorter-lived flatfish and roundfish life histories to every two years, while the frequency of assessment decreased for both rockfish life history types from every four years to every eight years. This sensitivity explored the relationship between assessment frequency and performance of the rebuilding plan (i.e. are there benefits to increased or decreased monitoring of the stock during rebuilding?).

4. "Flexible – assessment frequency" – As for "status quo – assessment frequency", but based on the "flexible" strategy. This sensitivity explored the interaction between reduced thresholds for updating the SPR during rebuilding and reduced or increased assessment frequency.

5. "Risk averse – flexible" – The SPR in the initial rebuilding plan was determined based on rebuilding by $\mathrm{T_{MAX}}$ with a 75% probability ($\mathrm{P_{INIT}}$). The SPR was adjusted upwards if the probability of rebuilding by $\mathrm{T_{TARGET}}$ fell below 40% to a new SPR that would rebuild with a 75% probability by the $\mathrm{T_{TARGET}}$. This sensitivity explored the interaction and rebuilding performance if the initial rebuilding SPR is set conservatively and the threshold for updating the SPR during rebuilding is reduced (i.e. do these two adjustments offset each other?).

6. "Risk neutral – maintain 50%" – The SPR in the initial rebuilding plan was determined based on rebuilding by $\mathrm{T_{MAX}}$ with a 50% probability. The SPR was adjusted upwards or downwards in each subsequent analysis to maintain a SPR that would rebuild with a 50% probability by $\mathrm{T_{TARGET}}$. The key difference in this sensitivity is that the SPR was adjusted each time the rebuilding analysis was conducted to maintain a 50% probability

of rebuilding by the T_{TARGET}, whereas the other strategies and sensitivities (except the "fixed" strategies) only adjusted the SPR if the probability of rebuilding by T_{TARGET} was less than the probability threshold value (P_{TARGET}). This sensitivity examined the impact of maintaining a 50% probability over the course of the rebuilding plan and whether this approach performed similarly to alternative approaches that allowed fluctuations in the rebuilding probability.

7. "Fixed – mid-course update" – The SPR in the initial rebuilding plan was determined based on rebuilding by T_{MAX} with a 60% probability (P_{INIT}). If the rebuilding period was > 10 years, an updated rebuilding analysis was conducted at the halfway point to the target rebuilding year to evaluate progress. The SPR was adjusted upwards if the probability of rebuilding by T_{TARGET} fell below 50% to a new SPR that would rebuild by T_{TARGET} with a 50% probability. If the stock failed to rebuild by T_{TARGET}, the SPR was set equal to either 125% of the SPR_{PROXY} value (flatfish: $SPR_{30\%}$, roundfish: $SPR_{45\%}$, and rockfish: $SPR_{50\%}$) or remained at the rebuilding SPR, whichever value was higher until the stock was estimated to have rebuilt. This sensitivity examined the impact of reevaluating the rebuilding performance and making any required adjustments to the rebuilding plan mid-course ($T_{TARGET}/2$ if rebuilding time was > 10 years) compared to the "fixed" strategy which did not apply any adjustments during rebuilding.

3.2.6 Performance measures

The following eight performance metrics were used to evaluate each alternative rebuilding strategy across the 100 simulations (using the median and 80% simulation interval):

1. The number of SPR changes during rebuilding which was used as a proxy measurement for predictability during rebuilding for management and stakeholders.

2. The number of times the value of the target rebuilding year was changed, an additional

measurement for predictability during rebuilding.

3. The number of times a rebuilding plan failed to recover the stock to the target stock size, requiring a new rebuilding plan.

4. The annual average variability of the catches (abbreviation AAV) over the whole projection period, defined as:

$$AAV = 100\frac{\sum_t |C_t - C_{t+1}|}{\sum_t C_t} \tag{3.6}$$

where C_t is the catch during year t. Decreased variability in catches would provide additional predictability for management and stakeholders.

5. The average catch during a set number of years when the resource was under a rebuilding plan (flatfish: 5, roundfish: 10, medium-lived rockfish: 25, and long-lived rockfish: 50 years), which was used as a measure of the average catch attained during rebuilding for each alternative strategy. The vast majority of simulated stocks were not yet rebuilt at the end of the defined number of years by the life history. However, in cases where a stock rebuilt more quickly, the average was calculated over the shortened period of rebuilding.

6. The average catch over the entire projection period (rebuilding period and recovery catches), which was a measure of the trade-off between the length of rebuilding with reduced catches and the benefit of increased catches from a recovered stock.

7. The "rebuilding ratio", the ratio of the number of years under rebuilding (until the stock was assessed to be rebuilt) divided by the number of years that it was expected that rebuilding would take place by the initial rebuilding plan (the initial $\mathrm{T_{TARGET}}$) to evaluate the ability of each alternative rebuilding strategy to meet the initial rebuilding estimates.

8. The number of years estimated for the overfished stock to rebuild to the target stock size from the initial rebuilding plan.

3.3 Results

3.3.1 Rebuilding strategies

The six alternative rebuilding strategies led to successful rebuilding for the majority of simulations across the life histories (Table 3.4). However, there were differences in performance across the strategies by life history type where varying adjustments to the SPR were required to meet the rebuilding timeline (Fig. 3.2: "*Status quo*" rebuilding example). The extent of increase (decrease in fishing effort) from the SPR_{PROXY} to the SPR applied during the initial rebuilding plan varied among life histories (Fig. 3.3). The most severe changes in the SPR occurred for the flatfish life history (Fig. 3.2a) and a small subset of the simulated roundfish stocks (Fig. 3.2b). These simulated flatfish and roundfish stocks were determined to be able to rebuild in < 10 years in the absence of fishing, triggering the 10-year rebuilding rule as required by the current US federal guidelines, requiring large adjustments to the SPR to meet the rebuilding timeframe.

Estimation error resulted in a number of simulated stocks being incorrectly declared rebuilt when the true operating model stock was still below the target biomass (Table 3.4). The difference between these values was most marked for the long-lived rockfish life history. This occurred when the true stock was close to being rebuilt, but still below the management proxy target stock size and the estimation method overestimated stock status resulting in the stock being declared rebuilt. Once rebuilt, catch was set using a SPR proxy value applied for all US west coast rockfish stocks. The SPR proxy value ($SPR_{50\%}$) will maintain the population at the management target ($SB_{40\%}$) when steepness is equal to 0.60 (based on the Beverton-Holt stock–recruit relationship), but the long-lived rockfish steepness was lower (0.50) resulting in an average stock size slightly below the management target biomass ($SB_{40\%}$).

Performance of the "*status quo*" and the "flexible" rebuilding strategies were nearly identical for both the flatfish and roundfish life histories (Figs. 3.4 and 3.5). The faster dynamics of each of these life histories resulted in shorter rebuilding times with little variance in the number of times the SPR was adjusted between the two strategies (Figs. 3.4a and 3.5a). The median number of SPR changes during rebuilding was higher for the roundfish life history. However, the rebuilding time was generally twice that required for the flatfish life history (Table 3.5). Across all life histories, both strategies resulted in median rebuilding times that were equal to or less than the rebuilding time estimates during the initial rebuilding analysis (Figs. 3.4f–3.7f). However, the slower dynamics and longer rebuilding periods associated with rockfishes led to differences between the "*status quo*" and the "flexible" rebuilding strategies for those life histories. The lower threshold probability of rebuilding by the target year ($\mathrm{P_{TARGET}}$) for the "flexible" strategy resulted in fewer SPR updates during rebuilding for the two rockfish life histories (median SPR changes – "*status quo*": 3, "flexible": 1). The average catch over the fixed period and the total average catch over all projection years did not vary greatly among strategies (Figs. 3.6e–3.7e and Table 3.6), and the "*status quo*" and "flexible" rebuilding strategies resulted in nearly identical rebuilding times for the medium- and long-lived rockfish life histories (Table 3.5).

The "risk averse" rebuilding strategy resulted in approximately a 10% faster rebuilding time relative to the "*status quo*" strategy for each of the life history types (Table 3.5). The faster rebuilding times of the "risk averse" strategy were achieved by having lower average catches during rebuilding (Figs. 3.4e–3.7e) compared to the "*status quo*" strategy. However, the average catch over all projection years for each strategy were comparable across all life histories (Table 3.6). This highlights the trade-off between the length of rebuilding with reduced catches and the benefit of increased catches from a recovered stock. The higher probability associated with rebuilding by the target year ($\mathrm{P_{TARGET}}$) for the "risk averse" strategy resulted in an increased number of SPR changes for both the rockfish life histories (Figs. 3.6a and 3.7a), but did not result in a median increase in the number of SPR changes for the flatfish or roundfish life histories (Figs. 3.4a and 3.5a).

The "risk neutral" rebuilding strategy defined the rebuilding SPR assuming a 50% probability of rebuilding by T_{MAX}. Relative to the "*status quo*" strategy, the "risk neutral" approach resulted in modest increases in the median average catch over the rebuilding period (Figs. 3.4e–3.7e), similar average catch over all projection years (Table 3.6), and lower median AAV (Figs. 3.4d–3.7d). However, compared to "*status quo*", the "risk neutral" strategy had an increased frequency of rebuilding failure (increased 80% simulation interval) for the rockfish life histories (Figs. 3.6c and 3.7c). Apart from flatfish, this strategy resulted in longer and more variable rebuilding times compared to the other strategies (Table 3.5).

The "fixed" rebuilding strategy resulted in a median average catch during rebuilding comparable to the other strategies (Figs. 3.4e–3.7e), but with longer median rebuilding periods relative to the "*status quo*" strategy for the rockfish and the roundfish life histories (Table 3.5). The extended rebuilding period resulted in lower average catches over all projection years (Table 3.6). Across life histories, the "fixed" rebuilding strategy estimated rebuilt stocks by the target rebuilding year for the majority of the simulations [88% (flatfish), 72% (roundfish), 58% (medium-lived rockfish), and 72% (long-lived rockfish)] without requiring additional adjustments to the SPR due to not rebuilding by the target year.

The "constant harvest rate" rebuilding strategy that did not apply a rebuilding plan, but rather reduced harvest by an increase in the SPR rate while the stocks were overfished resulted in a lower AAV in catches for all life histories (Figs. 3.4d–3.7d) and higher average catches during rebuilding (Figs. 3.4e–3.7e), except for the medium-lived rockfish life history. However, the average catch over the projection period was lower relative to the "*status quo*" strategy for all life histories (Table 3.6). The SPR rate applied while the stocks were overfished was lower (higher catches) for all life histories, except for the medium-lived rockfish (Fig. 3.3), which resulted in higher catches while the stocks were rebuilding, but also generally increased the number of years to rebuild (Table 3.5) and hence lower average catches in the projection period (Table 3.6). The "constant harvest rate" strategy resulted in rebuilding times that were greater than the "*status quo*" strategy initial rebuilding plan T_{MAX} year for 97, 55, 5, and 82% of the simulations for the flatfish, roundfish, medium-lived, and long-lived rockfish life histories, respectively. The maximum time allowed for rebuilding under a formal rebuilding plan is defined as the minimum time to rebuild in the absence of fishing plus one mean generation period, but the "constant harvest rate" strategy did not impose this requirement. This differing application resulted in the varying results for the medium-lived rockfish life history where the strategies that applied a rebuilding strategy were allowed a longer period to rebuild the stock and hence applied more aggressive SPR rates during rebuilding relative to the "constant harvest rate" strategy (Fig. 3.3).

3.3.2 Sensitivities

Impact of parameter misspecification

Both sensitivities that examined the impact of parameter misspecification, natural mortality ("*status quo* – natural mortality") and steepness ("*status quo* – steepness") resulted in an increase in the median times the SPR needed to be changed in the attempt to rebuild by the target year for each rockfish life history relative to the "*status quo*" strategy (Figs. 3.4a–3.7a). The majority of simulations for the shorter-lived flatfish and roundfish life histories failed to rebuild the stock by the initial estimate of T_{MAX}, requiring a new rebuilding plan (Figs. 3.4c and 3.5c). The longer-lived life histories only required adjustments to target rebuilding year for the sensitivity that misspecified natural mortality (Figs. 3.6b and 3.7b). The longer rebuilding times associated with the majority of rockfish life histories simulations allowed for sufficient time to adjust the rebuilding SPR to still rebuild in similar median rebuilding times relative to the "*status quo*" strategy once the misspecified parameter was corrected half-way through the initial rebuilding period (Table 3.5). Additionally, the misspecified parameters led to an overly optimistic estimate of the initial stock status resulting in estimated shorter rebuilding times relative to the true time required to rebuild the stock. This underestimate resulted in shortened rebuilding timelines relative to the "*status quo*" strategy, and even when the stock was not rebuilt by the target rebuilding year, the stock was often rebuilt in a similar timeframe as the strategies without parameter misspecification (Table 3.5).

Impact of assessment frequency

The impact of either increasing or decreasing the assessment frequency varied based on the life history. The sensitivity runs that examined assessment frequency for the "*status quo*" and the "flexible" strategies for the medium- and long-lived rockfishes resulted in similar median rebuilding times, with either the same or fewer changes to the SPR during rebuilding relative to the "*status quo*" and "flexible" strategies, which have assessments every fourth year (Figs. 3.6a and 3.7a). Reducing the frequency of assessment for the rockfishes from every fourth to every eighth year also resulted in a higher average catch during rebuilding for both the medium- and long-lived rockfish (Figs. 3.6e and 3.7e). However, increasing the assessment frequency for the fast dynamic life histories (flatfish and roundfish) from every fourth to every second year resulted in lower average catches during rebuilding (Figs. 3.4e and 3.5e), with a larger range of median SPR changes (Figs. 3.4a and 3.5a), and did not rebuild in shorter periods relative to each of the base strategies (Table 3.5).

Exploration of alternative threshold values for setting and changing the rebuilding SPR

The impact of altering the threshold probability that triggered a change to the SPR during rebuilding varied based on the initial probability of rebuilding selected to define the rebuilding timeline. The "risk averse – flexible" strategy that applied a high initial probability of rebuilding and allowed for increased flexibility prior to altering the rebuilding plan compared to the "risk averse" strategy resulted in generally fewer SPR changes during rebuilding (Figs. 3.4a–3.7a), while rebuilding the stock in a similar amount of time (Table 3.5) with comparable AAV (Figs. 3.4d–3.7d), average rebuilding catch (Figs. 3.4e–3.7e), and average catch over all projection years (Table 3.6) compared to the "risk averse" strategy.

Across all life histories, the "risk neutral – maintain" 50% rebuilding strategy, which maintained a 50% probability of rebuilding by $\mathrm{T_{TARGET}}$ for the duration of the rebuilding period, resulted in an increase in the median number of SPR and $\mathrm{T_{TARGET}}$ changes with an increase in failed rebuilding plans compared to the base "risk neutral" strategy (Figs.

3.4–3.7a–c). This result was most evident for the long-lived rockfish life history which resulted in 22 (median across simulation) SPR changes over the course of rebuilding (Fig. 3.7a). Across the alternative rebuilding strategies, this strategy led to the highest median average catch during the defined rebuilding periods (Figs. 3.4e–3.7e), but also had the longest median rebuilding times (Table 3.5) and the highest median AAV in catch for each of the life histories (Figs. 3.4d–3.7d).

Updating the SPR for the "fixed mid-course update" rebuilding strategy did not alter the overall results compared to the "fixed" rebuilding strategy where no update was performed (Figs. 3.4–3.7). The mid-course update was only performed if the rebuilding period was $>$ 10 years, which resulted in almost no updates for the flatfish life history (Fig. 3.4). A similar number of simulations successfully rebuilt by the target rebuilding year for the "fixed mid-course update" strategy [88% (flatfish), 71% (roundfish), 61% (medium-lived rockfish), and 74% (long-lived rockfish)] compared to the "fixed" strategy [88% (flatfish), 72% (roundfish), 58% (medium-lived rockfish), and 72% (long-lived rockfish)].

3.4 Discussion

The performance of alternative rebuilding plans that applied alternative values for the initial rebuilding probability and various threshold probabilities for updating the SPR during rebuilding were explored. The rebuilding plans that implemented a higher initial rebuilding probability ($\geq$ 60%) for determining rebuilding fishing mortality and targets generally resulted in fewer changes to the rebuilding plans and rebuilt by the target rebuilding year, particularly for stocks with the longer rebuilding plans (e.g. rockfishes). Punt and Ralston (2007) also determined that a key to a successful rebuilding plan was setting targets and fishing mortality at a rate that can buffer against future uncertainty to ensure that rebuilding deadlines are met. The strategies that not only incorporated a higher initial rebuilding probability, but also allowed for a lower threshold probability (P_{TARGET} = 40%) during rebuilding were less responsive to noise, resulting in fewer changes to fishing mortality and rebuilding targets while still successfully rebuilding stocks. Attempting to maintain a 50%

probability of rebuilding was the poorest performing rebuilding strategy based on the performance metrics, resulting in a high number of changes to the rebuilding plan, an increased frequency of failing to meet rebuilding targets, and a higher variation in catch.

The current US federal guideline to rebuild in $\leq$ 10 years were possible given that the biology of the stock impacted the performance of the alternative rebuilding plans for the flatfish and roundfish life histories. The majority of flatfish stocks and a select number of the simulated roundfish stocks were estimated to be able to rebuild in $\leq$ 10 years in the absence of fishing, requiring the maximum rebuilding time to be set at $\leq$ 10 years. An estimated minimum rebuilding time for a majority of simulated flatfish stocks was ca. 6–8 years across the simulations. The limited number of years between the minimum and maximum years for rebuilding resulted in extreme reductions in the fishing mortality rate at the start of rebuilding relative to the other life histories (Fig. 3.3). Overall, the results of the alternative rebuilding strategies for the flatfish life history showed little contrast across strategies, highlighting that the 10-year rule was the primary driver for rebuilding performance rather than the strategy applied. Additionally, the fact that only some roundfish simulations were estimated to be able to rebuild in < 10 years resulted in bimodal distributions for the average catch and rebuilding SPR rates, requiring relatively large reductions to harvest in order to rebuild within 10 years, compared to stocks that were allowed an average of 20 years to rebuild. This behavior highlights the discontinuity in the current US rebuilding guidelines, resulting in very different rebuilding plans for stocks where the minimum time for rebuilding is < or > 10 years. Patrick and Cope (2014) have outlined several alternatives for defining the maximum time allowed for rebuilding short-lived stocks that would be consistent across life histories. One suggestion made by Patrick and Cope (2014) would change the definition of the maximum year being set at twice the minimum rebuilding time (T_{MAX} = 2 x T_{MIN}), which for the flatfish life history would reduce the extreme initial reduction in fishing mortality rates required for rebuilding while only extending rebuilding timelines by ca. 2–6 years.

When a stock is declared rebuilt, an additional challenge is the potential substantial change in the catch level compared to the limited catches allowed during rebuilding. This was an issue for each of the life histories explored, but perhaps the most extreme for the flatfish life history where the rebuilding harvest rate was the most constrained relative to the management proxy harvest level. When a stock rebuilds, harvest predictions are based on the management harvest control rule and proxy harvest levels which are designed to obtain the maximum acceptable biological catch at the target biomass. This catch level can be substantially larger than the rebuilding catches (Thorson and Wetzel, 2015). An alternative rebuilding approach has been applied by the PFMC for West Coast petrale sole (Pacific Fishery Management Council, 2011). The stock was deemed able to rebuild in ≤ 10 years based on projections using the management harvest control rule which reduces catches linearly when the stock is below the relative target biomass to zero at a lower threshold relative stock size. In this instance, the Council adopted rebuilding catches based on the harvest control rule rather than the catches predicted by the traditional rebuilding plan. Predicting catches based on the harvest control rule that applied a linear reduction resulted in a smooth ramp between the rebuilding and rebuilt catch values while successfully rebuilding the stock.

Currently, the US federal rebuilding plans are required to contain specific components that define a rebuilding timeline, the probability of rebuilding by the target year, and harvest rate in order to achieve rebuilding. However, there are distinct trade-offs between rebuilding as quickly as possible through sometimes extreme harvest reductions and the economic and societal costs of doing so (Hilborn et al., 2012). The "constant harvest rate" strategy involved simple reductions in harvest that are consistent over the rebuilding period and also may limit the amount in lost yield by applying less extreme harvest restrictions. The "constant harvest rate" strategy resulted in higher average catches during rebuilding with lower annual variation (except for the medium-lived rockfish), but averaged longer rebuilding periods and lower average catch over the whole projection period, highlighting the trade-offs that should be considered by management and stakeholders when determining the strategy for rebuilding an overfished stock.

A rebuilding strategy will only be effective when management is responsive and is able to control fishing mortality to a rate at or below the level required for the stock to rebuild (Patrick et al., 2013). Once a stock is identified as being overfished, a delay in implementing a rebuilding plan can negatively impact rebuilding (Shertzer and Prager, 2007) if that stock is experiencing overfishing (i.e. removals exceed the maximum sustainable yield), and can be especially important for stocks that have a low intrinsic rate of growth (Neubauer et al., 2013). All rebuilding strategies in this paper were implemented and rebuilding harvest levels applied the year immediately following the overfished status determination, and catches were taken without error. The rebuilding timelines would have been extended requiring harvest restrictions for longer periods if there had been a delay in implementation of the rebuilding plan and if management was ineffective at reducing catch to a level at or below the rebuilding values.

The results here were designed to evaluate alternative rebuilding strategies for West Coast groundfish stocks. To determine the impact of each alternative rebuilding strategy, misspecification between the operating model and the estimation method was limited and was explored only in some sensitivity analyses, which allowed for the results to be attributed to the rebuilding plan rather than model misspecification. Although the assessment was fully simulated, the structural assumptions between the operating and estimation model matched, an attribute that is not commonly attainable in real-world assessments; hence, the assessment should not be considered entirely reflective of the uncertainties inherent in traditional stock assessments. Additionally, a relatively high effective sample size of age data that was informative for all life histories (although the sample size likely resulted in greater precision for the long-lived stocks) was provided to the estimation model. There are two types of error that could impact the results substantially which could be identified with model diagnostics in a real assessment, but something that is not easily done in a simulation framework. The first is poor estimation of key biological parameters (or parameters assumed known at incorrect values) such as natural mortality or steepness. The misspecification of each of these parameters was explored in separate sensitivity runs, but for only limited periods. Long-term misspecification or time-varying changes that are not accounted for in the assessment would reduce the performance of all the rebuilding strategies explored here.

The second type of error that could impact the interpretation of the results is not accounting for changes in the biology, productivity, or species interactions of the stock over time. Rebuilding projections depend on the values of biological parameters in the final year of assessment and predict future recruitment by sampling from historical recruitments from the stock–recruitment curve. There has been considerable concern about the impact of climate change on fish stocks, specifically how it may affect future recruitment (Hollowed et al., 2011; Ianelli et al., 2011; Mueter et al., 2011; Stachura et al., 2014). Decadal swings in productivity could result in rebuilding trajectories that deviate below or above the projected probability of rebuilding by the target year. Simulation testing has shown that detecting, predicting, and making the correct management adjustments to shifts in productivity can be very challenging (Haltuch and Punt, 2011; Szuwalski and Punt, 2012). Management strategies that specifically account for environmentally-driven recruitment have not always been shown to outperform those that assume a form of average recruitment in the future (A'mar et al., 2009; Punt, 2011; Punt et al., 2014a). Under these conditions over long rebuilding periods, the overall average recruitment may not deviate greatly from the forecasted levels even if there were periods that were above or below predicted levels. However, long-term rebuilding plans based on historical recruitment level will likely not perform well and could fail to reduce fishing mortality to a level required to rebuild the stock. Additionally, changing future conditions may result in long-term shifts in biological parameters (e.g. natural mortality, growth) (Swain and Benoît, 2015), altering the sustainable yield available from a stock and if not accounted for could impact the ability to rebuild a stock (Legault and Palmer, 2015).

This work focused specifically on rebuilding strategies applied for US west coast groundfish. However, the results can be informative for fisheries managers outside of this region to determine rebuilding plans that best meet management goals. There are trade-offs that must be considered when determining rebuilding fishing mortality levels for rebuilding overfished stocks. The development of rebuilding plans should consider the life history of the stock and the major sources of uncertainty. Future work applying selected rebuilding strategies identified here should be conducted to explore the impact of additional misspecification between the operating and estimation models and how varying future conditions may impact rebuilding performance.

Acknowledgments

This paper benefited greatly from comments by Ray Hilborn (University of Washington), Tim Essington (University of Washington), Trevor Branch (University of Washington), Ian Stewart (International Pacific Halibut Commission), Owen Hamel (Northwest Fisheries Science Center), Michelle McClure (Northwest Fisheries Science Center), Chris Legualt (Northeast Fisheries Science Center), and an anonymous reviewer. Additionally, feedback from the Pacific Fishery Management Council and the advisory groups (Groundfish Management Team, Groundfish Advisory Panel, Science and Statistical Committee) was instrumental in the development and analysis of this work.

3.5 Tables

Table 3.1: The US rebuilding plan required components and definitions.

Terminology	Definition
T_{MIN}	The minimum amount of time a stock could rebuild in the absence of fishing.
T_{MAX}	The maximum time allowed for a stock to rebuild, which cannot exceed the T_{MIN} plus one mean generation time.
T_{TARGET}	The target year for rebuilding, which must fall between T_{MIN} and T_{MAX}.
P_{INIT}	The initial probability for rebuilding T_{MAX}, which determines the appropriate rebuilding spawning potential ration (SPR) value for rebuilding by T_{MAX}. Defined by management, but must be greater than $\geq 50\%$.
P_{TARGET}	The probability of rebuilding by T_{TARGET} based on the intended spawning potential ratio (SPR). Set by management.

Table 3.2: Life history parameters used in the operating model for each life history.

Parameter	Sex	Flatfish	Roundfish	Medium-lived rockfish	Long-lived rockfish
Natural mortality (yr^{-1})	Female	0.15	0.20	0.08	0.05
	Male	0.17	0.20	0.09	0.06
Steepness (h)		0.85	0.70	0.65	0.50
Maximum length (L_∞) (cm)	Female	58	65	34	64
	Male	51	58	32	66
Growth coefficient (k)(yr^{-1})	Female	0.133	0.120	0.115	0.047
	Male	0.213	0.150	0.153	0.047
Body weight $w_l = \alpha L^\beta$ (kg)					
Growth coefficient (α)	Female	$2.06e^{-6}$	$8.50e^{-6}$	$7.40e^{-6}$	$9.76e^{-6}$
	Male	$3.05e^{-6}$	$7.70e^{-6}$	$8.30e^{-6}$	$8.70e^{-6}$
Growth exponent (β)	Female	3.50	3.10	3.17	3.17
	Male	3.40	3.05	3.13	3.10
Maturity slope (yr^{-1})		-0.75	-0.70	-0.67	-0.44
Length at 50% maturity (cm)		33	35	21	38
Mean generation time (yr)		18	28	40	50
Recruitment variation (σ_R)		0.60	0.60	0.60	0.60
Catchability coefficient (Q)		1	1	1	1
Survey standard error (σ_s)		0.20	0.20	0.20	0.20
Fishery selectivity (logistic)					
Age at inflection		7	5	7	15
Width for 95% selection		2	2	5	7
Survey selectivity (logistic)					
Age at inflection		5	3	3	10
Width for 95% selection		2	2	3	7
Initial relative stock size ($SB_{t=50}/SB_0$)		0.05	0.10	0.10	0.10

Table 3.3: The alternative initial rebuilding probability (P_{INIT}), threshold probability during rebuilding (P_{TARGET}), and the assessment frequency explored by each of the rebuilding strategies and sensitivities. See the *Alternative management actions* and *Sensitivities* subsections in the Material & Methods for additional details.

Rebuilding strategy	P_{INIT} (%)	P_{TARGET} (%)	Assessment freq. (yrs)	Special conditions
Status quo	60	50	4	See *Alternative management actions* for additional details
Flexible	60	40	4	See *Alternative management actions* for additional details
Risk averse	75	60	4	See *Alternative management actions* for additional details
Risk neutral	50	50	4	The *SPR* in the initial rebuilding plan was determined based on rebuilding by the T_{MAX} with a 50% probability ($T_{TARGET} = T_{MAX}$). See *Alternative management actions* for additional details
Fixed	60	-	4	During rebuilding the *SPR* was not updated until the rebuilding target year. See *Alternative management actions* for additional details
Constant harvest rate	-	-	4	No rebuilding plan, but allowed for rebuilding by reducing harvest by setting the *SPR* rate to 125% of the PFMC SPR_{PROXY} until the stock rebuilt. See *Alternative management actions* for additional details

Table 3.3 continued:

Sensitivity	P_{INIT} (%)	P_{TARGET} (%)	Assessment freq. (yrs)	Special conditions
Status quo - natural mortality	60	50	4	Natural mortality was biased high by 10% in the estimation model relative to the true value (operating model). See *Sensitivities* for additional details.
Status quo - steepness	60	50	4	Steepness was biased high by 10% in the estimation model relative to the true value (operating model). See *Sensitivities* for additional details.
Status quo - assessment frequency	60	50	2/8	Either increase or decrease the assessment frequency base on the life history (flatfish and roundfish assessed every second year, both rockfishes assessed every eighth year). See *Sensitivities* for additional details.
Flexible - assessment frequency	60	40	2/8	As for "*status quo* - assessment frequency" but based on the "flexible" strategy. See *Sensitivities* for additional details.
Risk averse - flexible	75	40	4	See *Sensitivities* for additional details.
Risk neutral - maintain 50%	50	50	4	SPR was adjusted every four years to maintain 50% probability of rebuilding by T_{TARGET}. See *Sensitivities* for additional details.

Table 3.3 continued:

Sensitivity	P_{INIT} (%)	P_{TARGET} (%)	Assessment freq. (yrs)	Special conditions
Fixed rebuilding - mid-course update	60	50	4	The SPR was adjusted upwards if the probability of rebuilding fell below 50% halfway through the initial estimated rebuilding period. See *Sensitivities* for additional details.

Table 3.4: The percentage of stocks that failed to rebuild according to the estimation method compared to the percentage of stocks that actually (i.e. within the operating model) failed to rebuild for each life history by alternative rebuilding strategy and sensitivity.

	Flatfish		Roundfish		Medium-lived rockfish		Long-lived rockfish	
Alternative strategy	True	Est.	True	Est.	True	Est.	True	Est.
Status quo rebuilding	4	0	6	0	6	0	32	0
Flexible rebuilding	4	0	6	0	5	0	32	0
Risk averse rebuilding	4	0	6	0	3	0	30	0
Risk neutral rebuilding	3	0	8	0	8	0	33	0
Fixed rebuilding	4	1	11	1	8	0	34	0
Constant harvest rate rebuilding	7	0	8	1	2	0	47	25
Sensitivity								
Status quo - natural mortality	4	0	7	1	7	0	36	1
Status quo - steepness	4	0	8	0	6	0	32	1
Status quo - assessment frequency	3	0	11	0	2	0	26	0
Flexible - assessment frequency	3	0	12	0	2	0	27	0
Flexible - risk averse	3	0	9	0	2	0	32	0
Risk neutral - maintain 50%	6	2	11	2	8	0	41	11
Fixed - mid-course update	4	0	10	1	7	0	34	1

Table 3.5: The median estimated number of years to rebuild the stock to the target relative stock size and the 80% simulation interval for each life history by alternative rebuilding strategy and sensitivity.

	Flatfish		Roundfish		Medium-lived rockfish		Long-lived rockfish	
Alternative strategy	Years	Interval	Years	Interval	Years	Interval	Years	Interval
Status quo rebuilding	10	(7-17)	20	(8-34)	41	(30-57)	87	(65-100)
Flexible rebuilding	10	(7-17)	20	(8-34)	41	(30-54)	87.5	(68-101)
Risk averse rebuilding	9	(6-20)	18	(9-30)	36	(26-45)	80	(68-95)
Risk neutral rebuilding	9	(6-20)	21.5	(9-36)	43	(32-56)	90	(73-105)
Fixed rebuilding	10	(6-17)	19	(8-41)	43.5	(33-60)	91	(72-111)
Constant harvest rate rebuilding	14	(8-27)	21	(11-43)	34	(24-53)	105	(79-119)
Sensitivity								
Status quo - natural mortality	13	(8-21)	16	(10-29)	38	(31-48)	88	(69-113)
Status quo - steepness	14	(9-26)	14	(9-32)	40	(33-48)	84	(71-100)
Status quo - assessment frequency	10	(7-16)	20	(8-32)	41	(31-57)	88	(69-100)
Flexible - assessment frequency	10	(7-16)	20	(9-32)	41	(31-56)	88.5	(72-100)
Flexible - risk averse	10	(6-18)	18	(8-29)	37	(27-49)	81	(66-95)
Risk neutral - maintain 50%	10	(6-28)	23	(9-40)	46	(32-64)	96	(77-115)
Fixed - mid-course update	10	(6-17)	20	(8-40)	43	(63-60)	91	(69-111)

Table 3.6: The median average catch over the projection period covering both the catches obtained during rebuilding and when the stock recovered along with the 80% simulation interval for each life history by alternative rebuilding strategy and sensitivity.

	Flatfish		Roundfish		Medium-lived rockfish		Long-lived rockfish	
Alternative strategy	Catch	Interval	Catch	Interval	Catch	Interval	Catch	Interval
Status quo rebuilding	462	(413-516)	338	(294-389)	741	(656-828)	33	(30-36)
Flexible rebuilding	462	(413-516)	338	(294-389)	742	(651-827)	33	(30-36)
Risk averse rebuilding	464	(414-516)	340	(295-387)	767	(675-847)	34	(31-37)
Risk neutral rebuilding	460	(405-519)	334	(287-379)	728	(643-822)	33	(30-36)
Fixed rebuilding	439	(384-502)	314	(266-369)	728	(620-837)	26	(23-29)
Constant harvest rate rebuilding	435	(385-498)	313	(269-363)	719	(621-822)	25	(22-29)
Sensitivity								
Status quo - natural mortality	455	(405-517)	356	(306-405)	763	(683-857)	34	(31-38)
Status quo - steepness	451	(399-504)	343	(298-400)	738	(651-829)	33	(30-35)
Status quo - assessment frequency	460	(411-519)	337	(290-383)	743	(656-828)	33	(30-36)
Flexible - assessment frequency	460	(411-519)	336	(290-383)	741	(653-827)	33	(30-36)
Flexible - risk averse	463	(414-522)	341	(296-387)	760	(673-849)	34	(31-37)
Risk neutral - maintain 50%	456	(402-514)	328	(282-372)	724	(634-808)	32	(29-35)
Fixed - mid-course update	439	(384-502)	314	(268-369)	730	(622-841)	26	(23-29)

3.6 Figures

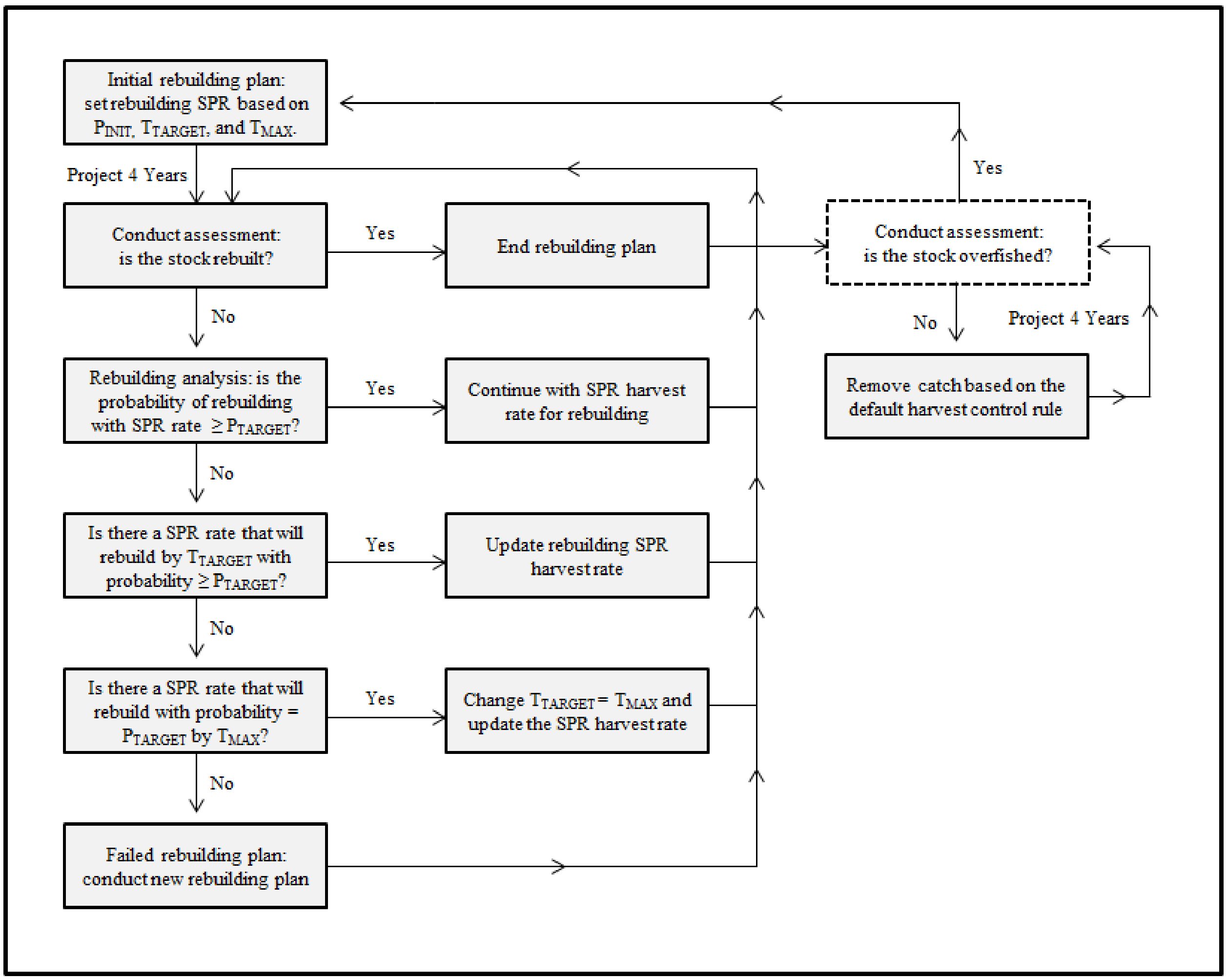

Figure 3.1: The process followed for determining when a rebuilding plan was implemented, how targets and harvest rates are adjusted during rebuilding, and the assessment for rebuilt stocks. The closed loop process starts by conducting the first assessment in year 50 (white box with dashed border) and continues for a fixed number of years for each life history.

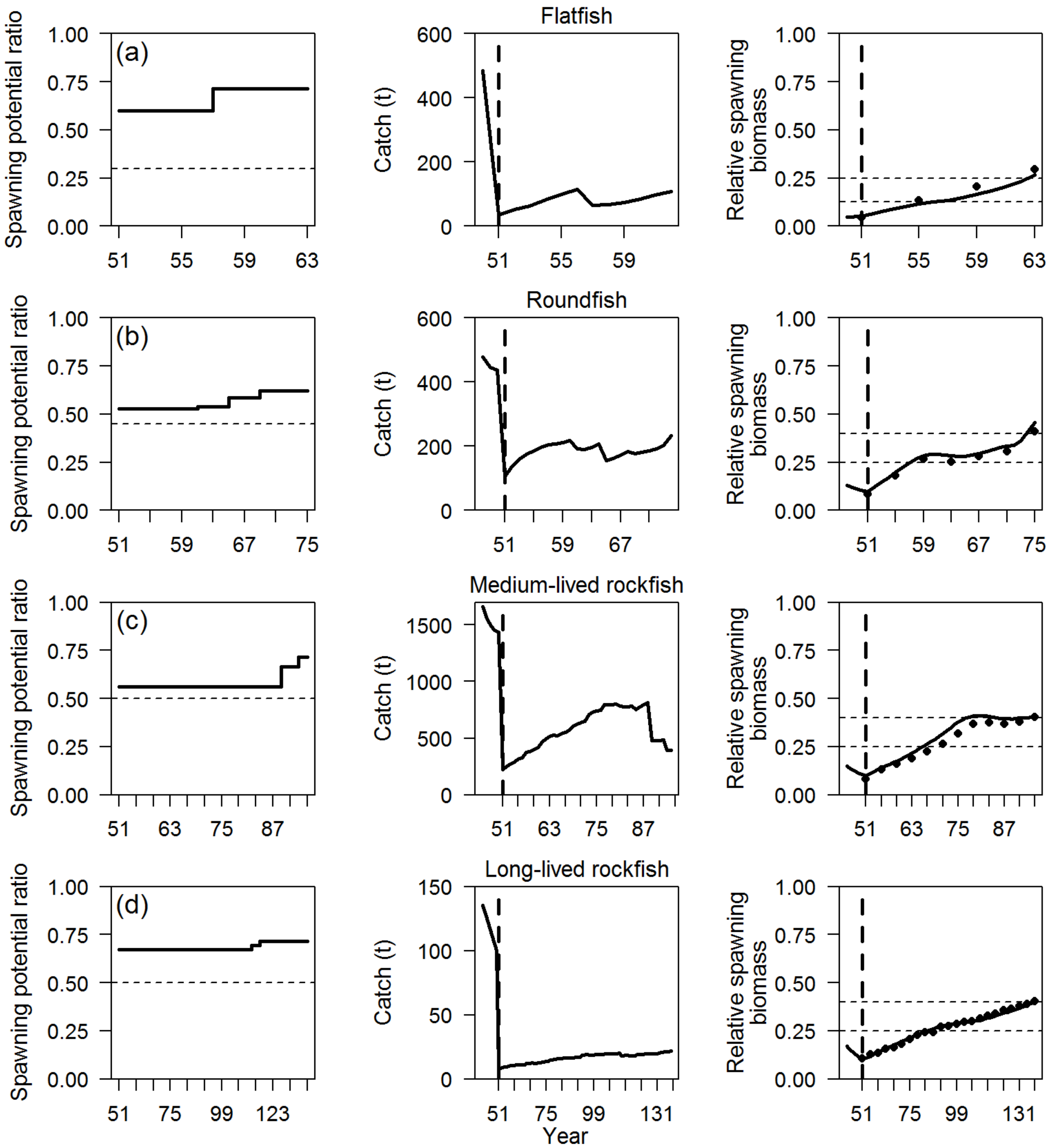

Figure 3.2: Results from an illustrative simulation for the "*status quo*" rebuilding strategy for each life history evaluated. The left set of panels summarizes the changes in the SPR during rebuilding, with the dashed line indicating the SPR_{PROXY} value for each life history. The middle panels give the trajectory of the catches during rebuilding for the example simulations. The right panels show the estimated relative spawning biomass (solid dots) and the true operating model relative spawning biomass trajectory (solid line), with the horizontal dashed lines indicating the overfished threshold and the target levels for each life history.

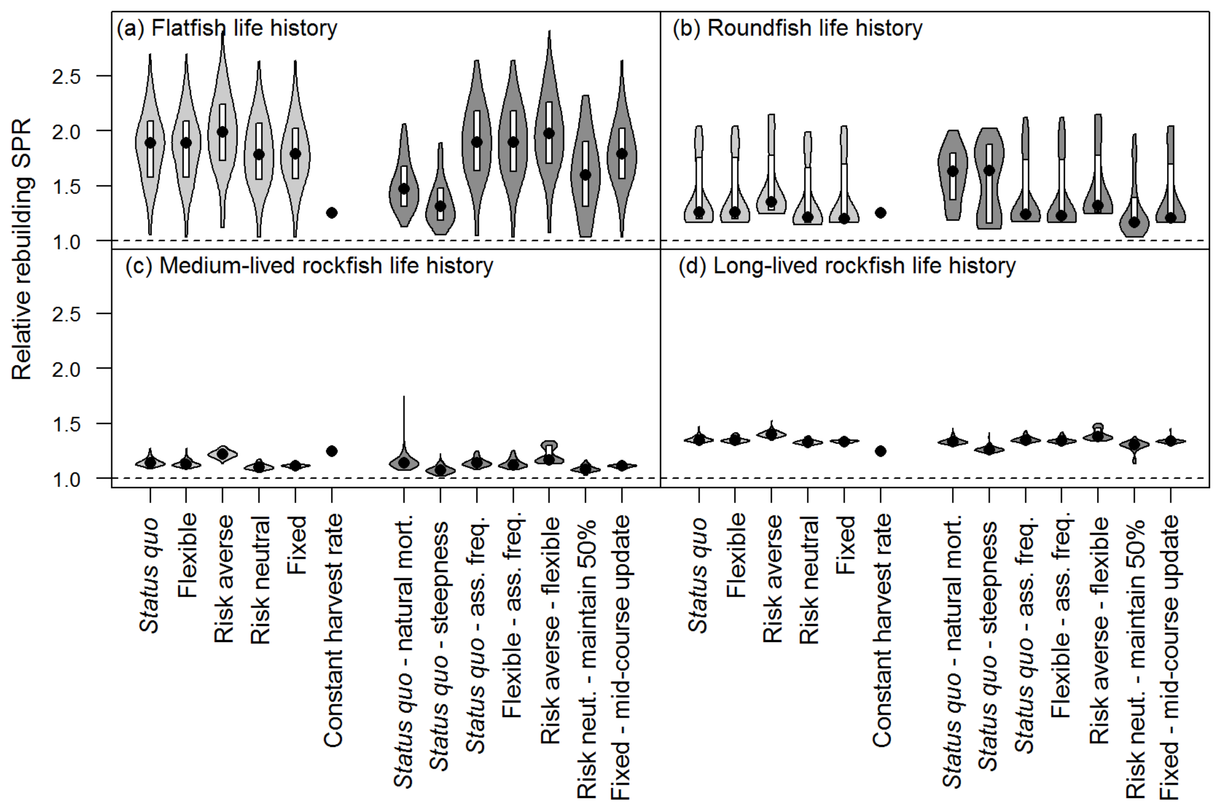

Figure 3.3: The mean SPR value during rebuilding relative to the management proxy SPR value [e.g. the ratio of the rebuilding SPR to the management target SPR (flatfish: 0.30, roundfish: 0.45, rockfish: 0.50)] for each alternative rebuilding strategy (light grey) and sensitivity (dark grey) for flatfish (a), roundfish (b), medium-lived rockfish (c), and long-lived rockfish (d) life history types, where the points are the median value and the white bars indicate the 80% simulation interval.

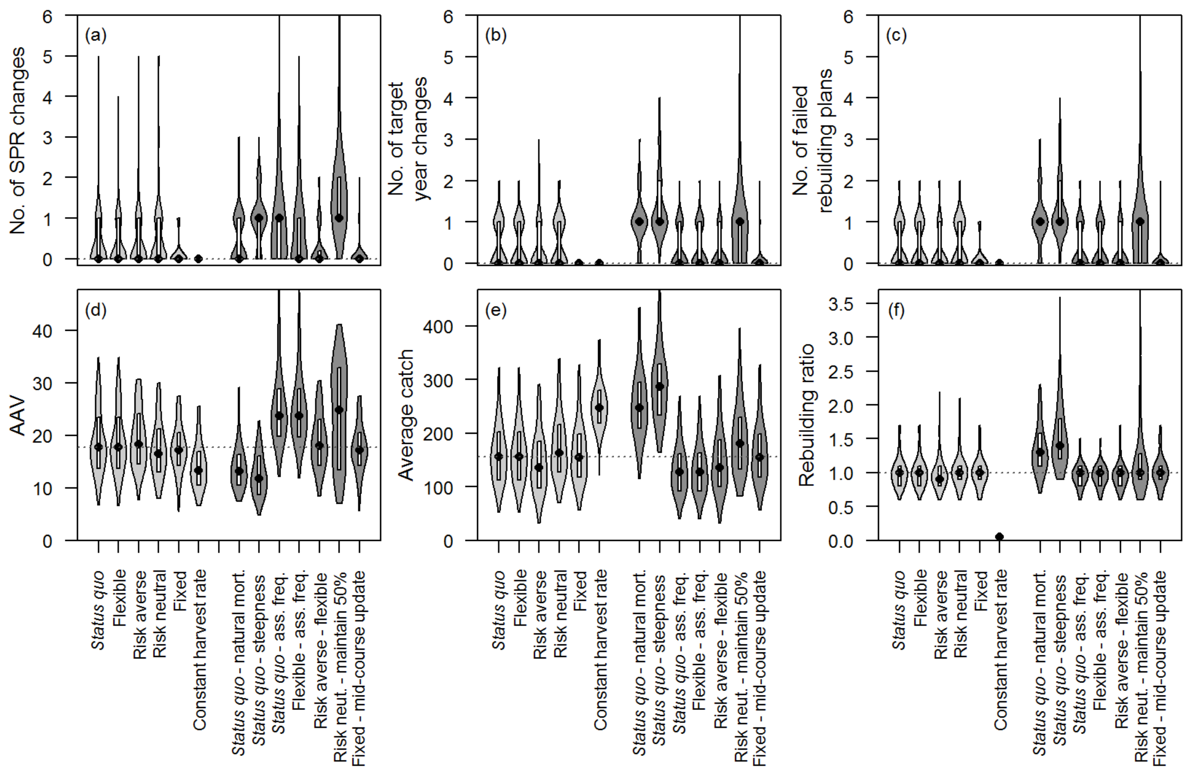

Figure 3.4: The flatfish life history median values (black points) and 80% simulation intervals (white bars) for the number of SPR changes (a), number of changes to the target rebuilding year (b), the number of failed rebuilding plans (c), the average annual variation in catch (AAV) (d), the average catch over the first five years of the rebuilding period (e), and the rebuilding ratio (f) for each of the alternative rebuilding strategies (light grey) and sensitivities (dark grey). The dashed horizontal line indicates the median value for the "*status quo*" strategy for visual reference for plots (a)–(e). The dashed horizontal line in (f) is set at 1.

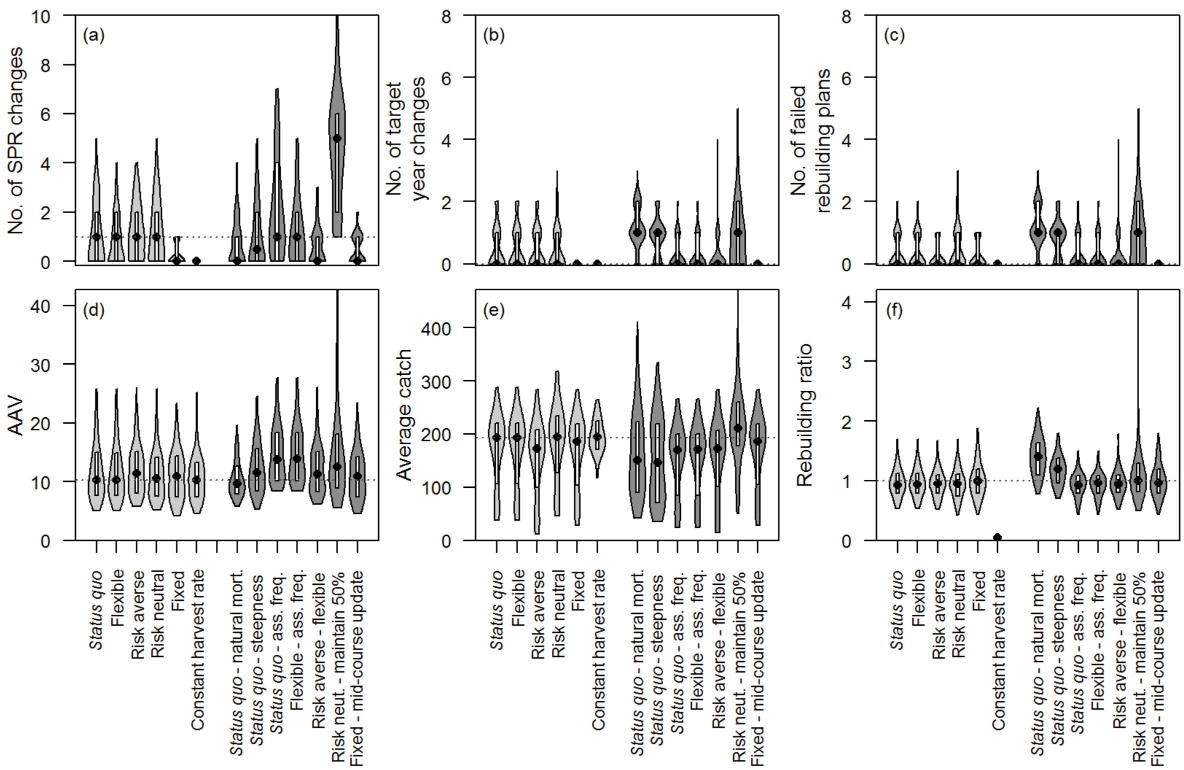

Figure 3.5: The roundfish life history median values (black points) and 80% simulation intervals (white bars) for the number of SPR changes (a), number of changes to the target rebuilding year (b), the number of failed rebuilding plans (c), the average annual variation in catch (AAV) (d), the average catch over the first five years of the rebuilding period (e), and the rebuilding ratio (f) for each of the alternative rebuilding strategies (light grey) and sensitivities (dark grey). The dashed horizontal line indicates the median value for the "*status quo*" strategy for visual reference for plots (a)–(e). The dashed horizontal line in (f) is set at 1.

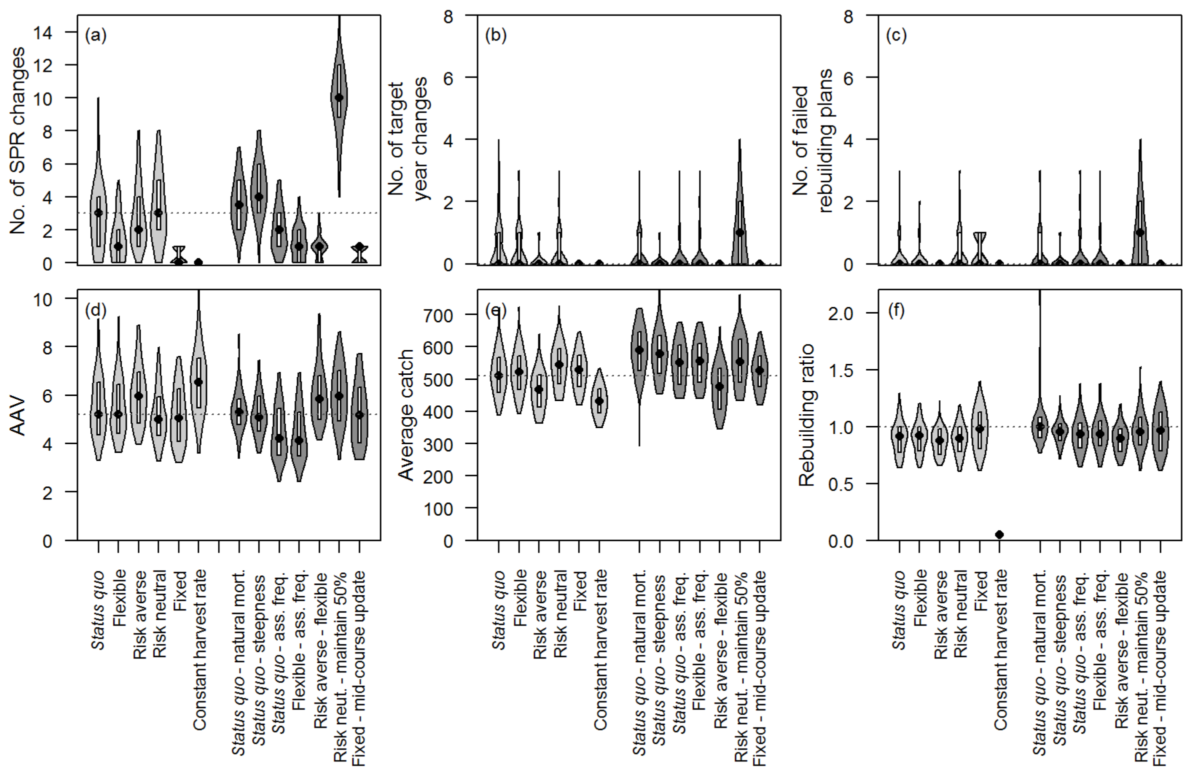

Figure 3.6: The medium-lived rockfish life history median values (black points) and 80% simulation intervals (white bars) for the number of SPR changes (a), number of changes to the target rebuilding year (b), the number of failed rebuilding plans (c), the average annual variation in catch (AAV) (d), the average catch over the first five years of the rebuilding period (e), and the rebuilding ratio (f) for each of the alternative rebuilding strategies (light grey) and sensitivities (dark grey). The dashed horizontal line indicates the median value for the "*status quo*" strategy for visual reference for plots (a)–(e). The dashed horizontal line in (f) is set at 1.

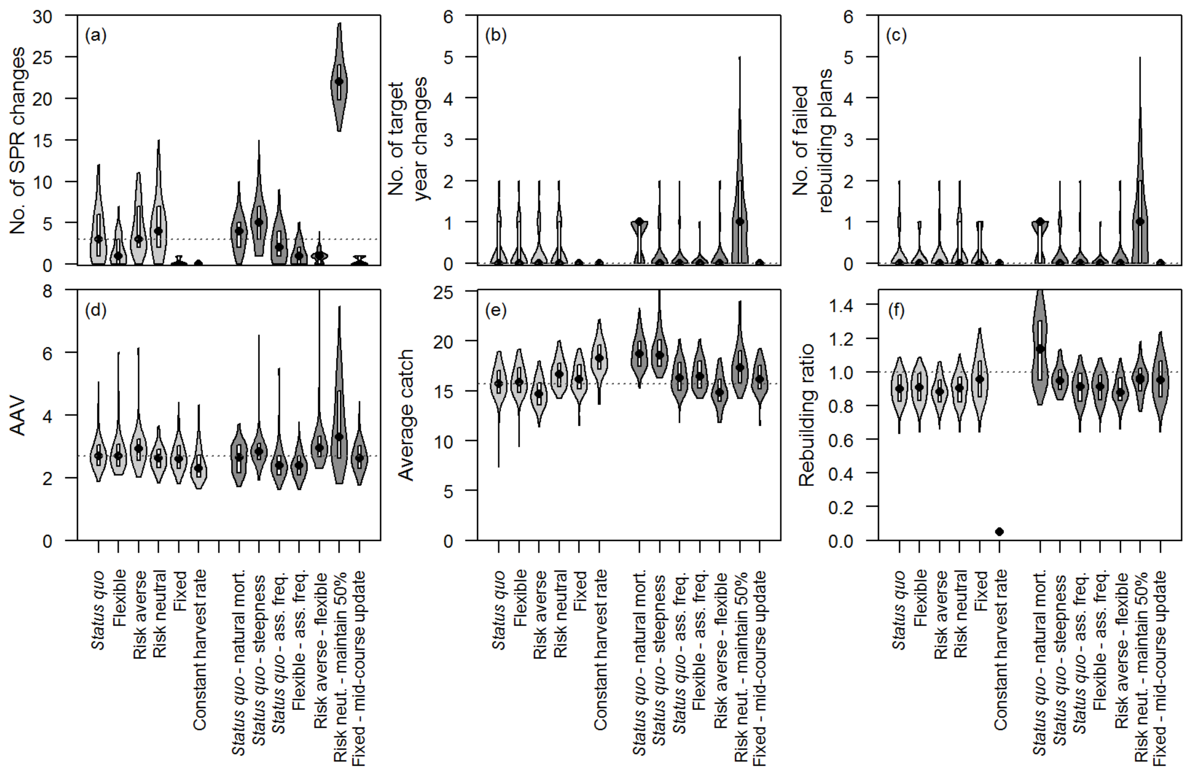

Figure 3.7: The long-lived rockfish life history median values (black points) and 80% simulation intervals (white bars) for the number of SPR changes (a), number of changes to the target rebuilding year (b), the number of failed rebuilding plans (c), the average annual variation in catch (AAV) (d), the average catch over the first five years of the rebuilding period (e), and the rebuilding ratio (f) for each of the alternative rebuilding strategies (light grey) and sensitivities (dark grey). The dashed horizontal line indicates the median value for the "*status quo*" strategy for visual reference for plots (a)–(e). The dashed horizontal line in (f) is set at 1.

Chapter 4

EVALUATING ALTERNATIVE HARVEST CONTROL RULES FOR US WEST COAST FLATFISH STOCKS

Abstract

US federal fisheries managers are mandated to obtain optimum yield while preventing overfishing. However, determining the maximum sustainable yield (MSY), identifying the relative biomass that produces MSY and the associated fishing rate required (F_{MSY}) is difficult. The Pacific Fishery Management Council, which manages groundfish stocks off the US west coast, has employed proxy targets in lieu of species-specific estimates of MSY, B_{MSY}, and F_{MSY}. The proxy targets are life history specific, with flatfish stocks managed using a target B_{MSY} of 0.25 of unfished biomass and a harvest control rule that applies an exploitation rate equal to a spawner-per-recruit harvest rate of $F_{0.30}$, with a linear reduction of catch to zero if the stock falls below 5% of unfished biomass (B_{THRES}), until a rebuilding plan is developed. A management strategy evaluation was performed to explore the performance of the current harvest control rule applied to flatfish stocks to meet management goals, along with alternative harvest control rules that explore varying the values for B_{PROXY}, B_{THRES}, and F_{SPR}. The harvest control rules maintained stocks at or near B_{PROXY} when stock-recruit steepness was 0.85 or greater, with very low probabilities of reducing relative biomass below a minimum stock size threshold. The most aggressive harvest control rule, which applied a B_{PROXY} of 0.20 and a target harvest rate of $F_{0.25}$, led to fishing rates that exceeded the operating model F_{MSY} values for low steepness (0.75), reducing the stock below B_{PROXY} with catches exceeding MSY. There was a tradeoff between the average catch obtained by each harvest control rule and the average annual variation in catch where higher estimated catches also led to an increase in annual variance.

4.1 Introduction

The concept of maximum sustainable yield (MSY) has a long history in fisheries management (Russell, 1931; Hjort et al., 1933; Graham, 1935). The theory behind sustainable yield within fishery science states that as stocks are fished down, a surplus yield, an amount beyond the replacement biomass, would be available for exploitation, and there exists a biomass that would produce MSY (B_{MSY}) and a fishing rate that results in MSY (F_{MSY}). However, estimating MSY and B_{MSY} can be difficult (Punt et al., 2002; Magnusson and Hilborn, 2007). To accurately determine these values one must have an understanding of the density-dependent behavior of the population, manifesting through a spawner-recruit relationship. However, recruitment data are notoriously noisy, making it difficult to determine the shape of this relationship with any confidence. Hence, there has been much debate on the viability of achieving MSY and its use in fisheries management (e.g. Larkin, 1977; Sissenwine, 1978).

The challenge of estimating B_{MSY}, MSY, and F_{MSY} with confidence has led to the development of alternative approaches to define targets to achieve optimum yield, including proxy target biomasses and harvest rates in the place of stock-specific estimates of B_{MSY} and F_{MSY}. Proxy values based on general life history attributes for a marine species can be applied to determine target stock sizes that produce yields that are close to MSY (Clark, 1991; Hilborn, 2010), avoiding the uncertainty and challenge of determining the stock-recruit relationship for each stock. The Pacific Fishery Management Council (PFMC) manages a range of groundfish life history types along the US west coast. It has employed the use of three types of proxy to manage groundfish stocks that are life history specific: 1) a B_{MSY} proxy defined in terms of biomass relative to the unfished level termed B_{PROXY}, 2) an F_{MSY} proxy harvest rate based on spawning biomass-per-recruit termed F_{SPR}, and 3) a minimum stock size threshold, MSST, which defines the relative biomass at which a stock would be declared overfished (Pacific Fishery Management Council, 2016).

The PFMC aims to maintain stocks at or near B_{PROXY}, and has adopted a threshold management strategy to achieve this aim. This strategy is in the form of a harvest control rule dictating that catch is a function of estimated biomass. Harvest control rules can be an effective management tool when they are defined appropriately based on the biology of the stock and management goals, providing explicit guidelines defining harvest based on stock biomass (Deroba and Bence, 2008; Punt et al., 2008, 2014b). The harvest control rule reduces the catch linearly when the relative stock biomass falls below B_{PROXY}, reducing catch to zero when the stock is at or below a lower threshold (B_{THRES}). However, a rebuilding plan is required which determines the catches until the stock recovers to the target biomass, if a stock is estimated below the management defined MSST.

Groundfish stocks along the US west coast are highly diverse in life history and productivity, and the PFMC has accounted for these differences in the proxy targets and harvest rates, along with the associated harvest control rule. Flatfish (e.g. petrale sole *Eopsetta jordani*, Dover sole *Microstomus pacificus*) which display a periodic life history type (Winemiller and Rose, 1992; Rose et al., 2001), are one of the more productive groundfish species groups on the US west coast. Historically, the PFMC applied the same B_{PROXY} and the harvest control rule for all stocks within the groundfish Fishery Management Plan (FMP) (although with life history specific proxy harvest rates, F_{SPR}) (Pacific Fishery Management Council, 2006). The Council amended the flatfish B_{PROXY}, F_{SPR}, and the harvest control rule in 2009 to account for the more productive nature of flatfish stocks relative to the other stocks within the groundfish FMP (e.g. rockfish: *Sebastes* spp., roundfish: sablefish *Anoplopoma fimbria*, lingcod *Ophiodon elongatus*) (Pacific Fishery Management Council, 2011). This changed the proxies from the previous targets of $0.40SB_0$ (40% of unfished spawning biomass) and a spawner-per-recruit harvest rate of $F_{0.40}$ to a B_{PROXY} of $0.25SB_0$, with a harvest rate of $F_{0.30}$. The harvest control rule was adjusted from the previous values of $0.40SB_0$ and $0.10SB_0$ (termed the "40-10" harvest control rule) to updated values that linearly reduce catch when the stock falls below $0.25SB_0$ to zero if the stock falls below $0.05SB_0$ (termed the "25-5" harvest control rule). Additionally, the MSST, the biomass that if a stock fell below would result in an overfished declaration requiring a rebuilding plan, was lowered to $0.125SB_0$ from the previous value of $0.25SB_0$.

The changes to the targets and harvest control rule were based on the theoretical deterministic relationship between stock size and density-dependent recruitment compensation (also known as steepness) for US west coast flatfish based on a meta-analysis, with the goal to manage towards a stock size that would produce maximum yield (Pacific Fishery Management Council, 2009). However, the relationship between spawning biomass and recruitment is highly complex and often unknown, which could result in unexpected and potentially undesirable behavior of the harvest control rule. This work uses management strategy evaluation (MSE) (Smith et al., 1999; Punt et al., 2014c) to evaluate the performance of the amended harvest control rule and the proxies for US west coast flatfish stocks. Other harvest control rules were explored to provide management and stakeholders with a suite of potential outcomes and trade-offs among approaches. The MSE performed here was generally parameterized based on petrale sole, a commercially important flatfish stock currently exploited off the US west coast. However, since the flatfish harvest control rule is applicable to all assessed flatfish stocks, a range of life history parameter combinations designed to encompass other flatfish species was explored. This MSE aims to address the ability of each harvest control rule to maintain stocks at or near the target level, the potential risks of each approach, and the trade-offs between potential catches and target stock sizes.

4.2 Materials and Methods

4.2.1 General approach

The MSE is based on an age- and sex-structured population dynamics model. In the simulations underlying the MSE, populations were sampled by two fisheries and a single survey, providing length- and age-composition data and an annual survey index of abundance, both with endogenous measures of uncertainty. The data were used to estimate stock size and hence future catch limits. Catch limits were adjusted using one of a suite of harvest control rules given the estimated relative stock biomass. The data generation, catch limit calculation, and updating of stock biomass was conducted in an iterative fashion for 75 years (termed the “management period”). This period was selected because it was long enough to eliminate the impact of the historical fishing pattern, such that results would be driven solely by the harvest control rule.

4.2.2 The operating model (OM)

The numbers-at-age at the start of the year are computed as:

$$N_{y+1,s,a} = \begin{cases} 0.5R_y & \text{if a} = 0 \\ N_{y,s,a-1}e^{-(M_s+\sum_f S_{f,s,a-1}F_{y,f})} & \text{if } 1 \leq \text{a} < \text{A-1} \\ N_{y,s,A-1}e^{-(M_s+\sum_f S_{f,s,A-1}F_{y,f})} + N_{y,s,A}e^{-(M_s+\sum_f S_{f,s,A}F_{y,f})} & \text{if a} = \text{A} \end{cases} \tag{4.1}$$

where $N_{y,s,a}$ is the number of fish of sex s and age a at the start of year y, R_y is the number of age-0 animals at the start of year y, $S_{f,s,a}$ is the selectivity by fishery, sex and age, A is the plus group, $F_{y,f}$ is the instantaneous fishing mortality rate at the peak selectivity for fishery f during year y, and M_s is the instantaneous rate of natural mortality for sex s.

The number of age-0 fish is related to spawning biomass according to the Beverton-Holt stock recruitment relationship with auto-correlated deviations:

$$R_y = \frac{4hR_0SB_y}{SB_0(1-h)+SB_y(5h-1)}e^{-0.5\sigma_R^2+\epsilon_y^R} \tag{4.2}$$

$$\epsilon = \rho_R\epsilon_{y-1} + \sqrt{1+\rho_R^2}\phi_y \quad \phi_y \sim N(0;\sigma_R^2) \tag{4.3}$$

where SB_0 is the unfished spawning biomass, SB_y is the spawning biomass at the start of the spawning season in year y, R_0 is the unfished equilibrium recruitment, σ_R is the standard deviation of recruitment in log space, h is the recruitment compensation (also known as steepness), ϕ_y is the recruitment deviation in year y, and ρ_R is the autocorrelation in recruitment.

A non-equilibrium starting condition was created by applying equations (4.1), (4.2), and

(4.3), prior to the start of fishing. The catch of fish of sex s and age a during year y in numbers was determined by:

$$C_{y,s,a} = \frac{S_{f,s,a}F_{y,f}}{M_s + \sum_f S_{f,s,a}F_{y,f}} N_{y,s,a}\left(1 - e^{-M_s - \sum_f S_{f,s,a}F_{y,f}}\right) \tag{4.4}$$

Historical catches for years 1- 50 were generated based on a fixed F pattern that linearly increased between years 1 and 24, and remained constant for years 25 to 50 for both fisheries (Fig. 4.1). The F pattern was selected to result in a variety of relative stock biomass levels in year 50 that ranged from highly depleted (well below the target biomass) to levels above the target biomass depending upon random recruitment deviations (ϵ_y in Eqn. 4.3). Fishery 1 was responsible for 75% of the historical fishing mortality, with fishery 2 accounting for the remaining 25%. The estimated catch during the management period was divided between the two fisheries by a 75-25 split, with catch removed without error from the operating model population.

Distributions by sex for natural mortality, the Brody growth coefficient (k), the parameters determining expected length at two specified ages were used to generate alternative replicates that would be representative of a spectrum of flatfish species off the US west coast (Table 4.1, Fig. 3.2). The distributions for the true B_{MSY}, F_{MSY}, and MSY resulting from the natural mortality and steepness values are given in Table 4.2. Each simulation drew a single value from each parameter distribution, along with a series of annual recruitment deviations to generate 200 unique population trajectories.

Each population was assessed for the first time at the start of year 50, and every 5th year thereafter. Assessments of US west coast groundfish are performed in two-year cycles. However, not every stock is assessed every two years. The frequency of assessment for flatfish stocks varies based upon commercial importance, which is an indicator of exploitation, the time since last assessment, and if there are data that raise concern for relative stock biomass (e.g. downward trend in the index of abundance) (Methot, 2015). The assessment frequency imposed in these simulations was selected to be a general representation of the average timing between assessments for a US west coast flatfish stock. Each assessment estimated relative stock biomass and catch limits based on the harvest control rule. The simulated future catches were equal to the catch limits. Annual fishery length- (Fishery 1: annual sample size (N) = 50, Fishery 2: N = 30) and age-composition data (Fishery 1: N = 25, Fishery 2: N = 15) with ageing error (10% error by age) were available from 25 years prior to the first assessment and annually thereafter, with catches known without error for all years. An annual survey index of abundance (CV = 0.20) with associated length- (N = 50) and age-composition data (N = 50) were available annually from 15 years prior to the first assessment and annually thereafter. The survey length, coefficient of variation, and composition sample sizes were designed to approximate the effective sample sizes for petrale sole. The operating model generated the annual index of abundance for each year y according to:

$$I_y = QB_y e^{-0.5\sigma_s^2 + \epsilon_y^2} \quad \epsilon_y^2 \sim N(0, \sigma_s^2) \tag{4.5}$$

where Q is the catchability coefficient for the survey and σ_s is the standard deviation of survey catchability in log space (see Table 4.4). The expected mid-year biomass index, B_y, is given by:

$$B_y = \sum_s \sum_{a=1}^{A} w_{s,a} S_{survey,s,a} N_{y,s,a} e^{-0.5(M_s + \sum_f S_{f,s,a} F_{y,f})} \tag{4.6}$$

where w_s is the weight of a fish of sex s at age a, and $S_{survey,s,a}$ is the selectivity for the survey for sex s and age a. The age-composition data for the fishery and survey catch were observed subject to ageing error and were assumed to be multinomially distributed.

4.2.3 The estimation method (EM)

Stock Synthesis (Methot and Wetzel, 2013), an integrated statistical catch-at-age model, was applied as the estimation method (EM) to assess the simulated stocks. Steepness was assumed to be known without error (simulations explored misspecification of steepness, see *Sensitivities*). R_0, annual recruitment deviations, growth, natural mortality and the selectivity parameters for the survey and each fishery were estimated. The relative stock biomass in the assessment year was estimated and the forecasted catches were projected according to a range of alternative harvest control rules (see *Alternative harvest control rules*) of the type used for US west coast groundfish (Fig. 4.3). The overfishing level (OFL) was set equal to the F_{SPR} multiplied by SB_y (Fig. 4.3). The OFL was reduced by a commonly used management buffer value that accounts for the uncertainty about current biomass for well-assessed stocks (0.956; Ralston et al., 2011) to determine the acceptable biological catch (ABC), i.e. ABC = 0.956 OFL. The PFMC then determines the annual catch limit (ACL), according to the harvest control rule.

4.2.4 Alternative harvest control rules

Harvest control rules were defined based upon three components; the B_{PROXY}, the lower threshold where the annual catch limit (the ACL) is set to zero if the stock falls below (B_{THRES}), and the F_{SPR} (the F_{MSY} proxy) used to determine harvest designed to maintain the stock at or near B_{PROXY}. Four harvest control rules were examined that applied alternative values for B_{PROXY}, B_{THRES}, and F_{SPR}. These four rules were (1) the current control rule applied by the PFMC for flatfish ($B_{\mathrm{PROXY}} = 0.25$, $B_{\mathrm{THRES}} = 0.05$, $F_{0.30}$), and three alternative rules consisting of (2) low biomass target and limit, with higher maximum fishing rate (0.20, 0.05, $F_{0.25}$), (3) larger biomass target and limit, with lower maximum fishing rate (0.30, 0.10, $F_{0.35}$), and (4) the previous harvest control rule and biomass target and limit which involves a higher biomass target and limit with lower maximum fishing rate (0.40, 0.10, $F_{0.40}$). The development of the B_{PROXY} and F_{SPR} values are based on the Beverton-Holt stock recruitment relationship where the SPR value that would maintain a deterministic population at B_{PROXY} dependent upon steepness (h):

$$SPR = \frac{1 - h + (5h - 1)B_{PROXY}}{4h} \tag{4.7}$$

Each of the alternative harvest control rules represents an alternative target stock size that ranges from more precautionary to more aggressive relative to the current harvest control rule.

The US west coast flatfish harvest control rule was updated in 2009. The Scientific and Statistical Committee of the PFMC provided guidance on determining a B_{PROXY} that would achieve approximate maximum yield and the associated F_{SPR} (Pacific Fishery Management Council, 2009). It recommended $B_{\mathrm{PROXY}} = 0.25$ and $F_{\mathrm{SPR}} = F_{0.30}$ assuming a steepness value of 0.80 based on the pleuronectid flatfish steepness value in Myers et al. (1999) meta-analysis (which is centered about 0.80) and the range of steepness values from historical assessments of flatfish stocks off the US west coast. While a steepness of 0.80 is not specifically considered here, the range of values evaluated with sensitivities (see *Sensitivities*) encompasses this value and Table 4.3 provides the theoretical SPR value associated with each alternative harvest control rule based on the ranges of steepness values explored in the base scenario and sensitivities.

The harvest control rules determined the ACL based on the B_{PROXY}, B_{THRES}, F_{SPR}, and the estimated relative stock size:

$$\mathrm{ACL}_y = \begin{cases} 0 & \text{if } SB_y/SB_0 < B_{\mathrm{THRES}} \\ 0.956C(F_{\mathrm{SPR}})_y \frac{B_{\mathrm{PROXY}}(SB_y/SB_0 - B_{\mathrm{THRES}})}{(B_{\mathrm{PROXY}} - B_{\mathrm{THRES}})(SB_y/SB_0)} & \text{if } B_{\mathrm{THRES}} \leq SB_y/SB_0 < B_{\mathrm{PROXY}} \\ 0.956C(F_{\mathrm{SPR}})_y & \text{if } SB_y/SB_0 \geq B_{\mathrm{THRES}} \end{cases} \tag{4.8}$$

where the $C(F_{\mathrm{SPR}})_y$ is the OFL in y. The forecasted catches (ACLs) were projected given the harvest control rule, the F_{SPR}, the estimated relative biomass in the assessment year, and each subsequent forecast year based on the population dynamics of the EM assuming full removal of the previous year's forecasted catch.

One major difference in the applications of the alternative strategies applied here and actual management practice on the US west coast was the omission of rebuilding plans that are implemented when a stock is assessed to have fallen below the MSST. Harvest for stocks below the MSST would no longer be determined based upon the standard harvest control rule, but rather a rebuilding plan would be used to determine catches until the stock is estimated to be rebuilt to the target stock size (i.e. B_{PROXY}). However, this work aims to examine the performance of alternative harvest control rules, hence a rebuilding plan was not imposed if a simulated stock was assessed to have fallen below the MSST and the catches were instead determined based on the associated harvest control rule for all estimated stock sizes.

4.2.5 Sensitivities

Sensitivities were conducted to determine the impact of alternative specifications of the operating and estimation models on the performance of the harvest control rules. The results for the sensitivities were compared against those from the base scenario (termed $h_{\mathrm{COR}\ 85}$).

Estimated catch based on the harvest control rules depends on the value assumed for steepness, where stocks with higher steepness values can sustain greater fishing mortality and allow populations to increase from lower stock sizes at higher rates relative to lower steepness values. The first suite of sensitivities altered the assumed steepness (productivity), where two alternative values (termed $h_{\mathrm{COR}\ 75}$ and $h_{\mathrm{COR}\ 95}$) were known without error in the OM and EM which bookended the assumed base value of 0.85 (Table 4.4). Assuming alternative values for steepness will evaluate whether a harvest control rule with the associated F_{SPR} would meet management goals of maintaining a population at or near the B_{PROXY} given the productivity of the stock.

The second set of sensitivities examined the performance of each harvest control rule when the value of steepness was assumed to be higher (0.95) or lower (0.75) than the true OM value (0.85) (Table 4.4, termed h_{LO} and h_{HI}). The estimates of steepness within the petrale sole assessment over the last three assessments have varied from 0.85 – 0.90 (Haltuch et al., 2011; Haltuch and Ono, 2013; Stawitz et al., 2015). The levels of steepness applied here were selected to be representative of the uncertainty associated with petrale sole, but also to be representative of other US west coast flatfish stocks (Kaplan and Helser, 2007; Stewart, 2007; Hicks and Wetzel, 2011).

The impact of data quantity and quality on the performance of each harvest control rule was examined. The sample sizes for the length and age data provided to the EM were reduced by 70% from the base scenario data levels (annual lengths: fishery 1, N = 15; fishery 2, N = 9; survey, N = 15; annual ages: fishery 1, N = 7; fishery 2, N = 4; survey, N = 15) and the uncertainty about the annual index of abundance was increased (from $CV = 0.20$ to $CV = 0.35$) (Table 4.4, termed Reduced Data [RD]).

The consequences of the extent of variability in recruitment and autocorrelation in recruitment was also examined. The first sensitivity related to recruitment variation evaluated the performance of the harvest control rules when the OM had higher variability among annual recruitments (σ_R= 0.60) and was known to the EM (Table 4.4, termed σ_R). The second recruitment sensitivity examined the impact of autocorrelation among annual recruitments within the OM ($\rho_R = 0.707$), but when recruitment autocorrelation was not accounted for by the EM (i.e. autocorrelation assumed to be 0) (Table 4.4, termed ρ_R). The last recruitment sensitivity combined the previous two sensitivities, i.e. there was higher variation in recruitment within the OM which was known by the EM, but the OM also included autocorrelation among the annual recruitment deviations, but this was not accounted for by the EM (Table 4.4, termed σ_R & ρ_R).

4.2.6 Performance Measures

The outcomes of the simulations for each harvest control rule were summarized using the following five metrics:

1. The probability (over simulations and years) that the spawning biomass was below the minimum stock size threshold level (MSST = $0.5B_{\mathrm{PROXY}}$ for each harvest control rule) during the last 25 years of the management period for each simulation;

2. The probability (over simulations) of being within 10% of the B_{PROXY} during the last year of the management period (abbreviation "P ± 0.10 B_{PROXY}");

3. The average (over simulations) of the total catch over the last 25 years of the management period;

4. The average (over simulations) of the relative biomass over the last 25 years of the management period;

5. The annual average variability of the catches (abbreviation AAV) over the last 25 years of the management period defined as:

$$AAV = 100 \frac{\sum_y |C_y - C_{y+1}|}{\sum_y C_y} \tag{4.9}$$

where C_y is the catch during year y.

The performance metrics were limited to the last 25 years of the management period (years 101-125) to eliminate the impact of the historical fishing pattern. Documenting effects over 25 years also allows for inference about the long-term performance of the harvest control rules.

4.3 Results

4.3.1 Harvest control rule behaviors

Each harvest control rule examined resulted in the median relative biomass of the OM populations at or near the B_{PROXY} for all scenarios (Fig. 4.4). All harvest control rules had very low probabilities of the stock falling below the MSST over the last 25 years of the management period (Table 4.5). The scenario that assumed the lowest steepness value ($h_{\text{COR 75}}$) resulted in the lowest median relative biomass levels across all harvest control rules relative to the scenarios with higher steepness values ($h_{\text{COR 85}}$ and $h_{\text{COR 95}}$) (Fig. 4.4a-c). This behavior arises from the relationship between B_{PROXY}, steepness, and F_{SPR} within each harvest control rule, which were determined based on an assumed higher flatfish steepness of 0.80 (Pacific Fishery Management Council, 2009). A minority ($< 50\%$) of the relative biomasses for the $h_{\mathrm{COR}\ 75}$ scenario were above B_{PROXY} during the final 25 years of the management period for the 20-5, 25-5, and 40-10 harvest control rules. However, the median relative biomass was within 10% of B_{PROXY} for 25-40% of the simulated populations (Fig. 4.5a [grey points]).

All harvest control rules where steepness was at least 0.85 within the OM and EM (except the 40-10 for $h_{\mathrm{COR}\ 85}$) resulted in $> 50\%$ (often much greater for $h_{\mathrm{COR}\ 95}$) of the relative biomasses being above B_{PROXY} during the last 25 years of the management period, with median relative biomass above target levels (Fig. 4.4a and 4.4c). Additionally, there was little to no probability of the stock falling below the MSST for each harvest control rule (Table 4.5).

Misspecifying steepness in the EM (h_{LO} and h_{HI}) generally resulted in stocks at or near the target relative biomasses for all alternative harvest control rules (Fig. 4.4d and 4.4e). The EM reconciled the misspecification of steepness by adjusting other estimated parameters. Estimates of SB_0, R_0, and natural mortality were all negatively biased relative to the operating model values (results not shown) when steepness was assumed to be lower than the true value. Conversely, the estimation model resulted in positively biased estimates of SB_0, R_0, and natural mortality when steepness was misspecified high (results not shown). This compensation within the EM to the misspecification of steepness was driven by data (e.g. index of abundance and composition data) that was informative about relative biomass.

4.3.2 Trade-offs among alternative harvest control rules

The percentage of stocks with relative biomasses within 10% of the B_{PROXY} over the last 25 years of the management period generally increased as the harvest control rule target increased (Fig. 4.5a). The 40-10 harvest control rule had the highest percentage of stocks within 10% of B_{PROXY} across all the scenarios, but this harvest control rule also had the widest range for "successful" relative stock sizes relative to the others examined (e.g. $0.9B_{\mathrm{PROXY}} <$ relative biomass $< 1.10B_{\mathrm{PROXY}}$; 40-10: 36% - 44% vs. 20-5: 18% - 22%). The misspecification of steepness in the EM did not result in a differing percentage of stocks within 10% of B_{PROXY} relative to the scenarios where steepness was correctly specified (Fig. 4.5a vs. 4.5d).

While the harvest control rules with higher B_{PROXY} led to a greater percentage of stocks within 10% of B_{PROXY}, the median average catch over the last 25 years of the management period generally decreased as B_{PROXY} increased (Fig. 4.5a and 4.5d). The one notable exception was the scenario where the OM assumed the low steepness ($h_{\mathrm{COR}\ 75}$). The median average catch for the 20-5 harvest control rule was lower than the 25-5 or the 30-10 options (Fig. 4.5a [grey points]). The 20-5 harvest control rule with an SPR of 0.25 was overly aggressive given the productivity of the stock when steepness was 0.75 (e.g. stock size would be depleted below the level that would achieve maximum yield) (Tables 4.2 and 4.3).

Across scenarios, the median AAV generally increased as the median average catch obtained by each harvest control rule increased (Fig. 4.5c and 4.5f). The relationship between AAV and the probability of the stock being within 10% of B_{PROXY} was consistent across assumed steepness values where higher steepness (assumed correctly or misspecified) resulted in smaller values for each metric compared to lower steepness values (Fig. 4.5b and 4.5e).

4.3.3 The impact of recruitment variability

Increased annual recruitment variation (sensitivity σ_R) reduced the probability of the stock being within 10% of B_{PROXY} compared to base scenario for all harvest control rules ($h_{\mathrm{COR}\ 85}$) (Fig. 4.5a [black points] vs. 4.5g [black points]). The median average catch did not differ greatly from the base scenario, although the AAV increased as variation in recruitment increased (Fig. 4.5c [black points] vs. 4.5i [black points]). Recruitment autocorrelation within the OM also increased the range of relative biomasses during the last 25 years of the management period (Fig. 4.4h). Autocorrelation in recruitment resulted in extended periods of time when recruitment was either below or above its expected value. Such patterns could impact the performance of harvest control rules that do not account for this behavior when estimating forecasted catches. However, each of the harvest control rules examined here resulted in median relative biomasses near B_{PROXY} (Fig. 4.4h). Nevertheless, compared to the scenarios without recruitment autocorrelation, the probabilities of the stock being within 10% of B_{PROXY} were lower and median AAVs higher for all harvest control rules (Fig. 4.5b [black points] vs. 4.5h [grey points]). The increased variability in relative biomasses also increased the probability that the stock would be below the MSST relative to the other sensitivities, but it still remained low (Table 4.5).

Increased variability in recruitment and autocorrelation (σ_R & ρ_R) resulted in the highest probabilities that the stock would be below MSST, although these remained low (6-13%) (Table 4.5). The distributions for the last 25 years of the management period for this case had the largest variation about the relative biomass compared to all other sensitivities (Fig. 4.4i), with larger AAVs, and low probabilities of being within 10% of B_{PROXY} (Fig. 4.5h). However, the median relative biomass for each harvest control rule was at or near B_{PROXY} (Fig. 4.4i).

4.3.4 Data quantity and quality

Reducing data quantity and quality did not greatly alter the performance of the harvest control rules compared to the base scenario. The median relative biomass of the OM populations and the percentage of simulations with biomass greater than B_{PROXY} during the last 25 years of the management period were marginally greater than the values for the base scenario for each harvest control rule when data quantity and quality was reduced (Fig. 4.4a vs. 4.4f). The estimated relative biomass was negatively biased relative to the OM relative biomass (not shown), resulting in a reduced median average catch and the probability of being within 10% of the B_{PROXY} relative to the base scenario (Fig. 4.5a vs. 4.5d). In addition, the median AAV increased when data quantity and quality was reduced (Fig. 4.5b vs 4.5d).

4.4 Discussion

The harvest control rules evaluated performed well in maintaining relative biomasses at or near B_{PROXY}, generally estimating catch at or lower than the OM median MSY value. The more aggressive the harvest control rule (i.e. lower B_{PROXY}), the higher the average catches (except the $h_{\mathrm{COR}\ 75}$ sensitivity). However, higher average catches coincided with increased variance among annual catches estimated by the assessment. The 20-5 harvest control rule was the only one that led to median catches that exceeded the theoretical MSY of the OM population when steepness was assumed to be the lowest value examined here (Table 4.2). Historically, US west coast flatfish stocks have been assumed to be highly productive (Stawitz et al., 2015; Hicks and Wetzel, 2011; Stewart, 2007). However, steepness can be notoriously difficult to estimate (Lee et al., 2012) and is a key parameter for determining stock productivity. One metric for defining a successful harvest strategy is whether it adequately balances the need for high average yield and maintaining an average relative biomass close to B_{MSY}/B_0. The median B_{MSY}/B_0 values in the OM ranged between 0.16–0.26 based on the assumed steepness (Table 4.2). The 25-5 and each subsequent harvest control rule with higher B_{PROXY} and F_{SPR} values (e.g. 30-10 and 40-10) maintained all OM stocks close to the true management targets. However, if steepness was below the lowest value examined here (0.75), the stock would experience exploitation rates exceeding F_{MSY} and the stock would be depleted below the target. This should be considered when selecting a harvest control rule that would be applied across all US west coast flatfish stocks.

The harvest control rules examined here resulted in low probabilities of the stock falling below the MSST. The linear reduction applied in each harvest control rule when biomass was estimated below the B_{PROXY} was one factor that led to this. The reduction in the ACL when the stock was estimated to be below the B_{PROXY} was sufficient to allow the stock to either stabilize or increase in biomass. However, assessment frequency and reduced lag in adjusting harvest based on assessed stock biomass is key to maintaining stocks close to the target stock size (Sylvia, 2015). The EM was applied every 5$^{\mathrm{th}}$ year which allowed for timely adjustment to catch if stock biomass had declined below the B_{PROXY} relative to the previous assessment. Applying forecasted catches for longer periods may result in serious declines in biomass if the population deviates from the expected behavior based on either estimation error in the previous assessment or long-term below average recruitment (Johnson et al., 2016).

Increased recruitment variability and or recruitment autocorrelation resulted in increased variance about relative biomass during the last 25 years of the management period, although the probability of falling below the MSST still remained low. Once again, the frequency of assessments most likely avoided stocks from declining greatly when recruitment was below average. Extended periods of below average recruitment would negatively impact the performance of the harvest control rule if the stocks were monitored less frequently and adjustments to the catch not made. Identifying drivers (e.g. climate), estimating their relationship with recruitment success, and incorporating this information into stock assessment has been difficult (Haltuch and Punt, 2011), and accounting for autocorrelation in recruitment in harvest control rules has not necessarily resulted in improved estimates of future catches to meet management goals (A'mar et al., 2009; Punt, 2011; Punt et al., 2014c). Additional data sources and an improved understanding of the complex relationship between environmental drivers and population dynamics are required to better address this issue. However, new approaches for estimating autocorrelation in recruitment have shown promise in improving forecasting of catches when the extent of autocorrelation can be estimated and accounted for correctly (Johnson et al., 2016). Managers may want to consider additional precautions within the harvest control rule and subsequent harvest specifications if assessment frequency is low and there is a reasonable belief that the population conditions have changed since the last assessment.

Accounting for uncertainty in estimates of biomass and applying precaution when setting harvest limits can improve the performance of harvest control rules to maintain stocks close to targets. All estimated harvests in this evaluation were reduced to account for uncertainty about assessment year biomass according to PFMC guidelines (Ralston et al., 2011), which has been shown to be effective to prevent stocks from becoming depleted below the limit reference point (Fulton et al., 2016).

The reduction of data can also lead to additional uncertainties that can impact the performance of assessment methods (Magnusson and Hilborn, 2007; Wetzel and Punt, 2011b) and harvest control rules which managers may want to consider. The sensitivity that explored data quantity and quality led to slight differences in performance that highlighted some of the potential impacts of data loss. There was a lower median catch estimate by the EM (Fig. 4.5d), leading to a higher percentage of OM stocks above B_{PROXY} (Fig. 4.4f) due to increased uncertainty about relative biomass. The reduced data led to negatively biased estimates of the relative biomass compared to the OM, leading to lower estimates of catch and hence less depleted OM stocks. However, the OM and EM made similar structural assumptions which limited the impact of reduced data on estimation and harvest control rule performance. Real-world stock assessments rarely have the advantage of knowing the correct formulations for the population dynamics (e.g. models for growth, mortality, recruitment) of an exploited stock and data are often not perfectly informative given process and observation error. Additionally, simulated data are generally more well behaved than real fishery data, a property that impacts the results for all the scenarios explored. The fishery and survey data were generally well behaved for all scenarios, and did not have properties that may have led to transient or systematic bias which can be encountered in real data. Simulations are a useful tool to understand the potential implication of alternative actions but often underestimate uncertainty due to the inability to capture the true complexity of real systems and data.

The MSE highlights some of the potential trade-offs between the alternative harvest control rules. The productivity of flatfish stocks could support the more aggressive harvest control rules examined here without significant loss of yield due to over-exploitation, assuming that the stock-recruit relationship was correctly specified, in a single species context. However, additional factors should be considered by managers when selecting a harvest strategy and the associated trade-offs (Little et al., 2016). The majority of US west coast commercial fisheries are considered mixed-stock, harvesting multiple species in tandem (Pacific Fishery Management Council, 2016). The co-occurrence of species has often been taken into consideration when setting harvest limits where attaining the full catch from a highly productive abundant stock may not be practical given the potential impacts of harvest on restricted or less productive stocks. Hence, managers may opt for a harvest strategy that balances the desire to minimize risk while maximizing yield.

This work focused on the performance of alternative harvest control rules for US west coast flatfish stocks in a single stock context with limited exploration of alternative assumptions about productivity and recruitment. The B_{PRXOY} and the associated F_{SPR} values applied in each harvest control rule assumed the correct stock recruitment relationship as applied in the OM and EM. Explorations of alternative relationships between stock size and recruitment should be explored to identify the robustness of each strategy. Additional work that incorporates more complex population dynamics that vary over time in a single stock context or the application of a multi-species OM with species interactions may offer deeper insight about the potential performance of each of the harvest control rule.

Acknowledgments

This paper benefited greatly from comments by Ian Stewart (International Pacific Halibut Commission), Allan Hicks (Internation Pacific Halibut Commission), Ian Taylor (Northwest Fisheries Science Center), and an anonymous reviewer. Additionally, feedback from Brad Pettinger (Oregon Trawl Commission) and Owen Hamel (Northwest Fisheries Science Center) was instrumental in the development and analysis of this work.

4.5 Tables

Table 4.1: The biological parameter values for the base simulations

Parameter	Female	Male	Standard Deviation	Distribution
Steepness (h)	0.85		-	-
Natural mortality (M, year^{-1})	0.15	0.16	0.15	lognormal
Minimum length (cm) at age 2	16	16	2.5	normal
Maximum length (cm), L_∞, at age 40	54	43	2.5	normal
Growth coefficient (year^{-1}), k	0.134	0.202	0.15	lognromal
Body weight (kg), $w_l = \alpha L^\beta$				
Growth coefficient, α	2.08e^{-6}	3.05e^{-6}	-	-
Growth exponent, β	3.47	3.36	-	-
Maturity slope (year^{-1})	-0.743	-	-	-
Length at 50% maturity (cm)	33	-	-	-
Recruitment variation, σ_R	0.40		-	-
Autocorrelation, ρ_R	0.0		-	-
Catchability coefficient, Q	3		-	-
Survey CV, σ_s	0.20		-	-

Table 4.2: The operating model median and 90% simulation interval for true values of B_{MSY}/B_0 and SPR corresponding to MSY for each steepness value.

	B_{MSY}/B_0		SPR	
Steepness	Median	Interval	Median	Interval
0.75	0.26	(0.25-0.29)	0.32	(0.31-0.35)
0.85	0.22	(0.20-0.25)	0.25	(0.24-0.28)
0.95	0.16	(0.13-0.18)	0.17	(0.14-0.19)

Table 4.3: Theoretical SPR value that corresponds to the B_{PROXY} given the range of steepness values evaluated.

	B_{PROXY}			
Steepness	0.20	0.25	0.30	0.40
0.75	0.27	0.31	0.35	0.45
0.85	0.24	0.28	0.33	0.43
0.95	0.21	0.26	0.31	0.41

Table 4.4: Sensitivities performed across each of the alternative harvest control rule options. If not specified, the base operating model parameters were assumed (Table 4.1).

Sensitivities	OM	EM
Steepness correctly specified ($h_{\text{COR }75}$, base scenario [$h_{\text{COR }85}$], $h_{\text{COR }95}$)	0.75, 0.85, 0.95	
Steepness misspecified (h_{LO}, h_{HI})	0.85	0.75, 0.95
Reduced data (RD)	Fishery and survey length and age-composition sample sizes reduced by 70% and survey CV increased to 0.35.	
Higher variation in recruitment deviations (σ_R)	0.60	
Autocorrelation in recruitment deviations (ρ_R)	0.707	0
Higher variation and autocorrelation in recruitment deviations (σ_R & ρ_R)	$\sigma_R = 0.60$ $\rho_R = 0.707$	$\sigma_R = 0.60$ $\rho_R = 0.0$

Table 4.5: The median probability and 90% simulation interval that the stock represented in the operating model fell below the minimum stock size threshold (20-5: 10% relative biomass, 25-5: 12.5% relative biomass, 30-10: 15% relative biomass, 40-10: 20% relative biomass) over the last 25 years of the management period for each harvest control rule and each scenario.

	20-5		25-5		30-10		40-10	
Scenario	Med.	Interval	Med.	Interval	Med.	Interval	Med.	Interval
$h_{\text{COR 85}}$	0.01	(0.00-0.01)	0.00	(0.00-0.00)	0.00	(0.00-0.00)	0.00	(0.00-0.00)
$h_{\text{COR 75}}$	0.02	(0.00-0.03)	0.01	(0.00-0.01)	0.00	(0.00-0.00)	0.00	(0.00-0.00)
$h_{\text{COR 95}}$	0.00	(0.00-0.00)	0.00	(0.00-0.00)	0.00	(0.00-0.00)	0.00	(0.00-0.00)
h_{LO}	0.01	(0.00-0.02)	0.01	(0.00-0.01)	0.00	(0.00-0.00)	0.00	(0.00-0.01)
h_{HI}	0.00	(0.00-0.01)	0.00	(0.00-0.00)	0.00	(0.00-0.00)	0.00	(0.00-0.00)
RD	0.01	(0.00-0.02)	0.01	(0.00-0.01)	0.00	(0.00-0.00)	0.00	(0.00-0.00)
σ_R	0.02	(0.01-0.04)	0.01	(0.00-0.02)	0.00	(0.00-0.01)	0.00	(0.00-0.00)
ρ_R	0.06	(0.04-0.08)	0.04	(0.02-0.06)	0.02	(0.01-0.03)	0.03	(0.01-0.04)
σ_R & ρ_R	0.13	(0.09-0.16)	0.10	(0.08-0.12)	0.06	(0.04-0.08)	0.09	(0.08-0.11)

4.6 Figures

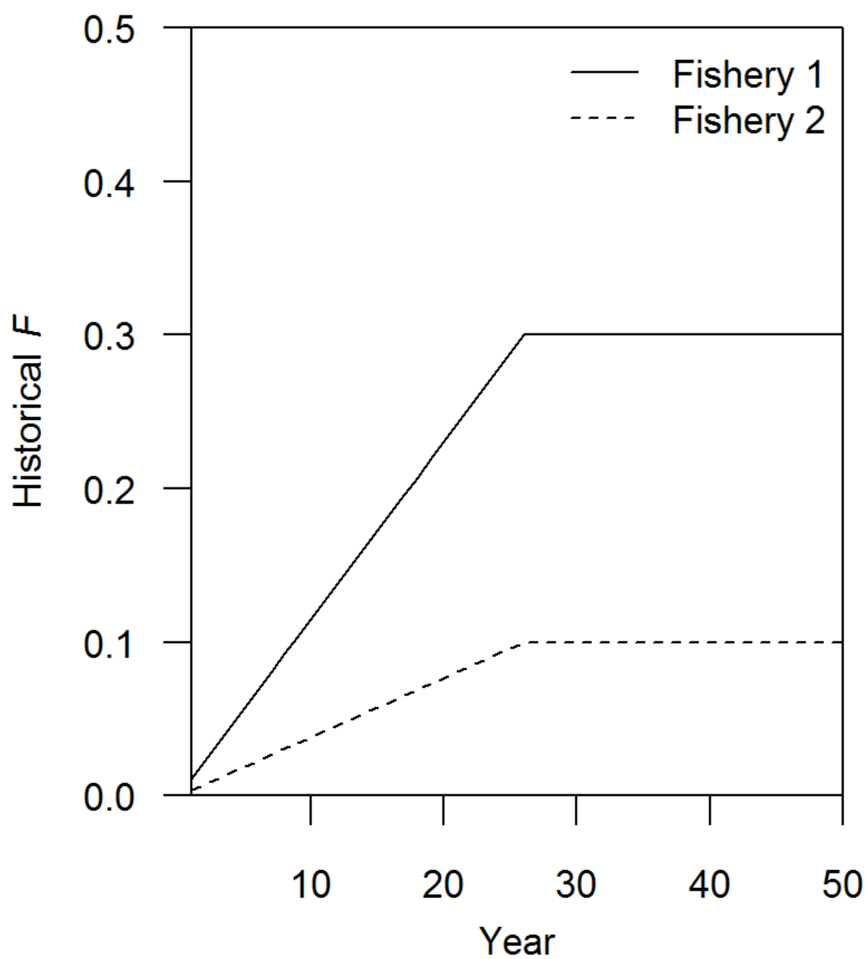

Figure 4.1: The historical F pattern for each fishery.

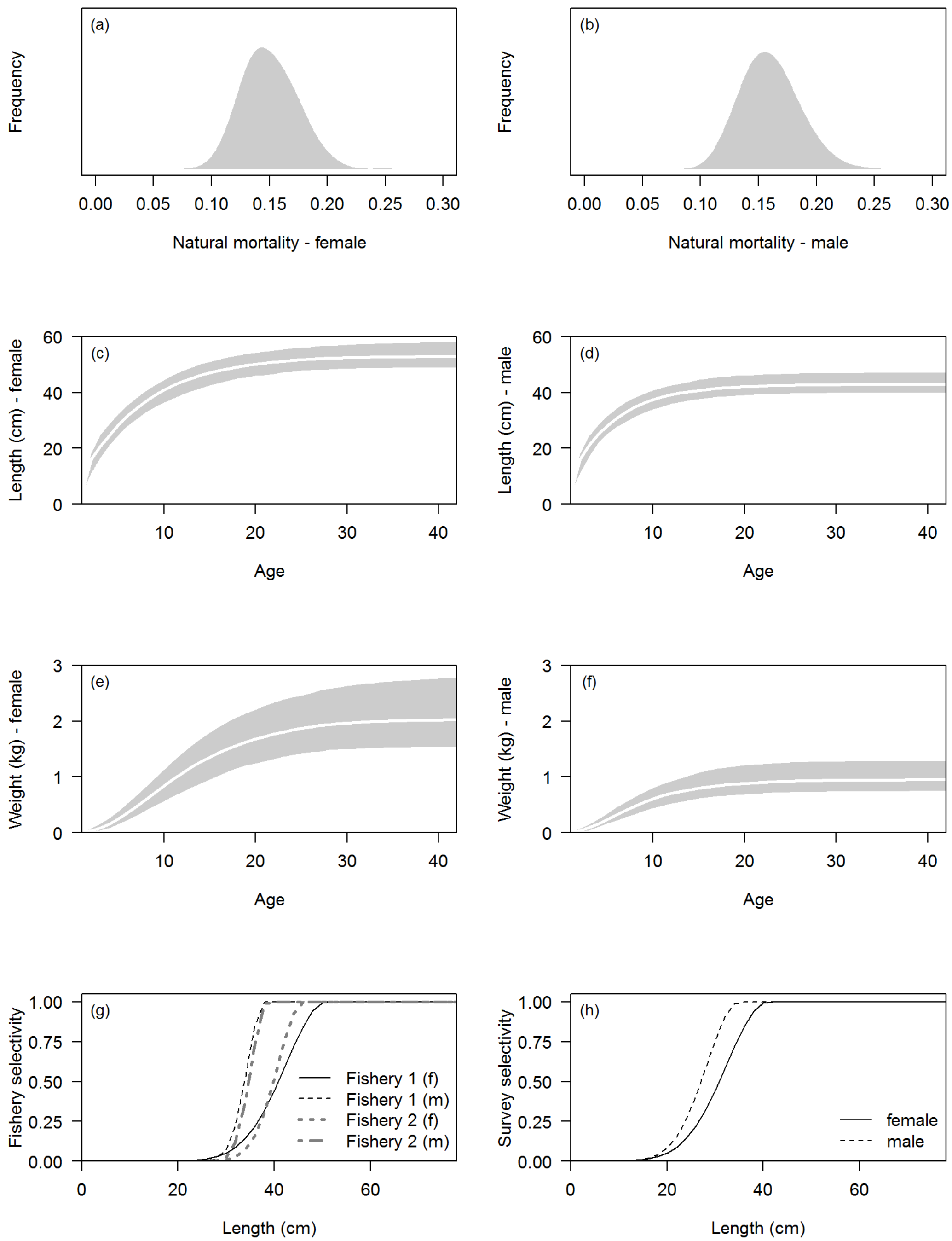

Figure 4.2: The biological and fishery parameters for the simulated realizations. Natural mortality by sex was drawn from a lognormal distribution. The 95% interval for length-at-age and weight-at-age are shown in grey and the white line indicates the median values.

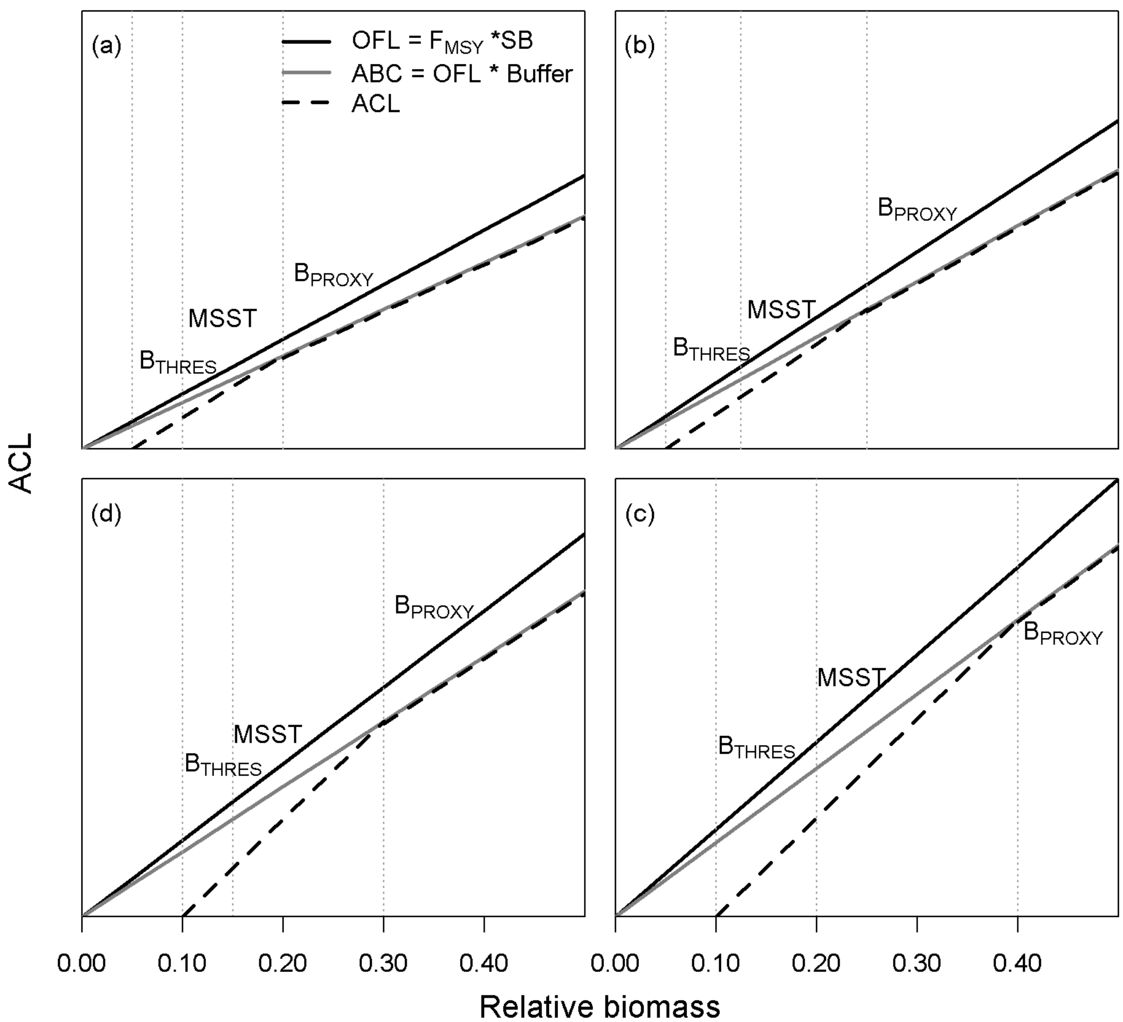

Figure 4.3: The alternative harvest control rules for determining the OFL (solid black line), the ABC (solid grey line), and the ACL (dashed black line) examined: a) 20-5: $B_{\text{PROXY}} = 0.20$, $B_{\text{THRES}} = 0.05$, and MSST = 0.10, b) 25-5: $B_{\text{PROXY}} = 0.25$, $B_{\text{THRES}} = 0.05$, and MSST = 0.125 c) 30-10: $B_{\text{PROXY}} = 0.30$, $B_{\text{THRES}} = 0.10$, and MSST = 0.15, and d) 40-10: $B_{\text{PROXY}} = 0.40$, $B_{\text{THRES}} = 0.10$, and MSST = 0.20.

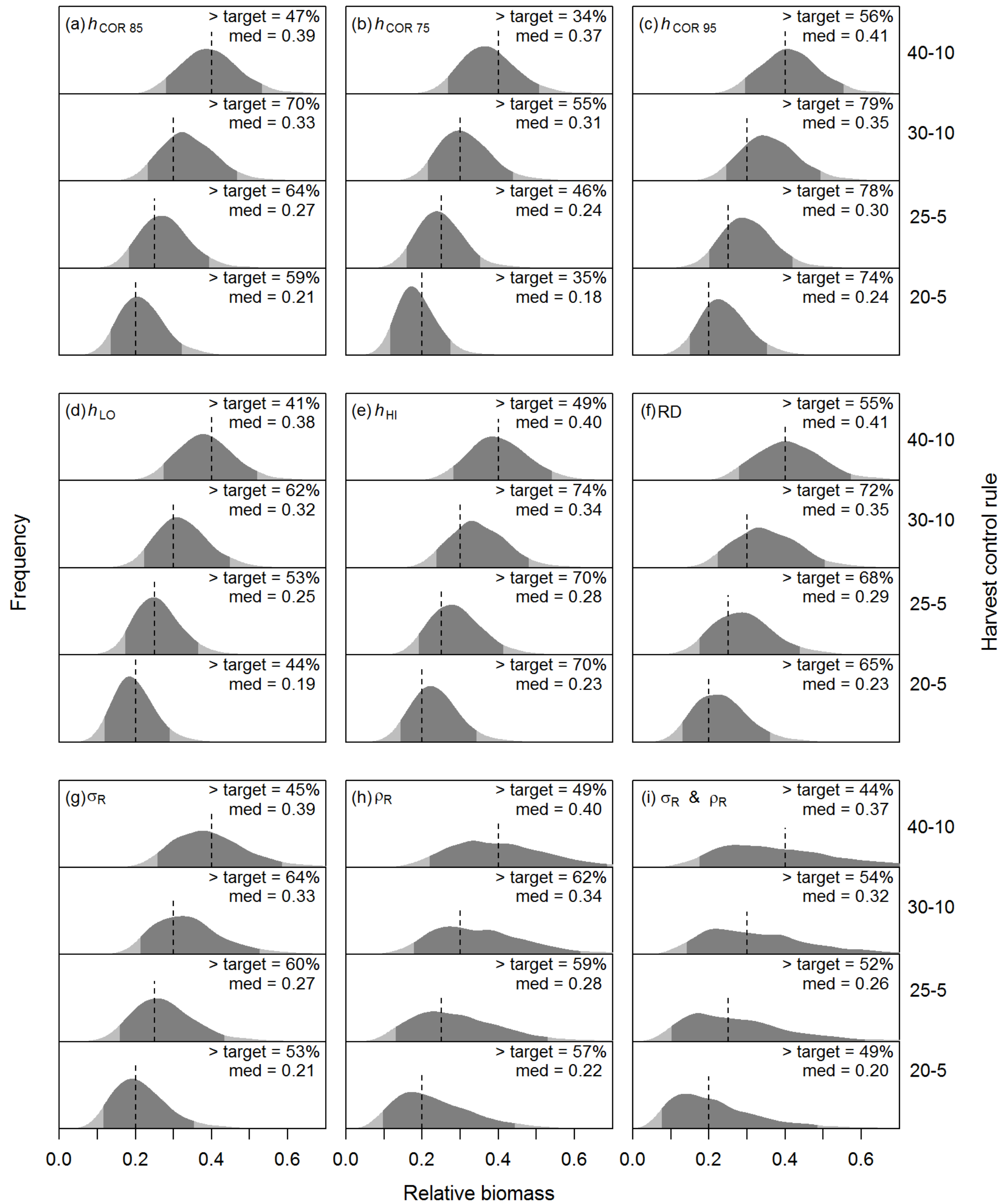

Figure 4.4: The distribution of relative biomass levels from the operating model during the last 25 years of the management period for the alternative harvest control rules for the base scenario and each sensitivity. The dark grey area indicates the 90% simulation interval about relative biomasses and the dashed black vertical line shows the harvest control rule B_{PROXY}.

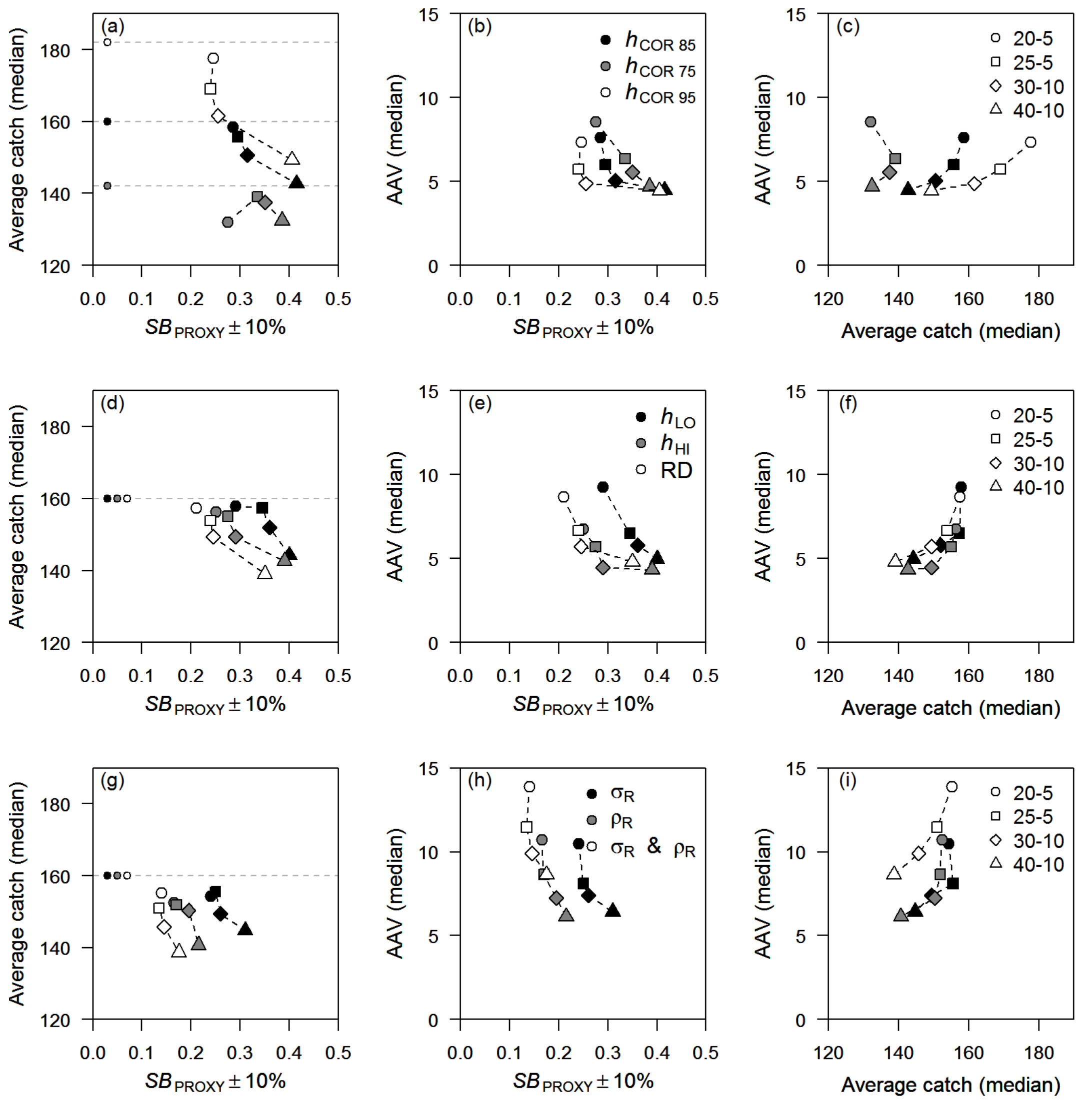

Figure 4.5: Trade-offs between the alternative harvest control rules and the average catch, probability that the stock is within 10% of B_{PROXY}, and annual average variation in catch (AAV). The median values over the last 25 years of the management period are plotted. Panels (a) – (c) show the trade-offs between the metrics for the alternative steepness scenarios ($h_{\text{COR 85}}$ [black], $h_{\text{COR 75}}$ [grey], and $h_{\text{COR 95}}$ [white]). Panels (d) – (f) show the trade-offs between the metrics for the steepness misspecification and increased recruitment deviation scenarios (h_{LO} [black], h_{HI} [grey], and reduced data [white]). Panels (g) – (i) show the trade-offs between the metrics for the autocorrelation in recruitment, recruitment deviation combined with autocorrelation, and reduced data scenarios (σ_R [black], ρ_R [grey], σ_R & ρ_R [white]). The dash grey horizontal line in (a), (c), and (g) provides the OM MSY with color circles corresponding to the associated scenario.

CONCLUSIONS

Each of the chapters evaluated key questions about data, modeling approaches, and evaluating the performance of management strategies for US west coast groundfish. The components of successful fisheries management begin the moment data are, or are not, collected, the model assumptions applied to represent a population, how the data are used to estimate parameters, and the management decisions made based on the assessment results. This dissertation examined each of these topics for specific issues concerning the management of US west coast groundfish.

Chapter 1 explored the impact of data on the ability to correctly monitor rebuilding of a stock. The performance of estimation methods to correctly estimate stock size and status depend on data quantity and quality. The loss of data due to harvest restrictions resulting from rebuilding efforts presents challenges for the ability to monitor rebuilding progress. Retaining data collection at historical levels allowed for improved parameter estimation, which resulted in reduced variability in estimated stock size with larger average catches during rebuilding. However, the median estimates of relative stock size were less than the true population biomass, resulting in a longer estimated time required to rebuild the stock to the target biomass. In contrast, when data were reduced during rebuilding, the estimates of relative stock size became more variable between assessments, resulting in stocks being declared rebuilt when the true population biomass was still below the target biomass. Declaring a stock rebuilt prematurely can have major implications for management. Catches are set based upon perceived stock size. An assessment that produces an overly optimistic view of stock size will set catches greater than the true population can support, resulting in overfishing, which may result in further reductions in stock size, potentially leading a rebuilt stock being subsequently declared overfished by a future assessment.

Producing appropriate harvest recommendations and accurately estimating the rate of recovery are driven by estimates of stock resilience (e.g. the productivity of a stock termed “steepness”). At the time of the first assessment, the operating model represented a one-way trip from unfished to low biomass, resulting in high variability in the estimates of steepness among simulations. Data scenarios that did not provide available data during rebuilding depended on the initial estimates of steepness, which were based on the generally uninformative historical data. As the operating model biomass began to increase from low stock sizes, continued data collection during the recovery period allowed for improved estimates of steepness, even with limited data collection. This work showed that a one-way trip scenario in stock size with limited data, may not be adequate to estimate steepness, but the inclusion of even limited data can, with contrast in stock size, improve the estimation of steepness and hence improve the estimate of stock biomass.

Continued data collection during rebuilding improved the ability to estimate key parameters, monitor rebuilding progress, and correctly identify when a stock had rebuilt to the target biomass. Historical data were not sufficient to achieve these key objectives in the absence of data during stock rebuilding. Beyond the specific questions examined in this chapter, data collection is crucial to identify misspecification within an estimation model. Model misspecification is an issue of great concern because incorrectly identifying key biological relationships or the interaction of the fishery on the stock can have major implications for the ability to correctly estimate biomass and hence the appropriate harvest. Ongoing data collection with reasonable sample sizes is key to the ability to identify and address the potential misspecifications.

While continued data collection will allow for improved understanding of fish stocks, historical and current data are not always available to support more advanced statistical models. The development of alternative estimation methods that depend on less data, through the application of additional modeling assumptions, can assist management in determining appropriate harvest levels when data are lacking. Chapter 2 evaluated the performance of alternative estimation methods for data-limited and data-moderate situations and current management practices to determine appropriate harvest levels for US west coast groundfish stocks through simulation. The Pacific Fishery Management Council (PFMC) has explicitly accounted for risk and uncertainty when setting harvest levels based on the data available and the assessment method applied. Estimates of harvest produced by each of the catch-only (data-limited) and data-moderate estimation methods used by the PFMC are reduced by a "buffer" value to account for the risk level selected by the Council and the predetermined uncertainty about current biomass. The ability of each estimation method and the applied buffers to produce harvest estimates that prevented overfishing varied based on life history, how informative the data were, and the assumptions regarding input parameter values. The data-moderate estimation method, Extended Depletion-Based Stock Reduction Analysis (XDB-SRA), which estimated stock size and status by fitting to an index of abundance, performed poorly for the flatfish life history when the index of abundance had limited contrast in population trend, resulting in overly optimistic estimates of stock status. However, both data-moderate methods which incorporated an index of abundance, Extended Simple Stock Synthesis (XSSS) and XDB-SRA, generally performed well for the rockfish life history due to its slow dynamics and the rate of change observed in the index of abundance.

The catch-only (data-limited) methods, Depletion-Corrected Average Catch (DCAC) and Depletion-Based Stock Reduction Analysis (DB-SRA), which only made use of assumed distributions for key parameters (i.e., productivity, natural mortality, stock status, and stock status producing maximum yield) also had mixed performances at preventing overfishing. DCAC, which applied a simple one-time calculation to determine a likely sustainable yield, coupled with the buffer, prevented overfishing for both the flatfish and rockfish life histories when parameter distributions were correctly specified, but resulted in periods of overfishing when key parameter distributions were misspecified. DB-SRA, which is based on a delay-difference model and similar parameters as DCAC, and produces annual harvest estimates for all projection years, was highly sensitive to parameter misspecification for both life histories where the buffer applied to reduce harvest failed to prevent overfishing.

The trade-offs between catch-only and data-moderate assessment methods should be considered by fishery managers. Given the general conclusion that both catch-only methods were sensitive to misspecification, especially DB-SRA, applying either for long-term management, could have undesirable properties, especially compared to the data-moderate options that have some ability to update priors based on data. Generally, the highest probability of overfishing among estimation methods (catch-only and data-moderate) occurred when the simulated stocks were at low stock sizes even when the estimation methods were correctly specified. Precaution should be taken if there is a reasonable possibility that the stock may have experienced exploitation levels that would reduce biomass to low sizes at the time of assessment. Based upon these results, the catch-only estimation methods could be applied to determine short-term harvest estimates for stocks that have low historical exploitation. Although caution should be taken when using data-moderate estimation methods, especially for stocks at low biomass levels, they have the advantage of potentially improving estimates about stock biomass and status over time as the amount of data increases. Ultimately, the application of estimation methods should be carefully determined based on the life history of the stock, the exploitation history, and the contrast in the data available. In addition, the use of buffers that account for the increased uncertainty inherent to estimation methods that are based on limited data can reduce the probability of overfishing.

Setting appropriate harvest limits can prevent overfishing and avoid stocks from becoming overfished in the future. However, there are numerous rockfish species off the US west coast where historical exploitation was unsustainable, driving populations below the minimum stock size threshold set by management, resulting in overfished declarations. US federal stocks that are declared overfished are required to implement a formal rebuilding plan dictating how management intends to recover the stock to the target biomass. Chapter 3 used management strategy evaluation (MSE) to examine the performance of alternative rebuilding strategies for overfished stocks. The rebuilding approaches that were most successful at rebuilding stocks involved a fishing mortality during rebuilding that buffered against future uncertainty. Rebuilding plans that set a higher initial probability to recover by the maximum rebuilding year, resulting in lower fishing mortality rates during rebuilding, were more successful in meeting rebuilding timelines with fewer changes to the rebuilding plan, especially for longer lived species. The current requirement to rebuild in $\leq$ 10 years when possible given the life history of a species, resulted in large reductions in catch during rebuilding for the flatfish life history relative to the historical and rebuilt catches to rebuild stocks within the 10-year timeframe. Additionally, the discontinuity within the current guidelines brought about by the 10-year rebuilding rule, resulted in bimodal distributions for rebuilding times and catches obtained during rebuilding. Depending on the age-structure of the simulated population, a subset of the roundfish life history stocks were determined to be able to rebuild within 10-years, while the remaining simulated stocks were allowed to set extended rebuilding times (approximately 20 years vs. $\leq$ 10 years), resulting in less severe harvest restrictions and higher average catches during rebuilding. Determining maximum allowable rebuilding times based on general life history traits or consistent factors related to the minimum time to rebuild (e.g. twice the minimum time) would eliminate the discontinuity in the current US federal rebuilding guidelines and eliminate the requirement for severe harvest restrictions required for a stock to rebuild in $\leq$ 10 years.

The MSE highlighted the trade-offs between rebuilding time and average catches during rebuilding. The more precautionary rebuilding approaches resulted in an increased likelihood of rebuilding on time, but required more conservative harvest estimates. Selecting risk averse rebuilding strategies can buffer against long-term changes in stock production or poor estimation (or assumptions) concerning key biological parameters. However, continued monitoring and assessment should occur to ensure that the population remains on track to rebuild and that there has not been a fundamental shift in the productivity or mortality of the stock.

Maintaining stocks at or near target biomass and preventing stocks from declining into an overfished state is achieved by determining appropriate harvest levels and applying effective harvest control rules. The final chapter, chapter 4, performed an MSE to understand the trade-offs associated with alternative harvest control rules for US west coast flatfish stocks and to evaluate the ability of each strategy to maintain stocks at or near target biomass and prevent stocks from becoming overfished. The application of proxy targets, harvest rates, and thresholds for management across a suite of species with similar life histories presents the challenge of determining an appropriate approach that balances the goal of maximizing harvest while limiting the risk of overfishing individual stocks. The current 25-5 (linear reduction in catch between the target biomass of $0.25SB_0$ and the lower limit of $0.05SB_0$) harvest control rule for US west coast flatfish maintained stocks at or near the target biomass across a range of alternative steepness values. Only the most aggressive harvest strategy evaluated, the 20-5 approach, resulted in exploitation rates that exceeded the fishing mortality rate corresponding to maximum yield when steepness was at the lowest assumed value (0.75). However, when a stock was estimated to have fallen below the target biomass, the harvest control rule reduced catches sufficiently to allow the stock biomass to increase or stabilize, reducing the probability of the stock dropping below the minimum stock size threshold.

While each alternative harvest control rule performed well at preventing the stocks from declining below the defined minimum stock size threshold, there were trade-offs between each strategy evaluated. The harvest control rules that involved higher proxy biomass targets resulted in higher probabilities of the relative spawning biomass being within 10% of the target biomass, with a lower average annual variation in catch, but also resulted in the lowest average catch over the last twenty-five years of the management period. The more aggressive harvest control rules (e.g. lower proxy biomasses) resulted in higher average catch, but with higher average annual variation in catch and lower probabilities of the stock biomass being within 10% of the proxy biomass.

Applying effective harvest control rules is only one component to preventing stocks from declining below threshold levels. Robustness tests examining the ability of independent vs autocorrelated recruitment variation to impact harvest control rule performance resulted in a wide distribution of stock sizes over the last twenty-five years of the management period. The increased variance in stocks sizes across simulations and over years was driven by fluctuations in recruitment and short-term changes in recruitment driven by autocorrelation, paired with the inability of the assessment method to account for these changes when forecasting future catches. The reapplication of the assessment method every 5th year during the management period reduced the probability of long-term overfishing due to unaccounted shifts in productivity, allowing for adjustments in harvest based on assessment results. The trade-offs, the risks, and the potential biology of a stock should be considered carefully by management when selecting a harvest strategy.

4.6.1 Conclusions, caveats, and future work

In conclusion, simulation (including management strategy evaluation) can be applied as an effective tool to evaluate complicated data, biological, and management actions. This dissertation demonstrated the importance of data for estimating biological parameters to correctly determine stock status, the importance of setting precautionary buffers to prevent overfishing based on both the estimation method and data availability, and the importance of management decisions to sustainably manage US west coast groundfish and the associated trade-offs among alternative approaches. Additionally, this work highlighted the interaction between data availability, modeling choices, and management actions. Each of these components can impact the management of fish stocks and should be considered in tandem when determining goals and measuring the success of achieving those goals.

Commonly trade-offs exist between alternative management actions. One trade-off present in multiple chapters was the relationship between higher average catches and the ability to meet management goals (e.g. prevent overfishing, rebuilding a stock quickly, reducing variability in stock size), where it was not feasible to maximize each simultaneously. MSE is a useful tool that can lay bare the risks and trade-offs among alternative actions and can be provide meaningful guidance to managers that will allow them to make informed decisions.

This dissertation highlighted the interaction between life history and management actions. The performance of estimation methods to estimate harvest limits that prevented overfishing were dependent upon the method applied, the life history, and the status of the stock. Additionally, the performance of alternative rebuilding strategies varied across life histories. The life history of a stock will determine the maximum sustainable yield, the ability for a stock to rebuild quickly from low sizes, and impact how informative data are to the estimation of growth and biomass. Management actions should explicitly consider life history and evaluations should be performed to determine if they can meet management's goals based upon the specific situation.

The methods applied and questions evaluated within this dissertation can be expanded upon to provide additional understanding about the importance of data, the selected estimation method, and management strategies. Simulation studies require many assumptions regarding the population dynamics of a simulated stock. To allow for interpretability and the ability to identify drivers of results, misspecification between the operating and the estimation models was limited within each chapter. The premise behind this strategic modeling choice was the assumption that if poor results were observed in the simplified world of simulation, then it would be reasonable to assume the application to a complex real-world situation would at best be similar, but likely much worse. Identifying approaches and strategies that in theory can meet management goals is a critical first step for successful fisheries management. The work performed here provides a baseline of performance, identifying when data are critical, when estimation methods perform well or fail, and the potential of management actions to buffer against undesirable outcomes. Each chapter could be extended by adding additional explorations that alter the assumptions within the operating model or estimation models. Rarely are we able to determine the functional form of key biological processes for real-world populations. Accounting for additional misspecification between the operating and estimation models would allow for increased ability to further bracket uncertainty and identify potential risks.

The work performed here was centered about questions faced for stocks off the US west coast. Extending the methods applied to other regions would provide additional insight. The available data and methods applied to assess exploited fish stocks vary dramatically within the US and globally. The complexities and challenges faced in other areas based on the varying stock biology, the fishery interactions on the stock, and available data would likely require alternative approaches or assumptions to those made within this work. In addition, exploring alternative harvest control rules and estimation methods that are currently being used in other regions could provide added context to the conclusion reached here.

BIBLIOGRAPHY

A'mar, Z.T., Punt, A.E., and Dorn, M.W. 2009. The evaluation of two management strategies for the Gulf of Alaska walleye pollock fishery under climate change. ICES Journal of Marine Science: Journal du Conseil **66**: 1614–1632.

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

Beddington, J.R. and Kirkwood, G.P. 2005. The estimation of potential yield and stock status using life-history parameters. Philosophical Transactions of the Royal Society B: Biological Sciences **360**: 163–170.

Beverton, R.J. and Holt, S.J. 1957. On the dynamics of exploited fish populations. Fishery Investigations Series II, Vol. XIX, Ministry of Agriculture, Great Britian.

Butterworth, D.S. and Punt, A.E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science: Journal du Conseil **56**: 985–998.

Cadrin, S.X. and Pastoors, M.A. 2008. Precautionary harvest policies and the uncertainty paradox. Fisheries Research **94**: 367–372.

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

Chen, Y., Chen, L., and Stergiou, K.I. 2003. Impacts of data quantity on fisheries stock assessment. Aquatic Sciences **65**: 92–98.

Clark, W.G. 1991. Groundfish exploitation rates based on life history parametersClark_1991.pdf. Canadian Journal of Fisheries and Aquatic Sciences **48**: 734–750.

Conn, P.B., Williams, E.H., and Shertzer, K.W. 2010. When can we reliably estimate the productivity of fish stocks? Canadian Journal of Fisheries and Aquatic Sciences **67**: 511–523.

Conway, F. and Shaw, W. 2008. Socioeconomic Lessons Learned from the Response to the Federally-Declared West Coast Groundfish Disaster. Fisheries **33**: 269–277.

Cope, J., Dick, E.J., MacCall, A., Monk, M., Soper, B., and Wetzel, C. 2013. Data-moderate stock assessments for brown, China, copper, sharpchin, stripetail, and yellowtail rockfishes and English and rex soles in 2013. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR.

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

Deroba, J.J. and Bence, J.R. 2008. A review of harvest policies: Understanding relative performance of control rules. Fisheries Research **94**: 210–223.

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

Federal Register 1998. Magnuson-Stevens Act Provision; National Standard Guidelines – Final Rule.

Francis, R.I.C.C. 1992. Use of Risk Analysis to Assess Fishery Management Strategies: A Case Study using Orange Roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. Canadian Journal of Fisheries and Aquatic Sciences **49**: 922–930.

Francis, R.C. and Shotton, R. 1997. "Risk" in fisheries managment: a review. Canadian Journal of Fisheries and Aquatic Sciences **54**: 1699–1715.

Francis, R.C. 1986. Two Fisheries Biology Problems in West Coast Groundfish Management. North American Journal of Fisheries Management **6**: 453–462.

Fulton, E.A., Punt, A.E., Dichmont, C.M., Gorton, R., Sporcic, M., Dowling, N., Little, L.R., Haddon, M., Klaer, N., and Smith, D.C. 2016. Developing risk equivalent data-rich and data-limited harvest strategies. Fisheries Research **183**: 574–587.

Givens, G.H. and Raftery, A.E. 1996. Local adaptive importance sampling for multivariate densities with strong nonlinear relationships. Journal of the American Statistical Association **91**: 132–141.

Graham, M. 1935. Modern theory of exploiting a fishery, and application to North Sea trawling. Journal du Conseil Conseil International pour l'Exploartions de la Mer **10**: 264–274.

Haltuch, M.A., Hicks, A.C., and See, K. 2011. Status of the U.S. petrale sole resource in 2010. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR.

Haltuch, M.A. and Ono, K. 2013. Status of the U.S. petrale sole resource in 2012. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Haltuch, M.A. and Punt, A.E. 2011. The promises and pitfalls of including decadal-scale climate forcing of recruitment in groundfish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences **68**: 912–926.

Haltuch, M.A., Punt, A.E., and Dorn, M.W. 2008. Evaluating alternative estimators of fishery management reference points. Fisheries Research **94**: 290–303.

Hicks, A.C. and Wetzel, C.R. 2011. The status of Dover sole (*Microstomus pacificus*) along the U.S. West Coast in 2011. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Hilborn, R. 2010. Pretty Good Yield and exploited fishes. Marine Policy **34**: 193–196.

Hilborn, R., Stewart, I.J., Branch, T.A., and Jensen, O.P. 2012. Defining Trade-Offs among Conservation, Profitability, and Food Security in the California Current Bottom-Trawl Fishery. Conservation Biology **26**: 257–268.

Hilborn, R.H. 1977. Comparison of fisheries control systems that utilize catch and effort data. Journal of Fisheries Research Board Canada **36**: 1477–1489.

Hixon, M.A., Johnson, D.W., and Sogard, S.M. 2014. BOFFFFs: on the importance of conserving old-growth age structure in fishery populations. ICES Journal of Marine Science **71**: 2171–2185.

Hjort, J., Jahn, G., and Ottestad, P. 1933. The optimum catch. Hvalradets Skrifter **7**: 92–127.

Hollowed, A.B., Barange, M., Ito, S.i., Kim, S., Loeng, H., and Peck, M.A. 2011. Effects of climate change on fish and fisheries: forecasting impacts, assessing ecosystem responses, and evaluating management strategies. ICES Journal of Marine Science: Journal du Conseil **68**: 984–985.

Ianelli, J.N., Hollowed, A.B., Haynie, A.C., Mueter, F.J., and Bond, N.A. 2011. Evaluating management strategies for eastern Bering Sea walleye pollock (*Theragra chalcogramma*) in a changing environment. ICES Journal of Marine Science: Journal du Conseil **68**(6): 1297–1304.

International Joint Technical Committee for Pacific Hake 2014. Status of the Pacific Hake (whiting) stock in the U.S. and Canadian waters in 2014 with a managment strategy evaluation. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

International Whaling Commission 1992a. Report of the Scientific Committee, Annex D. Report of the Sub-Committee on Managment Procedures. Reports of the International Whaling Commission **42**: 87–136.

International Whaling Commission 1992b. Report of the Third Comprehensive Assessment Workshop on Management Procedures. Reports of the International Whaling Commission **42**: 270–304.

Johnson, K.F., Councill, E., Thorson, J.T., Brooks, E., Methot, R.D., and Punt, A.E. 2016. Can autocorrelated recruitment be estimated using integrated assessment models and how does it affect population forecasts? Fisheries Research **183**: 222–232.

Johnson, K.F., Monnahan, C.C., McGilliard, C.R., Vert-pre, K.A., Anderson, S.C., Cunningham, C.J., Hurtado-Ferro, F., Licandeo, R.R., Muradian, M.L., Ono, K., Szuwalski, C.S., Valero, J.L., Whitten, A.R., and Punt, A.E. 2015. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. ICES Journal of Marine Science: Journal du Conseil **72**: 137–150.

Kaplan, I.C. and Helser, T. 2007. Stock assessment of the arrowtooth flounder (*Atheresthes stomias*) population off the West Coast of the United States in 2007. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Kimura, D.K., Balsiger, J.W., and Ito, D.H. 1984. Generalized stock reduction analysis. Canadian Journal of Fisheries and Aquatic Sciences **41**: 1325–1333.

Kinas, P.G. 1996. Bayesian fishery stock assessment and decision making using adaptive importance sampling. Canadian Journal of Fisheries and Aquatic Sciences **53**: 414–423.

King, J.R. and McFarlane, G.A. 2003. Marine fish life history strategies: applications to fishery management. Fisheries Management and Ecology **10**: 249–264.

Kuriyama, P.T., Branch, T.A., Bellman, M.A., and Rutherford, K. 2016. Catch shares have not led to catch-quota balancing in two North American multispecies trawl fisheries. Marine Policy **71**: 60–70.

Larkin, P.A. 1977. An Epitaph for the Concept of Maximum Sustained Yield. Transactions of the American Fisheries Society **106**: 1–11.

Lee, H.H., Maunder, M.N., Piner, K.R., and Methot, R.D. 2012. Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? Fisheries Research **125–126**: 254–261.

Legault, C.M. and Palmer, M.C. 2015. In what direction should the fishing mortality target change when natural mortality increases within an assessment? Canadian Journal of Fisheries and Aquatic Sciences **73**: 349–357.

Little, L.R., Punt, A.E., Dichmont, C.M., Dowling, N., Smith, D.C., Fulton, E.A., Sporcic, M., and Gorton, R.J. 2016. Decision trade-offs for cost-constrained fisheries management. ICES Journal of Marine Science: Journal du Conseil **73**: 494–502.

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

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

Martell, S. and Stewart, I. 2014. Towards defining good practices for modeling time-varying selectivity. Fisheries Research **158**: 84–95.

McAllister, M.K., Pikitch, E.K., Punt, A.E., and Hilborn, R. 1994. A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences **51**: 2673–2687.

Methot, R.D. and Piner, K.R. 2002. Rebuilding analysis for yelloweye rockfish: Update to incorporate results of coastwide assessment in 2002. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Methot, R.D. 2015. Prioritizing fish stock assessments. Technical Report NOAA Technical Memorandum NMFS-F/SPO 152, US Department of Commerce.

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

Milazzo, M.J. 2012. Progress and problems in US marine fisheries rebuilding plans. Reviews in Fish Biology and Fisheries **22**: 273–296.

Mora, C., Myers, R.A., Coll, M., Libralato, S., Pitcher, T.J., Sumaila, R.U., Zeller, D., Watson, R., Gaston, K.J., and Worm, B. 2009. Management effectiveness of the world's marine fisheries. PLoS Biol **7**: e1000131.

Mueter, F.J., Bond, N.A., Ianelli, J.N., and Hollowed, A.B. 2011. Expected declines in recruitment of walleye pollock (*Theragra chalcogramma*) in the eastern Bering Sea under future climate change. ICES Journal of Marine Science: Journal du Conseil p. fsr022.

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

National Reserach Council 2013. Evaluating the effectiveness of fish stock rebuilding plans in the United States. Technical report, The National Academies Press, Washington, DC.

Neubauer, P., Jensen, O.P., Hutchings, J.A., and Baum, J.K. 2013. Resilience and recovery of overexploited marine populations. Science **340**: 347–349.

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

Pacific Fishery Management Council 2006. Pacific Coast groundfish fishery management plan for California, Oregon, and Washington groundfish fishery through Amendment 19. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Pacific Fishery Management Council 2009. Scientific and statistical committee report on stock assessments and rebuilding analyses for 2011-2012 groundfish fisheries. Number http://www.pcouncil.org/wp-content/uploads/bb_2009_11_G2b_SUP_SSC_1109.pdf in Item G.2 b Supplemental SSC Report, November 2009. Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Pacific Fishery Management Council 2010. Amendment 23: Considerations for a new harvest specification framework that incorporates revised national standard 1 guidelines to prevent overfishing. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Pacific Fishery Management Council 2011. Proposed harvest specification and management measures for the 2011-2012 Pacific Coast Groundfish fishery and amendment 16-5 to the Pacific Coast Groundfish fishery management plan to update existing rebuilding plans and to adopt a rebuilding plan for petrale sole. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Pacific Fishery Management Council 2014a. Pacific coast groundfish fishery management plan. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Pacific Fishery Management Council 2014b. Pacific Coast groundfish fishery managment plan for the California, Oregon, and Washington groundfish fishery: Appendix F overfished species rebuilding plan. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Pacific Fishery Management Council 2014c. Status of the Pacific Coast groundfish fishery: Stock assessment and fishery evaluation. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Pacific Fishery Management Council 2016. Pacific Coast groundfish fishery management plan for California, Oregon, and Washington groundfish fishery. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Patrick, W.S. and Cope, J. 2014. Examining the 10-Year Rebuilding Dilemma for U.S. Fish Stocks. PLoS ONE **9**: e112232.

Patrick, W.S., Morrison, W., Nelson, M., and Marrero, R.L.G. 2013. Factors affecting management uncertainty in US fisheries and methodological solutions. Ocean & coastal management **71**: 64–72.

Peck, S.L. 2004. Simulation as experiment: a philosophical reassessment for biological modeling. Trends in Ecology & Evolution **19**: 530–534.

Punt, A.E., A'mar, T., Bond, N.A., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., Haltuch, M.A., Hollowed, A.B., and Szuwalski, C. 2014a. Fisheries management under climate and environmental uncertainty: control rules and performance simulation. ICES Journal of Marine Science **71**: 2208–2220.

Punt, A.E. and Butterworth, D.S. 1995. The effects of future consumption by the Cape fur seal on catches and catch rates of the Cape hakes. 4. Modelling the biological interaction between Cape fur seals *Arctocephalus pusillus pusillus* and the Cape hakes *Merluccius capensis* and *M. paradoxus*. South African Journal of Marine Science **16**: 255–285.

Punt, A.E. and Butterworth, D. 1996. Further remarks on the Bayesian approach for assessing the Bering-Chukchi-Beaufort Seas stock ob bowhead whales. Reports of the International Whaling Commission **46**: 481–491.

Punt, A.E. and Smith, A.D.M. 1999. Harvest strategy evaluation for the eastern stock of gemfish (*Rexea solandri*). ICES Journal of Marine Science: Journal du Conseil **56**: 860–875.

Punt, A.E., Smith, A.D.M., and Cui, G. 2002. Evaluation of management tools for Australia's southeastern fishery. Marine and Freshwater Research **53**: 631–644.

Punt, A.E., Smith, A.D.M., Smith, D.C., Tuck, G.N., and Klaer, N.L. 2014b. Selecting relative abundance proxies for B_{msy} and B<sub>MEY<s/ub>. ICES Journal of Marine Science **71**: 469–483.

Punt, A.E. 2011. The impact of climate change on the performance of rebuilding strategies for overfished groundfish species of the US west coast. Fisheries Research **109**(2): 320–329.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., and Haddon, M. 2014c. Management strategy evaluation: best practices. Fish and Fisheries **17**: 303–334.

Punt, A.E. and Donovan, G.P. 2007. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. ICES Journal of Marine Science: Journal du Conseil **64**: 603–612.

Punt, A.E., Dorn, M.W., and Haltuch, M.A. 2008. Evaluation of threshold management strategies for groundfish off the U.S. West Coast. Fisheries Research **94**: 251–266.

Punt, A.E. and Ralston, S. 2007. A management strategy evaluation of rebuilding revision rules for overfished rockfish stocks. In Biology, Assessment and Management of North Pacific Rockfishes pp. 329–351.

Ralston, S., Punt, A.E., Hamel, O.S., DeVore, J.D., and Conser, R.J. 2011. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fishery Bulletin **109**: 217–232.

Rose, K.A., Cowan, J.H., Winemiller, K.O., Myers, R.A., and Hilborn, R. 2001. Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis. Fish and Fisheries **2**: 293–327.

Rubin, D.B. 1987. The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The SIR algorithm. Journal of the American Statistical Association **82**: 543–546.

Rubin, D.B. 1988. Using the SIR algorithm to simulate posterior distributions. In Bayesian statistics 3: Proceedings of the Third Valencia International Meeting, June 1-5, 1987. Clarendon Press, Oxford.

Russell, E.S. 1931. Some theoretical considerations on the “Overfishing” problem. Journal du Conseil **6**: 3–20.

Shepherd, J. 1982. A versatile new stock-recruitment relationship for fisheries, and construcitons of sustainable yield curves. ICES Journal of Marine Science: Journal du Conseil **40**: 67–75.

Shertzer, K.W. and Prager, M.H. 2007. Delay in fishery management: diminished yield, longer rebuilding, and increased probability of stock collapse. ICES Journal of Marine Science: Journal du Conseil **64**: 149–159.

Sissenwine, M. 1978. Is MSY an adequate foundation for optimum yield? Fisheries **3**: 22–42.

Smith, A.D.M., Sainsbury, K.J., and Stevens, R.A. 1999. Implementing effective fisheries-management systems – management strategy evaluation and the Australian partnership approach. ICES Journal of Marine Science: Journal du Conseil **56**: 967–979.

Smith, A. 1994. Management strategy evaluation-the light on the hill. Australian Society for Fish Biology Workshop Proceedings, Perth, population dynamics for fisheries managment. d.a. hancok (ed) edition.

Smith, S.J., Hunt, J.J., and Rivard, D. 1993. Risk Evaluation and Biological Reference Points for Fisheries Management. Canadian Special Publication Fisheries and Aquatic Science 120.

Southward, G. 1968. A simulation of managment strategies in the Pacific halibut fishery. Ph.D. thesis, University of Washington, Seattle.

Stachura, M.M., Essington, T.E., Mantua, N.J., Hollowed, A.B., Haltuch, M.A., Spencer, P.D., Branch, T.A., and Doyle, M.J. 2014. Linking Northeast Pacific recruitment synchrony to environmental variability. Fisheries Oceanography **23**: 389–408.

Starr, P.J., Breen, P.A., Hilborn, R.H., and Kendrick, T.H. 1997. Evaluation of a management decision rule for a New Zealand rock lobster substock. Marine and Freshwater Research **48**: 1093–1101.

Stawitz, C.C., Hurtado-Ferro, F., Kuriyama, P.T., Trochta, J., Johnson, K.F., Haltuch, M.A., and Hamel, O.S. 2015. Stock assessment update: Status of the U.S. petrale sole resource in 2014. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Stewart, I.J. 2007. Updated U.S. English sole stock assessment: status of the resource in 2007. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Stewart, I.J., Wallace, J.R., and McGilliard, C.R. 2009. Status of the U.S. yelloweye rockfish resource in 2009. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Sustainable Fisheries Act 1996. Sustainble Fisheries Act of the Magnuson-Stevens Fishery Conservation and Management Act. Technical report, US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Silver Spring, MD.

Swain, D.P. and Benoît, H.P. 2015. Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Marine Ecology Progress Series **519**: 165–182.

Sylvia, A. 2015. Effects of assessment frequency and data-management lag on fishery management performance: strategies for improvement. University of Maryland, College Park.

Szuwalski, C.S. and Punt, A.E. 2012. Fisheries management for regime-based ecosystems: a management strategy evaluation for the snow crab fishery in the eastern Bering Sea. ICES Journal of Marine Science: Journal du Conseil **70**: 955–967.

Taylor, I.G. and Methot, R.D. 2013. Hiding or dead? A computationally efficient model of selective fisheries mortality. Fisheries Research **142**: 75–85.

Thorson, J.T. and Wetzel, C.R. 2015. The status of canary rockfish (*Sebastes pinniger*) in the California Current in 2015. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR.

Thorson, J.T. and Cope, J.M. 2015. Catch curve stock-reduction analysis: An alternative solution to the catch equations. Fisheries Research **171**: 33–41.

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

Wetzel, C.R. and Punt, A.E. 2016. The impact of alternative rebuilding strategies to rebuild overfished stocks. ICES Journal of Marine Science: Journal du Conseil **73**: 2190–2207.

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

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

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

Winemiller, K.O. and Rose, K.A. 1992. Patterns of life-history diversification in North American fishes: implications for population regulation. Canadian Journal of Fisheries and Aquatic Sciences **49**: 2196–2218.

Worm, B., Hilborn, R., Baum, J.K., Branch, T.A., Collie, J.S., Costello, C., Fogarty, M.J., Fulton, E.A., Hutchings, J.A., Jennings, S., and others 2009. Rebuilding global fisheries. Science **325**: 578–585.

Yin, Y. and Sampson, D.B. 2004. Bias and Precision of Estimates from an Age-Structured Stock Assessment Program in Relation to Stock and Data Characteristics. North American Journal of Fisheries Management **24**: 865–879.

Appendix A

SUPPLEMENTARY INFORMATION FOR CHAPTER 1

A.1 The impact of data loss in the presence of fishery independent data

Additional simulations were conducted to evaluate the impact of only having fishery information versus indices of abundance and length- and age-composition data available from both a fishery-independent survey and a fishery. The operating model generated a highly uncertain survey ($CV = 0.40$) that was conducted on a biennial basis with low length- and age-composition sample sizes, representative of a survey that poorly sampled the simulated stock (e.g. due to habitat and gear restrictions or restricted sampling areas) (Fig. A.1). The survey selectivity was assumed to be fixed at an asymptotic shape, selecting fish at smaller sizes relative to the fishery selectivity. All other specifications for the fishery within the operating model and the assumptions applied by the estimation method were the same as those detailed in the *Material and Methods* section.

The estimates of spawning biomass (Fig. A.2a-c) and relative spawning biomass (Fig. A.3a-c) for the time-invariant case were median unbiased at the time of the first assessment in year 50. The addition of a survey index and composition data for all data scenarios led to less among-simulation variability and reduced median bias over the management period relative to the simulations without survey data (Figs. 1.4a-c and 1.5a-c). The presence of survey data when fishery data were eliminated (eliminated data scenario) allowed the majority of the simulated stocks being estimated rebuilt by the end of the management period (Fig. A.2c) compared to the large fraction of simulations that failed to be estimated rebuilt when only historical data were available from the fishery (Fig. 1.4c).

The inclusion of survey data, in addition to fishery data when time-varying parameters were present led reduced the among-simulation variability in the estimates of spawning biomass and relative spawning biomass (Figs. A.2d-f and A.3d-f). The full data scenario had the lowest root mean squared error (RMSE) for relative spawning biomass during the early portion of the management period for both cases (time-invariant and time-varying), when the majority of simulations were rebuilding for both cases (Fig. A.4). However, midway through the management period, after a majority of the simulated stocks had rebuilt and data restrictions were removed, the data scenarios resulted in similar RMSEs (Fig. A.4). The inclusion of survey data for all data scenarios resulted in similar estimates of the median number of years to recover to the target biomass, which were similar to the median rebuilding time with the operating model (Table A.1).

A.2 Tables

Table A.1: The median and 90% simulation interval for the estimated number of years needed to rebuild to the target biomass, the operating model number of years needed to rebuild to target biomass, and the number of stocks that failed to rebuild to the target biomass determined by the estimation method (EM) and the operating model (OM) for each case and data scenario.

Selectivity/ data scenario	Estimated num. of rebuilding years		Operating model num. of rebuilding years		Num. of stocks that failed to rebuild	
	Median	90% SI	Median	90% SI	EM	OM
Time-invariant						
full data	37	(18-79)	39	(18-79)	2	2
reduced data	37	(19-67)	38	(16-78)	0	4
eliminated data	37	(18-74)	38.5	(16-84)	2	4
Time-varying						
full data	31	(13-85	41	(14-85)	15	3
reduced data	31	(13-85)	39	(13-80)	9	3
eliminated data	31	(13-88)	39	(13-80)	8	3

A.3 Figures

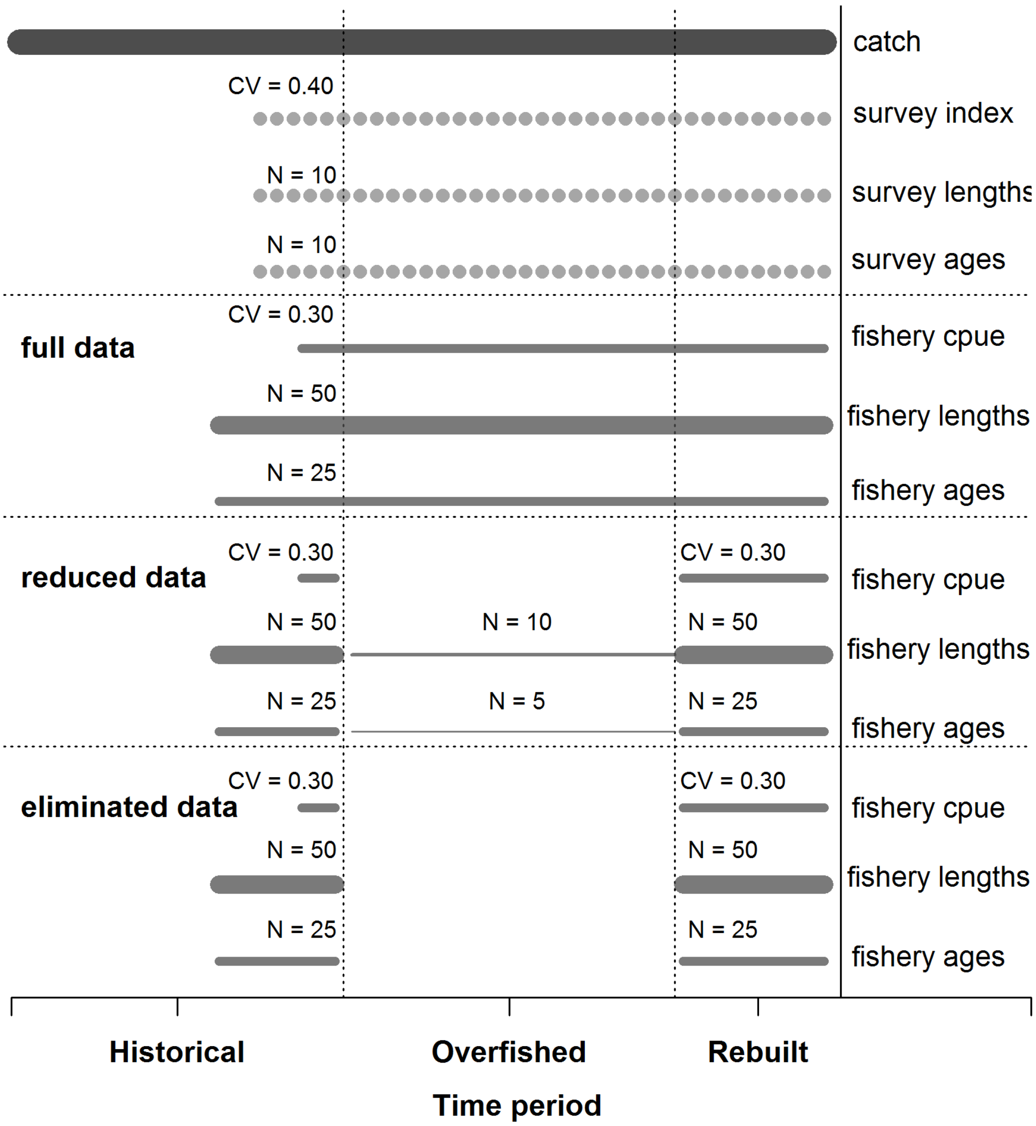

Figure A.1: Summary of the data available for each of the data scenarios. Catches, a fishery independent survey with length- and age-composition data were available for all data scenarios.

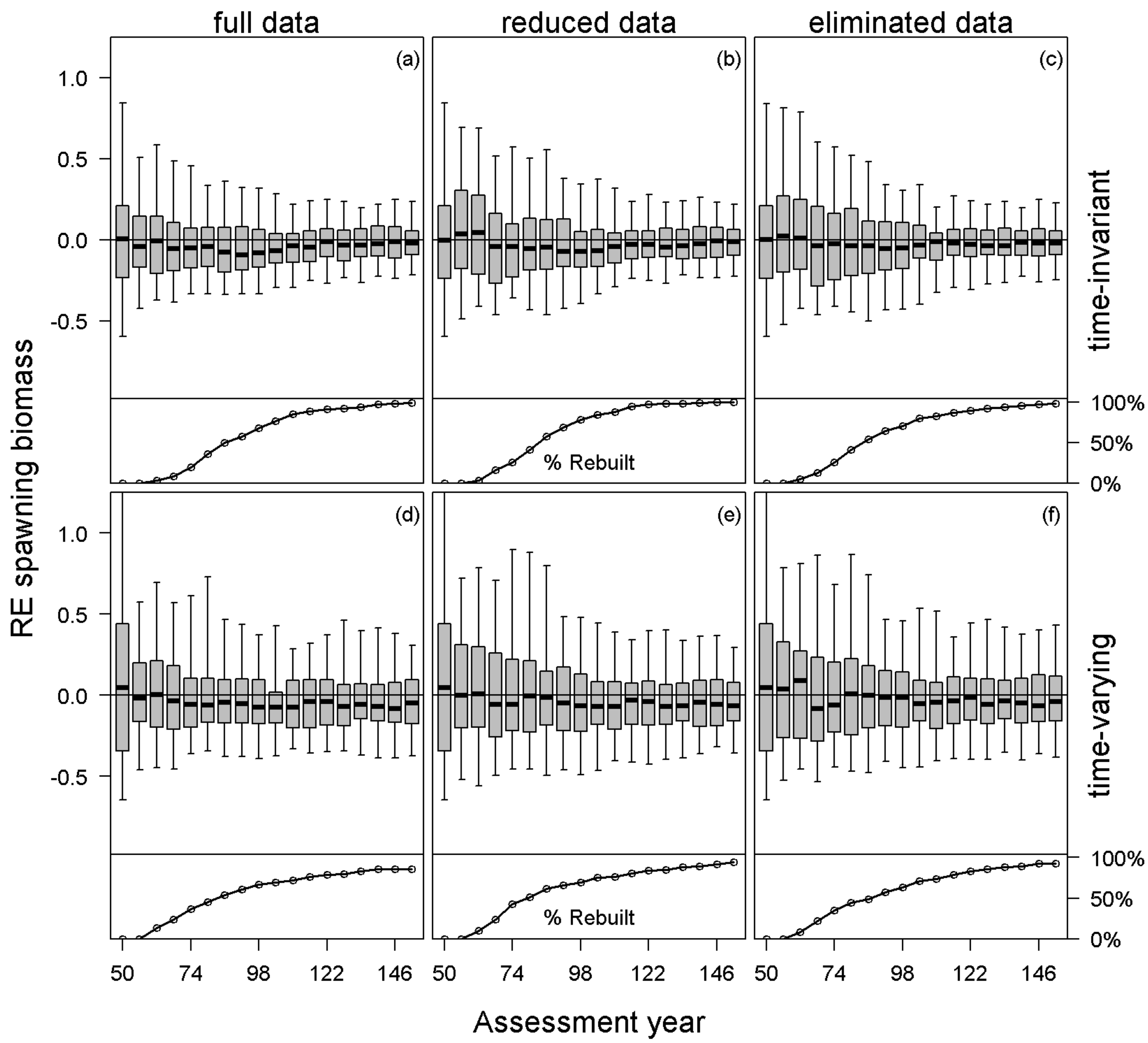

Figure A.2: Relative error of estimated spawning biomass in each assessment year for each case and data scenario for all simulations (top panel) and the percentage of stocks that had rebuilt to the target biomass during the management period (bottom panel), with data collection consequently returning to historical levels. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

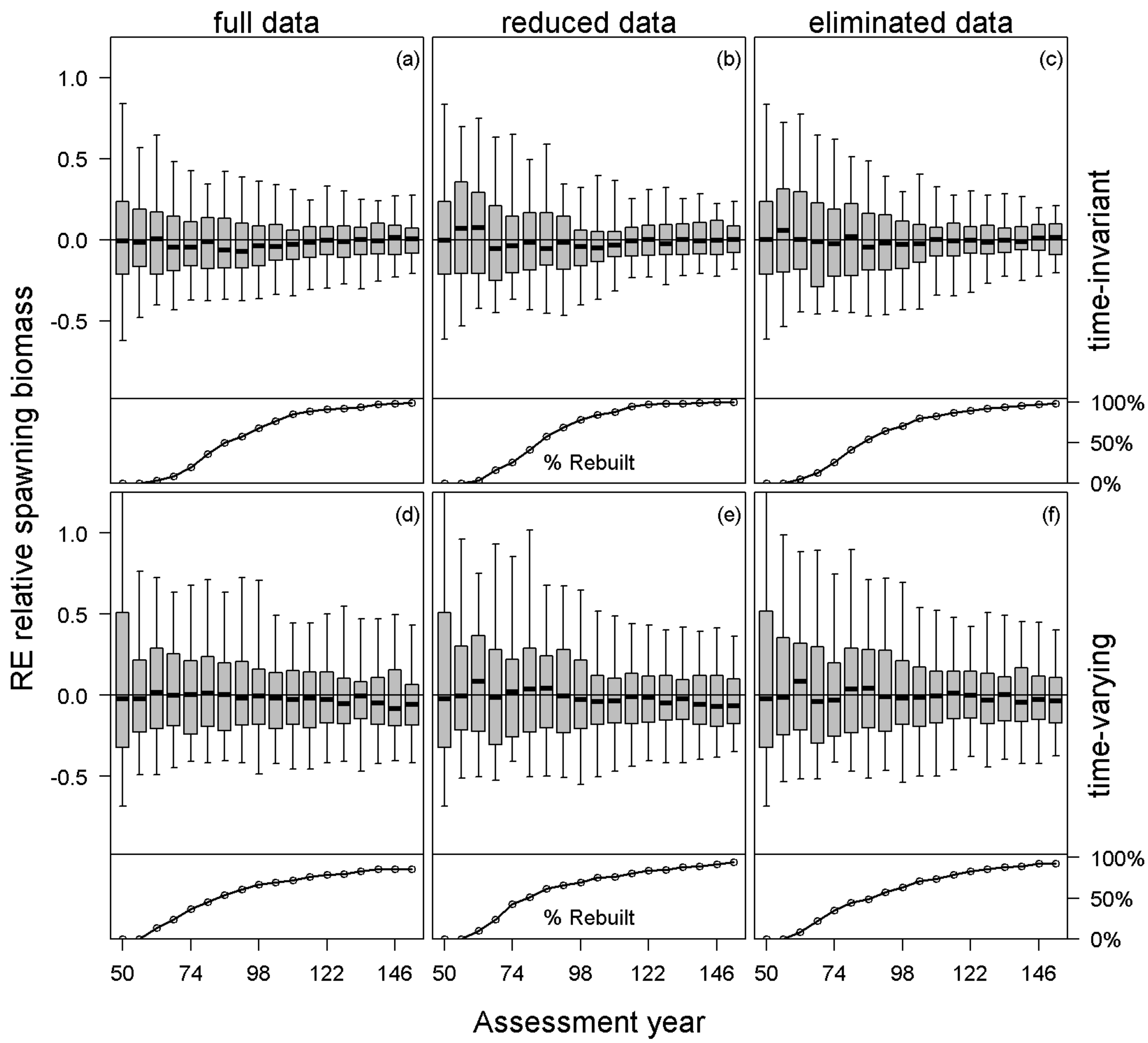

Figure A.3: Relative error of estimated relative spawning biomass in each assessment year for each case and data scenario for all simulations (top panel) and the percentage of stocks that had rebuilt to the target biomass during the management period (bottom panel), with data collection consequently returning to historical levels. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

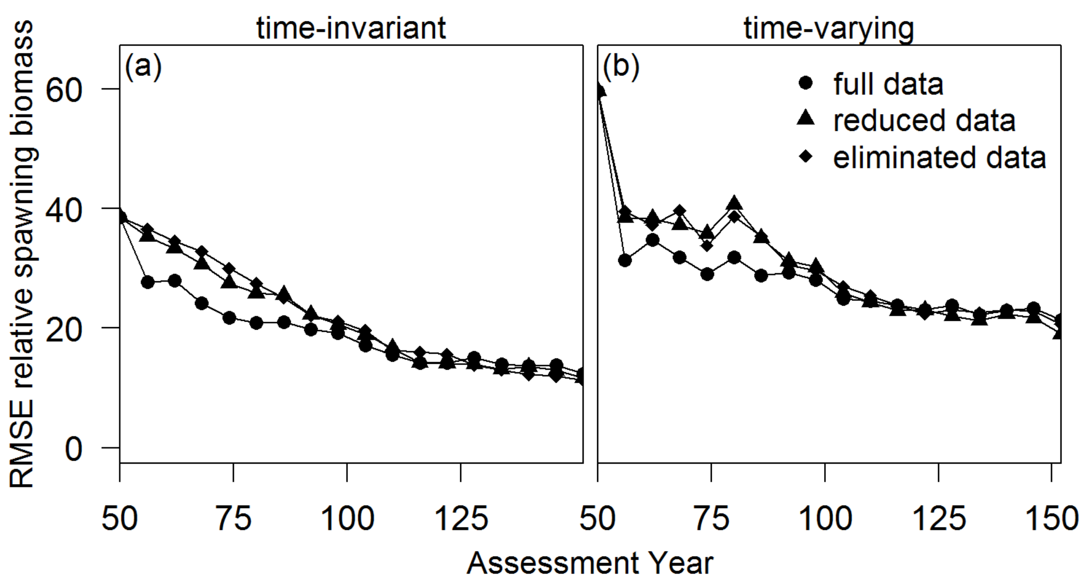

Figure A.4: The root mean square error (RMSE) about relative spawning biomass in the assessment year for each case and data scenario.

Appendix B

SUPPLEMENTARY INFORMATION FOR CHAPTER 2

B.1 The operating and observation models

The numbers-at-age at the start of the year are computed using the equation:

$$N_{t+1,\gamma,a} = \begin{cases} R_t & \text{if a} = 0 \\ N_{t,a-1}e^{-(M_t+S_{t,a-1}F_t)} & \text{if } 1 \leq \text{a} < \text{A-1} \\ N_{t,A-1}e^{-(M_t+S_{t,A-1}F_t)} + N_{t,A}e^{-(M_t+S_{t,A}F_t)} & \text{if a} = \text{A} \end{cases} \tag{B.1}$$

where $N_{t,\gamma,a}$ is the number of fish of age a and gender γ at the start of the year t, R_t is the number of age-0 animals at the start of year t, S_γ, a is the selectivity by gender and age, A is the plus group, F_t is the instantaneous fishing mortality rate during year t, and M is the instantaneous rate of natural mortality. The number of age-0 fish is related to spawning biomass according to the Shepherd (1982) stock recruitment relationship which applies a shape parameter to account for density dependence in recruitment survival at higher spawning biomass levels, and is able to take the shape of the Beverton Holt ($c = 1$) or a Ricker stock recruitment relationship ($c > 1$) based upon the shape parameter c which alters the level of compensation of the stock recruit curve:

$$R_t = \Big(\frac{SB_t}{SB_0}\Big)\Big(\frac{5hR_0SB_0^c(1-0.20^c)}{SB_0^c(1-5h0.20^c)+(5h-1)SB_t^c}\Big)e^{-0.5\sigma_2^2+\epsilon_t} \tag{B.2}$$

where SB_t is the spawning biomass at the start of the spawning season in year t, σ_R is the standard deviation of recruitment in log space, h determines the level of recruitment produced when the spawning biomass is reduced to 20% of its unfished size (also known as steepness), ϵ_t is the auto-correlated lognormal recruitment deviation for year defined as:

$$\epsilon_t = \rho\epsilon_{t-1} + \sqrt{1-\rho^2}\phi_t \quad \phi_t \sim \ N(0;\sigma_R^2) \tag{B.3}$$

where ρ is the level of autocorrelation associated with recruitment and ϕ_t is the recruitment deviation for year t. Both life histories assumed the same autocorrelation equal to one half of the total variability in recruitment being attributed to either noise or environmental factors.

A non-equilibrium starting condition was created by applying equations (B.1) and (B.2) for the number of years equal to the maximum age for each life history prior to the start of fishing with $F = 0$ and with autocorrelated recruitment deviations. The initial period of the fishery operated for 50 years, with the catch of fish of age a and gender γ during year t in numbers determined by:

$$C_{t,\gamma,a} = \frac{S_{\gamma,a}F_t}{M + S_{\gamma,a}F_t} N_{t,\gamma,a}\big(1 - e^{-M-S_{\gamma,a}}\big) \tag{B.4}$$

The observation model was used to generate data for a biomass index from the operating model that was used by the data-moderate estimation methods. The biomass available for observation during each year t is determined by:

$$\widetilde{B}_t = \sum_{\gamma} \sum_{a=1}^{A} \sum_{l=1}^{L} \phi_{\gamma,a,l} w_{\gamma,l} S_{s,\gamma,l} N_{t,\gamma,a} e^{-0.5(M+S_{f,\gamma,a}F_t)} \tag{B.5}$$

where $\phi_{\gamma,a,l}$ is the age-length-based transition matrix by gender (see Methot and Wetzel, 2013 for details), $w_{\gamma,l}$ is the weight by gender γ and length l, and $S_{s,\gamma,l}$ is the selectivity by the survey for gender γ and length l. The observed survey biomass index is related to the available population biomass according to:

$$I_t = Q\widetilde{B}_t e^{-0.5\sigma_s^2+\epsilon_t^s} \quad \epsilon_t \sim \; N(0, \sigma_s^2) \tag{B.6}$$

where Q is the catchability coefficient for the survey, and σ_s is the standard deviation of the survey catchability in log space.

B.2 Adaptive importance sampling

Initial population trajectories were created based on draws from each of the prior distributions. The parameter related to the relative biomass of the population (δ_{50}) was always assumed to pertain to year 50. The likelihood corresponding to a parameter vector is calculated as:

$$L_i(\theta_i|data) = (2\pi)^{(1/N)/2}\sigma^{1-N}e^{-\frac{1}{2\sigma^2}\sum_{t=1}^{N}(ln(I_t)-ln(\hat{q}_i\hat{B}_{t,i}))^2} \tag{B.7}$$

where N is the length of the biomass index, σ is the survey standard deviation, I_t is the observed biomass value for year t, $\hat{B}_{t,i}$ is the estimated biomass in year t for the i^{th} trajectory, and $\hat{q}_i$ is the maximum likelihood estimate of the catchability coefficient for the i^{th} trajectory calculated as:

$$\hat{q}_i = e^{\frac{1}{N}\sum_{t=1}^{N}(ln(I_i)-ln(\widetilde{B}_{t,i}))} \tag{B.8}$$

Setting $\hat{q}_i$ to the value from this equation means that equation B.7 is the marginal likelihood (Punt and Butterworth, 1996). The maximum likelihood estimate for $\hat{q}_i$ is applied to reduce the number of parameters to only those of particular interest over which to apply AIS. Punt and Butterworth (1996) show that this treatment of $\hat{q}_i$ does not differ from a strictly Bayesian analysis with a uniform prior on $ln(q)$.

The total survey standard deviation, σ, is defined as:

$$\sigma = \sqrt{1 + \sigma_s^2 + v} \tag{B.9}$$

where σ_s is the pre-specified survey standard deviation and v is an added uniformly distributed variance term from 0 to 1. The σ_s is a pre-specified input parameter, and additional variance is accounted for through the v, which is then integrated out.

The likelihood of the i^{th} trajectory given the data is combined with the prior probabilities of the parameter values to calculate the sampling envelope weights:

$$w_i = \frac{L_i(\theta_i|data)P(\theta_i)}{f(\theta_i)} \tag{B.10}$$

where $P(\theta_i)$ is the prior probability for parameter set i, and $f(\theta_i)$ is the value of the importance function for parameter set i. The probability under the importance function $f(\theta_i)$ in the first iteration of the SIR is equal to the prior probability $P(\theta_i)$. A sample with replacement size equal to 25% of the initial draw size and probability equal to the weights, w_i, was used to update the importance function $f(\theta_i)$. The mean and covariance of the parameters from the SIR draw were calculated and a Student's multivariate t-distribution applied to generate new parameter distributions that constructed new proposed posterior distributions. The iterative process of applying SIR, representing the resulting posterior using a Student multivariate-t distribution, and drawing new parameter vectors continued until a pre-specified entropy criterion was met. Entropy is a measure of uniformity about the sample weights, with values ranging between 0 and 1. The value of entropy will approach 1, which indicates a perfectly uniform distribution with each weight being equal to $1/O$ where O is the total number of draws as the importance sample function closes in on the target distribution. Entropy was calculated as:

$$e = -\sum_{i=1}^{O} w_i \frac{log(w_i)}{log(O)} \tag{B.11}$$

The AIS continued until an entropy criterion of 0.92, an entropy value that during model testing was a point where there was limited change in the resulting posterior distributions, was reached. A final sample of parameter vectors was drawn from the distribution of parameters that met the entropy criterion. This final sample of parameter vectors was then used to construct the posterior distribution for the time-trajectory of biomass.

VITA

Chantell Wetzel was born in Yakima, Washington. She obtained her B.S. in Fisheries Science at the School of Aquatic and Fishery Sciences at the University of Washington in Seattle. She continued her education at the School of Aquatic and Fishery Sciences and obtained her M.Sc. and Doctorate under the supervision of André Punt. Throughout her graduate career she has worked at the Northwest Fisheries Science Center as a member of Population Ecology Program with the Fishery Resource Analysis and Monitoring Division.