Reaching management objectives given uncertainty: A management strategy evaluation of the eastern Bering Sea snow crab (*Chionoecetes opilio*) fishery

Cody S. Szuwalski

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

Doctor of Philosophy

University of Washington
2014

Reading Committee:
André Punt, Chair
Ray Hilborn
David Armstrong

Program Authorized to Offer Degree:
School of Aquatic and Fisheries Science

University of Washington

Abstract

Reaching management objectives given uncertainty: A management strategy evaluation of the eastern Bering Sea snow crab (*Chionoecetes opilio*) fishery

Cody Szuwalski

Chair of the Supervisory Committee:
Director Professor André E. Punt
School of Aquatic and Fisheries Science

The snow crab (*Chionoecetes opilio*) population in the eastern Bering Sea historically supported a lucrative fishery, but was declared overfished in 1999. The subsequent rebuilding plan did not rebuild the fishery in the time specified, and low recruitment over this period likely contributed to the failure. A key component of the rebuilding plan for snow crab was the stock assessment on which management advice was based. In common with all stock assessments, there are numerous potential sources of uncertainty associated with the snow crab assessment, but it is not clear which source is most influential. Consequently, it is unclear how research should be focused and how management should respond to these uncertainties. This dissertation: 1) examines the influence of uncertainty on the ability of the assessment method to estimate quantities important to management, 2) reviews existing information on possible environmental drivers of recruitment, 3) presents an 'oscillating control' model that links recruitment to changes in the Pacific Decadal Oscillation (PDO), 4) evaluates management strategies given recruitment

dynamics driven by the PDO and 5) explores the influence of spatial structure the snow crab population on assessment.

The estimation ability of the assessment method was determined using management strategy evaluation methods. The results of the evaluation allowed biases that may have influenced the ability of the management strategy to rebuild the stock to be identified. When all assumptions are correct, the assessment method provides good estimates of management quantities, including mature male biomass and the overfishing level. Data to reduce uncertainty around growth and natural mortality would be most beneficial to improving confidence in assessment estimates of mature male biomass.

The 1988/89 regime shift in the North Pacific appears to have shifted the drivers of recruitment dynamics from female spawning biomass to environmental influences related to the winter PDO (winds at Unimak pass and sea surface temperature were also important variables). This conceptual model was evaluated using cross-validation and retrospective analysis, both of which indicate that it is relatively robust to varying levels of information. A potential mechanism for this change in driver is the influence of temperature during regimes on the timing of the ice retreat and its associated consequences (e.g. food availability, length of the pelagic stage in snow crab, and size of cold pool) for the survival of pelagic larvae and juvenile crab.

A management strategy evaluation framework using three scenarios (status-quo, regime-based and ‘oscillating control’) for future recruitment was used to evaluate the performance of harvest control rules, including the status-quo, a regime-based approach, constant proportion strategies, and Hilborn’s ‘hold-steady’ method. In regime-driven systems, regime-based harvest control rules increase yield and decrease variability at the cost of a higher probability of overfishing. In non-regime-based systems, regime-based harvest control rules slightly decrease yield (with no change in variability), and increase the probability of overfishing under the definitions imposed by US legislation.

Finally, the spatial assessment developed to estimate spatially explicit recruitment, fishing mortality and movement was unable to accurately estimate movement parameters, but reasonably good fits to the data were achieved. The fitted spatial population dynamics model was used to simulate data sets to test the ability of a non-spatially structured (aggregate) assessment method to estimate trends in abundance and fishing pressure in the presence of spatial structure. The aggregate assessment method performed well if survey selectivity was well known, but was unable to reliably capture the dynamics of the population without data on survey selectivity. Additional data on movement from tagging studies or winter surveys would be useful to improve the understanding of the influence of spatial structure on snow crab in the eastern Bering Sea.

Contents

Table of Figures and Tables:

Acknowledgements

This dissertation would not have been possible without the patient guidance of André Punt. He went above and beyond the norm as an advisor and showed me how to be a scientist. I will always be thankful he allowed me to study under him. Ray Hilborn shaped the way I approach problems by reminding me often to think of the big picture and encouraging me to question the status quo. The expertise of David Armstrong and Lobo Orensanz in all matters of crab biology was deep and indispensable in trying to understand the population dynamics of snow crab. Jack Turnock showed me the ins and outs of daily life of a stock assessment scientist and was always willing to discuss ideas with me. Anne Hollowed gave me many opportunities to discuss my research at meetings both academic and governmental and always managed to excite me with her forward-thinking talks. I'm grateful to the North Pacific Research Board and Washington SeaGrant for funding this work. I would like to dedicate this dissertation to my wife, Eura, whose patient support buoyed me through the process, and my daughter, Evelyn, whose smiles have almost made me a morning person.

Chapter 1: Introduction

Avoiding overfishing is one of the over-arching goals of fisheries managers. Overfishing occurs when exploitation rates exceed the reference rates, one of which is the fishing mortality that would achieve the maximum sustainable long term yield. Stocks that experience extended periods of overfishing may be designated 'overfished' if their biomass is estimated to be beneath some pre-agreed upon threshold. The specifics of reference point systems used to set management targets vary, but the end goal is similar: to prevent the stock from decreasing to levels that would diminish incoming recruitment.

An overfished status indicates a breakdown of a management system's ability to achieve its specified goals. Whether the failure is in implementation of the management strategy or in the estimation of the components (such as biomass and optimal exploitation rate) used to implement that strategy, an effort must be made to understand the root causes. Management systems are composed of models of underlying population dynamics of the harvested species, the data collection scheme, the method of stock assessment and the harvest control rule. Many methods of management often do not consider the uncertainty associated with each part of a management system and neglecting to include these key sources of uncertainty can lead to erroneous catch recommendations. The Management Strategy Evaluation (MSE) approach involves developing an operating model that captures all major aspects of the system, incorporating uncertainty associated with each of those aspects and then evaluating the ability of the management system to achieve goals using simulations.

The Alaskan snow crab (*Chionoecetes opilio*) fishery in the eastern Bering Sea (EBS) generated $100,000,000 from 62 million pounds of crab in 2008, but revenue from the fishery has been as high as $199,000,000 (1994) and catch has been as high as 325 million pounds (1991) (http://www.st.nmfs.noaa.gov/st1/commercial/landings/annual_landings.html). The fishery has been quite lucrative, but estimated biomass has been highly variable (Figure 1a). Only males are taken in this fishery and there is no discernable relationship of recruitment to spawning biomass. This stock is managed jointly by the North Pacific Fishery Management Council and the State of Alaska (Turnock and Rugulo, 2009). It was designated to be overfished in 1999 and has since been under a rebuilding plan. However, that plan failed to rebuild the stock within the allotted ten-year period and was consequently declared a failure in 2009. In 2011, the stock was declared rebuilt. Large variability in recruitment is hypothesized to drive changes in abundance (Kruse et al., 2007), and potentially confounds the ability to achieve management goals using strategies which manipulate fishing mortality. Also, estimates of the parameters governing population processes (e.g. growth, mortality, and recruitment) have a considerable amount of associated uncertainty due to a lack of data.

Evaluating the role of uncertainty around these processes in the decision making process can be accomplished through MSE (e.g. Smith, 1999). The MSE framework is generally composed of three components: an operating model, an estimation model and a harvest control rule. The operating model describes the ‘true’ underlying population dynamics of the system for a given simulation, the estimation model represents a managers perception of that system (informed by surveys and catch records), and the harvest control rule is the process by which a manager decides on catch limits (which for EBS crab relate to the Overfishing Level, OFL, Acceptable Biological Catch, ABC and Annual Catch Limit, ACL, as well as the State of Alaska harvest control rules).

Simulations can also be used to quantitatively evaluate the estimation ability of the stock assessment. In these simulations, the operating model determines the true values for the parameters governing the system and generates the data used in assessment (survey indices, estimates of landed and discarded catch, and information on survey and catch length compositions, all with uncertainty) and the stock assessment method (or estimation model) analyzes those data to provide estimates for the quantities used by management in harvest control rules to determine acceptable removals (for example, mature male biomass, vulnerable biomass, fishing mortality, OFL). The bias and precision associated with these estimates can be determined by comparing the estimated quantities with the true values.

The MSE approach can go further and evaluate the performance of the entire management system. Simplified, the operating model projects virtual populations that the estimating/assessment model "observes" with error. Using these observations, the combination of the stock assessment model and the harvest control rules lead to a catch limit, which is applied (possibly with error) to the population generated by the operating model, which is then 'observed' again. This cycle continues for the desired number of years, for the desired number of simulations, at the end of which, the collection of performance statistics from the operating model can be used to evaluate how often management goals are achieved. This cycle can be repeated for different operating models, stock assessment methods and harvest control rules to determine the performance of each in relation to the others.

An MSE for EBS snow crab is both timely and appropriate for several reasons. First, in some aspects, the Bering Sea is a relatively data rich system. Data for EBS snow crab are available spanning 1978-present from the summer National Marine Fishery Service (NMFS) survey that estimates the abundance, sex ratio, maturity state and location of the EBS snow crab population. Catch data from the directed fishery and the trawl fishery that extend to 1978 are also available. Second, in spite of being overfished, the snow crab fishery still earned an average of nearly 25% of the total earnings of all fisheries in the EBS in analyzed years during the rebuilding period from 1999-2009 (Garber-Yonts, 2009). Financial losses through poor management are consequently potentially very high. Finally, the red king crab (*Paralithodes camtschaticus)* fishery around Kodiak Island and the blue king crab (*Paralithodes platypus)* fishery around the Pribilof Islands collapsed nearly 25 years ago and neither has recovered appreciably, in spite of the absence of directed fishing (Pengilly et al, 1995). The potential for irreversible effects caused by extended periods of high fishing pressure underscores the urgency with which an evaluation of the management strategy for the snow crab fishery must be undertaken.

The key objectives of this dissertation are therefore to:

- explore how well it is possible to estimate quantities important in management (recruitment deviations, mature male biomass, etc.) given uncertainties regarding population processes (Chapter 2);
- build a model for snow crab recruitment incorporating environmental covariates to inform management strategies (Chapter 3);
- examine the ability of the management system to achieve management goals for the EBS snow crab, in particular given a changing climate (Chapter 4); and
- examine the performance of a non-spatial management system in the face of spatial population structure (Chapter 5).

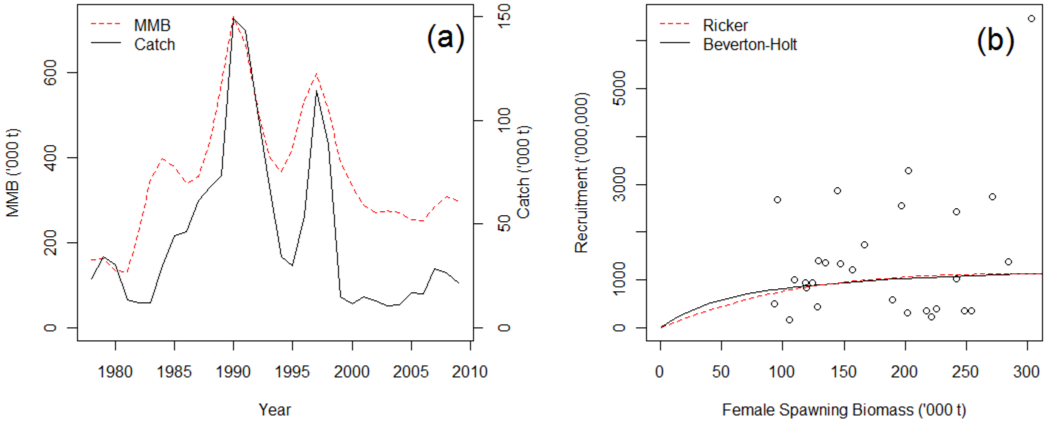

Figure 1.1. Snow crab catch ('000t) and estimated mature male biomass ('000t) in the eastern Bering Sea (panel a). Estimated spawner ('000t) – recruit (000,000; lagged 5 years) curves for snow crab in the eastern Bering Sea (panel b).

Chapter 2: Identifying research priorities for management under uncertainty: The estimation ability of the stock assessment method used for eastern Bering Sea snow crab (*Chionoecetes opilio*)

2.1. Introduction

Stock assessments aim to use the best available science to provide estimates of the biomass of a target species and then calculate acceptable removals based on a chosen harvest control rule. However, the best available science is often incomplete and/or uncertain. It is necessary to prioritize which research should be conducted when there are multiple uncertainties in a system and limited resources to address these uncertainties. Research priorities are often enumerated anecdotally (e.g. Jennings and Zigler, 2000), but cost-benefit analyses that evaluate the expected returns of a specific research project (e.g. a fishery-independent survey) are also found in the literature (e.g. McDonald et al, 1995; Powers and Restrepo, 1993). Simulation can be used to explore the impact of imperfect knowledge regarding the values of population parameters on the ability to estimate quantities used in management when experimental studies to determine the true parameter values are absent, infeasible and/or undesirable (Butterworth and Punt, 1999, Punt 2003b). In a simulation framework, perfectly-known (or ‘true’) simulated populations are produced and analysed using a stock assessment. The impact of incorrect assumptions can be quantified by evaluating the degree to which the ability to estimate the management quantities is affected. This approach differs from varying the values of the parameters of a stock assessment in a set of sensitivity tests because most of the parameters of a stock assessment are not specified by the analyst but are rather estimated by fitting a population dynamics model to a data set. Although simulation studies should be conducted for all modelling endeavours, they are often not. An exercise using simulations to identify research priorities when there are multiple sources of uncertainty is presented here for the snow crab (*Chionoecetes opilio*) fishery in the eastern Bering Sea (EBS).

The U.S. domestic fishery for EBS snow crab is all-male and historically productive, with a maximum estimated biomass of over 680,000 mt in the early 1990s (Turnock and Rugulo, 2009). However, both biomass and catch have been quite variable since the start of the domestic fishery, and EBS snow crab was declared overfished in 1999 when it was assessed to have dropped below its Minimum Stock Size Threshold (MSST) (Turnock and Rugulo, 2011). As required under the U.S. Magnuson-Stevens Fishery Management and Conservation Act, a plan was developed to rebuild the stock to the target biomass (a proxy for B_{MSY}, defined as the biomass corresponding to a fishing mortality of $F_{35\%}$, the fishing mortality rate which reduces spawning biomass-per-recruit to 35% of its unfished level) within 10 years (NPFMC/NMFS, 2000). Stock assessment estimates of spawning biomass (and later mature male biomass, MMB) were used to calculate levels of fishing mortality that would return the stock to the target biomass and these levels of fishing mortality were used to set Total Allowable Catches. This plan was, however, declared a failure in September 2009 (Turnock and Rugulo, 2009), and the inability to achieve management goals calls into question the efficacy of the management strategy used for this stock and also the stock assessment on which management advice has been based. (The stock was declared rebuilt in 2011, but this was longer than the specified time frame of ten years.)

The EBS snow crab stock is managed jointly by the North Pacific Fishery Management Council (NPFMC) and the state of Alaska. Each year, data from the directed pot fishery, the National Marine Fisheries Service (NMFS) summer trawl survey and bycatch data from the groundfish trawl fishery are used to estimate trends in fishing mortality, numbers-at-length, and MMB using a length-structured population dynamics model (Turnock and Rugulo, 2009). The estimate of MMB is used in conjunction with a catch control rule (equations 3.1-3.3) to determine the overfishing level (OFL) for the coming year (NPFMC/NMFS, 2007). The OFL is used to determine an Acceptable Biological Catch (ABC), which is less than the OFL to account for scientific uncertainty. A total allowable catch (TAC), to be divided among quota holders, is finally determined by the State of Alaska, and must be less than or equal to the ABC.

	Stock status level	F_{OFL}
(2.1)	$\frac{B_{current}}{B_{35\%}} > 1$	$F_{\text{OFL}} = F_{35\%}$
(2.2)	$\beta < \frac{B_{current}}{B_{35\%}} \leq 1$	$F_{\text{OFL}} = F_{35\%} * \frac{\frac{B_{current}}{B_{35\%}} - \alpha}{1-\alpha}$
(2.3)	$\frac{B_{current}}{B_{35\%}} \leq \beta$	Directed fishery F = 0

where α determines the fishing mortality used to compute the OFL, F_{OFL}, when biomass decreases to $\beta * B_{\text{MSY}}$ and β determines the threshold level of biomass at or below which directed fishing is prohibited. $B_{current}$ is the estimated MMB for the current year and $B_{35\%}$ is a proxy for the biomass at which maximum sustainable yield is achieved, B_{MSY}.

The currently-used catch control rules cannot be expected to perform adequately without a reasonably accurate and precise estimate for $F_{35\%}$, the time-series for MMB and the B_{MSY} proxy. The current stock assessment method consists of a length-based population dynamics model that accounts for sex, shell condition and maturity state, with an associated method for parameter estimation (Turnock and Rugolo, 2009). The processes modelled (e.g. growth, mortality, selectivity, catchability) in the assessment have considerable associated uncertainty because of a lack of data on snow crab biology and demographics. It is prudent to examine how well one would expect to be able to estimate the parameters used to predict quantities important to management given current levels of uncertainty around the parameters governing the biology, exploitation and monitoring of the EBS snow crab resource. A better understanding of the relative magnitude of impact on estimation performance caused by uncertainty will allow for appropriate allocation of finite management, research and monitoring resources.

Length-based stock assessments are used world-wide to model the population dynamics of stocks for which aging is difficult (e.g. American lobster in Maine (e.g., Chen et al. 2005), rock lobster in New Zealand and Australia (e.g., Hobday and Punt, 2001) and red king crab in Alaska (e.g., Zheng and Siddeek, 2010)). In addition to guiding future research directions for snow crab management, examining the impact of uncertainty on the estimation ability of size-based methods may identify general conclusions that could be useful for assessments of other stocks with size-based assessment methods.

Three categories of uncertainty are considered in this analysis: mis-specification of parameters, the amount or quality of data, and assumptions in the assessment related to fishing mortality. Questions related to mis-specification are: 1) what is the impact of uncertainty in growth on the ability of the EBS snow crab assessment to estimate MMB and other management-related quantities, 2) how does the assessment method perform when catchability and selectivity for the NMFS survey differ from the currently estimated values, and 3) how does assuming natural mortality at an incorrect value affect estimation ability? Changes in assumptions about quality and quantity of data will be used to answer the questions: 1) what are the benefits of additional growth data, and 2) how does the quality of survey data affect estimation ability? The final question (which addresses how the assessment is specified) examines how well assessment method is able to identify an overfished stock.

These questions (see below for further detail and rationale for their selection) have been selected because they are either contentious topics of discussion at management meetings between the fishing industry and assessment scientists (e.g. survey selectivity and natural mortality), there are few data available on the process (e.g. growth and natural mortality), the process has a large bearing on the ability of the management system to reach its goals (e.g. assumptions about fishing mortality), or because the process is one controllable by management (e.g. data quality in the survey). There are, of course, other questions (such as the values assigned to the factors used to weight different data sources) which could be addressed; those considered in the paper were selected because there is no straightforward method to explore them within the assessment process directly (e.g. sensitivity of model outputs to values of weighting factors). Answering these questions will aid in the process of identifying research priorities, be they experimental studies to more accurately determine parameter values, increased survey effort, or revision of the stock assessment model.

2.2. Methods

2.2.1 Overview

Briefly, the steps used to conduct the simulation study to identify research priorities here are: 1) Identify processes with uncertainty that may influence estimation ability, 2) Identify possible values for the quantities or parameters involved from the available literature, 3) Develop a simulation framework, and 4) Rank processes according to relative impact on the ability to estimate a quantity useful in management (e.g. the OFL). The word “process” is used loosely in this context. “Process” can mean not only the biological processes modelled in the assessment, but also the activities related to management (e.g. data collection or designation of allowable fishing mortality).

The processes to be examined in this study are those involved in the stated questions of interest. The Monte Carlo simulation framework used to address these questions is commonly used in the literature (e.g. Punt et al., 2002; Chen and Wilson, 2002; Kanaiwa et al. 2005; see Figure 2.1 for an overview) and is recommended over sensitivity testing (Hilborn and Walters, 1992; NRC, 1998). This simulation framework has two main components: an operating model and an estimation model. The operating model provides the 'truth' for a given simulation (which can change depending on the question being addressed) and is based on the population dynamics model that currently forms the basis for the stock assessment (See Appendix 2.A for details). Virtual populations are generated using the operating model, and the data necessary for assessment purposes are 'collected' from these populations with observation error. The data used for assessment purposes are numbers and length frequencies of crab caught in the directed pot fishery, the trawl fishery (as bycatch) and the NMFS summer survey.

The estimation model produces parameter estimates using the data 'collected' from the operating model, and represents a manager's perception of the status and productivity of the stock (i.e. produces estimates of mature male biomass and other quantities on which management advice is provided). In this study, the estimation model (see Appendix 2.B for the objective function used for parameter estimation) is very similar to the actual assessment method used to provide management advice for the EBS snow crab stock. The performance of the estimation model is evaluated in terms of the relative errors (REs) (Equation 2.4) of the estimates of key derived quantities (i.e. mature male biomass (MMB), the fishing mortality that will be used to compute the OFL (F_{OFL}), the fishing mortality that will result in a MMB that is 35% of virgin MMB ($F_{35\%}$), the MMB corresponding to the proxy for maximum sustainable yield (B_{MSY}) (i.e. the product of the MMB-per-recruit corresponding to $F_{35\%}$ multiplied by average recruitment from 1978 to 2009), and the overfishing level for the coming fishing year (OFL)). The full set of quantities (rather than just MMB or the OFL) are included to identify how error propagates through the estimation method. Moreover, these are all quantities which are reported to the decision makers.

$$E_t^{i,j} = (\hat{Q}_t^{i,j} - Q_t^{i,j})/Q_t^{i,j} \tag{2.4}$$

where $E_t^{i,j}$ is the relative error for quantity *i* during year *t* for simulation *j*, $Q_t^{i,j}$ is the true (i.e., based on the operating model) value for quantity *i* during year *t* for simulation *j*, and $\hat{Q}_t^{i,j}$ is the estimate of quantity *i* during year *t* for simulation *j*. The ratio of MMB to B_{MSY} is calculated for each scenario to explore how uncertainty in the processes of interest affects the ability to estimate relative versus absolute biomass. Relative errors and inter-simulation intervals for a given management quantity are presented in a common format, i.e. “median RE% [5% bound, 95% bound]”; in some instances, the lower and upper bounds are not presented. Abbreviations used for the various scenarios are listed in Table 2.1.

A different parameterization of the operating model is specified to address each of the questions. For each question, a population process is identified, and the parameters associated with that process are pre-specified. The parameters are pre-specified so that any estimation errors are the consequences of incorrect assumptions related to the process and not variation in the process. One hundred vectors of the remaining estimable parameters are then sampled from the posterior distribution obtained by fitting the operating model to the actual data for EBS snow crab using a Markov chain Monte Carlo (MCMC) algorithm given those pre-specified values. This involved conducting 1,000,000 cycles of the MCMC algorithm, implementing a 10% burn-in and selecting a thinning ratio so that 100 vectors are selected within the chain. Generating parameter vectors in this manner allows a parameter space that is consistent with both the actual fishery data and the question of interest to be represented (similar to Kanaiwa et al., 2005). Several diagnostic statistics (e.g. checking for lack of autocorrelation and calculating Geweke statistics) were used to check for evidence of convergence of the MCMC algorithm.

Each parameter vector is then used to produce a virtual population from which the data necessary for assessment purposes are generated. The estimation model differs from the operating model only in the assumptions about the process under examination and is then applied to the 'collected' data from each Monte Carlo replicate. This process is repeated for each parameter vector; the resulting 100 series of MMB estimated from the application of the estimation method and the other quantities of management importance are compared to the true series from the operating model, and measures of bias and precision are obtained.

The following sections outline how the operating model is specified to address each question. The operating models are all variants of a reference operating model which matches the specifications of the 2009 assessment.

2.2.2 Mis-specification of parameters

2.2.2.1 Survey selectivity

Analysis of tows conducted by the Bering Sea Fisheries Research Foundation (BSFRF) suggest that survey selectivity increases slowly and smoothly with size and asymptotes to values less than one (Somerton et al. 2010) and that survey catchability, q (the probability of capture for the most selected crab) is much lower than currently estimated. In contrast, the currently-estimated survey selectivity pattern is almost knife-edged with catchability close to 1. The ability to estimate survey selectivity within the stock assessment is examined by conducting simulations for scenarios in which survey selectivity for the most-recent group of years over which survey selectivity is assumed constant[1] (1989-present) is set to the curve estimated by Somerton et al (2010) and in which selectivity in all eras is set to Somerton's curve. There are three "eras" (referred to as "era 1", "era 2" and "era 3") which correspond to 1978-1981, 1982-1988 and 1989-present, respectively. In both scenarios, the estimation method estimates all selectivity parameters for all eras.

2.2.2.2 Natural mortality

Natural mortality (*M*) for mature males is specified in the assessment as 0.23yr^{-1} (Turnock and Rugolo, 2009) based on the approach of Hoenig (1983) which extrapolates natural mortality from maximum observed age. Although *M* for snow crab has been the subject of several studies (Somerton, 1981; Zheng, 2003), all of them have provided markedly different estimates (0.13-1.07 yr^{-1}), although the range of *M* estimates is considerably smaller when the assumption of a terminal molt is taken into account. The impact of incorrectly specifying natural mortality is assessed by fixing the true (i.e. operating model) value for natural mortality at 0.23yr^{-1} and then applying the estimation model when *M* is assumed to be +/- 50% of the true value. These scenarios will be referred to as "M+50%" and "M-50%" hereafter.

---

[1] Such a group of years will henceforth be referred to as an 'era'. Eras are selected given changes to the gear used by the survey.

2.2.3 Amount of data quantity or quality

The amount and quality of data available to the stock assessment method has the potential to impact estimation ability. Each of the data sources for the assessment have associated measures of uncertainty. A change in the assumptions about the quantity or quality of data for any of these sources would change the associated uncertainty, and this would impact the uncertainty around model estimates. In particular, data associated with the survey (e.g. coefficients of variation around area-swept estimates of biomass and length frequencies) and growth (i.e. the number of observations used to develop the growth curve) appear to have the highest associated uncertainty. These two processes are the focus of the simulations directed at determining the impact of the quality of data on estimation ability.

2.2.3.1 Survey data

The coefficients of variation, CV, around the survey data were increased and decreased by 0.08 in turn when generating data for assessment purposes to explore the relationship between model estimates and the precision of the survey data while the effective sample size used for the survey length data was held constant (abbreviated “CV+” and “CV-“, respectively). The fixed amount 0.08 was chosen because the smallest CV around the survey estimates is 0.09. Analyses were also conducted in which the effective sample sizes used to determine survey length frequencies were halved and doubled (from 200 to 100 and 400, and abbreviated “CV EffN-” and “CV EffN+“, respectively).

2.2.3.2 Growth

One of the key components of a length-based model is the size-transition matrix that describes how a crab grows each year after molting as a function of its current size. Until the 2010 assessment, growth was estimated outside of the assessment method and growth parameters pre-specified in the assessment (the assessment authors are currently experimenting with different ways of estimating growth). The size-transition matrix used in the 2009 assessment was based on only 14 observations of pre- and post-molt lengths taken from crabs collected during a NMFS survey (Rugulo, NMFS, unpublished) and the 2010 assessment used inferences from these 14 observations as the basis for a prior on growth rates. Error in the size-transition matrix has been demonstrated to lead to poor estimates of management-related quantities when those quantities are estimated using a length-based method of stock assessment (Punt, 2003a).

The impact of using only 14 data points to determine the size-transition matrix is assessed by randomly generating 14 pre- and post-molt lengths using equations 3.A.6-3.A.8 for each Monte Carlo replicate. These 14 data points are representative of the entire population (i.e. not only large males are sampled). The growth observations provide the data used to estimate the parameters of the size-transition matrix (intercept and slope of the growth equation and shape and scale parameter of the gamma distribution that describes the variance in growth increments; Eqn 2.A7a-b). The estimation method underlying each Monte Carlo replicate therefore has a unique size-transition matrix. Essentially, a virtual mark-recapture study is performed for each Monte Carlo replicate to define the growth curve to be used in the estimation model. Scenarios in which more length observations (25 and 100) are used to define the size-transition matrix are used to assess the utility of additional growth data.

2.2.4. Fishing mortality

It is important that an assessment method is able to detect when a stock is overfished given the need to achieve management goals. To evaluate the estimation model's ability to detect changes in fishing mortality, average fishing mortalities (catch, discard and trawl) are fixed at the values estimated from the current assessment method and vectors for the remaining estimable parameters are drawn from the posterior of the resulting model using MCMC. Average fishing mortality is then multiplied by a constant (i.e. 0.5, 2) from 1995 onward, data for assessment purposes are generated, and the estimation model is applied to the generated data. The method's ability to detect an overfished stock is summarized by the fraction of times the stock is correctly assessed to be above and/or below the true MSST (0.5 $B_{35\%}$). The stock assessment utilizes penalties on deviations around average fishing mortality in the objective function to constrain the inter-annual variation and overall magnitude of estimated fishing mortality. Analyses were conducted in which the penalties were eliminated, quartered, halved and doubled to assess their impact on parameter estimation and the ability of the assessment method to identify an overfished stock.

2.3. Results

2.3. 1 Reference scenario

Results are first shown for the 'reference' scenario in which the assumptions of the stock assessment are not violated. This scenario provides an impression of how well this method can be expected to perform in principle. For this scenario, M and the size-transition matrix are assumed to be known exactly, and data are generated based on sampling standard deviations and effective sample sizes which mimic those assumed when actually applying the stock assessment.

Mature male biomass (MMB) was relatively well estimated with an average median RE of 1% [-16%, 26%] over all years (Figure 2.2a). Variability in estimated MMB translates to a range relative errors for the OFL (1% [-38%, 52%]), F_{OFL} (6% [-16%, 26%]), and B_{MSY} (0% [-6%, 6%]) (Figure 2.2b). Recruitment was slightly negatively biased on average and relatively uncertain (-4%

[-39%, 49%], Figure 2.2c). On average, median estimated fishing mortality was positively biased by 7% [-20%, 35%] (Figure 2.2d).

Estimated selectivity parameters describing the length at which 95% (Sel 95) and 50% (Sel 50) of the fish were selected by the survey gear in the first era (1978-81) were estimated poorest of all eras (median relative errors of -10% for Sel 95 and 9% for Sel 50). Sel 95 in the second era (1982-88) was slightly biased, but Sel 50 and survey catchability were well estimated (-4% for Sel 95, 0% for Sel 50 and 0% for q (catchability)). Selectivity for era 3 (1989-present) was well estimated, with a median RE of 0 for all quantities and small inter-simulation intervals (Figure 2.3a).

2.3.2 Mis-specification of parameters

2.3.2.1 Natural mortality

Simulations in which natural mortality, M, differs from the true value by ±50% had the largest increases in bias and loss of precision in management quantities of all scenarios considered (Table 2.1). When M is negatively biased, all management quantities are negatively biased, except MMB and B_{MSY}, which are positively biased (Figure 2.4a-c). The opposite trend is apparent when M is positively biased (Figure 2.4d-f).

2.3.2.2 Survey selectivity

Fixing survey selectivity in the operating model at Somerton's estimated curve for 1989-present (referred to hereafter as the "Sel era 3" scenario) resulted in large biases in management quantities; MMB was negatively biased (-14%) (Figure 2.5a), as was the OFL (-16% [-54%, 36%] (Figure 2.5b) and recruitment was noisier to compensate (Figure 2.5c). Fixing survey selectivity in the operating model for all eras at Somerton's curve (referred to hereafter as the "Sel all era" scenario) led to a slightly larger impact on bias in MMB (-19%) and the OFL (-17%) (results not shown because of similarity to the "Sel era 3" scenario).

The slope of the survey selectivity curve for era 1 was estimated reasonably well for the "Sel era 3" scenario (Figure 2.5d). The estimated survey selectivity curve for era 2 was beneath the true curve (Figure 2.5e), and the median estimated selectivity curve for era 3 was above the true curve (Figure 2.5f) for this scenario. The median estimated curve for eras 1, 2 and 3 are respectively always beneath the true curve at larger sizes, reasonably close to the true curve and above the true curve for the 'Sel all era" scenario (results not shown).

2.3.3 Amount of data quantity or quality

2.3.3.1 Growth

Estimating growth for each Monte Carlo replicate based on 14 observations of pre- and post-molt lengths (instead of assuming perfect knowledge of the growth curve as in the reference scenario) led, as expected, to estimates of MMB with a wider range of relative errors and a positive bias on average (e.g. 8% [-23%, 106%]) ("Growth-14"; Figure 2.6a). Increasing the number of observations used to estimate growth to 25 ("Growth-25") and 100 ("Growth-100") observations resulted in narrower relative error distributions and a decrease in bias in MMB (25 observations: 7% [-23%, 56%], 100 observations: 3% [-19%, 35%]) (Figure 2.6a). The trend of decreasing ranges of relative error for MMB was mirrored in the estimates for the parameters for the growth equation, OFL, B_{MSY} and F_{OFL}. The estimates of B_{MSY} , OFL and F_{OFL} were positively biased when there was uncertainty associated with growth (3%, 18% and 10% respectively for the 14-tag scenario; Figures 2.6b-d), but these biases reduced with added data points for growth.

2.3.2 Survey data

The biases of the estimates of the management quantities are not sensitive to the CVs for the survey biomass estimates, although inter-simulation variation was slightly less for lower CVs. However, halving the effective sample sizes for the survey length-composition data when also increasing the CVs for survey estimates ("CV EffN-") had a negative impact on estimation ability (Figure 2.7a-b). The inter-simulation intervals of all management quantities were smaller when the effective sample sizes were doubled and the CVs decreased ("CV EffN+") than when the effective sample sizes were halved and the CVs increased. However, estimates for the "CV EffN-" scenario had a larger bias and loss of precision than the estimates from the scenario in which only the CV was increased (Figure 2.7a-b). Simply put, better survey data leads to better estimates of MMB and management quantities, with length frequency data being more important than the survey CV.

2.3.4 Assessment assumptions--Fishing mortality

Operating models were constructed by multiplying average fishing mortality (F_{mort}) by 0.5 ("Fx0.5" in Table 2.1), 1 and 2.0 ("Fx2.0" in Table 2.1) from 1995 onwards. The estimation model was then applied to the resulting data sets under the assumption that average fishing mortality had not changed. A decrease in absolute bias in MMB was seen with an decrease in the multiplier for average F_{mort} from 1995 (0.5 x F_{mort}: 0%[-17%, 27%], 1 x F_{mort}: 1%[-17%, 26%], 2 x F_{mort}: 5%[-13%, 28%]). Larger penalties applied to the fishing mortality deviations resulted in larger biases and narrower ranges of relative errors around MMB, while less severe penalties led to lower biases and less precision, with the scenario eliminating penalties altogether leading to the least bias in the management quantities (Figures 2.8 and 2.9).

Biases in management quantities translated to average false positive rates (cases in which the estimation model reports the population as overfished ($B_{current} < 0.5B_{\mathrm{MSY}}$), but the true population is not overfished) of 0.3%, 5.5%, and 1.2%, and average false negative rates (incorrectly identified overfished stocks as 'healthy') of 0%, 0.75%, and 5.4% for the 0.5 $F_{\mathrm{mort},}$ 1 F_{mort}, and 2 F_{mort} cases, respectively (Figure 2.10a-c). In the 1 x F_{mort} scenario (which is very similar to the reference case), more severe penalties result in lower false negative rates, but a large increase in the number of false positives (Figure 2.10b,e). The incidence of a false negative in the first year in which the stock is actually overfished in the 2 x F_{mort} scenario dropped from nearly 100% under the current penalties (Figure 2.10c) to approximately 25% when no penalties were applied (Figure 2.10f).

2.4. Discussion

The suite of estimated management quantities and parameters presented are useful for diagnosing the causes of biases and imprecision, but the overfishing level (OFL) is ultimately the quantity that is used in management decisions. An idea of the assumptions in the stock assessment that, when violated, have the largest potential impacts on the management of the fishery, and should hence be the focus for additional research, can consequently be formed through comparison of the relative errors in the OFL among scenarios. The ratio of the current biomass to B_{MSY} is also important because it defines whether a stock is considered to be overfished according to the US Magnusson-Stevens Act. B/B_{MSY} is used when calculating the OFL, so uncertainty in B/B_{MSY} is also reflected in the OFL. Smaller biases and losses of precision and less difference among scenarios appeared for B/B_{MSY} than for the OFL. Hence, in spite of the importance of B/B_{MSY} for management, discussion will focus on the OFL.

The OFL is reasonably close to unbiased (median RE=1%) when all of the assumptions are correct (Table 2.1). The OFL is estimated relatively imprecisely, but this is expected because the OFL is a function of several estimated quantities (F_{OFL}, current biomass, B/B_{MSY}). Processes ranked in order of decreasing median absolute RE (MARE) for the OFL when the associated assumptions were violated as described in this study are listed in Table 2.1. Rankings may have been different had the assumptions been violated differently, but Table 3.1 still provides a way to synthesize a large number of effects. The results from each category of question asked will now be considered briefly with a view toward identifying possible research priorities.

2.4.1 Scenarios

2.4.1.1 Parameter estimation

Uncertainty in natural mortality, *M,* led to the largest changes in the MARE for the OFL. The OFL is positively biased when the assumed value for *M* exceeds to the true value ("M +50%", Table 2.1) even though this results in a negatively-biased estimates of MMB. This result arises because of the interplay and relative magnitude of bias in recruitment, current biomass, $B_{\mathrm{MSY},}$ and $F_{35\%}$. Spawning biomass-per-recruit is negatively biased when natural mortality is assumed to be higher than it really is. This negative bias translates to a positive bias in $F_{35\%}$, which, in this case, was greater than the negative bias in MMB. Attempts to better incorporate this uncertainty into the assessment process and further studies to reduce uncertainty in *M* should be undertaken given the practice of specifying *M* rather than estimating it.

Incorporating the uncertainty associated with basing the size-transition matrix on a small sample resulted in a 4% increase to the MARE ("Growth-14"; Table 2.1). With only 100 samples (“Growth-100”; Table 2.1), the median RE for the OFL decreased from 18% to 7% (and the change in MARE from the reference scenario decreased to 1%). The precision of the OFL also began to approach that for the reference scenario. Considering the potential improvements in bias and precision, more accurately specifying the growth curve using an increased number of pre- and post-molt observations may be an efficient way of increasing confidence in the estimates produced by the stock assessment. Simulation could be useful in the experimental design for additional samples. Collection of data to better determine the growth parameters for snow crab began in 2011 in the EBS by the BSFRF.

The parameter estimation scenario in which selectivity was fixed at Somerton’s curve during 1989-present (“Sel era 3” in Table 2.1) led to large differences in the MARE of the OFL compared to the reference scenario. The observed differences between the estimated and true selectivity curves can explain the difference in MARE. Selectivity in the final era is overestimated, resulting in negatively-biased estimates of MMB, which in turn result in negatively-biased OFLs. A similar pattern is seen in the scenario in which the survey selectivity curves for all eras are set to Somerton’s curve (“Sel All Era”). Selectivity for the first era is underestimated, that for the second era is relatively well estimated and that for the third era is overestimated. Although a suitable explanation for this pattern is not immediately forthcoming, the fact that the assessment method is unable to recognize shifts in survey selectivity emphasizes the value of estimating survey selectivity using field data (to the extent that this possible). Starting in 2011, additional BSFRF data collected from side-by-side towing experiments with the NMFS survey using a nephrops trawl (which is assumed to have a survey catchability of 1) and the standard NMFS nets (Somerton et al., 2010) were used to inform survey selectivity and catchability in the stock assessment.

2.4.1.2 Data quality

The benefits of additional growth data have already been mentioned and additional growth data have the potential to improve model estimates greatly. Assuming the CV and effective sample sizes from the survey are better than they actually are leads to much larger variance in the estimated OFLs (resulting in a large MARE), but the median bias is small. The other survey data scenarios led to very little change in MARE from the reference scenario.

2.4.1.3 Assessment assumptions

The scenario in which average fishing mortality was multiplied by 2 from 1995 ("F x 2.0" in Table 2.1) led to the largest change in the MARE for the OFL compared to the reference scenario. Although this scenario is not likely in reality, it is instructive in evaluating the ability of the assessment method to detect overfished stocks. In the first year in which the simulated stock was overfished in this scenario, MMB was much less than MSST, but the assessment method was unable to identify a change in overfished status with penalties on fishing mortality in place. Detecting that MMB is less than MSST soon after MMB drops below MSST is very difficult when penalties are placed on changes on deviations in fishing mortality because the assessment method does not allow the large changes in fishing mortality to match the change in observed abundance. Consequently, penalties on fishing mortality can make achieving management goals more difficult when the stock is well below MSST. However, if MMB is just below the MSST on average ("F x 1.0" scenario; figures 2.10b, e), fishing mortality penalties reduce the number of overfished stocks that are assessed as healthy. Hence, there is a trade-off between incorrectly identified overfished and incorrectly identified 'healthy' stocks, with more 'healthy' stocks being identified as overfished (particularly in the years after the stock recovers to above MSST) with increased penalties.

Although decreasing (and ultimately eliminating) penalties generally resulted in the least amount of bias in management quantities for all fishing mortality scenarios, estimates varied more among simulations when the penalties were weaker. This is because penalties add artificial certainty around the estimates of deviations in fishing mortality (which then propagates through to the management quantities). Consequently, the selection of penalties on fishing mortality should be made after considering the associated risk, the history and variability of fishing pressure, and perceptions regarding biomass in relation to the reference biomass.

2.4.2 Conclusions and caveats

Considering the biases in OFL across scenarios and the existing body of data related to the processes considered here, future field research directed at decreasing uncertainty around growth and natural mortality would be most valuable to improve the estimation ability of the model. More accurately specifying the growth curve using an increased number of pre- and post-molt observations may be an efficient way of increasing confidence in the estimates produced by the stock assessment (the BSFRF has since begun collection of immature, molting crab). Results from the selectivity scenarios suggest that time spent incorporating all data available on selectivity into the stock assessment method (e.g. Somerton et al [2010], which is now used) could have large impacts on estimated management quantities.

This study has focused on EBS snow crab, but the results have implications for other assessments based on length-structured models. Our results support Punt (2003a) in demonstrating an incorrectly specified size-transition matrix (from small sample sizes or spatial heterogeneity in growth) can have large impacts on estimation ability. Potentially less specific to size-based models, penalties placed on deviations in fishing mortality were found to introduce bias and undue precision to management quantities and should therefore be carefully considered before implementation. Finally, relatively large relative errors around the OFL in the reference scenario emphasize the importance of precautionary management strategies.

Although our results are expressed in terms of how well quantities of management interest are estimated, this does not directly determine the implications on the ability to satisfy management goals. These implications can, however, be examined using Management Strategy Evaluation, MSE, (e.g. Punt and Smith, 1999; Smith et al., 1999; A'Mar et al., 2009a,b; Chapter 5).

Scenario	**MARE**	**MeARE**	**RE**	**5%**	**95%**
Reference	22%	23%	1%	-38%	52%
	(+/- reference)		(actual values)		
Mis-specification of parameters					
M-50%	23%	19%	-45%	-68%	-10%
M+ 50%	7%	14%	29%	-14%	91%
Growth-14	4%	12%	18%	-31%	93%
Sel era 3	3%	3%	-16%	-54%	36%
Sel all era	2%	3%	-17%	-52%	32%
Data quality/quantity					
CV EffN+	4%	9%	-1%	-46%	83%
CV+	2%	2%	1%	-40%	58%
CV EffN-	0%	3%	-3%	-45%	61%
CV-	-2%	-1%	-1%	-33%	49%
Growth 25	2%	6%	10%	-34%	81%
Growth 100	1%	3%	7%	-32%	59%
Assessment assumptions					
Fx2.0	7%	11%	12%	-40%	83%
Fx0.5	-2%	-1%	-1%	-33%	48%

Table 2.1. Median absolute relative errors (MAREs) and mean absolute relative errors (MeAREs) for the OFL in descending order for each scenario considered (left two columns) with average relative errors (RE) and 90% inter-simulation intervals for RE. Values for all MAREs and MeAREs, except those for the reference case, are reported as the difference between the MARE or MeARE for a given scenario and that for the reference case (i.e. if there was no change in MARE for a given scenario, the value in the MARE column would be zero for that scenario).

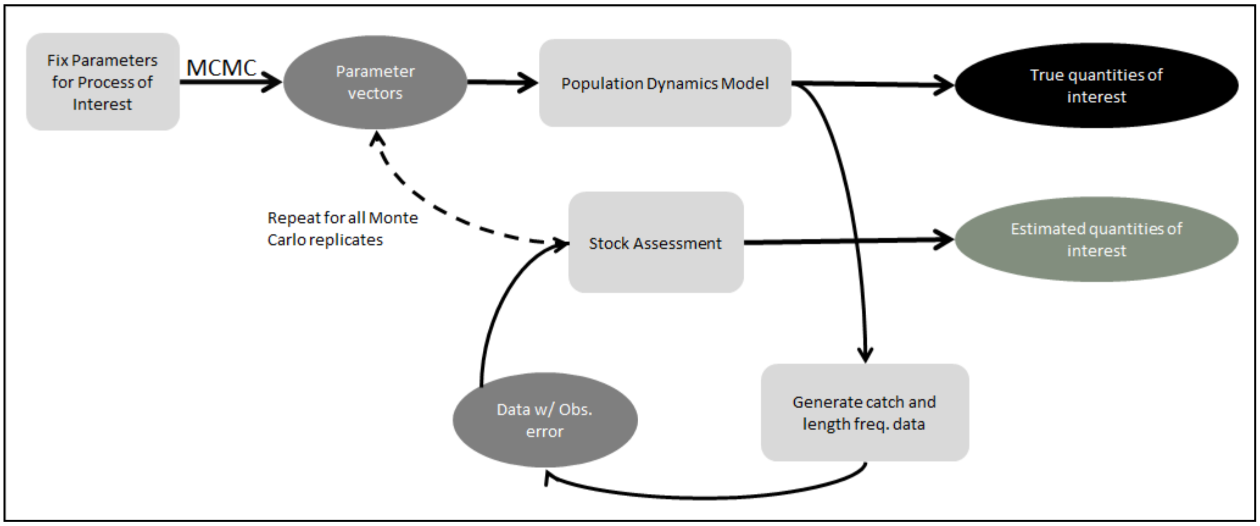

Figure 2.1. Flow of information in assessing the estimation ability of a stock assessment method.

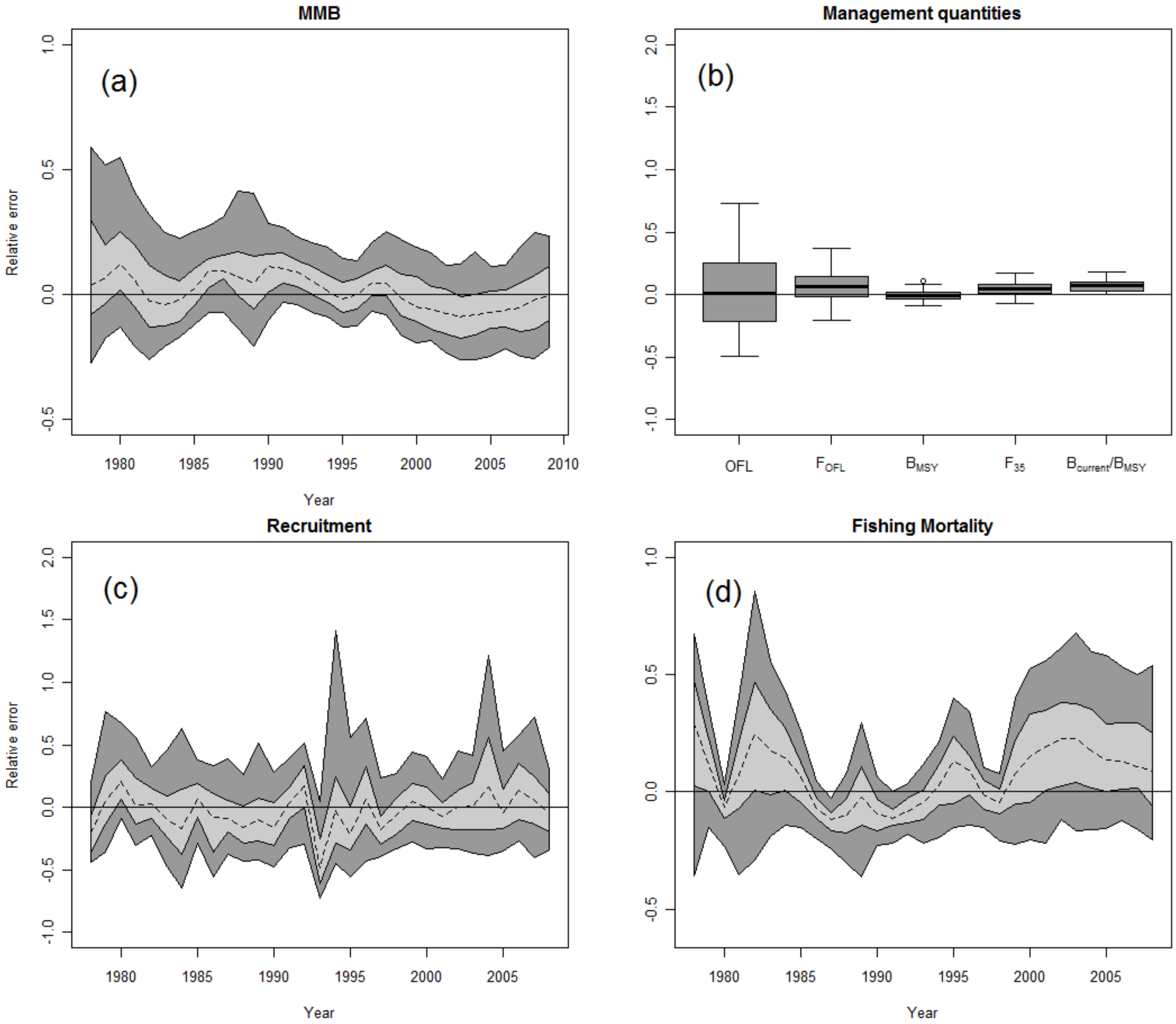

Figure 2.2. Relative error distributions for mature male biomass, recruitment and fishing mortality in the directed pot fishery for the reference scenario (median: dashed line; 5th and 95th quantiles: dark grey; 25th and 75th quantiles light gray) (a, c, d). Boxplots of relative error for key management quantities (b).

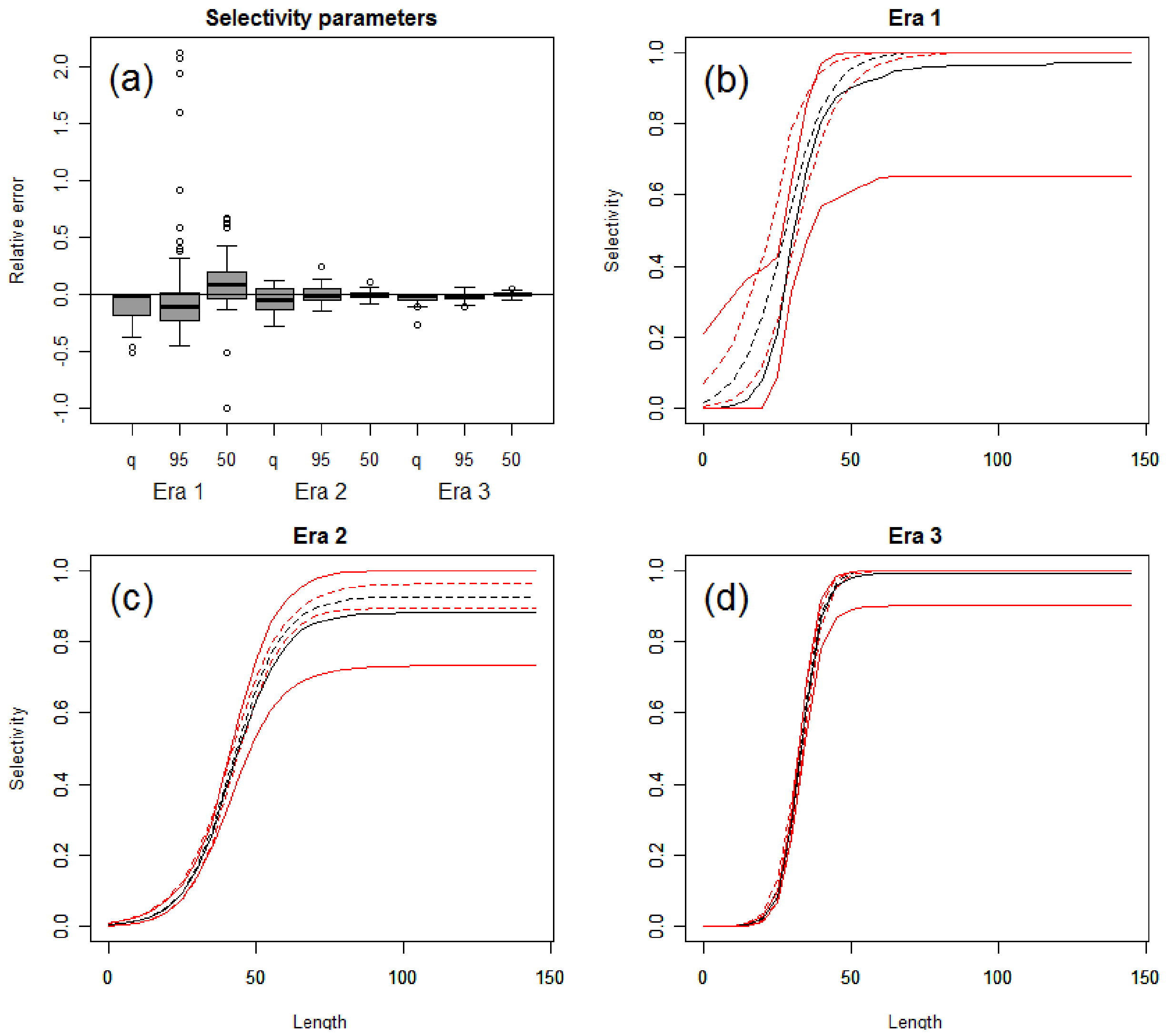

Figure 2.3. Relative error distributions for the parameters determining survey selectivity (a). True and estimated survey selectivity curves for eras 1, 2 and 3 (b,c,d). Dashed lines are the true curves and solid lines are the estimated curves, with the outer two dashed and solid lines respectively representing the 90% inter-simulation intervals for the true and estimated selectivity curves. The results in this figure are for the reference scenario.

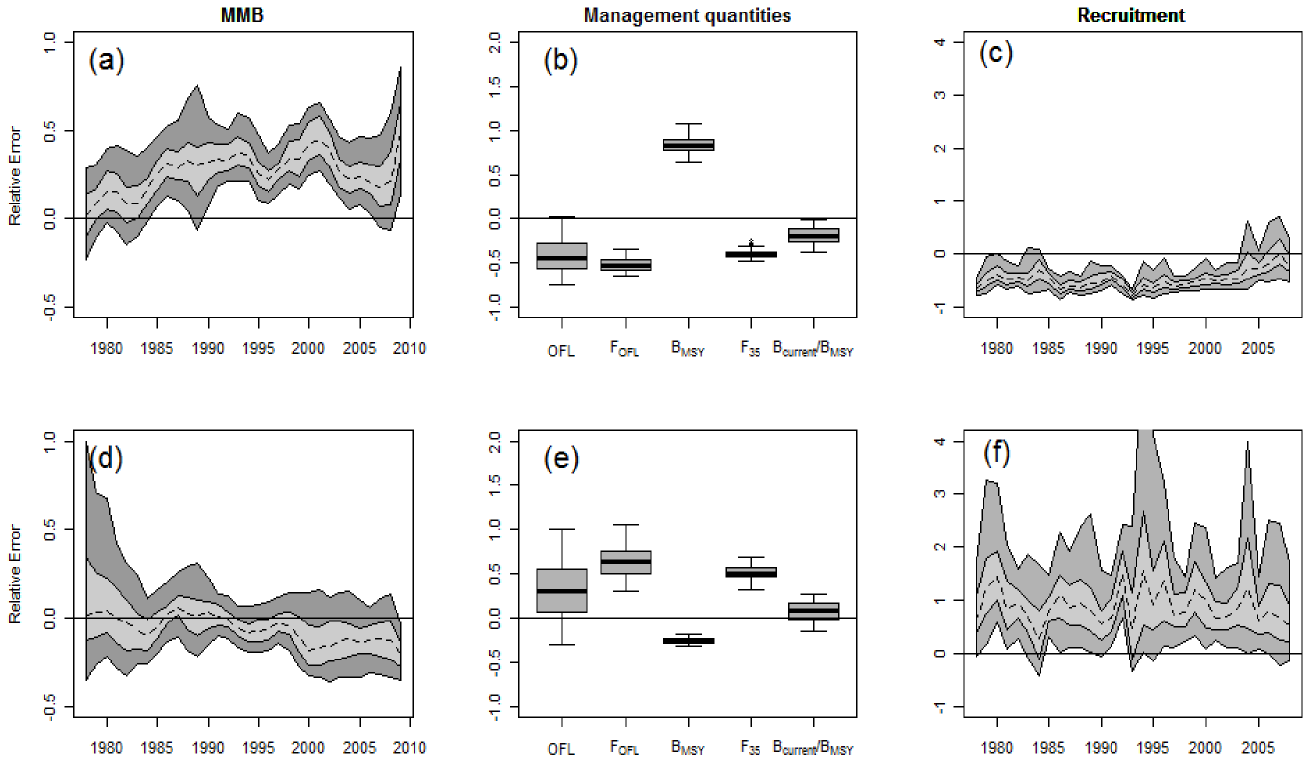

Figure 2.4. Results from natural mortality scenarios. Relative error in MMB (a and d) and recruitment (c and f) with 90% inter-simulation intervals in dark grey and 50% inter-simulation intervals in light grey, and management quantities (b and e). Panels a-c correspond to a scenario in which natural mortality is assumed to be 50% less than the truth in the estimation model and panels d-f to a scenarios in which natural mortality is assumed to be 50% higher.

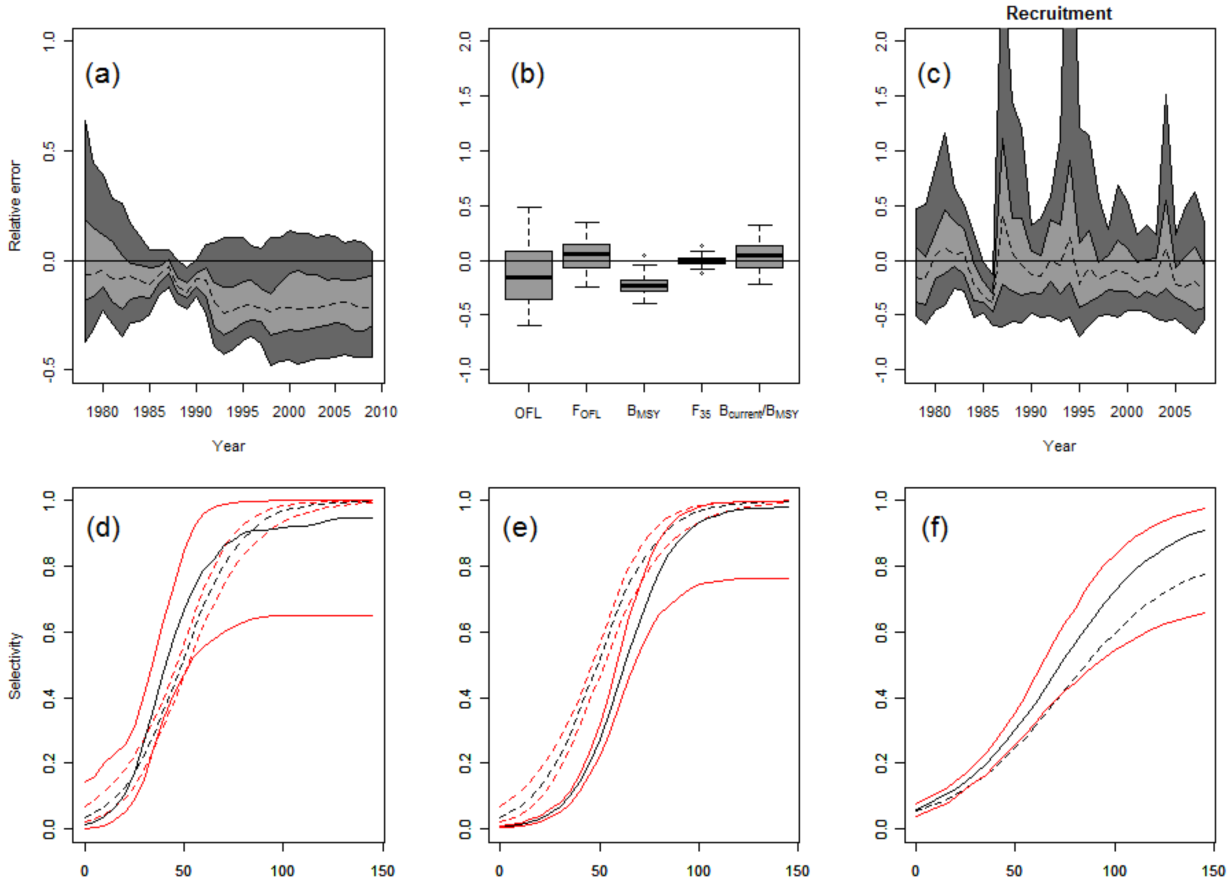

Figure 2.5. Results from selectivity scenarios. Relative error in MMB and recruitment with 90% inter-simulation intervals in dark grey and 50% inter-simulation intervals in light grey (a and c) and management quantities (b). Panels d-f are selectivity curves in era 1, 2, and 3, respectively. Dashed lines are true survey selectivity curves from the operating model and solid lines are estimated selectivity curves. Outer lines for each type (dashed/solid) are 90% inter-simulation intervals. The results in this figure are for the scenario in which survey selectivity is fixed at Somerton's curve for era 3 (1989-present).

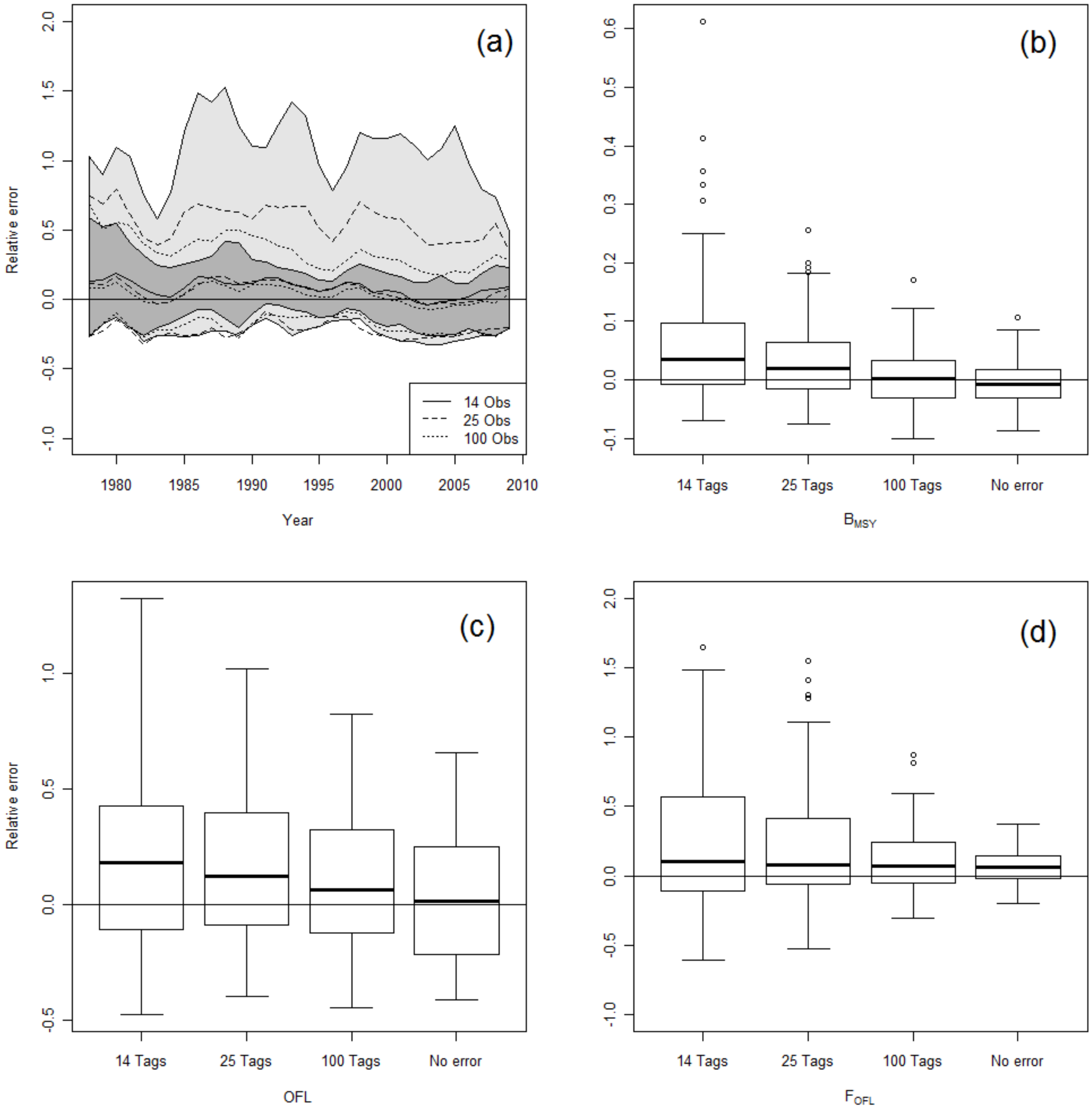

Figure 2.6. Results from growth scenarios. Median relative error in MMB for scenarios with different amounts of data used to estimate the parameters of the growth equation (a). The dark grey area is the 5th/95th quantiles for the reference scenario. The light grey area is the 5th/95th quantiles for the 14 observation scenario. Outermost dashed upper and lower lines are the 95th/5th quantiles for the 25 and 100 tag scenarios. The solid, dashed and dotted lines in the center of the dark grey area are the medians for the 14, 25 and 100 observation scenarios. Relative error distributions for the management quantities (B_{MSY} (b), OFL (c), F_{OFL} (d)) for increasing amounts of growth information.

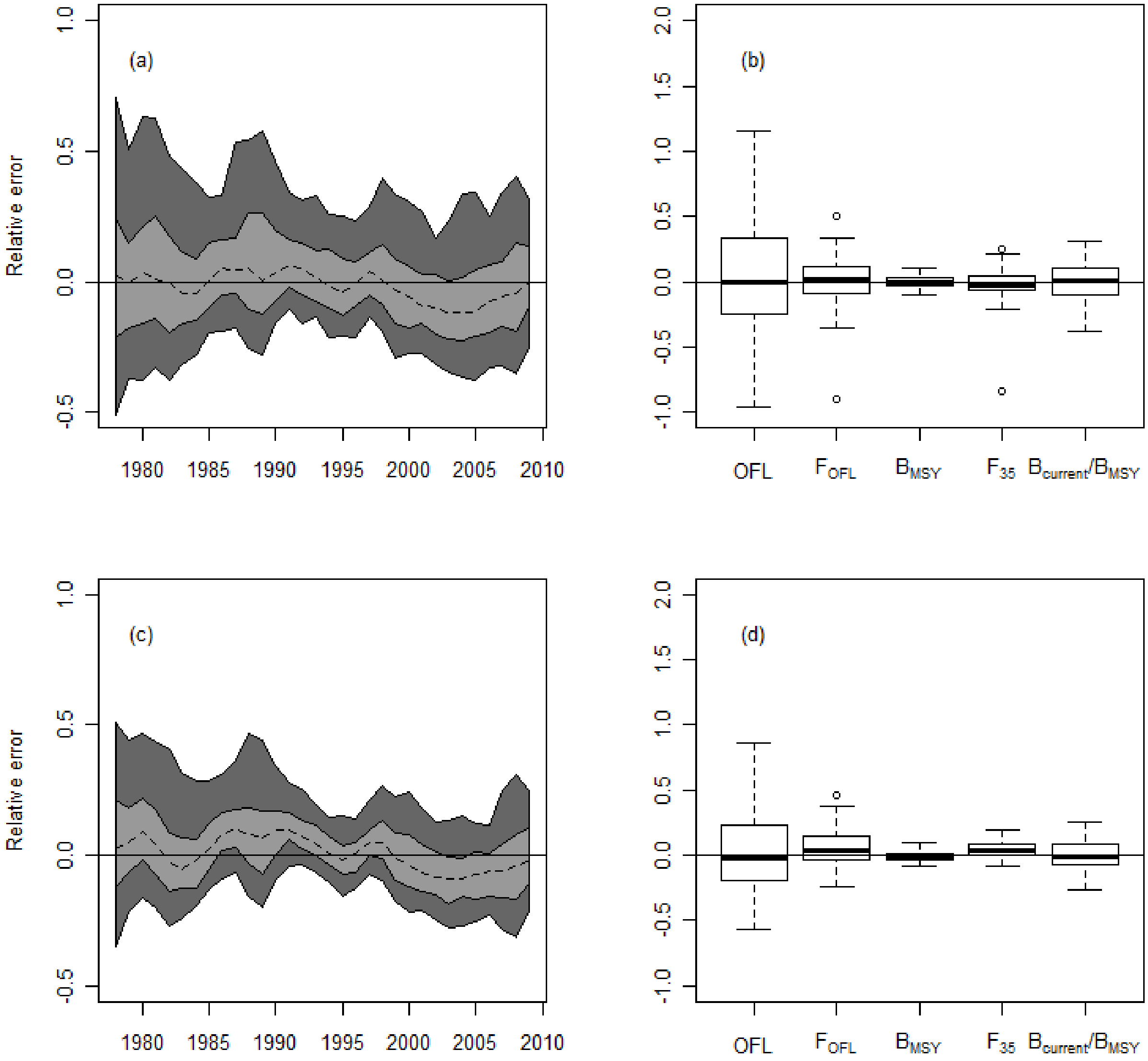

Figure 2.7. Results for survey data quality scenarios. Relative error distributions for mature male biomass (a and c) and management quantities (b and d) for scenarios in which survey coefficients of variation were increased and effective sample sizes (EffN) were decreased (top row; “CV EffN-“) and survey CVs were decreased by 0.08 and EffN was doubled (bottom row; “CV EffN+”) when generating data for assessment from the operating model .

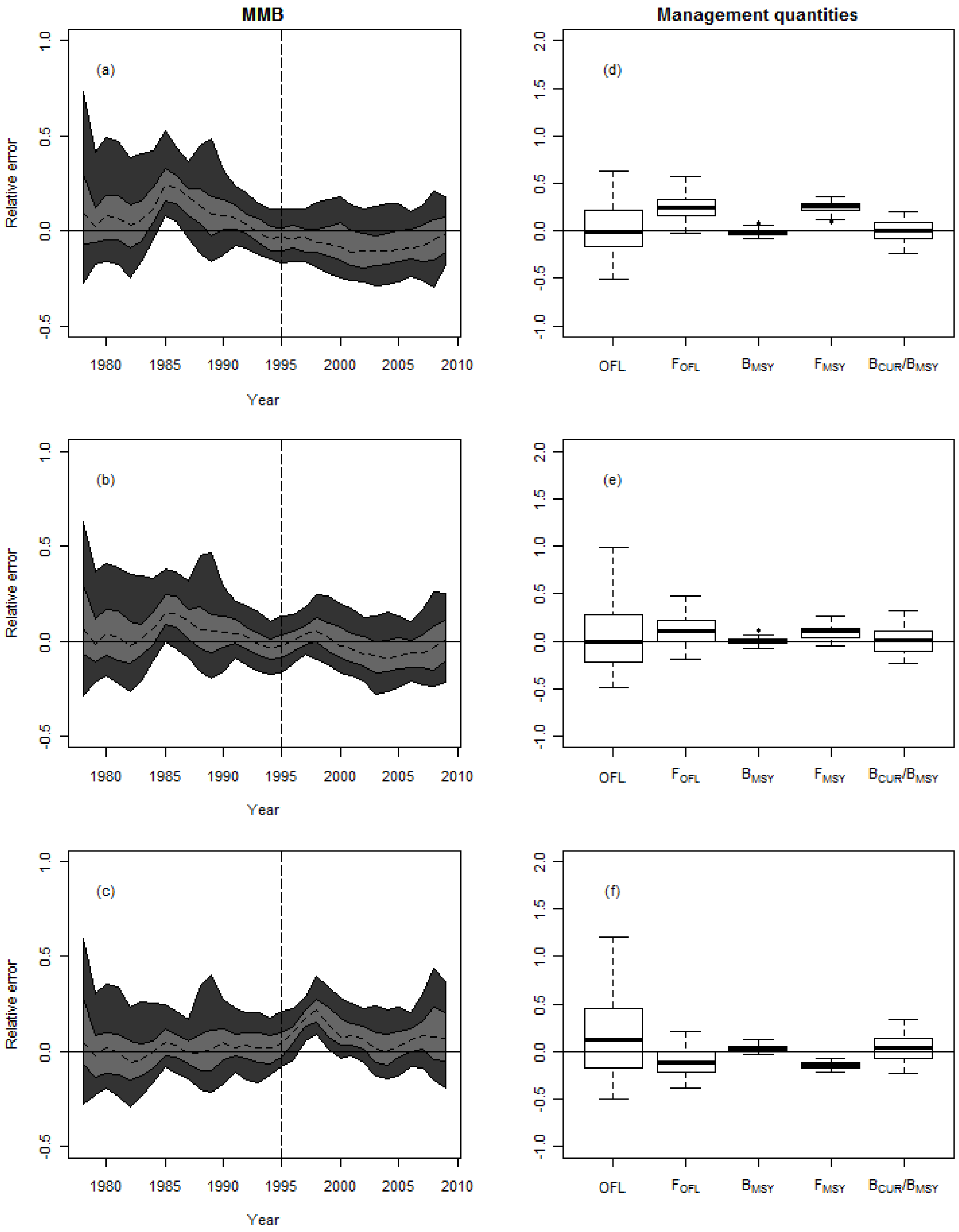

Figure 2.8. Results for fishing mortality scenarios. Relative error distributions for mature male biomass (MMB) and management quantities under scenarios in which average F_{mort} was multiplied by 0.5 (a and d), 1 (b and e), and 2 (c and f) starting in 1995. Quantities are estimated under the original fishing mortality deviation penalties.

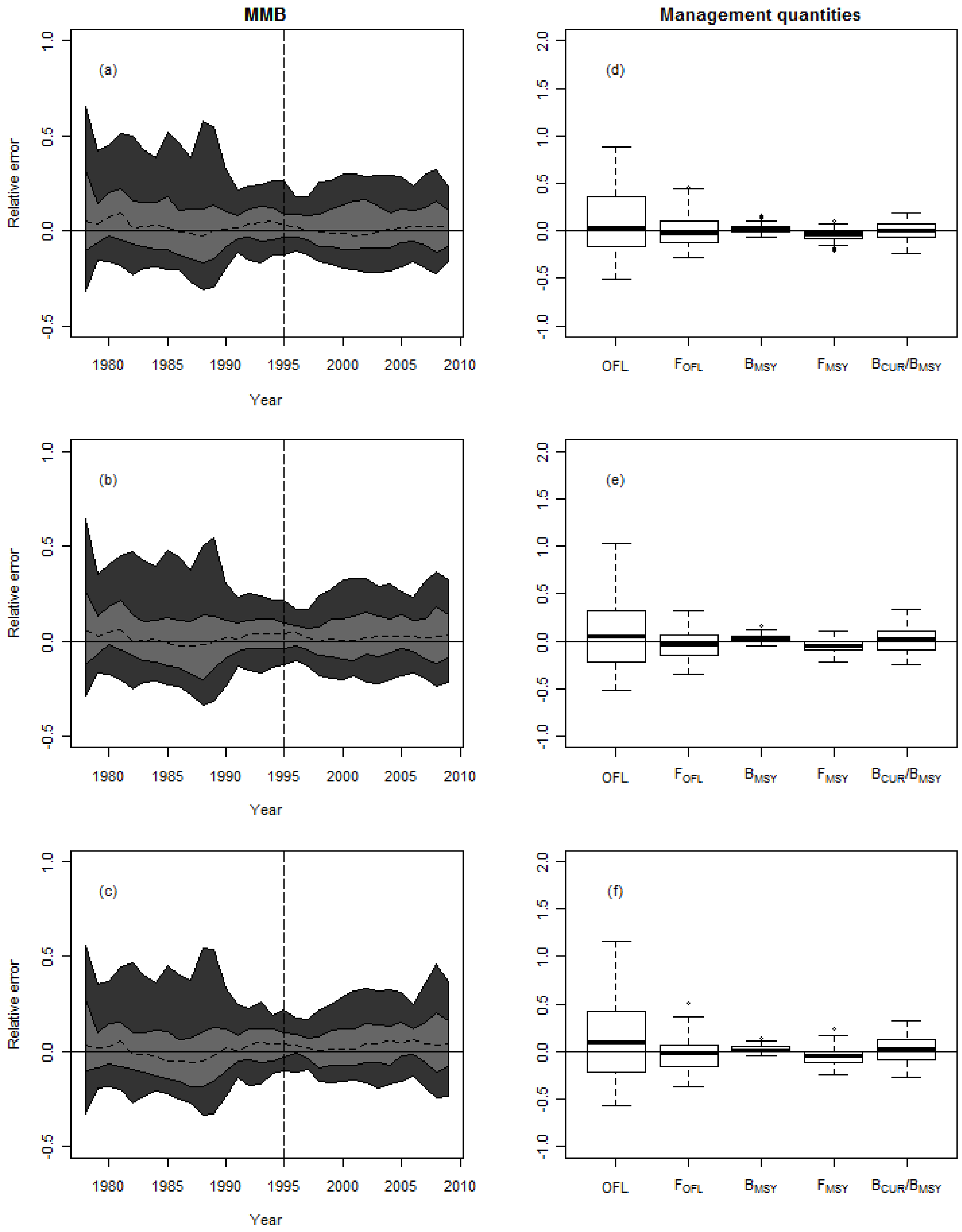

Figure 2.9. Results for fishing mortality scenarios. Relative error distributions for mature male biomass (MMB) and management quantities under scenarios in which average F_{mort} in the operating model was multiplied by 0.5 (a and d), 1 (b and e), and 2 (c and f) starting in 1995. Quantities are estimated with no penalties on fishing mortality deviations.

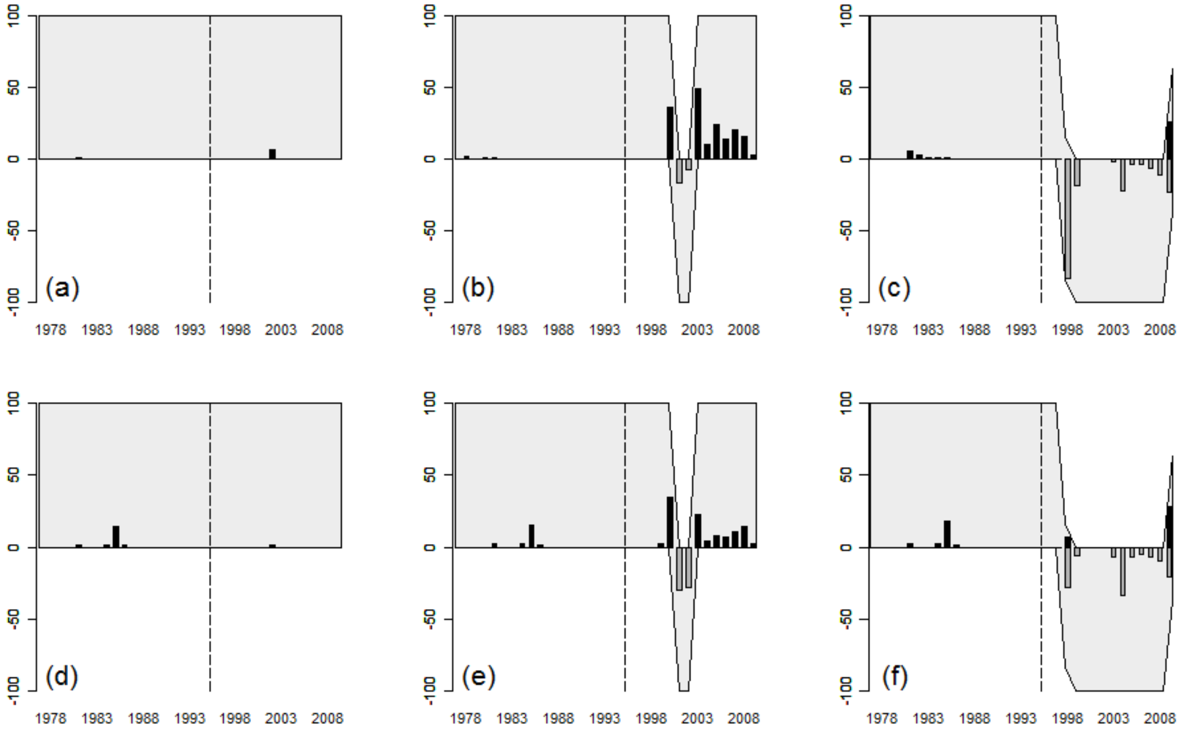

Figure 2.10. Evaluation of the ability to correctly identify overfished status under scenarios in which average F_{mort} in the operating model is multiplied by 0.5 (a and d), 1 (b and e) and 2 (c and f) from 1995 forward. Grey background shading indicates the number of simulations in which the true population was overfished (those below the "0" line) or not overfished (those above the "0" line). Black bars above the "0" line indicate the number of times a 'healthy' stock was identified as being overfished and grey bars beneath the "0" line indicate the number of times an overfished stock was identified as healthy. Panels a-c show results for scenarios in which the original penalties are applied and panels d-f results for scenarios in which penalties are eliminated.

Appendix 2.A: The Size-structured Population Dynamics Model

2.A.1 Basic dynamics

The model considers the dynamics of the population, grouping animals by sex *s*, shell condition (new/old) *v*, maturity state *m*, and length-class *l*, $N_{s,v,m,y,l}$:

$$N_{s,v,m,y+1,l}=\begin{cases} \emptyset_{s,l}\sum_{l'}\kappa_{s,imat,l'}Q_{s,imat,y,l'}X_{s,l',l} & if\ v=new; m=mat \\ (1-\emptyset_{s,l})\sum_{l'}\kappa_{s,imat,l'}Q_{s,imat,y,l'}X_{s,l',l}+\bar{R}e^{\varepsilon_y}Pr_l & if\ v=new;\ m=imat \\ Q_{s,imat,y,l'} & if\ v=old;\ m=mat \\ (1-\kappa_{s,imat,l'})Q_{s,imat,y,l'} & if\ v=old;\ m=imat \end{cases}$$

(2.A.1)

where $\kappa_{s,m,l}$ is the annual probability of an animal of sex *s* in maturity state *m* and length-class *l* moulting (the model includes a terminal moult to maturity so $\kappa_{s,\mathrm{mat},l}=1$), $Q_{s,m,y,l}$ is the number of animals of sex *s* in maturity state *m* and length-class *l* that survive fishing and natural mortality during year *y*:

$$Q_{s,m,y,l}=\sum_v N_{s,v,m,y,l}e^{-Z_{s,v,m,y,l}}$$

$\phi_{s,l}$ is the proportion of animals of sex *s* in length-class *l* that are mature, $X_{s,l',l}$ is the proportion of animals of sex *s* in length-class *l’* that moult into length-class *l* given that they moult (i.e. the size-transition matrix), $Z_{s,v,m,y,l}$ is the rate of total mortality on animals of sex *s*, shell condition *v*, maturity state *m*, and length-class *l* during year *y*:

$$Z_{s,v,m,y,l}=M_{s,m}+\sum_f F_{s,f,y,l}$$

$M_{s,m}$ is the instantaneous rate of natural mortality for animals of sex *s* and maturity state *m*, $F_{s,f,y,l}$ is the instantaneous rate of fishing mortality by “fleet” *f* (directed pot fishery and the groundfish trawl fishery) on animals of sex *s* in length-class *l* during year *y*, $\bar{R}$ is the median recruitment, ε_y is the recruitment deviation for year *y*, and Pr_l is the proportion of recruitment that occurs to length-class *l*.

2.A.2 Fishing mortality and selectivity

Fishing mortality is caused by directed fishing (for males), discard in the directed fishery (males and females) and bycatch in the groundfish trawl fisheries (the impact of the groundfish trawl fishery by length is assumed to be independent of sex). Fishing mortality is assumed to be independent of shell type and whether an animal is mature or not (this may not be the case in reality, but it is an assumption of the assessment model). Generically, fishing mortality due to each of these sources is given by:

$$F_{s,f,y,l} = S_{s,f,l}\,\tilde{F}_{s,f,y} = S_{s,f,l}\,\bar{F}_{s,f}e^{\eta_{s,f,y}} \qquad (2.A.2)$$

where $S_{s,f,l}$ is the selectivity of fleet *f* on animals of sex *s* in length-class *l*, $\bar{F}_{s,f}$ is the reference level of fully-selected fishing mortality for fleet *f* and sex *s*, and $\eta_{s,f,y}$ is the deviation during year *y* from the reference level of fully-selected fishing mortality for fleet *f* and sex *s*.

Fishery selectivity is assumed to be a logistic function of size and to be time-invariant. The selectivity patterns for the catch of males in directed fishery, the catch of females in the directed fishery, and the catch of males and females in the trawl fishery are given by:

$$\begin{aligned} S_{\mathrm{mal,dir},l} &= (1+\exp[-S_{\mathrm{slop,mal,dir}}(\bar{L}_l - S_{\mathrm{50,mal,dir}})])^{-1} \\ S_{\mathrm{fem,dir},l} &= (1+\exp[-S_{\mathrm{slop,fem,dir}}(\bar{L}_l - S_{\mathrm{50,fem,dir}})])^{-1} \\ S_{\mathrm{trawl},l} &= (1+\exp[-S_{\mathrm{slop,trawl}}(\bar{L}_l - S_{\mathrm{50,trawl}})])^{-1} \end{aligned} \qquad (2.A.3)$$

where $\bar{L}_l$ is the midpoint of length-class *l*, $S_{slop,\,,}$ is the slope of the selectivity curve and S_{50} is the length at which 50% of the individuals encountered are selected.

The probability of a male in length-class *l* being retained given that it was caught in the directed fishery, $\tilde{R}_l$, is given by:

$$\tilde{R}_l = (1+\exp[-\tilde{S}_{\mathrm{slop}}(\bar{L}_l - \tilde{S}_{50})])^{-1} \qquad (2.A.4)$$

2.A.3 Catches

The model predictions of the catch by fleet are given by:

$$\begin{aligned}
\hat{C}_{\mathrm{mal,dir},y} &= \sum_l \sum_v \sum_m w_{\mathrm{mal},l} \frac{\tilde{R}_l F_{\mathrm{mal,dir},y,l}}{F_{\mathrm{mal,dir},y,l} + F_{\mathrm{trawl},y,l}} N_{\mathrm{mal},v,m,y,l} e^{-\delta_y M_{s,m}} (1-e^{-(F_{\mathrm{mal,dir},y,l}+F_{\mathrm{trawl},y,l})}) \\
\hat{C}_{\mathrm{mal,tot},y} &= \sum_l \sum_v \sum_m w_{\mathrm{mal},l} \frac{F_{\mathrm{mal,dir},y,l}}{F_{\mathrm{mal,dir},y,l} + F_{\mathrm{trawl},y,l}} N_{\mathrm{mal},v,m,y,l} e^{-\delta_y M_{s,m}} (1-e^{-(F_{\mathrm{mal,dir},y,l}+F_{\mathrm{trawl},y,l})}) \\
\hat{C}_{\mathrm{fem,dir},y} &= \sum_l \sum_v \sum_m w_{\mathrm{feml},l} \frac{F_{\mathrm{fem,dir},y,l}}{F_{\mathrm{fem,dir},y,l} + F_{\mathrm{trawl},y,l}} N_{\mathrm{fem},v,m,y,l} e^{-\delta_y M_{s,m}} (1-e^{-(F_{\mathrm{fem,dir},y,l}+F_{\mathrm{trawl},y,l})}) \\
\hat{C}_{\mathrm{trawl},y} &= \sum_s \sum_l \sum_v \sum_m w_{s,l} N_{s,v,m,y,l} (1-e^{-F_{\mathrm{trawl},y,l}})
\end{aligned} \qquad (2.A.5)$$

where $\hat{C}_{\mathrm{mal,dir},y}$ is the model-estimate of the retained catch (in mass) during year *y* of males by the directed fishery, $\hat{C}_{\mathrm{mal,tot},y}$ is the model-estimate of the total (retained and discarded) catch (in mass) during year *y* of males by the directed fishery, $\hat{C}_{\mathrm{fem,dir},y}$ is the model-estimate of the catch (in mass) during year *y* of females by the directed fishery, $\hat{C}_{\mathrm{trawl},y}$ is the model-estimate of the catch (in mass) during year *y* of animals of both sexes by the trawl fishery, $w_{s,l}$ is the weight of an animal of sex *s* in length-class *l*, and δ_y is the mid-point of the fishery (in years).

2.A.4 Growth

The probability of moulting as a function of maturity state for immature animals is a declining logistic function of length (mature animals are assumed not to moult given the assumption of a terminal moult at maturity):

$$\kappa_{s,m,l} = 1-(1+\exp[-\theta^1_{s,m}(\bar{L}_l-\theta^2_{s,m})])^{-1} \tag{2.A.6}$$

The growth increment for animals that do moult is based on the gamma function, i.e.:

$$X_{s,i,j} = Y_{s,i,j} / \sum_k Y_{s,i,k} \tag{2.A.7a}$$

$$Y_{s,i,j} = (\Delta_{i,j})^{(\hat{L}_{s,i}-(\bar{L}_i-2.5))/\beta_s} e^{-\Delta_{i,j}/\beta_s} \tag{2.A.7b}$$

where $\hat{L}_{s,i}$ is the expected length for an animal of sex *s* in length-class *l* given that it moults:

$$\hat{L}_{s,i} = \gamma^1_s + \gamma^2_s \bar{L}_l \tag{2.A.8a}$$

γ^1_s, γ^2_s are the parameters of the relationship between length and growth increment, $\Delta_{i,j}$ is the difference in length between midpoints of length-classes *i* and *j*:

$$\Delta_{i,j} = \bar{L}_j + 2.5 - \bar{L}_i$$

β_s is the parameter which defines the variability in growth increment.

3.A.5 Recruitment

The fraction of the annual recruitment which recruits to length-class *l* is based on a gamma distribution, i.e.:

$$\Pr_l = (\Delta_{1,l})^{\upsilon^1/\upsilon^2} e^{-\Delta_{1,l}/\upsilon^2} / \sum_j (\Delta_{1,j})^{\upsilon^1/\upsilon^2} e^{-\Delta_{1,j}/\upsilon^2} \tag{2.A.9}$$

where υ^1, υ^2 are the parameters that define the recruitment fractions.

The annual recruitments are treated as estimable parameters.

3.A.6 Initial conditions

The numbers by length-class at the start of the first year considered in the model are treated as estimable parameters, i.e.:

$$N_{s,v,m,l,y_1} = \begin{cases} \phi_{s,l}\,\lambda_{s,v,l} & \text{if } v = \text{new}; m = \text{mat} \\ (1-\phi_{s,l})\,\lambda_{s,v,l} & \text{if } v = \text{new}; m = \text{imat} \\ \lambda_{s,v,l} & \text{if } v = \text{old}; m = \text{mat} \\ 0 & \text{if } v = \text{old}; m = \text{imat} \end{cases} \qquad (2.A.10)$$

where $\lambda_{s,v,l}$ is the proportion of animals of sex s which are of shell condition v in length-class l.

Appendix 2.B: The Objective Function for the Size-structured Model

2.B.1 Likelihood components

The model is fitted to the length-frequency of the retained catch of males, the length-frequency of the total catch of males, the length-frequency of females in the directed fishery, and the length-frequency of the catch by the trawl fishery:

$$L_{1,a} = \lambda_{1,a} \sum_y N_{1,y}^{\text{eff}} \sum_l p_{1,y,l}^{\text{obs}} \ell n(\hat{p}_{1,y,l} / p_{1,y,l}^{\text{obs}}) \qquad \text{(2.B.1a)}$$

$$L_{1,b} = \lambda_{1,b} \sum_y N_{2,y}^{\text{eff}} \sum_l p_{2,y,l}^{\text{obs}} \ell n(\hat{p}_{2,y,l} / p_{2,y,l}^{\text{obs}}) \qquad \text{(2.B.1b)}$$

$$L_{1,c} = \lambda_{1,c} \sum_y N_{3,y}^{\text{eff}} \sum_l p_{3,y,l}^{\text{obs}} \ell n(\hat{p}_{3,y,l} / p_{3,y,l}^{\text{obs}}) \qquad \text{(2.B.1c)}$$

$$L_{1,d} = \lambda_{1,d} \sum_s \sum_y N_{4,s,y}^{\text{eff}} \sum_l p_{4,y,s,l}^{\text{obs}} \ell n(\hat{p}_{4,y,s,l} / p_{4,y,s,l}^{\text{obs}}) \qquad \text{(2.B.1d)}$$

where $\lambda_{1,\text{a-d}}$ are weighting factors (see Table App 3.1), $N_{i,y}^{\text{eff}}$ is the effective sample size for year *y* and data-type *i* ($i=1$: male retained catch; $i=2$: male total catch; $i=3$: female discards in the directed fishery; $i=4$: catches in the trawl fishery), $p_{i,y,l}^{\text{obs}}$ is the observed proportion of the catch during year *y* of data-type *i* that is in size-class *l*, $\hat{p}_{1,y,l}$ is the model-estimate of the proportion of the retained catch of males in the directed fishery during year *y* that in length-class *l*:

$$\hat{p}_{1,y,l} = \tilde{R}_l \, S_{\text{mal,dir},l} \sum_v \sum_m N_{\text{mal},v,m,y,l} e^{-\delta_y M_{\text{mal},m}} \Big/ \sum_{v'} \sum_{m'} \sum_{l'} \tilde{R}_{l'} \, S_{\text{mal,dir},l'} N_{\text{mal},v',m',y,l'} e^{-\delta_y M_{\text{mal},m'}} \qquad \text{(2.B.2a)}$$

$\hat{p}_{2,y,l}$ is the model-estimate of the proportion of the total catch of males in the directed fishery during year *y* that is in length-class *l*:

$$\hat{p}_{2,y,l} = S_{\text{mal,dir},l} \sum_v \sum_m N_{\text{mal},v,m,y,l} e^{-\delta_y M_{\text{mal},m}} \Big/ \sum_{v'} \sum_{m'} \sum_{l'} S_{\text{mal,dir},l'} N_{\text{mal},v',m',y,l'} e^{-\delta_y M_{\text{mal},m'}} \qquad \text{(2.B.2b)}$$

$\hat{p}_{3,y,l}$ is the model-estimate of the proportion of the catch of females in the directed fishery during year *y* that is in length-class *l*:

$$\hat{p}_{3,y,l} = S_{\text{fem,dir},l} \sum_{v} \sum_{m} N_{\text{fem},v,m,y,l} e^{-\delta_y M_{\text{fem},m}} / \sum_{v'} \sum_{m'} \sum_{l'} S_{\text{fem,dir},l'} N_{\text{fem},v',m',y,l'} e^{-\delta_y M_{\text{fem},m'}} \tag{2.B.2c}$$

$\hat{p}_{4,y,s,l}$ is the model-estimate of the proportion of the trawl catch during year *y* that is sex *s* and is in length-class *l*:

$$\hat{p}_{4,y,s,l} = S_{\text{trawl},l} \sum_{v} \sum_{m} N_{s,v,m,y,l} e^{-\delta_y M_{s,m}} / \sum_{s'} \sum_{v'} \sum_{m'} \sum_{l'} S_{\text{trawl},l'} N_{s',v',m',y,l'} e^{-\delta_y M_{s',m'}} \tag{2.B.2d}$$

The model is fit to the survey length-frequency data by sex, shell-condition, and maturity state, i.e.:

$$L_2 = \lambda_2 \sum_{s} \sum_{v} \sum_{m} \sum_{y} N^{\text{eff}}_{5,s,v,m,y} \sum_{l} p^{\text{obs}}_{5,s,v,m,y,l} \ell n(\hat{p}_{5,s,v,m,y,l} / p^{\text{obs}}_{5,s,v,m,y,l}) \tag{2.B.3}$$

where $N^{\text{eff}}_{s,v,m,y}$ is the effective sample size for the survey length-frequency data for animals of sex *s* in shell condition *v*, and maturity state *m* during year *y*; $p^{\text{obs}}_{5,s,v,m,y,l}$ is the observed proportion of the survey catch for animals of sex *s* in shell condition *v* and maturity state *m* that is in length-class *l*; and $\hat{p}_{5,s,v,m,y,l}$ is the model-estimate corresponding to $p^{\text{obs}}_{5,s,v,m,y,l}$:

$$\hat{p}_{5,s,v,m,y,l} = S^I_{s,y*} N_{s,v,m,y,l} / \sum_{s'} \sum_{v'} \sum_{m'} \sum_{l'} S^I_{s',y*} N_{s',v',m',y,l'} \tag{2.B.4}$$

where $S^I_{l,y*}$ is the selectivity of the survey gear for animals in length-class *l* during year *y** (years are grouped into three epochs: before 1982, 1982-88, and 1989+):

$$S^I_{y*,l} = q^I_{y*}(1+\exp[-S^I_{\text{slop},y*}(\bar{L}_l - S^I_{50,y*})])^{-1} \tag{2.B.5}$$

$q^I_{y_*}$ is the survey catchability coefficient for the years represented in the set *y**, $S^I_{50,y*}$ is the length-at-50%-survey selectivity for the years represented in the set *y**, and $S^I_{\text{slop},y*}$ is the slope of the survey selectivity ogive for the years represented in the set *y**.

The model is fit to the survey indices separately by sex, i.e.

$$L_3 = 0.5\lambda_3 \sum_s \sum_y \left(\ell\mathrm{n} I_{s,y} - \ell\mathrm{n} \hat{I}_{s,y} \right)^2 / \sigma_{s,y}^2 \qquad (2.B.6)$$

where $I_{s,y}$ is the survey index of abundance for sex *s* and year *y*, $\hat{I}_{s,y}$ is the model-estimate corresponding to $I_{s,y}$:

$$\hat{I}_{s,y} = \sum_l \sum_v S_{y*,l}^I \, w_{s,l} \, N_{s,v,\mathrm{mat},y,l}$$

$\sigma_{s,y}$ is the standard error of $I_{s,y}$.

The contribution of the “large” males (102mm+) to the likelihood function is given by:

$$L_4 = \lambda_4 \sum_y \left(\ell\mathrm{n} J_y - \ell\mathrm{n} \hat{J}_y \right)^2 \qquad (2.B.7)$$

where J_y is the number of “large” males in the survey during year *y*, and $\hat{J}_y$ is the model-estimate corresponding to J_y:

$$\hat{J}_y = \sum_v \sum_m \left(0.5 S_{\mathrm{mal},16}^I N_{s,v,m,y,16} + \sum_{l>16} S_{\mathrm{mal},l}^I N_{s,v,m,y,l} \right) \qquad (2.B.8)$$

The contribution of the catch data to the likelihood function is given by:

$$L_{5,a} = \lambda_{5,a} \sum_y \left(\ell\mathrm{n} C_{\mathrm{mal,dir},y}^{obs} - \ell\mathrm{n} \hat{C}_{\mathrm{mal,dir},y} \right)^2 \qquad (2.B.9a)$$

$$L_{5,b} = \lambda_{5,b} \sum_y \left(\ell\mathrm{n} C_{\mathrm{mal,tot},y}^{obs} - \ell\mathrm{n} \hat{C}_{\mathrm{mal,tot},y} \right)^2 \qquad (2.B.9b)$$

$$L_{5,c} = \lambda_{5,c} \sum_y \left(\ell\mathrm{n} C_{\mathrm{fem,dir},y}^{obs} - \ell\mathrm{n} \hat{C}_{\mathrm{fem,dir},y} \right)^2 \qquad (2.B.9c)$$

$$L_{5,d} = \lambda_{5,d} \sum_y \left(\ell\mathrm{n} C_{\mathrm{trawl},y}^{obs} - \ell\mathrm{n} \hat{C}_{\mathrm{trawl},y} \right)^2 \qquad (2.B.9d)$$

where $C_{f,y}^{\mathrm{obs}}$ is the observed catch (in mass) by fleet *f* during year *y*.

3.B.2 Penalty components

There are several penalty components in the objective function.

A penalty is placed on the deviations in recruitment from average recruitment:

$$P_1 = \lambda_6 \sum_{y=1979}^{2008} \varepsilon_y^2 \qquad (2.B.10)$$

A penalty is placed on the between-length-class variation in the initial size-structure:

$$P_2 = \lambda_7 \sum_s \sum_v \sum_{l=1}^{n_{L,s}-1} (\ell n\lambda_{s,v,l} - \ell n\lambda_{s,v,l+1})^2 \qquad (2.B.11)$$

where $n_{L,s}$ is the number of size-classes in the first year for sex s (12 for females and 22 for males).

Penalties are placed on the extent of inter-annual variation in deviations in fishing mortality and change depending on the phase of estimation:

$$P_{3,a} = \lambda_{8,a} \sum_y \eta_{\mathrm{mal,dir},y}^2 \qquad (2.B.12)$$

Table App.3.1. The weighting factors applied to the penalties and likelihood components.

Description	Value
Likelihood components	
Catch proportions-at-length	
Direct male catch, $\lambda_{1,a}$	1
Total male catch, $\lambda_{1,b}$	1
Direct female catch, $\lambda_{1,c}$	0.2
Trawl catch, $\lambda_{1,d}$	0.25
Survey proportions-at-length, λ_2	1
Survey biomass data, λ_3	0.25
Survey proportions-at-length (large males), λ_4	0.001
Catches	
Direct male catch, $\lambda_{5,a}$	1000
Total male catch, $\lambda_{5,b}$	1000
Direct female catch, $\lambda_{5,c}$	10
Trawl catch, $\lambda_{5,d}$	100
Penalty components	
Recruitment deviations, λ_6	2
Initial size-structure, λ_7	1
Deviations in fishing mortality	
Direct male fishery (phase 1), $\lambda_{8,a}$	10
Direct male fishery (phase 2), $\lambda_{8,b}$	50
Direct male fishery (final), $\lambda_{8,c}$	0.1

Chapter 3: Regime shifts and oscillating control of snow crab, *Chionoecetes opilio,* recruitment in the eastern Bering Sea

3.1. Introduction

The snow crab (*Chionocetes opilio*) fishery in the eastern Bering Sea (EBS) is a historically lucrative fishery on male crabs that became strictly domestic in 1980 (Turnock and Rugolo, 2011). It is characterized by large fluctuations in both catch and biomass (Fig. 3.1a), which are hypothesized to be the result of highly variable recruitment (Kruse et al., 2007). The highest historical catch of 149,000 tons was harvested in 1991, which coincided with an estimated male mature biomass of 426,000 tons (Turnock and Rugolo, 2011). The stock was declared overfished in 1999 when survey estimates of total mature biomass were below the agreed Minimum Stock Size Threshold (208,710 t), and a rebuilding plan was enacted (NPFMC, 2000). The goal of this rebuilding plan was to return mature male biomass (MMB) to above the MMB at which Maximum Sustainable Yield is achieved, B_{MSY} (133,246 t) for two consecutive years, but the plan was deemed a failure in 2009 when this goal was not met. Estimated recruitments during the rebuilding period of 1999-2009 were substantially lower than the average recruitment over the history of the fishery. Average recruitment is used when calculating quantities used in the harvest control rules on which management advice is based. Low recruitment was therefore likely a contributing factor to the failure of the rebuilding plan. A better understanding of recruitment processes would have been useful in formulating the rebuilding plan (and designation of stock status) and may have impacted its success. This paper develops models for snow crab recruitment that incorporate both environmental effects and the influence of spawning biomass. These models will be used in conjunction with IPCC projections for management strategy evaluations to assess the performance of different harvest strategies.

Large female spawning biomasses during the 1990s failed to produce strong year classes and this is reflected in the absence of an obvious spawner-recruit relationship for EBS snow crab when the entire time series of spawner and recruitment data is considered (Fig. 3.1b). There are many hypotheses concerning the mechanisms governing year-class strength for EBS snow crab. An exhaustive treatment, broken down by influence on life history stage, is presented in Kruse et al. (2007) and key points are summarized in Table 3.1. However, efforts to correlate recruitment to large scale environmental indices and other factors have been unsuccessful. For example, Zheng and Kruse (2000) attempted to find relationships between snow crab recruitment and barometric pressure, upwelling, coastal sea level, sea surface temperature (SST), subsurface temperature and groundfish biomass, but no significant relationships were found. Zheng and Kruse (2006), using a longer time series of data, compared the trends in the Aleutian Low Pressure Index (ALPI) and the Pacific Decadal Oscillation (PDO) and snow crab recruitment, but again found no significant correlation.

Snow crab life history strategies are thought to be relatively conservative. For example, snow crab follow a consistent ontogenetic migration (Ernst et al., 2005) and have relatively narrow preferences for environmental conditions during early life stages (Dionne et al., 2003). This leads to the possibility that environmental factors could be involved in determining year-class strength. Temperature, in particular, has the potential to affect many life history stages. For example, size-at-maturity appears to be influenced by temperature (Orensanz et al., 2007). Females also may retain their eggs for one or two years after mating, depending on water temperature at the time of mating (Moriyasu and Lanteigne, 1998). Specifically, temperatures below 1.5° C appear to trigger biennial spawning in female snow crab in the Bering Sea (Rugolo et al., 2005). Furthermore, early benthic stages have shown a preference in the laboratory for cold (0 to 1.5° C) water (Dionne et al., 2003). Survival of the pelagic larval stages may depend on the timing, location and intensity of primary production, which are related to sea surface temperature, ice-edge retreat and wind mixing (Jin et al., 2007). It has also been hypothesized that low temperature may be necessary for successful mating; if water temperatures are too warm, mating may be disrupted (Kruse et al., 2007).

Recently, Hinckley et al. (2010) used individual-based models linked to a Regional Ocean Modeling System to predict the tracks of individual larvae in relation to modeled plankton blooms (given a release point estimated from survey estimates of female spawning biomass) to explore the impact of the overlap of trajectories of larvae and the bloom. Temperature at settlement, location of hatching and settlement, and cod predation were also included in the model. The results of this study were summarized in ‘The Gates Hypothesis”, which proposes snow crab larvae must go through a series of ‘gates’ before recruiting to the population. Important ‘gates’ were determined to be settlement on the middle shelf and encountering plankton blooms. The ‘r^{2}’ s of the models presented by Hinckley et al. (2010) range from 0.43 to 0.54.

Potential influences on recruitment due to regime shifts have not yet been included in any recruitment analyses for EBS snow crab. “Regime shift” has several definitions (King, 2005; Rudnick and Davis, 2003; Duffy-Anderson et al., 2005; see Overland et al., 2008 for a discussion). A common theme throughout these definitions is a persistent change in environmental conditions that are reflected in the biological components of the system, and it is in this context that ‘regime shift’ is used here. Overland et al. (2008) tested a suite of climate indices (e.g. PDO and Arctic Oscillation (AO)), physical indices (e.g. Bering Sea Pressure Index, surface temperature and ice cover index) and biological indices (e.g. indices related to salmon, walleye pollock and jellyfish) for shifts in average value. Most shifts occur in these indices around 1976-1978, 1988-1990 and 1998-1999. Temperature and differences in pressure systems are the largest distinguishing characteristics among regimes in the North Pacific, with winter conditions during 1977-1988 being exceptionally warm with a low Arctic Oscillation (AO) and Aleutian Low Pressure Index (ALPI), and 1989-2003 relatively cool with a high AO and ALPI. The 1998 shift in summer PDO was primarily in sea surface temperature (Bond et al., 2003) and ecosystem reorganization was not observed (Litzow, 2003). A potential shift in the winter PDO appears to have occurred in 2003 (Overland et al., 2008).

Shifts in climate have been shown to influence recruitment of several species in the North Pacific. For example, Pacific cod (*Gadus macrocephalus*), rock sole (*Lepidopsetta. bilineatea*), flathead sole (*Hippoglossoides elassodon*) and walleye pollock (*Theragra chalcogramma*) in the Bering Sea all experienced increases in recruitment after the 1977 regime shift (Conners et al., 2002). Recruitment of Bristol Bay sockeye salmon was also observed to increase after the 1977 shift (Adkison et al., 1996). Hare and Mantua (2000) noted shifts in recruitment of many species in 1977 and 1989. Control of walleye pollock recruitment in the Gulf of Alaska appeared to shift from environmental effects on larvae to top-down predation on juveniles after the 1977 regime shift (Bailey, 2000). Similarly, control of Pacific Cod abundance in the North Pacific changed from top-down effects to bottom-up effects after the regime shift in the late 1970s (Litzow and Cianelli, 2007). Estimated snow crab recruitment also sharply declined after the 1989 regime shift at which time the previously relatively close relationship between female spawning biomass and recruitment breaks down (Fig. 3.1c).

Shifts in climate regime could affect year-class strength in snow crab given the evidence for the influence of regime shifts on other species in the EBS, the coincidence of the decline of snow crab recruitment and the 1989 regime shift and the environmental preferences of snow crab. This paper explores models for snow crab recruitment that incorporate shifts in climate regime and suggests a hypothesis to explain the observed trends.

3.2. Methods

3.2.1 Data

Snow crab recruitment estimates (lagged to year of fertilization) are available from the stock assessment for the years 1974-2005 (Turnock and Rugulo, 2011). For this analysis, estimates from 1979-2003 were used because estimates of female spawning biomass are only available starting in 1978. The last two years of recruitment are not included in the analysis because of the uncertainty associated with the most recent estimates of recruitment from a stock assessment. This uncertainty is not apparent in the probability intervals for estimated recruitment (Fig. 3.1c). However, those intervals are overly narrow as reflected by the extent to which recent recruitment estimates change from one assessment to the next. Assessment estimates are the most appropriate quantities to be modeled (as opposed to estimates generated from the survey) because they are based on all available sources of data. Assessment estimates are often used in similar analyses (e.g., Arregui et al., 2006; Fiksen and Slotte, 2002).

Although the assessment method is designed to track mature male biomass, female spawning biomass, FSB, is used as a proxy for reproductive capacity for this analysis. There is a negative relationship between mature male biomass and recruitment (discussed briefly below) and total spawning biomass does not appear to be related to recruitment. Determining which measure of FSB from the assessment with which to predict recruitment is an important consideration. Estimated FSB at mating could be used, but it depends on assumptions about survey catchability that are not easily verified. Survey selectivity and catchability are estimated in three blocks (1978-1982, 1983-1988, 1989-present) in the stock assessment because of changes in survey gear. Estimated survey catchability is 1 during the period 1978-1982, but is estimated at 0.51 during 1983-1989 under the data weighting used to estimate the recruitment and female spawning biomass series used here. Given the manner of change in survey gear, this difference does not seem plausible and suggests that assessment-based estimates of FSB at mating are not appropriate for this analysis.

The only source of data available that informs estimated FSB are the survey data and those data could also be used as a covariate. However, each year's estimate has an associated and variable amount of uncertainty which the assessment method smoothes using auxiliary information, in particular the information on the survey length-composition. All considered, the most appropriate measure is the predicted FSB at the time of the survey estimated in the assessment. This quantity is similar to estimated FSB at mating, but it is not subject to assumptions about survey catchability. It is somewhat questionable to use dependent quantities from one model in another model, but using this measure of FSB as a covariate represents the best compromise between incorporating as much information into a predictor as possible and at the same time avoiding too many untestable assumptions.

Finally, eighteen regional and local scale environmental indices that reflect the environmental conditions of the EBS taken from www.beringclimate.noaa.gov (Table 3.2) were considered as potential covariates. The covariates selected have the potential to influence one or more of the processes hypothesized to influence recruitment dynamics outlined in Table 3.1. All variables were scaled and centered on zero.

3.2.2 Selection of covariates and cross-validation

Recruitment was broken into two periods (1979-1989 and 1990-2003) corresponding to the climate regimes in the EBS. Correlations between FSB, environmental indices and recruitment during were examined to identify candidate covariates in each regime. Francis (2006) suggests the process of selecting the covariates should be included in any cross validation of environment-recruitment models. Cross-validation ensures that single observations do not unduly influence the estimated relationship between the data and a covariate. A year of data was dropped and the significance of correlation between the indices and recruitment was recalculated. This process was repeated for each year to identify robust relationships between recruitment and covariate.

Recruitment/environment relationships often collapse when new data are added to an analysis (Myers, 1998), and regressions in which the entire data set is used to estimate model parameters are prone to overestimating the power of a relationship (Francis, 2006). In addition to using ‘leave-one-out’ cross-validation for identifying covariates, cross validation was also applied to the model(s) formulated from the selected covariates. A retrospective analysis was also performed in which the model was fit using data up to a certain year (e.g. 1996) then the remaining years (e.g. 1997-2003) were predicted using the model. The predictive ability of the model was assessed by comparing mean absolute relative errors (RE) over the period 1996-2003 for each step of the retrospective analysis. Absolute RE was calculated using:

$$RE_t^j = |\hat{Q}_t^j - Q_t|/Q_t \tag{3.1}$$

where RE_t^j is the absolute relative error for year *t* based on predictions of recruitment from model *j*, Q_t is the observed recruitment for year *t*, and $\hat{Q}_t^j$ is the estimate of recruitment for year *t* from model *j*. There are 8 retrospective models, based on truncating the data in each of the years 1996-2003.

3.2.3 Model structure

Piece-wise regression is used to fit the data in each regime separately, testing the covariates that exhibit robust relationships with recruitment within a regime. A model estimating regime-specific average recruitments with the breakpoint at the observed regime shift (1989) was fit to the data to provide a baseline for comparison. Additionally, a series of models were fit that tested each year of the time series as the breakpoint to verify the regime shift in 1989 can explain a large proportion of the variation in the data.

3.3 Results

3.3.1 Selected covariates

Recruitment is highly correlated ($r=0.81$; $p<0.005$) with only FSB in the first regime, suggesting it is an appropriate covariate for that regime. The correlation between FSB and recruitment breaks down coincident with a decline in recruitment after 1989. Correlations between recruitment and the winter Pacific Decadal Oscillation (wPDO; $p<0.04$), winter sea surface temperature near the Pribilof islands (PribSSTw; $p<0.005$), winter wind stress along the Aleutian peninsula (wWind; $p<0.06$), sea surface temperature in May in the Bering Sea ($p<0.10$), bottom temperature ($p<0.10$) and the ice cover index ($p<0.10$) were significant in the second regime and were considered as potential explanatory covariates.

Cross-validation indicated that only the correlations between recruitment and the wPDO, PribSSTw, and wWind were significant ($p<0.10$) during the second regime (i.e. above the cutoff) in a large percentage of the iterations (wPDO: 93%, PribSSTw: 100%, wWind: 78%). The models considered below only consider these covariates further.

3.3.2 Models tested

Each model uses FSB to predict recruitment in the 'warm' regime from 1979-1989 (Equation 3.2).

$$\log(Rec_t) = \alpha_1 + \beta_1 * FSB_t + \varepsilon \qquad (3.2)$$

The covariates used to predict recruitment in the 'cool' regime from 1990 to 2003 change among models (Equations 3-6). The indices PribSSTw and wPDO are not included in the same model because they are highly correlated. The following models were compared and model selection was performed by Akaike's Information Criterion adjusted for small samples, AIC_c.

Model 1: $\log(Rec_t) = \alpha_2 + \beta_2 * wPDO_{t+1} + \varepsilon$ (3.3)

Model 2: $\log(Rec_t) = \alpha_2 + \beta_2 * wPDO_{t+1} + \beta_3 * wWind_{t+1} + \varepsilon$ (3.4)

Model 3: $\log(Rec_t) = \alpha_2 + \beta_2 * PribSSTw_{t+1} + \varepsilon$ (3.5)

Model 4: $\log(Rec_t) = \alpha_2 + \beta_2 * PribSSTw_{t+1} + \beta_3 * wWind_{t+1} + \varepsilon$ (3.6)

where $\varepsilon = N(0,\sigma^2)$ (3.7)

In all equations, α_1, α_2, β_1, β_2, β_3 and σ are estimated parameters. Therefore, recruitment in year t depends on the environmental conditions in the year during which settlement occurs t+1 (usually one year after fertilization) [Equations 3.3-7] *or* female spawning biomass during the year of fertilization t [Equation 2].

3.3.3 Model fits

Fitting an average to the log-recruitment data and allowing for a change in average recruitment in only 1989 resulted in an AIC_c of 52.56. Shifts in average recruitment during 1990 and 1989 lead to the largest r^2s when allowing the change in average recruitment to occur in each year of the time series (results not shown), indicating that models with shifts at these points explain the most variation in recruitment.

Model 2 (FSB, wPDO and wWind) and model 4 (FSB, PribSSTw, and wWind) led to the lowest AICcs (37.93 and 35.50, respectively; Table 3.3) and ANOVA indicated all variables included are significant (at least $p < 0.02$). The difference in AICc between models 2 and 4 suggests that model 2 is a significant improvement over model 4 (fits are in Fig. 3.2 and parameter estimates and standard errors are listed in Table 3.4). In both models, log-recruitment and the covariates representing sea surface temperature (wPDO and PribSSTw) exhibit a strong negative relationship. Wind stress along the Aleutian peninsula (wWind) and log-recruitment are positively related.

Leave-one-out cross-validation resulted in only small changes in model fit (Fig 3.3). Prediction of the years 1996-2003 from retrospective models also changed very little with additional data. Model 2 changed most with the exclusion of data (Fig. 3.4), but even then the mean absolute REs over the period 1996-2003 were at most 5%. Relatively small changes in model fit under cross-validation and retrospective analysis are likely a result of selecting the covariates via cross-validation, using a small number of covariates in each regime as predictors, and assuming a linear relationship between log-recruitment and the variables used.

3.4. Discussion

Two models are presented in which drivers of recruitment dynamics change from female spawning biomass to environmental effects coincident with a shift in climate regime. Measures of sea surface temperature and winter winds were significant predictors of log-recruitment in both models. One model uses the wPDO, a regional scale index of SST that exhibits a 'regime shift'; the other uses local measure of SST (PribSSTw) that does not display a regime shift. The model with locally-derived covariates is selected by a narrow margin via AIC_c , but, for the purpose of projection, the model including the wPDO could be preferred because the break point that determines when control of recruitment dynamics switches from FSB to environmental effects is derived from a 'regime shift' reflected in the wPDO. Additionally, indices of large-scale climatic variation such as the wPDO often predict ecological relationships as well or better than local variables, possibly because they are often correlated to many local processes that would be very difficult to capture in a single model (Hallet et al., 2004).

Regardless of the choice of presented models, snow crab recruitment appears to be strongly related to both atmospheric and oceanic environmental conditions in the winter in the EBS. Measures of winter SST are likely good predictors of recruitment because they reflect the oceanic conditions closest to the point when larvae enter the pelagic phase. Winter winds at Unimak Pass may reflect to both current patterns and timing and location of the spring bloom (Jin et al. 2007). However, there seems to be a contradiction in the presented models--warm winter regimes appear to result in better overall recruitment, but, in cool regimes, colder years result in better recruitment. A (difficult to test) hypothesis to explain this phenomenon draws from and adds to ideas put forth in the Oscillating Control Hypothesis (Hunt et al., 2002) and the Gates Hypotheses.

In cold regimes, the spring algal bloom often occurs early (March-April) and in association with the ice-edge (Hunt et al., 2002). Larval crab are abundant in the water column during April-June (Incze et al., 1987), and larval survival could be negatively impacted by missing these early blooms. In warm regimes, the algal bloom often occurs later (May-June) and may overlap better with the pelagic stage of the larvae. Additionally, zooplankton biomass and chlorophyll concentration was higher through the 1980s than the 1990s (Sugimoto and Tadokoro, 1997). Cushing's match/mismatch theory (e.g. Cushing (1990)) suggests that abundant food availability could improve survival through the pelagic stage and year-classes in warm regimes might therefore be more directly related to spawning biomass. Durations of larval stages are also inversely related to temperature, so less time would be spent in the vulnerable pelagic stage during warmer years (Kon, 1970). The combination of higher food availability and a shorter time in the water column may explain why recruitment is well-related to FSB during warm regimes.

Why, then, would relatively warm years in cool regimes not also result in better recruitment? Temperature is not the only variable that changes between climate regimes; atmospheric conditions change as well. After 1989, the ALPI changed from intense Aleutian Lows (above average south-westerly and westerly circulation patterns and warming of coastal sea surface temperatures) to average Aleutian Lows (less frequent south-westerly and westerly circulation and slightly cooler coastal sea surface temperatures in winter) (McFarlane et al., 2000). The resulting changes in oceanic currents could influence the proportion of larvae advected from the outer domain farther onto the shelf to suitable nursery grounds. Juvenile crab prefer colder water (Dionne et al., 2003), so colder years may result in more 'optimal' nursery habitat available on the shelf. Low bottom temperatures limit the numbers and growth of groundfish, which may ease predation pressure on juvenile crab (Wyllie-Echeverria & Wooster, 1998). So, in cool regimes, even larval survival through the pelagic phase is relatively good due to a warm year, larvae may not be advected to suitable nursery grounds. Cold years may increase the area of suitable nursery grounds, increasing the probability of settling in favorable conditions in spite of suboptimal advection. Figure 3.5 presents a cartoon representation of this hypothesis. Ultimately, this hypothesis (and any other that attempts to explain recruitment through larval survival) is speculative and will remain so until detailed larval survival, timing, and density data are available.

Making predictions of recruitment useful for tactical management with either of these models may be difficult. The relationship between a covariate and recruitment is only known after a regime is underway. In general, however, warmer winter regimes appear to have a higher potential for recruitment. If the wPDO model is followed, what appears to be a recent shift back to a warmer regime (as seen through the winter PDO in 2003) suggests a short-term return to higher recruitment starting in 2003 (settlement year), with a return to poorer recruitment during the following cooler wPDO regime. These predictions are supported by the most recent stock assessment, but the estimates of recruitment for the most recent years are still fairly uncertain. Irrespective of the predictive ability of these models, the fact that recruitment appears to be related to the current climate regime may warrant reconsideration of the approaches used to calculate reference biomasses and set overfishing limits (NPFMC, 2008).

The models presented here will be useful in evaluating management strategies using operating models coupled to climate predictions to assess long term yield and the potential for overfishing (*sensu* Ianelli et al. 2011). Intuitively, methods based on long-term average recruitment will overharvest the population in one regime and underharvest in the next. However, previous simulation studies for the pollock, *Theragra chalcogramma,* fishery in the

Gulf of Alaska have shown that management strategies that incorporate shifts in recruitment regime can increase catches, but do so at a higher risk of overfishing (A'mar et al., 2009).

Distinguishing between the influences of fishing versus those of the environment on recruitment dynamics is problematic. Estimated fishing mortality was very high around the change in average recruitment around 1989. However, including mature male biomass in the model during the period 1980-1989 resulted in a negative estimated coefficient linking MMB to recruitment (results not show). While this is not conclusive evidence that fishing mortality does not influence recruitment dynamics over the range of observed stock sizes, it does suggest that fishing pressure may play a smaller role in determining recruitment. Ernst et al. (2012) also appear to question the impact of the fishery on recruitment dynamics, emphasizing the role of primiparous females and sub-legal mature and adolescent males.

Finally, this analysis does not explicitly consider the impact of spatial dynamics in recruitment processes. There is a considerable body of literature that addresses this (e.g. Ernst et al., 2005, 2012; Hinckley et al., 2010). The goal of this analysis was to develop a relatively simple model that could be used in conjunction with IPCC projections for management strategy evaluations, so inclusion of a spatial dimension was not feasible. The recruitment time-series is still relatively short and a better understanding of the dynamics of the system may develop as more years of data are added. Considering the track record of environment-recruitment relationships (Myers, 1998), the presented models are considered starting points and, as with the explanation of walleye pollock recruitment in the Bering Sea that invokes the Oscillating Control Hypothesis (Hunt et al., 2002), may need revision (Hunt et al. 2011).

Mechanism	Time lag	Influence on recruitment	Notes
1. Temperature	0-6	dome-shaped, positive, negative	Temperature can affect growth (Orensanz et al., 2004, 2007), length of brooding period (Moriyasu and Lanteigne, 1998), settlement patterns, migrations (Ernst et al., 2005), food availability, time in the pelagic phase (Kon, 1970) and predation (Wyllie-Echeverria and Wooster, 1998).
2. Predator density	0-4	negative	Predation by cod is hypothesized to affect crab most heavily until their terminal molt.
3. Cannibalism among early Juveniles	0-4	positive	Density-dependent cannibalism has been hypothesized to influence year class strength (Comeau and Conan, 1992).
4. Size frequencies	5-6	positive relationship to primipara	Primipara (females mating for the first time) have been hypothesized to contribute most to recruitment because of their spatial distribution in the oceanic current of the EBS (Parada et al., 2010).
5. Environmental cues for hatch	4	positive	Larval survival increases when the hatch coincides with favorable currents and food availability (Hinckley et al., 2010).
6. Favorable advection of pelagic larvae	4	positive	The position of egg extrusion in the currents of the Bering Sea influences both the location of settlement and the intersection with plankton blooms (Hinckley et al, 2010).
7. Availability of prey for juveniles and larvae	4	positive	Plankton blooms are a major food source for larval crabs. Their location and timing in relation to the hatch may be important for survival (Hinckley et al, 2010).
8. Interaction of food availability and temperature	0-4	positive	The effects of increasing temperatures on food availability can be complicated. There is a negative relationship between the date of ice retreat and primary productivity (Mueter et al., 2007).

Table 3.1. A summary of possible mechanisms behind variation in snow crab recruitment. Time lag is the number of years before recruitment the variable would have influenced year-class strength and is approximate.

Environmental index	**Influence**
Regional scale	
Winter Pacific Decadal Oscillation	1,2,5,6,7,8
Arctic Oscillation	5,6,7,8
Aleutian Low Pressure Index	5,6,7,8
North Pacific Index	5,6,7,8
Local Scale	
Winter sea surface temperature (Pribilof Islands)	1,2,5,6,7,8
Optimal winds	5,6,7,8
North/South winds	5,6,7,8
Wind stress along Aleutian peninsula (winter and summer)	5,6,7,8
Ice cover	1,2,5,6,7,8
Ice retreat	5,6,7,8
Wind mixing (summer)	5,6,7,8
Bottom temperature (summer)	1,2
Sea level pressure (summer and winter)	5,6,7,8
Strong winds	5,6,7,8
Wind mixing (summer)	5,6,7,8
Sea surface temperature (Bering Sea average; summer)	1,2,5,6,7,8

Table 3.2. Regional and local scale environmental indices considered as covariates to explain recruitment. See beringclimate.noaa.gov for specific definitions. Processes hypothesized to influence recruitment dynamics from Table 1 that may be influenced by an index are noted by number in the 'influence' column.

Model	**AICc**	**Parameters**
Average	68.86	2
Regime Average	52.56	3
Model 1	43.26	5
Model 2	37.93	6
Model 3	39.54	5
Model 4	35.50	6

Table 3.3. Models considered, AICc and number of parameters estimated.

	α_1	β_1	α_2	β_2	β_3	σ
Model 2	14.56	0.51	13.17	-0.38	0.34	0.39
	(0.11)	(0.11)	(0.13)	(0.13)	(0.13)	
Model 4	14.56	0.51	13.17	-0.42	0.30	0.37
	(0.11	(0.11)	(0.11)	(0.12)	(0.12)	

Table 3.4. Parameter estimates and standard errors (in parenthesis) for the two best models.

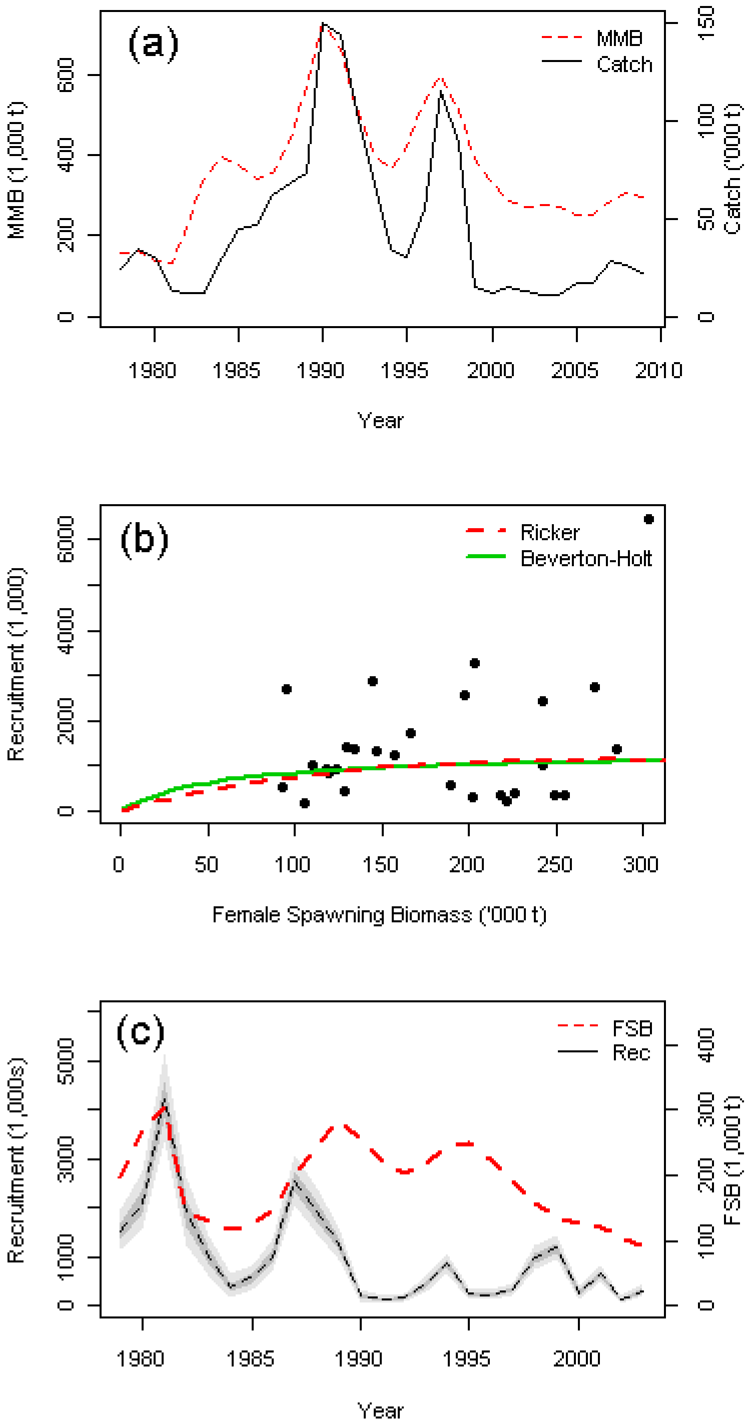

Figure 3.1. History of the snow crab fishery. A: Snow crab catch (‘000t) and estimated mature male biomass (‘000t) in the eastern Bering Sea. B: Estimated spawner (‘000t) – recruit (‘000,000t; lagged 5 years) curves for snow crab in the eastern Bering Sea. C: Estimated recruitment (dashed black line; 5th and 95th quantiles in grey) with estimated female spawning biomass at the time of the survey (dashed line).

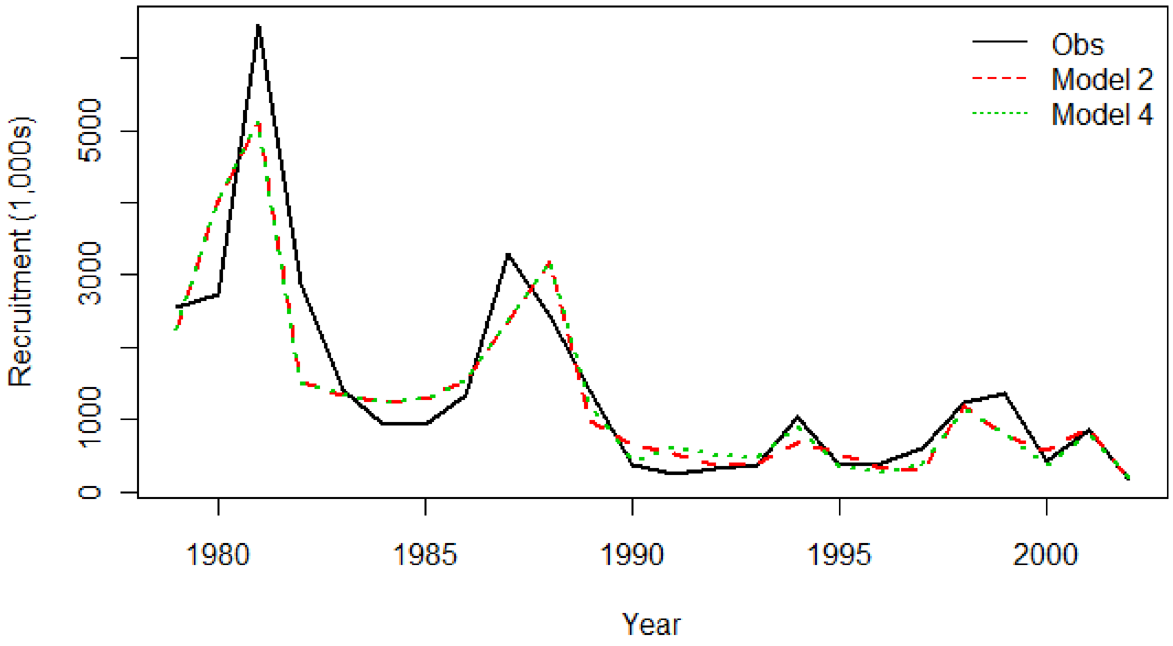

Figure 3.2. Fits to observed recruitment (solid line) from model 2 (FSB+wPDO+wWind; dashed line) and model 4 (FSB+PribSSTw+wWind; dotted line).

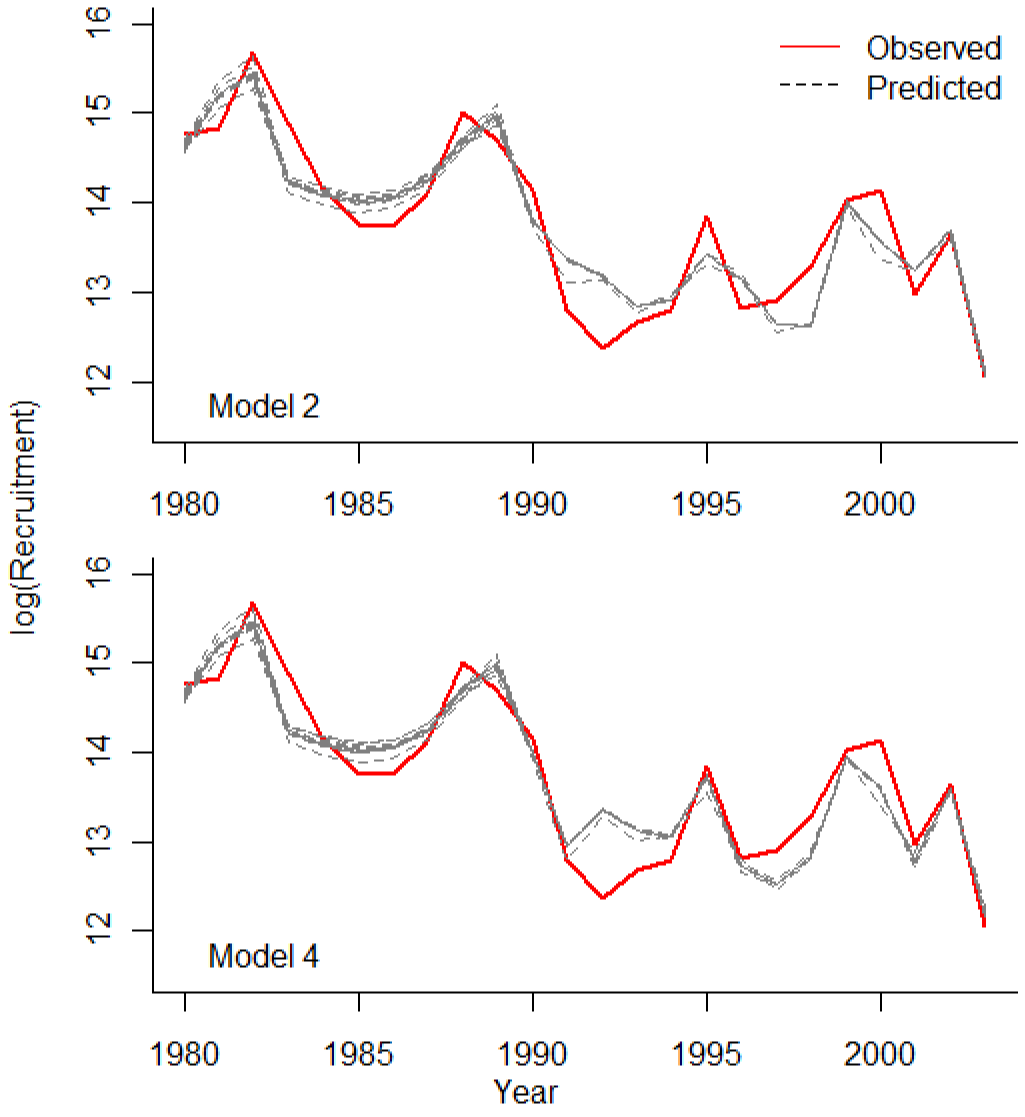

Figure 3.3. Fits from leave-one-out cross validation (dashed grey lines) for model 2 and 4 to observed recruitment (solid red line).

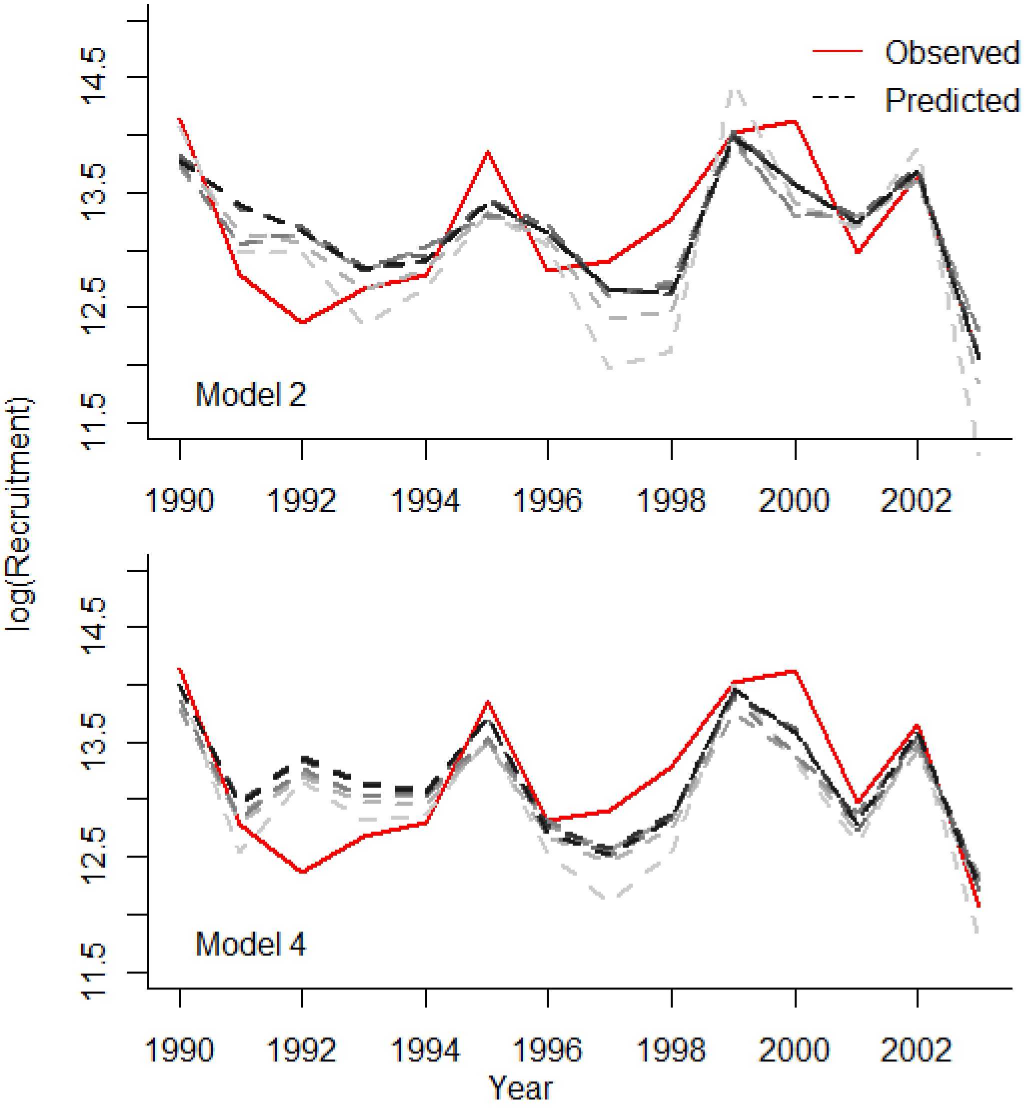

Figure 3.4. Predictions and fits from the retrospective analysis (dashed lines) for the model 2 (a) and model 4 (b) to recruitment from the years 1990-2003 (solid line). The darker the dashed line, the more years of data were used to develop the model used to predict recruitment.

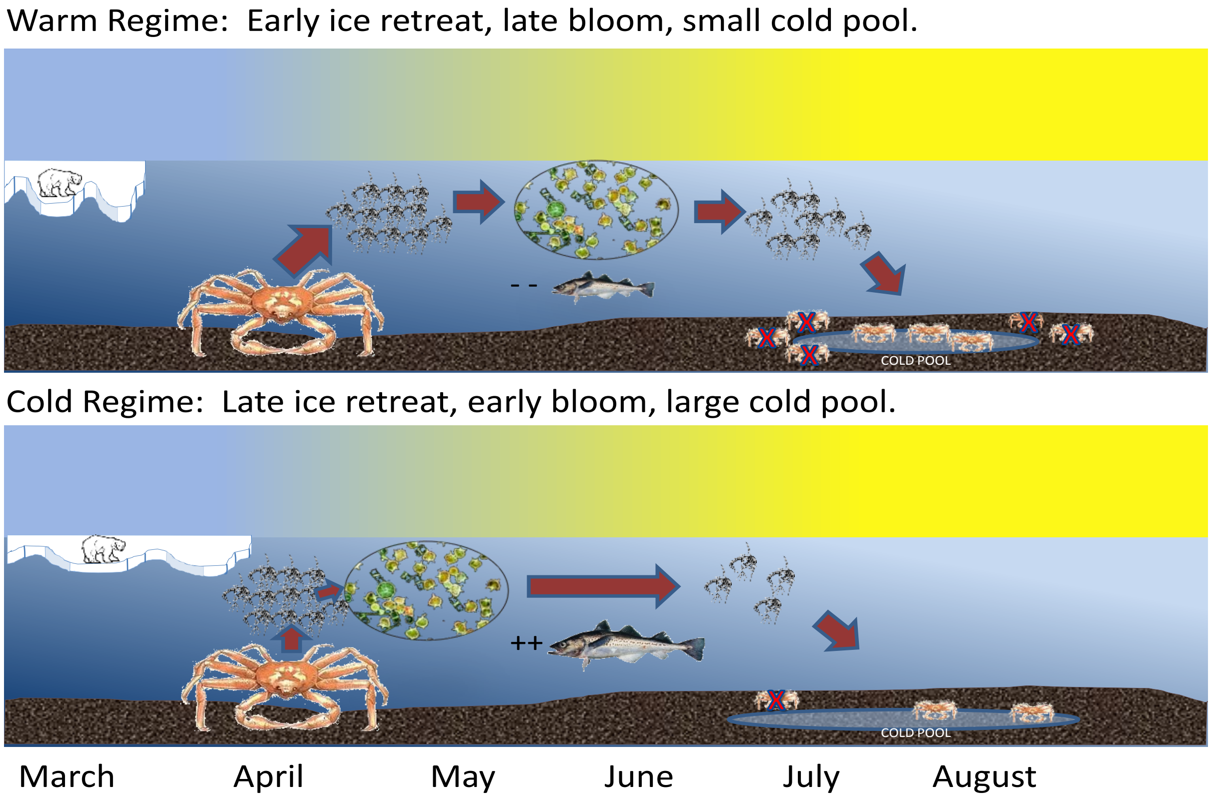

Figure 3.5 A cartoon representation of a potential hypothesis explaining the mechanism behind the presented models. On average, in warm regimes, early ice retreat leads to a bloom that coincides better with the pelagic stage of snow crab larvae. Less time spent in the water column, predation is lessened and recruitment is generally higher and related to FSB. In cold regimes, ice retreats later on average and the crab larvae miss the earlier ice-edge bloom. This mismatch, coupled with suboptimal advection, leads to poor recruitment in cold regimes. However, more available nursery grounds may result in better recruitment in colder years of cold regimes.

Appendix 3.A: Incorporating environmental relationships into the stock assessment

A framework for incorporating environmental data into the stock assessment to inform recruitment was also developed to test the relationship between recruitment and the wPDO described in Chapter 3. Recruitment during the period of time environmental effects were hypothesized to drive recruitment was forced in the assessment method by an index generated from the wPDO by adding the following negative log-likelihood component to the objective function.

$$L_{env} = 0.5\lambda_{env} \sum_v (EnvInd - RecDev)^2 / \sigma_{env}^2 \qquad (3.4)$$

where, L_{env}, is the quantity added to the objective function, λ_{env} is the weight applied to the data (a value of 1 was used for this analysis), v is the number of years of data that are used to force the recruitment deviations, *EnvInd* are the normalized predicted recruitments from Chapter 4, *RecDev* are the estimated recruitment deviations and σ_{env} is a measure of the closeness of the relationship between the environmental index and recruitment (here a value of 0.4 was used).

Seventy five 'dummy' indices were generated from a standard normal distribution to test the significance of the relationship between the PDO and recruitment. The stock assessment was fit to the data (including the environmental data) for both the 'dummy' indices and the true environmental index and the resulting likelihoods recorded. An environment-recruitment relationship similar to that outlined above is 'significant' at the $\alpha = 0.1$ level when incorporated inside the assessment method. The model incorporating the environmental data used in the rest of Chapter 4 had a negative log-likelihood two points lower than ~93% of the models that used a random series of environmental data. Estimates of recruitment and their standard deviations did not change with the inclusion of the environmental data. However, the weighting of the environmental data was fairly light because σ_{env} was relatively large. The worst fits to random data added ~50 to a base value for the objective function of 4044 (i.e. the objective value for the model without the environmental data), while the best added ~10. Different weighting may result in changes in confidence of estimates of recruitment. For example, setting σ_{env} to 0.01 forces recruitment to follow the environmental index, but also greatly increases the asymptotic standard deviations around the estimates (Fig. App. 3.1). Simulation testing beyond the scope of this study would be required to understand appropriate weighting for any given environment-recruitment relationship.

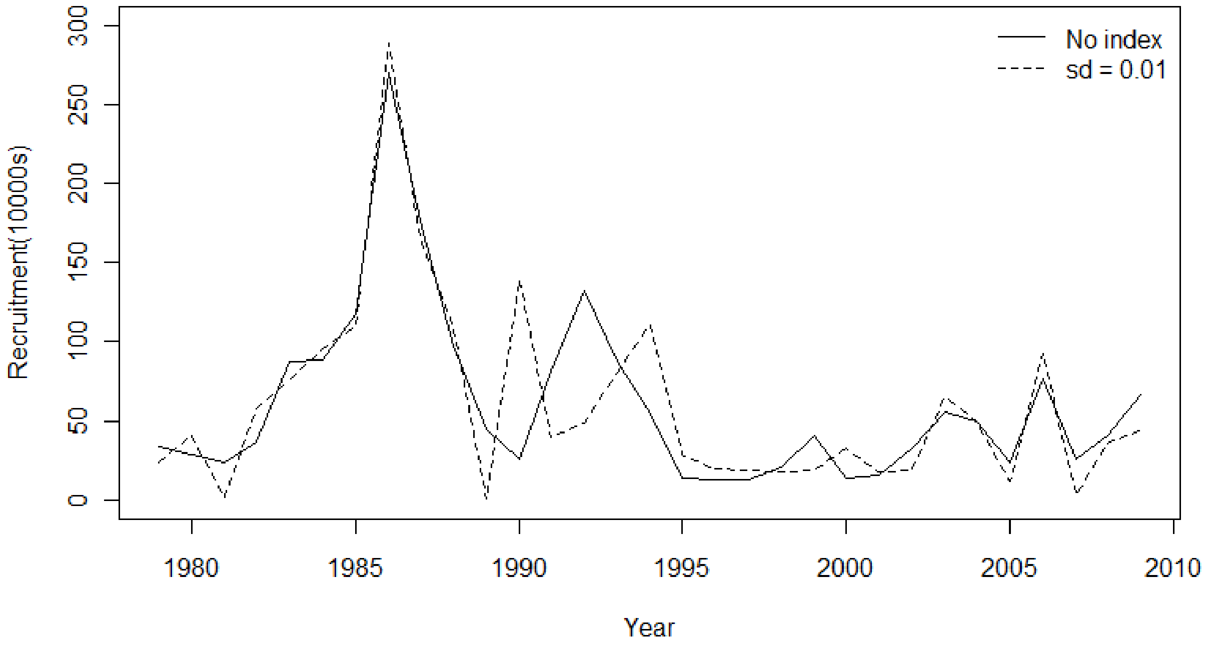

Figure App. 3.1. Estimated recruitment for assessments with no environmental index (solid line) and with an environmental index weighted with a standard deviation set to 0.01 (dashed line).

Chapter 4: Fisheries management strategies for regime-based systems: A management strategy evaluation for the snow crab fishery in the eastern Bering Sea

4.1.Introduction

Climate influences oceanic conditions which in turn influence the population dynamics of marine species and their fisheries (e.g. Mantua and Hare, 2002; Hollowed et al. 2001). The impact of "regime shifts" (the rapid reorganization of an ecosystem driven by changes in large-scale climate forcing (Overland et al., 2008)), on marine fisheries has received considerable attention in recent years (e.g. Litzow, 2006; Mueter et al., 2007). Regime shifts appear to occur on a decadal time scale and can influence the inferred productivity of stocks (e.g. Mantua and Hare, 2002; Rodionov and Overland, 2005). However, changes in productivity are rarely considered in stock assessments and when providing management advice.

Fishery managers often seek to maintain a population at or above a target biomass considered to provide the maximum sustainable yield (i.e. B_{MSY}). B_{MSY} is strongly influenced by the productivity of a stock. Target biomasses for stocks managed by the North Pacific Fisheries Management Council (NPFMC) are either set using an estimated stock-recruitment relationship or by specifying a proxy for B_{MSY} (NPFMC, 2007). The observations (or estimates) of spawning biomass and recruitment used to fit a stock-recruitment relationship or calculate the proxy for B_{MSY} are key to this process. Proxies for B_{MSY} are often calculated as the product of 35% of virgin spawning-biomass-per-recruit (i.e. $SBPR_{35\%}$) and average recruitment. Average recruitment is usually based on the full time series of spawning biomass and recruitment (e.g. Turnock and Rugulo, 2011). However, some estimates of recruitment may not be derived from the current 'regime' if recruitment is a function of climate regime. Target biomasses calculated from the entire time series will be higher than the ‘true’ target biomass during ‘low’ recruitment regimes and vice versa.

Management strategy evaluation (MSE) can assess the impact of assumptions regarding productivity on the ability of a management system to achieve its goals (Smith et al., 1999; Smith, 1994). An MSE consists of three components: an operating model, an estimation model and a harvest control rule (HCR). Operating models simulate populations which can be ‘observed’ and ‘assessed’ using an estimation model. An estimation model is a collection of equations that attempt to describe the dynamics of the population. Population parameters within the estimation model are estimated using non-linear optimization, and the estimates of these parameters are used to apply the HCRs, which are frameworks for deciding how catch limits for the fishery are to be calculated. Removals from the simulated populations are set to the catch limit from the HCR. A variety of operating models can be used to evaluate the impact of incorrect assumptions about the population dynamics in the estimation model (e.g. Punt, 2003; A’mar, 2009a, b). The relative performance of combinations of estimation methods and HCRs (often referred to as management strategies) across operating models can also be compared using metrics important to management such as long term yield and the probability of being overfished (NPFMC, 2007).

MSE can also assess the impact of projected climate change on the ability of management strategies to achieve their goals. This involves identifying the population dynamics processes influenced by environmental conditions and linking them in the operating model to projections of the environmental covariates that best describe the region of study (Hollowed et al., 2009).

Ianelli et al. (2011) and Amar et al (2009b) performed MSEs for walleye pollock (*Theragra chalcogramma*) in which operating models were linked to Intergovernmental Panel on Climate Change (IPCC, 2007) climate projections for the North Pacific. They concluded that calculating target biomasses based on expectations generated from only recent recruitments performs similarly to using all observations of recruitment to set target biomass under stationary recruitment, but may offer advantages when environmental conditions change over time.

Management should consider the possibility that the productivity of a stock can change over time (and sometimes suddenly) in ecosystems with regime-based climate. We seek to understand the risk, relative performance and trade-offs associated with HCRs that incorporate changes in expected recruitment induced by climate regime shifts compared to HCRs that ignore changes in average recruitment. These questions are explored in the context of the snow crab (*Chionoecetes opilio*) fishery in the eastern Bering Sea (EBS). The U.S. domestic fishery for EBS snow crab only lands large male crab and is historically productive, with a maximum estimated biomass of over 680,000 tonnes in the early 1990s (Turnock and Rugulo, 2011). However, both biomass and catch have been quite variable, and EBS snow crab was declared overfished in 1999 when it was assessed to have dropped below its Minimum Stock Size Threshold (MSST) (Turnock and Rugulo, 2011). A plan to rebuild the stock to its B_{MSY} proxy within 10 years (NPFMC/NMFS, 2000) was implemented in 1999 but declared a failure in 2009 (Turnock and Rugulo, 2011). The stock was declared rebuilt in 2011 when the stock was estimated to be above the $\mathrm{B}_{\mathrm{MSY}}$ proxy.

"Regime shifts" have occurred in the Bering Sea most recently in 1977, 1989, and possibly 1999 (Overland et al., 2008). These shifts influenced recruitment of species in the region (e.g.

Adkison et al., 1996; Hunt et al., 2011; Mantua et al, 1997; Wilderbuer et al., 2002), and a recently suggested relationship between snow crab recruitment and the winter Pacific Decadal Oscillation (wPDO) proposes that snow crab recruitment is also influenced by regimes shifts. Szuwalski and Punt (2012a) propose that drivers of recruitment oscillate between female spawning biomass and the wPDO, with the change point linked to shifts in the average wPDO (see below for further discussion). The stock was declared overfished ten years after the 1989 regime shift (seen strongly in the wPDO) at which time recruitment decreased markedly and ten years is the time a male crab takes to enter the mature population after fertilization. It is therefore possible that the overfished declaration and initial failure of the rebuilding plan was not solely the result of fishing pressure, but was also due to a change in productivity related to the 1989 shift in the wPDO.

The current HCR may not be appropriate if regime shifts influence recruitment of EBS snow crab. In order to evaluate the impact of regime-based dynamics on management performance, the trade-offs associated with two potential HCRs for the snow crab fishery (one of which considers shifts in productivity) are examined under three operating models that use different methods to simulate future recruitment (Fig. 4.1).

4.2.Methods

4.2.1 Estimation method

The estimation method for all simulations is closely related to the currently-used estimation method for EBS snow crab; it is size-based with considerations for sex, maturity state and shell condition (Turnock and Rugulo, 2011; see Szuwalski and Punt 2012b, Appendix 4.A for details). The estimation method involves fitting a population model to data from the directed fishery, bycatch from the trawl fishery and data from the National Marine Fisheries Service (NMFS) summer survey. One major difference between this estimation method and that applied by Turnock and Rugulo (2011) is the penalties on deviations in fishing mortality applied by Turnock and Rugulo (2011) are dropped, as suggested by Szuwalski and Punt (2012b). Removing these penalties results in poorer fits to the data, but considering the large bias introduced into estimates of the fishing mortality corresponding to MSY (F_{MSY}) and selectivity identified by Szuwalski and Punt (2012b), including these penalties in the estimation method would only demonstrate why penalties on fishing mortality should not be used. The most recent stock assessment included two years of data from an additional survey, but this survey is unlikely to take place regularly in the future so it is ignored here. Simulated data representing the directed pot fishery, the NMFS summer trawl survey, and bycatch data from the groundfish trawl fishery are used during each year of the projection period to estimate trends in fishing mortality, numbers-at-length, and mature male biomass (MMB) using the estimation method. Production models and other simpler estimation methods were not considered because they estimate B_{MSY} and F_{MSY} poorly in this system (Punt and Szuwalski, 2012).

4.2.2 Operating models

The operating models are based on the population dynamics model on which the estimation method is based and represent the ‘true’ population dynamics. Six “scenarios” are considered where a “scenario“ is the set of simulations resulting from the application of an HCR to a specific operating model. Sixty-five simulations were performed for each scenario (all projected 50 years). The parameters used to define the population dynamics model for each of these simulations were generated by sampling from the posterior distribution obtained by fitting the operating model to the actual data for EBS snow crab using a Markov chain Monte Carlo

(MCMC) algorithm. Parameter vectors were generated by implementing a 10% burn-in on 1,000,000 cycles of an MCMC algorithm and selecting a thinning ratio that returned the desired number of parameter vectors. Evidence of non-convergence of the MCMC algorithm was checked using several diagnostic statistics (e.g. lack of autocorrelation and the Geweke statistics; Gelman et al., 2004). Only the manner in which recruitment is projected changes among operating models. These three methods of projecting recruitment are described below.

4.2.2.1 Status quo recruitment

The first operating model ("Status quo"; Fig 4.2a.) projects recruitment using a Beverton-Holt spawner-recruitment relationship with steepness set such that F_{MSY} is equal to $F_{35\%}$ because this is an implicit assumption of the current HCR. Recruitment is related to MMB (also an implicit assumption of the current HCR) with a five-year lag. The projected recruitments are subject to bias-corrected lognormal error with a standard deviation of the log similar to that estimated by fitting the population dynamics model to the actual data for EBS snow crab ($\sigma = 0.75$). The performance of the current HCR under this operating model serves as a reference for the other operating model/HCR combinations because it is consistent with the assumptions of the current HCR.

4.2.2.2 Regime-based recruitment

A shift in average recruitment in 1995 divides the fishery's history into a period of relatively 'high' recruitment (1984-1995) and one of lower recruitment (1996-2008). The second operating model ("Regime-based"; Fig 4.2b) shifts between a 'high' and 'low' recruitment regime every ten years. The average recruitments during these regimes are similar to the observed averages (the high regime average recruitment is roughly double that of the low regime) and the overall average recruitment is equal to the average recruitment in the *status quo* operating model. Process error is again generated from a bias-corrected lognormal distribution with $\sigma = 0.75$. This operating model assumes no relationship between spawning biomass and recruitment, and represents an idealized regime-based system to evaluate the performance of regime-based HCRs. The true B_{MSY}, and hence the basis for evaluating whether the stock is really overfished, for this operating model is defined as $SBPRF_{35\%}$ multiplied by the average recruitment within a regime.

4.2.2.3 Oscillating control recruitment

The third operating model ("Oscillating control"; Fig. 4.2c) is also regime-based, but incorporates the mechanism suggested by Szuwalski and Punt (2012a) in which regimes are demarcated by shifts in climate regime. Recruitment is 'high', and related to female spawning biomass (FSB) when the wPDO is 'warm'; recruitment is 'low' and related the environment when the wPDO is 'cool' (see Fig. 4.3a for fits of this recruitment model to actual data for EBS snow crab). Hence, control of recruitment 'oscillates' from spawning biomass to environmental effects depending on the current climate regime. Model 1 from Szuwalski and Punt (2012a) was selected to project recruitment because it uses only the wPDO as a predictor during 'cool' regimes. Incorporating the wPDO in the projection model is important because it defines the point at which recruitment dynamics shift.

Recent recruitments during a warm period not used when developing the original wPDO/recruitment model indicate an increase in recruitment (Turnock and Rugulo, 2011) larger than would be predicted by parameter estimates in Szuwalski and Punt (2012a) given the recent low level of FSB. This suggests a density-related effect on recruitment not originally observed, and higher productivity at lower spawning biomasses. Consequently, the relationship between recruitment and FSB is modeled using the function in Figure 4.3b. Productivity is higher for low spawning biomasses, but recruitment generation reverts to the model from Szuwalski and Punt (2012a) when FSB is greater than 350,000 tonnes. A cap of three million recruits is imposed before error is added. Ten IPCC climate models (all under the A1B emission scenario) were used to project the wPDO. Overland and Wang (2007) selected these ten models given their ability to simulate large-scale aspects of the climate in the eastern Bering Sea (e.g., sea ice area). Control of recruitment oscillates from FSB to the wPDO at the midpoint between the average wPDO of the 'cool' and 'warm' regimes in the oscillating control model (horizontal dotted line in Figs. 5.3c-f). FSB determines recruitment when the average projected wPDO is above the horizontal line in Figures 5.3c-f; the wPDO determines recruitment when the wPDO is less than the cutoff. The standard deviation of log-recruitment for this operating model is based on the fit of the model of Szuwalski and Punt (2012a) to the actual estimates of recruitment ($\sigma = 0.41$).

The true B_{MSY} during cool regimes is calculated as the average recruitment for the recruitment regime multiplied by $SBPRF_{35\%}$. Determining the 'true' B_{MSY} for warm phases is not straightforward because recruitment is related to FSB, but only males are fished. The number of males necessary to inseminate the females present is an appropriate way to define the true B_{MSY} under these conditions. However, this number is unknown. Males expend less than 2.5% of sperm reserves with each ejaculation and can copulate with more than one female during a mating season. Additionally, males and females have asynchronous maturity schedules and can store sperm from year to year (Rondeau and Sainte-Marie, 2001). Considering this, and the flexibility of male mating strategies as seen through time spent guarding a receptive female (Rondeau and

Sainte-Marie, 2001), the required male to female sex ratio is likely less than one. In laboratory studies, Rondeau and Sainte-Marie used effective sex ratios from 0.06 to 0.39 to examine the influence of sex ratio on sperm allocation and guarding time. Spermathecal load (the amount of sperm a female stored) declined with sex ratio. The definition of B_{MSY} for this operating model for cold regimes (and also the current management strategy) leads to a sex ratio of roughly 0.33 on average (3 females for each male), which is at the high end of the range tested by Rondeau and Sainte-Marie and is relatively conservative. Therefore, the true B_{MSY} during warm regimes is calculated in the same manner as during cool regimes for this operating model.

4.2.3 Harvest Control Rules

HCRs determine removals using estimates of biomass and calculated target biomasses. This study considers two HCRs that differ in the assumptions made regarding the productivity of the stock as seen through recruitment.

4.2.3.1 Status quo harvest control rule

The North Pacific Fisheries Management Council (NPFMC) and the Alaska Department of Fish and Game (ADFG) jointly manage EBS snow crab. The NPFMC primarily sets the overfishing level (OFL), the Acceptable Biological Catch (ABC), and the MSST, and the ADFG sets the total allowable catch (TAC). The TAC must be lower than the ABC (which for Alaska crab is essentially the same as the OFL) so that catches removed from the stock for these simulations are the lower of the OFL and the TAC from the ADFG HCR.

The OFL is determined using estimates of MMB from the NPFMC assessment and a HCR (Eqns. 4.1-3) (NPMFC, 2007).

	Stock status level	F_{OFL}
(4.1)	$\frac{MMB_{current}}{MMB_{35\%}} > 1$	$F_{\text{OFL}} = F_{35\%}$
(4.2)	$\beta < \frac{MMB_{current}}{MMB_{35\%}} \leq 1$	$F_{\text{OFL}} = F_{35\%} * \frac{\frac{MMB_{current}}{MMB_{35\%}} - \alpha}{1-\alpha}$
(4.3)	$\frac{MMB_{current}}{MMB_{35\%}} \leq \beta$	Directed fishery F = 0

where α determines the fishing mortality rate used to compute the OFL, F_{OFL}, as MMB decreases to β*MM$B_{35\%}$ and β determines the threshold level of biomass at or below which directed fishing is prohibited. $\text{MMB}_{35\%}$ is a proxy for B_{MSY}, equal to average recruitment multiplied by $\text{SPPR}_{35\%}$.

The TAC is determined by the State of Alaska using the following HCR to determine fishing mortality on mature males:

	Stock status level	*Fishing mortality*
(4.4)	$TMB < \gamma$	0
(4.5)	$\gamma \leq TMB \leq \rho$	$0.75 * \text{F}_{\text{MSY}} * \frac{\frac{TMB}{AvgTMB} - \delta}{1-\delta}$
(4.6)	$TMB \geq \rho$	$0.75 * \text{F}_{\text{MSY}}$

where TMB is total mature (male and female) biomass at the time of the survey (note that the OFL control rule uses only MMB), γ is the minimum TMB threshold for opening the fishery, F_{MSY} is assumed to equal to the assumed natural mortality rate (0.3yr^{-1} according to ADFG), AvgTMB is the average estimate of TMB at the time of the survey from the stock assessment over the years 1983 – 1997 (Turnock and Rugulo, 2011), ρ is the TMB at which the full exploitation rate is applied (currently 921.6 million lbs), and δ is a constant (0.35) that determines the slope of the relationship between fishing mortality and TMB. The fishing mortality is zero if TMB is less than γ. The TAC computed according to the ADFG harvest strategy is the lesser of the outcome from the above HCR and 58% of exploitable legal male abundance.

The federal and ADFG HCRs depend on target biomasses. The federal HCR is based on a proxy for B_{MSY} which is $SPPR_{35\%}$ multiplied by average estimated recruitment for 1979-present. The ADFG HCR depends on the TMB during the years 1983-97, which corresponds to a period of relatively good recruitment. Assumptions about incoming recruitment can affect the validity of these target biomasses.

The federal HCR is applied without the ADFG HCR for the status quo operating models to demonstrate the conservative nature of the ADFG HCR compared to the federal HCR. All other scenarios are based on the combination of the two HCRs.

4.2.3.2 Regime-based harvest control rule

HCRs that limit the observations of recruitment used to calculate expected recruitment can be formulated. Averages calculated in sliding windows might be used to calculate expected recruitment (e.g. A'mar et al., 2009a), but shifts in productivity can be sudden in regime-based systems. A sliding window approach would identify gradual changes in productivity, but sudden shifts would not be well-captured until many years after the regime shift. Consequently, estimated target biomasses may be higher than the productivity of the stock would imply if average recruitment dropped, which would increase the likelihood of falsely declaring a stock overfished, and vice versa.

Sudden shifts in productivity are addressed here by incorporating an algorithm such as Rodionov's sequential t-test analysis for regime shifts (STARS; Rodionov, 2004) into the HCR. STARS assumes a length of regime and defines a 'previous regime' based on the assumed regime length and available data. Next, the deviations of each new year's data from the previous regime's average are compared to a t-distribution defined by the mean and variance of the observed data for the previous regime. A new regime is considered to have possibly begun when the deviation for the new year is significantly different ($p<0.1$) from the mean of the previous regime. A shift in regime is 'confirmed' when the algorithm has progressed the number of years into the 'new' regime that is the assumed length of the regime, without encountering observations that are inconsistent with a shift.

STARS was incorporated into an HCR in this analysis to detect changes in recruitment 'regime'. HCRs with target biomasses based on expected recruitment for the current recruitment regime (with "recruitment regime" defined by STARS) are used here to capture sudden changes in productivity and will be referred to as "regime-based". An increase in overfishing due to poor identification of regime shifts can occur when HCRs incorporate STARS (A'mar et al., 2009a). However, snow crab are observed as 'recruits' to the survey several years before recruiting to the exploitable population, so it is possible that changes in recruitment regimes will be identified more easily. The regime-based HCR presented only changes the way expected recruitment to the smallest size-class in the model is calculated and the hence the average recruitment term when computing the B_{MSY} proxy. Spawning biomass-per-recruit calculations used to calculate B_{MSY} therefore do not depend on regime, i.e. regime shifts are assumed only to influence recruitment to the smallest size-class in the population dynamics model, not natural mortality, growth or any other biological processes. This is reasonable because environmental conditions are thought to impact early life stages most heavily (Kruse et al., 2007).

A regime-based alternative for the application of the ADFG HCR is difficult to formulate because the target biomass is based on TMB during 1983-97. However, the relevant laws (Alaska Statutes, 2012) state that the target biomass is meant to be analogous to B_{MSY}. Regime-based versions of the ADFG HCR were developed in a similar manner to the federal HCR by estimating a proxy for both F_{MSY} and B_{MSY}, with the exception that B_{MSY} is based on total mature biomass instead of MMB. F_{MSY} is set as $F_{35\%}$ in the same manner as the federal HCR (i.e. the spawning biomass referenced is MMB, not TMB) because only males are fished. These proxies for B_{MSY} and F_{MSY} are used in place of the ADFG reference points.

4.2.4 Performance metrics

Management strategies applied by the NPFMC seek to achieve maximum sustainable yield while avoiding overfishing and overfished stocks. Consequently, the most important metrics for measuring performance of an HCR are long-term yield, the probability of a stock becoming overfished, and the probability of overfishing. The probability of correctly identifying overfishing and overfished statuses can be evaluated because both the ‘true’ and ‘estimated’ states of the fishery are known in the operating model. Performance metrics are calculated over the last 40 years of the simulation period to allow two full cycles of 10-year alternating regimes. Finally, relative error in the estimates of B_{MSY}, MMB and the OFL are examined to determine if the nature of biases in these quantities noted by Szuwalski and Punt (2012b) are compounding, self-correcting or persistent. Relative error is formulated as:

$$E_t^{i,j} = (\hat{Q}_t^{i,j} - Q_t^{i,j})/Q_t^{i,j} \tag{4.7}$$

where $E_t^{i,j}$ is the relative error for quantity *i* during year *t* for simulation *j*, $Q_t^{i,j}$ is the true (i.e., based on the operating model) value for quantity *i* during year *t* for simulation *j*, and $\hat{Q}_t^{i,j}$ is the estimate of quantity *i* during year *t* for simulation *j* from the estimation method. The median absolute relative error (MARE, a measure of error and bias) and the mean-median [mean (over simulations) median (over years)] relative error over the last forty years of the projections (MMRE, a measure of bias) for selected management quantities summarize the results and follow the format: B_{MSY} [0.02; 0.03], where '0.02' is the MARE and '0.03' is the MMRE. An MMRE and MARE of 0.0 represent perfectly accurate (MMRE and MARE) and precise (only for MARE) estimates; MARE is always positive, but MMRE can be negative.

4.3.Results

4.3.1 Performance of the status quo HCR

The status quo HCR returned the mean-median MMB during the years 2020 to 2059 to 108% of $MMB_{35\%}$ (Fig. 4.4a). The true population was not overfished in any of the scenarios because 1) the ADFG HCR combined with the federal HCR is conservative in comparison to the federal HCR on its own (Fig. 4.4a) and 2) the estimation method underestimates $F_{35\%}$ (Fig. 4.4b). The negative bias in $F_{35\%}$ gradually lessens over the course of the projections to -5% and is accompanied by a decreasing estimated B_{MSY} proxy ($MMB_{35\%}$) [0.07; 0.0] (Fig. 4.4c). This pattern in $F_{35\%}$ is nearly identical in all simulations. The median estimated $MMB_{35\%}$ over the projection period was essentially unbiased by the end of the projection period. The OFL had an overall bias of 4% (MARE = 0.18, Fig. 4.4d). The 'true' OFL is the catch that would result from applying the 'true' $F_{35\%}$ as calculated from the operating model. Overfishing (based on a comparison of the true OFL and the actual removals) only occurred in 4% of the simulations (Fig. 4.5a). The mean-median catch over the projection period was 40,600t, with a standard deviation of 2,200t.

Mean-median MMB fluctuated widely around the estimated $MMB_{35\%}$ when applying the status quo HCR to data generated from the "regime-based" operating model (Fig. 4.6a). MMB was well below the estimated $MMB_{35\%}$, but above the true $MMB_{35\%}$ in low recruitment regimes; MMB was nearly at the true $MMB_{35\%}$ on average and well above the estimated $MMB_{35\%}$ in high regimes. This is intuitive because the estimation method uses all available recruitment observations (rather than those that pertain to the current regime) to calculate the B_{MSY} proxy. Estimates of $MMB_{35\%}$ [0.70; 0.87] were highly biased because the status quo HCR could not track the true target biomass (Figs. 4.6a & 4.6c). This bias, coupled with positive bias in MMB [0.10; 0.03] (Fig. 4.6b) resulted in a slightly negatively biased and relatively imprecise estimate of the OFL [0.23; -0.03] (Fig. 4.6d). The mean-median catch over the projection was 27,600t (SD 12,800t). Overfishing never occurred within high regimes, but occurred 19% of the time during low regimes (Fig. 4.5b).

Projected MMB and catch under the status quo HCR in the oscillating control operating model are uncertain because of the variation among IPCC models used for projection (Fig. 4.7a).

Recruitment during periods in which recruitment is related to FSB can be much higher than when it is driven by the wPDO (see figs. 4.3c-f for examples). Large variability in recruitment translated to a large loss of precision and positive bias in the estimated $MMB_{35\%}$ [0.73; 0.67] (Fig. 4.7c). Positive bias in $MMB_{35\%}$ resulted in a negative bias in the OFL and relatively low precision [0.19,-0.10] (Fig. 4.7d). The median catch over the projection for this scenario was 20,200t (SD 5,200t) and overfishing occurred with an average probability of 3% (Fig. 4.5c).

4.3.2 Performance of the 'regime-based' HCR

The utility of regime-based HCRs in non-regime based systems is determined by the magnitude of bias and imprecision caused by using a selection of observations to calculate average recruitment. The bias in $MMB_{35\%}$ did not increase markedly under the status quo operating model, but the precision deteriorated [0.29; 0.05] (Fig. 4.8c). This translated to a loss of precision and additional bias in the estimate of the OFL [0.21; 0.08], and also to an increased probability of overfishing (to 23%; Fig. 4.5a). The median catch over the projection period was 39,700t (SD 2,130t), 3% lower than under the status quo HCR.

The value of a regime-based HCR is most easily seen when they are applied to the regime-based operating model. Compared to the status quo HCR, estimation of $MMB_{35\%}$ was nearly unbiased (-0.02 for regime-based vs 0.87 for status quo) and much more precise (MARE = 0.35 vs 0.70) under the regime-based HCR because it is able to track the true $MMB_{35\%}$ (Figs. 4.9a and c). The precision of the OFL did not change, but the bias changed sign and increased to 0.09 (Fig. 4.9d). This was reflected in an increased probability of overfishing (19%; 33% in high regimes, 3% in low regimes; Fig. 4.5b), with particularly high probability around the years when the regime switches from high to low. There was an increase in mean-median catch (28,500t) and a decrease in variability (sd = 10,100t) when compared to the status quo HCR.

Projected MMB, B_{MSY} and catch were highly variable in the scenarios in which recruitment followed the oscillating control model and the regime-based HCR was applied (Fig. 4.10a). Estimates of $MMB_{35\%}$ [0.27; 0.03] improved compared to the status-quo HCR (Fig. 4.10c), while the bias in the OFL did not change magnitude, but changed signs [0.16; 0.05] (Fig. 4.10d). There was a slight increase in the probability of overfishing compared to the status quo HCR (from 3% to 6%, respectively; Fig. 4.5c), but the mean-median catch increased to 22,600t (SD 3,600t) compared to the status quo HCR.

4.4. Discussion

Incorporating changes in inferred productivity in management strategies can reduce bias in calculated target biomasses, increase catches, and reduce variability in catches in regime-based systems. However, imprecision in target biomasses increases, the probability of overfishing increases and yield is lost when applied to non-regime-based systems. Although observing recruits several years before entering the fishery was a potential opportunity to better identify shifts in regime, overfishing still increased under regime-based HCRs, mirroring the results of A’mar et al. (2009a) for walleye pollock in the Gulf of Alaska. There are two reasons for this: 1) shifts in regime were not captured in exactly the year they occurred because within-regime variability in recruitment was high; and 2) estimates of the B_{MSY} proxy were noisy because average recruitment was calculated using relatively few data points.

Regime-based target biomasses are most useful when the underlying dynamics of a system are truly regime-like and shifting from a ‘high’ regime to a ‘low’ regime. A stock may be falsely declared overfished under these circumstances if the status quo HCR is applied, which could result in potentially costly rebuilding plans. This did not happen in this study because 1) the fishing mortality from the ADFG HCR is conservative, and 2) the differences in average recruitment among regimes were not large enough for this to occur without higher fishing mortality. However, high fishing mortalities coincided with the overfished declaration for the actual fishery. Shifting from a ‘low’ to a ‘high’ regime under regime-based management strategies usually leads to a reduction in fishing pressure because the estimate of the B_{MSY} proxy increases. This potentially reduces the OFL, which is counter to the actions taken by a non-regime-based management strategy. Consequently, regime-based HCRs have some utility, but only when the dynamics are truly regime-based. The deeper (and more complicated) issue that needs to be addressed before implementation of regime-based HCRs is determining when population dynamics are truly regime-like.

The accuracy of estimates (or proxies) of $\mathrm{F}_{\mathrm{MSY}}$ is potentially more important than designation of a target biomass. The target biomass determines what proportion of $\mathrm{F}_{\mathrm{MSY}}$ is applied: the full $\mathrm{F}_{\mathrm{MSY}}$ is applied when the current mature biomass is above $\mathrm{B}_{\mathrm{MSY}}$; some fraction of $\mathrm{F}_{\mathrm{MSY}}$ is applied when the current mature biomass is below $\mathrm{B}_{\mathrm{MSY}}$. This analysis assumed that the population processes that determine the proxy for $\mathrm{F}_{\mathrm{MSY}}$ (e.g. natural mortality, growth, and selectivity for crabs large enough to have recruited to the survey gear) do not change from regime to regime. This may be a poor assumption if a stock has documented changes in one of these processes in response to environmental changes. However, studies (tagging or otherwise) required to identify these changes every time a regime shift is suspected for EBS snow crab are not currently feasible. Similarly, applying the estimation method over a much shorter time period to generate 'regime-specific' estimates of F_{MSY} is not likely to lead to reliable estimates because of the lack of contrast over short time series. Many years of data that span high and low stock sizes that have undergone high and low fishing pressure are necessary to truly understand the response of a stock to fishing (Magnusson and Hilborn, 2007).

As for the future of the snow crab fishery, levels of catch will likely never reach those experienced in the 1980s due to the well-defined harvest strategy. Future recruitment may be typified by increasingly frequent periods of good recruitment which will be seen as higher levels of MMB. However, some IPCC models suggested that MMB may never reach historical highs. In general, these projections should not be interpreted too strongly—their main utility here is for identifying HCRs which are robust to uncertainty and not for making predictions. Environmental-recruitment relationships often collapse with the addition of new data (Myers, 1998), and a spatial component not accounted for in this analysis may influence recruitment (e.g. Parada et al., 2010). Additional years of recruitment estimates (particularly after the recent juxtaposition of very cold and very warm years in the Bering Sea) should be informative as to the veracity of the oscillating control model.

Regime-based HCRs are risky when they are applied to non-regime-based systems and there is a positive bias in F_{MSY}; this may result in an unchecked downward trajectory in the mature biomass (i.e. recruitment overfishing—this did not occur in any of the scenarios presented here because F_{MSY} was under-estimated). Regime-based dynamics cause population dynamics to be non-stationary, but most management strategies are based on stationary population dynamics. Incomplete understanding of non-stationary dynamics in a system requires a precautionary approach if stationary dynamics are assumed in the management strategy. Future research on methods to definitively declare regime shifts and identify stocks within those systems influenced by these shifts are likely required before regime-based HCRs will be implemented by management agencies.

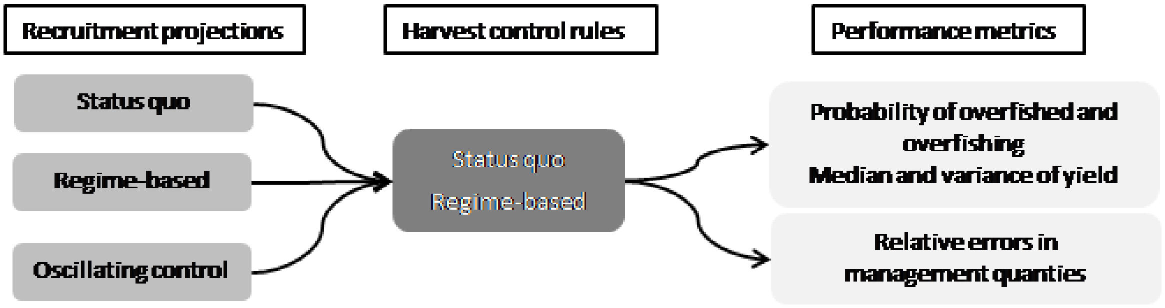

Figure 4.1. Schematic of the components and flow of the management strategy evaluation.

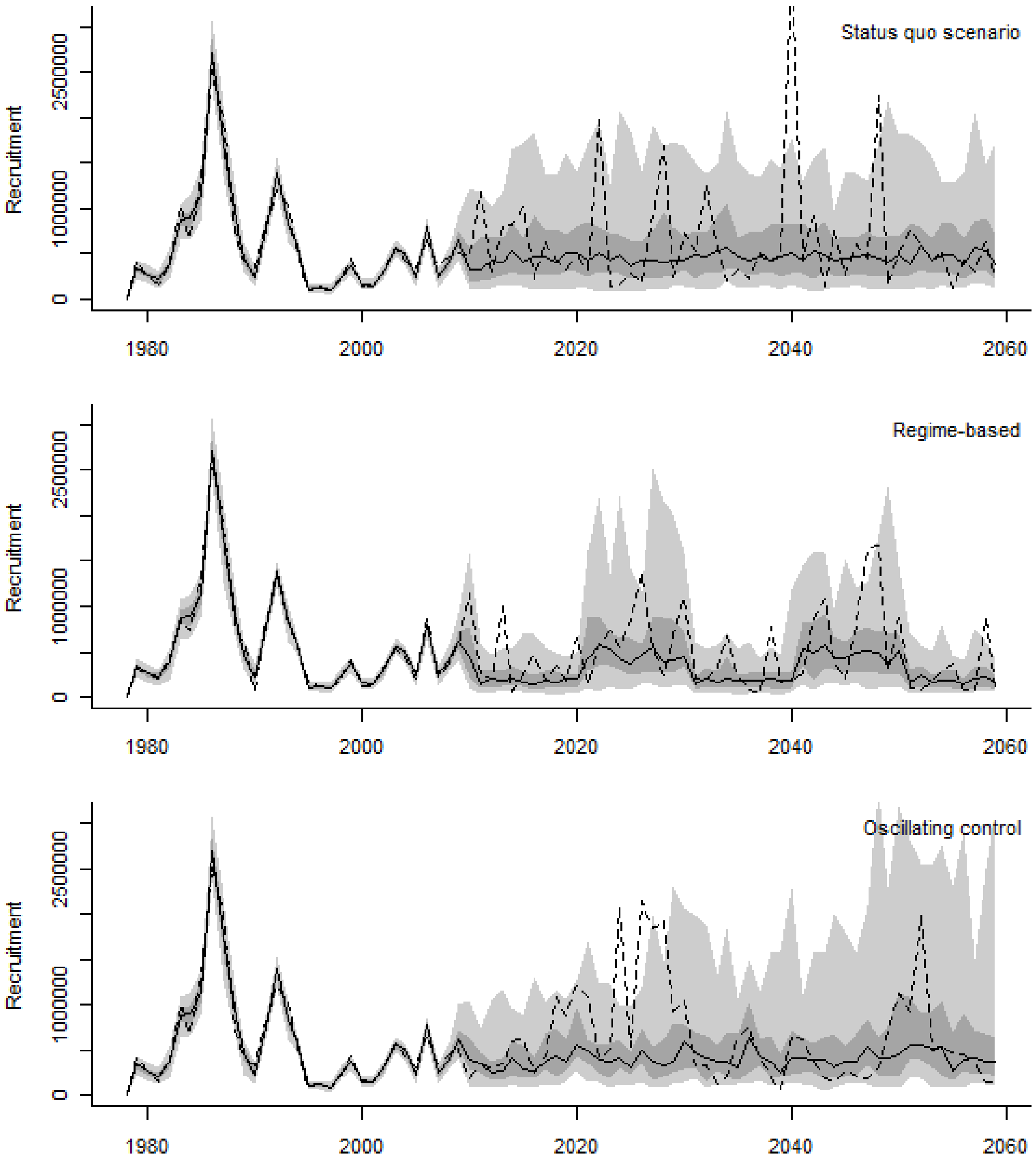

Figure 4.2. Projected recruitment for each operating model. The "Status quo" operating model (a) represents a system in which there is a single relationship between MMB and recruitment over the projection period. Recruitment is not related to spawning biomass and alternates between 'high' and 'low' recruitment regimes in the "Regime-based" operating model (b). Recruitment in the "Oscillating control" operating model (c) shifts from a period of 'high recruitment' related to female spawning biomass to periods of 'low' recruitment related to projections of wPDO. Shifts between drivers occur with changes to the average value of the wPDO (Fig. 5.3). Light grey represents the 5th and 95th inter-simulation interval; dark grey is the 25th and 75th. The black line is the median, and the dashed line is one random realization of future recruitment.

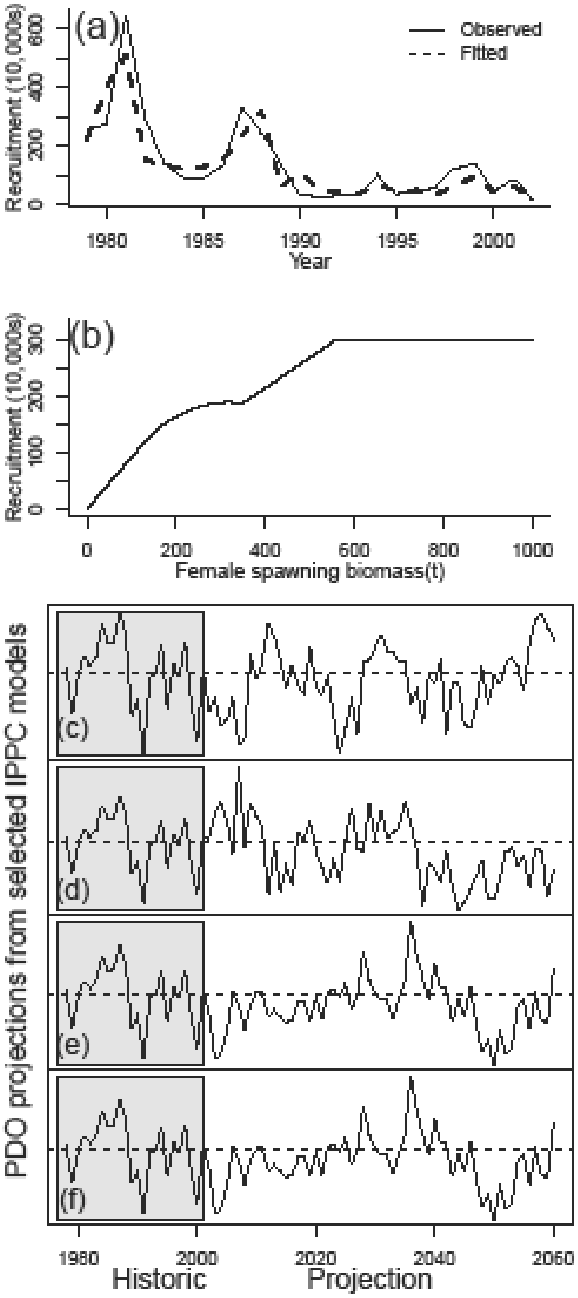

Figure 4.3. Stock-recruitment-environment relationships. Panel a: fits of the oscillating control model from Chapter 4 (recruitment is in the year crabs recruit to the survey i.e. 5 years after fertilization). Panel b: realized stock recruitment relationship used in the 'oscillating control' operating model. Panels c-f: An arbitrary selection of 4 of the 10 projected Pacific Decadal Oscillation from Overland and Wang (2007). The horizontal dashed line denotes the cut-off at which control of recruitment changes from female spawning biomass (above the line) to the wPDO. Note that the 'historic' period had a sustained time period in which the wPDO was anomalously high and this pattern is seen very few times in the projections.

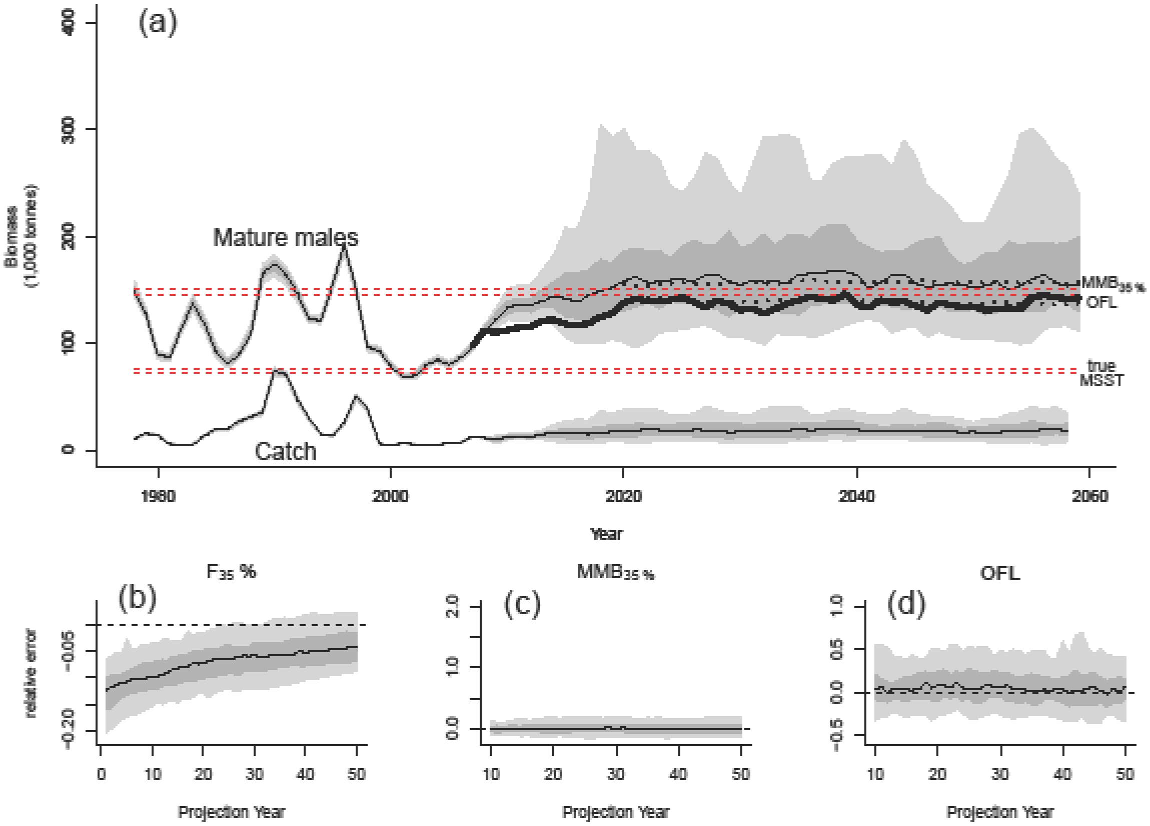

Figure 4.4. Results for the scenario in which the operating model is 'status quo' and the estimation method is also 'status quo'. Graph (a): MMB and catch in thousands of metric tonnes. Thick solid line is the median (over simulations) from applying only the OFL HCR. Dashed lines are ranges for true $MMB_{35\%}$ (top) and minimum stock size threshold, MSST (bottom); each simulation had a slightly different 'true' $MMB_{35\%}$, hence the 'range'. Dotted lines are the range for the estimated $MMB_{35\%}$. Graphs (b), (c) and (d) are the distributions of relative error for $F_{35\%}$, $MMB_{35\%}$ and the OFL, respectively. The scale for $MMB_{35\%}$ and the OFL will remain the same across all scenarios for ease of comparability. For all graphs, light grey outlines 5th and 95th quantiles, dark grey outlines 25th and 75th quantiles. Thin solid lines are the medians.

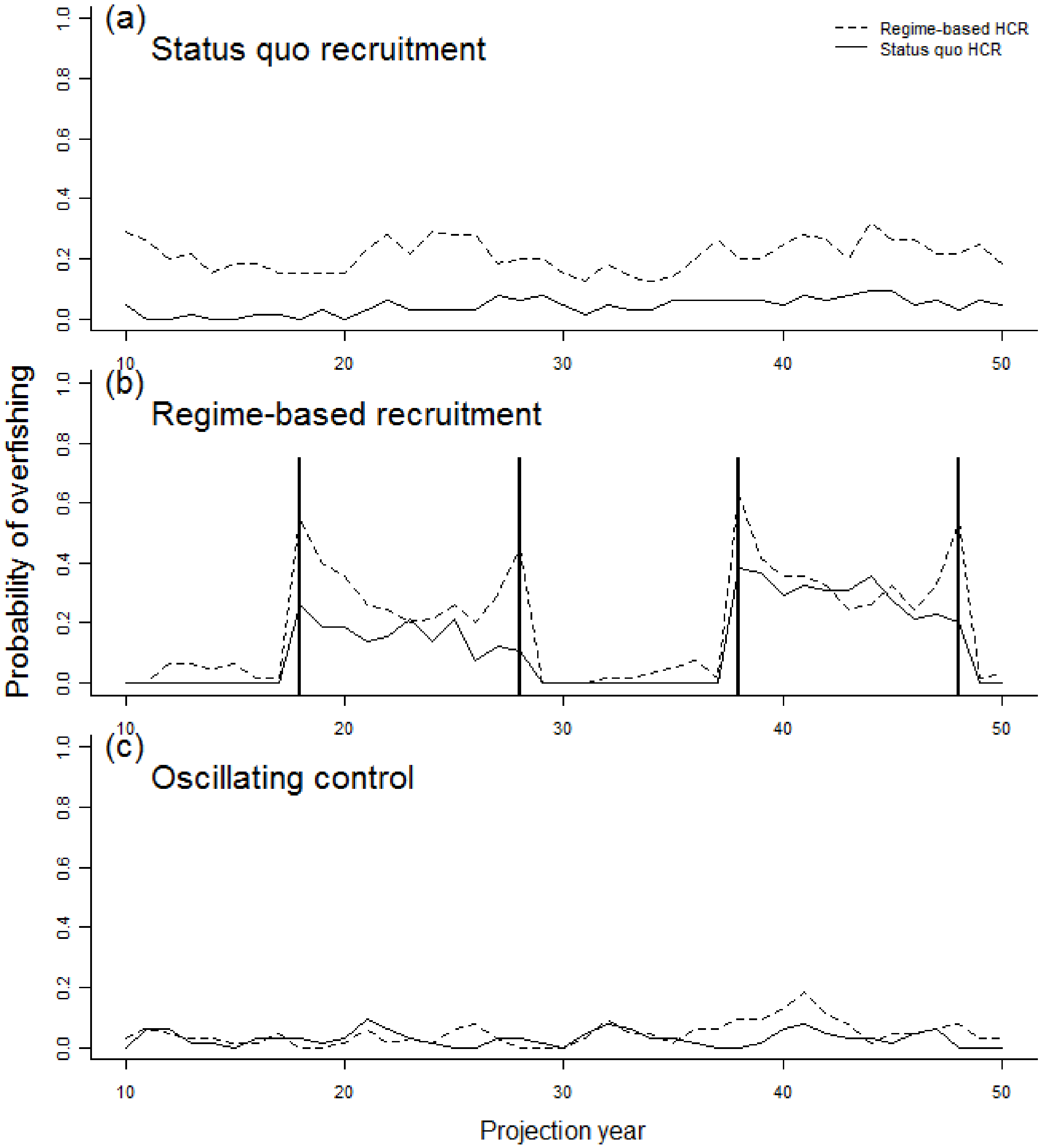

Figure 4.5. Probability of overfishing as the percentage of simulations in a given year in which the removed catch exceeded the overfishing limit (OFL) for a regime-based HCR and the status quo HCR under three future recruitment models. Vertical lines in panel (b) indicate the years in which the regime switched.

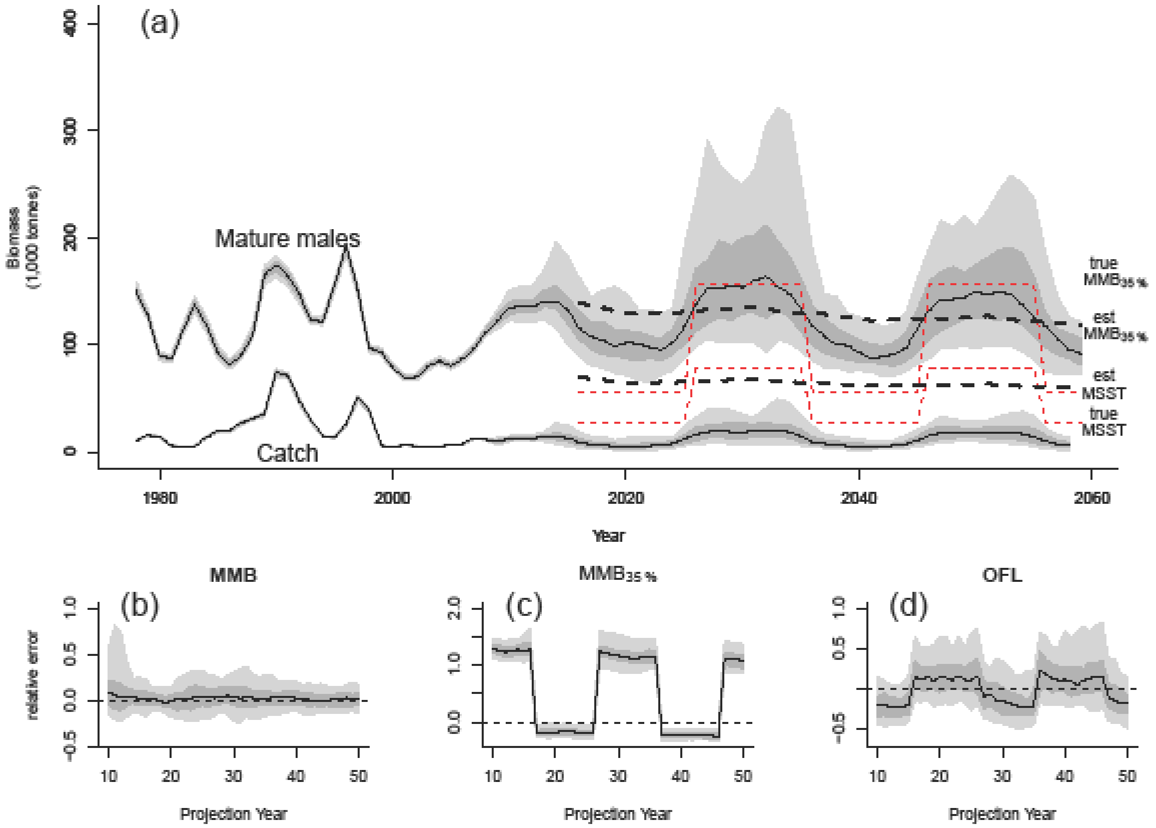

Figure 4.6. Results for a scenario in which the operating model is ‘regime-based’ and the estimation method is ‘status quo’. Graph (a): MMB and catch in thousands of metric tonnes. Thick dashed line is the median estimated MMB$_{35\%}$. Thin dashed lines are ranges for the true MMB$_{35\%}$ (top) and MSST (bottom). Graphs (b), (c) and (d) are distributions of relative error in MMB, MMB$_{35\%}$ and the OFL, respectively. For all graphs, light grey outlines 5th and 95th quantiles; dark grey outlines 25th and 75th quantiles. Thin solid lines are the medians.

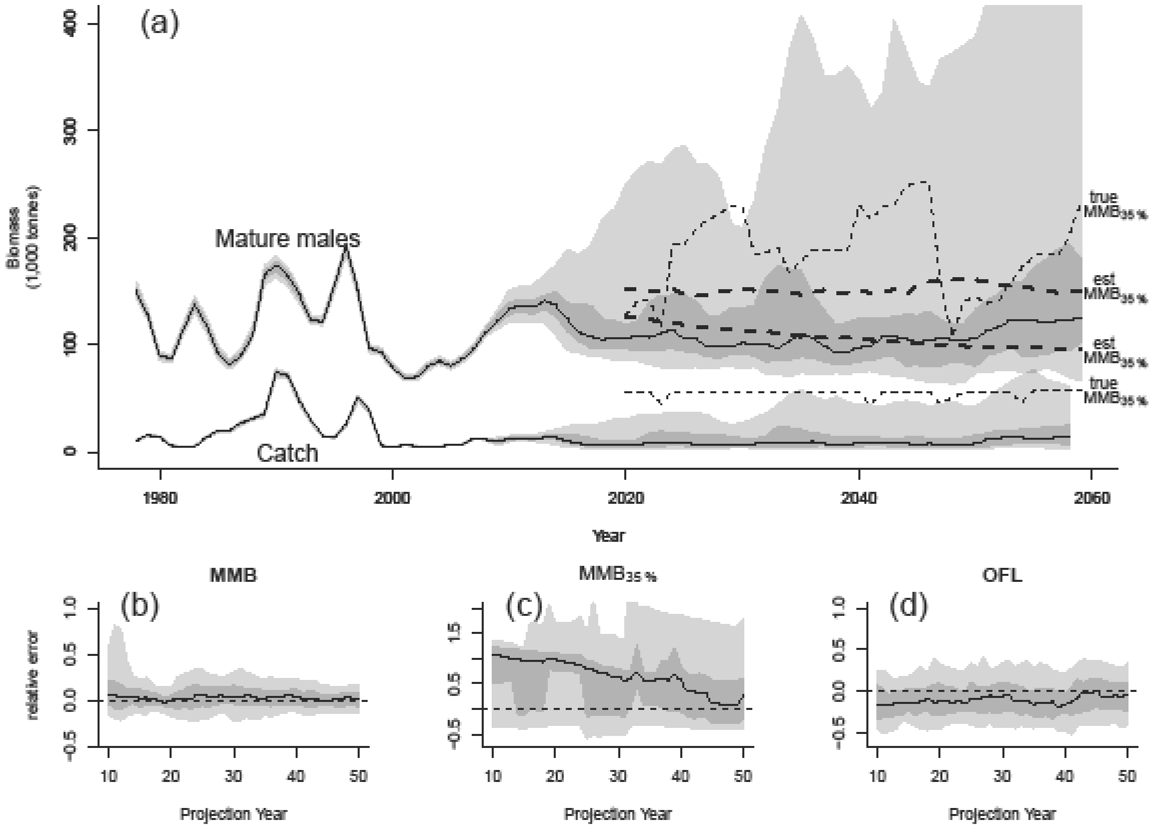

Figure 4.7. Results for the scenario in which the operating model is ‘oscillating control’ and the management strategy is ‘status quo’. Graph (a): MMB and catch in thousands of metric tonnes. Thick dashed line is the range of estimated MMB$_{35\%}$. Thin dashed lines are ranges for the true MMB$_{35\%}$ (top). Graphs (b), (c) and (d) are distributions of relative error in MMB, MMB$_{35\%}$ and the OFL, respectively. For all graphs, light grey outlines 5th and 95th quantiles; dark grey outlines 25th and 75th quantiles. Thin solid lines are the medians.

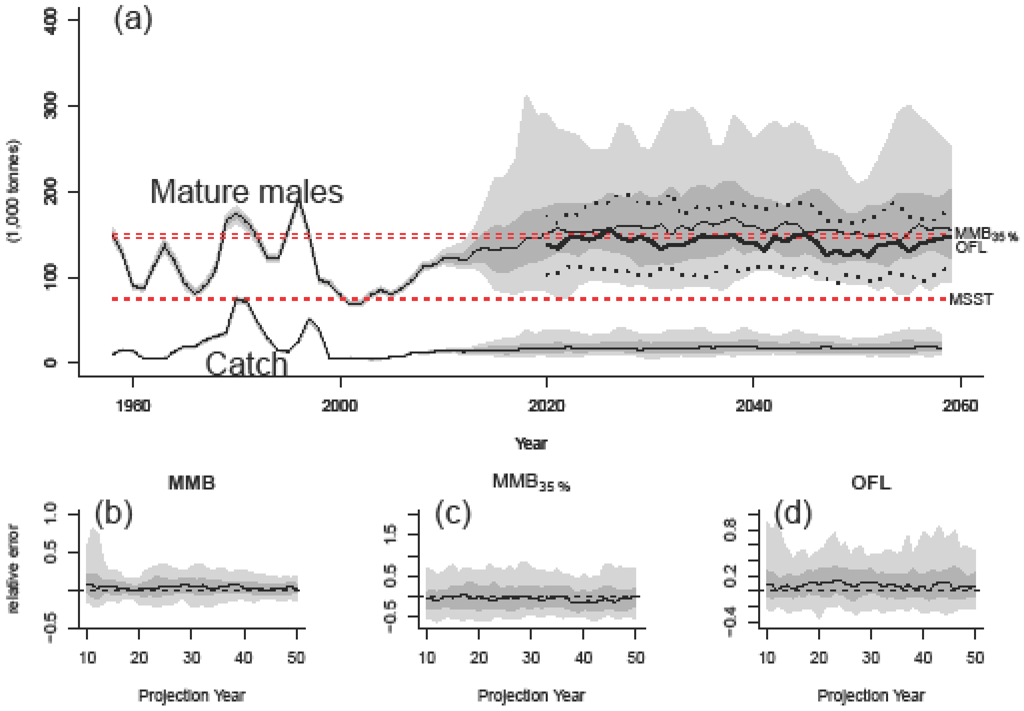

Figure 4.8. Results for a scenario in which the operating model is ‘status quo’ and the management strategy is ‘regime-based’. Graph (a): MMB and catch in thousands of metric tonnes. Thick solid line is the median from applying only the OFL HCR. Dashed lines are ranges for MMB$_{35\%}$ (top) and MSST (bottom). Dotted lines are the range for estimated B$_{MSY}$. Graphs (b), (c) and (d) are distributions of relative error in MMB, MMB$_{35\%}$ and the OFL, respectively. For all graphs, light grey outlines 5th and 95th quantiles; dark grey outlines 25th and 75th quantiles. Thin solid lines in all plots are the medians.

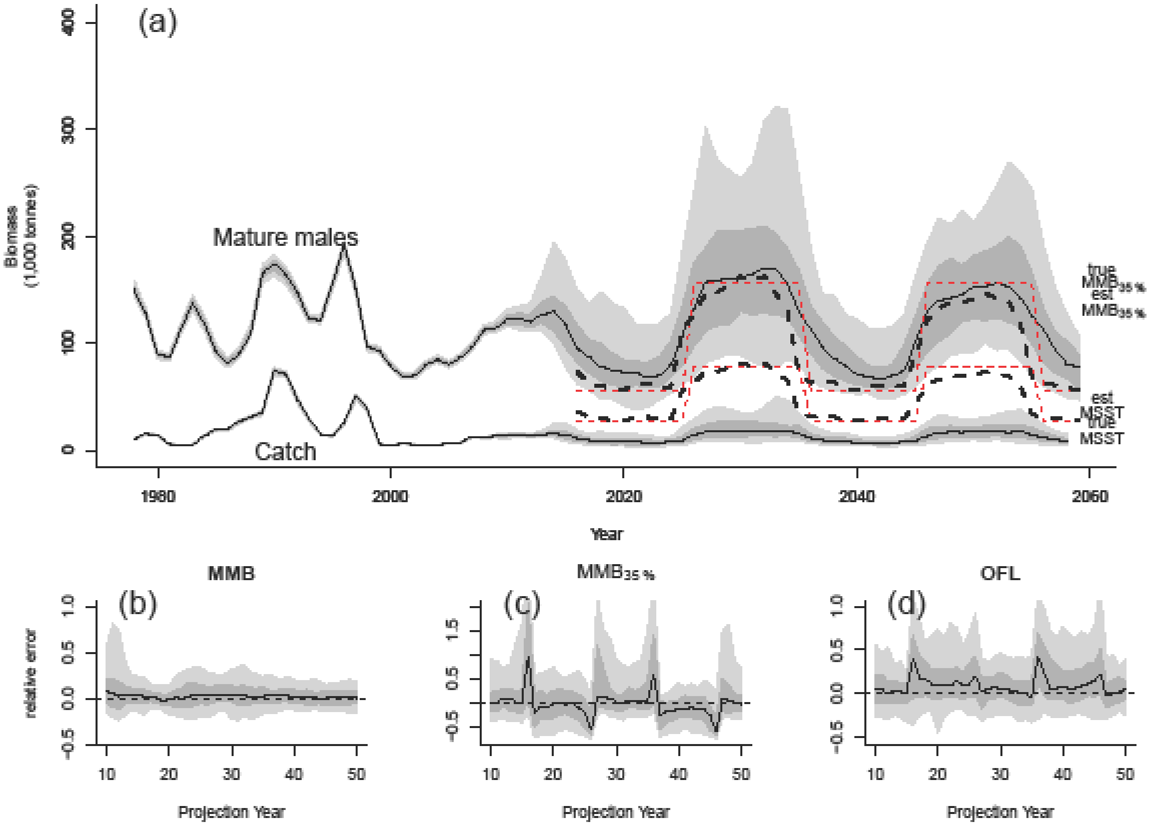

Figure 4.9. Results for a scenario in which the operating model is 'regime-based' and the management strategy is 'regime-based'. Graph (a): MMB and catch in thousands of metric tonnes. Thick dashed line is the median estimated $MMB_{35\%}$and MSST. Thin dashed lines are the median for the true $MMB_{35\%}$ (top) and MSST (bottom). Graphs (b), (c) and (d) are distributions of relative error in MMB, $MMB_{35\%}$ and the OFL, respectively. For all graphs, light grey outlines 5th and 95th quantiles; dark grey outlines 25th and 75th quantiles. Thin solid lines are the medians.

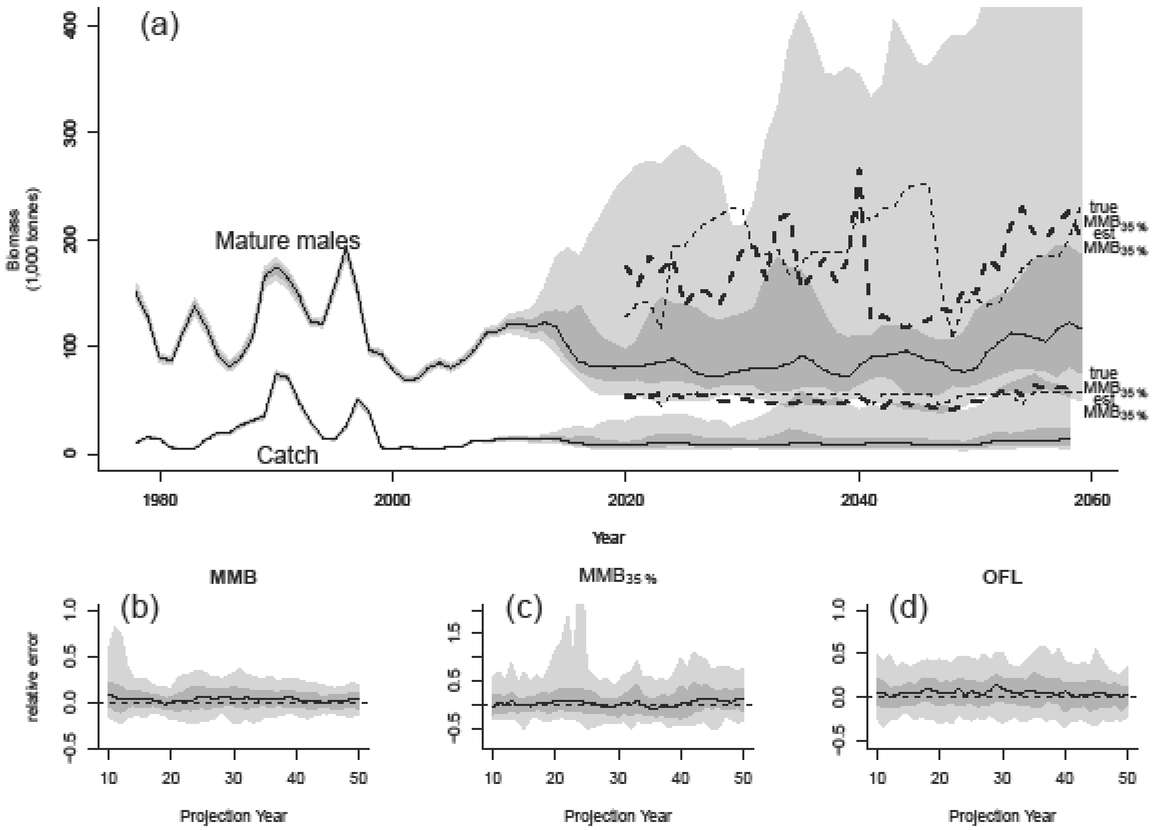

Figure 4.10. Results for a scenario in which the operating model is 'oscillating control' and the management strategy is 'regime-based'. Graph (a): MMB and catch in thousands of metric tonnes. Thick dashed line is the range of estimated MMB$_{35\%}$. Thin dashed lines are range for the true MMB$_{35\%}$ (top). Graphs (b), (c) and (d) are distributions of relative error in MMB, B$_{MSY}$ and the OFL, respectively. For all graphs, light grey outlines 5th and 95th quantiles; dark grey outlines 25th and 75th quantiles. Thin solid lines are the medians.

Appendix 4.A: Additional harvest strategies

The proposal referred to additional harvest strategies, the evaluation of which was not included in the above manuscript for the sake of brevity. The performance of a fixed proportion strategy, the Hilborn method (Hilborn et al., 2002; see below for details), and the Tier 2 strategy (NPFMC, 2007) were also compared to the status quo strategy.

Fixed proportion:

A fixed proportion strategy is attractive because an assessment is not required. Here, a range of fixed proportions (10% - 40%) of the survey mature male biomass were set as the TAC over a 50-year projection period under status-quo recruitment (Fig. App.4.1). An exploitation rate of 0.25 led to the highest mean-median yield over the last 40 years of the projection period (39,500 tonnes) which was less than the status-quo yield (40,600 tonnes) (Table App.4.1). The standard deviation of the mean-median yield over this period for the fixed exploitation rate strategy (2,100 tonnes) was comparable to the status quo strategy (2,200 tonnes). A fixed exploitation rate of 0.2 maintained the population closest to what is currently considered B_{MSY} (Fig App.4.1).

Hilborn ("hold steady") method:

Hilborn et al (2002) proposed a 'hold-steady' method that is essentially a fixed escapement policy:

$$C_{t+1} = C_t * cLag + (1 - cLag) * D$$

$$D = \begin{cases} 0 & if\ I_t \leq I_{target} \\ I_t - I_{target} & if\ I_t > I_{target} \end{cases}$$

Each year an index (I_t) for the stock is taken (e.g. a survey) and compared to a target index (I_{target}). The TAC is the amount of biomass over the target, buffered by a factor to stabilize the catch from year to year (cLag). This method again has the advantage of not requiring a stock assessment, and Hilborn et al suggest that 'hold-steady' policies may provide a better alternative for some stocks.

This method was used to set the TAC for a range of target indices and buffering factors for the status-quo and oscillating control operating models. The estimated mature male biomass from the NMFS survey was used as the index, and simulations were projected 50 years into the future.

For the status-quo operating model, each combination of target index and catch lag was able to maintain the median MMB at the proposed level, but for the higher targets, catches in the first several years of the projection period were very low to allow the stock to increase (Fig. App.4.2). The highest mean-median catch over the final 40 years of the projection period with the hold-steady rule (38,700 tonnes) approached that for the status quo rule (40,600 tonnes), but again had much higher variability (sd = 4,000 tonnes; Table App.4.2). Once the desired target biomass was reached, there was little difference in median yield between the different targets. Variability in MMB increased with the buffer factor, but variability in catch was generally negatively related to the buffer factor (Table App.4.2).

Patterns in catch and MMB under the oscillating control operating model were similar to those for the status quo operating model (i.e., lower target biomasses led to higher catches), with the exception that there was much more variation between scenarios for the oscillating control model (Fig. App.4.3). This variation is expected given the difference between projections of the wPDO among climate models.

Tier 2 vs Tier 3:

The question of when management advice should be based on estimates of B_{MSY} and F_{MSY} (Tier 2) instead of proxies values for these quantities for snow crab was explored by estimating the stock-recruitment relationship, and using model selection techniques to determine when a two-parameter (i.e., Beverton-Holt, B-H) model outperformed a one-parameter (constant recruitment) model (the Beverton-Holt model was the true model for the simulations). For 65 simulations based on the status-quo operating model, projected 50 years with catch determined by the status-quo harvest strategy, one- and two-parameter models were fit to the actual data used to assess the stock (catches, surveys, trawl bycatch and observer data from 1979-2009) and the AICc was calculated. A year of data was added and the assessment models were refit, along with the stock recruitment relationship and AICc recalculated. This process was repeated until the end of the time series. The two-parameter model was never selected and the mean change in AICc was 2, in favor of the model which sets recruitment at B_{MSY} to average recruitment.

There are two reasons why AICc never selects the stock-recruitment relationship even though it is the true model. First, the historical data show a negative relationship between MMB and recruitment (Fig. App.4.4), so the B-H relationship will not be selected based on the actual data. Second, the HCR maintains the biomass around B_{MSY} during the projection period. There are consequently many observations of recruitment and MMB around B_{MSY} and only a few for higher biomasses. These two reasons, compounded by the high value for the variability in recruitment, lead to a negative relationship between MMB and recruitment and hence little support for the (true) Beverton-Holt curve.

If, instead of including the historical data (which does not display an obvious stock-recruitment relationship that can be well fit by a Beverton-Holt curve), only the projected data (which have an underlying Beverton-Holt stock-recruitment relationship) are used to compare stock-recruitment models, the single parameter model is also always selected. The projection starts at a relatively high biomass and the HCR doesn't allow biomass to decrease to a point where recruitment would decline to allow the extent of curvature in the stock-recruitment relationship to be detected. So, again, there is insufficient contrast in the data to select a two-parameter model once a management procedure designed to maintain a stock at a specific biomass is in place. If the stock-recruitment relationship suggested by these data were used in a Tier 2 rule, MSY would be reached by applying a very high fishing mortality because steepness is 1 and crabs do not grow after maturity. This is an undesirable outcome.

		Exploitation rate						
	Status quo	**0.1**	**0.15**	**0.2**	**0.25**	**0.3**	**0.35**	**0.4**
MMB mn md	349	584	449	348	270	208	161	129
(sd)	11	30	17	11	11	11	12	12
Catch mn md	41	31	37	40	40	38	35	32
(sd)	2	2	2	2	2	2	3	3

Table App.4.1. Results from the fixed exploitation HCR. The mean-median MMB and catch over the final 40 years of the projection and the mean standard deviation over the same period.

Target		150			200			250			300			350		
cLag	Status quo	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9
MMB mn md	349	255	256	265	330	334	336	405	405	404	474	471	468	542	539	575
(sd)	11	12	15	17	17	20	25	21	27	35	25	31	45	26	32	40
Catch mn md	41	37	37	38	36	37	38	34	35	36	31	33	34	28	30	29
(sd)	2	4	4	4	5	5	4	5	5	5	6	6	5	5	6	7

Table App.4.2. Results from ‘Hilborn method’ HCR under the status-quo operating model. The mean-median (‘mn md’) MMB and catch over the final 40 years of the projection and the mean standard deviations over the same period.

Target		150			200			250			300			350		
cLag	Status quo	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9
MMB mn md	280	263	265	266	332	333	331	394	396	391	450	447	441	495	492	503
(sd)	16	19	23	33	20	27	33	23	27	35	24	32	42	27	35	48
Catch mn md	29	30	29	28	24	23	23	18	19	19	12	13	14	7	8	8
(sd)	5	7	6	5	6	5	4	5	5	3	5	4	2	3	3	4

Table App.4.3. Results from ‘Hilborn method’ HCR under the oscillating control operating model. The mean-median (‘mn md’) MMB and catch over the final 40 years of the projection and the mean standard deviations over the same period.

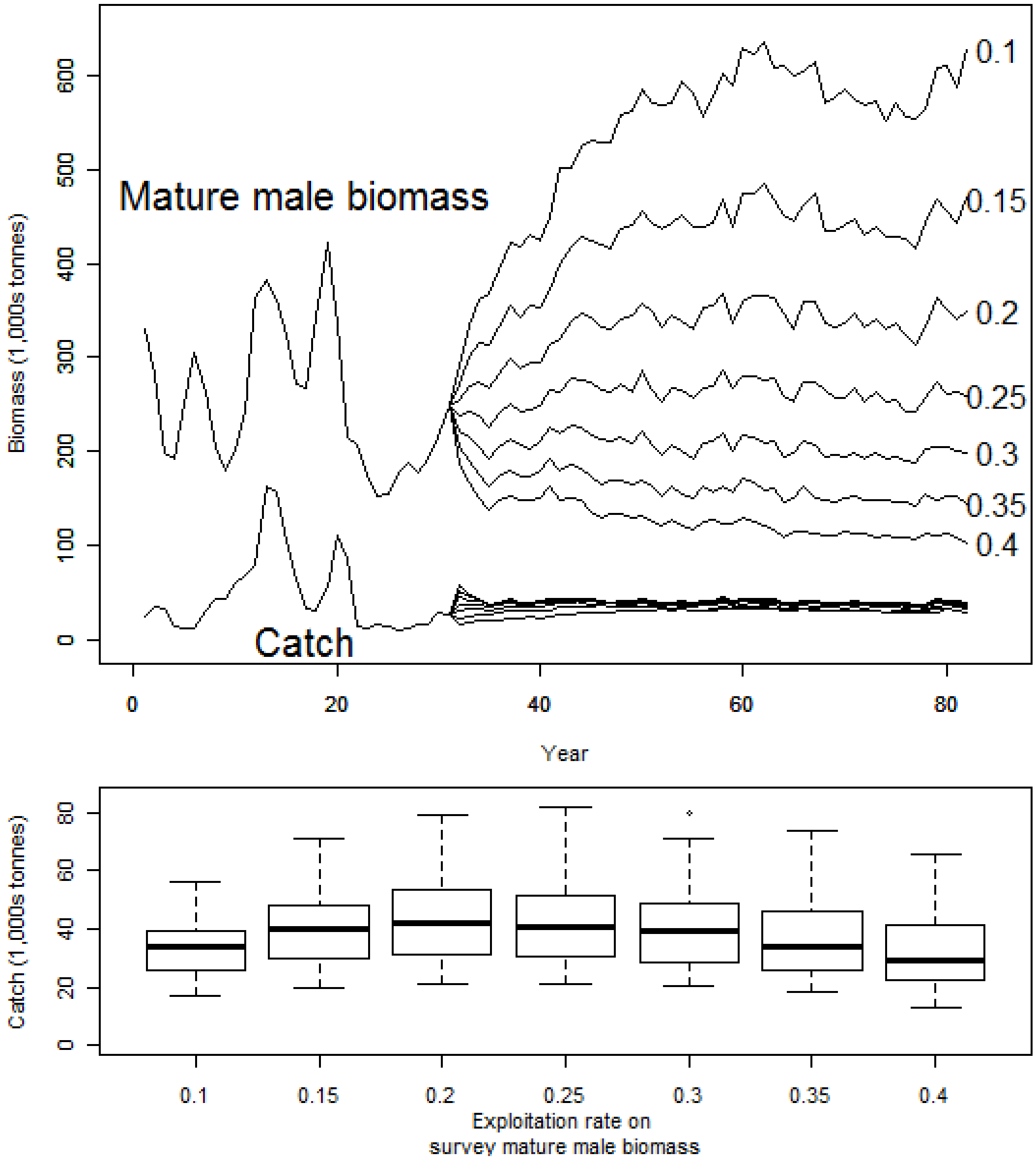

Figure App.4.1. Results from the fixed exploitation HCR for the status-quo operating model. Top: MMB and catch projected under a range of fixed exploitation rates. Bottom: Boxplot of catch in the last year of the projection for a range of exploitation rates.

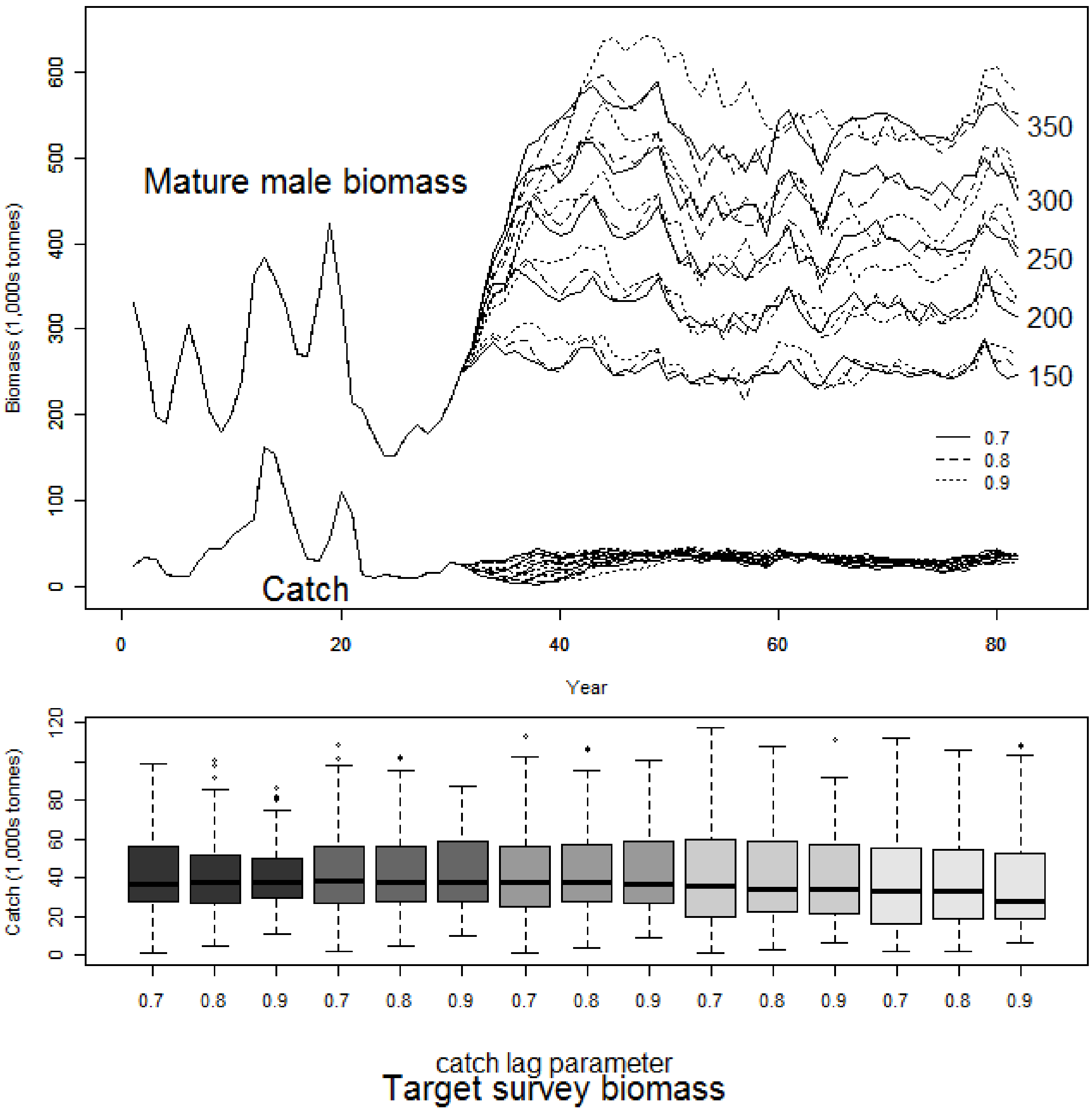

Figure App.4.2. Results from the ‘hold steady’ HCR for the status-quo operating model. Top: MMB and catch projected under a range of values for the target biomass and the catch buffering factor. Bottom: Boxplot of catch in the last year of the projection for the range of values for the target biomasses and the catch buffering factor.

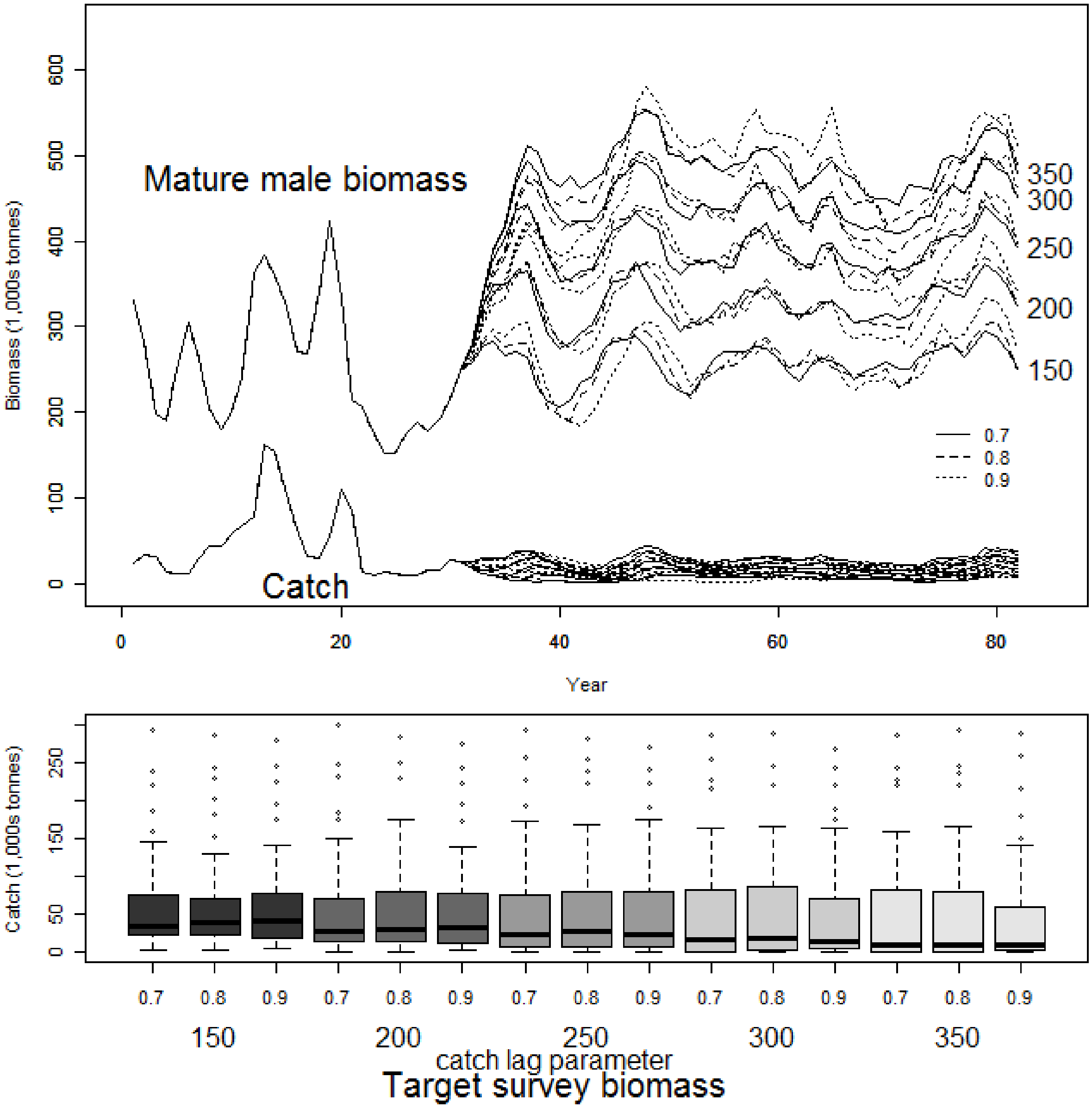

Figure App.4.3. Results for the hold-steady HCR for the oscillating control operating model. Top: MMB and catch projected under a range of values for the target biomass and the catch buffering factor. Bottom: Boxplot of catch in the last year of the projection for the range of values for the target biomasses and the catch buffering factor.

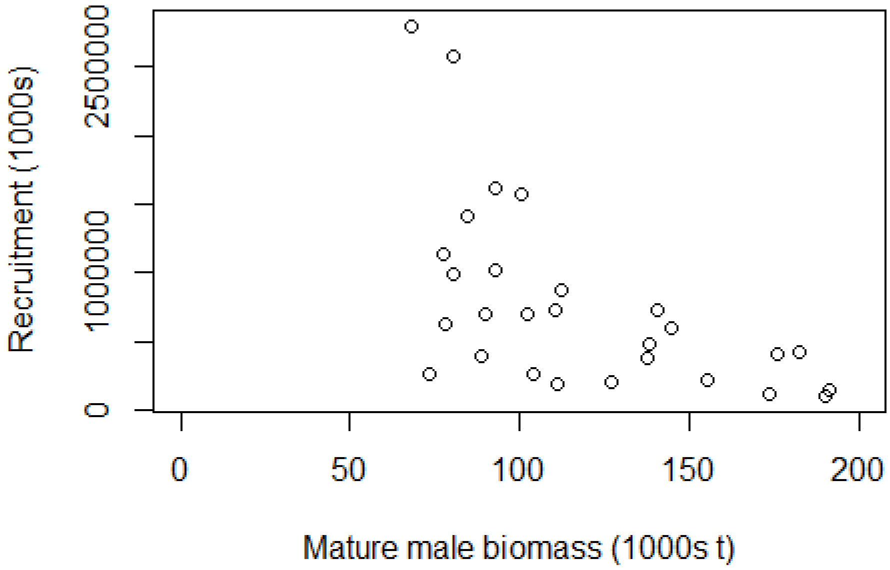

Figure App.4.4. Relationship between MMB and recruitment.

Chapter 5: Can an aggregate stock assessment capture the dynamics of a spatially-structured stock? Snow crab in the eastern Bering Sea as a case study

5.1 Introduction

Spatial structure in exploited marine populations comes in many varieties: local densities can vary through space and/or time in a single panmictic population (e.g. Bentley et al., 2004), sink/source dynamics can be present in which metapopulations feed satellite populations (e.g. Porch et al., 2001), or genetically distinct stocks can be present in the same location (e.g. Taylor et al., 2009). Unaccounted for spatial structure in stock assessments may influence the ability of management systems to reach their goals (Cadrin and Secor, 2009; Ying et al., 2011; Guan et al., 2013). Attempts to incorporate spatial structure into stock assessments have been made to correct the potential biases in quantities important in management introduced by spatial structure (e.g. Booth, 2000; Goethel et al., 2011; Punt et al., 2000). However, accurately describing spatial structure in marine stocks is often difficult given commonly available data. Movement, selectivity, mortality, growth and recruitment are confounded in stock assessment models, and determining whether a population in a given area decreased because of higher natural mortality or emigration (for example) is often impossible without studies targeted at understanding each of these processes (e.g. Hilborn, 1990, Anganuzzi et al., 1993; Mullen, 1994).

Spatial structure is ignored when conducting stock assessments of the population of snow crab in the eastern Bering Sea (EBS). Recruitment appears to occur primarily on the northeastern portion of the shelf (based on length frequencies from surveys, Ernst, 2005) and there is evidence from Regional Ocean Monitoring System-linked individual-based models that larvae released in the middle domain, the area between 50-100 m depth, disproportionately contribute to overall recruitment (Parada, 2010). An ontogenetic migration is inferred from survey data in which snow crab move from the northeastern portion of the EBS shelf to the southwest (Ernst, 2005), resulting in a gradient of both maturity and size across and down the shelf. Additionally, the fishery occurs primarily in the southwestern portion of the surveyed-area (Figure 1), and the distribution of spawning females and males at the time of the survey do not completely overlap (Turnock and Rugulo, 2013). However, the population appears to be genetically panmictic (Hardy et al., 2011), which offers some counter-balance to the cited spatial structure.

The current assessment method for snow crab in the EBS (Turnock and Rugulo, 2013) implicitly assumes: 1) no spatial structure, 2) recruitment is equally distributed over the range of the population, 3) the mature animals contribute equally to recruitment irrespective of their location in the EBS, and 4) fishing mortality is applied equally to the entire population subject to a selectivity ogive (which also incorporates availability). It is unclear whether these simplifications have impacted management performance, but the fishery was declared overfished in 1999. A rebuilding plan was subsequently enacted with the goal of returning the stock to its target biomass in ten years. However this plan was declared a failure in 2009 (in 2011, the stock was declared rebuilt following a recalculation of target biomasses).

Attempts have been made to assess this stock using a spatially-structured model (Murphy, 2012) and such models have been applied to other stocks in the Bering Sea (e.g. walleye pollock, *Theragra chalcogramma*, by Hulson et al. 2013) in the past. This paper explores the feasibility of developing a spatially-explicit assessment method for EBS snow crab which estimates movement, local densities, recruitment and trends in fishing mortality for male crab. A stock assessment method with four areas is presented, the limitations of this method are discussed and recommendations are made for future research. The results from the spatial assessment are then used as an operating model to evaluate the ability of a non-spatially structured (aka ‘aggregated’) assessment method to estimate quantities important for management when the underlying population is spatially structured.

5.2 Methods

5.2.1. A spatially-structured assessment method

The assessment model (coded in Auto-differentiation Model Builder (C++), Fournier et al. 2012) tracked numbers of male crab by size, maturity state and area (see the appendix for model structure, likelihood components and estimated parameters). Only males are modeled because only males are retained in the fishery and management reference points are based on mature male biomass (NPFMC, 2008). Growth, natural mortality and fully-selected discard mortality were fixed to the values used for males in a recent National Marine Fishery Service's (NMFS) assessment (Turnock and Rugulo, 2011) because they are difficult to estimate with available data. Length frequencies and numbers from the NMFS summer survey of the Bering Sea, catch records of retained crab, and length frequencies of the total catch (i.e. all crab caught irrespective of whether they were discarded) from observers informed parameter estimates (Table 1). Data from side-by-side trawls to determine the catchability and selectivity of the NMFS survey gear performed by the Bering Sea Fisheries Research Foundation (BSFRF) were also used for parameter estimation. Snow crab are also caught as bycatch in the groundfish fishery in the Bering Sea, but bycatch accounts for only 6% (on average) of total snow crab catch and obtaining spatially-explicit bycatch data is difficult. Consequently, groundfish fishery bycatch data were not included in the analyses of this paper. Excluding these data will not impact the qualitative conclusions of the analyses.

The portion of the Bering Sea shelf over which snow crab are surveyed and assessed was partitioned into four areas demarcated by 58.5 degrees N latitude and the 100 m bathymetric line. The north/south division of 58.5 degrees N was chosen because a large portion of the historical catch occurs south of this line (Figure 1). The east/west division of the 100m depth contour represents the divide between the 'middle' and 'outer' domains of the Bering Sea shelf, which are thought to play distinct roles in recruitment dynamics (e.g. Parada et al., 2010).

The total number of male crab at length from the NMFS survey were split into maturity state using observed proportions of mature crab at length based on chela height taken from subsamples of the survey hauls. The density of crab at a given sampling station was calculated as the number of individuals caught divided by the area swept and averaged over repeated trawls (if present; equation 1):

$$D_{y,s} = \frac{1}{n_{y,s}} \sum_{i=1}^{n_{y,s}} \frac{d_{y,s,i}}{a_{y,s,i}}$$

where $D_{y,s}$ is the density in year y at station s, $n_{y,s}$ is the number of hauls during year y at station s, $d_{y,s,i}$ is the number of crab collected during year y at station s during haul i, and $a_{y,s,i}$ is the area swept during year y at station s during haul i. The number of crab in each of the four areas, x, was calculated as the mean density at the stations in the area x multiplied by the nautical square miles contained in area x, A_x:

$$N_y^x = \frac{\sum_{j=1}^{j=m_x} d_{y,j}^x}{m_x} A_x$$

where N_y^x is the number of crab during year y in area x, m_x is the number of stations in area x, and $d_{y,s}^{x}$ are densities during year y that are located in area x.

Catch data were available at a spatial resolution of 1 degree longitude x 0.5 degrees latitude, and the approximate depth of catch at this resolution was inferred by using depth and location observations from the survey. There were slight differences between the total estimated number of crab from the survey used in the current analyses and those used in the actual stock assessments. The differences are likely due to the division of crab caught in the survey to ‘strata’ in the NMFS estimates, but including these strata was beyond the scope of this paper and the general trends in abundance are preserved.

5.2.2. Priors for movement

Convergence of the optimization algorithm is likely to be an issue for the spatial assessment given the number of confounded processes, but placing priors on movement parameters may help stabilize the optimization (Kurota et al. 2009). Priors for the parameter that scales the movement curve (i.e. determines the proportion of crab-at-length that are fully-selected for movement that moves from one area to another) were developed by assuming that the centroids of the locations of large males (>101mm carapace width) caught in the fishery represent the true centroids of large males in a given year at the time of the fishery. This assumption may be reasonable given that fishers have large incentive to target areas with a high density of large males. Some directions of movement are improbable based on observed changes in the centroid of abundance for large males from the survey to the fishery (figure 2) and the proportions moving from these directions were fixed to zero. The proportions of crab moving from (for example) the southeast area to the southwest area after the survey was calculated by first finding the centroids of abundance at the time of the survey and the fishery (equations not shown):

$$\theta_{y,Surv,lat} = \frac{\sum_{j=1}^{N_y} Lat_j D_{y,j}}{\sum_{j=1}^{N_y} D_{y,j}};$$

$$\theta_{y,Surv,lon} = \frac{\sum_{j=1}^{N_y} Lon_j D_{y,j}}{\sum_{j=1}^{N_y} D_{y,j}}$$

where, $\theta_{y,Surv,lat}$ is the latitudinal component of the centroid during year y at the time of the survey, N_y is the total number of stations for which densities are available during year y, Lat_j is the latitude for station j, and $D_{y,j}$ is the density at station j during year y. Centroids were calculated for the fishery in a similar manner. The differences among the latitudinal and longitudinal components of the centroids between the fishery and the survey, $\Delta_{y,lat}$ and $\Delta_{y,long}$ during year y were calculated as:

$$\Delta_{y,lat} = \theta_{y,Surv,lat} - \theta_{y,Fish,lat}$$

$$\Delta_{y,lon} = \theta_{y,Surv,lon} - \theta_{y,Fish,lon}$$

The differences in centroids were added to the raw survey observations of latitude and longitude, and the proportion of crab moving from (for example) the southeast area to the southwest area (after the addition of the differences) during year y, $Pr_y^{SE\to SW}$, was calculated as:

$$Pr_y^{SE\to SW} = \frac{N_y^{SE\to SW}}{N_y^{SE}}$$

where $N_y^{SE\to SW}$ is the number of crab that moved during year y to the southwest area from the southeast area after the addition of $\Delta_{y,lat}$ and $\Delta_{y,lon}$ to the location data and N_y^{SE} is the total number of crab in the southeast area during year y. This process was repeated for all areas and the period of movement from the fishery to the survey. The resulting proportions (e.g. $Pr_y^{SE\to SW}$) were used as priors for movement in some fittings (see below) of the assessment model. This method doesn't account for the curvature of the earth when calculating distance traveled, but the estimates of the proportions moving among areas offer plausible priors given the scale of the data, the potential magnitude of the errors introduced by ignoring the curvature and the weighting given the priors.

The spatial assessment model was fit in four steps to attempt to allow information in the length frequencies to guide the estimates of proportions moving between years. First, the model was fit without incorporating the movement prior provided by the changes in centroids, and all movement parameters were estimated (see appendix for estimated parameters). Second, the curves describing the proportion of crab at length moving (see appendix for description of the parameterization of movement) were estimated, while the scaling factors (for which $Pr_y^{SE\rightarrow SW}$, for example, is a prior) for those curves were held constant at the value with the highest likelihood from the priors from the centroids. Next, the movement curves were held constant at the estimates from the second fitting, while the scaling factors were estimated without the priors and finally, the parameters of the model were estimated fixing the movement curves and the scaling factors at the estimated values in the second and third fitting to calculate variances for the other parameters in the model. In each fitting, all parameters other than those associated with movement that are noted in the appendix as estimated were estimated.

5.2.3. Simulation evaluation

The fitted spatial population dynamics model was used as an operating model to generate data to evaluate the performance of a non-spatially structured (aggregate) assessment when the underlying population is spatially structured. The aggregated assessment model was identical in structure to the spatial model, except that there was a single area, rather than four areas. Spatial data generated from the spatially-explicit population dynamics model were aggregated (i.e. retained catch biomass, survey numbers, and length frequencies for total catch, retained catch and survey numbers were aggregated over the four areas) and the aggregated model was fit to 100 simulated data sets in which observation error was added to the survey abundance and catch data (CVs of 0.22 and 0.05, respectively) and length frequencies were generated by sampling 100 random lengths from the expected length frequency distributions calculated using the spatial model. The data used to inform selectivity were not changed because selectivity is assumed to be the same among areas. The simulations were run twice: once when data were available to inform survey selectivity and once when those data were absent.

The performance of the aggregate assessment method was evaluated based on its ability to reproduce trends in survey abundance, total recruitment, intensity of exploitation and selectivity patterns. It is difficult to compare fishing mortality or reference points directly between spatial and aggregate models, so the ratio of the predicted catch to the predicted exploitable biomass (calculated as the sum of the product of biomass at length at the time of the fishery and fishery selectivity) was used to indicate an overall exploitation rate.

5.3 Results

5.3.1 Spatial assessment

The Hessian matrix was not positive-definite and the maximum gradient component was greater than desired for the spatial assessment method when movement was estimated, regardless of the inclusion or exclusion of the centroid-derived priors for movement. This is not unexpected, considering the confounding between movement among multiple areas, selectivity/catchability, and recruitment. However, it was impossible to simultaneously fit the length frequencies from the survey, the BSFRF selectivity data and the observed catch unless movement was allowed between the fishery and the survey (this is a result of nearly all recruitment occurring in the NE area and most of the catch occurring in the southwest and southeast). Additionally, the assumption that the centroid of abundance of large males as a proxy for movement appeared to be good for some years, but not for others (compare rows in Figure 3). Estimated relative movement at length and scaling factors were slightly different when including or excluding the centroid-derived movement priors (Figure 4). Although movement could not be estimated reliably, parameter values could be found that led to adequate fits to the data, and excluding the centroid-derived data resulted in the best fits. Consequently, the parameter estimates from the model fit that excluded the centroid-derived priors were used to generate data sets to test the aggregated assessment, and are discussed in detail below.

Movement occurred twice in each model year, once after the fishery and once after the survey. Movement tended to be towards the south and west between the time of the survey and the midpoint of the fishery, and mostly large crabs moved (figure 4). Movement tended to be towards the north and east between the fishery and the survey, but larger crabs were less likely to move (figure 4). Fits to the survey numbers were good both by area (figure 5) and survey length frequencies were generally well fit (Figures A1-4). The northeast portion of the shelf was dominated by small, immature crab. Relatively few crab were observed in the southwest at the time of the survey, but those present were more likely to be mature and large. Crab in the northwest and southeast were primarily large mature individuals, with occasional peaks in small crab. Catch by area and catch length frequencies are slightly underestimated, likely because the weighting of the BSFRF selectivity data informs catchability very strongly (figure A5-9). Fishery and discard selectivity (i.e. probability of discard) were very similar to those estimated by the current assessment model (figure 6a).

Recruitment occurred primarily in the northeast portion of the shelf (figure 6b), with a few good years in the southeast and a large reduction in recruitment occurred around 1990. Fishing mortality was highest in the southwestern portion of the shelf, with the southeast having occasional years of high fishing mortality (figure 6c). Fishing mortality in the area with the highest estimated levels of fishing mortality was less than those thought to represent F_{MSY} for the stock according to the most recent assessment (Turnock and Rugulo, 2013).

5.3.2. Simulation evaluation of an aggregated assessment method

The aggregated model accurately fit the generated survey abundances both when BSFRF data to inform survey selectivity were available and when such data were absent (figures 7a and 7d). However, the assessment was unable to estimate survey catchability accurately when the data from the BSFRF-NMFS side-by-side trawls were not used to inform survey selectivity, which has been noted as an issue for the current assessment (Szuwalski and Punt, 2012) and can be difficult for stock assessment in general (Arreguin-Sanchez, 1996) . The shape of the estimated selectivity ogive was still roughly correct when BSFRF data were not available (figure 7e), but the estimates of catchability were positively biased which resulted in underestimates of exploitable biomass, and translated to a positive bias in the perceived exploitation rate (i.e. the aggregated assessment reports a higher exploitation rate than is occurring when there is no data informing survey selectivity; figure 7f). Estimated recruitment is positively biased when there are no data on selectivity (not shown) as would be expected given the bias in survey catchability.

5.4 Discussion

Snow crab must move between when the survey occurs and the time of the fishery (relatively high rates of movement and low catchability (0.38) are estimated here), but it is difficult to reliably determine how much with the currently available data. A crab would have to move approximately a half of a kilometer a day to travel the average distance covered by the change in centroids used in this analysis, and this seems reasonable given rates of movement in breeding migrations observed in Atlantic populations (e.g. Lefebvre and Brethes, 1991). The shape of the estimated movement curve changes little when the centroid priors are included or excluded, which suggests the length frequencies do hold some information about movement. The directions of movement after the survey and fishery are consistent with what is known about the temperature preferences of snow crab at different stages in their life history: smaller, younger crab prefer colder water (which occurs in shallower water) and larger, older crab prefer relatively warm water (Dionne, 2003).

Estimated fishing mortality is higher in the southwest portion of the shelf, but, given the estimated movement rates, less than the average estimated F_{MSY} for ten scenarios presented in a recent assessment ($1.17\mathrm{yr}^{-1}$; Turnock and Rugulo, 2011, their Table 10). Estimated patterns in recruitment support changes in recruitment noted by Szuwalski and Punt (2013), who suggest decreased recruitment after the year 1990 is related to a shift in climate in the Bering Sea and also likely played a role in the overfished status of the fishery.

Previous attempts at spatial assessment of this population included only two areas (Murphy, 2012). However, four areas are the minimum required to address the validity of leading hypotheses describing snow crab population dynamics (e.g. Ernst et al., 2012; Parada et al., 2010). Estimates of the numbers, movement, and location of females are also key pieces of information for appropriate management, but there is little information on the movement of females because they are not targeted by the fishery. Current hypotheses on the contribution of primiparous vs. multiparous females to recruitment are predicated on the assumption that eggs are released near the location of the females at the time of the survey (Ernst et al. 2012). These hypotheses suggest that primiparous females are more important to recruitment than multiparous females, but this may be counter-intuitive from an evolutionary standpoint. Small changes in the location and timing of release of larvae could result in different settlement patterns and different inferred contributions of female spawning biomass to recruitment depending on location. Long term tagging studies with archival tags or winter surveys (at the time of the fishery in the area of the Bering Sea not covered by ice) would provide much-needed data to begin to accurately determine the densities of both sexes of snow crab at the times critical to understanding recruitment processes and determining the impact of fishing on reproductive dynamics.

This spatial assessment should not be used to provide management advice. Results of simulation testing with data generated from the spatial model [not shown] show that the spatial assessment is able to reliably return parameter values when optimization is started at the correct values, but when started at random values, it converges on a solution other than the parameter values used to generate the data. This suggests multiple local minima in the objective function, which is not unexpected given confounding among processes. Although estimates from the spatial model should not be used to provide management advice, the spatial model provided a way to test existing stock assessment methods. The currently-used assessment is non-spatial and it appears that it is possible for such assessments to capture the dynamics of spatially-structured populations under certain circumstances. The modeled population is panmictic in spite of having spatial structure (Hardy, 2011), so the difficulties arising from applying a single allowable catch to two different spawning stocks (e.g. Guan et al. 2013) do not apply to snow crab in the Bering Sea.

Unaccounted for spatial structure can bias quantities important in management estimated by aggregate assessment methods under many circumstances, but we show that a spatially-aggregated assessment can capture the dynamics of certain types of spatially-structured populations. Capturing the dynamics of a population is, however, only the first step in sustainable management. Management targets and tactics to reach those targets can be more difficult to calculate and define when spatial structure is present, and studies aimed at identifying spatial structure and exploring the impact of incorrect assumptions on management (as suggested by Begg et al., 1999 and Stephenson, 1999) should be routinely performed for exploited stocks.

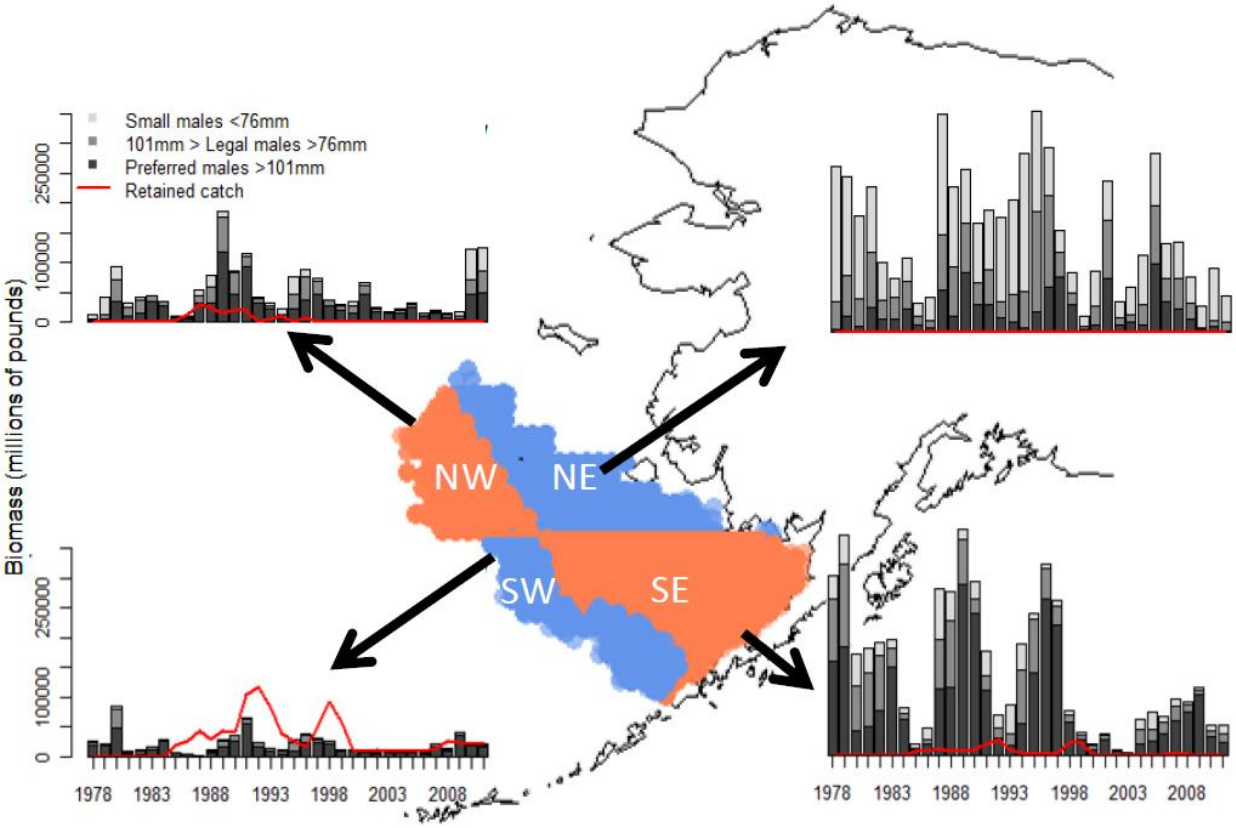

Figure 5.1. Map of spatial areas in relation to Alaska coastline (center) with raw area swept estimates of biomass of male crab from the survey (i.e. catchability equals 1) binned into ‘small’, ‘legal’ and ‘preferred’ size (i.e. the size preferred by fishermen and processors) classes by spatial area (outer). Observed catch is overlaid in the red line.

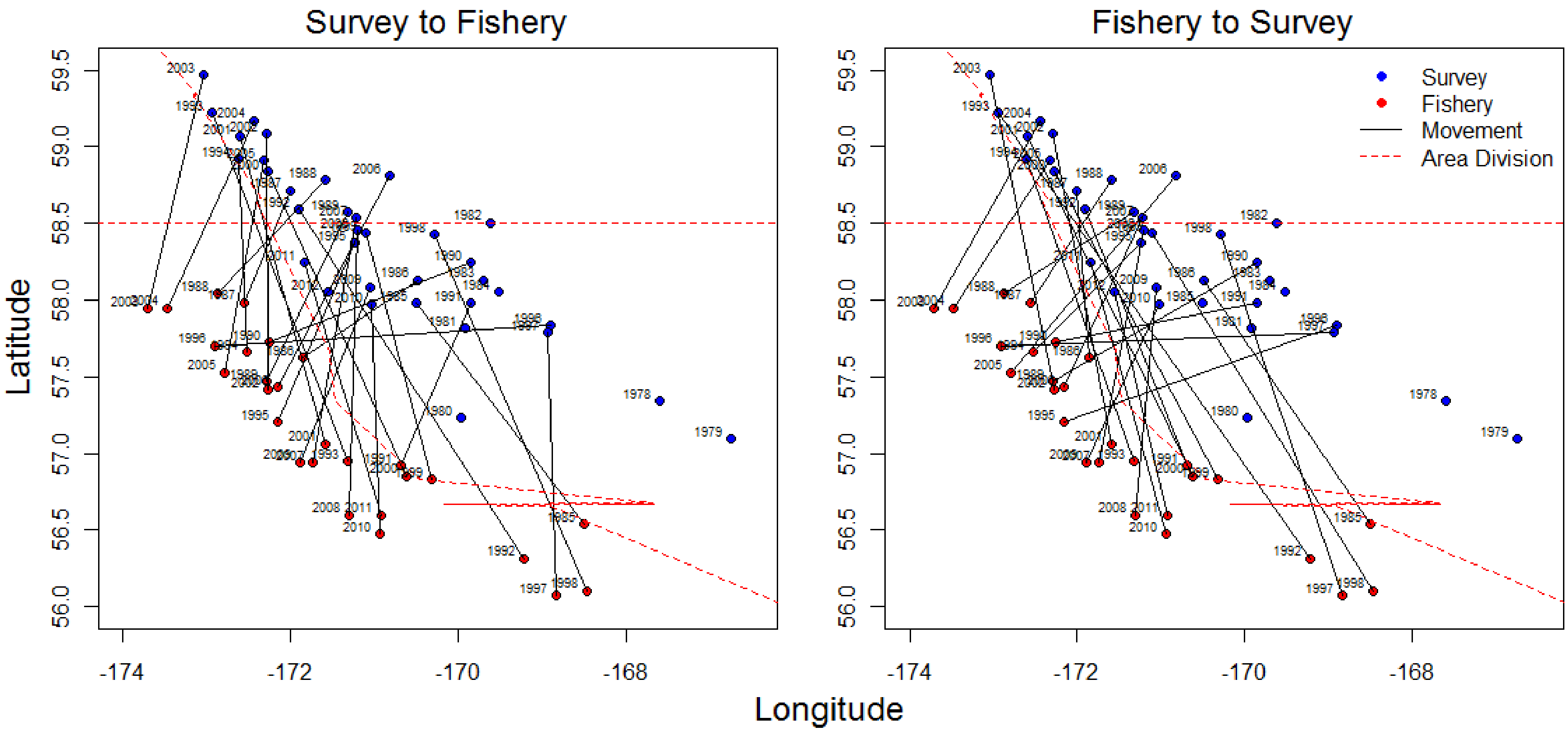

Figure 5.2. Movement of the centroid of abundance for large males (>101 mm) from the time of the survey to the fishery and back. Spatial catch data for some years (1978-1983) were unavailable and therefore no direction of movement was able to be inferred. Red dashed lines indicate the divisions between areas.

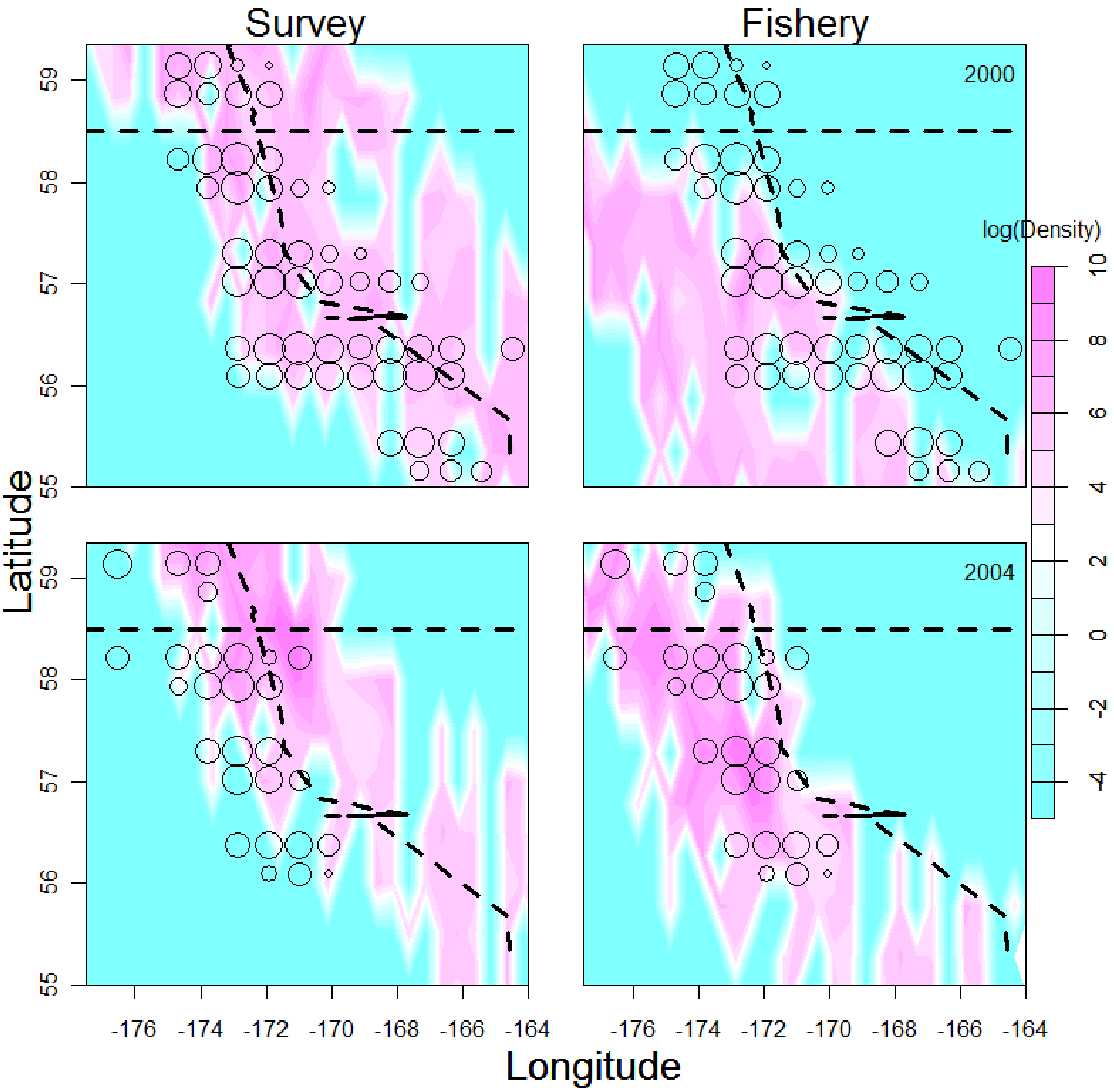

Figure 5.3. Density of large male crab (>101 mm carapace width) for the years 2000 (top row) and 2004 (bottom row). Heat maps in the left column show the density of crab observed at the survey; heat maps in the right column show the density of crab given the assumption that the change in the centroid of abundance for large crab from the survey to the fishery is a proxy for movement in the population. Circles in both columns represent the observed densities.

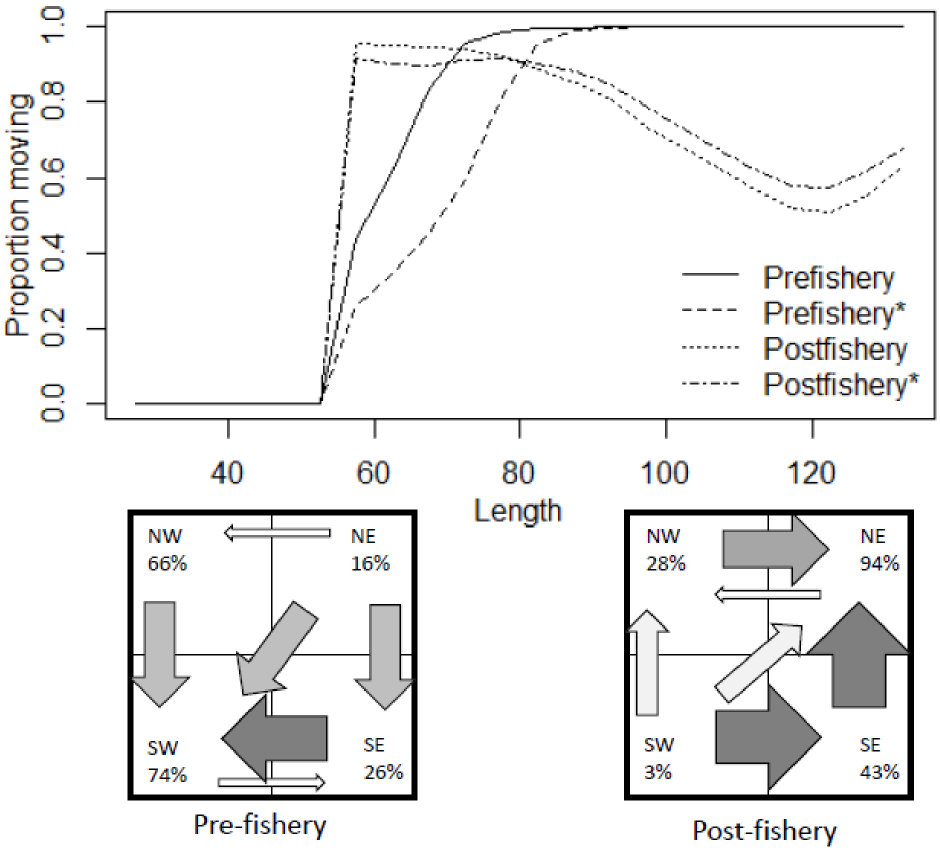

Figure 5.4. Estimated movement-at-length curves (top) for movement before and after the fishery when the centroid priors are included (no asterisk) and when they are excluded (asterisk). Bottom row shows the average proportion in an area moving (which is multiplied by the movement-at-length curve to obtain the proportion of crab at length moving). Percentages in boxes representing an area are the average 'scaling factor' (i.e. the proportion of crab moving) for crabs not moving; the size of arrows leading out of an area indicate the average magnitude and direction of movement over years.

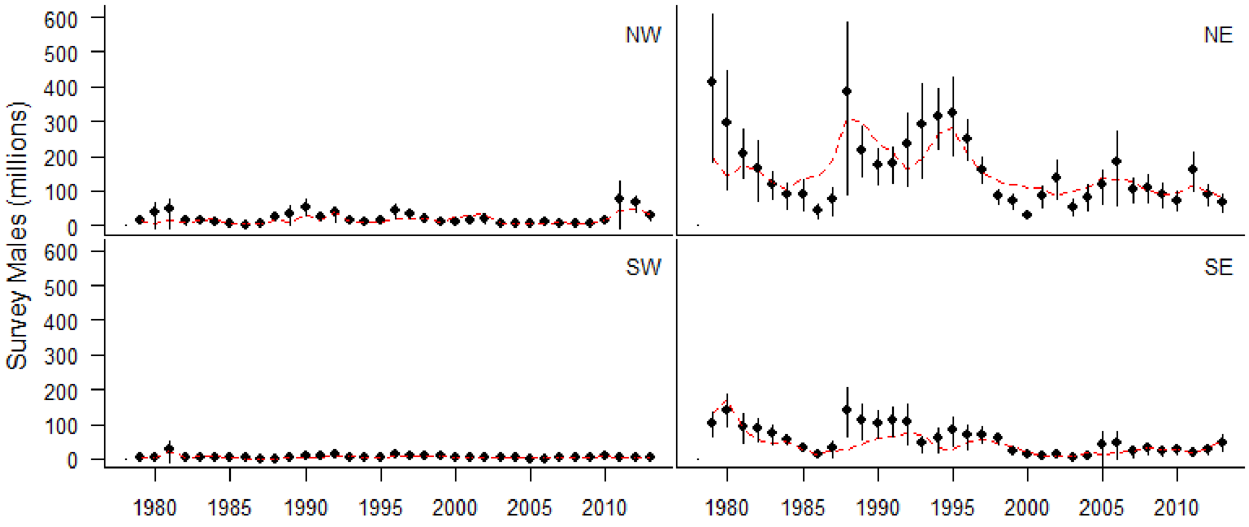

Figure 5.5. Fits (dashed red line) to spatially-disaggregated survey data (points). Lines around the points are boot-strapped 95% confidence intervals.

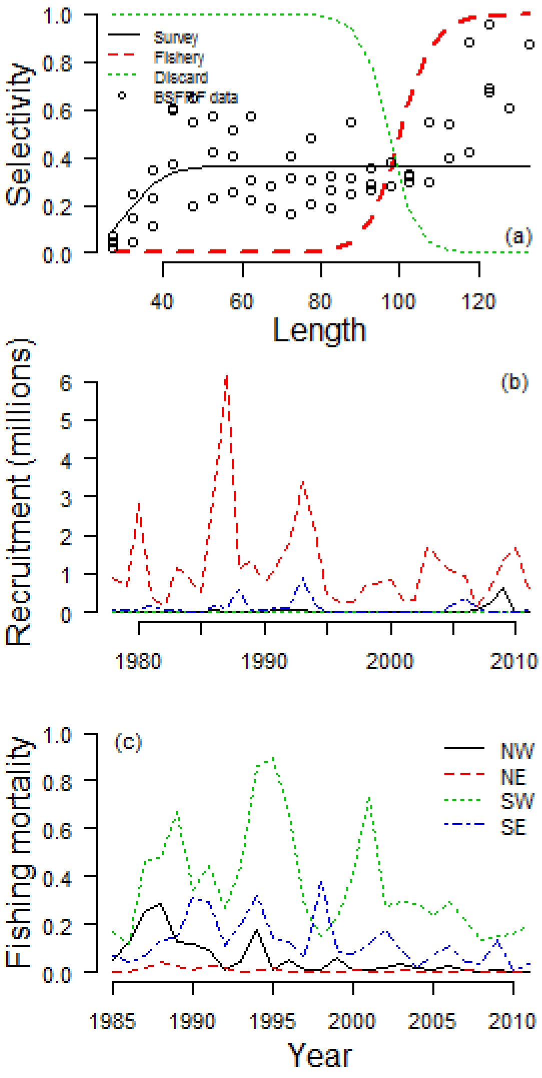

Figure 5.6. Estimated population processes. Selectivities (top) (discard is discard probability rather than retention probability), recruitment by area (middle) and fishing mortality by area (bottom) from the spatial assessment were used to simulate data for testing the aggregate assessment.

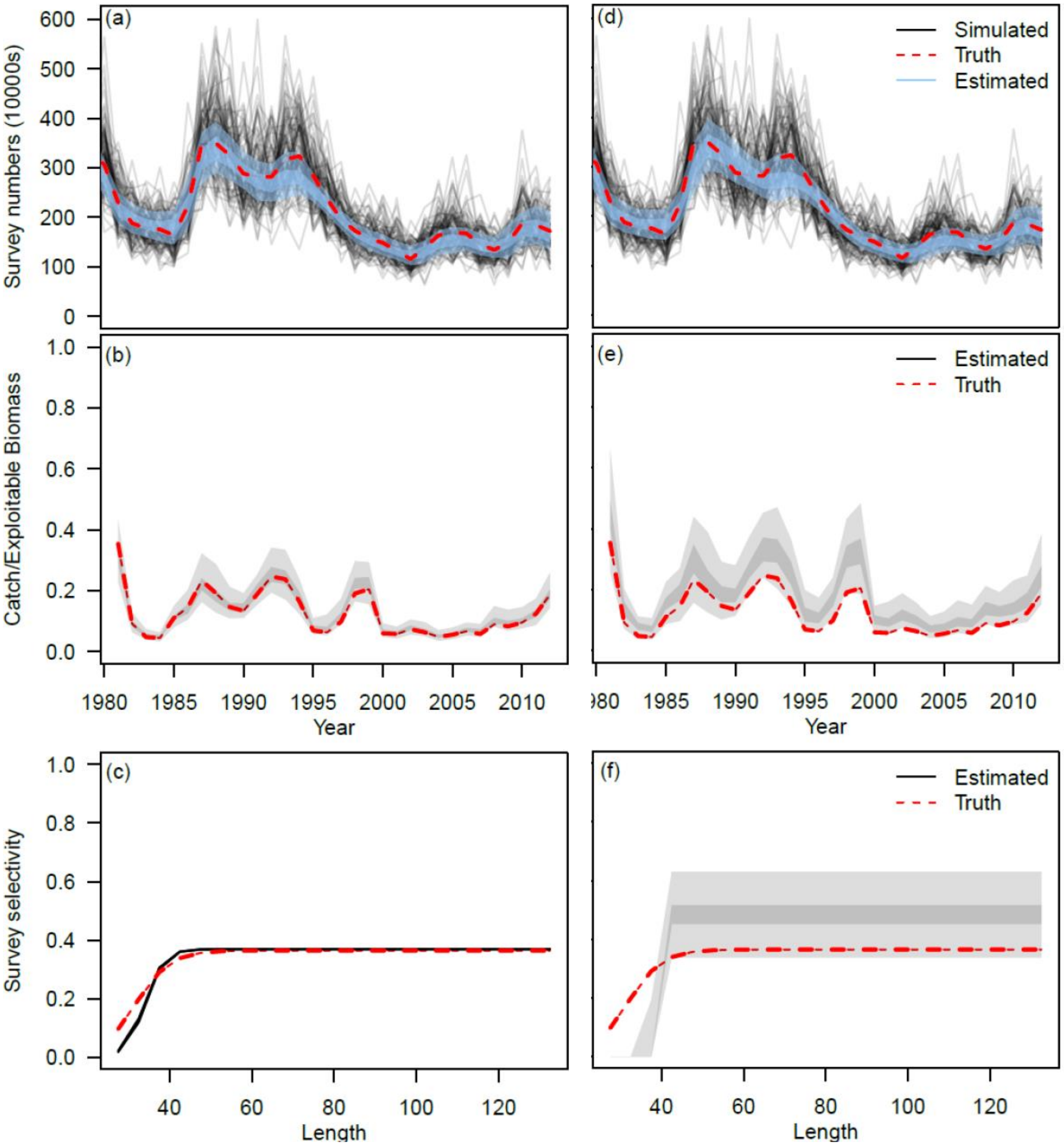

Figure 5.7. Performance of spatially-aggregated assessment method when the assessment data include information to inform survey selectivity (left column) and exclude selectivity information (right column). Light grey (or blue) polygons represent the 95th and 5th quantiles; darker grey (or blue) polygons represent the 25th and 75th quantiles.

Appendix 5.5: Spatial assessment model description

Population dynamics

The model tracks the biannual dynamics of male crabs between the fishery (occurring at variable times during the winter, defined by *midpt*) and the survey (occurring during the summer and indicated by subscript *y*). Animals are grouped by maturity state *m* (mature individuals do not molt, *m=i* denotes immature crab, *m=mat* denotes mature crab), length-class, *l*, and area, *a*:

Survey to fishery dynamics for immature crab (equation A.5.1):

$$N_{a,m=i,y+midpt_y,l} = (1-\varphi_{l)} \sum_{l'} \kappa_{m,l'}\Theta_{a,m,y,l'}\mathrm{X}_{l,l'} + I_{a,y,m,l} - E_{a,y,m,l} + \; Pr_l R_{a,y}$$

Survey to fishery dynamics for mature and maturing crab (equation A.5.1a):

$$N_{a,m=mat,y+midpt_y,l}$$

$$= \begin{cases} \varphi_l \sum_{l'} \kappa_{m,l'}\Theta_{a,m,y,l'}\mathrm{X}_{l,l'} + I_{a,y,m,l} - E_{a,y,m,l} \quad + & if\ m = i \rightarrow mat \\ \Theta_{a,m,y,l'} \; + I_{a,y,m,l} - E_{a,y,m,l} & if\ m = \; mat \rightarrow mat \end{cases}$$

Fishery to survey dynamics (both mature and immature):

$$N_{a,m,y+1,l} = \sum_{l'} \Theta_{a,m,y+midpt_y,l'} + I_{a,y\,\_midpt,m,l} - E_{a,y+midpt,m,l} \tag{A.5.2}$$

where φ_l is the probability of an animal in length-class *l* molting to maturity, $\kappa_{m,l'}$ is the probability of an immature animal molting, $\Theta_{a,m,y,l'}$, is the number of animals of maturity state *m* in length-class *l* at time step *y* (or $y+midpt_y$):

$$\Theta_{a,m,y,l'} = N_{a,m,y,l} e^{-Z_{a,t,l}} \qquad (A.5.3)$$

$X_{l,l'}$ is the proportion of animals in length-class *l*' that molt into length-class *l* given they molt (i.e. the size transition matrix), $Z_{a,t,l}$, is the rate of total mortality on animals in area a and length class *l* during time step *t* (time step *t* is a subdivision of year *y*):

$$Z_{a,t,l} = M_t + \sum F_{a,t,l} \qquad (A.5.4)$$

M_t is the instantaneous rate of natural mortality, and is dependent upon the timing of the fishery, $midpt_y$, $F_{a,t,l}$ is the instantaneous rate of fishing mortality by the directed pot fishery on animals in area a and length class *l* during time step *t*, $R_{a,y}$ is the recruitment to area a during year y Pr_l is the proportion recruiting to length-class *l*, $I_{ax,y}$ is the immigration from area x to area a during time step y (or $y+midpt_y$), and $E_{ax,y}$ is the emigration from area a to area x during time step y (or $y+midpt_y$).

Fishing mortality and selectivity

Fishing mortality is caused by landings and discard in the directed fishery. Fishing mortality is assumed to be independent of whether an animal is mature or not (this may not be the case in reality, but it is assumed to be so in the assessment). Fishing mortality is given by:

$$F_{a,t,l} = S_{l,dir} e^{\bar{F}_a + n_{y,a}} \qquad (A.5.5)$$

where $S_{l,dir}$ is the selectivity of the fishery on animals in length-class l, $\bar{F}_a$ is the average (over time) fully-selected fishing mortality in area a, and $n_{y,a}$ is the deviation in fishing mortality for year y from the average fishing mortality for area a.

Fishery selectivity is assumed to be a logistic function of size and constant over time and area:

$$S_{l,dir} = \left(1 + \exp\left(-\frac{\log(19)\left(\bar{L}_l - L_{50,dir}\right)}{L_{95,dir} - L_{50,dir}}\right)\right)^{-1} \tag{A.5.6}$$

where $L_{50,\mathrm{dir}}$ is the length at which 50% of animals are selected, $\bar{L}_l$ is the midpoint of length-class l, and $L_{95,\mathrm{dir}}$ is the length at which 95% of animals are selected.

Fishery selectivity is used to calculate total catch (i.e. crabs of all lengths brought up in pots). Some portion of the total catch is discarded to give the retained catch. The retained proportion as a function of length is modeled as a logistic function:

$$S_{l,disc} = \left(1 + \exp\left(-\frac{\log(19)\left(\bar{L}_l - L_{50,disc}\right)}{L_{95,disc} - L_{50,disc}}\right)\right)^{-1} \tag{A.5.7}$$

where $L_{50,\mathrm{disc}}$ is the length at which 50% of caught animals are retained, and $L_{95,\mathrm{disc}}$ is the length at which 95% of caught animals are retained. Fifty percent of animals that are discarded are assumed to survive and rejoin the population (Turnock and Rugulo, 2013).

Survey selectivity is used to calculate predicted numbers-at-length in the survey and is modeled as a logistic function:

$$S_{l,surv} = q * \left(1 + \exp\left(-\frac{\log(19)\left(\bar{L}_l - L_{50,surv}\right)}{L_{95,surv} - L_{50,surv}}\right)\right)^{-1} \tag{A.5.8}$$

where $L_{50,\mathrm{surv}}$ is the length at which 50% of animals are caught in the survey gear, $L_{95,\mathrm{surv}}$ is the length at which 95% of animals are caught, and q is the survey catchability coefficient.

Movement

A proportion of crab can immigrate to and emigrate from an area each year. The proportion moving into an area is given by:

$$I_{a,t,m,l} = \sum_{x} N_{a_x,t,m,l}\, move_{l,s}\, Prop_{a_x,t} \tag{A.5.9}$$

where $I_{a,t,l}$ are the numbers in length-class l moving to area a during time step t (which can be y or $y+midpt_y$), $move_{l,s}$ is the proportion moving as a function of length-class and timing, s, referring to whether the movement occurs after the survey or after the fishery. $Prop_{a_x,t}$ is the proportion of crab 'fully-selected' for movement (referred to as a 'scaling factor' in the main text) which move from area x to area a at time step t (which can be y or $y+midpt_y$). The summation occurs over the three areas that are not the area into which crab are immigrating. Emigration (E) is calculated in the same manner, but is subtracted from the total in an area Six movement parameters are estimated for each year (210 in total) and two movement curves (22 parameters each) are estimated (one for before the survey and one for before the fishery; Table A1).

Catch at length

The model prediction of the total and retained catch at are given by:

$$\hat{C}^{tot}_{y,a,l} = \sum_{m} S_{l,dir}\, N_{a,m,y,l} e^{-midpt_{y}*M}(1 - e^{-F_{a,t,l}}) \qquad \text{(A.5.10)}$$

$$\hat{C}^{ret}_{y,a,l} = \hat{C}^{tot}_{y,a,l} S_{l,disc} \qquad \text{(A.5.11)}$$

where $\hat{C}^{tot}_{y,a,l}$ is the model estimate of the total catch of animals in length-class *l* in area *a* during year *y* in numbers, $N_{a,m,y,l}\, e^{-midpty*M}$ is the number of animals in length-class *l* of maturity state *m* when the fishery occurs during year *y*. $(1\text{-}e^{-Fa,t,l})$ is the proportion of crab in area *a* taken by the fishery during year *y*. $\hat{C}^{ret}_{y,a,l}$ is the model estimate of the number of animals in area *a* and length-class *l* which are retained during year *y*.

Growth

Growth is the same among areas. The probability of moulting for immature animals is a declining logistic function of length (mature animals are assumed not to moult given the assumption of a terminal moult at maturity):

$$P_l = \frac{1}{\left(1 + \exp\left(\frac{\log(19)\left(L_{50,moult} - \bar{L}_l\right)}{L_{95,moult} - L_{50,moult}}\right)\right)} \qquad \text{(A.5.12)}$$

where $L_{50,\text{moult}}$ is the length at which 50% of immature animals moult, and $L_{95,\text{moult}}$ is the length at which 95% of animals moult. The growth increment for animals that do moult is based on the gamma function, i.e.:

$$X_{l,l'} = Y_{l,l'} / \sum_{l'} Y_{l,l'} \qquad \text{(A.5.13)}$$

$$Y_{l,l'} = (\Delta_{l,l'})^{(L_l-(\bar{L}_l-2.5))/\beta} e^{-\Delta_{l,l'}/\beta} \tag{A.5.14}$$

where L_l is the expected length for an animal in length-class *l* given that it moults:

$$L_l = \delta_1 + \delta_2 \bar{L}_l \tag{A.5.15}$$

δ_1, δ_2 are the parameters of the relationship between length and growth increment, $\Delta_{l,l'}$ is the difference in length between midpoints of length-classes *i* and *j*:

$$\Delta_{l,l'} = \bar{L}_{l'} + 2.5 - \bar{L}_l \tag{A.5.16}$$

β is the parameter which defines the variability in growth increment and “2.5” is half a length bin’s length.

Recruitment

The fraction of the annual recruitment in an area which recruits to length-class *l* is based on a gamma function, i.e.:

$$Pr_l = (\Delta_{l,l'})^{\mu_1/\mu_2} e^{-\Delta_{l,l'}/\mu_2} / \sum_{l'} (\Delta_{l,l'})^{\mu_1/\mu_2} e^{-\Delta_{l,l'}/\mu_2} \tag{A.5.17}$$

where υ^1, υ^2 are the parameters that define the recruitment fractions. These parameters are the same for all four areas.

The annual recruitments by area are treated as estimable parameters, resulting in four estimated averages and 140 yearly deviations (35 years for each of four areas; Table A1).

Likelihood components

The model is fitted to spatially-disaggregated length frequency for the total catch (L_1), spatially-disaggregated survey length frequency data (L_2), spatially-aggregated length frequency data for the retained catch (L_3), and spatially-aggregated survey length frequency data (L_4):

$$L_1 = \gamma_1 \sum_y \sum_a \sum_l p_{1,l,a,y}^{obs} \ln(p_{1,l,a,y}^{pred} / p_{1,l,a,y}^{obs}) \tag{A.5.18}$$

$$L_2 = \gamma_2 \sum_y \sum_a \sum_l p_{2,l,a,y}^{obs} \ln(p_{2,l,a,y}^{pred} / p_{2,l,a,y}^{obs}) \tag{A.5.19}$$

$$L_3 = \gamma_3 \sum_y \sum_l p_{3,l,y}^{obs} \ln(p_{3,l,y}^{pred} / p_{3,l,y}^{obs}) \tag{A.5.20}$$

$$L_4 = \gamma_4 \sum_y \sum_l p_{4,l,y}^{obs} \ln(p_{4,l,y}^{pred} / p_{4,l,y}^{obs}) \tag{A.5.21}$$

where L_k is the contribution to the objective function of the fit to data type k; $p_{k,l,a,y}^{pred}$ is the model-estimate of the length-frequency from data source k for length-class l in year y by area a (if not aggregated); $p_{k,l,a,y}^{obs}$ is the observed length-frequency from data source k for length-class l in year y by area a. Retained catch length frequencies are not available disaggregated spatially because the retained catch length-frequencies are based on measurements at the processing plants. In contrast, total catch length frequencies are derived from observer data which are taken at sea (and therefore have associated spatial data). Aggregated survey length frequencies are fit in addition to spatially-disaggregated in an attempt to lend stability to the optimization.

The model is also fit to the number of crab by spatial area from the survey (L_5) and the aggregated number of crab (L_6):

$$L_5 = \gamma_{5,y} \sum_{a} \sum_{y} \frac{\left(\ln\left(N_{a,y}^{pred}\right) - \ln\left(N_{a,y}^{obs}\right) + smallNum\right)^2}{\sqrt{\ln\left(\frac{CV_{a,y}}{Shrink}\right)^2 + 1}} \qquad \text{(A.5.22)}$$

$$L_6 = \gamma_{6,y} \sum_{y} \frac{\left(\ln\left(N_{y}^{pred}\right) - \ln\left(N_{y}^{obs}\right) + smallNum\right)^2}{\sqrt{\ln\left(\frac{CV_{y}}{Shrink}\right)^2 + 1}} \qquad \text{(A.5.22)}$$

where $N_{a,y}^{pred}$ is the model-estimate of the number of crab caught in the survey in area *a* during year *y*, $N_{a,y}^{obs}$ is the observed number of crab in the survey in area *a* during year *y* (if spatially-disaggregated), $CV_{a,y}$ is the observed coefficient of variation for $N_{a,y}^{obs}$, *Shrink* is a constant by which the CV is reduced to allow the survey indices to be fit better. *Shrink* does not depend on year and is set to 2. *smallNum* is a small number (equal to 0.001 here) added to avoid taking the log of zero. The data for the last 10 years are weighted more heavily (twice as heavy) as the earlier data to avoid poor fits to these data, which is reflected in the year subscript (*y*) of the weighting factors $\gamma_{5,y}$ and $\gamma_{6,y}$.

Aggregated retained catch and retained catch by area (in biomass) are fit using the following likelihoods:

$$L_7 = \sum_{area} \sum_{y} \frac{\left(\ln\left(C_{a,y}^{pred}\right) - \ln\left(C_{a,y}^{obs}\right) + smallNum\right)^2}{\sqrt{\ln\left(CV_{a,y}\right)^2 + 1}} \tag{A.5.24}$$

$$L_8 = \sum_{y} \frac{\left(\ln\left(C_{y}^{pred}\right) - \ln\left(C_{y}^{obs}\right) + smallNum\right)^2}{\sqrt{\ln\left(CV_{y}\right)^2 + 1}} \tag{A.5.25}$$

where $C_{a,y}^{pred}$ is the catch biomass predicted by the model for year *y* in area *a*, $C_{a,y}^{obs}$ is the observed catch biomass for year *y* in area *a* (if spatially disaggregated), $CV_{a,y}$ is the assumed coefficient of variation for the observed data for year *y* in area *a* (here 0.05 is used), and *smallNum* is a small number added to avoid taking the log of zero when catches do not occur in an area. Aggregate data were fit in addition to spatial data to lend stability to the optimization.

Survey selectivity was estimated from BSFRF and NMFS side-by-side trawls in which BSFRF boats deployed nephrops trawl gear alongside NMFS' gear (Somerton, 2010). The nephrops trawl gear is assumed to have a catchability of 1 and therefore selectivity-at-length of the NMFS gear can be calculated as the ratio of the density of crab caught in the NMFS gear to the density caught in the nephrops gear. This ratio was taken as data which are measures of survey selectivity. Data were collected over two years (2009 and 2010). The likelihood component for the survey selectivity data is:

$$L_9 = \sum_{n} \sum_{l} \gamma_{8,l} (S_{l,surv} - S_{n,l}^{BSFRF})^2 \tag{A.5.26}$$

where, γ_9 is the weighting factor for the selectivity likelihood (the weighting factor is different for each length and proportional to the numbers of observations at each length), $S_{l,surv}$ is the estimated survey selectivity for length-class l (equation 8), and $S_{n,l}^{BSFRF}$ is the nth observed selectivity at length l (there are four replicates (n) of experiments to determine selectivity—immature and mature animals in 2009 and 2010).

Penalty components

Several penalty components were added to the objective function to stabilize estimation. A penalty was placed on the deviations in recruitment and fishing mortality from their respective averages:

$$P_1 = \gamma_j \sum_a \sum_y \varepsilon_{y,a}^2 \tag{A.5.27}$$

where γ_j is the weight for the penalty and (all are equal to 1 for the assessment presented here) $\varepsilon_{y,a}$ is the recruitment *or* fishing mortality deviation for year y by area a.

A penalty is placed on the between-length-class variation in the initial size-structure and the movement curve:

$$P_2 = \gamma_w \sum_l \sum_a \left(\ln(\eta_{l,a}) - \ln(\eta_{l-1,a})\right)^2 \tag{A.5.28}$$

where, $\eta_{l,area}$ is the initial number for length-class l and area a or the proportion moving at length l for area a and γ_w is the weighting factor (equal to 1 in the assessment presented).

Fixed parameters	Number	Equation numbers
Growth	3	13-16
Mortality of discards	1	Not shown
Natural mortality	1	4
Proportion recruiting	2	17
Moulting probability	2	12
Estimated parameters		
Log average recruitment	4	1
Log recruitment deviations	140	1
Log average fishing mortality	4	5,10
Log fishing mortality deviations	140	5,10
Survey selectivity	3	8
Fishery selectivity	2	6
Discard selectivity	2	7
Movement (pre-fishery)	22	9
Movement (post-fishery)	22	9
Proportion moving (pre-fishery)	210	9
Proportion moving (post-fishery)	210	9
Numbers at length in start year	22	1,28

Table A.5.1. Fixed and estimated parameters for the spatial assessment and equations in which they appear.

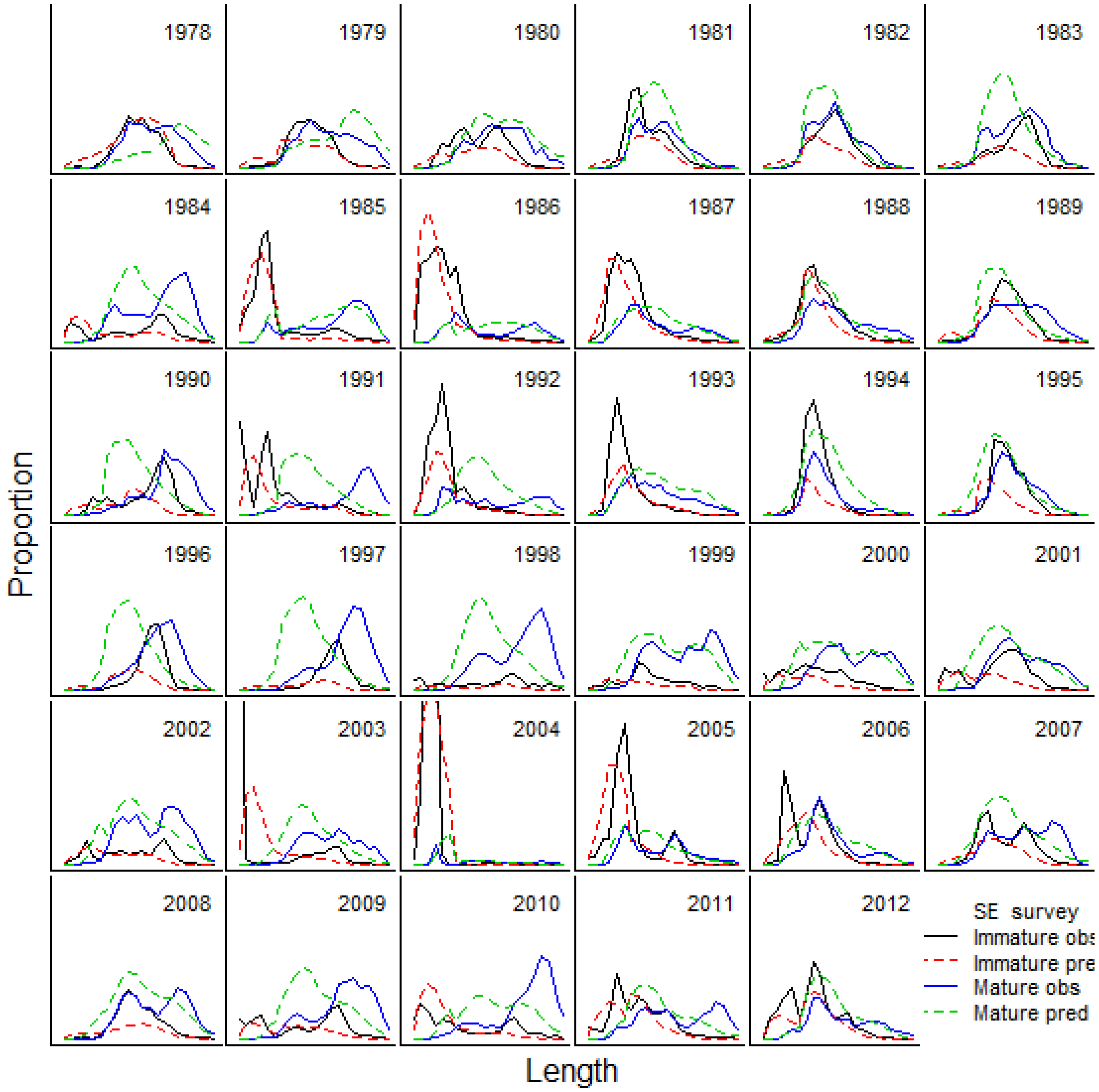

Figure A.5.1. Comparison of observed (solid lines) and predicted (dashed lines) survey length frequencies for the southeast area.

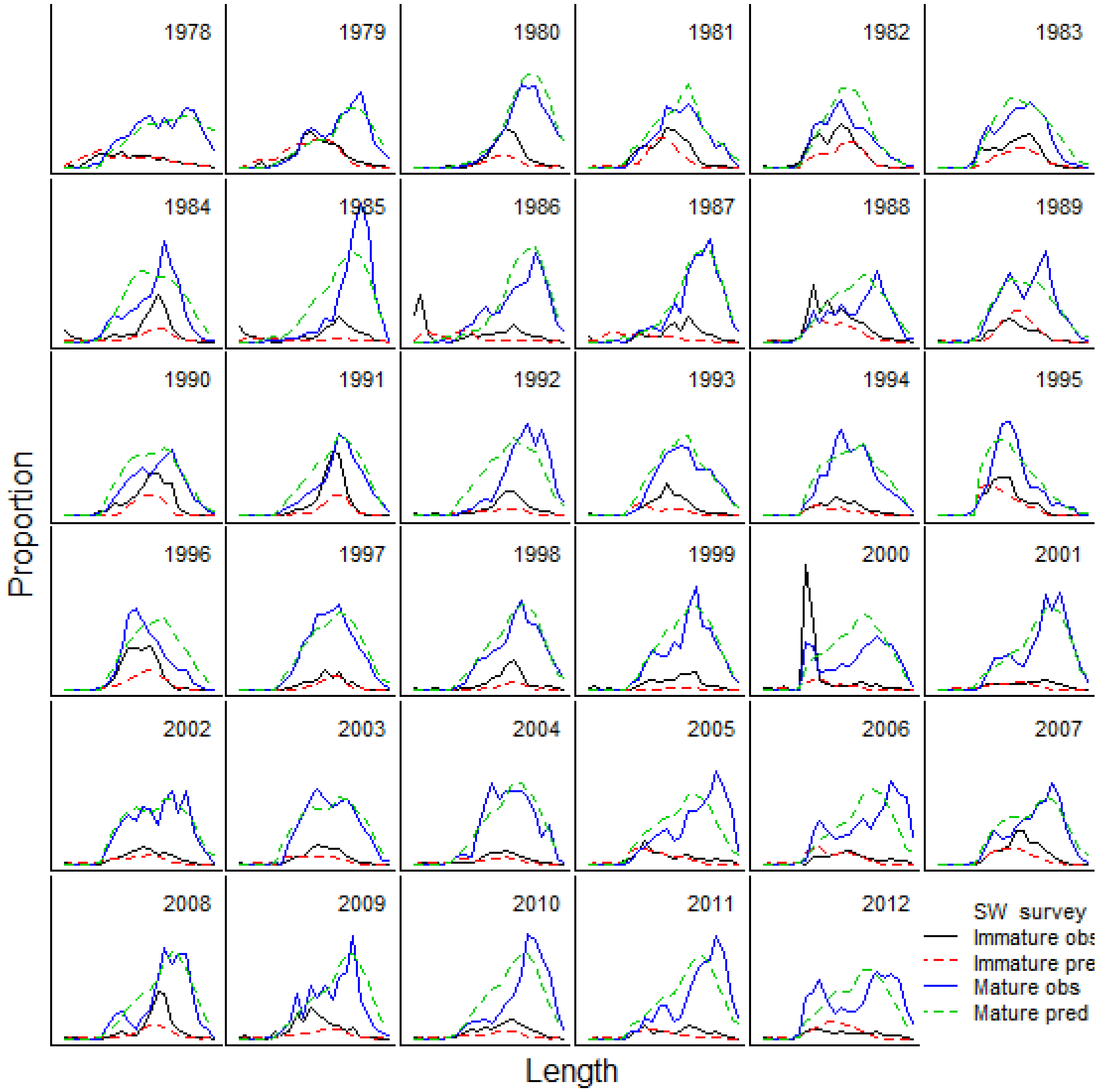

Figure A.5.2. Comparison of observed (solid lines) and predicted (dashed lines) survey length frequencies for the southwest area.

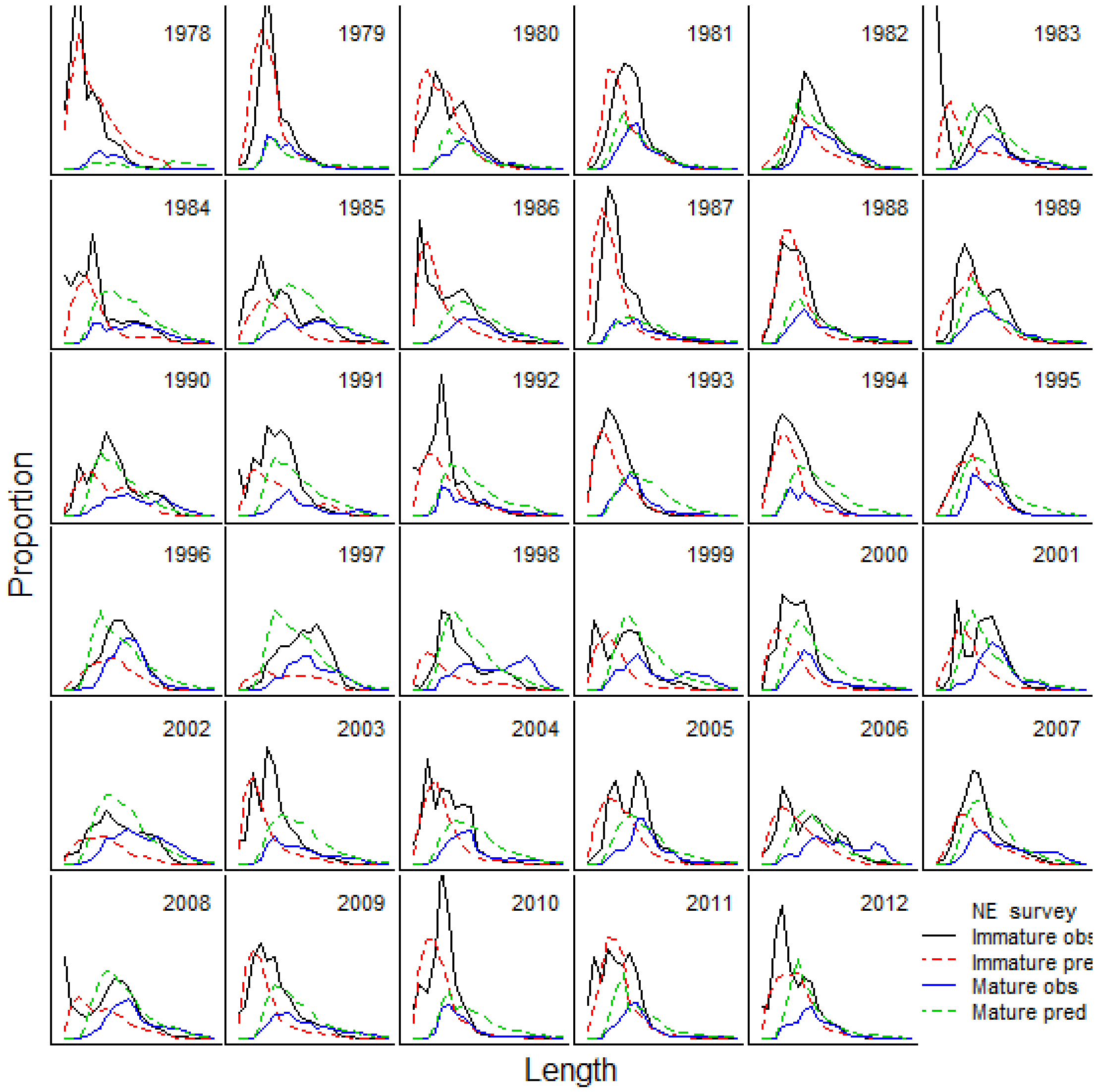

Figure A.5.3. Comparison of observed (solid lines) and predicted (dashed lines) survey length frequencies for the northeast area.

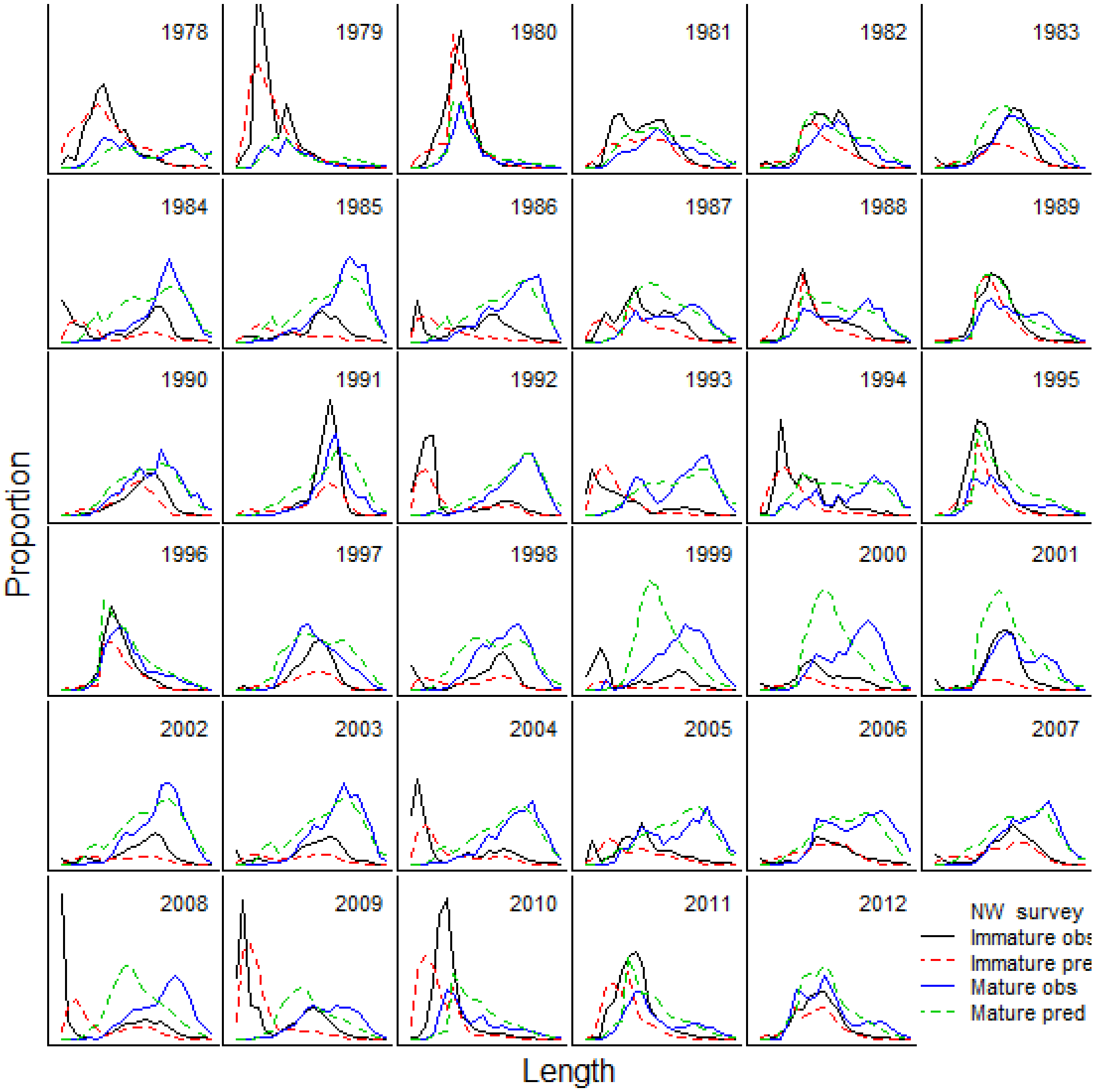

Figure A.5.4. Comparison of observed (solid lines) and predicted (dashed lines) survey length frequencies for the northwest area.

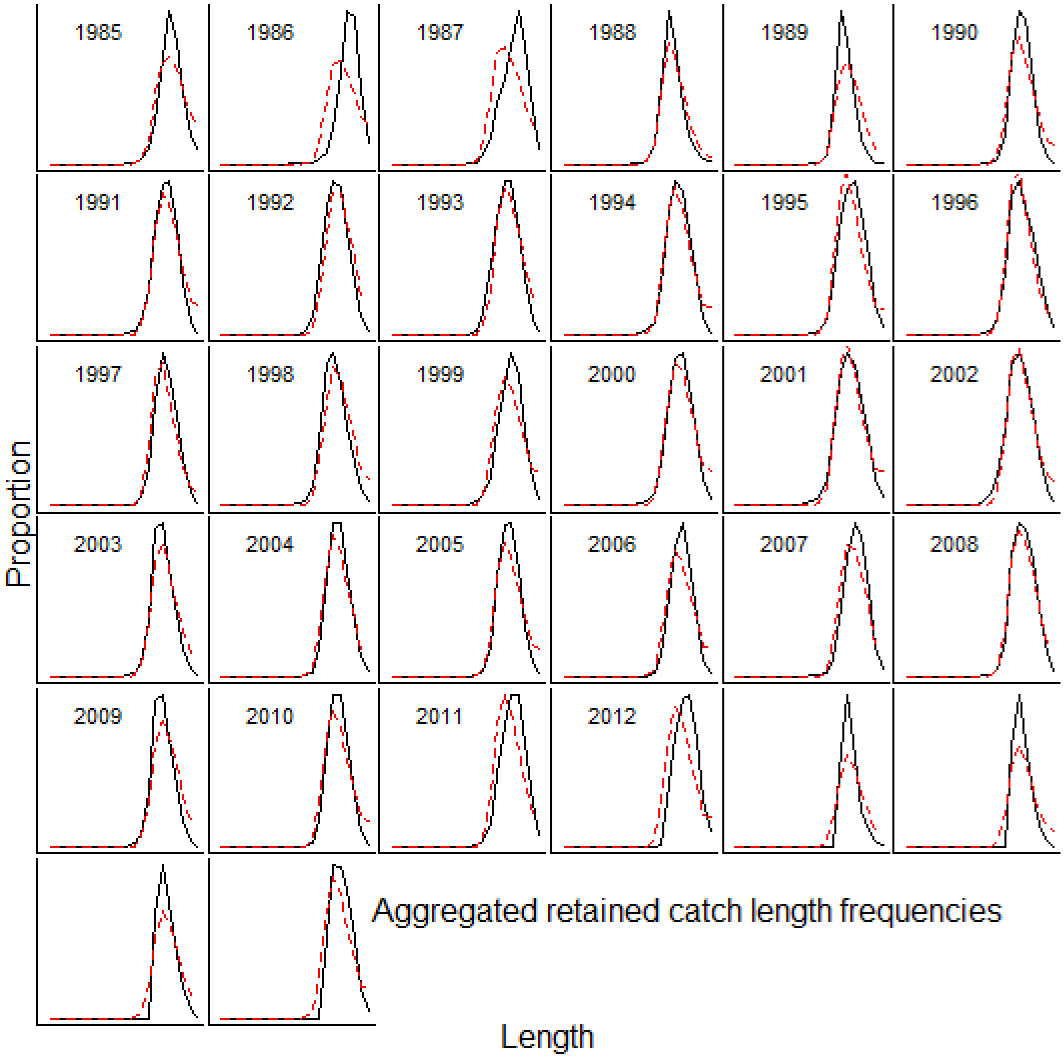

Figure A.5.5. Comparison of observed (solid black) and predicted (dashed red) spatially-aggregated retained catch length frequencies.

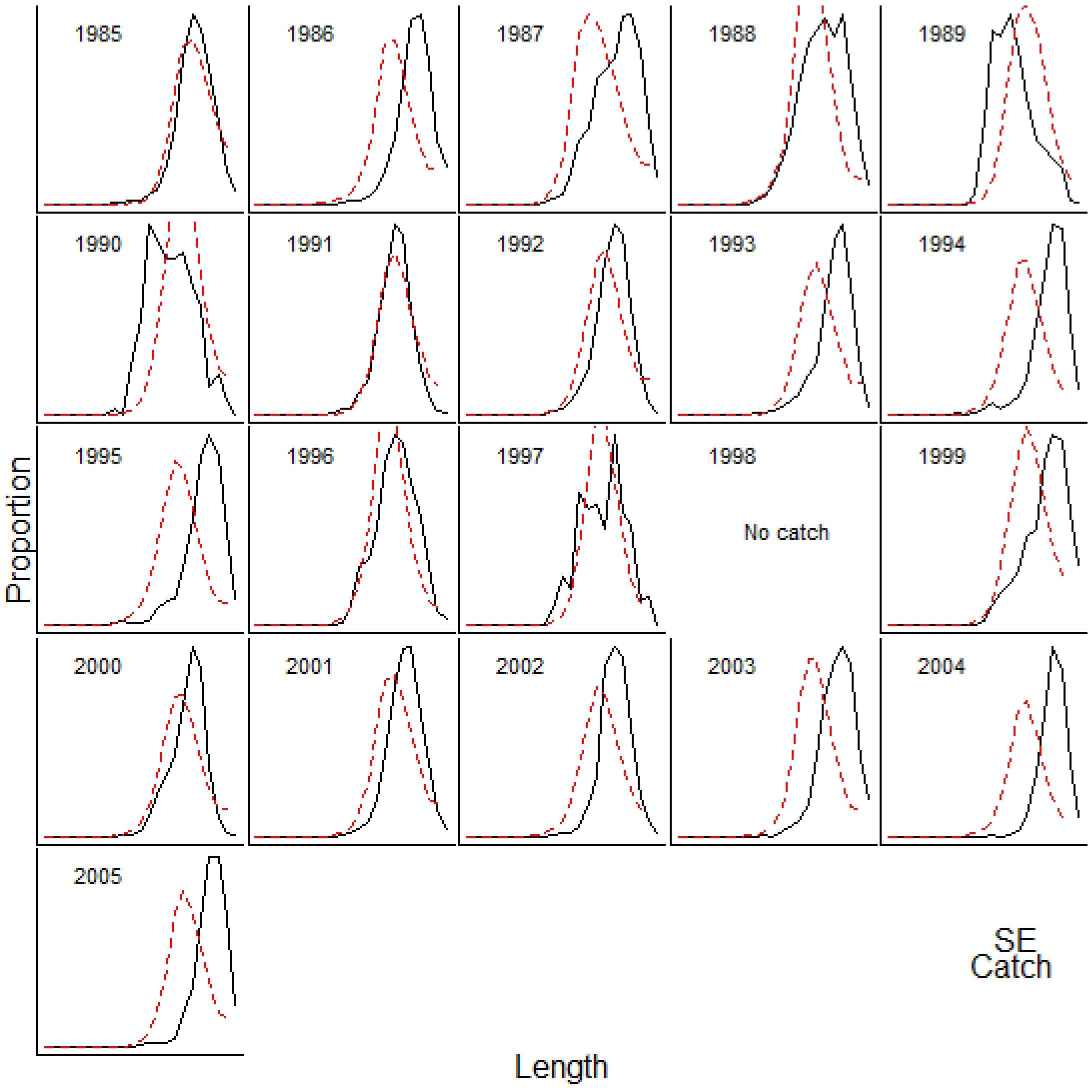

Figure A.5.6. Comparison of observed (solid lines) and predicted (dashed lines) total catch length frequencies for the northwest area.

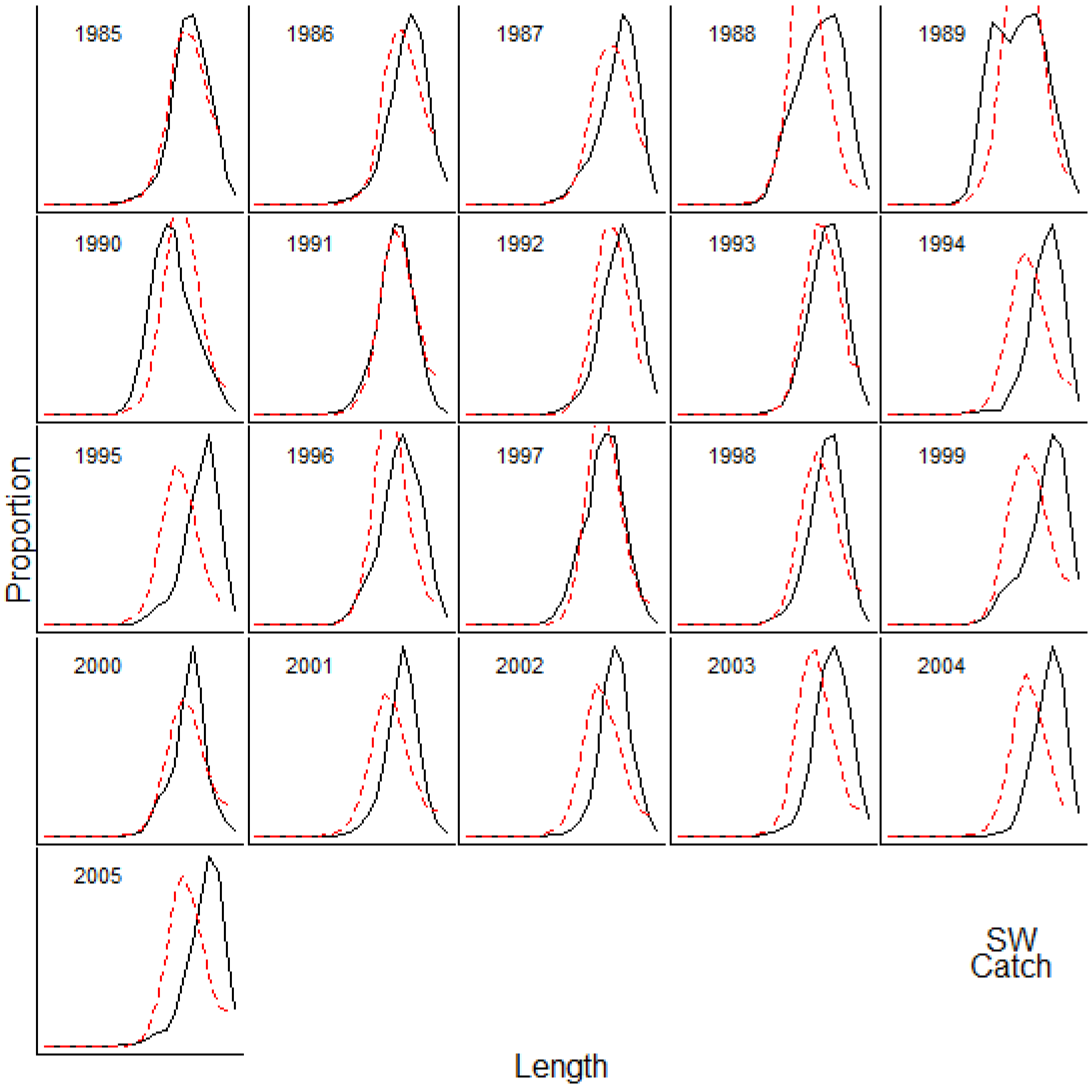

Figure A.5.7. Comparison of observed (solid lines) and predicted (dashed lines) total catch length frequencies for the southwest area.

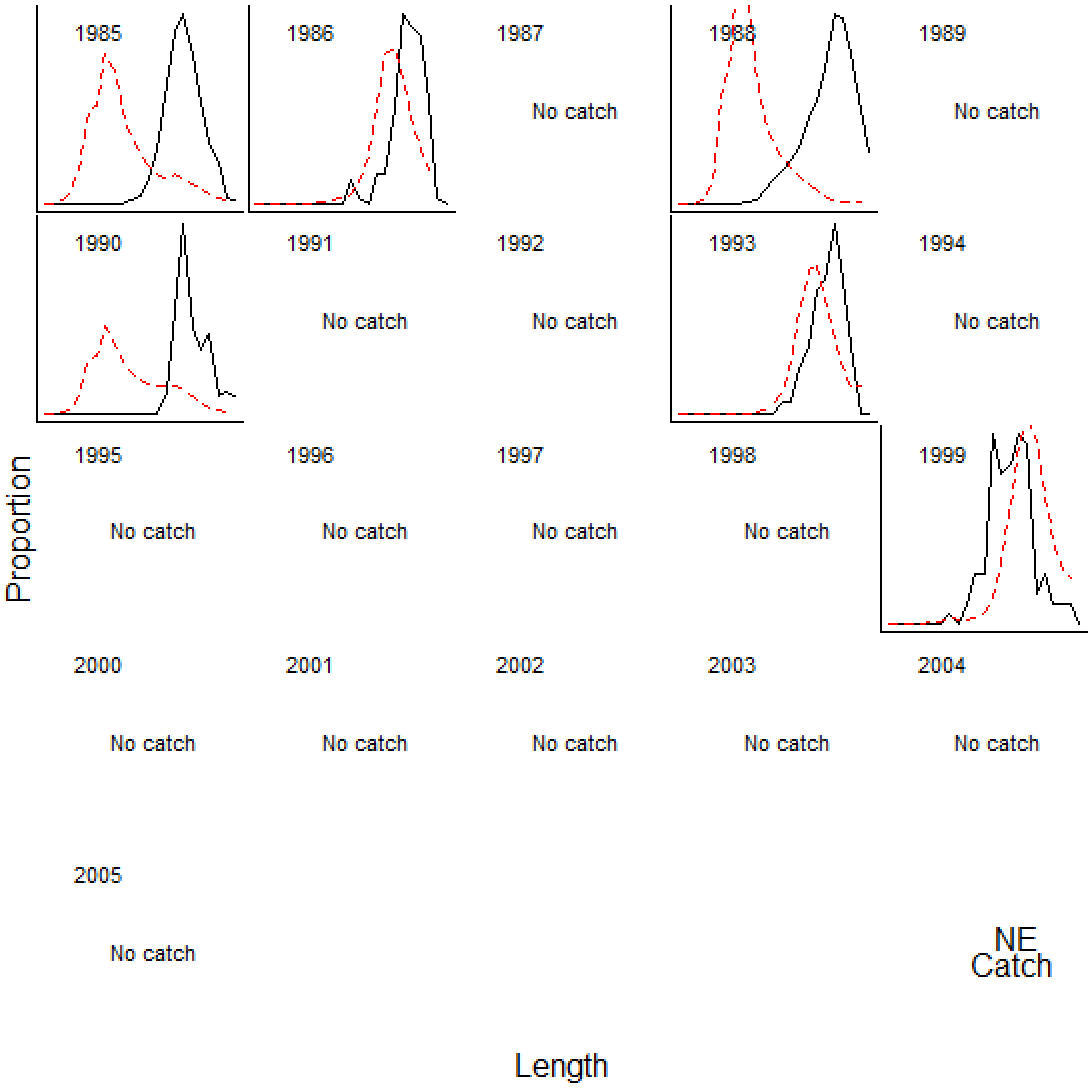

Figure A.5.8. Comparison of observed (solid lines) and predicted (dashed lines) total catch length frequencies for the northeast area.

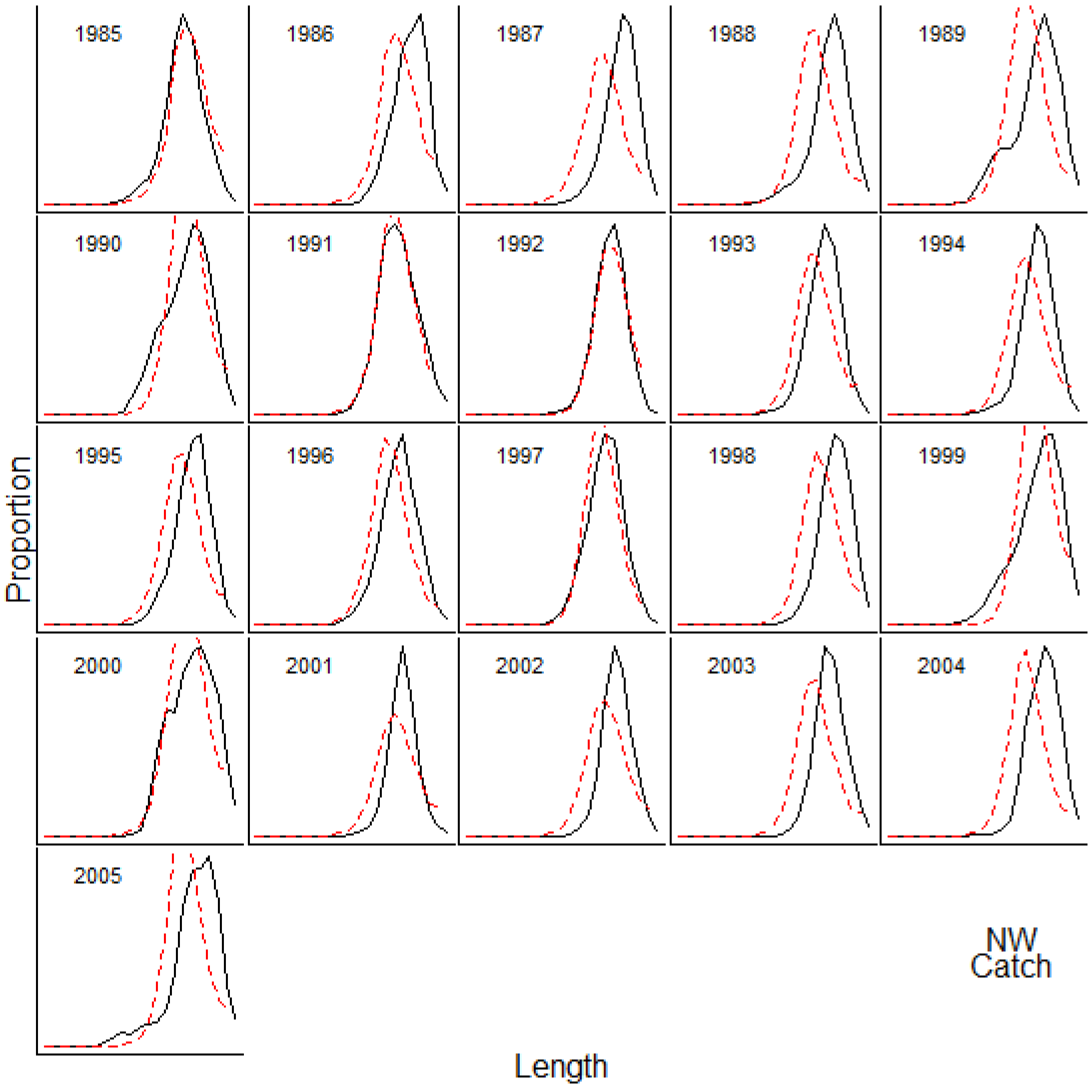

Figure A.5.9. Comparison of observed (solid lines) and predicted (dashed lines) total catch length frequencies for the northwest area.

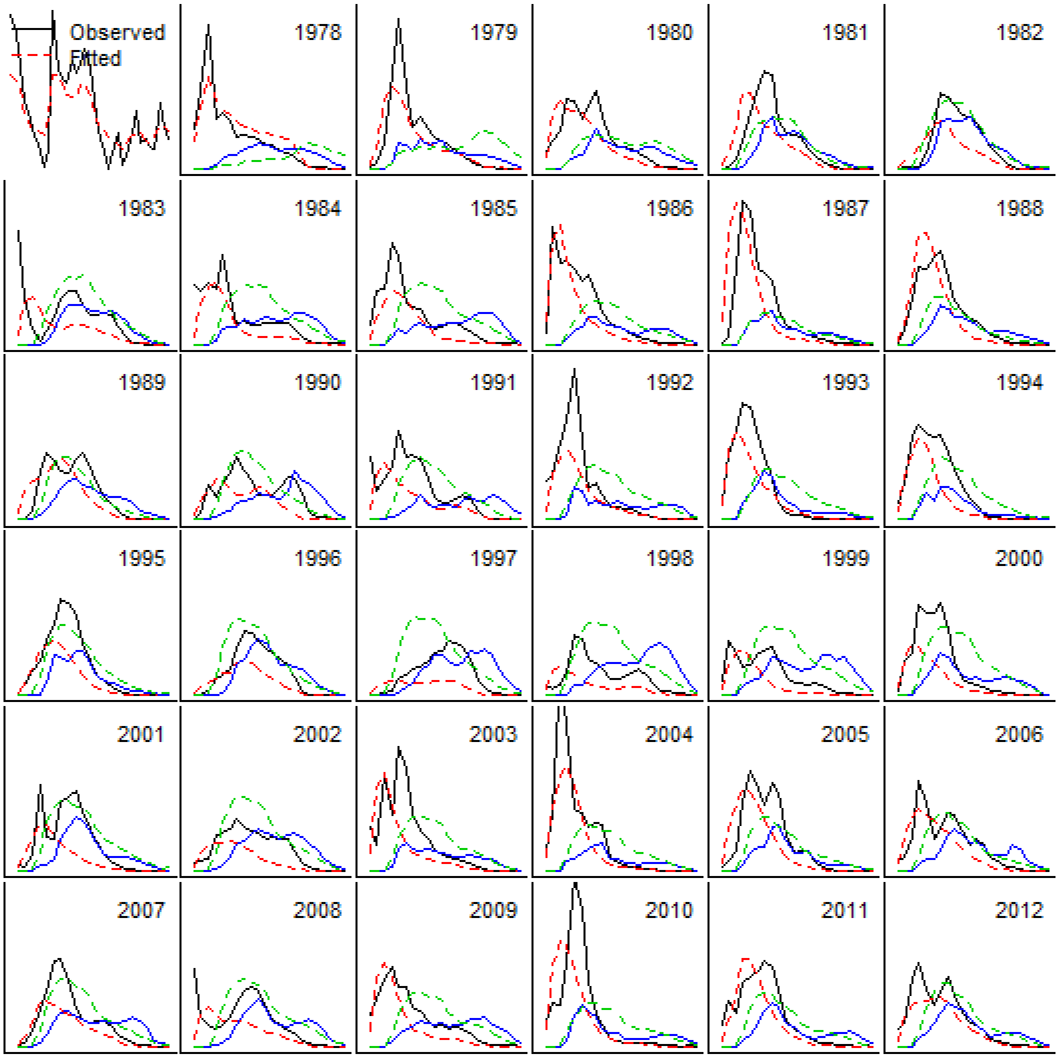

Figure A.5.10. Comparison of observed (solid lines) and predicted (dashed lines) survey length frequencies in aggregate for immature and mature crab. Top left panel shows the observed (solid) and predicted (dashed) time series of numbers observed during the survey

Chapter 6: Conclusion

An overfished status is declared when the estimated biomass of a stock declines beneath a predefined threshold (often based on the inferred productivity of the stock; NPFMC, 2008), but it is not always clear why the biomass declines beneath that threshold (as for snow crab in the eastern Bering Sea, for example). Overfished statuses often result in allowable catches being reduced to rebuild a stock to a higher biomass on the reasoning that cutting catches today can improve long-term yield (Wallace and Cope, 2011). However, there are many potential sources of a decline in estimated biomass to overfished levels. For example, the stock assessment methods may have produced biased estimates of management quantities (e.g. mature biomass) and reference points (e.g. F_{MSY}) on which catch recommendations were based for an extended period of time. Even in the absence of biased management quantities, poor enforcement of allowable catches may stymie management efforts to maintain biomass at a given level. It is also possible that environmental conditions changed such that the processes that drive the productivity of the stock (e.g. growth, recruitment, catchability) have changed (Wildebuer et al. 2013). Changes in these processes can make identifying appropriate management targets against which the ‘overfished’ designation is measured increasingly difficult. Disentangling the relative influences of fishing and climate on the dynamics of any exploited population is one of the key challenges of fisheries management and this problem will become progressively more challenging as climate change moves environmental conditions beyond the range which have been experienced in the natural decadal variability inherent to many systems (e.g. the PDO in the North Pacific, Hare and Mantua, 2000).

This dissertation was undertaken to explore the causes and potential solutions to the overfished status of the snow crab fishery in the eastern Bering Sea. The first potential source of error explored was the stock assessment method and uncertainty associated with the processes modeled within the assessment. Correct assumptions about growth, natural mortality and selectivity were essential for the unbiased estimation of management quantities by the stock assessment method (Chapter 2 and 5). Action is being taken both by the government (NMFS) and the industry (the Bering Sea Fisheries Research Federation, BSFRF) to decrease the uncertainty associated with these processes. For example, the BSFRF has completed side-by-side surveys with the NMFS ship to better understand the selectivity of the NMFS trawl gear (Somerton, 2010). This study resulted in data that were incorporated into the assessment and changed the estimated values of parameters associated with survey selectivity. The BSFRF is also working on tagging studies to inform growth and movement, another area this dissertation identified as an important uncertainty. The assessment method estimated quantities important in management relatively well when the assumptions of the model were correct, so decreasing uncertainty around the processes noted should result in better estimates of mature male biomass and overfishing levels.

A key finding of the simulation work in Chapter 2 was the influence of the penalties on fishing mortality on management. Penalties on fishing mortality can substantially increase the probability of falsely concluding a stock is in an overfished state when this is not the case and vice versa. This finding has prompted recommendations at crab plan team meetings that may lead to revisions to the way assessments are conducted (Punt, 2014) and is reflected in how the management strategy evaluation was conducted.

Other potential sources of error in management relate to the assumptions made about the productivity of the stock in the management strategy. Currently, the management strategy uses proxies to calculate target biomasses that use average recruitment over the entire history of assessment to inform expectations of future recruitment. The recruitment model presented in chapter 3 linking snow crab recruitment to a large-scale index of environmental variation incorporates many of the leading hypotheses on the drivers of year-class strength, and suggests that productivity changes over time dependent upon the environmental conditions of the Bering Sea. Regime-based recruitment suggests that management targets based on the entire time series of recruitment incorporate observations that are not relevant to the current environmental conditions and therefore may misrepresent the productivity of the stock. Although the presented model should not be taken as a predictive model yet, it provided potential future scenarios for management strategy evaluation. Additional years of recruitment observations will be instructional in revising or validating the model. However, if recruitment continues to follow the winter PDO and exhibit regime-like behavior, the practice of basing inference about B_{MSY} on the entire history of the fishery and manipulating fishing mortality to maintain the stock at that level may need to be rethought, particularly considering the projected upward trajectory of sea surface temperature.

A management strategy evaluation (Chapter 4) was undertaken to determine if alternative harvest strategies (that include alternates ways of calculating B_{MSY}) can perform better than the *status quo* in systems with regime-based dynamics. First (and importantly), the MSE showed that the *status quo* assessment method can maintain the stock around B_{MSY} if assumptions about productivity are correct. However, if recruitment dynamics are regime-based, B_{MSY} calculated by *status quo* methods were strongly biased and overfishing occurred more often during ‘low’ recruitment regimes. A regime-based control rule could increase yield in the long term, but it did so at the cost of an increased probability of overfishing, particularly during the ‘high’ recruitment regimes. An increased probability of overfishing occurred because it was difficult to accurately determine the timing and magnitude of shifts in productivity regimes using the currently-available data and techniques. Field research aimed at understanding the recruitment dynamics of snow crab may bolster confidence in the current management strategy (or offer definitive evidence to change). Particularly useful pieces of information which can be measured in the field include: the number of fertilized eggs in relation to sex ratio and spatial overlap of mature males and females at the time of mating (rather than only the time of the survey), larval surveys, and more complete diet data for predators of juvenile snow crab. Future efforts towards both developing methods to detect regime-like behavior and developing management strategies for populations with non-stationary dynamics will be important to managing marine stocks under a changing climate. The results of projections of recruitment in the presented MSE were sensitive to the choice of model for forecasting future recruitment. It would seem prudent therefore to consider a broader set of recruitment models when evaluating harvest control rules than was done in Amendment 24 (NPFMC, 2008).

This dissertation offers an example of the difficulty with which management targets are calculated when a stock appears to be influenced by the environment. This may occur more often for stocks at higher latitudes (which are often subject to 'regime-like' environmental conditions) that are precautionarily managed because we have less data near the origin of the stock/recruit relationship to inform reference points. Ultimately, although the overfished status for eastern Bering Sea snow crab may have been a result of changes in the environment, it is worrying that the status came immediately after a period of intense fishing pressure. A possibility that was not conclusively dismissed by the studies presented in this dissertation was that fishing pressure restricted to the south-western extent of the snow crab range disrupted the reproductive dynamics of the snow crab by removing too many mature males to fully fertilize the available mature females. However, with the current data, we are unable to fully differentiate between the two hypotheses. Consequently, this dissertation underscores the importance of monitoring and data collection at times important to the reproductive dynamics of a given species.

Chapter 7: References:

Adkison, M.D., Peterman, R.M., Lapointe, M.F., Gillis, D.M., and Korman, J. (1996) Alternative models of climatic effects of Sockeye salmon, *Oncorhynchus nerka*, productivity in Bristol Bay, Alaska, and the Fraser River, British Columbia. Fish. Oceanogr. 5(3/4): 137-152.

Alaska Statues. 16.05.251. (2012) "5 AAC 35.517. Bering Sea C. opilio Tanner crab harvest strategy."

A'mar, Z.T., Punt, A.E., Dorn, M.W. (2009a) The impact of regime shifts on the performance of management strategies for the Gulf of Alaska walleye pollock (*Theragra chalcogramma*) fishery. Can. J. Fish. Aquat. Sci. 66: 2222-2242.

A'mar, Z.T., Punt, A.E., and Dorn, M.W. (2009b) The evaluation of two management strategies for the Gulf of Alaska walleye pollock fishery under climate change. ICES J. Mar. Sci. 66: 1614-1632.

Anganuzzi, A., Hilborn, R., Skalski, J.R. (1994) Estimation of size selectivity and movement rates from mark-recovery data. Can J. Fish Aquat. Sci. 51: 734-742.

Arregui, I., Arrizabalaga, H., Kirby, D.S. and Martin-Gonzalez, J.M. (2006) Stock-environment-recruitment models for North Atlantic albacore (*Thunnus alalunga*). Fish. Oceanogr. 15**(**5): 402-412.

Arreguin-Sanchez, F. (1996) Catchability: a key parameter for fish stock assessment. Rev. in Fish. Bio. Fish. 6: 221-242.

Bailey, K.M. (2000) Shifting control of recruitment of walleye pollock after a major climate and ecosystem change. Mar. Ecol. Prog. Ser. 198: 215–224.

Begg, G.A., Friedland, K.D., Pearce, J.B. (1999) Stock identification and its role in stock assessment and fisheries management: an overview. Fish. Res. 43: 1-8.

Bond, N.A., Overland, J.E., and Stabeno, P. (2003) Recent shifts in the state of the North Pacific. Geophys. Res. Lett. 30: 2183e2186.

Bond, N.A., and Overland, J.E. (2005) The importance of episodic weather events to the ecosystem of the Bering Sea shelf. Fish. Oceanogr. 14: 97– 111.

Booth, A.J. (2000) Incorporating the spatial component of fisheries data into stock assessment models. ICES J. Mar. Sci. 57: 858–865.

Butterworth, D.S. and Punt, A.E. (1999) Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56: 985-998.

Cadrin, S.X., Secor, D.H. (2009) Accounting for Spatial Population structure in stock assessment: past, present, and future. In: Beamish, R.J., Rothschild, B.J. (Eds.) The Future of Fisheries Science in North America. Springer, Netherlands, pp.405–426.

Chen, Y., Kanaiwa M., and Wilson, C. (2005) Developing and evaluation a size-structure stock assessment model for the American lobster, *Homarus americanus*, fishery. New Zeal. J. Mar. Fresh. 39: 645-660.

Chen, Y. and Wilson, C. (2002) A simulation study to evaluate impacts of uncertainty on the assessment of American lobster fishery in the Gulf of Maine. Can. J. Fish. Aquat. Sci. 59: 1394-1403.

Comeau, M and.Conan, G.Y. (1992) Morphometry and gonad maturity of male snow crab, *Chionoecetes opilio*. Can. J. Fish. Aquat. Sci. 49: 2460-2468.

Conners, M.E., Hollowed, A.B., and Brown, E. (2002) Retrospective analysis of Bering Sea bottom trawl surveys: regime shift and ecosystem reorganization. Prog. Oceanogr. 55: 209-222.

Cushing, D. H. (1990) Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. Adv. Mar. Biol. 26:250–293.

Deser, C, Phillips, A.S., Hurrell, J.W. (2004) Pacific Interdecadal Climate Variability: Linkages between the Tropics and the North Pacific during Boreal Winter since 1900. J. Climate. 17: 3109–3124.

Dionne, M., B. Sainte-Marie, E. Bourget, and D. Gilbert. (2003) Distribution and habitat selection of early benthic stages of snow crab *Chionoecetes opilio*. Mar. Ecol. Prog. Ser. 259: 117-128.

Duffy-Anderson, J.T., Bailey, K.M., Ciannelli, L., Cury, P., Belgrano, A., Stenseth, N.C. (2005) Phase transitions in marine fish recruitment processes. Ecol. Comp. 2: 205–213.

Ernst, B., Armstrong, D.A., Burgos, J., and Orensanz, J.M. (2012). Life history schedule and periodic recruitment of female snow crab (*Chionoecetes opilio*) in the eastern Bering Sea. Can. J. Fish. Aquat. Sci. 69(3): 532-550.

Ernst, B., Orensanz, J. M., and Armstrong D.A. (2005) Spatial dynamics of female snow crab (*Chionoecetes opilio*) in the eastern Bering Sea. Can. J. Fish. Aquat. Sci. 62: 250–268.

Fisken, O. and Slotte, A. (2002) Stock-environment recruitment models for Norwegian spring spawning herring (*Clupea harengus*). Can. J. Fish. Aquat. Sci. 59: 211–217.

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

Francis, R.I.C.C. (2006) Measuring the strength of environment-recruitment relationships: the importance of including predictor screening within cross-validations. ICES J Mar. Sci. 63: 594-599.

Garber-Yonts, B. (2009) Preliminary economic analysis of alternative snow crab harvest strategies. National Marine Fisheries Service.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (2004) Bayesian Data Analysis, 2nd ed. Chapman and Hall, London. 668 pp.

Goethel, D.R., Quinn, T.J., and Cadrin, S.X. (2011) Incorporating spatial structure in stock assessment: movement modeling in marine fish population dynamics. Rev. Fish.Sci. 19: 119–136.

Gravel K. (2002) Bilan energetique des juveniles du crabs des neiges, *Chionoecetes opilio*, a differentes temperatures. MSc, Universite du Quebec a Rimouski.

Guan, W, Cao, J. Chen, Y. and Cieri, M. (2013) Impacts of population and fishery spatial structures on fishery stock assessment. Can. J. Aquat. Sci. 70: 1178-1189.

Hallett, T.B., Coulson, T., Pilkington, J.G., Clutton-Brock, T.H., Pemberton, J.M., Grenfell, B.T. 2004. Why large-scale climate indices seem to predict ecological processes better than local weather. Nature. 430:71-75

Hardy, S.M., Iken, K., Hundertmark, K., and Albrecht, G. (2011) Defining genetic structure in Alaskan populations of snow crab, Chioenocetes opilio. Final report: OCS Study BOEM 2011-060. http://www.boem.gov/Environmental-Stewardship/Environmental-Studies/Alaska-Region/Index.aspx

Hare, S.R. and Mantua, J.M. (2000) Empirical evidence for North pacific regime shifts in 1977 and 1989. Prog. Oceanogr. 47:103-145.

Hilborn, R. (1990) Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. Can. J. Fish. Aquat. Sci. 47: 635-643.

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

Hinckley, S., Armstrong, D.A., Parada, C., Orensanz, L., Ernst, B., Hermann, A., Napp, J., Megrey, B., Kruse, G., (2010) Modeling transport and survival of larval crab: Investigating the contraction and variability in snow crab stocks in the EBS using individual-based models. North Pacific Research Board Project Final Report 624.

Hobday, D. and Punt, A.E. (2001) Size-structured population modelling and risk assessment of the Victorian southern rock lobster, *Jasus edwardsii*, fishery. Mar. and Freshwater Res. 52: 1495–1507.

Hoenig, J. (1983) Empirical use of longevity data to estimate mortality rates. Fish. Bull. 82: 898-903.

Hollowed, A.B., Bond, N.A., Wilderbuer, T.K., Stockhausen, W.T., A'mar, Z.T., Beamish, R.J., Overland, J.E., and Schirripa, M.J. (2009) A framework for modeling fish and shellfish responses to future climate change. ICES J. of Mar. Sci. 66(7): 1584-1594.

Hollowed, A.B., Hare, S.R., and Wooster, W.S. (2001) Pacific Basin climate variability and patterns of Northeastern pacific marine fish production. Prog. Oceanogr. 49(1-4): 257-282.

Hulson, P.F., Quinn T.J., Hanselman, D.H., and Ianelli, J.N. (2013) Spatial modeling of Bering Sea walleye Pollock with integrated age-structure assessment models in a changing environment. Can. J. fish. Aquat. Sci. 70(9): 1402-1416.

Hunt Jr., G.L., Stabeno, P., Walters, G., Sinclair, E., Brodeur, R.D., Napp, J.M., Bond, N.A. (2002) Climate change and control of the southeastern Bering Sea Pelagic ecosystem. Deep-Sea Research Pt. II. 49 (26): 5821–5853.

Hunt, G.L., Coyle, K.O., Eisner, L.B., Farley, E.V., Heintz, R.A., Mueter, F., Napp, J.M., Overland, J.E., Ressler, P.H., Salo, S., and Stabeno, P.J. (2011) Climate impacts on eastern Bering Sea foodwebs: a synthesis of new data and an assessment of the Oscillating Control Hypothesis. ICES J. of Mar. Sci. 68:1284-1296.

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 J. of Mar. Sci. 68(6): 1297–1304.

Incse, L.S., Armstrong, D.A., and Smith, S.L.. (1987) Abundance of larval Tanner crabs (*Chionoecetes spp*.) in relation to adult females and regional oceanography of the southeastern Bering Sea. Can. J. Fish. Aquat. Sci. 44: 1143-1156.

IPCC. (2007) Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Ed. by R. K. Pachauri, and A. Reisinger. IPCC, Geneva. 104 pp.

Jennings C.A. and Zigler S.J. (2000) Ecology and biology of paddlefish in North America: historical perspectives, management approaches, and research priorities. Rev. Fish Biol. Fish. 10: 167-181.

Jin, M., Deal, C., Wang, J., Alexander, V., Gradinger,R., Saitoh, S.,Iida, T., Wan, Z., and Stabeno, P. (2007) Ice-associated phytoplankton blooms in the southeastern Bering Sea. Geophys. Res. Lett., 34: L06612.

Kanaiwa, M., Chen, Y., and Hunter, M. (2005) Assessing a stock assessment framework for the green sea urchin *Strongylocentrotus drobachiensis* fishery in Maine, USA. Fish. Res. 74: 96-115.

King, J.R (Ed.), (2005) Report of the study group on fisheries and ecosystem responses to recent regime shifts. PICES Scientific Report 28, 162 pp.

Kon, T. (1970) Fisheries biology of the Tanner crab. III. The duration of the planktonic stages estimated by rearing experiments of larvae. Bulletin of the Japanese Society of Scientific Fisheries 36(3): 219-224 (English translation by fisheries Research Board of Canada, Translation Series No. 1602).

Kruse. G.H., Tyler, A.V., Sainte-Marie, B., Pengilly, D. (2007) A workshop on mechanisms affecting year-class strength formation in snow crabs *Chionecetes opilio* in the eastern Bering Sea. Alaska Fishery Research Bulletin. 12: 2.

Kurota, H., McAllister, M.K., Lawson, G.L., Nogueira, J.I., Teo, S.L.H, Block, B.A. (2009) A sequential Bayesian methodology to estimate movement and exploitation rates using electronicic and conventional tag data: application to Atlantic bluefin tuna (*Thunus thynnus*) Can. J. Fish. Aquat. Sci. 66: 321-342.

Lefebvre, L. and Brethes, J.C. (1991) Orientation des deplacements de crabs neiges males (*Chionoecetes opilio*) marques dan le sud-ouest du golf du Saint-Lauren. Can. J. Fish. Aquat. Sci. 48: 1167-1175.

Litzow, M.A. and Ciannelli, L. (2007) Oscillating trophic control induces community reorganization in a marine ecosystem. Ecol. Lett. 10: 1-11.

Litzow, M.A. (2003) Climate regimes shifts and community reorganization in the Gulf of Alaska: how do recent shifts compare with 1976-1977. ICES J. of Mar. Sci. 63: 1386-1396.

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

Mantua, N.J., Hare, S.R, Zhang, Y., Wallace, J.M., Francis, R.C. (1997) A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production. Bulletin of American Meteorological Society. 78: 1069–1079.

Mantua, N.J. and Hare, S.R. (2002) The Pacific decadal oscillation. Journal of Oceanography. 58(1): 35-44.

Mantua, N. (2011) The Pacific Decadal Oscillation. Retrieved August 31, 2011 from http://jisao.washington.edu/pdo/PDO.latest.

McFarlane, G.A., King J.R. and Beamish, R.J. (2000) Have there been recent changes in climate? Ask the fish. Prog. Oceanogr. 47**:** 147-169.

McDonald, A.D., Smith, A.D., Punt, A.E., Tuck, G.N., and Davidson, A.J. (1997) Empirical evaluation of expected returns from research on stock structure for determination of total allowable catch. Nat. Res. Mod. 10: 3-29.

Moriyasu, M., and C. Lanteigne. (1998) Embryo development and reproductive cycle in the snow crab, *Chionoecetes opilio* (Crustacea: Majidae), in the southern gulf of St. Lawrence, Canada. Can. J. Zoo. 76:2040-2048.

Mueter, F.J., Boldt, J.L., Megrey, B.A. and Peterman, R.M. (2007) Recruitment and survival of northeast pacific Ocean fish stocks; temporal trends, covariation and regime shifts. Can. J. of Fish. Aquat. Sci. 64(6): 911-927.

Mullen, A.J. (1994) Effects of movement on stock assessment in a restricted-range fishery. Can. J. Fish Aquat. Sci. 51: 2027-2033.

Murphy, J. (2011) Spatial Dynamics of the Eastern Bering Sea Snow Crab (*Chionoecetes opilio*). (Doctoral dissertation) University of Washington.

Myers, R. A. (1998) When do environment-recruitment correlations work? Rev. Fish Biol. Fisher. 8: 285-305.

NMFS (National Marine Fisheries Service) (2014) Annual Catch Landings. Website: http://www.st.nmfs.noaa.gov/st1/commercial/landings/annual_landings.html

NPFMC (North Pacific Fishery Management Council)/NMFS (National Marine Fisheries Service. (2000) Environmental assessment for the Bering Sea snow crab rebuilding plan. Amendment 14. North Pacific Fishery Management Council, 605 W, 4th Ave, Suite 306, Anchorage, AK 99501, USA or www.dakr.noaa.gov.

NPFMC (North Pacific Fishery Management Council)/NMFS (National Marine Fisheries Service). (2007) Environmental Assessment for Amendment 24. Overfishing definitions for Bering Sea and Aleutian Islands King and Tanner crab stocks. North Pacific Fishery Management Council, 605 W, 4th Ave, Suite 306, Anchorage, AK 99501, USA or www.dakr.noaa.gov.

NRC (National Research Council). (1998) Improving fish stock assessments. Committee on fish stock assessment methods, ocean studies board, commission on geosciences, environment and resources. National Academy Press. Washington, D.C.

Orensanz, J.M, Ernst, B., Armstrong, D.A., Stabeno, P. and Livingston, P. (2004) Contraction of the geographic range of distribution of snow crab (Chionoecetes opilio) in the eastern Bering Sea: An environmental ratchet? CalCOFI Rep. 45: 65-79.

Orensanz, J.M., Ernst, B. and. Armstrong, D.A. (2007) Variation of female size- and stage-at-maturity in snow crab (*Chionoecetes opilio*) (BRACHYURA: MAJIDAE) from the eastern Bering Sea. J. Crustacean Biol. 27: 576–591.

Overland, J. Rodionov, S., Minobe, S., Bond, N. (2008) North Pacific regime shifts: Definitions, issues, and recent transitions. Prog. Oceanogr. 77: 92-102.

Overland, J.E. and Wang, M. (2007) Future climate of the North Pacific Ocean. EOS Transactions of the American Geophysical Union. 88: 178-182.

Parada, C., Armstrong, D.A., Ernst, B., Hinckley, S., Orensanz, J.M. (2011) Spatial dynamics of snow crab (*Chionoecetes opilio*) in the eastern bering Sea—Putting together the pieces of the puzzle. Bul. Mar. Sci. 86(2):413-437.

Porch, C., Turner, S.C., Powers, J.E. (2001) Virtual population analyses of Atlantic Bluefin tuna with alternative models of transatlantic migration: 1970-1997. Col. Vol. Sci. Pap. ICCAT. 52: 1022-1045.

Powers, J.E. and Restrepo, V.R. (1993) Evaluation of stock assessment research for Gulf of Mexico king mackerel: Benefits and costs to management. N. Am. J. Fish. Man. 13:1, 15-26.

Punt, A.E. and Smith, A.D.M. (1999) Harvest strategy evaluation for the eastern stock of gemfish (Rexea solandri). ICES J. Mar. Sci. 56: 860-875.

Punt, A.E., Pribac, F., Walker, T.I., Taylor, B.L., and Prince, J.D. (2000) Stock assessment of school shark, Galeorhinus galeus, based on a spatially explicit population dynamics model. Mar. Fresh. Res. 51(3): 205-220.

Punt, A.E., Smith, A.D.M., and Cui, G. (2002) Evaluation of management tools for Australia's south-east fishery, 2. How well can management quantities be estimated? Marine and Freshwater Research 53, 631–644.

Punt, A.E. (2003a) The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. Fish. Res. 65: 391-409.

Punt, A.E. (2003b) Evaluating the efficacy of managing West Coast groundfish resources through simulations. Fish. Bull. 101: 860-873.

Punt, A.E. (2014) Report of technical crab modeling workshop. North Pacific Fisheries Management Council and National Marine Fisheries Service.

Rodionov, S., 2004. A sequential algorithm for testing climate regime shifts. Geophys. Res. Lett. 31: L09204.

Rodionov, S., Overland, J.E. (2005) Application of a sequential regime shift detection method to the Bering Sea ecosystem. ICES J. Mar. Sci. 6: 328–332.

Rondeau, A. and Sainte-Marie, B. (2001) Variable mate-guarding time and sperm allocation by male snow crabs (*Chionoecetes opilio*) in response to sexual competition and their impact on the mating success of females. Biol. Bul. 201: 204-217.

Rudnick, D.L., Davis, R.E. (2003) Red noise and regime shifts. Deep-Sea Research I. 50: 691–699.

Rugolo, L., D. Pengilly, R. MacIntosh, and K. Gravel. (2005) Reproductive dynamics and life-history of snow crab (Chionoecetes opilio) in the eastern Bering Sea. Final Completion Report to the NOAA, Award NA17FW1274, Bering Sea Snow Crab Fishery Restoration Research. Technical report.

Sainte-Marie, B., Raymond, S., and Brethes, J. (1995) Growth and maturation of the male snow crab, *Chionoecetes opilio* (Brachyura: Majidae). Can.J.Fish.Aquat.Sci. 52:903-924.

Sambrotto, R.N., Niebauer, H.J., Goering, J.J., and Iverson, R.L. (1986) Relationships among vertical mixing, nitrate uptake, and phytoplankton growth during the spring bloom in the southeast Bering Sea middle shelf. Cont. Shelf Res. 5(1/2): 161-198.

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 J. Mar. Sci. 56: 967-979.

Smith, A.D.M. (1994) Management strategy evaluation: the light on the hill. In Population Dynamics for Fisheries Management, pp. 249–253. Ed. by D.A. Hancock. Australian Society for Fish Biology, Perth, Western Australia.

Somerton, D.A. (1981) Life history and population dynamics of two species of Tanner crab, *Chionoecetes bairdi* and *C. opilio*, in the eastern Bering Sea with implications for the management of the commercial harvest. Ph.D. Dissertation. University of Washington, Seattle, p. 220.

Somerton, D., Wenberg, K. and Goodman, S. (2010) Review of the research to estimate snow crab selectivity by the NMFS trawl survey. Report to NPFMC. January 2010.

Stephenson, R.L. (1999) Stock complexity in fisheries management: a perspective of emerging issues related to population sub-units. Fish. Res. 43: 247-249.

Sugimoto, T. and Tadokoro, K. (1997) Interannual-interdecadal variation in zooplankton biomass, chlorophyll concentration and physical environment in the subarctic pacifi and Bering Sea. Fish. Oceanogr. 6**:** 74-93.

Taylor, N., McAllister, M.K., Block, B., Lawson, G. (2009) A multistock age-structured tag integrated model (MAST) for the assessment of Atlantic Bluefin tuna. Col. Vol. Sci. Pap. ICCAT. 64: 513-531.

Turnock, B. J. and Rugolo, L. J. 2009. Stock Assessment of eastern Bering Sea snow crab. In: NPFMC (North Pacific Fishery Management Council). Stock Assessment and Fishery

Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Region: 2009 Crab SAFE. September 2009. North Pacific Fishery Management Council, 605 W, 4th Ave, Suite 306, Anchorage, AK 99501, USA.

Turnock, B.J. and Rugolo, L.J., 2011. Stock Assessment of eastern Bering Sea snow crab. In: NPFMC (North Pacific Fishery Management Council). Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Region: 2011 Crab SAFE. September 2011. North Pacific Fishery Management Council, 605 W, 4th Ave, Suite 306, Anchorage, AK 99501, USA.

Wallace, J.R., Cope, J.M. (2011) Status update of the U.S. canary rockfish resource in 2011. www.pcouncil.org. (last access: September 2013)

Wilderbuer, T.K., Hollowed, A.B., Ingraham Jr., W.J., Spencer, P.D., Connors, M.E., Bond, N.A., and Walters, G.E. (2002) Flatfish recruitment response to decadal climatic variability and ocean conditions in the eastern Bering Sea. Progress in Oceanography. 55: 235–247.

Wildebuer, T., Stockhausen, W. and N. Bond. (2013) Updated analysis of flatfish recruitment response to climate variability and ocean conditions in the Eastern Bering Sea. *Deep-Sea Research II* 94: 157-164.

Wyllie-Echeverria T, Wooster WS (1998) Year-to-year variations in Bering Sea ice cover and some consequences for fish distributions. Fish. Oceanogr. 7:159–170

Ying, Y., Chen, Y., Lin, L., Gao, T., and Quinn, T. (2011) Risks of ignoring fish population spatial structure in fisheries management. Can. J. Fish. Aquat. Sci. 68: 2101–2120.

Zheng, J. and Siddeek, M.S.M. (2010) Bristol Bay red king crab assessment in spring 2010. In: NPFMC (North Pacific Fishery Management Council). Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Region: 2010 Crab SAFE. September 2010. North Pacific Fishery Management Council, 605 W, 4th Ave, Suite 306, Anchorage, AK 99501, USA.

Zheng, J., Kruse, G.H. (2006) Recruitment variation of eastern Bering Sea crabs: Climate-forcing or top-down effects? Prog. Oceanog. 68: 184-204.

Zheng, J., and Kruse, G.H. (2003) Stock-recruitment relationships for three major Alaskan crab stocks. Fish. Res. 65: 103-121.

Zheng, J., and Kruse, G.H. (2000) Recruitment patterns of Alaskan crabs in relation to decadal shifts in climate and physical oceanography. ICES J. Mar. Sci. 57: 438–451.