Fisheries stock assessments for commercial Alaskan species, accounting for age-size-structured population dynamics

Caitlin I. Allen Akselrud

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

Master of Science

University of Washington
2017

Committee
André E. Punt, Chair
Trevor A. Branch
William T. Stockhausen

Program Authorized to Offer Degree:
Aquatic and Fishery Sciences

University of Washington

Abstract

Stock assessment for commercial Alaskan species, accounting for age-size-structured population dynamics

Caitlin I. Allen Akselrud

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

This thesis develops a framework for conducting stock assessments that unifies age- and size-structured assessment models, creating what we term "age-size-structured models." Several studies have explored the application of age-size-structured models to fisheries data. However, most are specific applications or were limited in scope. There is a dearth of generalized single-stock age-size-structured assessment models, which led to the creation and testing of such a model. An age-size-structured assessment method is developed, and examined first using a simulation study, and then using data for two species in the Eastern Bering Sea: Tanner crab (*Chionoecetes bairdi*) and Pacific cod (*Gadus macrocephalus*). Simulations are used to test under which conditions the assessment method is biased or imprecise. The simulations are summarized using a measure of the relative error between true and estimated spawning stock biomass. The simulation study suggested little bias and relatively high precision, such that the estimation method was considered acceptable for use in management. The results from the example application were compared to those from the official NOAA assessments. The estimates of mature male biomass for Tanner crab were found to be similar to those from the most recent stock assessment, but the estimates of spawning stock biomass for Pacific cod differed between the age-size-structured model and the latest official assessment, particularly after the mid-2000s. Some of the dissimilarity may be because maturity and growth are modelled based on size-dynamics in the age-size-

structured model, while the official assessment considers only age-dynamics. Nevertheless, the age-size-structured model produces relatively robust estimates, and could be used as an alternative model to which current stock assessments could be compared. This generalized age-size-structured fishery stock assessment method could be applied to any data-rich stock globally, particularly where a species exhibits a mix of age- and size-based life history and fishery dynamics.

Table of Contents

List of Figures

List of Tables

Acknowledgements

My sincere thanks go to my adviser, André Punt, for bringing me into the world of quantitative ecology and assessment science, for his support and guidance, particularly through steep learning curves, and for always pushing me to become a better scientist. I am truly grateful to have had this experience.

I would also like to thank my committee members: Trevor Branch and William “Buck” Stockhausen. Trevor and Buck were both instrumental in reminding me of the bigger picture of this project. Additional thanks to Buck, from whom I have learned a great deal about crab dynamics, and for the unique opportunity to work with him at the Alaska Fisheries Science Center.

Thank you to Grant Thompson, who took the time to answer numerous questions about Pacific cod and Stock Synthesis. Thank you to Doug Pengilly, Laura Stichert, and Chris Long, who all provided data, references, and answers to many questions about Tanner and blue king crab.

My appreciation also goes out to Anne Hollowed, Owen Hamel, and Melissa Haltuch, who gave me opportunities to work directly on assessments and with assessment scientists at both the Alaska and Northwest Fisheries Science Centers.

I also extend my thanks to the supportive SAFS community, particularly the Punt Lab (including current students, post docs, and former lab members), as well as the Branch, Hilborn, and Essington labs. The collaboration and encouragement from my fellow graduate students has greatly contributed to my programming skills and general knowledge of quantitative methods. I also want to acknowledge all participants who survived FISH 558 and 559 with me, and from whom I learned a great deal.

The North Pacific Research Board funded this work under Project 1304, and I am thankful to the Board for their support.

Dedication

For my husband,
thank you for your eternal support and encouragement.

For my family,
thank you for your boundless love.

Chapter 1. Introduction

Stock assessments have evolved over the last century from simple methods using age or length data to complex methods that rely on multiple data sources and types. Methods for estimating population numbers using catch-at-age data were originally developed in 1949 for an Ontario lake trout fishery (Fry, 1949). Population ecologists have worked since to develop more accurate and precise methods to estimate stock size. Fournier and Archibald (1982) improved upon previous methods to estimate biomass by developing an age-based model that included observational (or measurement) error within the context of a maximum likelihood estimation framework. In contrast, Sullivan et al (1990) proposed a length-based model, which was based on a size-transition matrix. This model eliminated three assumptions made in previous length-based models: (1) length distributions are assumed to be in a steady state over time, (2) recruitment to the population is constant, and (3) growth is deterministic, implying that length-at-age is well defined.

Model development has progressed from simple population dynamics models to age- or length-structured models, and more recently to integrated age- and length-structured models. Fournier and Archibald (1982) were among the first to use "integrated analysis" by combining several sources of information into a single analysis to improve parameter estimation. De Leo and Gatto (1995) made use of an age-size model for European eel (*Anguilla anguilla*) to address several life-history characteristics that are functions of both size and age, including sexual dimorphism, skewed sex ratio, highly plastic body growth, knife-edged selectivity for mature adults, and ontogenetic habitat shifts between juvenile and adult phases. However, that study only used 1.5 years of fishery-independent data. To account for stochastic adult survival and growth, Deriso and Parma (1988) introduced a fully stochastic age-length model, which included recursive updating of parameters to account for stochasticity. They also introduced a likelihood function that could trade-off between fitting to data for age-composition, length-composition, and catch biomass. Gilbert et al. (2006) fitted to more data sources, including fishery-dependent and independent age- and length-compositions, fishery-caught proportions-at-age, tag-recapture estimates of numbers at length, catch per unit effort, and an environmental index: sea surface temperature. They built on previous models, and represented the population using a matrix of numbers at age and at length, allowing for age-length-dependent growth to vary over time.

Information about fish size, life stage, and age can be used to construct models of increasing complexity. Fish growth is often modelled according to a standard growth curve, such as the von Bertalanffy growth equation (von Bertalanffy, 1934). The von Bertalanffy growth curve is used to represent the relationship between the age of an individual fish and its expected size. Age-structured models track cohorts over time, relating a specific age of fish to length using a growth equation. Length-structured models track length-distributions over time using a transition matrix to determine the probability of growth to a specific size between years. Additionally, there are stage-structured models, which use life-history stage information to estimate population growth rate over time.

Relationships between size and age allow researchers to follow the distribution of length within a population. Tracking age and length distributions over time is important because population dynamics parameters, such as fecundity and mortality, are driven by both age and length dynamics. Fish populations are assessed using age-structured and length-structured techniques, depending on the data available. Fish are typically easier to age than invertebrates because fish have hard parts, such as otoliths or scales, which add a countable growth layer each year (Chilton and Beamish, 1982; Boehlert, 1985). Invertebrates present a problem because they are often soft-bodied or molt, and do not have a sequential record of age over time (W. Stockhausen, pers. comm., January 2015).

Invertebrate species are often measured by length or categorized by life-stage. Consequently, stock assessments for invertebrates tend to be based on length-structured or stage-structured models. However, the population dynamics of vertebrates and invertebrates are driven by age and length (or stage) processes. This is particularly important to take into account when estimating fishing and natural mortality.

Age- and length-structured models make assumptions about expected length at a given age or time. Age-structured models use a growth equation that predicts expected length based on age. One major challenge with age-structured models is the need to age species. The ageing process is expensive and time-consuming, and often cannot be done for invertebrate species. Length-structured models depend on a size-transition matrix, which determines the probability of an individual growing to a specific size over a specific time interval given its current size. A potential problem both model types face is confounded parameters, such as natural mortality and selectivity (Punt et al., 2013). A model that unifies both age- and length-structured processes may overcome the limitations of each individually, and account for population dynamics controlled by both.

This thesis develops a framework for conducting stock assessments which unifies age- and size-structured assessment models, creating what we term “age-size-structured models.” The framework was developed collaboratively for NPRB Project 1307 (Appendix 1). Current assessment methods can include many data sources and factors in an integrated analysis (e.g. Maunder and Punt, 2013), including fishing effort, the effect of environmental factors, and accuracy of ageing techniques. The stock assessment method developed in this project can use multiple data sources, including age- and length-compositions, and has three components: (1) an age-structured model, (2) a size-structured model, (3) and an age-size-structured model. By comparing results between model types, stock assessment scientists can explore sensitivities among models, and select the most appropriate model.

The population dynamics in the age-size-structured model are governed using a size-transition matrix. Growth as a function of age can be estimated using the von Bertalanffy equation, with variability in the growth increment assumed to be normal. Additional options, which are particularly useful for crab stocks, include a size-increment of either a linear or log-linear function of size, with molt probability a declining logistic function of size. The distribution around the size increment can either be normal, log-normal, or gamma distributed. Individual variation in growth may not be captured by a single size-transition matrix, so the age- and length-structured model uses an option called “platoon-structure,” similar to that used in the Stock Synthesis fisheries stock assessment model (Taylor and Methot, 2013). Platoon-structure divides each age class into distinct groups, each with its own growth curve. This allows organisms within an age-class to display different rates of growth, leading to different length-at-age distributions and hence a more appropriate account of length-specific mortality (Lee, 1920).

Expected recruitment to a population, at age 0, is based on the Beverton-Holt stock-recruit relationship (Beverton and Holt, 1957). This relationship determines the expected number of age-0 individuals that will join the population in the next year, based on the mature biomass during the current year. The impact of different fleets on the population dynamics is included through fleet-specific selectivity patterns that account for differential impact of gear types on population size and age structures (Hurtado-Ferro et al., 2014). Selectivity can be age-specific, length-specific or both age- and length-specific.

The age-size-structured model includes two options for representing fishing mortality associated with catch. For data-poor fleets, or industries with low observer coverage (and thus more uncertain catch and discard values), the model can estimate annual fishing mortality as a parameter. However, for fleets where there is near-perfect reporting of catches, the "hybrid method" (Methot and Wetzel, 2013) should be used to calculate full-selection fishing mortality since it is more efficient than the parameter estimation method. The age-size-structured model also takes discard mortality into account, with the probability of a fish being discarded modelled as a function of length or age.

Time blocks are an option in the modeling framework, where parameters for biology, selectivity, or catchability are assumed to be the same within each block, but can vary between blocks. This is useful when there are time-varying processes, such as time-varying natural mortality (caused by regime shifts), or changes in selectivity over time (caused by changes in gear technology or fishing behavior).

Initial testing of the model framework included a simulation evaluation varying the effective sample size for length-composition data from catches, discards and a survey, and varying the coefficient of variance (CV) for the index of abundance and effort data. The framework allows for an evaluation of the impact of age-size-dynamics on the assessment outcomes. The framework is applied to data for two stocks that differ in terms of data availability and quantity: Eastern Bering Sea Tanner crab (*Chionoecetes bairdi*) (Stockhausen, 2015), and Eastern Bering Sea Pacific cod (*Gadus macrocephalus*) (Thompson, 2015; Cohen et al., 1990). The empirical data sets used to fit the model come from the NOAA Alaska Fisheries Science Center, and include time series of catch and length-composition (fishery-dependent data), and NOAA trawl survey index of abundance, length-composition, and age-composition (for cod) data (fishery-independent data). The model results are compared to those from the official assessments of each stock, and used to evaluate the likely magnitude of change in assessment outcomes arising from the adoption of this framework.

The objectives of this thesis include: 1) using simulation evaluation to test for bias and imprecision in the estimation method under several configurations; 2) applying the age-size-structured model to an invertebrate species: Eastern Bering Sea Tanner crab, and comparing estimates of recruitment and spawning biomass against the current assessment (Stockhausen, 2015); and 3) applying the age-size-structured model to a vertebrate species with data for both length and age: Eastern Bering Sea Pacific cod, and comparing estimates of recruitment and spawning biomass against the current assessment (Thompson, 2015). The simulation evaluation is important in determining whether there is inherent bias or imprecision in the estimates of biomass from the estimation method, and under what configurations accuracy in the assessment can be improved. The goal of the last two objectives is to assess whether the estimation method is viable for multiple categories of stocks with differing data types.

Chapter 2: Simulation Study

2.1 Introduction

Simulations are used to create circumstances where the truth is known and estimation methods can be evaluated based on their ability to approximate the truth. In this way, simulation studies identify strengths and weaknesses of estimation methods and allow for a better understanding of their limitations. In this chapter, the performance of the assessment method is examined under varying levels of data precision. An operating model is used to represent the truth for the simulations, and is used to generate pseudo datasets that could have arisen if the operating model was real (Fig. 2.1; Punt, 2003). It is important to examine alternative configurations of the operating model to identify potential biases and examine robustness to multiple plausible conditions (Punt et al., 2006). The amount of deviation in estimates from the truth can show under which configurations the assessment method is imprecise or biased. In this study, precision is defined as the amount of variability about the simulated true value and bias is characterized by estimated values consistently occurring above (or consistently below) the true value.

Assessment methods take disparate pieces of information about a fish stock and put them together to try to identify the underlying dynamics of the population and the true population biomass (Hilborn and Mangel, 1997). There are many ways to approach population assessment, ranging from a simple population model that aggregates biological processes, to more complex models including age- and size-structure (Haddon, 2010). While these models can provide helpful information about population biomass, they can also have inherent biases, particularly when life history or stock dynamics are not well known. It is important to understand when population estimates may be biased or imprecise because estimates of population biomass are critical to managing populations sustainably under fishing pressure. Simulation testing is often used to identify under which conditions the assessment results are biased or imprecise.

The simulation framework consists of an operating model, which produces the “truth” for the simulations, and an estimation method, which selects parameter values to produce the best fit to the data. It is generally recommended that the operating model be sufficiently complicated such that conclusions are robust to key biological processes and uncertainties that would likely occur (Punt et al., 2006).

Managers often track the amount effort fishers put in to catching fish, such as the number of times crab pots are pulled from the ocean or the number of hooks on a line in longline boats. Information about effort can inform biomass estimates when examined relative to catch. If, for instance, catch increases over time but effort remains the same, it can indicate that the fishery fleet is becoming more efficient at catching fish over time rather than indicating an increase in biomass over time. The survey forms the basis from which fishery-independent biomass is estimated, and uncertainty in the survey biomass estimates and the value of the effective sample size for survey length-composition may result in biased or imprecise estimates of population biomass. The operating model configurations examined here include varying uncertainty and effective sample size for several types of data.

2.2 Methods

2.2.1 Model Configuration: Base Model

The simulation framework takes input information and generates a “true” population. The estimation method is then used to estimate the parameters of a population dynamics model, given data generated from the operating model (Fig. 2.1). The specifications of the operating model are available in Punt et al. (2017), and the estimation method is detailed in Appendix 1.

2.2.1.1 Operating model

The inputs required to generate a "true" population include the effective sample sizes for the length-compositions for catches, discards, and survey samples by fleet and by year, sampling error associated with survey estimates of biomass, expressed as a coefficient of variance (CV) for the index of abundance, and the uncertainty of the fishery effort (also expressed as a CV). The operating model generates catch, discard, survey index, effort, size-composition, and age-at-length-composition data. A 45-year fishery and survey period is established in the operating model. Discards include any individuals that were caught in the directed fishery, but not kept for reasons such as being below a minimum-size threshold or females in a male-only fishery. Retention is modeled using a logistic curve to represent the proportion of animals at size that are expected to be kept by the fishery, with mean size-at-50%-retention as the inflection point, and slope indicating the degree to which the proportion changes by size.

The values for the biological parameters, recruitment, selectivity patterns, and retention in the operating model are, in general, based on parameters estimated in the NOAA National Marine Fisheries Service (NMFS) assessment for Eastern Bering Sea (EBS) Tanner crab (*Chionocoetes bairdi*) (Stockhausen, 2015). The operating model generates a population with two sexes, 20 age classes, and 37 length groups (in increments of 5 mm). The generated population is fished by four fishery fleets: one directed fishery and three bycatch fisheries, and there is one survey.

Sex-specific values based on the NMFS assessment of Tanner crab (Stockhausen, 2015) are used in the operating model for natural mortality, growth (based on a von Bertalanffy growth curve; von Bertalanffy, 1934), length-weight relationship, length-at-50%-maturity, and the slope of length-at-maturity (Table 2.1). Time-varying selectivity, which is governed by a logistic curve (Equation A.1.9), is parameterized using 31 blocks (Table 2.2), resulting in 31 selectivity patterns (Table 2.3), which correspond to changes in the fleets and gear used in the fishery for Tanner crab and match the assumptions on which the current NMFS assessment (Stockhausen, 2015) is based. Fishery retention for the directed fishery is logistic (Equation A.1.11), with a mean size-at-50%-selection of 150 mm and a slope of 0.1. A time-series of effort is also specified for each of the four fishery fleets over the 45 years of the model (Table 2.4).

2.2.1.2 Estimation Method

Artificial data sets are generated for each of the 100 replicates of each operating model configuration (step 3 in Fig. 2.1). Catch and effort time series for the directed fishery and the three bycatch fleets are generated for years 1-45, and are based on the true spawning stock biomass from the operating model (Fig. 2.2). A time series of discarded animals (i.e. all females and immature males) for the directed fishery is also generated. An index of abundance time series, based on the survey, is generated for the full 45 years. Length-composition data for each fleet, including the survey, are also generated for the full time series. While the specific values generated for each fleet differ between replicates, the CV of the index and effort data, as well as the effective sample size (N) for the length-composition data from the catch, discard, and survey time series are all pre-specified in the operating model.

The initial values used in the estimation method for natural mortality and growth are sampled from a distribution around the true values in the operating model, while the length-weight intercept and slope, length-at-50%-maturity, and maturity slope are fixed at their true (i.e. operating model) values. All recruitment parameters except for unfished recruitment, $R_{0,}$ are fixed at their true values (see Appendix 1 for equation parametrizations). The selectivity blocks in the estimation method match those in the operating model. The inflection point and slope of the logistic selectivity curves are estimated by fleet and by sex in the estimation method. The mean size-at-50%-retention and slope in the directed fishery are estimated for males only (because retention of females is not allowed in this configuration). There are 93 parameters estimated in the estimation method.

2.2.2 Experimental Design

Several alternative configurations of the operating model were investigated, including varying the CV for the survey index of abundance, the CV for the effort for the fisheries, and the effective sample size (*N*) for lengths from catches, discards, and the survey by fleet and by year. The index of abundance and effort CV were fixed at 0.2 in the base operating model and in the estimation method (Section 2.2.1). Effective sample sizes for catch lengths were fixed at 400 in the base operating model, while the effective sample sizes for discard and survey lengths were set at 100 in that operating model.

The values for the effective sample sizes and the CV were changes to values larger and smaller than those of the base model (Table 2.5), resulting in 10 alternative configurations of the operating model. Uncertainty about the survey indices of abundance involved a high CV of 0.5, while the low uncertainty scenario involved a CV of 0.1. Uncertainty about fishery effort involved a high CV of 0.3 and a low CV of 0.1. High effective sample size for the catch length-composition data was examined with 500 samples, while smaller effective sample size for catch lengths was examined with 300 samples. High average discard and survey effective sample sizes for length-composition data were set to 150, and low effective sample sizes for discards and the survey length-compositions were set at 50. All other inputs remained the same throughout each simulation test.

Each model configuration consisted of 100 replicates, and the resulting percent relative error was used to examine the level of relative bias and precision: $\delta x = 100\frac{\hat{x}-x}{x}$, where x is the true value generated by the operating model, and $\hat{x}$ is the estimated value for each simulated replicate. The median of the percent relative error and 50% and 90% simulation intervals about the median percent relative error of spawning stock biomass (SSB) were used to quantify estimation performance for each model configuration.

The mean and median percent relative error was calculated across all years and all replicates. The results were also summarized by the mean, median and range for the distribution over years of the 5th, 25th, median, 75th and 95th percentiles over years of percent relative error.

2.3 Results

Mean percent relative error in SSB generally remained close to zero in most of the operating model configurations (Table 2.6; Fig. 2.3). The mean percent relative error for base model was -0.3% across all years and all 100 replicates, and the median relative error in the base model was -1.1% (Table 2.6).

The mean percent relative error was close to zero (within ±3% of the truth) across both years and replicates for all scenarios except the high survey CV case, where the mean relative error was +33%, and the median relative error was +19% (Table 2.6). The mean relative error for the high catch *N* case was just slightly positive (+0.02%), but the median relative error was slightly negative (-0.9%; Table 2.6). Low discard *N* resulted in a slight positive bias (mean relative error=1.2%; median relative error=0.1%; Table 2.6), unlike all other cases which resulted in no or slightly negative mean percent relative error across years and replicates (except the high biomass CV case).

While median percent relative error remained relatively close to zero in the base model, median relative error ranged between -15.8% for the lower 5th percentile (over years) and +17.2% for the upper 95th percentile over all years (Table 2.7). Mean relative error ranged from -16% for the lower 5th percentile to +19% for the upper 95th percentile across all years.

The mean percent relative error was close to zero for most scenarios. However it varied considerably across years. Notably, high uncertainty in the indices of abundance resulted in the greatest range of mean relative error values (-39.8% for the lower 5th percentile to +139.5% for the upper 95th percentile) of SSB over years (Fig. 2.3, Table 2.7). Conversely, low uncertainty in the index of abundance resulted in the narrowest range of mean relative error of SSB across replicates (-11.3% for the lower 5th percentile to +7.6% for upper 95th percentile; Table 2.7). All remaining model configurations led to a range of relative error estimates for SSB quite similar to those for the base model (Table 2.7).

2.4 Discussion

The estimation method was considered acceptable for use in a management context given the simulation study resulted in little bias and relatively high precision when estimating biomass. For example, mean relative error was less than 2.6% in absolute terms across all years and replicates, except when the survey index of abundance was highly uncertain. The sensitivity to uncertainty related to the index of abundance data provides some context for the conditions under which this assessment method might be considered appropriate. In a data-rich context, where relative abundance is measured with relatively high precision, the estimation method could be expected to perform well, such as for Tanner crab or Pacific cod in the Eastern Bering Sea. However, this assessment method was not tested for data-poor stocks and its use would not be recommended for populations with little fishery-independent information on abundance.

Fishing history scenarios can be grouped into four categories, where each category leads to different primary trends in biomass: one-way trip, no change, good contrast, and rebuild only (Magnusson and Hilborn, 2007). Here, the trends in the “true” spawning stock biomass are consistent with a ‘good contrast’ scenario, in which the population initially declines, but begins to rebound toward the end of the time series (Fig. 2.2). The one-way trip and good contrast scenarios provide more reliable data with which to estimate reference points (Magnusson and Hilborn, 2007). This is important because both the EBS Tanner crab (Chapter 3) and EBS Pacific cod (Chapter 4) populations exhibit similar patterns of good contrast in estimated spawning stock biomass over time, both from the estimation method developed here and from their respective NMFS assessments (Stockhausen, 2015; Thompson, 2015).

While simulation studies for size-based assessments, and to an even greater extent age-size-based assessments, are scarce in the literature, some of the results here show similarities to other studies. Punt (2003) found that even when estimates of error from the assessment are relatively large, they nevertheless underestimate the true extent of uncertainty, which is often due to assumptions made in the model. The index of abundance can be an important source of uncertainty in estimating biomass and can have major implications for the results of an assessment. If the CV of the index of abundance is not well-known, but is assumed to be lower than its true value, estimates of spawning stock biomass could be more imprecise or biased than indicated by the uncertainty in the assessment.

Additional sources of uncertainty relate to model misspecification, including the assumption that key relationships are known, such as natural mortality rates, growth, and maturity. Ignoring size-selective mortality can have much more consequential impacts on assessment results than mis-specifying whether selectivity is age- or size-based (Punt et al., 2017).

While the age-size-structured estimation model performed well in most model configurations, there are still many factors that could be examined to better understand model performance, and its implications for estimating population biomass relative to management reference points. Future work should examine the assumption that growth is known and is time-invariant, the implications of fixing versus estimating natural mortality for SSB estimates, and assuming the stock-recruit relationship is known. The growth curve and some of the stock-recruit relationship parameters are fixed to their true values in the estimation method, although this is clearly unrealistic as the true values are never known in actuality. Another important concern is the assumption that growth is time-invariant, which can be particularly problematic for northern-latitude species that are primarily controlled by size-structured dynamics (Frøysa et al., 2002). Similarly, natural mortality rates are not known exactly, and may also change through time due to fluctuations in environmental conditions or changes in predator-prey dynamics.

Overall, the simulation study shows that the estimation method is able to closely approximate estimates of spawning stock biomass to their true values, as quantified by the very low rates of mean percent relative error in the base model and alternative operating model configurations. The high precision and low bias of estimated spawning stock biomass suggests that the age-size-structured assessment model would be an acceptable alternative to current stock assessment models for management.

2.5 Tables

Table 2.1. Biological and recruitment parameters used in the operating model.

Parameter	Sex	
	Females	Males
Natural Mortality	0.29 yr^{-1}	0.29 yr^{-1}
Prior	0.29 yr^{-1}, 0.05 SD	0.29 yr^{-1}, 0.05 SD
Growth		
$L\infty$	120 mm	218.65 mm
K	0.1 yr^{-1}	0.1 yr^{-1}
CV	0.1	0.1
Length-weight		
Intercept	0.00064 kg/mm	0.00016 kg/mm
Slope	2.794	3.136
Maturity		
Length-at-50%-maturity	120 mm	120 mm
Slope	0.2 number mature/mm	0.2 number mature/mm
Recruitment		
Log(R_0)	9	
H	0.99	
σ_R	0.6	
Mean of gamma function (α)	0.1	
Variance of gamma function (β)	4	

Table 2.2. Selectivity blocks defined as specific ranges of years in the operating models and estimation method.

Fleet	Sex	
	Females (years)	Males (years)
Directed Fishery	1-45	1-23, 24, 25, 26, 27, 28, 29-37, 38, 39, 40, 41, 42, 43-45
Bycatch Fishery 1	1-45	1-29, 30-37, 38-45
Bycatch Fishery 2	1-45	1-29, 30-37, 38-45
Bycatch Fishery 3	1-45	1-19, 20-29, 30-45
Survey	1-14, 15-45	1-14, 15-45

Table 2.3. Selectivity inputs used in the operating model.

Fleet	Sex	Blocks	Intercept	Slope
Directed Fishery	Female	1	100	0.1
Directed Fishery	Male	1-12	100	0.1
Bycatch Fishery 1	Female	1	110	0.1
Bycatch Fishery 1	Male	1-3	110	0.1
Bycatch Fishery 2	Female	1	100	0.1
Bycatch Fishery 2	Male	1-3	100	0.1
Bycatch Fishery 3	Female	1	100	0.1
Bycatch Fishery 3	Male	1-3	100	0.1
Survey	Male	1-2	70	0.1
Survey	Female	1-2	70	0.1

Table 2.4. Time trajectories of fishing effort by fleet.

Year	Directed Fishery	Bycatch Fleet 1	Bycatch Fleet 2	Bycatch Fleet 3
1	0.1	0.05	0.41	1.59
2	0.2	0.41	0.05	1.31
3	0.3	1.00	1.00	1.00
4	0.4	1.59	1.95	0.69
5	0.5	1.95	1.59	0.41
6	0.6	1.95	0.41	0.19
7	0.7	1.59	0.05	0.05
8	0.8	1.00	1.00	0.00
9	0.9	0.41	1.95	0.05
10	1	0.05	1.59	0.19
11	1	0.05	0.41	0.41
12	1	0.41	0.05	0.69
13	1	1.00	1.00	1.00
14	1	1.59	1.95	1.31
15	1	1.95	1.59	1.59
16	1	1.95	0.41	1.81
17	1	1.59	0.05	1.95
18	1	1.00	1.00	2.00
19	1	0.41	1.95	1.95
20	1	0.05	1.59	1.81
21	1	0.05	0.41	1.59
22	1	0.41	0.05	1.31
23	1	1.00	1.00	1.00
24	1	1.59	1.95	0.69
25	1	1.95	1.59	0.41
26	1	1.95	0.41	0.19
27	1	1.59	0.05	0.05
28	1	1.00	1.00	0.00
29	1	0.41	1.95	0.05
30	1	0.05	1.59	0.19
31	1	0.05	0.41	0.41
32	1	0.41	0.05	0.69
33	1	1.00	1.00	1.00
34	1	1.59	1.95	1.31
35	1	1.95	1.59	1.59
36	1	1.95	0.41	1.81
37	1	1.59	0.05	1.95
38	1	1.00	1.00	2.00
39	1	0.41	1.95	1.95
40	1	0.05	1.59	1.81
41	1	0.05	0.41	1.59
42	1	0.41	0.05	1.31
43	1	1.00	1.00	1.00
44	1	1.59	1.95	0.69
45	1	1.95	1.59	0.41

Table 2.5. Base case, and low and high configurations of the operating method for effective sample size (*N*) for fishery, discard and survey length-compositions and the coefficient of variation (CV) for the index of abundance (biomass) and fishing effort (effort).

	Low	Base Model	High
Biomass CV	0.1	0.2	0.5
Effort CV	0.1	0.2	0.3
Catch lengths *N*	300	400	500
Discard lengths *N*	50	100	150
Survey lengths *N*	50	100	150

Table 2.6. Mean and median of percent relative error in spawning stock biomass estimates across all years and all 100 replicates for each model configuration.

Configuration	Mean	Median
Base Model	-0.3	-1.1
High Biomass CV	32.6	18.8
Low Biomass CV	-2.6	-2.8
High Effort CV	-0.3	-1.2
Low Effort CV	-0.3	-1.1
High Catch *N*	0.0	-0.9
Low Catch *N*	-0.3	-1.3
High Discard *N*	-0.7	-1.4
Low Discard *N*	1.2	0.1
High Survey *N*	0.0	-0.8
Low Survey *N*	-0.6	-1.5

Table 2.7. Summary statistics for percent relative error for spawning stock biomass. Results show the mean, median and range for the distribution over years of the 5th, 25th, median, 75th and 95th percentiles over simulations of percent relative error; i.e. -15.9 is the mean of the year-specific lower 5th percentile, where the lower 5th percentile is over 100 simulations.

	Quantile:	5%	25%	50%	75%	95%
Base Model	Mean	-15.9	-7.8	-1.1	6.1	18.8
	Median	-15.8	-7.8	-1.0	5.3	17.2
	Minimum	-22.6	-10.9	-4.5	3.0	10.6
	Maximum	-10.7	-4.9	1.3	14.1	33.4
High Biomass CV	Mean	-22.5	-2.9	19.1	52.5	139.5
	Median	-22.4	-3.1	19.1	52.1	134.1
	Minimum	-39.8	-8.0	9.9	32.1	82.6
	Maximum	-14.4	1.8	28.9	82.8	209.4
Low Biomass CV	Mean	-11.3	-6.9	-2.9	1.1	7.6
	Median	-11.3	-7.0	-2.8	0.9	7.7
	Minimum	-14.3	-9.9	-5.3	-1.9	3.6
	Maximum	-7.5	-4.2	-1.1	4.1	14.5
High Effort CV	Mean	-16.1	-7.9	-1.1	6.2	19.2
	Median	-16.4	-8.0	-1.0	5.3	17.5
	Minimum	-22.9	-11.3	-4.6	2.9	10.5
	Maximum	-11.1	-4.9	1.3	14.0	33.5
Low Effort CV	Mean	-16.0	-7.9	-1.1	6.1	18.7
	Median	-16.0	-8.0	-1.0	5.5	17.0
	Minimum	-22.6	-11.6	-4.5	2.9	10.4
	Maximum	-10.9	-4.8	1.6	13.9	33.6
High Catch *N*	Mean	-15.1	-7.5	-0.7	6.3	19.6
	Median	-15.4	-7.7	-0.8	5.5	18.3
	Minimum	-21.8	-10.4	-3.7	3.1	11.7
	Maximum	-10.2	-4.2	1.5	13.9	33.2
Low Catch *N*	Mean	-15.7	-7.9	-1.2	6.3	19.3
	Median	-15.5	-8.0	-1.2	5.4	17.8
	Minimum	-21.3	-10.7	-4.8	2.7	9.8
	Maximum	-11.5	-4.9	1.4	13.0	33.7
High Discard *N*	Mean	-16.0	-7.8	-1.2	5.8	18.0
	Median	-16.1	-7.8	-1.1	4.8	16.6
	Minimum	-22.3	-10.6	-4.1	2.9	10.8

	Maximum	-10.8	-5.4	1.4	13.0	29.6
Low Discard *N*	Mean	-15.5	-7.4	0.4	8.2	22.3
	Median	-15.2	-7.3	0.2	8.1	20.7
	Minimum	-20.4	-11.0	-3.3	4.7	12.6
	Maximum	-10.2	-4.0	2.7	16.8	35.5
High Survey *N*	Mean	-15.8	-7.6	-0.7	6.4	19.0
	Median	-15.7	-7.9	-0.6	5.7	17.6
	Minimum	-22.3	-10.5	-4.1	3.3	11.1
	Maximum	-11.0	-4.6	1.9	13.9	32.8
Low Survey *N*	Mean	-16.2	-8.1	-1.4	5.7	18.4
	Median	-16.0	-8.1	-1.2	5.2	16.9
	Minimum	-23.0	-11.3	-4.7	2.6	10.3
	Maximum	-10.7	-4.9	1.0	13.6	32.7

2.6 Figures

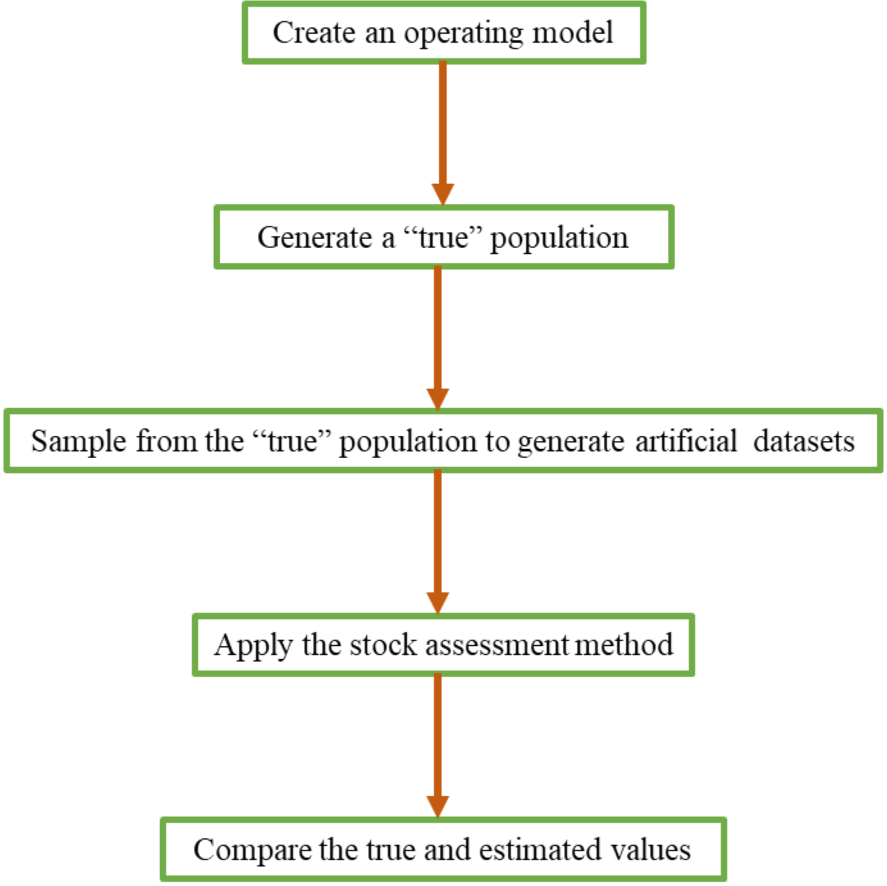


Figure 2.1. Flowchart of the process used to generate data and evaluate the performance of stock assessment methods. Modified from Punt (2003).

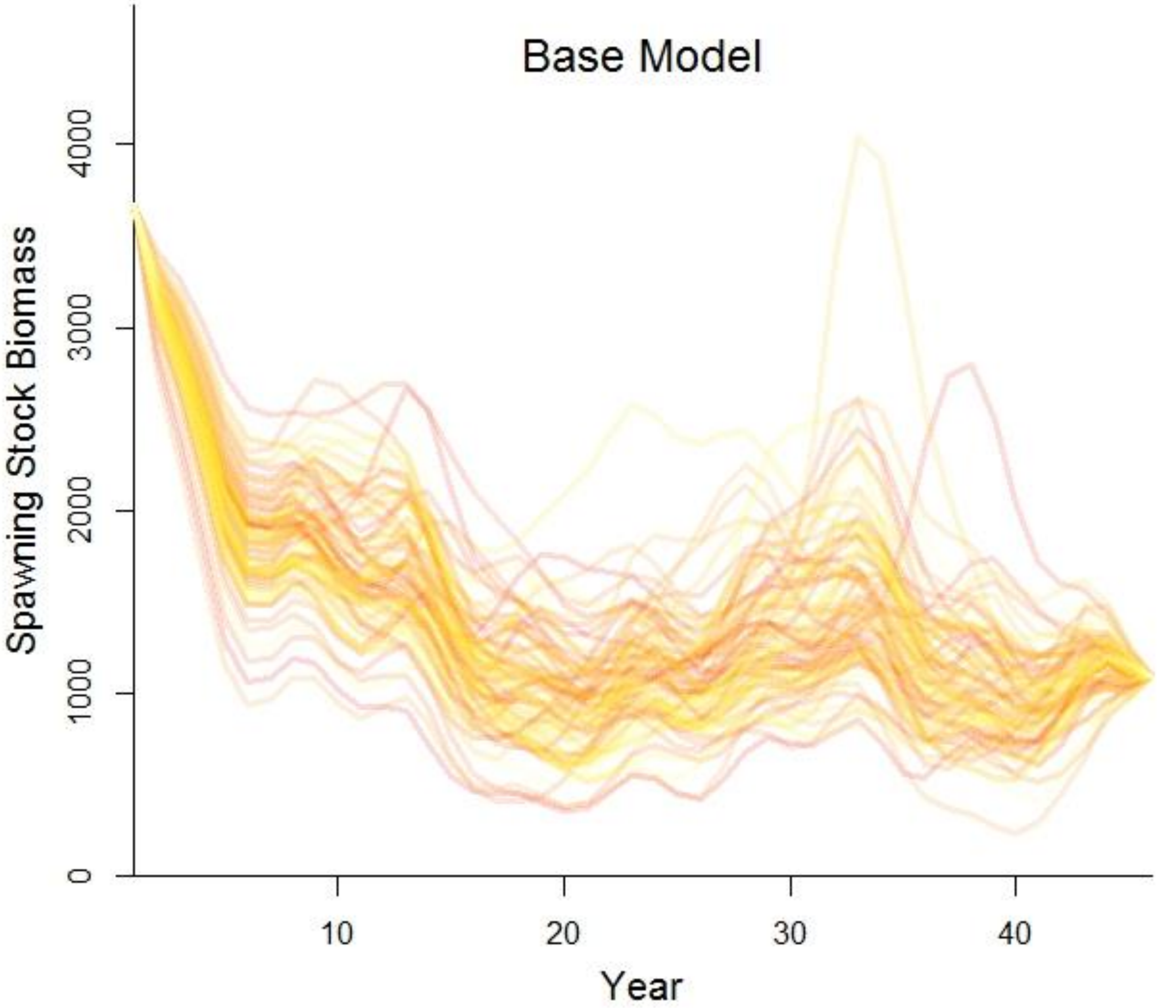


Figure 2.2. Projections of spawning stock biomass for the base operating model (for all 100 iterations).

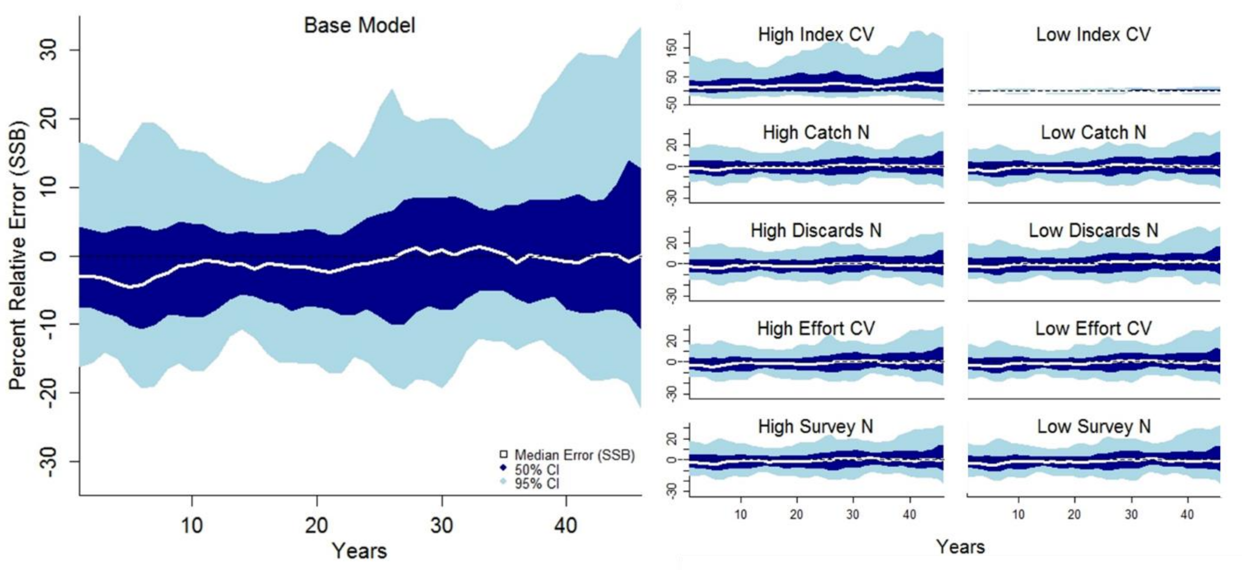


Figure 2.3. Percent relative error in spawning stock biomass (SSB) for the base case operating model (left panel) and for 10 operating model configurations related to uncertainty in biomass CV, effective sample size (*N*) for catch, *N* for discards, fishing effort CV, and *N* for survey length-composition samples (top to bottom of the right panel), with the thick white line depicting the median percent relative error over 100 replications and the 50% (dark blue) and 90% (light blue) percentiles of relative error. Specifications for each scenario are listed in Table 2.5.

Chapter 3: Eastern Bering Sea Tanner Crab

3.1 Introduction

3.1.1 Life History

Southern Tanner crab (*Chionocoetes bairdi*) are found across the North Pacific, from Oregon (Hosie and Gaumer, 1974), north through the Bering Sea (Somerton, 1981), and west across to Japan (Kon, 1996). Southern Tanner crab in the Eastern Bering Sea (EBS) of Alaska are assumed to be separate from Tanner crab in the Aleutian Islands and elsewhere, and management of EBS Tanner crab is further divided into east and west sectors at 166°W. However, evidence of stock structure demonstrating two distinct, non-mixing, and non-breeding populations is lacking (Stockhausen, 2015), and therefore in this assessment, EBS Tanner crab (hereafter, Tanner crab) are modelled as one population.

Tanner crab, like most crustaceans, molt as they grow. For both sexes, Tanner crab larger than 25 mm carapace width (CW) molt annually, followed by a terminal molt at maturity, beyond which no further growth occurs (Tamone et al., 2007). Females Tanner crab undergo terminal molt after their last instar phase (Donaldson and Adams, 1989). Male Tanner crab are considered to have undergone terminal molt when they have a chela to carapace ratio of 0.18 or greater, though this can occur at different sizes for individual crab (Tamone et al., 2007). Tanner crab caught in the survey are assigned a shell condition and a maturity state. While Tanner crab shell condition is still recorded, little correlation between post-molt age and degree of encrusting epifauna can be found (Nevisi et al, 1996), and shell condition is not used in the age-size-structured assessment, beyond assigning mature versus immature designations in the assessment. Growth parameters for male and female Tanner crab are based primarily on the findings from a growth study by Rugolo and Turnock (2012), and growth parameters in the age-size-structured model are selected to mimic growth from the assessment (Stockhausen, 2015).

In relation to maturity, Rugolo and Turnock (2012) estimated 50% maturity to be 68.8-74.6 mm CW for females and 91.1-104.4 mm CW for males. While research suggests that female reproductive output is a dome-shaped function of age(NMFS, 2004), spawning biomass in Alaskan crab assessments is measured using mature male biomass (MMB) because only males are retained in the fishery (NPFMC, 2007).

Very little is known about natural mortality (M) of Tanner crab. Somerton (1981) estimated mortality for individual cohorts, and obtained estimates of M ranging between 0.13 and 0.35 yr^{-1}, but concluded that natural mortality rates of 0.22-0.28 yr^{-1} were most representative of the population in models that used both fishery and survey data. In contrast, Rugolo and Turnock (2011) reasoned that Tanner crab would exhibit traits similar to those seen in snow crab (*C. opilio*) because both populations exhibit analogous population dynamics, life history characteristics, and are similar enough to hybridize in regions where the two species overlap (Karinen and Hoopes, 1971). Karinen and Hoopes (1971) used the approximately 20-year life span of snow crab as a proxy for Tanner crab longevity, and Hoenig's (1983) method to estimate the natural mortality rate at 0.23 yr^{-1}. This is the base value currently used in the assessment (Stockhausen, 2015) and is the starting value for estimating M in the age-size-structured model.

3.1.2 Fishing History

Fishing has been primarily dominated by the domestic fleet since foreign fishing for Tanner crab ended in 1980. Historically, Tanner crab landings were recorded in both Japanese (1965-1978) and Russian (1965-1971) tangle net fisheries (Stockhausen, 2015). Tanner crab were targeted domestically starting in the 1970s, peaking in the fall 1977 though spring 1978 season (Fig. 3.1). However, landings subsequently fell, and the fishery was closed for the 1985/86-1986/87 seasons. A similar pattern was repeated, with a peak harvest during the 1989/90-1991/92 seasons, followed by a fishery closure during the 1997/98-2004/05 seasons (Fig. 3.1). The fishery remained open from the end of the 2005/06 season through 2009/10, then closed from the end of the 2010/11 season through the 2012/13 season. The fishery was again reopened in fall 2013, and remained open until 2016/17, when it was most recently closed. During open seasons, the fishery has generally been able to fully catch the total allowable catch (TAC) in both the east (98% of TAC on average) and the west (75% of the TAC) regions (Stockhausen, 2015).

Discards occur in the Tanner crab fishery both due to minimum size limits and because the fishery is a male-only fishery. Minimum legal size limits for retained males have changed over time: they were 138 mm CW prior to March 2011, but since then are 122 mm CW in the east and 112 mm CW in the west (Alaska Board of Fisheries, 2011). Discard data for Tanner crab from the directed fishery are available from 1992/93-2015/16 for years when the fishery was open. Tanner crab are also caught incidentally in other crab pot fisheries, including those for snow crab and red king crab, and the groundfish fisheries. Bycatch has occurred since the groundfish and crab pot fisheries began, but data are only available starting in 1973 for the groundfish fishery and 1992 for the pot fisheries. All bycatch has to be discarded, but some of the discarded crab are assumed to survive. Discard survival is accounted for in the stock assessment by assigning a probability of survival for each discarded individual, assuming there is some information about survival based on the gear type used in each fleet.

3.1.3 Data for Assessment Purposes

The data available for Tanner crab are catches by fleet, including: 1) the retained catch of males in the directed fishery, 2) estimates of discarded males and females in the directed fishery, 3) estimates of bycatch from the fisheries for Bristol Bay red king crab, Eastern Bering Sea snow crab, and groundfish, 4) effort in the crab (pot) fisheries, 5) survey indices of numbers of Tanner crab by sex, and 6) size-composition data for males in the directed fishery, males and females in the bycatch fleets, and in the survey (Tables 3.1-3.2). Data sources and collection methods are described in Stockhausen (2015).

Retained catch in the directed fishery is available from 1965 (foreign fleets) through the present (domestic fleets) (Table 3.1). Total catch is recorded in both numbers and biomass, and effort is quantified as the number of potlifts (Stockhausen, 2015). Discard mortality is assessed assuming handling mortality rates of 32.1% for the crab fisheries and 80% for groundfish fisheries (Stockhausen, 2015). From 1992-2004, groundfish trawls accounted for the largest proportion of Tanner crab bycatch. Specifically, groundfish trawls caught an average of 75% of the total catch by the directed fishery between 1992/93 and 1996/97, and the directed fishery was closed between 1997/98 and 2004/2005 seasons. However, crab fisheries including the Red King crab (RKC) and Snow crab (SC) fisheries have produced the highest amount of Tanner crab bycatch since 2005/06, averaging 167% of the total catch by the directed fishery, for years that the directed fishery was open. Bycatch and discard totals are provided by observer sampling in the RKC, SC, and groundfish fleets. Observer sampling also provides catch-at-size data for Tanner crab bycatch and discards. Effort is available for the directed and crab bycatch fisheries.

Fishery-independent estimates of abundance are provided by the National Marine Fisheries Service (NMFS) trawl surveys. Sex-specific survey indices of abundance are estimated annually using survey size-compositions and weight-at-size regressions from Rugolo and Turnock (2012). The NMFS trawl survey also provides size-composition data for males and females, and the corresponding sample size for each survey. Tanner crab are aggregated into 37 length classes, in 5mm CW bins for assessment purposes.

3.2 Methods: Specifications for the Assessment

3.2.1 Base Model

Blocks are used in the Tanner crab assessment to allow for time-varying selectivity. The blocks are based on changes to the fleets, such as gear changes, and to fishing regulations, are independent of each other, and mimic the block structure for which the most recent assessment is based (Stockhausen, 2015). In the directed fishery, blocks extend from 1965-1991, and there are single-year blocks for each year from 1992-2012. The groundfish bycatch blocks are 1965-1986, 1987-1996, and 1997-2012. The blocks for the crab bycatch fisheries (snow and red king crab) are 1965-1996, 1997-2004, and 2005-2012. The survey blocks are 1965-1981, 1982-1987, and 1988-2012.

Natural mortality is estimated for both mature and immature male and female crab. The natural mortality values for immature males, mature females, and immature females are estimated as offsets from natural mortality for adult males. All initial and estimated parameter values are available in Appendix A.2.1.

Growth is parameterized based on expected crab growth, which occurs only during molting. The growth curve, based on size-increments, is a log-linear function with a gamma-distributed size-increment distribution (Table 3.3; Appendix A.1.4.1). Molt probability is a declining logistic function of length. Growth for each sex is based on growth increments at two pre-specified lengths, the variability in the growth increment (parameterized as a CV), and the probability of molting at the pre-specified lengths (Appendix A.2.1). The growth parameters are tuned to closely match trends in numbers-at-size found in the virgin state, based on the assessment conducted by Stockhausen (2015), which accounts for a terminal molt at maturity (Fig. 3.2). The length-weight intercept and slope are fixed in the model based on values from Stockhausen (2015).

Maturity is a logistic function of length, with the length-at-50%-maturity and the slope of the maturity-at-length relationship pre-specified (Appendix A.2.1). The mating time for mature biomass is set at 0.625 (i.e. 15 February), the time of year at which mating is conventionally assumed to occur for crab (NPFMC, 2007). The sex ratio at recruitment is assumed to be 1:1 between males and females.

Initial recruitment (R_0) is estimated. Other parameters related to recruitment (see Appendix A.1 for details), including steepness (h) and the extent of variation in recruitment (σ_R), are pre-specified (Stockhausen, 2015). The parameters determining the distribution of crab lengths at recruitment (α and β) are estimated (Appendix A.2.1). The last recruitment size class, or crab entering the juvenile phase at about 6 months of age, is approximately 10 mm CW (Donaldson, et al., 1981). The age-size assessment assumes juveniles recruit to the fishery at age 0, whereas the NMFS assessment assumes juveniles recruit at age 5 to the fishery.

The model makes a simplifying assumption that no fishing occurred before recorded catch data began in 1965. The maximum fishing mortality for Tanner crab is assumed to be 5 yr^{-1}. The hybrid method (Methot and Wetzel, 2013) is used to calculate fishing mortality.

Selectivity is estimated by block and sex (Appendix A.2.2). The size-selectivity curve is logistic for all the fishery fleets (directed and bycatch), except for bycatch of males in the snow crab fishery, which has been shown to be dome-shaped (Stockhausen, 2015), and is modelled here using a double-logistic function. Size-selectivity for the survey is logistic, with a fixed asymptote at 1. The size-at-50%-retention and the slope of the retention versus length relationship in the directed fishery are estimated, with initial values based on legal limits for Tanner crab (Appendix A.2.1). Males make up 100% of retained catch, since retention of females is illegal. Catchability is estimated for the survey, based on priors provided by the assessment author (Stockhausen, *pers. comm.*).

The overall objective function is a weighted sum of each data source's contribution to the likelihood function (Appendix A.1.7). Iterative weighting has not been performed for this model, although sensitivity to some weights is explored. Pre-specified weights for all data, by type, are used, with some additional weight assigned to the survey data for the base model (Appendix A.2.3). Additional weight was assigned to fitting the survey index to better fit to the fishery-independent index of abundance rather than the fishery-dependent data.

3.2.2 Sensitivity Analysis

A sensitivity analysis was conducted for the Tanner crab assessment. Uncertainty in the index came out as one of the factors that had the largest impact on estimation performance in the simulation evaluation (see Chapter 2), so the CVs for the survey index of abundance used in the base model were doubled and halved in two sensitivity analyses to evaluate the extent to which this uncertainty might impact the assessment results. The effect of the effective sample sizes (*N*) for the length-composition data was also considered in this sensitivity analysis, with *N* doubled and halved for the survey and separately for the fishery fleets. These tests resulted in six sensitivity analyses (Table 3.4).

3.3 Results

3.3.1 Parameter Estimation

In total, 159 parameters are estimated in the Tanner crab model, all except five of which fall within the assumed bounds. Natural mortality, *M*, for males is expected to be closer to 0.23 yr^{-1} (Stockhausen, 2015; Somerton, 1981). However, it is estimated at the pre-specified upper bound of 0.35 yr^{-1}, although this may not be unreasonable due to additional work by Somerton (1981) that concluded that the instantaneous *M* is 0.35 yr^{-1} for age five crab, or the first age group to be fully selected by the NMFS trawl survey gear. The variance of the proportion of the total number of age-0 animals settling to each size-class (the stock-recruit relationship β parameter) is initialized at 4, but is estimated at its lower bound of 1, which reduces the variance in the gamma function (Equation A.1.6). The remaining three parameters estimated at pre-specified boundaries are used to compute fleet-specific selectivity functions. The slope of the selectivity of females in the directed fishery is estimated at its upper bound of 0.4, although initialized at 0.14. However, the only data available to inform selectivity of females in the directed fishery are length-frequency data for small amounts of bycatch (Fig. 3.3). The double-logistic selectivity function for male crab caught as bycatch in the snow crab fishery estimated parameters 2 and 5 at their lower bounds of -2 and -10, respectively. Double-logistic selectivity is set for male Tanner crab to mimic the assessment (Stockhausen, 2015), and this result suggests that this relationship may require additional exploration or sensitivity testing.

3.3.2 Model Fit Diagnostics

The model tends to fit male discards within the assumed confidence interval (CI) of the data for about half of the years with reported discards (Fig. 3.3). The model over-estimates discards in years with very low male discards (e.g. 1996 and 2009) in an attempt to fit the other data points. In contrast, discards for 1994, 1995, 2006, and 2007 are all under-estimated. Female discards are underestimated prior to 1995, when discard rates in the directed fishery were reportedly higher. The model fits more closely to the latter years, which have very low discards. Discard rates are markedly over-estimated for female Tanner crab in the last two years of the model. However, the amount of discarded females is always low in absolute terms. The model is able to match the general pattern of effort for the directed fishery (Fig. 3.4), but the model predictions tend not to fluctuate as much as the actual data. The effort data for the red king and snow crab bycatch fleets are generally followed after the mid-1990s, but the model overestimates effort for the earlier years.

The model fits closely to the index data from the NMFS survey, although with some under-estimation (Fig. 3.5). The model under-estimates the index of abundance for both males and females during 1989-1992, 2005-2006, and for males in 2012-2014. The model is able to fit the size-composition data quite well for most fleets. The model over-predicts the frequency of male crab between about 45-75 mm CW in the survey, and underestimates the frequency of larger sizes between 85-135 mm, the size groups just before male Tanner crab reach legal fishing size (Fig. 3.6). The fits to the annual length-composition data are generally good in the directed fishery, and red king crab and snow crab bycatch fleets, with mixed results for the groundfish bycatch and survey (Appendix A.2.4). The directed, red king crab, and snow crab fleets generally have good fits to the data across all years, with the directed fishery only showing a relatively small over-estimation of the largest size classes in 1995-1996, 2009, and 2013-2014. The estimation of length-compositions in the survey for both males and females was reasonable, although fits are often better when the trends in the length frequencies were similar to the previous year. The groundfish bycatch fits were generally poor, alternating between over- and under-estimation of length frequencies, often with a peak in the smallest, largest, or both smallest and largest size classes that did not fit well to the data.

The selectivity patterns for the Tanner crab fisheries are primarily logistic, with the exception being dome-shaped selectivity for bycatch of males in the snow crab fishery. The directed fishery exhibits steeply increasing selectivity around the minimum legal size (Fig. 3.7). The bycatch fleets show some variation in selectivity over time, based on the pre-assigned blocks for each fleet (see Section 3.2.1 for block details). Tanner crab are nearly fully-selected by the time they reach 45 mm CW in the survey.

3.3.3 Comparisons with Previous Assessments

Tanner crab model-estimated recruitment is for age-0 animals. However, the official NMFS assessment considers recruitment to the stock to be at approximately age-5 (Stockhausen, 2015). The values for recruitment from the two assessments are compared (Fig. 3.8), comparison is between two different age classes of Tanner crab. It is interesting to note that the patterns of recruitment are closely matched since 1990, with peaks and valleys in the official assessment occurring just after similar peaks and valleys in the age-size-structured assessment.

Mature male biomass (MMB), or the predicted spawning stock biomass of males, is estimated slightly larger in the age-size-structured assessment than in the official assessment in many years (Fig. 3.9). However, both estimates follow approximately the same pattern through time, with the age-size-structured assessment having less extreme peaks and valleys than the official assessment. The estimates from the age-size-structured model and the assessment model are highly uncertain prior to the start of the survey in 1974. Following the addition of the survey data into the model (in 1975), the differences between the two estimates range from 400 t to 54,500 t (or 0-22%) on an annual basis, with an average difference between the two of about 17,600 t (or 8.5%) (Fig. 3.9).

3.3.4 Sensitivity Analysis

Results from the sensitivity analyses indicate that the Tanner crab assessment is relatively stable, with changes to the inputs having little impact on the final results. Estimates of directed fishery discards remained mostly the same as for the base model, except when the effective sample size (*N*) for the fishery length-compositions was increased, which resulted in a slight decrease in estimated discards of males (Fig. 3.10). Effort in the directed fishery, and red king crab and snow crab fisheries changed little from the base model for each sensitivity test (Fig. 3.11).

The model estimates of the index of abundance for 1980 to 1995 were affected by survey CV and effective sample size (*N*; Fig. 3.12). High survey CV and *N* led to the model not capturing as much of the peaks and valleys in the data. Lower survey CV, or more certainty about the index of abundance, as well as lower survey *N*, or lower weight on the length-composition data, generally led to better fits to the index data. Changing the survey CV or effective sample size for the length compositions had little effect on the fit to length frequency data for the fishery fleets (Fig. 3.13). However, changing the survey *N* (in either direction) led to slightly poorer fits to the data for male and female survey length frequencies. The sensitivity analyses did not lead to markedly changes to the estimates of selectivity (Fig. 3.14).

Patterns of recruitment and mature male biomass (MMB) from the sensitivity analyses were generally very similar to those from the base model. For recruitment, two particular years stand out: 1978 and 1986. In 1978, the low survey CV and *N* sensitivity configurations estimate a high recruitment while all other configurations estimated a continued decline from the mid-1970s through the early 1980s (Fig. 3.15). The high survey CV and *N* sensitivity configurations estimate a sharp decline in recruitment for 1986, while all other configurations estimate a local peak in recruitment for that year. Patterns in MMB for the sensitivity analyses remain very similar to those for the base model (Fig. 3.16). The sensitivity configuration leading to the greatest differences from the base model was ‘high survey CV’, which resulted in a MMB estimate for 2016 that was 7,000 t larger than the base model estimate. High survey *N* also slightly over-estimated MMB compared to the base model, while low survey CV and *N* resulted in lesser under-estimates of MMB compared to the base model. The fishery *N* sensitivity configuration led to essentially no difference from the base model after the first year of data (1965).

The estimated overfishing limit (OFL) for the directed fishery for each sensitivity configuration was within 5,000 t of the base model estimate of 29,404 t (Table 3.5; Fig. 3.17). The largest difference in OFL estimates from the base model (5,000 t) occurred when the survey CVs were increased. Larger *N* for the survey length-composition data resulted in a 2,800 t increase in the OFL. Reducing the survey CV and *N* resulted in lower OFL estimates compared to the base model by 4,800 t and 3,100 t, respectively. Changing the *N* for the fishery fleets resulted in less than a 1,000 t change in the OFL for the directed fishery.

3.4 Discussion

Estimates of model outputs for Tanner crab are similar to those found in the most recent assessment (Stockhausen, 2015). Both assessments exhibit similar predictions of biomass trends following the start of the survey data in 1974. However, the differences between the two assessments, varying between 400-54,500 t of MMB, may impact management reference points. This is a fishery with a historical maximum catch of 35,500 t (in 1977), and a historical average catch of 8,000 t (1965-2015).

In terms of importance to management, the OFL is a key metric when applying the acceptable biological catch (ABC) control rule, and to providing an upper limit for the total allowable catch (TAC). Overall, the OFL does not change much in each sensitivity configuration compared with the base model. The estimated OFL decreased 5,000 t when more weight was given to information provided by the survey, which could have an impact on catch limits. When the fishery-independent indices of abundance were assumed to be less certain, OFL estimates increased.

Females cannot be landed in Alaskan crab fisheries, so mature male biomass (MMB) is used as a metric for the reproductive population. However, male biomass may not be the best indicator for population-level reproduction, and using data from females may provide a more direct estimate of reproduction on an annual basis. An Egg Production Index (EPI) was developed by Rugolo and Turnock (2010) from females caught in the bottom trawl survey. An interesting result of this study is that estimates of male and female biomass increased after 2005, but proportional increases in egg production were not observed (Rugolo and Turnock, 2010). Additional concerns for estimating reproduction that are not addressed in the age-size-structured model include the possibility that there are climate-driven impacts on reproduction and the fact that older females have been observed to senesce (Stockhausen, 2015). A future study could examine the impacts of regime shifts on reproductive success by using the available block structure to mimic regime shifts in the environment.

Tanner crab experience strong size-selective fishing, similar to many other crustaceans (e.g. Carver et al., 2005), and fishery-driven evolution toward smaller male crabs is a concern for management (Garcia et al., 2012; Fenberg and Roy, 2007; Conover and Munch, 2002). Tanner crab undergo a terminal molt at maturity, adding an additional complication in assessing the population biomass and establishing management reference points. Stockhausen (2015) states that length-at-50%-maturity is 91-104 mm, but the minimum catch size (until 2011) was 138 mm. This means that some of the mature males never enter the legal fishery, and could result in additional reproductive success, eventually leading to more small males in the population. In terms of management, this may complicate efforts to establish OFL based on an $F_{35\%}$ target rate of fishing mortality because total mature biomass is not equivalent to legally catchable mature male biomass and may result in over-harvest of larger size males. However, it does mean that there are mature males that are never retained and can contribute to perpetuating the species. An excess of small, breeding males means managers must be aware of the potential for fishery-induced genetic selectivity towards smaller male crab, as fisheries based on harvesting the largest fish have been shown to decrease in individual size over evolutionary time-scales (Conover and Munch, 2002). Alternatively, a model such as that proposed by Hilborn and Minte-Vera (2008) could be used to test hypotheses regarding the rate of change in growth patterns over time due to size-selective fishing mortality. Information regarding the proportion of mature males over the legal size limit and whether that proportion changes through time is critical to provide to management so that total allowable catch (TAC) limits can be set based on harvestable biomass.

Though terminal molt at maturity is an issue for modelling and management, it is also an opportunity for a tagging study. Mature Tanner crab cease to molt after reaching maturity, so adults in that life-stage would not shed tags. A tagging study could provide more definitive information about crab lifespan and movement after maturing. Tagging studies also provide information for growth. An update to the 1981 tagging study for Tanner crab (Donaldson et al., 1981) could provide new information on growth over time. The Donaldson et al. (1981) study captured Tanner crab around their time of molt and kept them in captivity to measure pre- and post-molt size for multiple size classes of crab, and documented the average size of crab post-terminal molt. An update might provide information as to whether growth rates are changing over time or under different environmental conditions. A sensitivity configuration that could be explored in a future iteration of this assessment is the impact of time-varying growth on biomass estimates.

Overall, the Tanner crab age-size-structured model resulted in biomass estimates similar to the current assessment model. Results remained fairly consistent throughout each sensitivity test, indicating that the current model specifications are relatively stable, although fits to data may improve slightly if the fishery-independent index of abundance is given more weight. It is likely that biomass estimates would be more accurate and precise with additional information about trends in growth and reproduction over time.

3.5 Tables

Table 3.1. Specifications for data by type and fleet, and the years for which each data source is available for Tanner crab.

Fleet	Data Type	Years
Directed Fishery	Catch biomass, Discard biomass, Effort, Size-composition	1965-2012, 1992-2009, 1968-2009, 1980-2009
Groundfish Bycatch	Bycatch biomass, Size-composition	1992-2013, 1973-2012
Red King Crab Bycatch	Bycatch biomass, Effort, Size-composition	1992-2012, 1992-2012, 1992-2012
Snow Crab Bycatch	Bycatch biomass, Effort, Size-composition	1992-2012, 1992-2012, 1992-2012
Survey	Abundance, Size-composition	1974-2012, 1975-2012

Table 3.2. Data available for the Tanner crab assessment by fleet and type. The designation “na” indicates that there are no data available, or that the data type does not exist for the specified fleet. * Tanner crab effort data are available for the red king and snow crab fleets, but not the groundfish fleet.

Fleet	Directed	Bycatch	Survey
Number of Fleets:	1	3	2
Catch	■	■	na
Index	░	░	■
Discard	■	░	na
Effort	■	RKC, SC*	na
Size-Frequency	■	■	■
Age-Frequency	░	░	░

Table 3.3. Overview of the options used for parameter estimation in the Tanner crab assessment.

Initial parameter values are available in Appendix A.2.1.

Process	**Setting**
Maturity	Logistic
Growth curve: size increment	Log-linear
Growth curve: molt probability	Declining logistic
Growth curve: size-increment distribution	Gamma
Selectivity: sex 2: estimate or offset from sex 1	Estimate parameters for sex 2 (females)
Length Selectivity: Fishery Fleets	Logistic
Length Selectivity: Male Snow Crab Bycatch	Double-logistic
Length Selectivity: Survey	Logistic, asymptotic at 1
Catchability: Survey	Estimate, accounting for the prior: 0.95 (Stockhausen, 2015)

Table 3.4. Specifications for each sensitivity configuration.

Sensitivity:	Test:	Change:
Survey CV	High	Doubled initial values
Survey CV	Low	Halved initial values
Effective Sample Size: Survey Length-Composition	High	Doubled initial values
Effective Sample Size: Survey Length-Composition	Low	Halved initial values
Effective Sample Size: Fishery Length-Composition	High	Doubled initial values
Effective Sample Size: Fishery Length-Composition	Low	Halved initial values

Table 3.5. Final OFL estimates for the base model and each sensitivity configuration for the directed fishery and bycatch fleets (in thousands of tons of Tanner crab).

Fleet:	Base	Survey CV (H)	Survey CV (L)	Effective Sample Size: Survey Length-Composition (H)	Effective Sample Size: Survey Length-Composition (L)	Effective Sample Size: Fishery Length-Composition (H)	Effective Sample Size: Fishery Length-Composition (L)
Directed Fishery (t)	29,400	34,600	24,500	32,200	26,300	30,100	28,600
Groundfish Bycatch (t)	0.1	0.1	0.2	0.1	0.2	0.1	0.1
RKC Bycatch (t)	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Snow Crab Bycatch (t)	1.5	1.5	1.6	1.5	1.6	1.5	1.5

3.6 Figures

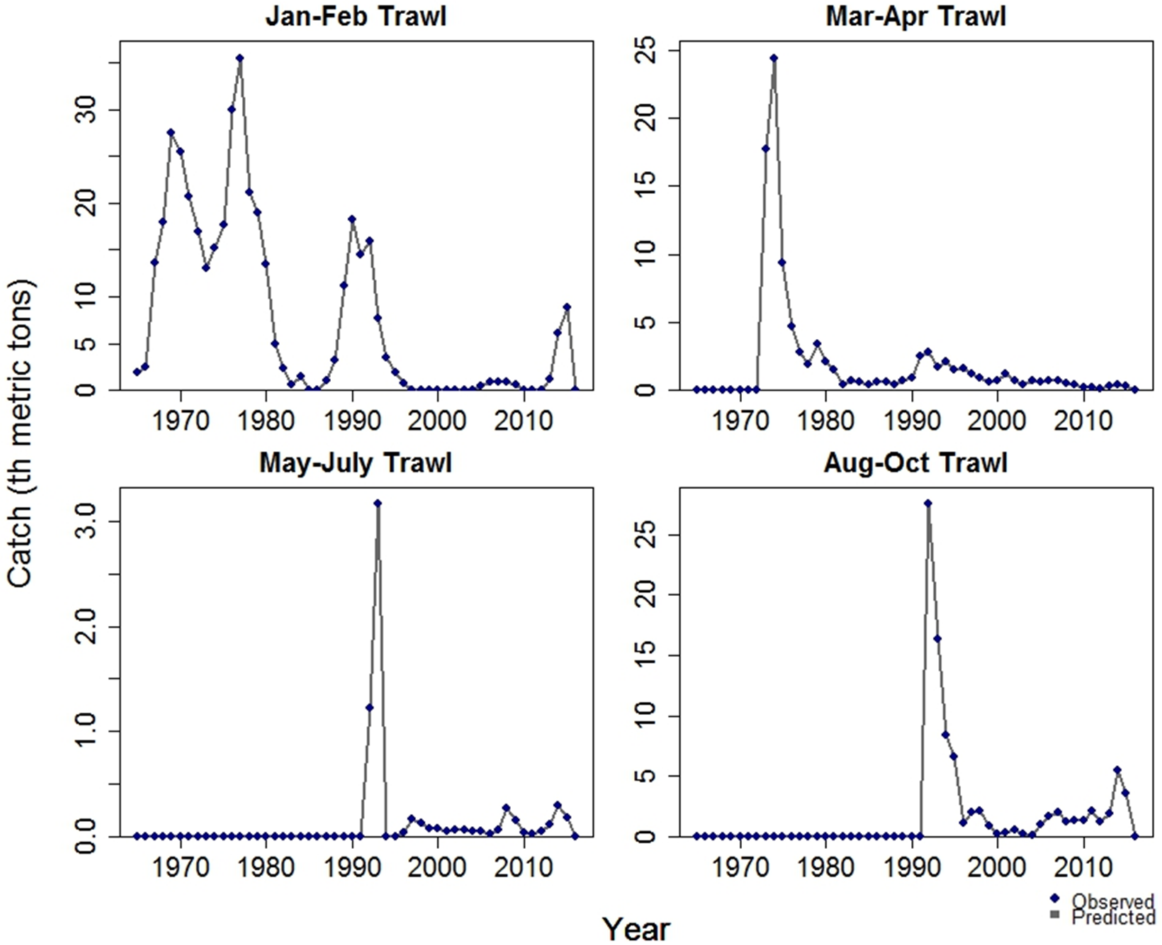


Figure 3.1. Tanner crab catch (thousands of metric tons) in the directed fishery, and recorded bycatch in the groundfish, red king crab, and snow crab fisheries from 1965-2017.

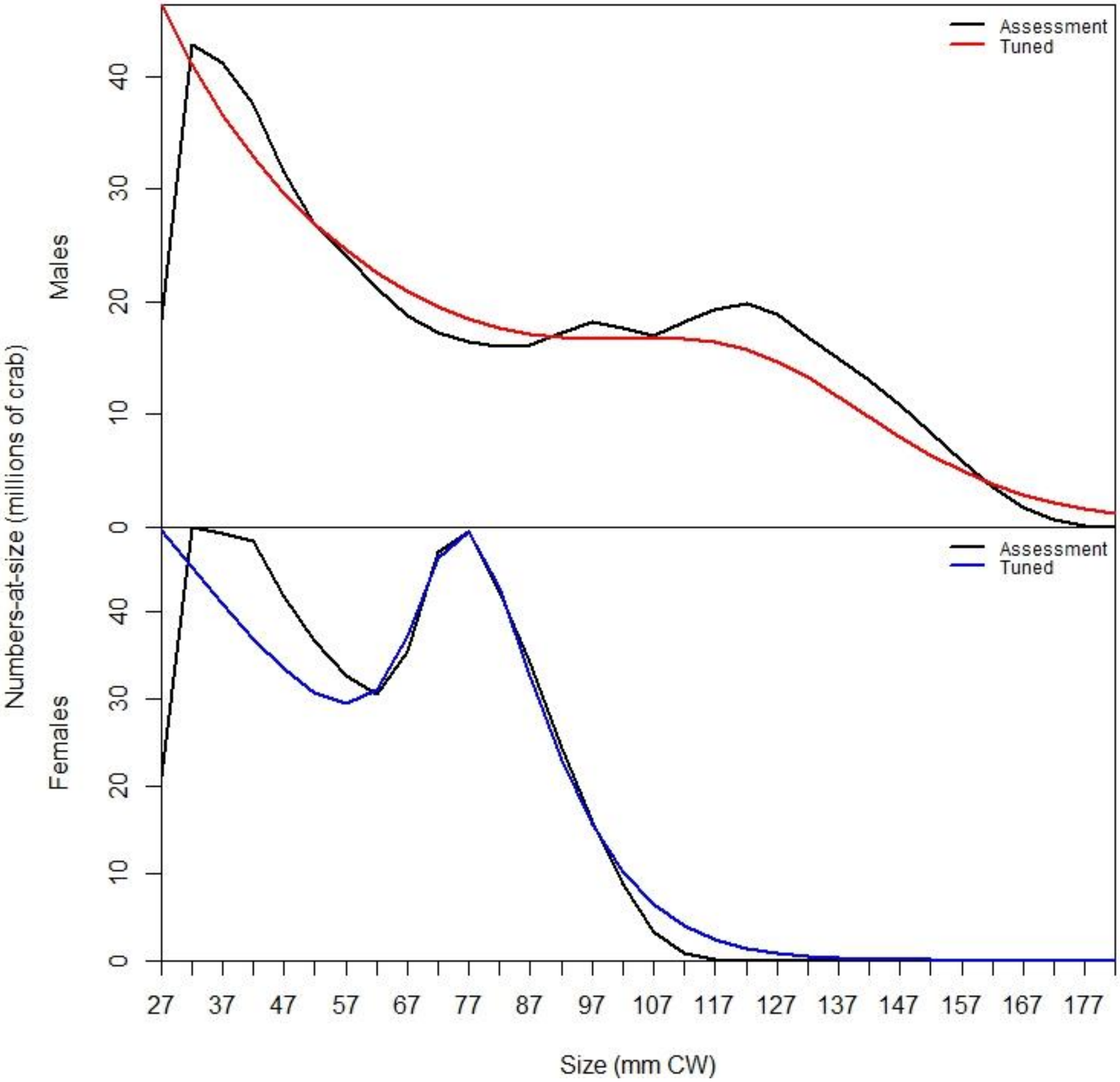


Figure 3.2. Growth in the age-size-structured assessment was tuned to match numbers-at-size in the unfished state, based on the most recent assessment (Stockhausen, 2015). Numbers-at-size for males (red line) and females (blue line) are tuned to the assessment values (black lines) for males (top panel) and females (bottom panel).

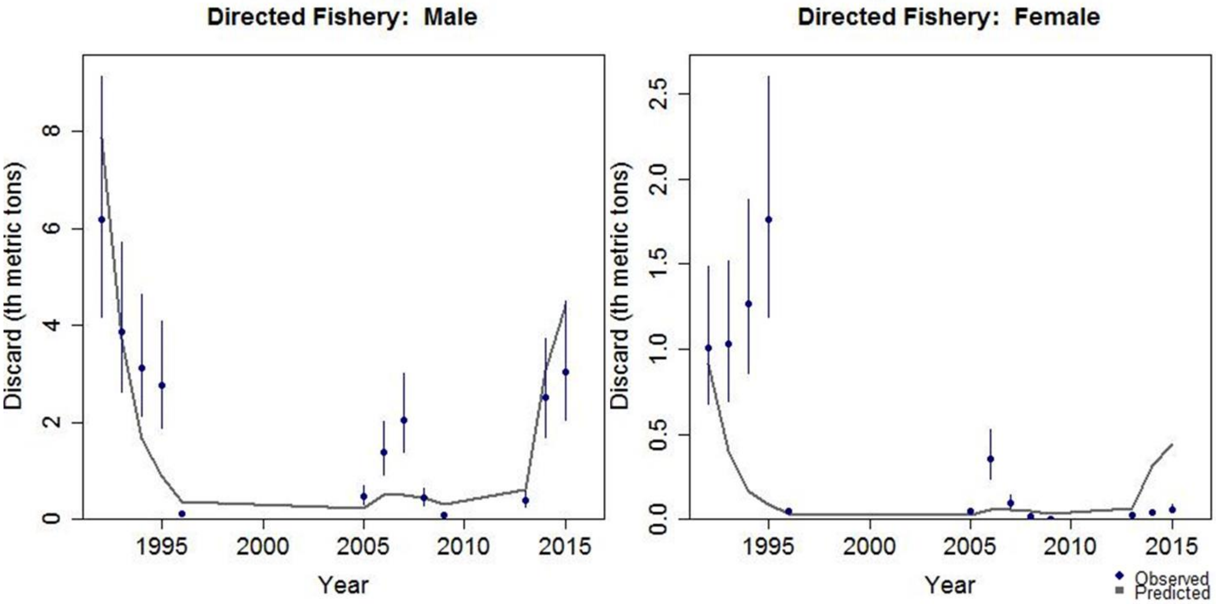


Figure 3.3. Observed (points) and model-predicted (lines) values of discards (in thousands of metric tons) in the male and female Tanner crab directed fishery.

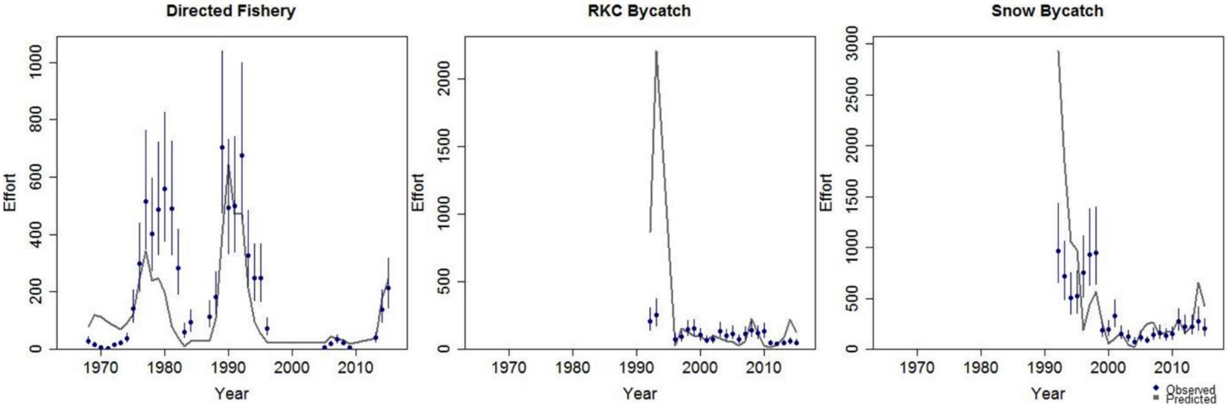


Figure 3.4. Observed (points) and model-predicted (lines) effort for the directed fishery for Tanner crab, and Bristol Bay red king crab and snow crab bycatch fisheries.

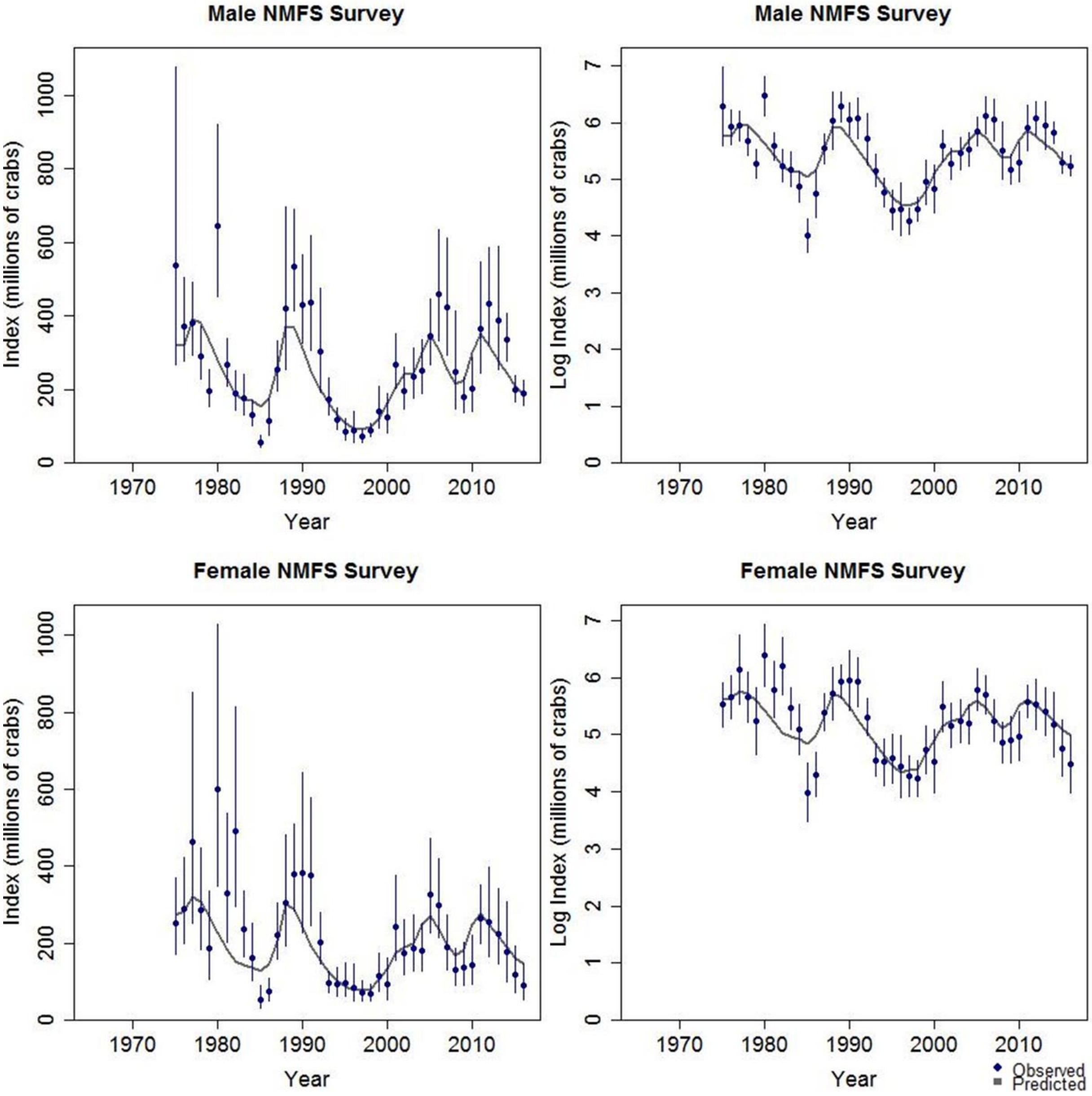


Figure 3.5. Index and log index plots (in millions of crabs) with observed (points) and model-predicted (lines) abundance estimates for the NMFS survey for Tanner crab.

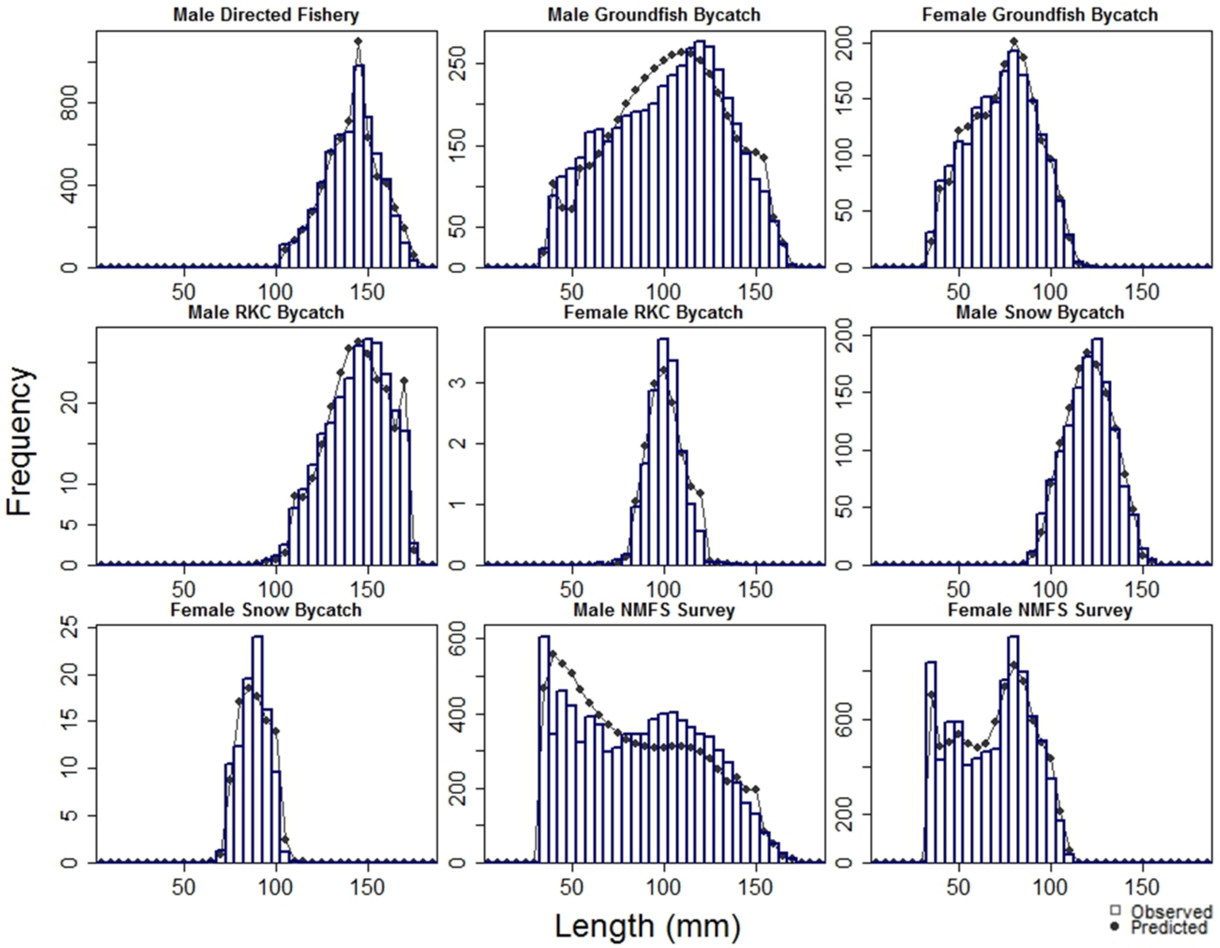


Figure 3.6. Observed length-composition data (bars) and predicted frequencies (points) for male and female Tanner crab from the survey and bycatch fleets, and male data for the directed fishery.

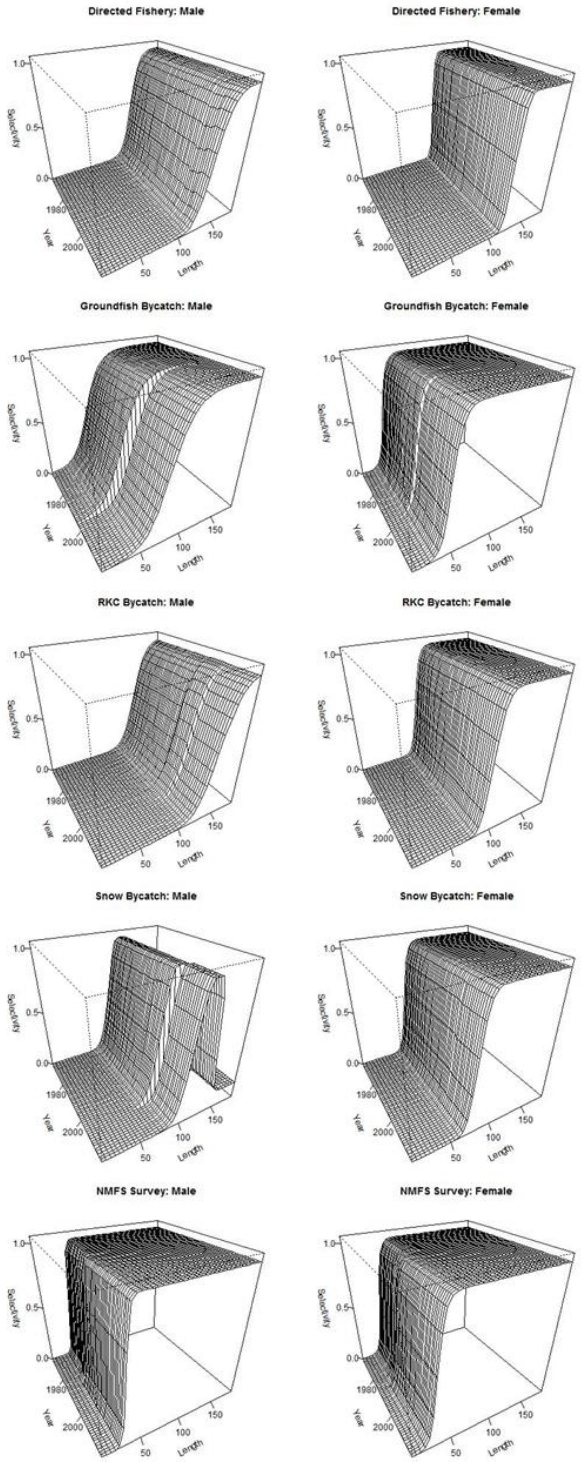


Figure 3.7. Estimated selectivity patterns for Tanner crab, by sex and each fleet. Each panel depicts the level of selectivity at each length, with one line for each year.

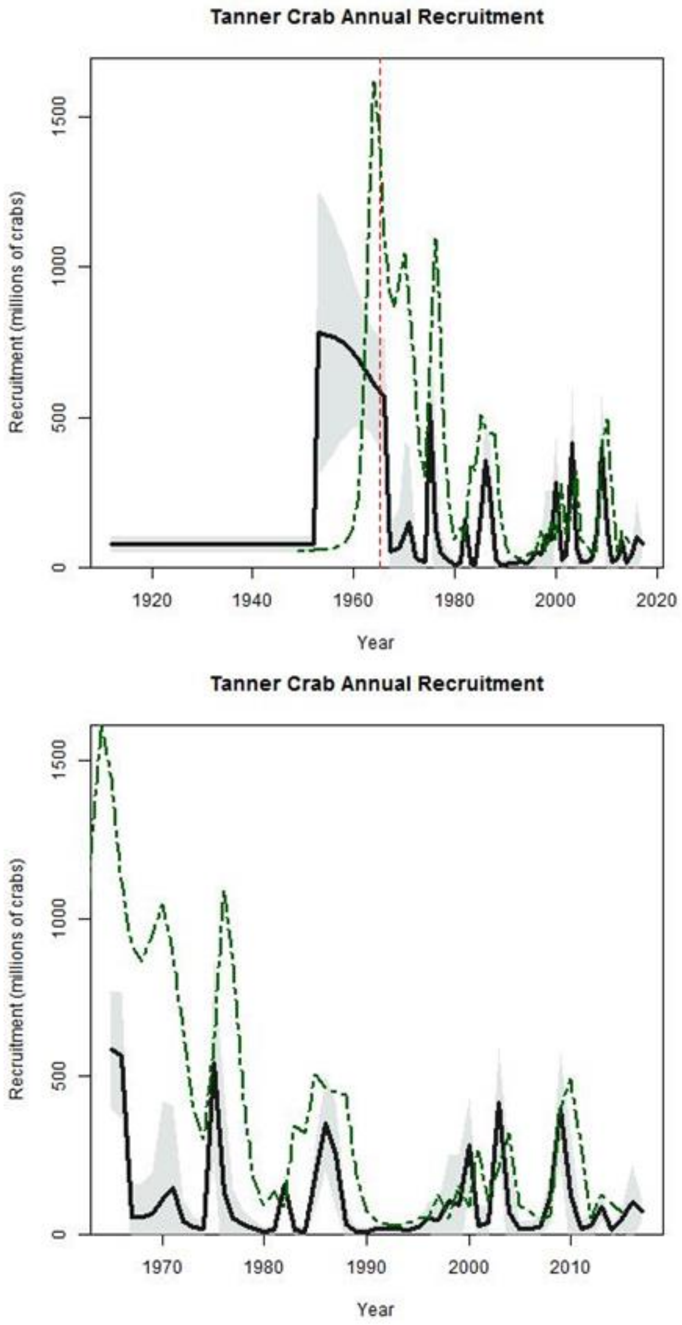
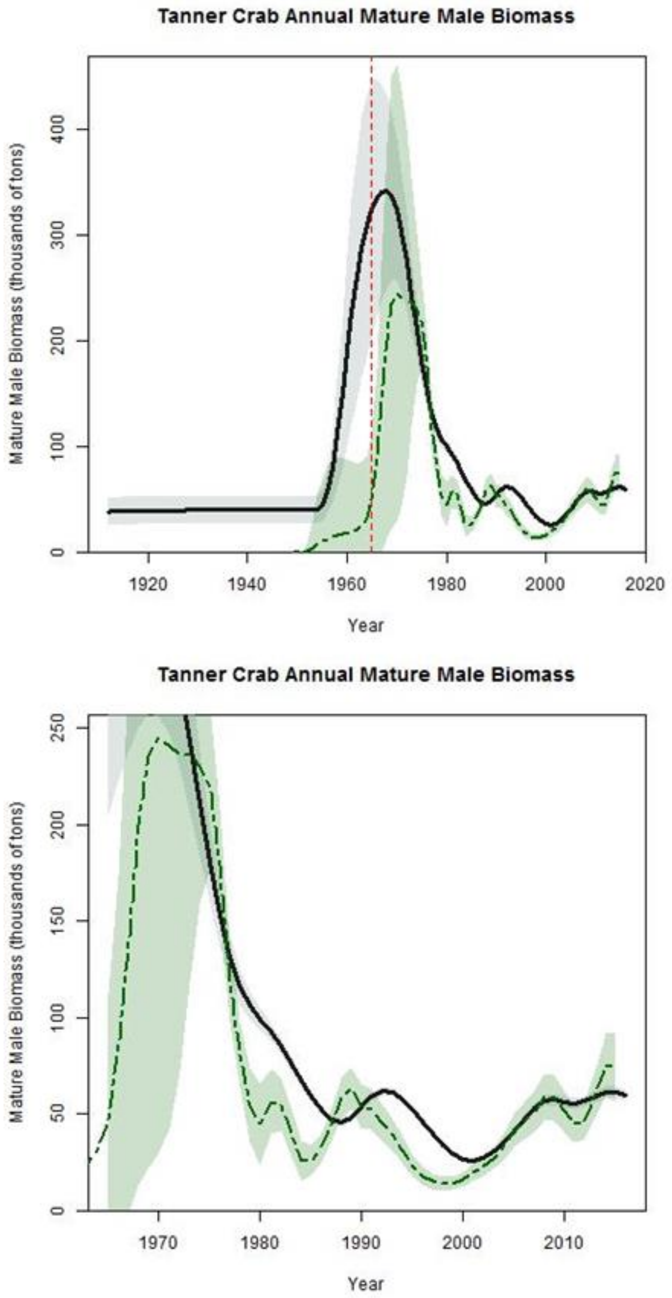


Figure 3.8. Model-predicted Tanner crab recruitment (black solid line) and asymptotic 95% confidence intervals for age-0 crab. The top panel shows predicted recruitment across all years of the assessment (starting in 1912), with the first year of data in 1965, indicated by the dotted red line. The bottom panel shows recruitment from the first year of data (1965). *Current assessment values for recruitment are shown (green dotted line), but are for age-5 crab (Stockhausen, 2015).



Tanner Crab Annual Mature Male Biomass

Tanner Crab Annual Mature Male Biomass

Figure 3.9. Model-predicted Tanner crab mature male biomass (MMB) (solid black line) across all years of the assessment (1953-2017) (top), including a 95% confidence interval (grey), with the first year of data (1965) indicated by the dotted red line, and in the bottom panel from the first year of data (1965-2017). The dotted green line indicates the assessment-estimated MMB (Stockhausen, 2015).

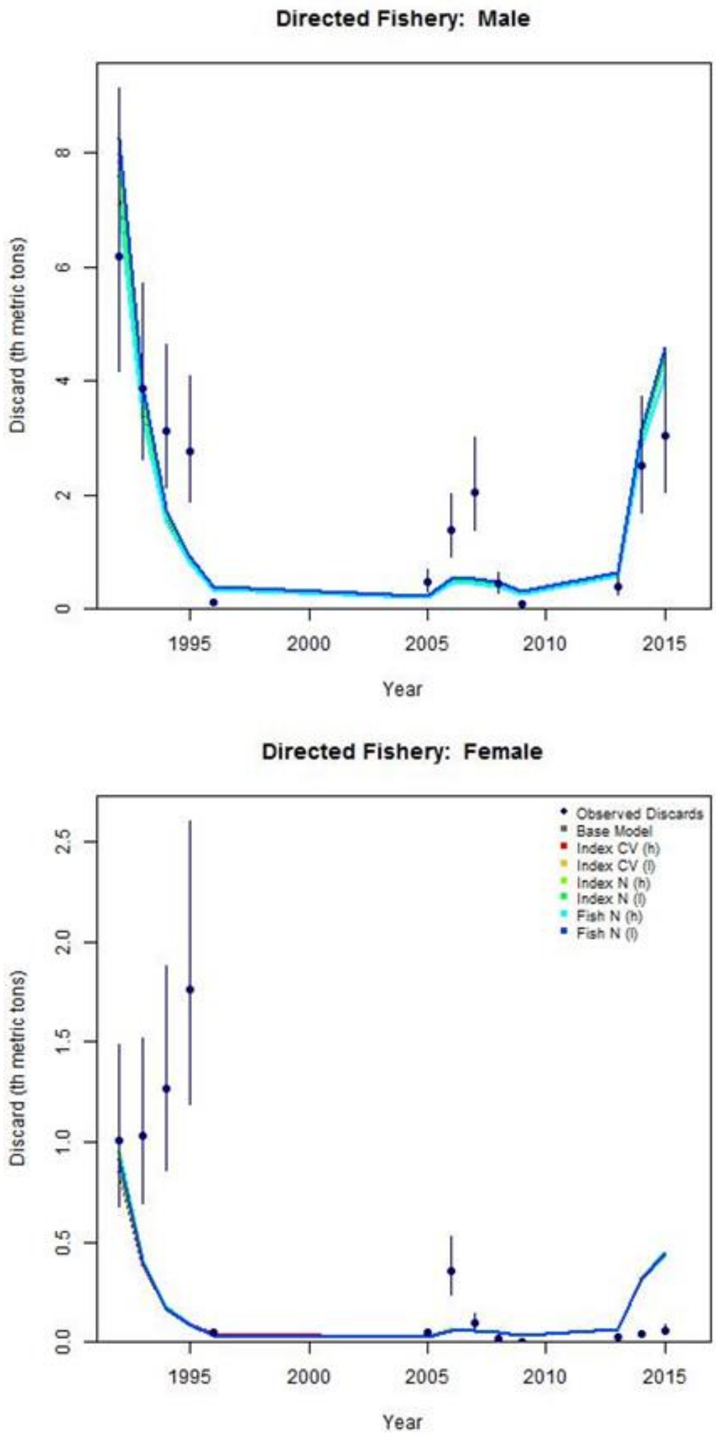


Figure 3.10. Observed (points) and model-predicted (lines) values of discards in the male and female Tanner crab directed fishery (in thousands of metric tons) for the base model and each sensitivity configuration.

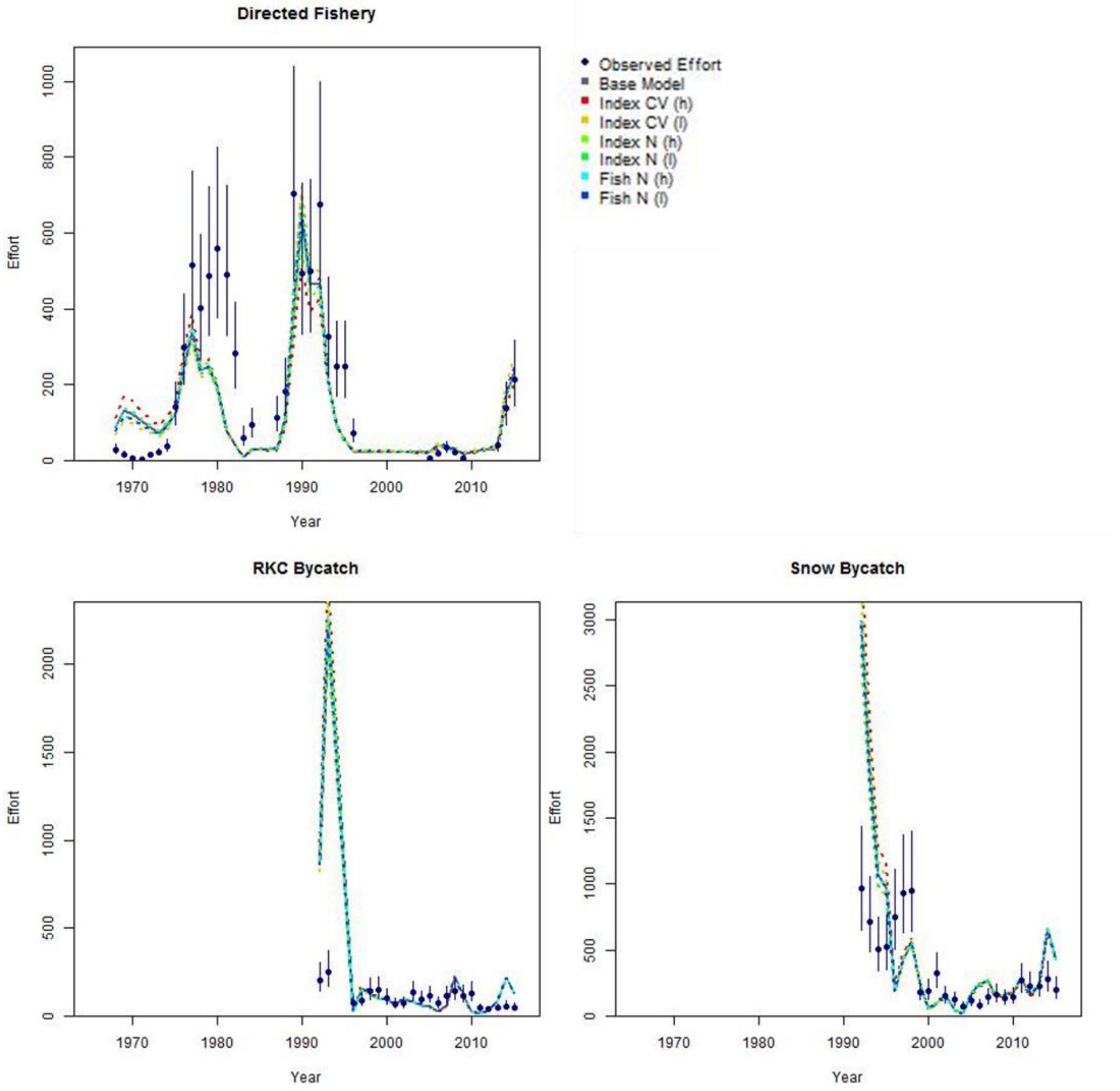


Figure 3.11. Observed (points) and model-predicted (lines) effort for the directed fishery for Tanner crab, and Bristol Bay red king crab and snow crab bycatch fisheries, including the base model and each sensitivity configuration.

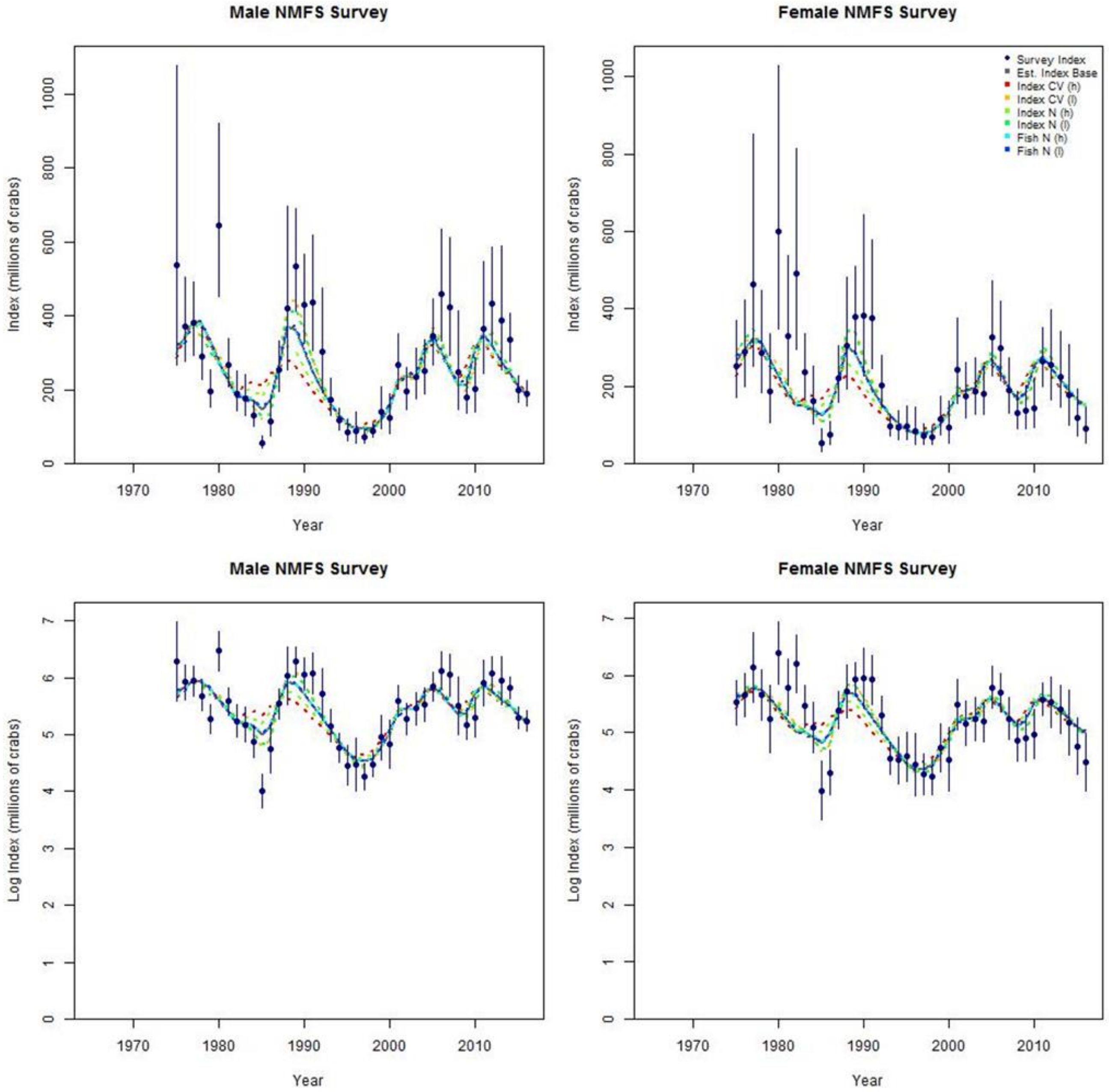


Figure 3.12. Tanner crab index and log index plots (in millions of crab) with observed (points) and model-predicted (lines) abundance estimates for the NMFS survey, including the base model and each sensitivity configuration.

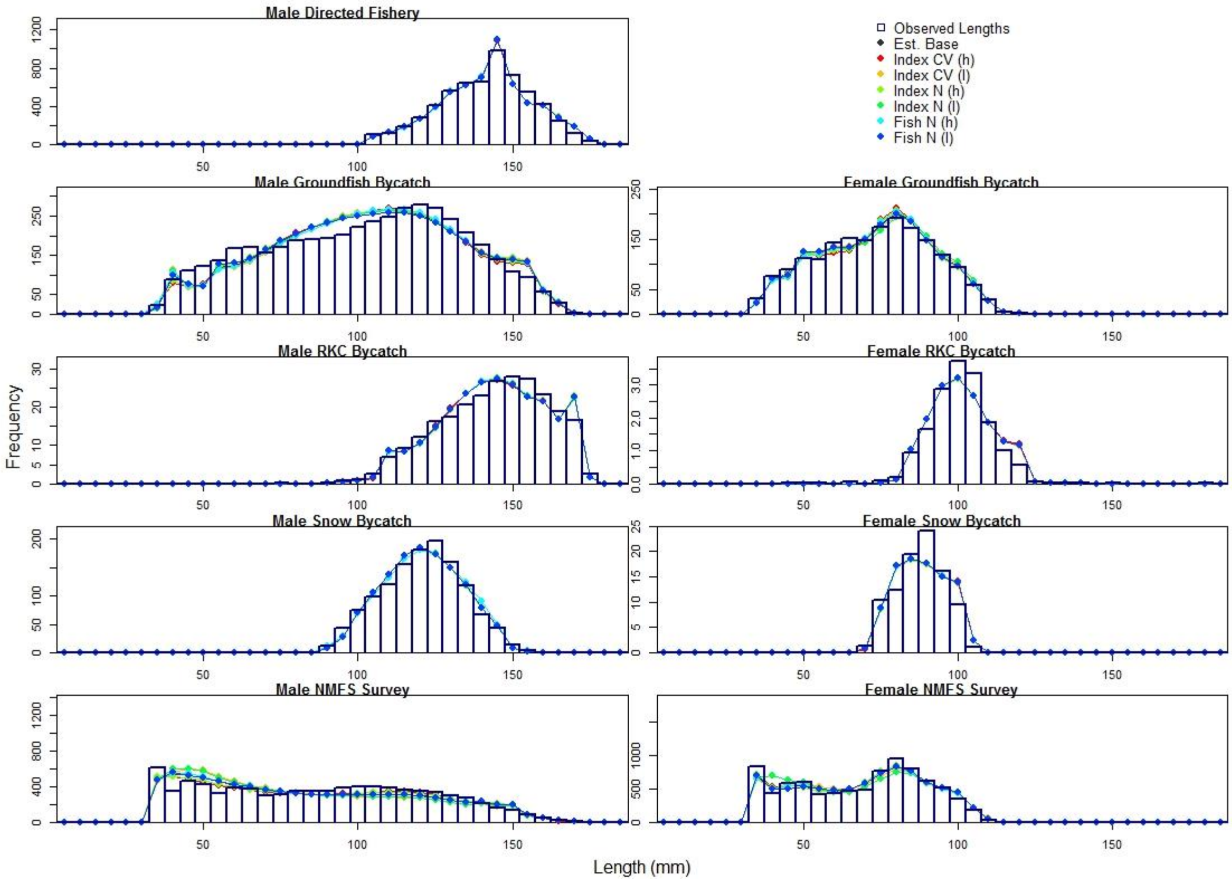


Figure 3.13. Observed length-composition data (bars) and predicted frequencies (points) for male and female Tanner crab from the survey and bycatch fleets, and male data for the directed fishery in the base model and each of the sensitivity configuration.

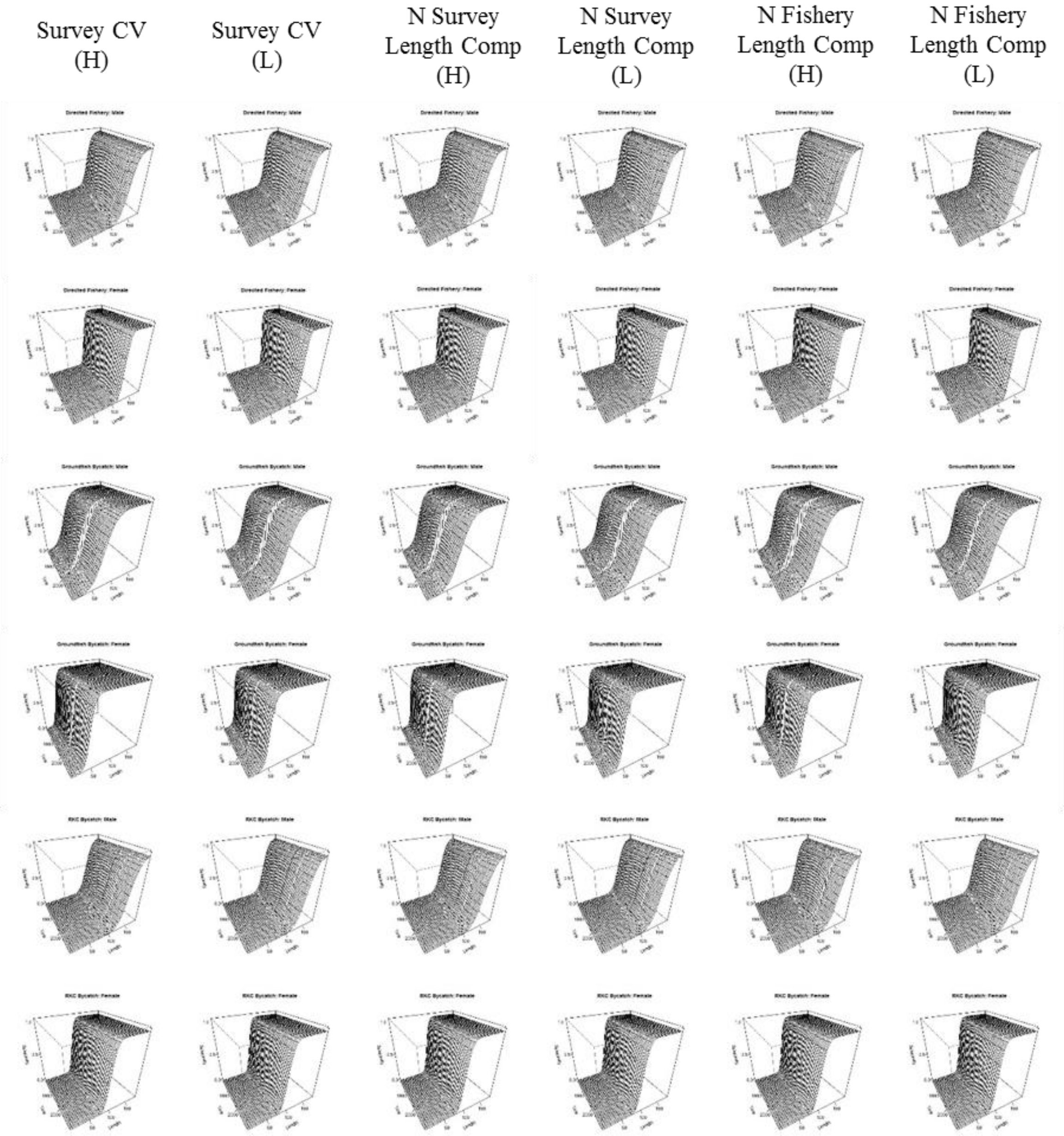


Figure 3.14. Estimated selectivity patterns for Tanner crab, by sex and by fleet for the base model and each sensitivity configuration.

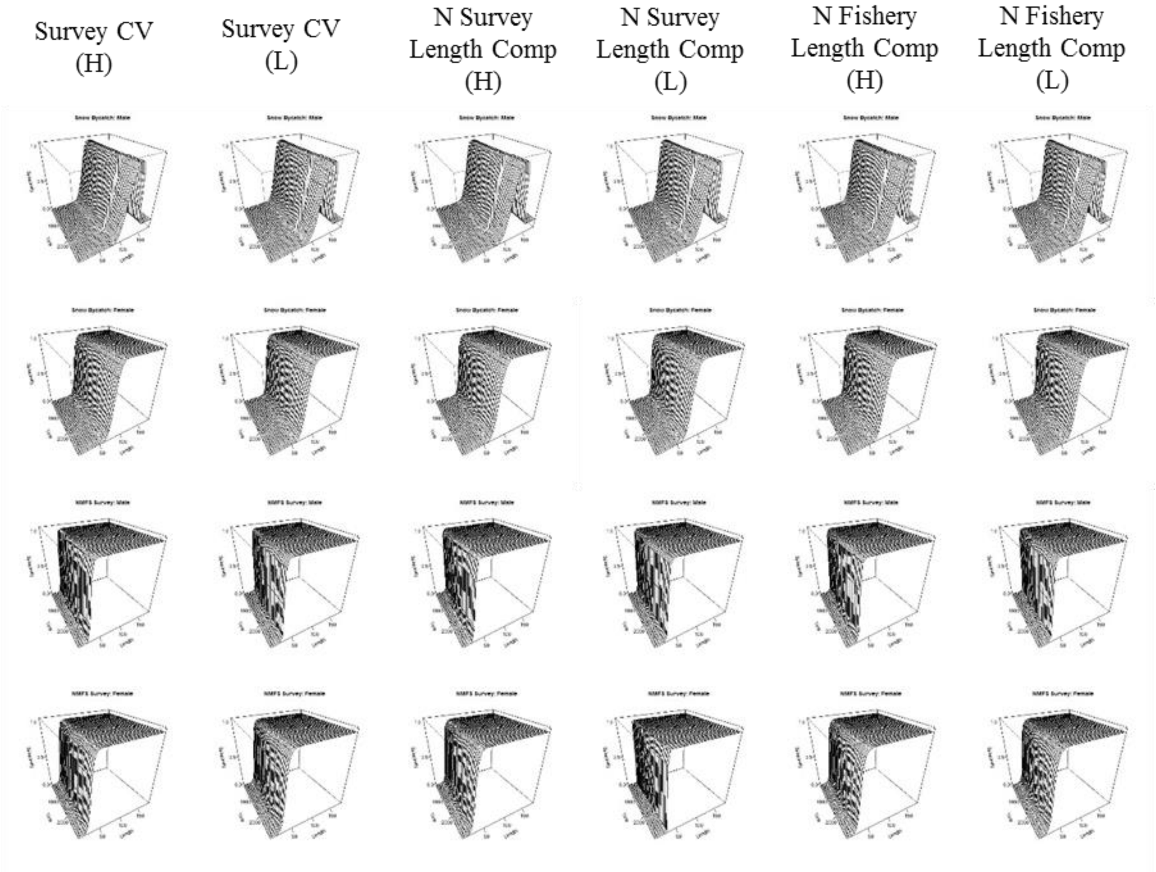


Figure 3.14. Continued.

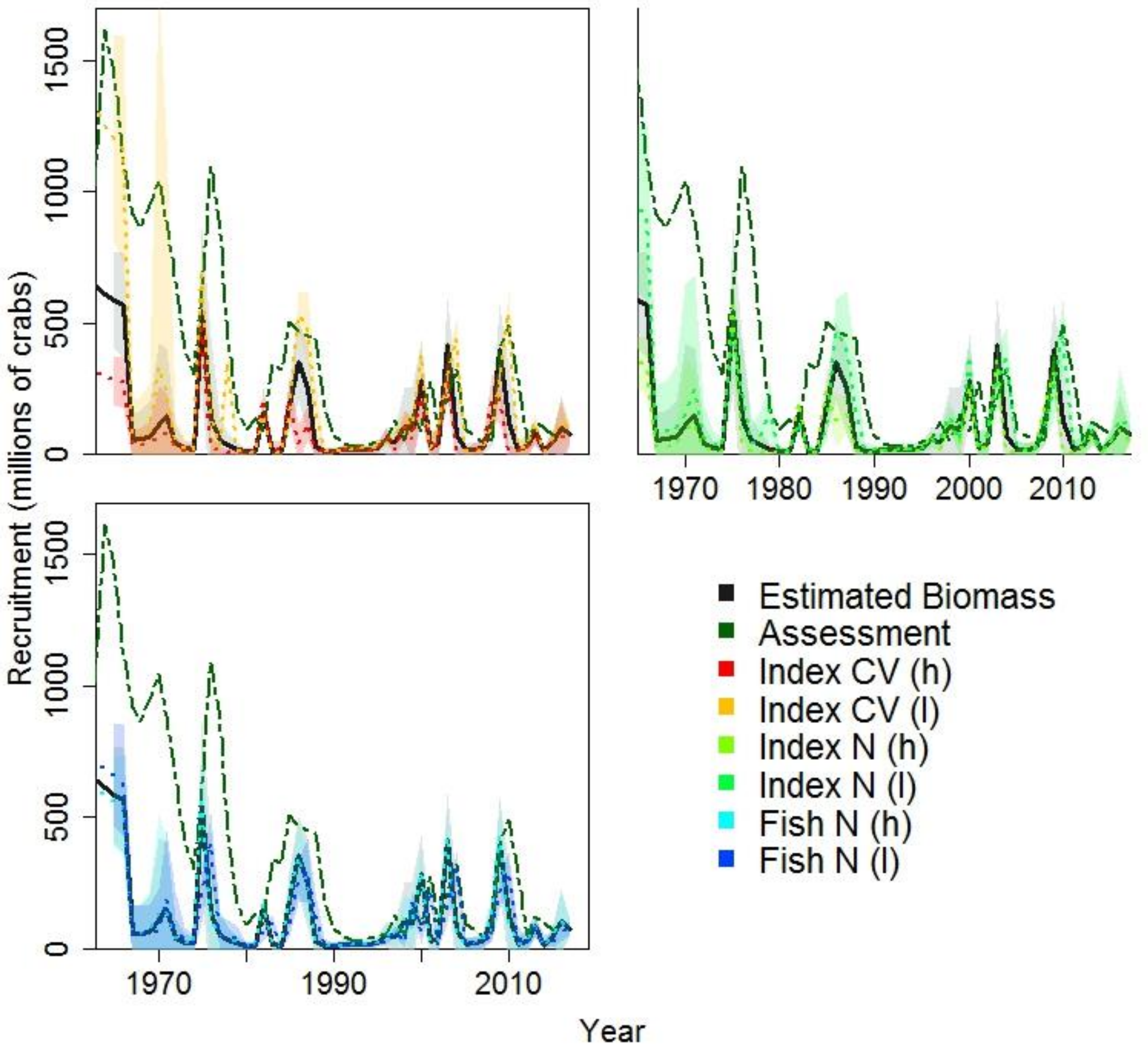


Figure 3.15. Model-predicted Tanner crab recruitment (colored solid lines) for the base model and each sensitivity configuration, and asymptotic 95% confidence intervals for age-0 crab. The panels show recruitment from 1965-2017. *Current assessment values for recruitment are shown (dark green dotted line), but are for age-5 crab (Stockhausen, 2015).

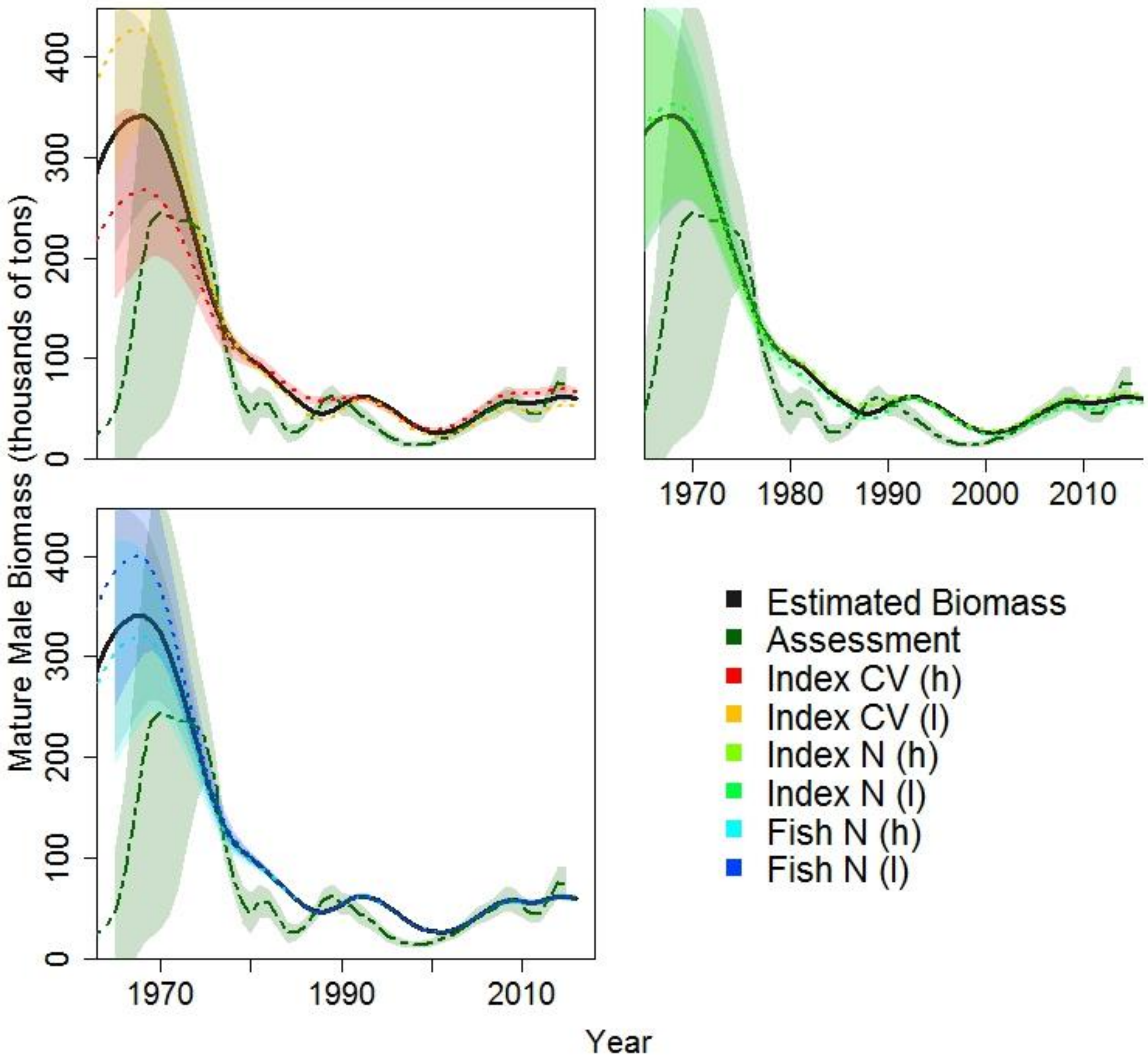


Figure 3.16. Model-predicted mature male biomass (MMB) for Tanner crab (colored solid lines) from the first year of data for the assessment (1965-2017) (left), including 95% confidence intervals. The dark green dotted line indicates the assessment-estimated MMB (Stockhausen, 2015).

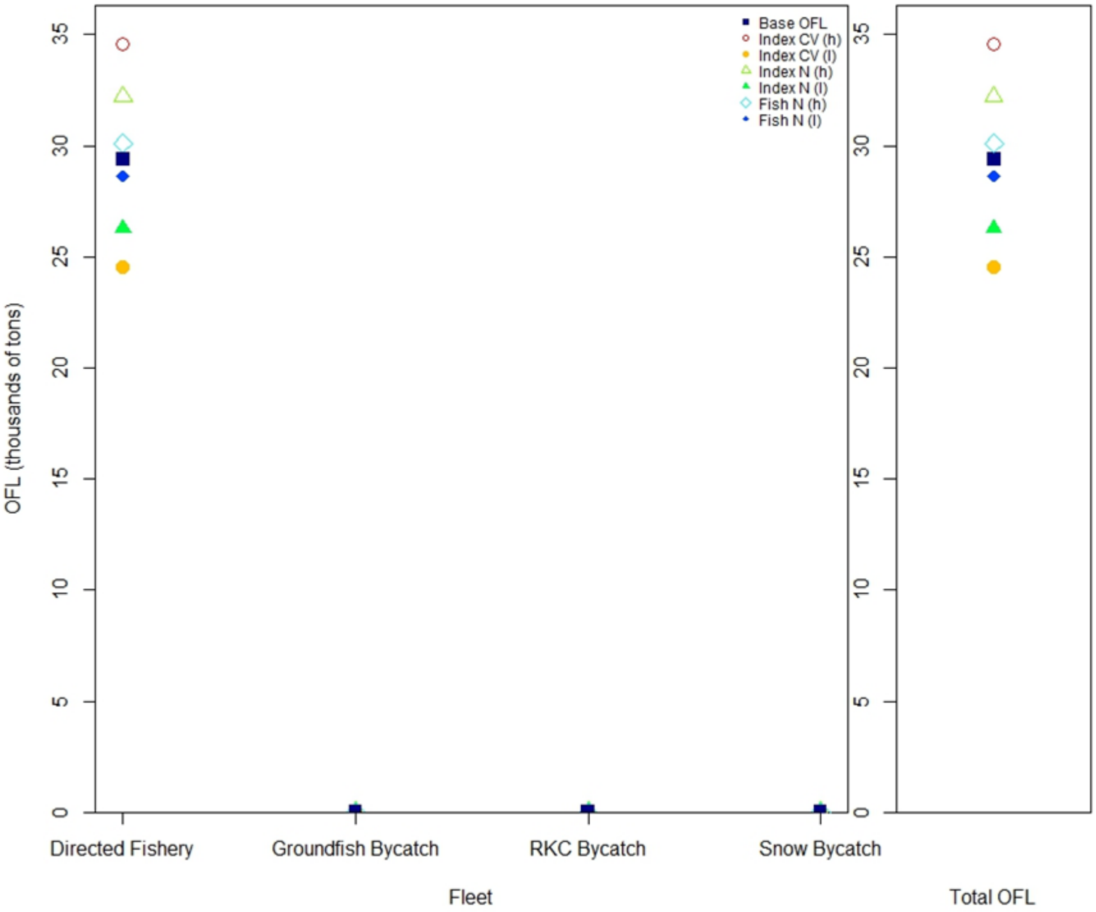


Figure 3.17. Estimated OFL for the base model and each sensitivity configuration in the directed Tanner crab fishery and for each of the bycatch fleets, as well as the total OFL for all Tanner crab.

Chapter 4: Pacific cod in the Eastern Bering Sea and Aleutian Islands

4.1 Introduction

4.1.1 Life History

Pacific cod (*Gadus macrocephalus*) ranges across the North Pacific Ocean, from Southern California to the Arctic Circle, and is distributed across the Eastern Bering Sea (EBS) and Aleutian Islands (AI) to about 500 m (Thompson, 2015). Recent research indicates a discrete stock occurs in the EBS and AI regions (Canino et al., 2005, 2010; Cunningham et al., 2009; Spies, 2012). The EBS-AI population was managed as one stock from 1977-2013, but was split into two separate harvest areas starting in 2014 (Thompson, 2015). The EBS-AI Pacific cod stocks (hereafter "Pacific cod") are treated as one population in the age-size-structured assessment. Spawning takes place near the ocean floor, and eggs sink after fertilization (Thompson, 2015). Larvae are epipelagic, and undergo metamorphosis around 25 to 35 mm length (Thompson, 2015). Juveniles are typically found along the inner continental shelf, with adults ranging from the shoreline to a maximum of 500 m (Thompson, 2015). Age-2 Pacific cod may school over rocky benthic terrain, unlike most other ages that are found in sandy bottom habitats (Ueda et al., 2006). This species reaches a maximum length of about 120 cm, with sexual maturity occurring at about 4.9 years (Stark, 2007). Adult natural mortality rates for Pacific cod are poorly estimated, but are likely age-dependent. National Marine Fisheries Service (NMFS) stock assessments have used a natural mortality rate 0.34 yr^{-1} since 2007, with a 95% confidence interval of 0.30 to 0.38 yr^{-1} (Thompson, 2015) based on published studies (Stark, 2007; Jensen, 1996). Estimates of age-0 natural mortality for other cod stocks are 2.12% to 4.17% per day (Thompson, 2015).

4.1.2 Fishing History

Pacific cod have been harvested in a directed fishery since the 1960s that has transitioned from foreign fleets to entirely domestic fleets. Initially, the Japanese longline fishery targeted Pacific cod in the EBS, and they were also caught as bycatch in the Japanese trawl fishery for walleye pollock (*Theragra chalcogramma*) (Thompson, 2015). Domestic trawling for Pacific cod started in 1981 and displaced all foreign fleets within a decade (Thompson, 2015). Currently, the Pacific cod fishery is dominated by longline, trawl, and pot gear, which fish in five fishing seasons: January-February, March-April, May-July, August-October, and November-December to represent differences in intra-annual fleet operations (Thompson et al., 2010). In addition, there is a single survey that starts in 1982, the NMFS trawl survey (Fig. 4.2). It is assumed that all catches of Pacific cod are retained so discards are not modelled. The trawl and longline fisheries for Pacific cod have never closed since recorded catch data began in 1977, although the pot fishery has seen some years of no catch. The total catch has not exceeded the overfishing limit (OFL) since 1993 (Thompson, 2015).

4.1.3 Data for assessment purposes

Though foreign fishing for Pacific cod started in the early 1960’s, information for retained catch is only available from 1977 for the longline, trawl, and pot fisheries (Tables 4.1 and 4.2). For the assessment, there are 16 designated fleets: 15 fishing fleets (three gear types for each of the five seasons), and one survey fleet. The data for males and females are combined for consistency with the actual assessment of Pacific cod (Thompson, 2015).

Size-composition data are available for some of the fishery fleets and the survey. Size-composition data for at least one gear type per year are available from 1977 through the present. Fishery-independent survey data are available from 1982 to the present. Length-composition data are available in 5 cm bins, with 34 total length classes. In addition, the survey provides annual age-composition data for Pacific cod ages 0-20 starting in 1994. The survey data also include mean size-at-age for ages 0-20 from 1994 to 2013. Survey data for the Pacific cod are from the EBS shelf bottom trawl survey, and include strata 1-6, which have been sampled annually since 1982 (Thompson, 2015).

Fishery effort data are available from 1991 through the present. Effort is measured in kg/minute for the trawl fleet, kg/hook for longline gear, and as kg/pot in the pot fishery (Thompson, 2015). However, because there is little validation of the accuracy of the fishery effort data, the CPUE indices from the fishery fleets are ignored in the NOAA assessment (Thompson, 2015) and in this assessment.

4.2 Methods: Specifications for the Assessment

4.2.1 Base Model

Blocks are defined for Pacific cod for the same reasons as for Tanner crab (see Chapter 3), and mimic the structure used in the most recent assessment by Thompson (2015). The blocks for the trawl fishery are 1977-1984, 1985-1989, 1990-1994, 1995-1999, 2000-2004, and 2005-2014 for seasons 1 and 2, 1977-1984, 1985-1989, 1990-1999, 2000-2004, and 2005-2014 for season 3, and 1977-1979, 1980-1984, 1985-1989, 1990-1994, 1995-1999, and 2000-2014 for seasons 4 and 5. The blocks for the longline fishery are 1977-1979, 1980-1984, 1985-1989, 1990-1994, 1995-1999, 2000-2004, and 2005-2014 for seasons 1 and 2, 1977-1979, 1980-1984, 1985-1989, 1990-1999, 2000-2004, 2005-2014 for season 3, and 1977-1979, 1980-1984, 1985-1989, 1990-1994, 1995-1999, 2000-2004, and 2005-2014 for seasons 4 and 5. The blocks for the pot fishery are 1977-1994, 1995-1999, 2000-2004, and 2005-2014 for seasons 1 and 2, and 1977-1994, and 1995-2014 for seasons 3, 4 and 5. Selectivity is assumed to be constant over time for the NMFS survey.

Natural mortality is set to 0.344 yr^{-1} (Thompson, 2015) for both mature and immature cod, and implicitly for males and females as this is a single-sex model. The parameters of the growth increment are determined by the assumption that the growth increment is a linear function of current size (i.e. growth follows the von Bertalanffy curve; Thompson, 2015; Appendix A.2.5). The length-weight intercept and slope are pre-specified at values used in the 2015 stock assessment, but asymptotic size (L_∞), growth rate (K), and the CV of growth are estimated. The proportion of the year before spawning, Φ, is fixed to 0.169, the start of March (the start of season 2; Thompson, 2015).

Length-at-50%-maturity, maturity slope, recruitment steepness (h), and the extent of variation in recruitment (σ_R) are pre-specified at values from the current assessment (Thompson, 2015). Estimated parameters are the logarithm of number of fish at unfished equilibrium biomass, R_0, an R_0 offset (a scalar for the initial year), and the α (mean) and β (variance) parameters for the gamma function, which assigns the proportion of age-0 animals that settle into each size class (Equation A.1.6; Appendix A.2.5). Recruitment is allowed to occur up through the 10 cm size-class, based on age-length results in Thompson's assessment (2015). Initial fishing mortality is estimated for the January-February trawl fishery, and fixed at 0 for all remaining fishery fleets to maintain consistency with the assessment setup of initial fishing mortality (Thompson, 2015). The maximum fishing mortality rate is set to 3.0 yr^{-1} and fishing mortality is tuned five times using the hybrid method (Methot and Wetzel, 2013).

Selectivity is estimated separately for each gear type. Selectivity is a function of length for the fishery fleets, and of age for the survey. Within gear types, selectivity is mirrored based on the curve type (logistic or double logistic) assumed in the current Pacific cod assessment (Table 4.4;

Thompson, 2015). All individual parameter values can be found in Appendix A.2.6. Catchability is pre-specified at 0.77.

The overall objective function is a weighted sum of each data source's contribution to the likelihood function (Appendix A.1.7). Iterative weighting has not been performed for this model, so the weights are pre-specified by data type (Appendix A.2.7). The age-size-structured model includes ages 0 to 20, each in 1 year bins, and includes sizes from 5 to 120 cm in 5 cm size bins. The final age- and size-class are treated as plus groups.

4.2.2 Sensitivity Analysis

Sensitivity analyses were conducted for the Pacific cod base model described above. Uncertainty in the index emerged as one of the factors that had the most impact on estimation performance in the simulation evaluation (see Chapter 2), so the CVs assumed in the base model were doubled and halved in the sensitivity configurations to detect the extent to which this uncertainty might impact the assessment. Effective sample size (N) for the length-composition data was also included in this sensitivity analysis, with N doubled and halved for the survey and separately for the fishery fleets. As natural mortality, M, in the base model is fixed, a sensitivity analysis was conducted estimating M, as well as changing fixed M by ±0.02 yr^{-1} (to 0.32 yr^{-1} and 0.36 yr^{-1}). Selectivity was set as double-logistic in the base model, with some fleets mimicking a logistic curve, so an additional sensitivity configuration involved all fleets assigned logistic selectivity. However, the selectivity mirroring patterns used in the base model (Table 4.4), were the same for the sensitivity configuration. These tests resulted in eight sensitivity analyses (Table 4.5) for the Pacific cod base model.

4.3 Results

4.3.1 Parameter Estimation

The Pacific cod model estimates 200 parameters, three of which are estimated at pre-specified boundaries in the double-logistic selectivity pattern. The maximum age parameter for double-logistic age-structured selectivity in the survey is estimated at a lower bound of -9.99 in negative log space, when the initial value was 10 (i.e. survey selectivity was zero at the largest age). The January-February longline and pot fisheries were estimated at the lower bound of -9.99 for the parameter designating the size at which selectivity decreases from 1, or the width of the peak in log-space in the length-structured, double-logistic selectivity pattern. The low estimated value indicates that selectivity increases to the estimated peak value (76 cm, in this case) and then decreases immediately thereafter.

4.3.2 Model Fit Diagnostics

The model fits the survey index data well, but fails to predict several year-to-year changes that are more extreme than expected from the model (Fig. 4.2). The survey estimates for 1989 and 2008 are smaller than expected from the data. The population index is underestimated in 1986, 1994, 1995, and 2001, where the index data exhibit much larger increases in population biomass than are estimated.

The age-composition data from the NMFS survey are closely mimicked on average (Fig. 4.3), although the model expected slightly more age-2, age-5, and age-7 fish and slightly underestimated the number of age-1, age-3, and age-4 fish than are caught in the survey (Fig. 4.3). The length-composition data are well-fitted for the fishery fleets and the survey (Fig. 4.4). The model expects a slightly lower frequency of fish in the 20 cm range for the NMFS survey, which is likely due to an over-emphasis on fitting to the survey index and the survey age-composition data. The model is able to mimic the length-frequency data for the fishery fleets well on average (see Appendix A.2.8 for the fits by year). However, the estimated peaks in size for the trawl fishery are about one size class smaller than expected in the early part of the year (January-April). The annual fits to the survey age-composition data are generally very good, with estimates closely approximating the frequencies of each age group seen in each year (Appendix A.2.8a). The annual estimates of the survey length-composition are reasonable, but with some over-estimates of smaller size classes (Appendix A.2.8b), likely due to a conflict related to the need to fit the survey age data. The fits to the length-composition of the fishery fleets is generally good, with only the August-October trawl fishery regularly showing over- or under-estimates of the frequencies of multiple size classes (Appendix A.2.8b).

Length selectivity patterns for the fishery fleets are mirrored within gear type, based on information for the shape of selectivity for each fleet from Thompson (2015) (Table 4.4; Fig. 4.5). All fleets for fishery trawl gear show a logistic selectivity pattern, with an asymptote between 0.99-1, and selecting for fish 82-97 cm or greater with selectivity greater than 0.90. The longline fishery exhibits dome-shaped selectivity, with an optimal length at 72-82 cm. The January-February, March-April, and May-July pot fishery seasons have dome-shaped selectivity, while the fishery for the remainder of the year shows logistic selectivity (Thompson, 2015). The winter pot fisheries optimally select for fish 72-87 cm. The remaining pot fisheries select for fish 77 cm or larger, with selectivity greater than 0.99. Age selectivity for the NMFS survey is dome-shaped, and ages 3-5 are selected for with selectivity greater than 0.99 (Fig. 4.6).

4.3.3 Comparisons with Previous Assessments

The estimates of age-0 recruitment are quite similar between this assessment and the most recent official assessment in Thompson (2015), with most annual estimates falling within the 95% confidence intervals of estimated recruitment from the age-size-structured assessment (Fig. 4.7). The official assessment estimates for female spawning stock biomass (SSB) are within the 95% confidence intervals of estimated SSB from the age-size-structured assessment through the 1980s, but begin to diverge after 1990, resulting in the SSB in the age-size-structured assessment being much larger than the estimates from the official assessment (Fig. 4.8). Trends between the two assessments are similar, and differences in estimates rarely exceed 300,000 t.

4.3.4 Sensitivity Analysis

A number of inputs substantially impact the final outcome of the assessment, notably assuming logistic selectivity for all fleets instead of allowing some to have dome-shaped selectivity, and estimating natural mortality instead of fixing it.

Logistic Selectivity

Assuming logistic selectivity instead of dome-shaped selectivity did not result in dramatic changes from the base model. Less of the year-to-year variability in the survey index of abundance was captured in the logistic selectivity sensitivity configuration (Fig. 4.9). The logistic selectivity pattern resulted in an underestimate of the frequency of ages 1-3 in the survey and the poorest fit to the age-composition data of all sensitivities tested (Fig. 4.10).

Assuming logistic selectivity resulted in slightly poorer fits to length-composition data (Fig. 4.11), though much less than might have been expected, given all the longline and January through

May pot fisheries have dome-shaped selectivity in the base model and the assessment (Fig. 4.12; Thompson, 2015). While selectivity patterns remain relatively similar among the sensitivity tests, assuming logistic selectivity generally results in a shallower slope, and fully-selected fish occurring at larger sizes in the trawl fleets (Fig. 4.13). The first fully-selected age class occurs at 6 years old for the survey with logistic selectivity, rather than at 3 years old as is seen in the base model (Fig. 4.14).

Changing the selectivity pattern to logistic led to similar estimates of recruitment until 2013, when the logistic sensitivity configuration resulted in a low point that was opposite a local maximum in the base model (Fig. 4.15). The recruitment from the logistic selectivity sensitivity configuration was estimated at 1 million more individuals than in the base model in 2014. Overall, however, logistic selectivity resulted in smaller estimates of SSB (Fig. 4.16).

Natural Mortality

When M is estimated instead of being fixed, model estimates of M occur at the pre-specified upper-bound of 0.4 yr^{-1}, and the M offset for immature animals is also estimated at the upper-bound of 1.0. Estimating M results in an increase to the estimate for L_∞ from about 90 cm in the base model to 109 cm, and the estimate for K changes from 0.2 yr^{-1} in the base model to 0.15 yr^{-1}. The log R_0 estimate increases from 13 to 17, which, in normal space, is more than a 5000% increase in initial recruitment. The estimates of α and β values in the Beverton-Holt stock-recruit relationship also change, resulting in α increasing from -42 in the base model to 9, and β decreasing from 1.3 in the base model to 0.4. These changes indicate a major, and probably unrealistic, shift in the underlying population dynamics for both growth and recruitment. Nevertheless, estimating natural mortality resulted in the largest improvement in the fit to the survey data (Fig. 4.9).

Many of the age and length fits are quite similar to the base model for both the survey and the fishery data (Fig. 4.10-4.12). The selectivity patterns for the fishery fleets are generally similar to the base model (Fig. 4.13), but differ markedly for the survey fleet (Fig 4.14). Estimating natural mortality results in more fully-selected ages, but starting at an older age (6-year-old cod) than the base model configuration (3-year-old cod).

Much higher estimates and recruits result from allowing natural mortality to be estimated. Estimates of recruits were up to 35 times larger than the base model (Fig. 4.15). Most estimates of SSB from the sensitivity analyses fell within the 95% confidence intervals from the base model, although estimated natural mortality resulted in much larger SSB estimates overall (Fig. 4.16), and an estimated 195,000 t more SSB in the final year of the assessment (2014). However, estimating natural mortality at fixed rates of 0.32 yr^{-1} and 0.36 yr^{-1} (±0.02 yr^{-1} from the base model) resulted in little change from the base model, but larger estimates of spawning stock biomass were the result of the lower natural mortality value, and higher natural mortality resulted in smaller estimates of spawning stock biomass (Figure 4.16).

Uncertainty (CV) and Effective Sample Size (N)

Changing uncertainty (CV) and effective sample size (*N*) for the survey and fishery fleets resulted in fewer changes from the base model than using logistic selectivity and estimating *M*. Decreasing uncertainty around the survey estimates resulted in a better fit to data, although fits overall remained similar to those for the base model. The fits to the age- and length-composition data for the survey remained similar for most sensitivities. The number of fully-selected ages increases when uncertainty and effective sample size are increased for the survey (Fig. 4.14).

Changing index CV and effective sample size for the survey and fishery fleets has little effect on the estimates of recruitment and SSB. Decreasing survey uncertainty resulted in slightly larger peaks in recruitment, although with the same pattern as the base model (Fig. 4.15). Low survey uncertainty resulted in higher SSB estimates, though not as high as when *M* was estimated (Fig. 4.16).

4.4 Discussion

The estimates of spawning stock biomass differ between the age-size-structured assessment and the official assessment, particularly after the mid-2000s, even though both use the same data. With the age-size-structured model estimating 158,000 t more mature female adults than the assessment in 2015, the Acceptable Biological Catch (ABC) was 30% higher (91,700 t) than the current assessment for 2015 (based on Pacific cod as a Tier 5 stock; NPFMC, 2017).

Process and measurement uncertainty are to be expected, but the difference in spawning stock biomass estimation in the last years of the two assessments suggests that some uncertainty is currently uncounted for. In cases such as this, where estimation of the ABC in the most recent year modelled differed by 30%, managers should consider model averaging the results from multiple models in the assessment process. While fishery assessments may sometimes be averaged over several versions of the same model with varied parameter values, Buckland et al. (1997) suggests a method for model selection between models with fundamental differences. Allen Akselrud et al. (2017) propose using this approach to weight multiple models, and use a weighted model average of spawning biomass between several models that use the same data, as could be applied in this case.

Several factors could account for the difference between the results from the age-size-structured assessment and the official assessment (Thompson, 2015), including maturity and growth being modelled as length-specific instead of age-specific dynamics. The same Pacific cod data used here and applied in the age-size-structured, and age-structured options of this modelling framework yield different results (Allen Akselrud et al., 2017). Maturity in the age-size-structured model was based on length, while maturity in the age-structured model was based on age. How maturity is modelled has an important impact on estimates of spawning biomass, which are larger when maturity is length-specific, rather than age-specific. This difference in maturity modelling is likely an important factor causing estimates of SSB to be larger in the age-size-structured model than the assessment model by Thompson (2015).

Whether growth is based on size-structure (e.g. represented using a size-transition matrix), or age-structure (e.g. following a von Bertalanffy curve), may contribute to the different estimates of SSB between the two assessments. The age-based growth curve results in a better fit to the data than the age-size-structured model with its size-transition matrix (Allen Akselrud et al., 2017). Improving the fit to the data where possible should lead to more accurate estimates of biomass and management reference points. However, if the primary goal is to provide management reference points, model-predicted survey indices provide more robust estimates and are less affected by model uncertainty in this case (Allen Akselrud et al., 2017).

The greatest changes from the base model occurred when natural mortality was estimated instead of being pre-specified. This suggests that more information may be needed to inform the model about natural mortality or information is missing that might affect estimation of total mortality. With the backing of studies for cod natural mortality (e.g. Jensen, 1996; Stark, 2007), a strong prior could also be introduced when allowing the model to estimate natural mortality, to keep it restricted within a range that is biologically plausible. Interestingly, spawning stock biomass estimates are higher than in the base model when natural mortality is estimated at an upper limit of 0.4 yr^{-1}, but fixed natural mortality at lower rates than the base model resulted in higher spawning stock biomass estimates, while higher fixed natural mortality rates than the base model result in lower spawning biomass estimates. Estimated natural mortality is likely confounded with other parameters in the model, where increasing natural mortality leads to a better fit to the data than the base model, but also unrealistic parameter, biomass, and recruitment estimates.

Additional future versions of this model could also examine the effect of including effort data from the fishery. There is information available for fishery catch per unit effort (CPUE). However, because there is little validation of the accuracy of the fishery effort data, the CPUE index from the fishery fleets is not used for parameter estimation. It may be beneficial to allow fishery effort data as a data source, but with sensitivity tests at different weightings to determine how much this information may or may not drive the model. A common issue associated with including CPUE data is that CPUE may remain high while abundance declines, resulting in hyperstability of estimates of abundance (Harley et al., 2001).

Another useful addition to this model may be to incorporate platoon structure (Methot, 2013), which would allow for two different size-transition matrices. Male and female growth is evaluated separately in many assessments, so allowing for two growth curves may provide the potential to see differences between males and females in the population.

While there are some caveats to each method of assessment, it is clear that the age-size-structured assessment demonstrates relatively good fits to the data, and could be used with the NMFS model to better understand sources of model uncertainty. Comparing among multiple stock assessments can provide a better understanding of model uncertainty due to differences in model structure, and differences in fixed and estimated parameters (Ralston et al., 2011). Variation among assessments with a CV of up to 37% is common (Ralston et al., 2011), indicating the importance of examining multiple models for a single species.

4.5 Tables

Table 4.1. Specifications for data by type and fleet, and the years for which each data source is available for Pacific cod.

Fleet	**Data Type**	**Years**
Directed Fishery	Catch biomass, Size-composition	1977-2014, 1977-2014 (start year varies by fleet)
Survey	Abundance, Size-composition, Age-composition, Mean size-at-age	1982-2014, 1982-2014, 1994-2013, 1994-2013

Table 4.2. Data available for the Pacific cod assessment by fleet and type. The designation “na” indicates that there are no data available, or that the data type does not exist for the specified fleet.

Fleet	Directed	Survey
Number of Fleets:	15	1
Catch	■	na
Index	░	■
Discard	░	na
Effort	░	na
Size-Frequency	■	■
Age-Frequency	░	■
Mean Size-at-Age	░	■

Table 4.3. Overview of the options used for parameter estimation in the Pacific cod assessment.

Initial parameter values are available in Appendix A.2.5.

Process	**Settings**
Maturity (age-based)	Logistic
Growth curve	Von Bertalanffy
Selectivity: sex 2	NA (single-sex model)
Age Selectivity: Survey	Double-Logistic
Length Selectivity: Fishery	Double-Logistic
Length Selectivity: Survey	Constant
Catchability: Survey	Fixed at 0.77

Table 4.4. Specifications for how selectivity is set by gear type for Pacific cod. The "Mirror" option copies the selectivity pattern of the fleet number indicated in the Mirror Fleet column, and is not estimated independently. "NA" in the final column means that the selectivity parameters for that fleet are estimated.

Fleet Number	Fishing Months	Gear	Season	Type	Mirror Fleet
1	Jan-Apr	Trawl	1	Double-Logistic	NA
2	Jan-Apr	Trawl	2	Mirror	1
3	May-Jul	Trawl	3	Double-Logistic	NA
4	Aug-Dec	Trawl	4	Double-Logistic	NA
5	Aug-Dec	Trawl	5	Mirror	4
6	Jan-Apr	Longline	1	Double-Logistic	NA
7	Jan-Apr	Longline	2	Mirror	6
8	May-Jul	Longline	3	Double-Logistic	NA
9	Aug-Dec	Longline	4	Double-Logistic	NA
10	Aug-Dec	Longline	5	Mirror	9
11	Jan-Apr	Pot	1	Double-Logistic	NA
12	Jan-Apr	Pot	2	Mirror	11
13	May-Jul	Pot	3	Double-Logistic	NA
14	Aug-Dec	Pot	4	Double-Logistic	NA
15	Aug-Dec	Pot	5	Mirror	14
16		Survey		Double-Logistic	NA

Table 4.5. Specifications for each sensitivity configuration.

Sensitivity:	Test:	Change:
Survey CV	High	Doubled initial values
Survey CV	Low	Halved initial values
Effective Sample Size: Survey Length-Composition	High	Doubled initial values
Effective Sample Size: Survey Length-Composition	Low	Halved initial values
Effective Sample Size: Fishery Length-Composition	High	Doubled initial values
Effective Sample Size: Fishery Length-Composition	Low	Halved initial values
Natural Mortality		Estimated
Selectivity		All logistic

4.6 Figures

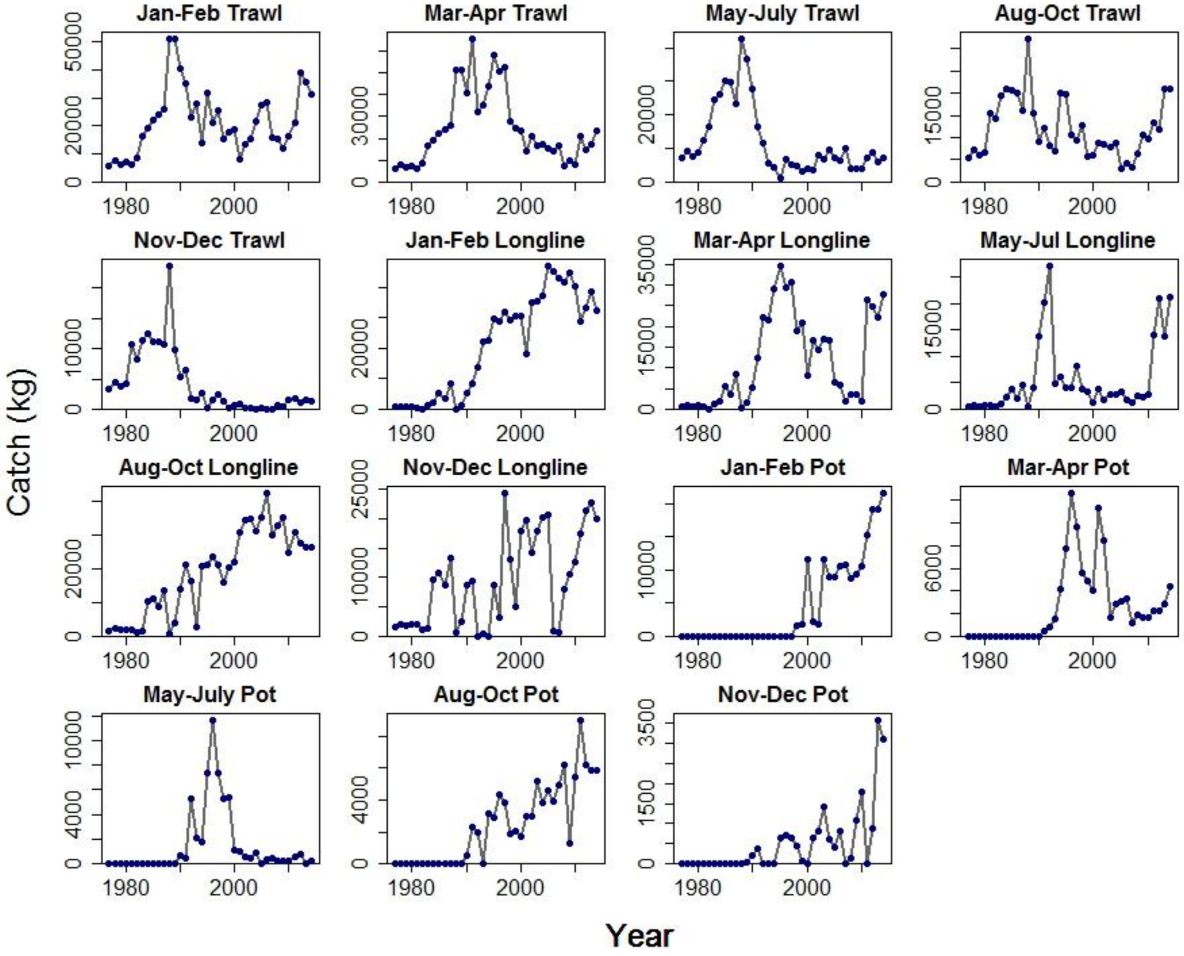


Figure 4.1. Reported catch (kilograms) for each Pacific cod fleet (1977-2014).

◆ Observed
■ Predicted

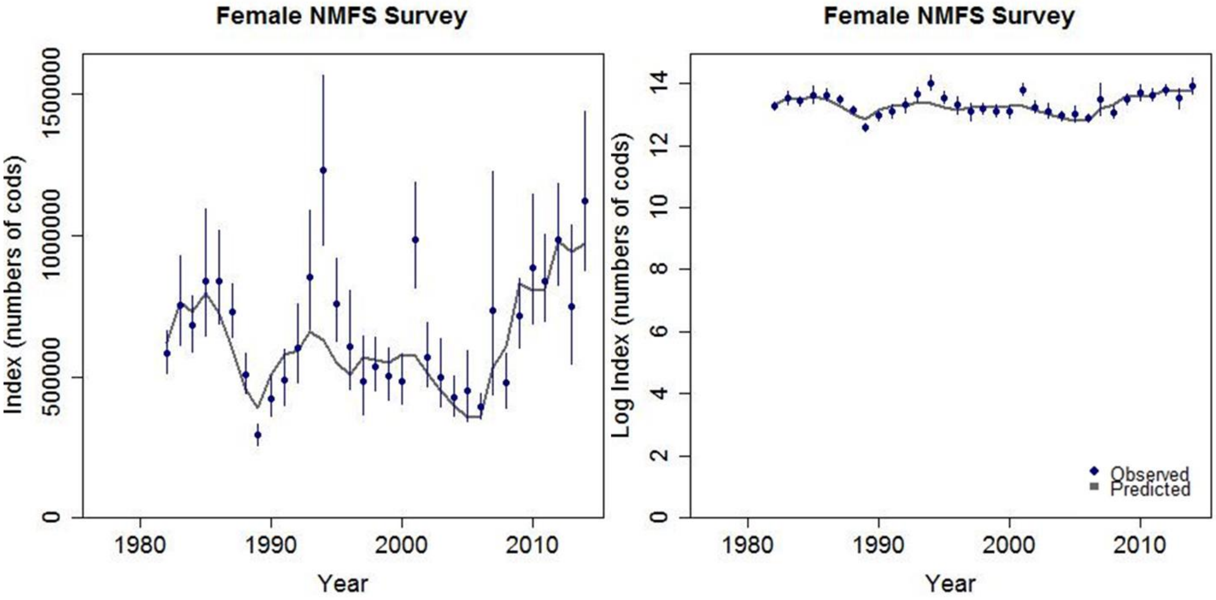


Figure 4.2. Index and log index plots (in millions of cod) with observed (points) and model-predicted (lines) abundance estimates for the NMFS survey for Pacific cod.

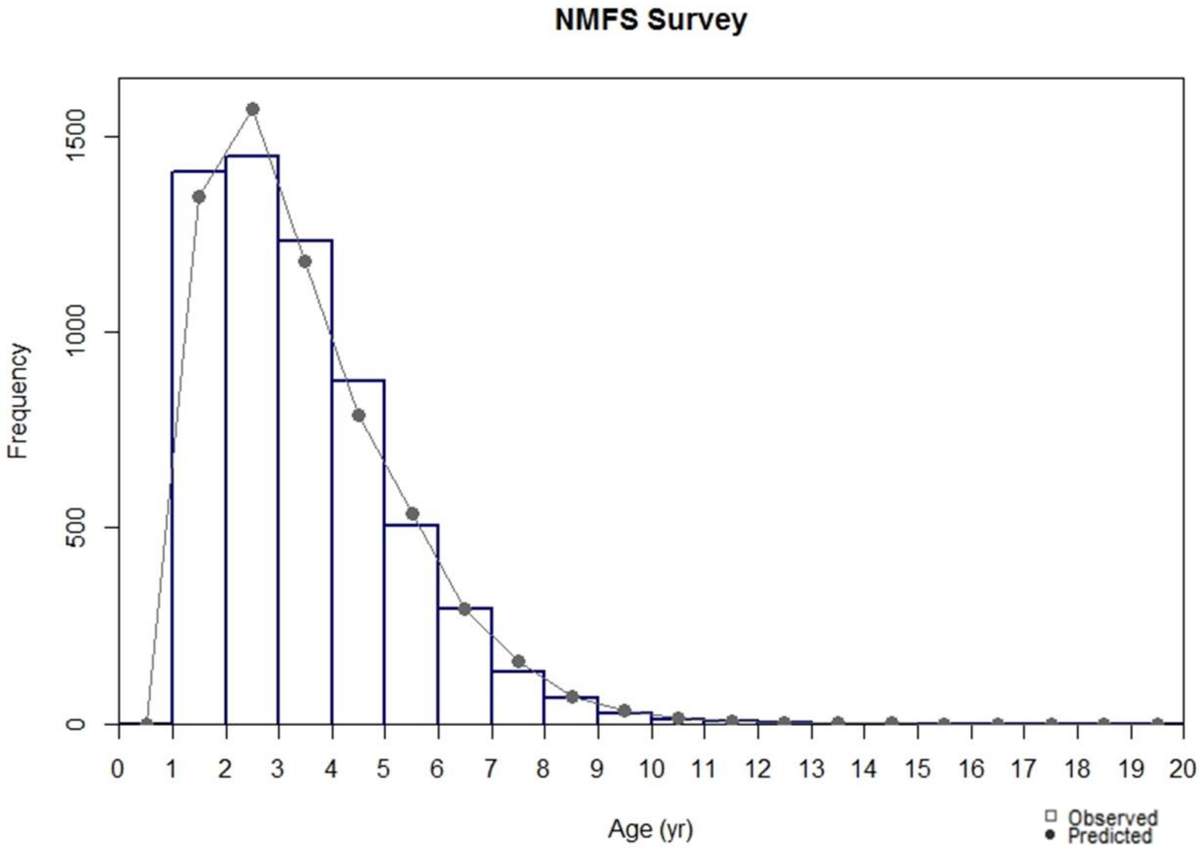


Figure 4.3. The age-composition data averaged over years (bars) and model-based predicted frequencies (points) for Pacific cod in the NMFS EBS trawl survey.

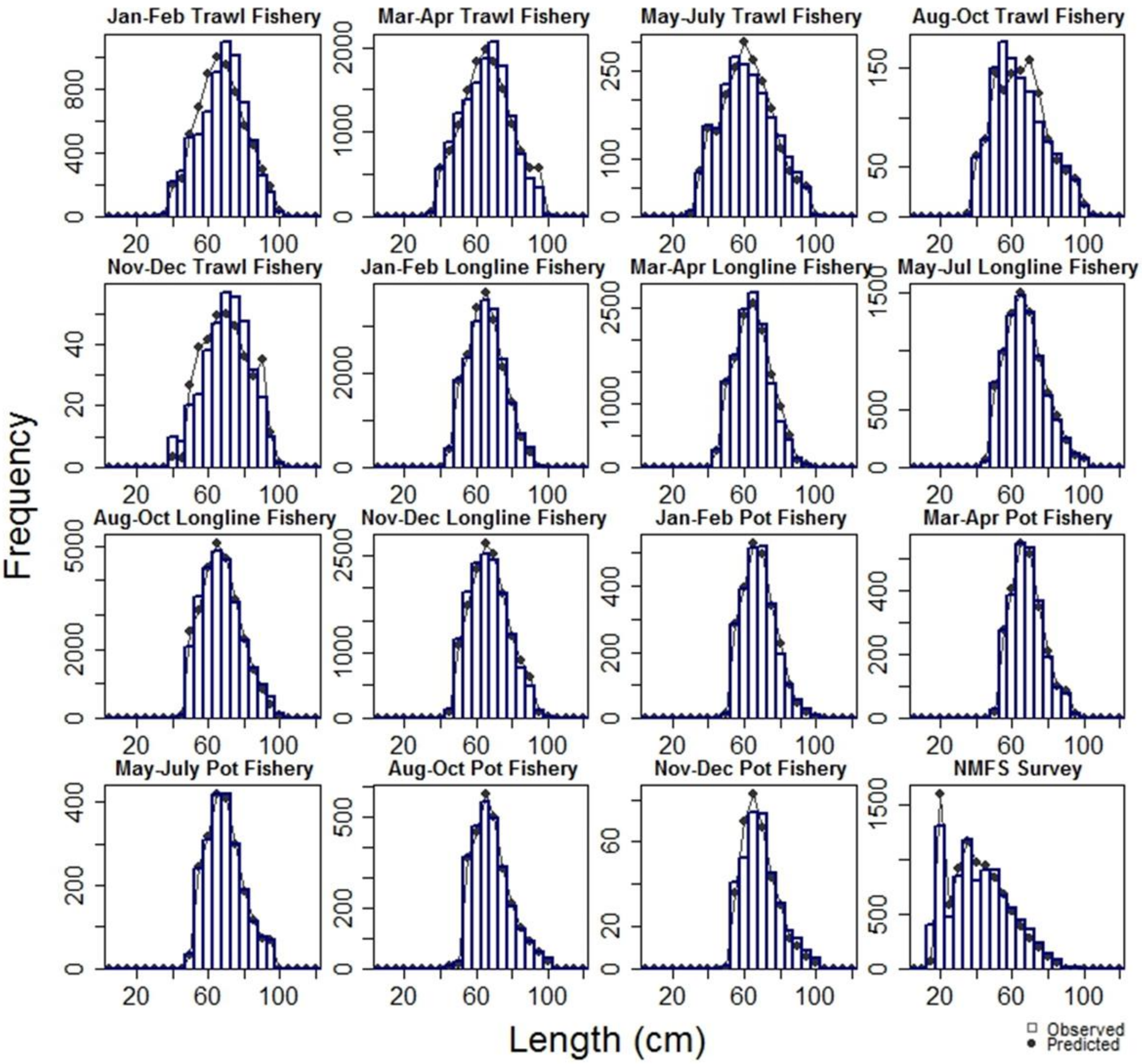


Figure 4.4. The length-composition data averaged over years (bars) and the model-predicted frequencies (points) for Pacific cod from the fishery fleets and the NMFS EBS trawl survey.

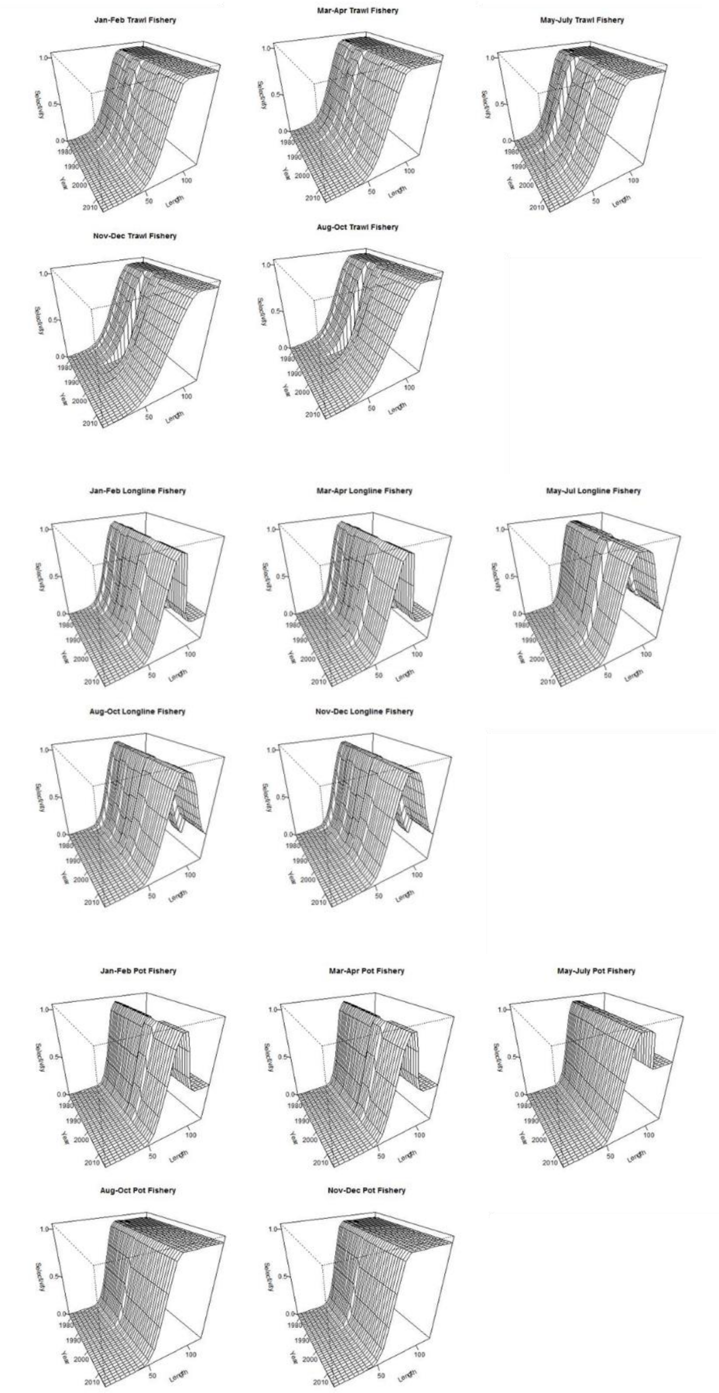


Figure 4.5. Length selectivity patterns for Pacific cod in each fishery fleet, by gear type. Table 4.4 describes the patterns and mirroring designated for each fleet. Each panel depicts the level of selectivity at each length, with one line for each year.

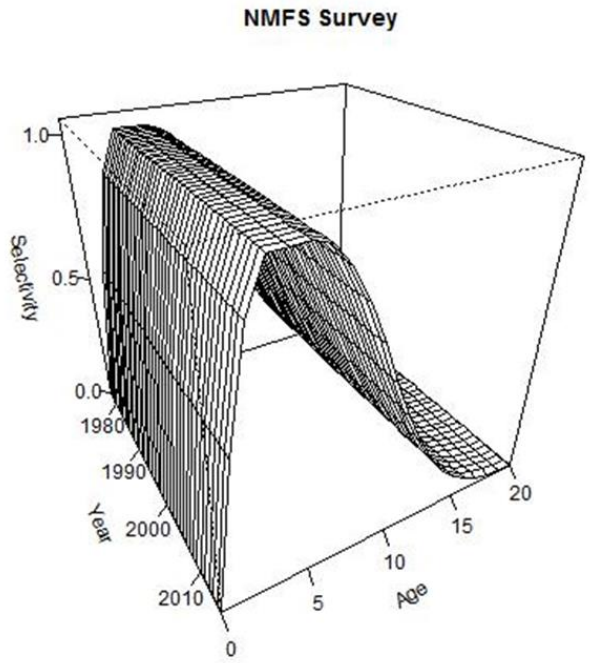


Figure 4.6. Age selectivity pattern for Pacific cod in the NMFS EBS trawl survey, depicting the level of selectivity at each age, with one line for each year.

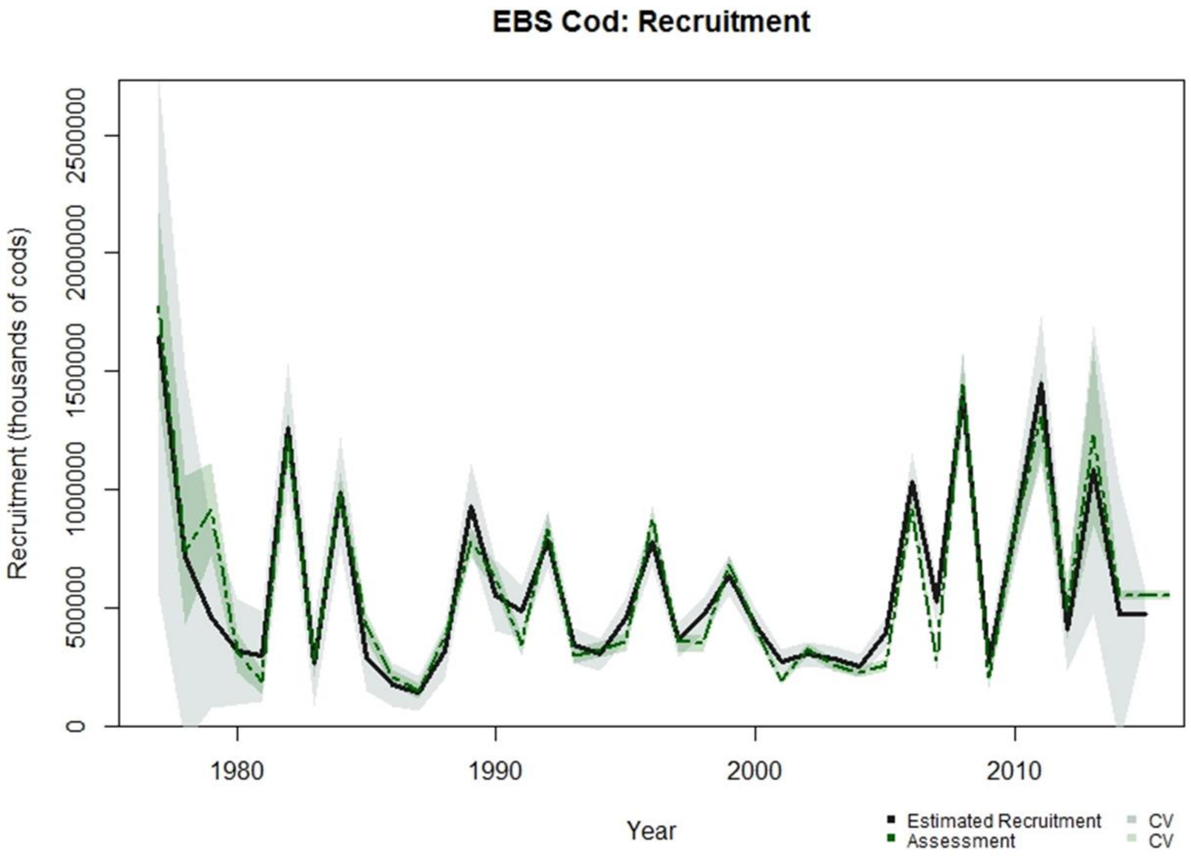


Figure 4.7. Model-predicted recruitment of Pacific cod (black solid line) and 95% confidence intervals (grey) for age-0 cod. The panel shows predicted recruitment across all years with data for the assessment (1977-2014). Current assessment values and 95% confidence intervals for recruitment are shown (green dotted line and green shading) (Thompson, 2015).

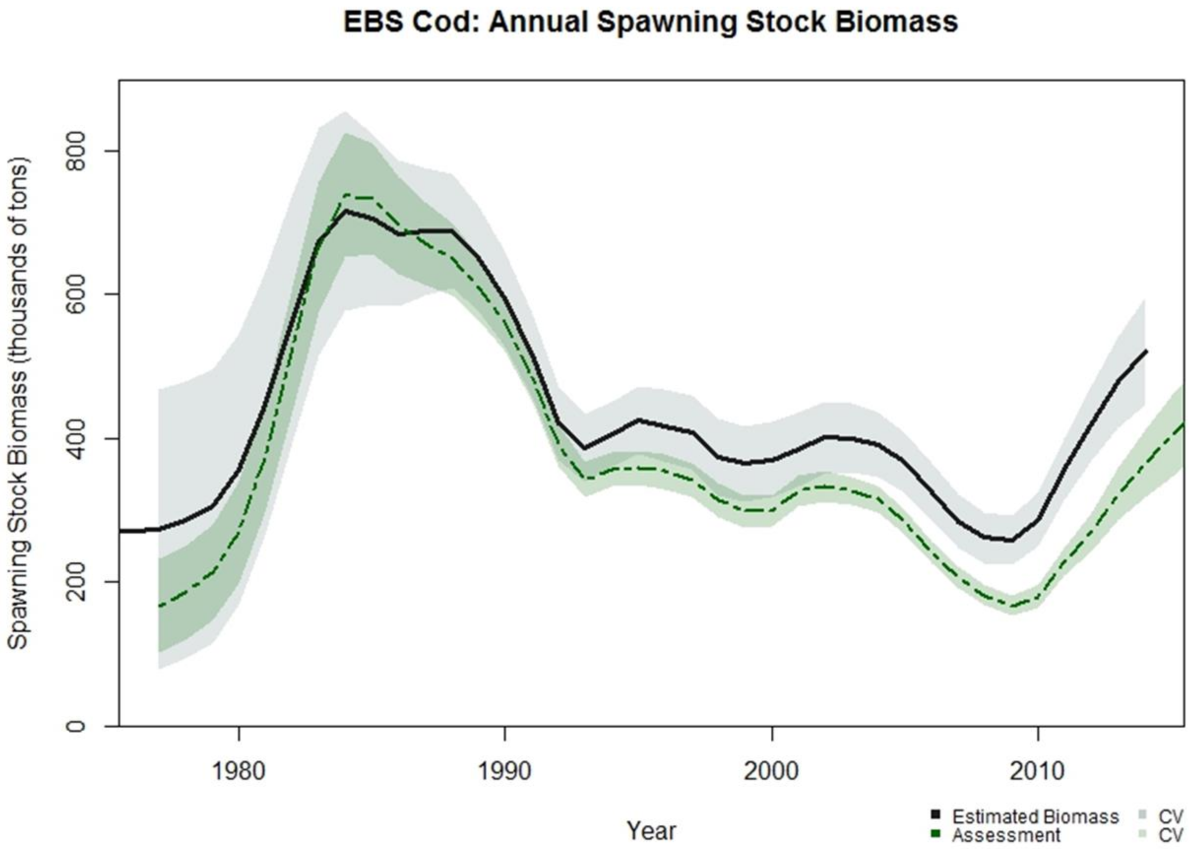


Figure 4.8. Model-predicted spawning stock biomass (SSB) of Pacific cod across all years with data for the assessment (1977-2014), with 95% confidence intervals (grey). The current assessment-predicted SSB and 95% confidence intervals are denoted by the dotted green line and green shading (Thompson, 2015).

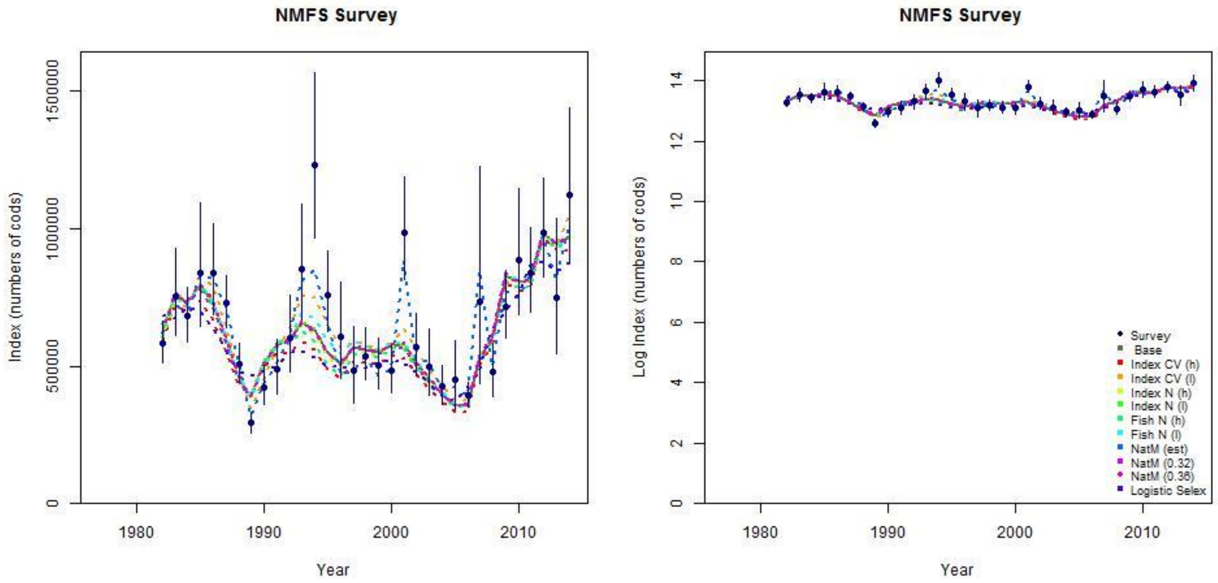


Figure 4.9. Index and log index plots (in millions of cod) for the base model and each sensitivity configuration, with observed (points) and model-predicted (lines) abundance estimates for the NMFS survey for Pacific cod.

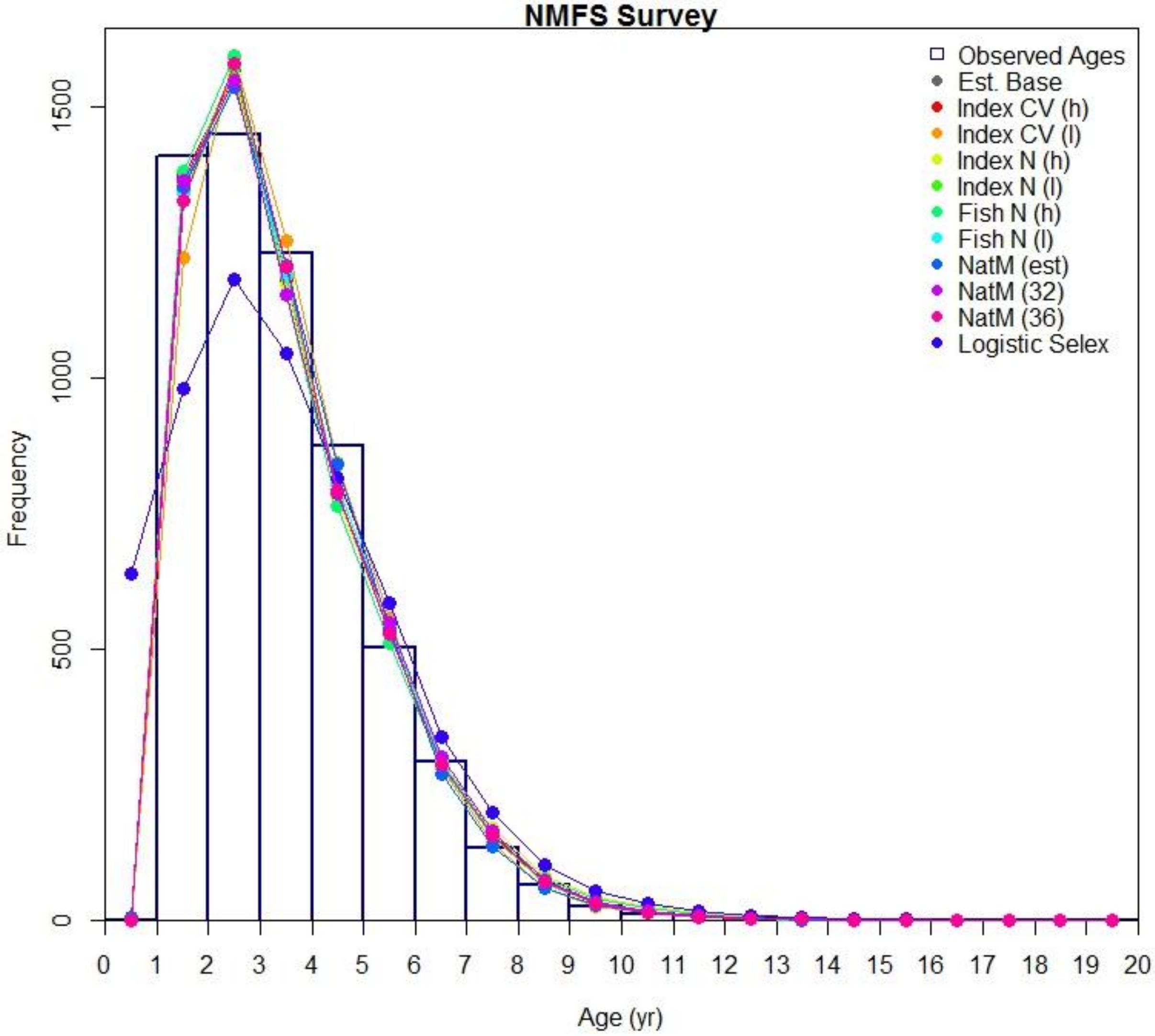


Figure 4.10. The age-composition data averaged over years (bars) and model-based predicted frequencies (points) for the base model and each sensitivity configuration of Pacific cod in the NMFS EBS trawl survey.

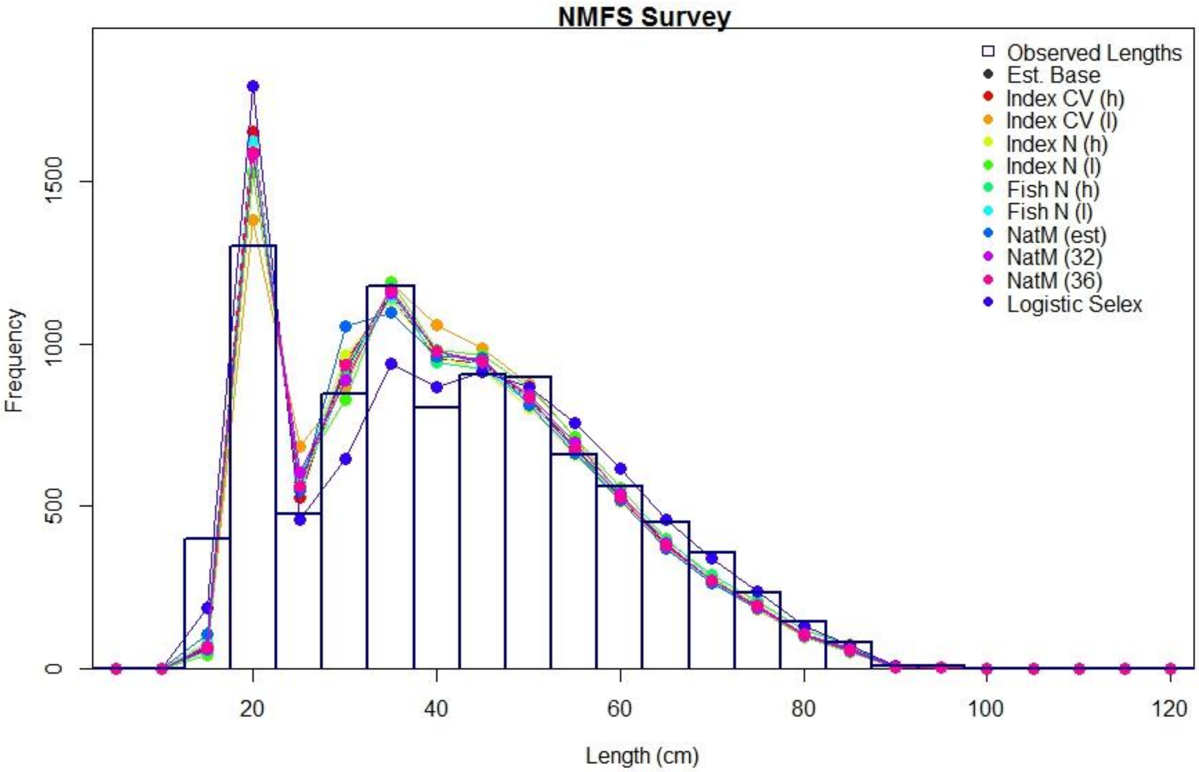


Figure 4.11. The length-composition data averaged over years (bars) and the model-predicted frequencies (points) for the base model and sensitivity configurations of Pacific cod from the NMFS EBS trawl survey.

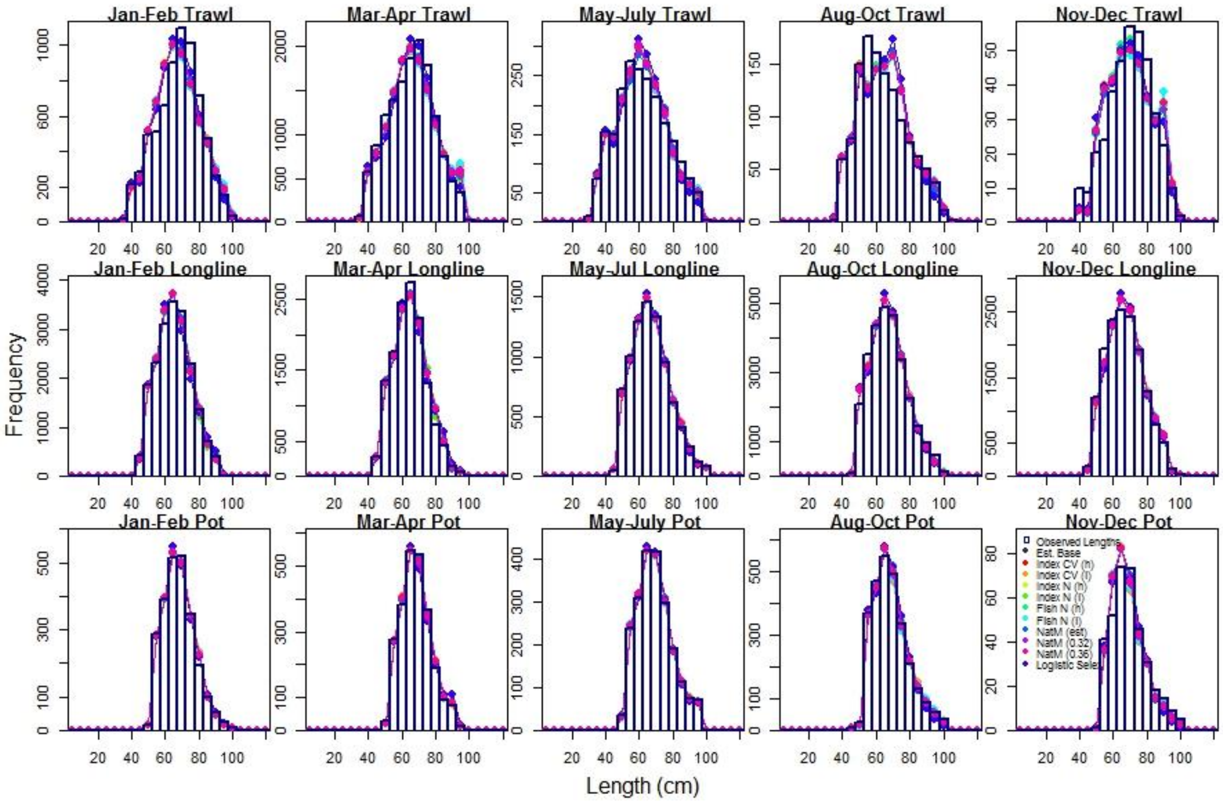


Figure 4.12. The length-composition data averaged over years (bars) and the model-predicted frequencies (points) for the base model and sensitivity configurations of Pacific cod from the fishery fleets.

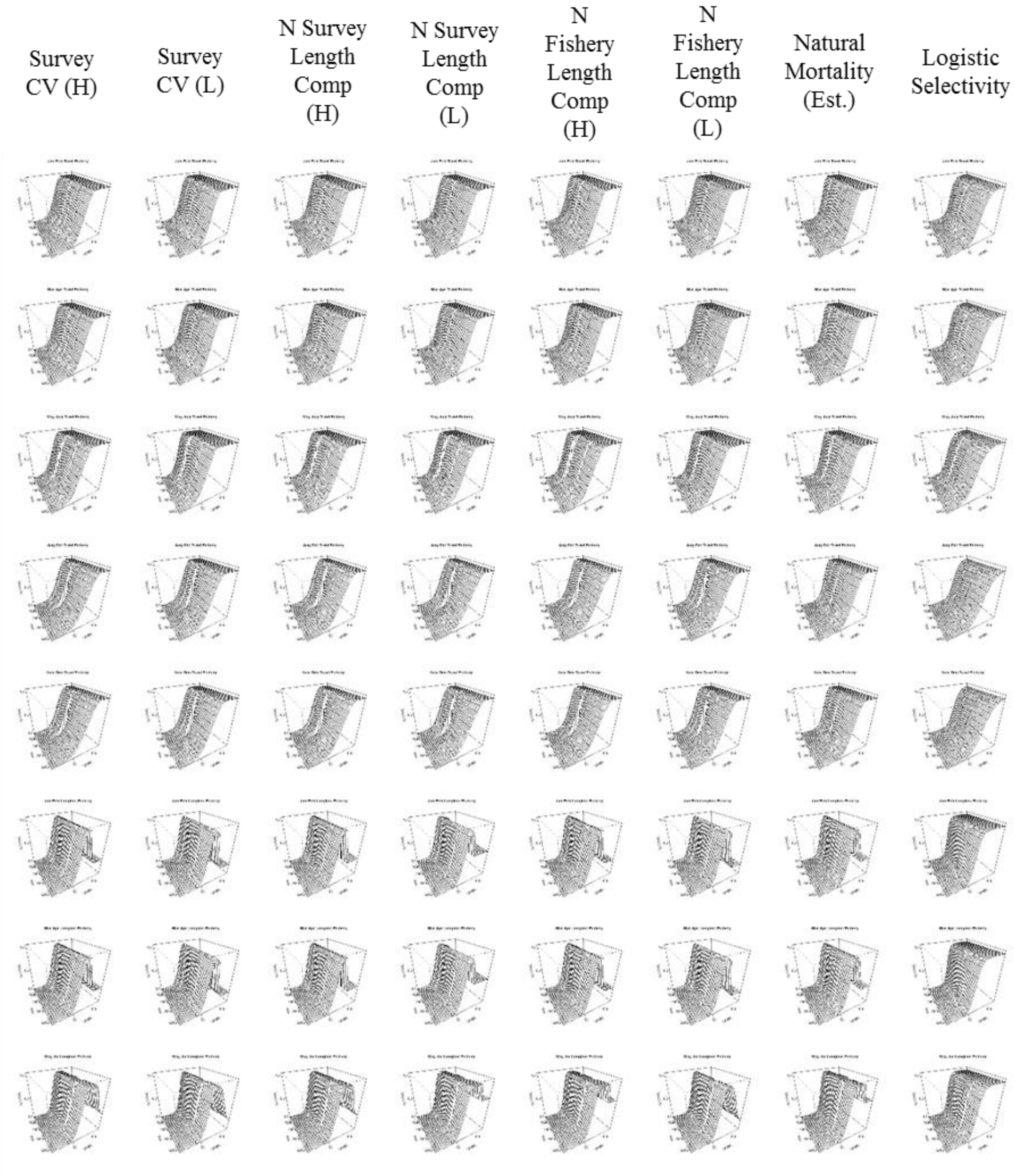


Figure 4.13. Length selectivity patterns for Pacific cod in each fishery fleet, by gear type, and sensitivity configuration. Table 4.4 describes the patterns and mirroring designated for each fleet.

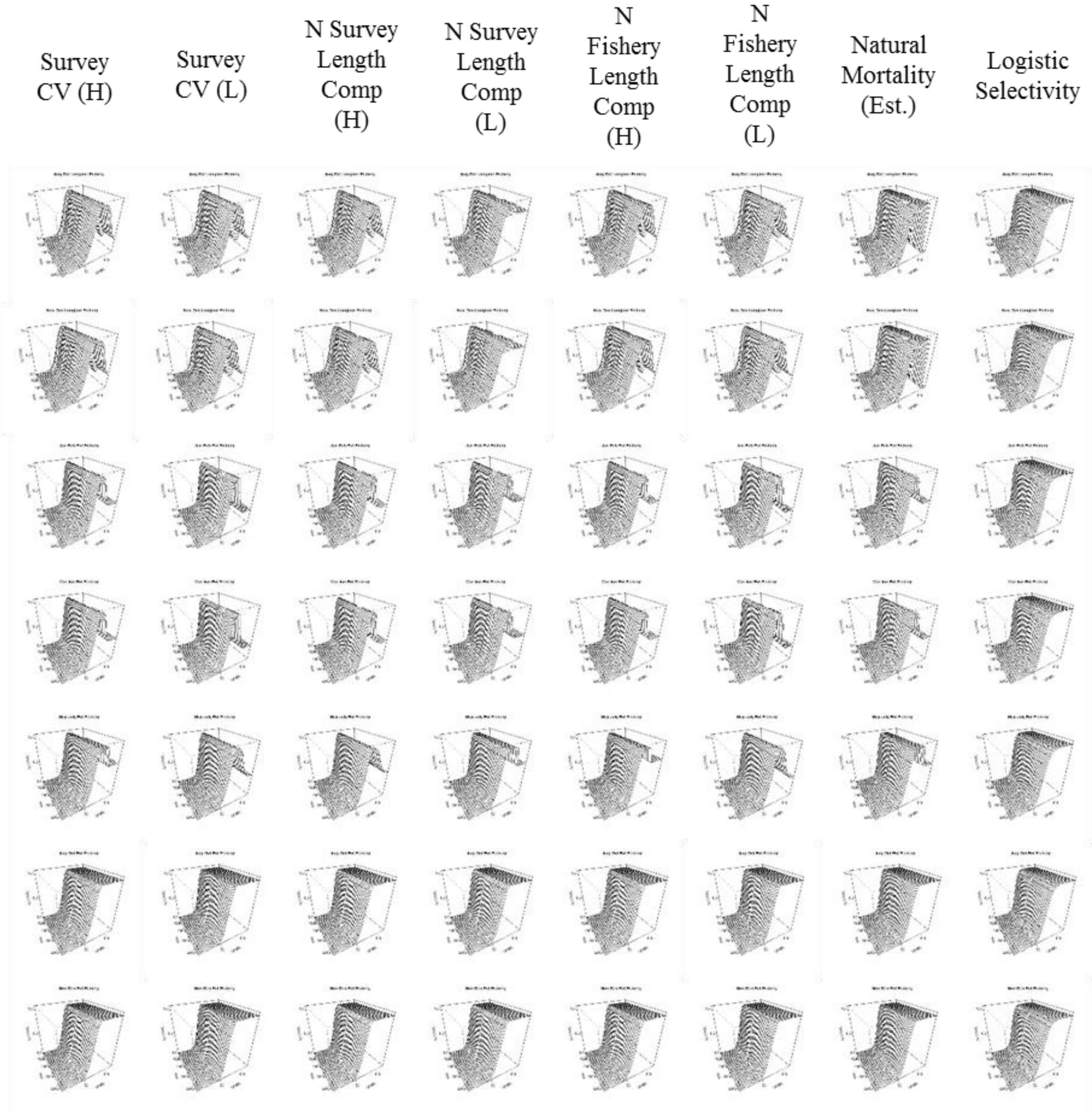


Figure 4.13. Continued. Pacific cod length selectivity patterns for each sensitivity configuration of the fishery fleets.

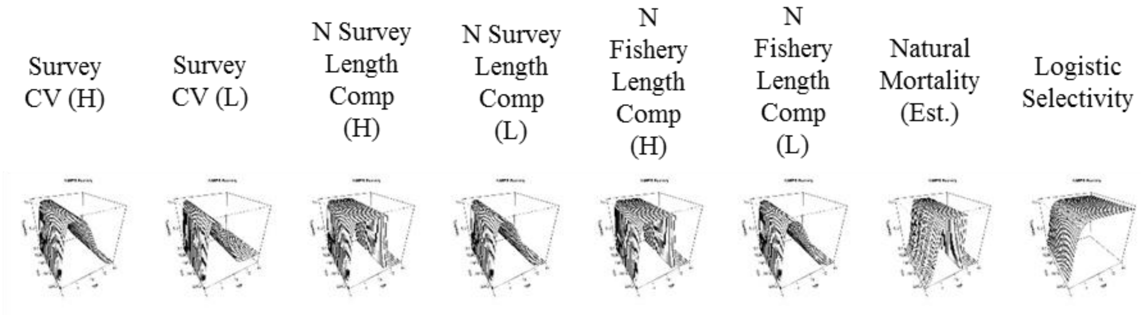


Figure 4.14. Age selectivity pattern by sensitivity configuration for Pacific cod in the NMFS survey.

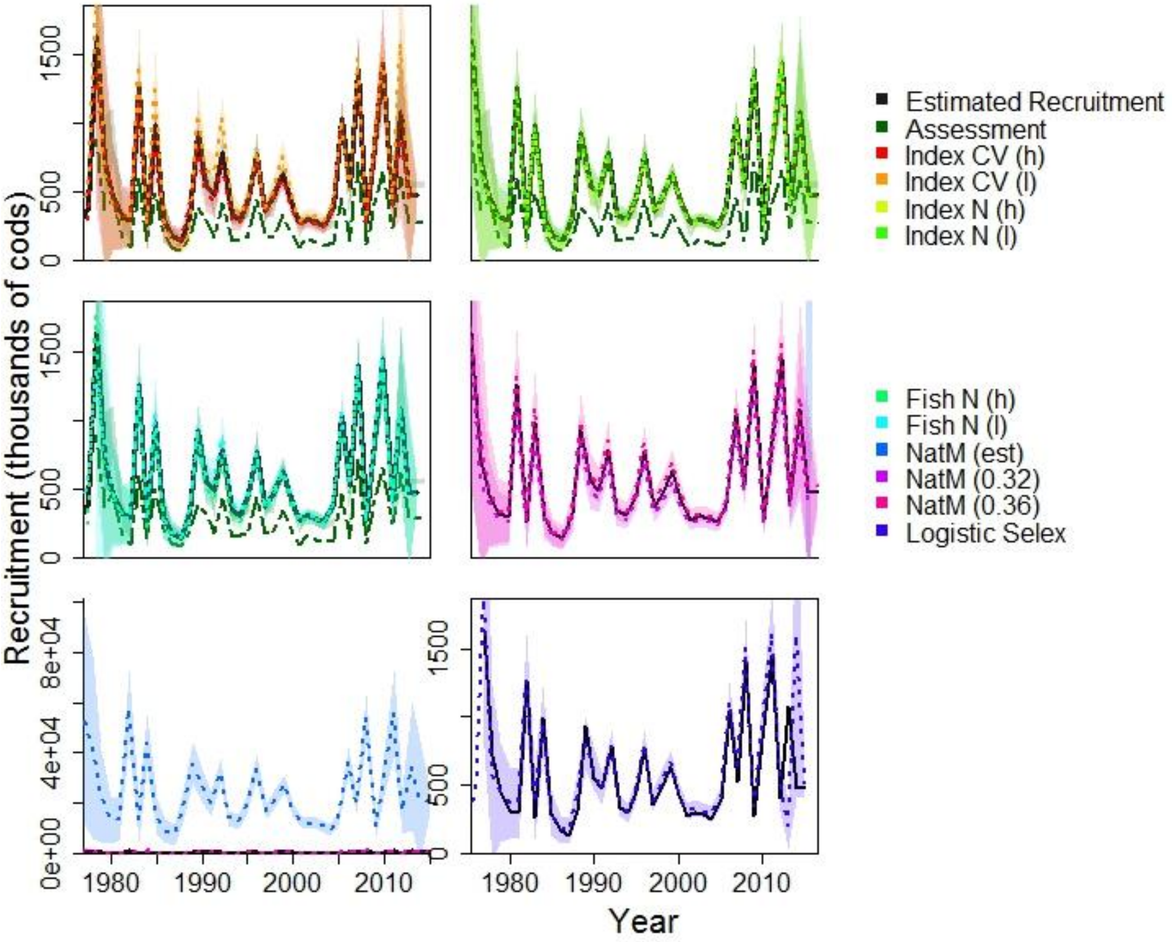


Figure 4.15. Model-predicted recruitment of Pacific cod (solid colored lines) and 95% confidence intervals for age-0 cod. Predicted recruitment is shown for years with data (1977-2014). Current assessment values for recruitment are shown (green dotted line) (Thompson, 2015).

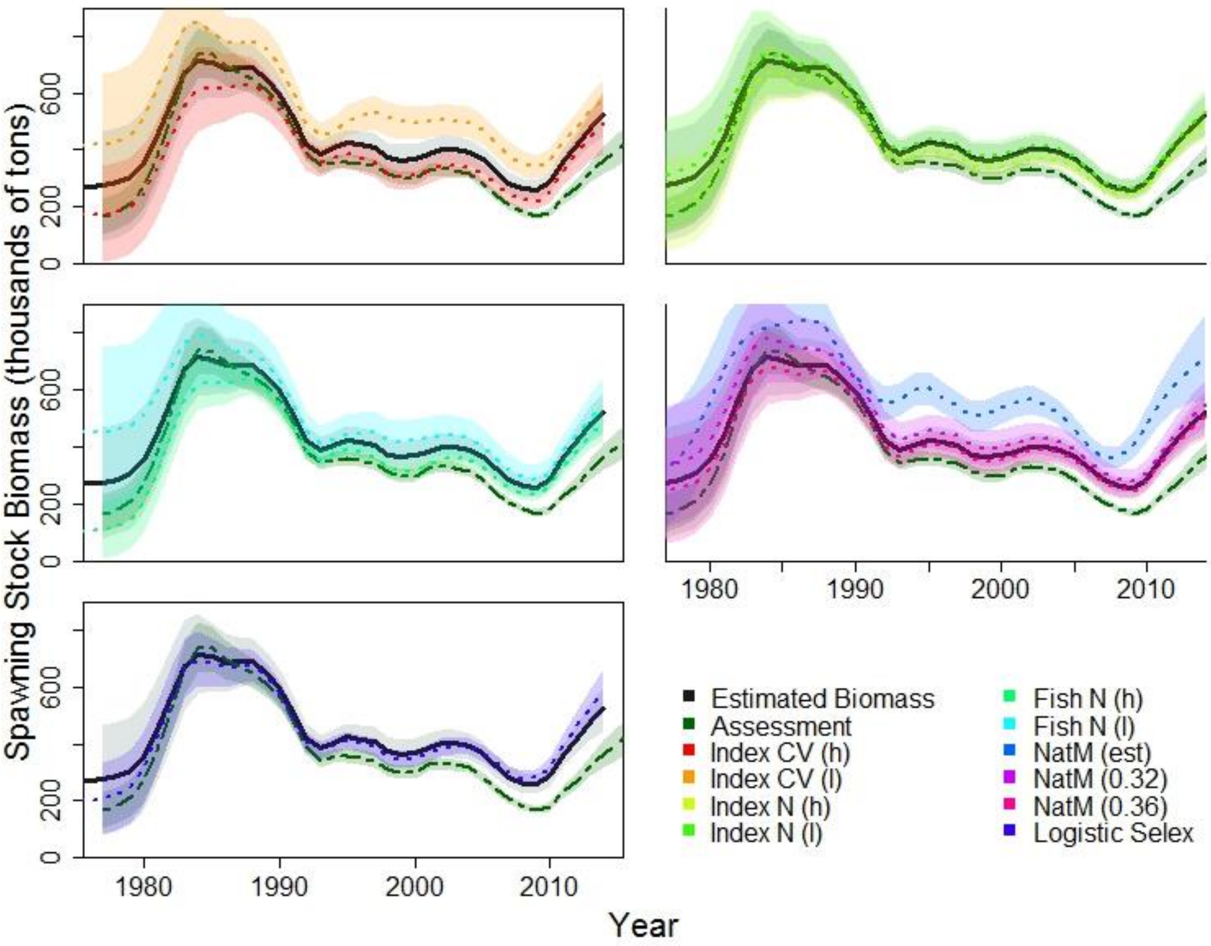


Figure 4.16. Model-predicted spawning stock biomass (SSB) of Pacific cod across all years with data (1977-2014), including 95% confidence intervals. The current assessment-predicted SSB is denoted by the dotted dark green line (Thompson, 2015).

Chapter 5: Concluding Remarks

The age-size-structured assessment performed well in the basic simulation tests conducted (Chapter 2), and generally produced biomass estimates similar to those found in the current NOAA assessments for the example applications (Chapters 3 and 4). Some of the differences in performance for each of the assessments is likely due to the different assumptions made in each. The Eastern Bering Sea Tanner crab (Tanner crab; *Chionoecetes bairdi*) assessment (Chapter 3) estimated natural mortality (*M*), while it is fixed for Eastern Bering Sea Pacific cod (Pacific cod; *Gadus macrocephalus*) (Chapter 4). The sensitivity analysis exploring estimation of *M* for Pacific cod resulted in a better fit to the survey index of abundance, but estimated a much larger spawning stock biomass (SSB) than both the base model and the current assessment (Thompson, 2015). In both the Tanner crab and Pacific cod models, steepness (*h*) is a fixed parameter, while *M* is estimated in the base model for Tanner crab and estimated in a sensitivity analysis for Pacific cod. However, Magnusson and Hilborn (2007) find that, generally, models estimating *h* perform better on average than those estimating *M*. Punt et al. (2013) note that *M* can be confounded with selectivity in age- and sex-structured models, but that this problem can be alleviated with survey data. Fortunately, both the Tanner crab and Pacific cod models have several decades of survey data that are sampled annually. It is extremely important to have a robust fishery-independent index of abundance when confounding variables are present, to inform the model on general trends in abundance over time. The sensitivity analyses suggest that the fit to the index data improved when the survey index of abundance data had a smaller coefficient of variation.

Robust survey estimates of abundance, accurately documented catch data, and some length-composition data are minimum data requirements for the age-size-structured model. Survey data are important to inform the model when confounding parameters are simultaneously estimated.

Catch is assumed to be known exactly for the age-size-structured model, so knowledge of catch is crucial to achieving robust estimates of population biomass. It would be inappropriate to apply this assessment method in conditions under which catch is highly uncertain, such as for some data-poor stocks. Length-composition data are also essential because the population model is based on a size-transition matrix, for which there must be some information as to how animals grow through time.

Integrated models, such as the age-size-structured model evaluated here, combine multiple sources of information into a single analysis (Maunder and Punt, 2013). A number of challenges exist in creating, using, evaluating, and implementing such models. Computational requirements can limit the level of complexity for which an integrated model can be conducted (Maunder and Punt, 2013), although improvements in computer resources are constantly emerging. Model mis-specification, complexity, convergence, and confounding parameters can make using integrated models difficult (Maunder and Punt, 2013). Integrated models require considerable amounts of data, and models with length-based dynamics perform poorly if the size-transition matrix is uncertain (Punt, 2003). Contradictory data can also present major problems when trying to integrate multiple data sources into a single analysis (Punt et al., 2013). Multiple data sources can also be problematic when trying to weight data (Maunder and Punt, 2013). The level of complexity and many potential problems associated with integrated analysis can deter potential users, and result in assessments that are not transparent to biologists and policy-makers (Hilborn, 1997).

When assessments do use integrated analysis, often age-structured or size-structured models are chosen, while integrated age-size-structured models are rarely implemented. Problems arise when having to choose between age- or size-structured models because fishery processes have to be either age- or length-based. This assumption can be violated, for example, when animals mature or when they migrate ontogenetically (Punt et al., 2013), or when estimation of year-class strength from size data alone requires adequate characterization of growth (Punt, 2003).

General assessment frameworks, such as CASAL (Bull et al., 2012), Stock Synthesis (Methot, 2013), and MULTIFAN-CL (Kleiber et al., 2012) provide age- and size-structured options, but without fully integrating age-size-dynamics, as is done here. CASAL is a generalized platform for assessment that allows for either an age-structured or size-structured model to be used (Bull et al., 2012). Problems can occur, for example, when an age-structured population also exhibits length-based traits, such as when growth varies over time or survival is length-based. Stock Synthesis is based in an age-structured model, but includes length capabilities by allowing for platoon structure (Methot, 2013). However, Punt et al. (2017) examined the impact of model mis-specification for age-structured, size-structured, and age-size-structured models, and found that excluding platoon structure in the assessment, even when warranted, resulted in the best estimation approaches. MULTIFAN-CL implements a size-based, age-structured and spatially-structured model for assessment, which models population abundance by age class and region, and is fit to catch and size-composition data (Kleiber et al., 2012). Strong assumptions must be made that the age-at-size relationship is known and unchanging through time when setting up an age-structured model, but fitting to catch and length data.

Several assessment methods have attempted to integrate age-size-dynamics, although many with either very specific applications or limited scope (De Leo and Gatto, 1995; Deriso and Parma, 1998; Quinn et al., 1998; Gilbert et al., 2006). In terms of generalized stock assessments, Fleksibest (Frøysa et al., 2002) is one of the only options that exists for single-stock integrated age-size-structured assessment until now. Fleksibest structures population size at any given point as an abundance matrix, or the number of fish in each age and length group (Frøysa et al., 2002).

Fleksibest arose due to age being a poor proxy for animal size in a boreal system, where large inter-annual variation led to highly variable size-at-age (Frøysa et al., 2002). In addition, Fleksibest considers mature and immature animals, models fishing mortality, selectivity, natural mortality based on length, and uses data from catches, surveys, age- and length-compositions, and diet data (Frøysa et al., 2002). Unlike the age-size-structured model developed here, Fleksibest estimates catches instead of assuming they are known exactly, and finds, in an assessment for Northeast Arctic cod (*Gadus morhua*), that there is more dynamic stock development, although catch estimates are larger than the data would suggest but fits to survey index are better (Frøysa et al., 2002). The age-size-structured model here assumes perfect reporting of catch because misreporting catch is illegal for US stocks

The age-size-structured model could be used to conduct assessments for stocks with known age- and size-dynamics, as well as act as an alternative assessment to which current official assessments can be compared. The model framework includes options for an age-structured assessment, a size-structured assessment, and an age-size-structured assessment, so hypothesis testing for stocks that are currently modelled in age- or size-structured models, but that may also have size- or age-dynamics (respectively), could be easily conducted given that the likelihood function remains the same throughout. Allen Akselrud et al. (2017) explored some of these options for the Pacific cod population, finding that variation in estimates of population biomass is larger among model types than within model types, suggesting that assessment scientists may need to present results for multiple model types, rather than only conducting sensitivities for one specific model. While this is an important result, future work should include a broader analysis for the magnitude and direction of change in population estimates using all three model forms across multiple taxa, and across multiple regions. This future work could serve as a reference for fishery assessment scientists as to how estimates of biomass might change due to model mis-specification of age- or length-dynamics.

References

Alaska Board of Fisheries. 2011. King and Tanner Crab, and supplemental issues. March 22-26, 2011.

Allen Akselrud, C.I., Punt, A.E. and L. Cronin-Fine. 2017. Exploring model structure uncertainty using a general stock assessment framework: the case of Pacific cod in the Eastern Bering Sea. *Fisheries Research* 193: 104-120.

Beverton, R.J.H. and S.J. Holt. 1957. *On the dynamics of exploited fish populations*, Fishery Investigations Series II Volume XIX, Ministry of Agriculture, Fisheries and Food, UK.

Boehlert, G.W. 1985. "Using objective criteria and multiple regression models for age determination in fishes." *Fishery Bulletin* 83: 103-117.

Buckland, S.T., Burnham, K.P. and N.H. Augustin. 1997. Model selection: an integral part of inference. *Biometrics* 603-618.

Bull, B., Francis, R.I.C.C., Dunn, A., McKenzie, A., Gilbert, D.J., Smith, M.H., Bian, R. and D. Fu. 2012. CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual v2. 30-2012/03/21. *NIWA Technical Report 135*: 280.

Carver, A.M., Wolcott, T.G., Wolcott, D.L. and A.H. Hines. 2005. Unnatural selection: effects of a male-focused size-selective fishery on reproductive potential of a blue crab population. *Journal of Experimental Marine Biology and Ecology*, 319(1): 29-41.

Canino, M. F., Spies, I. B. and L. Hauser. 2005. Development and characterization of novel di- and tetranucleotide microsatellite markers in Pacific cod (*Gadus macrocephalus*). *Molecular Ecology Notes* 5: 908-910.

Canino, M.F., Spies, I.B., Cunningham, K M., Hauser, L. and W. S. Grant. 2010. Multiple ice-age refugia in Pacific cod, *Gadus macrocephalus*. *Molecular Ecology* 19: 4339-4351.

Chilton, D.E. and R.J. Beamish. 1982. *Age determination methods for fishes studied by the groundfish program at the Pacific Biological Station*. Canadian Special Publication of Fisheries and Aquatic Sciences 60. Department of Fisheries and Oceans. Ottawa, Ontario, Canada.

Cohen, D. M., Inada, T. and N. Scialabba. 1990. FAO species catalogue. Vol. 10. Gadiform fishes of the world (Order Gadiformes). An annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to data. *FAO Fisheries Synopsis* 125: 452.

Conover, D.O. and S.B. Munch. 2002. Sustaining fisheries yields over evolutionary time scales. *Science* 297: 94-96.

Cunningham, K. M., Canino, M. F., Spies, I. B. and L. Hauser. 2009. Genetic isolation by distance and localized fjord population structure in Pacific cod (*Gadus macrocephalus*): limited effective dispersal in the northeastern Pacific Ocean. *Canadian Journal of Fisheries and Aquatic Sciences* 66: 153-166.

De Leo, G.A. and M. Gatto. 1995. A size and age-structured model of the European eel (*Anguilla anguilla* L.) *Canadian Journal of Fisheries and Aquatic Sciences* 52: 1351–1367.

Deriso, R.B. and A.M. Parma. 1988. Dynamics of age and size for a stochastic population model. *Canadian Journal of Fisheries and Aquatic Sciences* 45: 1054–1068.

Donaldson, W.E., Cooney, R. T. and J. R. Hilsinger. 1981. Growth, age and size at maturity of Tanner crab, *Chionoecetes bairdi* MJ Rathbun, in the northern Gulf of Alaska (Decapoda, Brachyura). *Crustaceana* 40.3: 286-302.

Fenberg, P.B. and K. Roy. 2008. Ecological and evolutionary consequences of size-selective harvesting: how much do we know? *Molecular Ecology 17*(1): 209-220.

Fournier, D.A. and C.P. Archibald. 1982. A general theory for analysing catch-at-age data. *Canadian Journal of Fisheries and Aquatic Sciences* 39: 1195–1207.

Frøysa, K.G., Bogstad, B. and D.W. Skagen. 2002. Fleksibest—an age–length structured fish stock assessment model. *Fisheries Research*, *55*(1): 87-101.

Fry, F.E.J. 1949. Statistics of a lake trout fishery. *Biometrics* 5: 27-67.

Garcia, S.M., Kolding, J., Rice, J., Rochet, M.J., Zhou, S., Arimoto, T., Beyer, J.E., Borges, L., Bundy, A., Dunn, D. and E.A. Fulton. 2012. Reconsidering the consequences of selective fisheries. *Science* 335: 1045-1047.

Gilbert, D.J., Davies, N.M. and J.R. McKenzie. 2006. Development of an age-length structured model of the Hauraki Gulf-Bay of Plenty snapper (*Pagrus auratus*) population. *Marine and Freshwater Research* 57: 553–568.

Haddon, M. 2010. *Modelling and quantitative methods in fisheries*. CRC press. Boca Raton, FL, USA.

Harley, S.J., Myers, R.A. and A. Dunn. 2001. Is catch-per-unit-effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences* 58(9): 1760-1772.

Hilborn, R. 1998. Lobster stock assessment: report from a workshop; II. *Marine and Freshwater Research* 48(8): 945-947.

Hilborn, R. and M. Mangel. 1997. *The ecological detective: confronting models with data*. Princeton University Press. Princeton, NJ, USA.

Hilborn, R. and C.V. Minte-Vera. 2008. Fisheries-induced changes in growth rates in marine fisheries: are they significant? *Bulletin of Marine Science* 83(1): 95-105.

Hoenig, J. 1983. Empirical use of longevity data to estimate mortality rates. *Fishery Bulletin* 82: 898-903.

Hosie, M. J. and T. F. Gaumer. 1974. Southern range extension of the Baird crab *(Chionoecetes bairdi* Rathbun). California Department of Fish and Game, CA, USA 60: 44-47.

Hurtado-Ferro, F., Punt, A.E. and K.T. Hill. 2014. Use of multiple selectivity patterns as a proxy for spatial structure. *Fisheries Research* 158: 102–105

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

Karinen, J. F. and D. T. Hoopes. 1971. Occurrence of Tanner crabs (*Chionoecetes* sp.) in the eastern Bering Sea with characteristics intermediate between *C. bairdi* and *C. opilio*. *Proceedings of the National Shellfisheries Association* 61: 8-9.

Kleiber, P., Hampton, J., Davies, N., Hoyle, S., and D.A. Fournier. 2012. MULTIFAN-CL User's Guide. < http://www.multifan-cl.org/>

Kon, T. 1996. Overview of Tanner crab fisheries around the Japanese Archipelago. *High Latitude Crabs: Biology, Management and Economics* 13-24. Alaska Sea Grant Report, AK-SG-96-02, University of Alaska Fairbanks, AK, USA.

Lee, R. 1920. A review of the methods of age and growth determination in fishes by means of scales. Ministry of Agriculture and Fisheries, *Fisheries Investigations* 2: 1-32.

Magnusson, A. and R. Hilborn. 2007. What makes fisheries data informative? *Fish and Fisheries 8*(4): 337-358.

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

Methot, R.D. 2013. User manual for stock synthesis. NOAA Fisheries, Seattle, WA, USA.

Methot, R.D. and C.R. Wetzel. 2013. Stock Synthesis: Providing a biological and statistical framework for fishery management forecasts across a data-poor to data-rich continuum. *Fisheries Research* 142: 86–99.

Nevisi, A., Orensanz, J. M., Paul, A. J. and D. A. Armstrong. 1996. Radiometric estimation of shell age in *Chionoecetes* spp. from the eastern Bering Sea, and its use to interpret shell condition indices: preliminary results.. *High Latitude Crabs: Biology, Management and Economics* 389-396. Alaska Sea Grant Report, AK-SG-96-02, University of Alaska Fairbanks.

NMFS. 2004. Final environmental impact statement for Bering Sea and Aleutian Islands crab fisheries. National Marine Fisheries Service, P.O. Box 21668, Juneau, AK, USA..

NPFMC. 2007. Initial review draft environmental assessment, Amendment 24 to the fishery management plan for Bering Sea and Aleutian Islands King and Tanner crabs to revise overfishing definitions. North Pacific Fishery Management Council, 605 W. 4th Avenue, 306, Anchorage, AK, USA.

NPFMC. 2017. Fishery management plan for groundfish of the Bering Sea and Aleutian Islands management area. North Pacific Fishery Management Council, Anchorage, AK, USA.

Punt, A.E. 2003. The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. *Fisheries Research 65*(1): 391-409.

Punt, A.E. 2006. The FAO precautionary approach after almost 10 years: have we progressed towards implementing simulation-tested feedback-control management systems for fisheries management? *Natural Resource Modeling 19*(4): 441-464.

Punt, A.E. and R.B. Kennedy. 1997. Population modelling of Tasmanian rock lobster, *Jasus edwardsii*, resources. *Marine and Freshwater Research* 48: 967-980.

Punt, A.E., Smith, D.C., Tuck, G.N. and R.D. Methot. 2006. Including discard data in fisheries stock assessments: Two case studies from South-eastern Australia. *Fisheries Research* 79: 239-250.

Punt, A.E., Haung, T-C. and M.N. Maunder. 2013. Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. *ICES Journal of Marine Science* 70: 16-33.

Punt, A.E, Allen Akselrud, C.I. and L. Cronin-Fine. 2017. The effects of applying mis-specified age- and size-structured models. *Fisheries Research* 188: 58-73.

Quinn, T.J., Turnbull, C.T. and C. Fu. 1998. A length-based population model for hard-to-age invertebrate populations. International Symposium on Stock Assessment Models for the 21st Century. Anchorage, AK, USA.

Ralston, S., Punt, A.E., Hamel, O.S., DeVore, J.D. and Conser, R.J. 2011. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. *Fishery Bulletin*. 109(2): 217-232.

Rugolo L,J. and B.J. Turnock. 2010. 2010 Stock assessment and fishery evaluation report for the Tanner crab Fisheries of the Bering Sea and Aleutian Islands regions. Draft Report to the North Pacific Fishery Management Council Crab Plan Team, AK, USA 61.

Rugolo, L.J. and B.J. Turnock. 2011. Length-based stock assessment model of eastern Bering Sea Tanner crab. Report to Subgroup of NPFMC Crab Plan Team, AK, USA 61.

Rugolo, L.J. and B.J. Turnock. 2012. Length-based stock assessment model of eastern Bering Sea Tanner crab. Report to Subgroup of NPFMC Crab Plan Team, AK, USA 69.

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. PhD Thesis, University of Washington, WA, USA 220.

Spies, I. 2012. Landscape genetics reveals population subdivision in Bering Sea and Aleutian Islands Pacific cod. *Transactions of the American Fisheries Society* 141: 1557-1573.

Stark, J. W. 2007. Geographic and seasonal variations in maturation and growth of female Pacific cod (*Gadus macrocephalus*) in the Gulf of Alaska and Bering Sea. *Fishery Bulletin* 105: 396-407.

Stockhausen, W.T., Turnock, B.J. and L.J. Rugulo. 2013. Draft 2013 Stock assessment and fishery evaluation report for the Tanner crab fisheries of the Bering Sea and Aleutian Islands regions. In: Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. North Pacific Fishery Management Council, Anchorage, AK, USA 342–449.

Stockhausen, W.T. 2015. 2015 Stock assessment and fishery evaluation report for the Tanner crab fisheries of the Bering Sea and Aleutian Islands regions. Pg. 295-440. In: Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. North Pacific Fishery Management Council, Anchorage, AK, USA.

Sullivan, P.J., Lai, H-L. and V.F. Gallucci. 1990. A catch-at-length analysis that incorporates a stochastic model of growth. *Canadian Journal of Fisheries and Aquatic Sciences* 47: 184–198.

Tamone, S. L., S. J. Taggart, A. G. Andrews, J. Mondragon, and J. K. Nielsen. 2007. The relationship between circulating ecdysteroids and chela allometry in male Tanner crabs: Evidence for a terminal molt in the genus *Chionoecetes*. *Journal of Crustacean Biology* 27: 635-642.

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

Thompson, G.G. 2015. 2015 Assessment of the Pacific Cod Stock in the Eastern Bering Sea. In: 2015 North Pacific Groundfish Stock Assessment and Fishery Evaluation Reports for 2016. North Pacific Fishery Management Council, Anchorage, AK, USA 251-470.

Thompson, G., Ianelli, J. and R. Lauth. 2010. Assessment of the Pacific cod stock in the Eastern Bering Sea and Aleutian Islands Area. Plan Team for Groundfish Fisheries of the Bering Sea/Aleutian Islands (compiler), Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions. North Pacific Fishery Management Council, Anchorage, AK, USA 243-424.

Ueda, Y., Narimatsu, Y., Hattori, T., Ito, M., Kitagawa, D., Tomikawa, N. and T. Matsuishi. 2006. Fishing efficiency estimated based on the abundance from virtual population analysis and bottom-trawl surveys of Pacific cod (Gadus macrocephalus) in the waters off the Pacific coast of northern Honshu, Japan. *Nippon Suisan Gakkaishi* 72: 201-209.

von Bertalanffy, L. 1934. *Modern theories of development (JH Woodger, trans.).* Oxford University Press. Oxford, UK

Supplemental Files

File 1. Appendix 1: Specifications for the estimation model

File 2. Appendix 2: Supplemental tables and figures for Chapters 2 and 3.