

ESTIMABILITY OF TIME-VARYING NATURAL MORTALITY IN GROUND FISHES:  
COVARIATES AND HIERARCHICAL MODELS

By

Philip D. Ganz, B.S.

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Fisheries

University of Alaska Fairbanks

August 2017

APPROVED:

Dr. Terrance J. Quinn II, Committee Chair  
Dr. Peter-John F. Hulson, Committee Member  
Dr. Gordon H. Kruse, Committee Member  
Dr. Gordon H. Kruse, Chair  
*Graduate Program in Fisheries*  
Dr. S. Bradley Moran, Dean  
*College of Fisheries and Ocean Sciences*  
Dr. Michael A. Castellini,  
*Dean of the Graduate School*

## Abstract

Natural mortality,  $M$ , has historically been a difficult parameter to estimate in conjunction with other stock assessment parameters. Time-varying  $M$ , while likely to be experienced by a population, is a particularly difficult process to estimate with the data and methods currently available to most stock assessments. Although auxiliary information in the form of a covariate to  $M$  has been shown to improve model fit for some stocks, such data are rarely available. Meanwhile, hierarchical models continue to be utilized in capturing processes that vary in time and space. I tested both the covariate and hierarchical methods in their ability to estimate time-varying  $M$ . I attempted to fit hierarchical models by two different methods: penalized likelihood and the integrated likelihood approach associated with mixed effects models. Mixed effects models performed poorly in comparison to penalized likelihood. Including a covariate to natural mortality aided the estimability of time-varying  $M$ , regardless of the observation error associated with the covariate. Estimating a constant value of  $M$  resulted in biased estimates when  $M$  was time-varying in the simulated population. I showed that the Akaike information criterion (AIC) is a useful metric for comparing models although it does not necessarily align with the accuracy of estimates that are of most interest to managers, such as terminal year spawning stock biomass. In addition to showing empirically that incorporating a covariate is a robust approach to estimating time-varying  $M$ , I conclude that this approach is also advantageous to stock assessment on theoretical grounds, as it is more amenable than hierarchical models to making predictions.



## Table of Contents

Title Page .....	i
Abstract .....	iii
List of Figures .....	vii
List of Tables .....	ix
Acknowledgements .....	xi
1. Introduction .....	1
1.1 The problem .....	1
1.2 Existing knowledge .....	3
1.3 Potential solutions .....	4
1.3.1 Covariates .....	4
1.3.2 Hierarchical models .....	5
1.4 Objectives of this study .....	6
2. Methods .....	7
2.1 Simulation .....	7
2.2 Estimation .....	12
2.2.1 Covariates .....	12
2.2.2 Random effects: uncorrelated process .....	13
2.2.3 Random effects: correlated process .....	14
2.3 Objective function .....	15

2.4	Model scenarios.....	18
2.5	Model performance criteria .....	23
2.6	Sensitivity study .....	24
3.	Results.....	25
3.1	Penalized likelihood and random effects performance.....	25
3.2	The influence of population dynamics and data quality on estimation .....	27
3.2.1	Species .....	27
3.2.2	Natural mortality trend.....	29
3.2.3	Survey biomass variability.....	30
3.2.4	Ratio of fishing to natural mortality.....	31
3.3	Evaluating the performance of different estimation methods .....	32
3.3.1	All estimation methods.....	32
3.3.2	Sensitivity study.....	38
3.3.3	Model comparison .....	39
4.	Discussion.....	42
5.	Conclusions.....	45
6.	References.....	47

## List of Figures

	Page
Figure 3.1 Convergence proportions after an initial 25 runs for models fit by penalized likelihood and random effects.....	25
Figure 3.2 Simulated and estimated values for natural mortality by species and likelihood method.....	26
Figure 3.3 Relative errors (REs) in spawning stock biomass for pollock and sablefish..	28
Figure 3.4 Relative errors of parameters for pollock and sablefish.....	29
Figure 3.5 Relative errors (REs) in spawning stock biomass by species and natural mortality trend.....	30
Figure 3.6 Relative errors (REs) in spawning stock biomass by species and survey precision..	31
Figure 3.7 Relative errors (REs) in spawning stock biomass by species and ratio of fishing mortality to natural mortality.....	32
Figure 3.8 Relative errors (REs) in spawning stock biomass by species and estimation model for the case in which natural mortality is constant over time.....	33
Figure 3.9 Relative errors (REs) in spawning stock biomass by species and estimation model for the case in which natural mortality increases linearly over time.....	34
Figure 3.10 Relative errors (REs) in spawning stock biomass by species and estimation model for the case in which natural mortality decreases linearly over time.....	34
Figure 3.11 Relative errors (REs) in spawning stock biomass by species and estimation model for the case in which natural mortality has a normal random distribution. ....	35
Figure 3.12 Relative errors in estimates of terminal year spawning stock biomass by species and natural mortality trend.....	36
Figure 3.13 Relative errors in estimates of mean natural mortality by species and natural mortality trend.....	37
Figure 3.14 Relative errors in estimates of survey catchability by species and natural mortality trend. ....	38
Figure 3.15 Relative errors (REs) in spawning stock biomass by species and covariate precision.....	39

Figure 3.16 Proportion of instances in which AIC selected each estimation model.. ..... 40

Figure 3.17 Median  $\Delta$ AIC values for each estimation model by natural mortality trend..... 41

Figure 3.18 Median relative errors (REs) in spawning stock biomass for each estimation model..  
..... 42

## List of Tables

	Page
Table 2.1 Notation used in model description. ....	9
Table 2.2 Description of population dynamics equations used in the operating and estimation models. ....	11
Table 2.3 Parameter values used in the operating models. ....	20
Table 2.4 Quantities for proportion mature and weight-at-age. ....	21
Table 2.5 Operating model cases, having a total of 32 possible combinations. ....	22





## Acknowledgements

This publication is the result in part of research sponsored by the Cooperative Institute for Alaska Research with funds from the National Oceanic and Atmospheric Administration under cooperative agreement NA13OAR4320056 with the University of Alaska. The source of funds was the Stock Assessment Improvement Program of the National Marine Fisheries Service.

I would like to acknowledge and thank my committee for guiding me through the process of writing this thesis. Dr. Terrance Quinn II, I'm grateful for your hands-off, but open door approach. Thank you for encouraging my writing and for giving me opportunities to teach. Your wicked sense of humor was greatly appreciated, especially when I crashed your car in Chile. In my defense, that other guy was going way too fast.

Dr. Peter Hulson, it is hard to believe I was your first student. You have a knack for mentoring and I think science will only stand to benefit if you continue to guide aspiring scientists. Thank you for keeping me sane by listening to my frustrations with all things programming, even if you could not always help.

Dr. Gordon Kruse, I checked in with you least often, but was consistently amazed by your ability to dive into my thesis and give incredibly insightful feedback. Thank you for the meticulous edits. Thank you also for teaching an extremely educational course on the operations of the North Pacific Fisheries Management Council. I found this course invaluable for understanding how natural resource management actually works.



## 1. Introduction

### 1.1 The problem

The fundamental objective of most fisheries stock assessments is to estimate population abundance and related parameters, to inform management actions. As stock assessment methodology and computational power have progressed over decades, many methods of varying complexity have been developed to achieve this end. Simpler models do not account for population dynamics explicitly, but instead consider changes in the population as a whole. Surplus production models (Quinn and Deriso 1999) are an example of this approach, as they model the weight (biomass) of total growth against total subtractions rather than accounting directly for all the processes that affect growth (recruitment, individual growth) and subtractions (natural mortality, fishing mortality).

The current gold standard in fisheries stock assessments is the age-structured assessment (ASA) model. This model lies on the other end of the complexity spectrum, accounting for age structure, multiple sources of mortality, and recruitment. These models commonly attribute the total mortality ( $Z$ ) experienced by a population to two sources: fishing mortality ( $F$ ) and natural mortality ( $M$ ). Estimates of  $F$  are informed by the catches reported in the various fisheries present. Fishing mortality is expressed as a function of age-specific gear selectivity ( $s$ ), with full-recruitment  $F$  experienced by fully-selected age classes. In parsing natural mortality from total mortality, data directly related to natural mortality are rarely available. Thus,  $M$  tends to be confounded with  $F$  (Quinn and Deriso 1999).  $M$  also has a tendency to be confounded with gear selectivity, catchability ( $q$ ), and recruitment ( $R$ ), (Thompson 1994, Schnute and Richards 1995) as all of these parameters affect population scale and change over time. These issues of

confounding hinder the ability of models to accurately estimate  $M$  simultaneously with other parameters. It is therefore commonplace for stock assessments to estimate  $M$  separate from the assessment model, and then to fix  $M$  at that single value or set of values that vary with age or sex (Quinn and Deriso 1999). This likely does not represent the true underlying dynamics, as one would expect  $M$  to not only vary with age but also with size, time, and due to density-dependent effects (Beverton and Holt 1957).

Although common practice, treating natural mortality as a constant comes with risk. Simulations have shown that misspecification of  $M$  results in bias and that high harvest recommendations can result when  $M$  is set constant at a rate higher than that experienced by the population (Clark 1999). The collapse of red king crab (*Paralithodes camtschaticus*) in Bristol Bay, Alaska may have been due in part to high rates of natural mortality; allowing for time-varying natural mortality significantly improved the assessment model fit to survey abundance when compared to a model that assumed constant natural mortality (Zheng et al. 1995). In this case, allowing for time-varying  $M$  improved the quality of results that could be used for management. However, time-varying  $M$  is not always estimable due to the reasons discussed above and limited available data.

The difficulty in estimating  $M$ , combined with potential bias associated with model misspecification, prompted the National Marine Fisheries Service (NMFS) to hold a workshop on the topic of estimating natural mortality in 2011. The resultant report stressed the difficulty in estimating  $M$  and its influence on management reference points, and defined a set of best practices regarding the treatment of  $M$  within stock assessment models (Brodziak et al. 2011). Central to the best practices proposed is a tier system for categorizing which approaches for

estimating  $M$  are possible for a given stock based on data available. Estimates involving greater complexity—such as spatially or temporally varying  $M$ —are placed in the tier corresponding to the most data-rich stocks. The report encouraged the development of methods that account for the effect of temporal changes in  $M$  on management reference points.

## 1.2 Existing knowledge

Several studies have investigated alternative approaches for addressing time-varying  $M$ . Time-varying  $M$  was shown to be estimable in a length-based assessment with precise survey data, and this was the preferred method when true natural mortality is time-varying (Fu and Quinn 2000). However, estimating a constant value for  $M$  can be a robust approach in an ASA model when  $M$  experiences a sharp increase or decrease in time (Johnson et al. 2015).

Allowing parameters to vary over time is not only of interest because it may result in more accurate estimates of the population, but acknowledging these additional dynamics may allow for a more accurate accounting of the uncertainty in a model's predictions and resultant management recommendations (Hilborn 2003). Accounting for this additional uncertainty is of particular importance for a parameter as influential in an assessment as  $M$ . Further, if trends in  $M$  are detected by a model, one can then begin to form hypotheses about potential biological or environmental drivers of those trends (Zheng et al. 1995). However, there are limits to the number of parameters that can be estimated by models, so tradeoffs must be recognized when complexity is increased.

Natural mortality is also just one population process that is likely to be time-varying. Individual growth rates can change over time, as has been observed with Pacific halibut (*Hippoglossus stenolepis*, Clark et al. 1999). Selectivity can in turn be affected by changes in growth (Clark et al. 1999) and by modifications to fishing gear. Catchability can vary with factors including

density and temperature, as has been observed with cape hake (*Merluccius capensis*, Gordo and Hightower 1991).

### 1.3 Potential solutions

Two general approaches have been proposed to make time-varying  $M$  more broadly estimable.

One approach is to include auxiliary information related to  $M$  into the model as a covariate to the parameter. The other approach is to estimate time-varying deviations in  $M$  as parameters and to make assumptions regarding the error structure of  $M$  within the model.

#### 1.3.1 Covariates

The theoretical basis for including covariates in stock assessment models is to reduce uncertainty in parameter estimates associated with key population processes such as growth, mortality, or recruitment. Typically, the correlation between an environmental time-series and a population process is estimated separately from all other model parameters. However, although correlations are often identified in this way, it is rare that these relationships are formally incorporated into an assessment, because it has been found that such relationships often break down over time (Myers 1998).

The alternative is to estimate these relationships internally in the model in conjunction with all other model parameters. A typical ASA model utilizes maximum likelihood estimation and allows for the incorporation of multiple data sets, making them very amenable to the inclusion of environmental covariates (Maunder and Watters 2003, Deriso et al. 2008). This approach has been applied experimentally with sablefish (*Anoplopoma fimbria*) recruitment in the Gulf of Alaska (Shotwell et al. 2012). It has also been applied in relating multiple covariates, including indices of disease (Marty et al. 2010) and humpback whale (*Megaptera novaeangliae*)

abundance (Teerlink 2011) to the natural mortality of Pacific herring (*Clupea pallasii*) in Prince William Sound, Alaska. Including each of these covariates was shown to improve model fit, and indices of disease are now included in the stock assessment for herring in Prince William Sound. I test the same methodology, estimating time-varying natural mortality as a function of a covariate during the estimation process of an ASA.

### 1.3.2 Hierarchical models

Modeling time-varying natural mortality within an ASA using hierarchical models contrasts with the covariate approach in that no auxiliary information needs to be included. Instead, an additional assumption is made regarding the error structure of a dependent variable or parameter of interest, in this case  $M$ . This often replaces the assumption of independence with one of temporal correlation among unobserved states (Thorson and Minto 2015). Hierarchical models can be structured in various ways that differ in the statistical distribution, objective function, and fitting method used.

Mixed effects models are perhaps the most recognized example of hierarchical model structure. A mixed effects model is any model that contains both fixed and random effects. Random effects approaches have been applied in the estimation of various stock assessment parameters.

Modeling time-varying catchability as a random walk is a robust estimation method across various states of nature (Wilberg and Bence 2006). Movement can be modelled as a random walk in an ASA model (Hulson et al. 2013). Recruitment has been modeled using uncorrelated random effects and total mortality as a random walk (Trenkel et al. 2012). A Bayesian ASA model with random effects was able to detect trends in the natural mortality of Atlantic weakfish (*Cynoscion regalis*) over time (Jiao et al. 2012). The authors emphasized that, beyond simply



detecting trends in  $M$ , it is important to be able to predict changes in  $M$  so that a changing fishery can be managed more effectively. Toward this end, the detected trends in  $M$  were compared to long-term indices of climate–ocean oscillation, and a correlation was identified. Although the authors did not incorporate this relationship into the ASA, they noted that doing so is another potential method of predicting changes in  $M$ .

Another example of hierarchical models are those fitted by penalized likelihood (De Valpine and Hilborn 2005). Penalized likelihood models, like mixed effects models, are hierarchical in the sense that they make a distributional assumption regarding the deviations of a variable or a parameter. The core difference between mixed effects and those fit by penalized likelihood is in the objective function that is maximized.

#### 1.4 Objectives of this study

Following the recommendations of Brodziak et al. (2011) and Jiao et al. (2012), I investigated the estimability of natural mortality under a wide range of estimation approaches and population characteristics. The objective of my study is to compare performance between dominating methods of accounting for time-varying processes in their ability to estimate time-varying  $M$ . I expected estimation to be affected by both underlying population characteristics and the particular estimation methodology being applied. However, I also hypothesized that there are factors, such as the precision of survey data, which broadly influence model performance across conditions. Similarly, I hypothesized that some estimation methodologies, such as inclusion of a precise covariate, will consistently outperform others across underlying conditions. I used the simulation–estimation approach identified in Fu and Quinn (2000) in evaluating model performance. Computer simulation has been widely used in fisheries to evaluate the ability of

models to quantify population dynamics and the potential impacts of different management actions (Walters 1969, Hilborn and Walters 1987, Schnute and Richards 1995). Qualitative aspects of model performance were investigated graphically. Relative errors and a formal model selection criterion were used in quantitatively evaluating model performance.

## 2. Methods

### 2.1 Simulation

The basic simulation-estimation framework involves creating an operating model (OM) that simulates the underlying dynamics of the system of interest, in this case a population, and then testing the precision and accuracy with which an estimation model (EM) is able to estimate those same processes. The EM requires data that are generated with error from the known dynamics of the OM. Here, I consider the step of data generation to be part of the OM.

The OM and corresponding EM are ASA models, which originate from work by Fournier and Archibald (1982) and Deriso et al. (1985). In addition to general structure, parameterization between the OMs and EMs is largely the same. Both models structure survival ( $S$ ) as an exponential function of instantaneous rates of  $F$  and  $M$ . Year and age specific fishing mortality ( $F_{t,a}$ ) is considered separable between annual components of fishing mortality ( $F_t$ ) and selectivity-at-age ( $s_a$ ), the latter of which is assumed to remain constant over time. A logistic function is used to describe the age-specific selectivity with which survey and fishing gears catch individuals. Both models structure annual recruitment and fishing mortality as means ( $\mu_R$  and  $\mu_F$ , respectively) with annual deviations ( $\tau_t$  and  $\phi_t$ , respectively). Examples where the OM and EM differ include cases for the estimation of  $M$  in which mismatches are of particular interest (e.g., the case in which true  $M$  is time-varying, but estimated as a constant). Estimates are made for the

30 years of data simulated by the OM, with no projections made beyond those years. The notation used in describing both models is summarized in Table 2.1. The structure of the OM is summarized in column 1 of Table 2.2.

Stochastic process error is incorporated into the OM in the way that recruitment, fishing mortality, and initial abundance are generated. Observation error is generated in order to produce observed age composition and abundance data from known values. It is commonly assumed that age composition data have a multinomial distribution, while error in catch and survey abundance data is lognormally distributed (Quinn and Deriso 1999). I simulated observation error using R functions that draw random numbers from the distributions assumed for these data types. Unique draws were taken in each of the 100 replications. The stochastic elements of the operating model and the distributions used to simulate them are summarized in column 1 of Table 2.2. The variability of the distributions is defined by the sample size and CV values listed in Table 2.3.

Each OM also contains deterministic components. These include the processes of selectivity and survey catchability ( $q$ ). Maturity and weight-at-age are also incorporated deterministically and are discussed further in section 2.4.

Table 2.1 Notation used in model description.

Symbol	Description
$t$	Time (year), $t=1,2,\dots,T$
$T$	Terminal year of the model
$a$	Model age class, $a=a_0, a_0+1, \dots, a_+$
$a_0$	Age at recruitment to the model
$a_+$	Oldest age considered plus all older ages
$w_a$	Average weight at age $a$
$m_a$	Proportion of females mature at age $a$
$\mu_R$	Average log-recruitment
$\mu_F$	Average log-fishing mortality
$\phi_t$	Annual fishing mortality deviation
$\tau_a$	Initial log-abundance at age $a$
$\tau_t$	Annual recruitment deviation
$\sigma_R$	Recruitment coefficient of variation
$\sigma_F$	Fishing mortality coefficient of variation
$\sigma_\delta$	Covariate process error standard deviation for white noise
$\sigma_\gamma$	Covariate observation error standard deviation
$\sigma_C$	Catch coefficient of variation
$\sigma_I$	Survey coefficient of variation
$N_{t,a}$	Numbers of fish at age $a$ in year $t$
$R_t$	Recruitment ( $N_{t,a_0}$ ) in year $t$
$F_{t,a}$	Fishing mortality for year $t$ , age class $a$
$M_t$	Natural mortality for year $t$
$Z_{t,a}$	Total mortality for year $t$ and age class $a$ ( $=F_{t,a}+M_t$ )
$s_a$	Selectivity at age $a$
$a_{50\%}$	Age at 50% selection
$\delta$	Slope parameter for logistic selectivity curve
$q$	Abundance index catchability
$C_t$	Catch biomass in year $t$
$\omega_t$	Observation error for catch biomass in year $t$

Table 2.1 cont.

$I_t$	Survey biomass in year $t$
$v_t$	Observation error for survey biomass in year $t$
$\hat{\phantom{a}}$	Estimated value

---

Table 2.2 Description of population dynamics equations used in the operating and estimation models.

Operating Model	Description	Estimation model
$R_t = e^{\mu_R + \tau_t - \frac{\sigma_R^2}{2}}$	Recruitment in year $t$	$\hat{R}_t = e^{\mu_R + \tau_t}$
$\tau_t \sim N(0, \sigma_R^2)$		
$N_{1,a} = e^{(\mu_R + \tau_a - \frac{\sigma_R^2}{2})} e^{-(a-a_0)M}$	Numbers at ages $a_0 < a$ in year $t=1$	$\hat{N}_{1,a} = e^{\tau_a}$
$N_{1,a_+} = e^{\mu_R} e^{-(a-a_0)M} (1 - e^{-M})^{-1}$		
$\tau_a \sim N(0, \sigma_R^2)$		
$N_{t,a} = N_{t-1,a-1} e^{-Z_{t-1,a-1}}$	Numbers at ages $a_0 < a < a_+$ in years $t > 1$	$\hat{N}_{t,a} = N_{t-1,a-1} e^{-Z_{t-1,a-1}}$
$N_{t,a_+} = N_{t-1,a_+-1} e^{-Z_{t-1,a_+-1}} + N_{t-1,a_+} e^{-Z_{t-1,a_+}}$	Numbers at ages $a = a_+$ in year $t$	$\hat{N}_{t,a_+} = N_{t-1,a_+-1} e^{-Z_{t-1,a_+-1}} + N_{t-1,a_+} e^{-Z_{t-1,a_+}}$
$s_a = (1 + e^{-\delta(a-a_{50\%})})^{-1}$	Logistic selectivity at age $a$	$\hat{s}_a = (1 + e^{-\delta(a-a_{50\%})})^{-1}$
$F_t = e^{\mu_F + \phi_t - \frac{\sigma_F^2}{2}}$	Fishing mortality in year $t$	$\hat{F}_t = e^{(\mu_F + \phi_t)}$
$\phi_t \sim N(0, \sigma_F^2)$		
$F_{t,a} = s_a F_t$	Fishing mortality in year $t$ and age $a$	$\hat{F}_{t,a} = s_a F_t$
$C_t = \sum_{a_0}^{a_+} N_{t,a} F_{t,a} (1 - e^{-Z_{t,a}}) Z_{t,a}^{-1}$	Catch abundance in year $t$	$\hat{C}_t = \sum_{a_0}^{a_+} N_{t,a} F_{t,a} (1 - e^{-Z_{t,a}}) Z_{t,a}^{-1}$
$\tilde{C}_t = \left( \sum_{a_0}^{a_+} w_a N_{t,a} F_{t,a} (1 - e^{-Z_{t,a}}) Z_{t,a}^{-1} \right) e^{\omega_t - \frac{\sigma_C^2}{2}}$	Catch biomass in year $t$	$\hat{C}_t = \sum_{a_0}^{a_+} w_a N_{t,a} F_{t,a} (1 - e^{-Z_{t,a}}) Z_{t,a}^{-1}$
$\omega_t \sim N(0, \sigma_C^2)$		
$I_t = \sum_{a_0}^{a_+} q s_a N_{t,a}$	Survey abundance in year $t$	$\hat{I}_t = \sum_{a_0}^{a_+} q s_a N_{t,a}$
$\tilde{I}_t = \left( \sum_{a_0}^{a_+} w_a q s_a N_{t,a} \right) e^{\nu_t - \frac{\sigma_I^2}{2}}$	Survey biomass in year $t$	$\hat{I}_t = \sum_{a_0}^{a_+} w_a q s_a N_{t,a}$
$\nu_t \sim N(0, \sigma_I^2)$		
$\tilde{P}_{t,a} = \text{Multinomial} \left( N_{t,a} s_a \left( \sum_{a_0}^{a_+} N_{t,a} s_a \right)^{-1} \right)$	Proportion at age $a$ in year $t$	$\hat{P}_{t,a} = N_{t,a} s_a \left( \sum_{a_0}^{a_+} N_{t,a} s_a \right)^{-1}$

## 2.2 Estimation

Next, data generated by the OM are used by the EM to estimate population parameters (e.g.  $\mu$ ,  $M$ ,  $F$ ) and other values of interest, such as selectivity and spawning stock biomass (SSB).

Numerical estimation is performed in AD Model Builder (ADMB) version 11.6. ADMB optimizes using automatic differentiation; it is fast and stable in optimizing highly non-linear problems, including many fisheries stock assessments (Fournier et al. 2012). ADMB with random effects (ADMB-RE) is used for parameter estimation with mixed effects models.

In addition to the parameterization of natural mortality, there is a mismatch between the way that the EM and OM construct initial abundance. The OM simulates initial ( $t=1$ ) cohort abundances by essentially generating past stochastic recruitment in years before  $t=1$  and then progressing the cohorts to time  $t=1$  by applying  $M$  for the appropriate number of years. In contrast, the EM estimates initial log-abundance for each age as a free parameter. Furthermore, the estimation model makes no assumptions regarding the distributions of recruitment and fishing mortality, estimating these values as simple bounded vectors on the natural log scale that sum to 0. These differences, along with all other estimation model structure, are summarized in column 3 of Table 2.2. Parameterizations of  $M$  are discussed in sections 2.2.1, 2.2.2, and 2.2.3.

### 2.2.1 Covariates

I assume that there is a covariate,  $x_t$ , that is a measurable variable related to  $M$ . As in Marty et al. (2010),  $M$  is parameterized as a linear function of the covariate:

$$(1) \quad M_t = M_0 + \beta x_t,$$

in which  $M_0$  represents a base rate of natural mortality attributable to sources other than the mechanisms indicated by the covariate and  $\beta$  is the coefficient specifying the magnitude of the effect that the covariate,  $x_t$ , has on natural mortality. This relationship between the covariate and  $M$  was estimated during the process of fitting the ASA.

The covariate is generated as either a constant or linear function of time, or as normally distributed random values (Equations 2a, 2b, and 2c). The simulated values of natural mortality used in the OM are then generated according to Equation 1, so that  $x_t$  is a covariate to  $M_t$ . The covariate is observed with error, however, so the values actually passed to the model are

$$(2a) \quad x_t = a + bt + \gamma_t,$$

$$(2b) \quad x_t = a - bt + \gamma_t,$$

$$(2c) \quad x_t = a + \delta_t + \gamma_t,$$

in which  $\gamma_t \sim N(0, \sigma_\gamma^2)$  and represents observation error and  $\delta_t \sim N(0, \sigma_\delta^2)$  and represents process variability of the covariate (and therefore  $M_t$ ). A sensitivity analysis was performed on  $\sigma_\gamma$  to investigate the ways in which observation error in the covariate affects estimation. This study is detailed in section 2.6.

### 2.2.2 Random effects: uncorrelated process

I tested four random effects parameterizations in their ability to estimate  $M_t$ . The parameterization presented in this section assumes that values of  $M_t$  are uncorrelated. The three parameterizations presented in section 2.2.3 allow for correlation between values of  $M_t$ . As opposed to assigning the variation estimated in  $M$  to some environmental process, hierarchical models allow for variability in  $M$  without making any assumptions about the mechanism behind the variability. Time-varying  $M$  can be written as:



$$(3) \quad M_t = M_0 + \varepsilon_t \text{ and} \\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2),$$

where  $M_0$  is a parameter equal to the mean rate of  $M_t$  and  $\sigma_\varepsilon^2$  is the variance term for the annual uncorrelated deviations,  $\varepsilon_t$ . Following the convention of Wilberg and Bence (2006), I refer to this parameterization as “white noise”. I make the assumption that these annual deviations fall under a normal distribution, and a term for this distribution is added to the likelihood, as described in section 2.3.

### 2.2.3 Random effects: correlated process

Instead of assuming that parameter values are uncorrelated about a mean value, one can model an autocorrelated process known as a random walk in which the parameter value in the current year is the result of a random deviation from the value of that parameter in the previous year. Wilberg and Bence (2006) estimated time-varying catchability as a random walk. Fu and Quinn (2000) used the same methodology for estimating the logarithm of  $M_t$ .  $M_t$  represented as a random walk is written as:

$$(4) \quad M_t = M_{t-1} + \varepsilon_t \text{ and} \\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2).$$

A generalization of the random walk is an autocorrelated process referred to here as a correlated walk.  $M_t$  as a correlated walk is written:

$$(5) \quad M_t = \phi M_{t-1} + \varepsilon_t \text{ and} \\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2),$$

where  $\phi$  specifies the degree of correlation between successive estimates of  $M$ , reverting to a basic random walk when  $\phi = 1$ .

The random walk structure can also be modified by adding a constant term for the average distance of each annual step:

$$(6) \quad \begin{aligned} M_t &= M_{t-1} + \alpha + \varepsilon_t \text{ and} \\ \varepsilon_t &\sim N(0, \sigma_\varepsilon^2), \end{aligned}$$

where  $\alpha$  is known as drift and specifies the mean distance of each step. If  $\alpha = 0$ , the process reverts to a basic random walk.

### 2.3 Objective function

All EMs are fitted by maximum likelihood. Negative log-likelihoods are given in this section, as this is what ADMB minimizes in order to find a solution. Likelihood components are included for abundance data and age composition from both the survey and fishery, making a total of four likelihood components for non-hierarchical models. The likelihood components for the survey index and fishery catch assume the lognormal distribution:

$$(7) \quad \ell_1 = \sum_t (\ln C_t - \ln \hat{C}_t)^2 / (2\sigma_C^2) ,$$

$$(8) \quad \ell_2 = \sum_t (\ln I_t - \ln \hat{I}_t)^2 / (2\sigma_I^2) ,$$

where  $C_t$  and  $I_t$  are the observed and  $\hat{C}_t$  and  $\hat{I}_t$  are the predicted survey index and fishery catch, respectively, in year  $t$ .

The negative of the log-likelihood for age composition data comes from the multinomial distribution:

$$(9) \quad \ell_3 = -\sum_t \sum_a n_{a,t,f} \ln \hat{p}_{a,t,f} ,$$

$$(10) \quad \ell_4 = -\sum_t \sum_a n_{a,t,s} \ln \hat{p}_{a,t,s} ,$$

where  $n_{a,t}$  is the number of fish observed at age  $a$  from a sample with effective size  $n_t$  and  $\hat{p}_{a,t}$  is the estimated proportion of fish aged  $a$  in year  $t$ . Subscripts  $f$  and  $s$  denote fishery and survey, respectively. Methods for estimating effective sample size are an active area of research (Hulson et al. 2012, Thorson et al. 2017, Truesdell et al. 2017). However, for the purposes of this study, it is assumed that effective sample size is known.

Hierarchical models have a fifth component to the likelihood ( $\ell_5$ ), which is the prior or penalty associated with the normal distribution:

$$(11) \quad \ell_5 = \frac{1}{2} \sum_t (\varepsilon_t)^2 .$$

The objective function by which the model is fitted is the negative log-likelihood and is equal to the sum of all the relevant likelihood components:

$$(12) \quad -\ln(L) = \sum_i \ell_i .$$

The method by which the likelihood is maximized differs between fixed effect and mixed effects models. For fixed effect models, the maximum sought is for:

$$(13) \quad L(\boldsymbol{\theta} | \mathbf{D}) ,$$

where  $\theta$  represents all model parameters and  $\mathbf{D}$  represents the data. If deviations, such as those on natural mortality,  $\varepsilon_i$ , are included as fixed effects, the additional likelihood component (Equation 12) acts as a penalty. The likelihood being maximized in this case is therefore referred to as a “penalized likelihood”:

$$(14) \quad L(\theta, \varepsilon | \mathbf{D})P(\varepsilon | \tau).$$

where  $L(\theta, \varepsilon | \mathbf{D})$  is the joint likelihood of the parameters  $\theta$  and  $\varepsilon$  given the data  $\mathbf{D}$  and  $P(\varepsilon | \tau)$  represents the probability of the deviation parameters given the parameters that control their distribution,  $\tau$ . This differs from the likelihood maximized in the case that  $\varepsilon$  are treated as random effects. If included as random effects,  $\varepsilon$  are integrated out of the likelihood in order to produce the marginal likelihood:

$$(15) \quad \int_{\varepsilon} L(\theta, \varepsilon | \mathbf{D})P(\varepsilon | \tau)d\varepsilon.$$

Parameters that control the distribution of the random effects—such as  $\sigma_{\varepsilon}^2$  in this study—are referred to as hyper-parameters and are estimated when the marginal likelihood is maximized. When  $\varepsilon$  are instead included as fixed effects and the penalized likelihood is maximized,  $\sigma_{\varepsilon}^2$  is chosen *a priori* rather than being estimated by the model (Thorson et al. 2015). When fitted by penalized likelihood, the *a priori* assumption made in this study is that  $\sigma_{\varepsilon} = 0.5M_0$  for white noise and  $\sigma_{\varepsilon} = 0.1M_0$  for random walk parameterizations and its variants. Penalized likelihood methods have been shown to produce biased parameter estimates (De Valpine and Hilborn 2005). This result, along with the preference to minimize the number of *a priori* assumptions made by pulling inference about hyper-parameters directly from the data, has made mixed effects models the preferred hierarchical approach when tenable. However, approximating the integral

that produces the marginal likelihood is computationally intensive, and the difficulty of fitting complex mixed effects models has been well-documented (Bolker et al. 2013). Therefore, one aspect of this study is to compare the performance of penalized likelihood and mixed effects models.

## 2.4 Model scenarios

In addition to the assumptions on  $M$  and modeling approaches presented above, various scenarios are tested in order to determine how species life-history and fisheries dynamics influence the estimation of time-varying  $M$ . The true value of  $M$  is inherently related to the maximum age of the species, as longer-lived species must have a lower rate of mortality. Therefore, I hypothesize that estimability of time-varying  $M$  could be affected by the maximum age of the species for which it is being estimated. I simulate populations of Alaska sablefish and eastern Bering Sea walleye pollock (*Gadus chalcogrammus*), which have considerably different maximum ages—94 and 28 years, respectively, in Alaskan waters (Kimura et al. 1998, Munk 2001). Sablefish and walleye pollock are also chosen for this study because both species are data-rich. This quality would likely be important if time-varying  $M$  were to be estimated for a wild population.

Operating models take parameter values from existing assessments for Alaska sablefish (Hanselman et al. 2015) and eastern Bering Sea walleye pollock (Ianelli et al. 2015b). These values are found in Table 2.3. The OMs for this study are simplified versions of these full assessments, using parameters associated with one survey and one fishery for each species rather than the multiple surveys and fisheries used in the actual assessments. The assessment for sablefish also contains sex-specific parameters, which I do not consider. Sex-specific parameter values are instead averaged in the OM and estimated as a single value. Using R version 3.3.2 (R

Development Core Team 2016), 30 years of data are simulated for sablefish aged 2-31+ and walleye pollock aged 3-15+, with the plus group encompassing all older ages in addition to the one specified.

Maturity and weight-at-age schedules are deterministic within the OM and EM for each species, as they are in the sablefish assessment (Table 2.4). In the assessment for walleye pollock (Ianelli et al. 2015b), maturity-at-age is deterministic and weight-at-age is estimated using variance terms that are estimated by a separate random effects model outside of the model used for assessment. The deterministic values used for weight-at-age in the walleye pollock OM and EM (Table 2.4) are the mean weights-at-age from the bootstrapped data used by the separate random effects model (Table 1.20 in Ianelli et al. (2015b)). Table 2.4 summarizes the deterministic values for maturity and weight-at-age used in all models for the respective species.

Table 2.3 Parameter values used in the operating models. Deviation vector values are not included because they are randomly generated for each replication.

Name	Symbol	Value	
		Sablefish	Pollock
Mean natural mortality	$M_0$	0.10	0.30
Catchability	$q$	7.56	1.22
Mean recruitment	$\mu_R$	4.48	19253.12
Recruitment CV	$\sigma_R$	1.36	0.64
Mean fishing mortality	$\mu_F$	0.10, 0.20	0.30, 0.60
Fishing mortality CV	$\sigma_F$	0.19	0.25
Age at 50% selection for survey	$a_{50\%I}$	2.74	5.30
Slope parameter for logistic survey selectivity curve	$\delta_I$	2.00	1.06
Age at 50% selection for fishery	$a_{50\%C}$	3.90	5.30
Slope parameter for logistic fishery selectivity curve	$\delta_C$	2.00	1.06
Survey biomass CV	$\sigma_I$	0.05, 0.15	0.05, 0.15
Catch biomass CV	$\sigma_C$	0.01	0.01
Survey sample size	$n_I$	160	69
Catch sample size	$n_C$	200	87

Table 2.4 Quantities for proportion mature-at-age and weight-at-age. These values were incorporated deterministically within the operating and estimation models for sablefish and walleye pollock. Values were estimated outside of the assessment model for each species.

Age	Pollock		Sablefish	
	Proportion Mature (Smith 1981)	Weight at age (Ianelli et al. 2015b)	Proportion Mature (Sasaki 1985)	Weight at age (Hanselman et al. 2007)
2	-	-	0.021	0.946
3	0.290	0.357	0.046	1.468
4	0.642	0.510	0.101	1.966
5	0.842	0.647	0.207	2.408
6	0.902	0.772	0.377	2.784
7	0.948	0.899	0.583	3.095
8	0.964	1.030	0.764	3.349
9	0.970	1.159	0.882	3.552
10	1.000	1.272	0.946	3.715
11	1.000	1.369	0.976	3.843
12	1.000	1.454	0.989	3.945
13	1.000	1.530	0.995	4.024
14	1.000	1.556	0.998	4.087
15	1.000	1.643	0.999	4.136
16	-	-	1.000	4.175
17	-	-	1.000	4.205
18	-	-	1.000	4.229
19	-	-	1.000	4.247
20	-	-	1.000	4.262
21	-	-	1.000	4.273
22	-	-	1.000	4.282
23	-	-	1.000	4.289
24	-	-	1.000	4.295
25	-	-	1.000	4.299
26	-	-	1.000	4.303
27	-	-	1.000	4.305
28	-	-	1.000	4.307
29	-	-	1.000	4.309
30	-	-	1.000	4.310
31	-	-	1.000	4.311

In the OMs, four  $M$  conditions were investigated for each species: two in which  $M$  varies linearly over time(both an increase and decrease), one in which  $M$  is simulated as white noise ( $CV=0.3$ ), and one in which  $M$  is constant. For the case in which  $M$  is simulated as white noise, an existing



R function was used to draw new values from the normal distribution for each run. Current stock assessments fix  $M$  at 0.3 for age 3+ pollock in the eastern Bering Sea (Ianelli et al. 2015b) and 0.1 for age 2+ sablefish in Alaska (Hanselman et al. 2015). Therefore, I allowed  $M$  to vary about these respective values for each operating model that simulates time-varying  $M$ . The effects of low ( $CV=0.05$ ) and high ( $CV=0.15$ ) survey biomass variability and ratio (1 or 2) of  $F$  to  $M$  are also investigated. These OM characteristics are listed in Table 2.5. All 32 possible combinations of the different species,  $M$  trend, survey biomass CV, and  $F$  to  $M$  ratios were tested. Within the age-structured EMs, I tested seven parameterizations of  $M$ . The three EMs without hierarchical structure treat  $M$  as either a fixed constant (FC, specified as the mean of true  $M$ ), estimated constant (EC), or linear function of the covariate over time (COV). The four EMs with hierarchical structure treat  $M$  as either a white noise (WN), random walk (RW), random walk with drift (DW), or correlated walk (CW) process. To ensure reproducibility, the R and ADMB code used to implement all scenarios is provided at [github.com/philganz](https://github.com/philganz).

Table 2.5 Operating model cases, having a total of 32 possible combinations.

	Trend in $M$	Survey biomass variability	Ratio of fishing to natural mortality
Sablefish	Linear increase (LI)	$CV=0.05$	1
Pollock	Linear decrease (LD)	$CV=0.15$	2
	White noise (WN)		
	Constant (C)		

## 2.5 Model performance criteria

Runs differed only in that new values for each stochastic process were generated. The seed number used by R to generate these random numbers is documented so that results are reproducible. The results of each combination are evaluated based on the precision (standard error) and accuracy (bias) with which parameters and quantities of interest are estimated. These measures of performance are quantified using relative errors:

$$(16) \quad RE = (\hat{\theta} - \theta) / \theta.$$

The range of relative error (RE) is used to evaluate precision and median RE is used to quantify the direction of any bias. REs are evaluated for model parameters and SSB, the latter of which is often used to calculate biological reference points used in management.

I used the Akaike information criterion (AIC, Akaike 1973) to compare model fit while accounting for model complexity:

$$(17) \quad AIC = -2\ln(L) + 2p$$

where  $p$  is the number of parameters being estimated. The preferred model is one that strikes a balance between high likelihood (low negative likelihood) and few parameters. Therefore, given two models, the one with the lowest AIC is preferred.

Mixed effects models were removed from the set being compared for reasons detailed in section 3.1. Had mixed effects been included in the final set of candidate models, it would not have been appropriate to use AIC as a model selection criterion due to the fact that random effects are not considered as parameters in the frequentist sense upon which AIC relies. Random effects could

therefore not be included in  $p$ , but the way in which they add to model complexity would still need to be accounted for. There exist other model selection criteria that can accommodate mixed effects models. One such criterion is the deviance information criterion (DIC, Spiegelhalter et al. 2002), which calculates an “effective number of parameters” from the behavior of the objective function during Markov chain Monte Carlo sampling.

It is also important to note that the penalty term for  $M$  deviations (Equation 11) was removed from the likelihood when calculating AIC. The rationale for removing this term is two-fold. Firstly, I am interested in the model fits to the data. This is quantified by likelihood components 1-4, which fit the model to the survey and fishery biomass and age composition. The purpose of the penalty term is to improve this fit, but it does not represent a data source itself. Secondly, removing the penalty makes the likelihood identical across the models being compared, which is a requirement of AIC.

## 2.6 Sensitivity study

When working with a natural system, one never knows the true amount of error present in the processes being measured (process error) or in the observations of those processes (observation error). Much of the modeling endeavor entails making assumptions about how these errors enter the natural system so that they can be minimized appropriately when searching for expected values and quantifying uncertainty in the estimation process. One assumption I make in this study with regard to how error enters the simulated system is that the covariate is observed with error. The magnitude of this error is likely to affect the amount with which the covariate aids in estimation, but would never be known when working with data from a natural system. The purpose of the sensitivity study, therefore, is to investigate the way in which different amounts of

covariate observation error affect estimation. The objective is to identify any breaking points at which the covariate no longer aids in estimation. Estimation is performed at four levels of observation error variability:  $\sigma_\gamma = 0.0$ ,  $\sigma_\gamma = 0.2$ ,  $\sigma_\gamma = 0.4$ , and  $\sigma_\gamma = 0.6$ . Additive error is applied in simulating observed covariates (Equations 2a-2c).

### 3. Results

#### 3.1 Penalized likelihood and random effects performance

An initial set of 25 runs for each scenario revealed stark differences in the performance of penalized likelihood models compared with their mixed effects counterparts. Convergence, as defined by the production of a positive-definite hessian matrix of second derivatives, is necessary to produce the covariance matrix and associated standard errors. Convergence proportions for the mixed effects models were considerably lower than those for penalized likelihood (Figure 3.1).

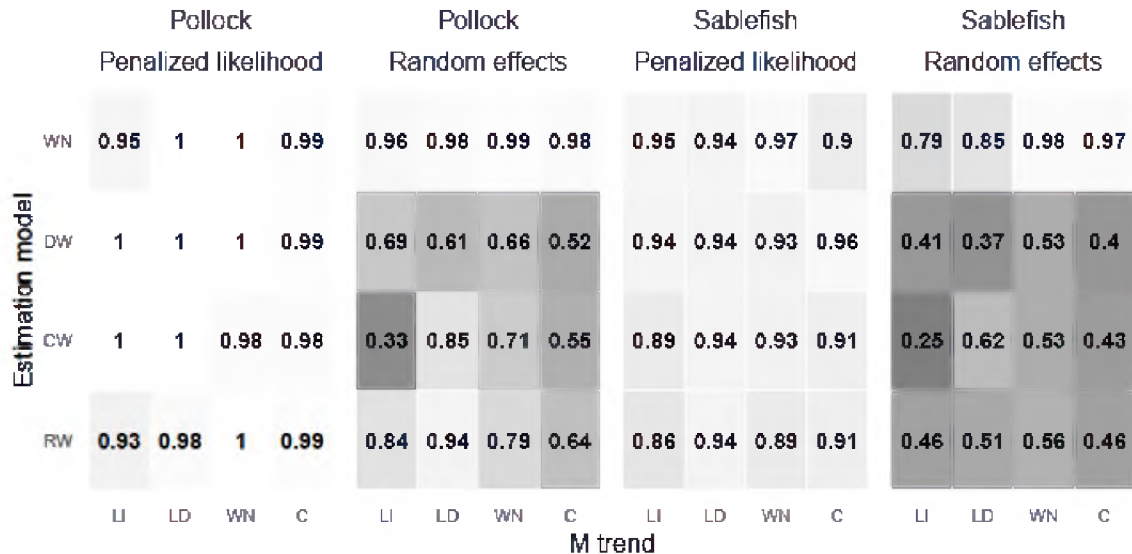


Figure 3.1 Convergence proportions after an initial 25 runs for models fit by penalized likelihood and random effects. Lighter shading corresponds with higher rates of convergence.

Further, even under scenarios for which convergence was not an issue for mixed effects models, such as the case in which  $M$  is estimated as white noise, it was found that the variance term tended to 0. This means that a constant  $M$  was estimated by mixed effects models for the same scenarios in which the analogous penalized likelihood model was able to more closely approximate the true underlying trend in mortality (Figure 3.2). The tendency of the variance term to go to 0 has been documented elsewhere (Skaug and Fournier 2015), and is interpreted as an indication that random effects are not supported. For this reason, and because penalized likelihood models better approximated the true underlying trends in  $M$ , the results presented below for hierarchical EMs are those that are fit by penalized likelihood.

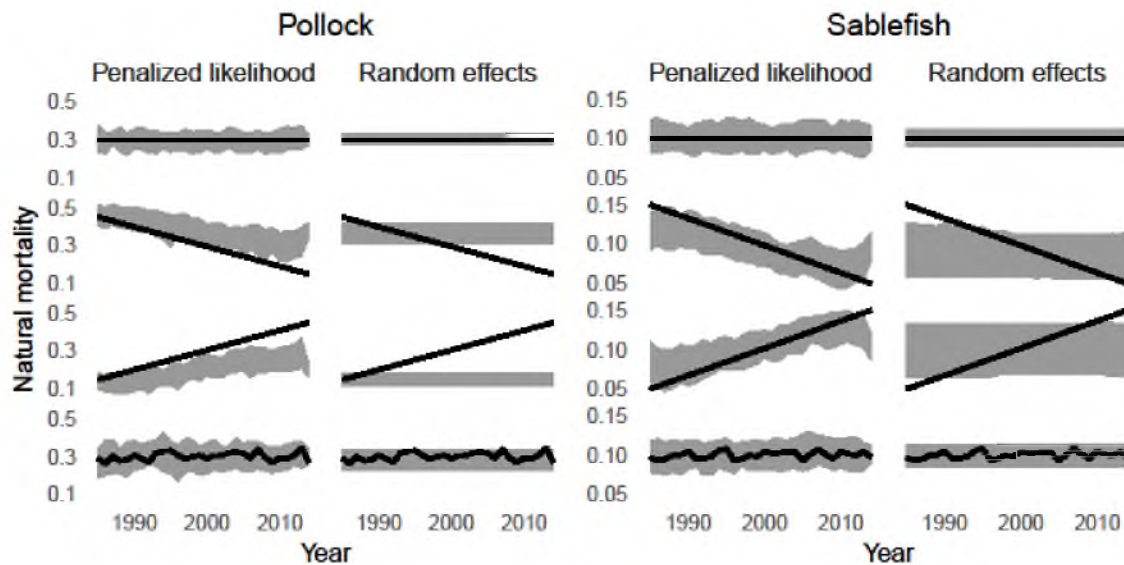


Figure 3.2 Simulated and estimated values for natural mortality by species and likelihood method. Results are for the white noise estimation model only. The solid lines and light grey areas correspond to simulated values of natural mortality and the interquartile range of estimates, respectively. The median is shown for simulated values of white noise mortality.

## 3.2 The influence of population dynamics and data quality on estimation

I present results for each population characteristic being tested. For clarity, I present results for each population or data quality characteristic individually, rather than presenting results for every possible combination, of which there are many. This necessitates combining results across the OM cases not being investigated and across estimation methods. There are drawbacks to looking at results in this way; for instance, the same estimation model may be responsible for the highest relative errors in each OM case. However, combining results in this manner is a straightforward way of investigating the effects that underlying dynamics have on estimation. Within sections 3.2.1 to 3.2.4, I expand on each of these population characteristics to see how they affect estimation. In section 3.3, I separate results by estimation method.

### 3.2.1 Species

I did not observe substantial differences between species alone in the ability of EMs to estimate SSB. The interquartile range of REs in SSB estimates are almost identical (Figure 3.3). The total number of parameters estimated for the case in which a single value of  $M$  is estimated is 80 for walleye pollock and 97 for sablefish. Pollock and sablefish data sets contained 840 and 1860 observations, respectively.

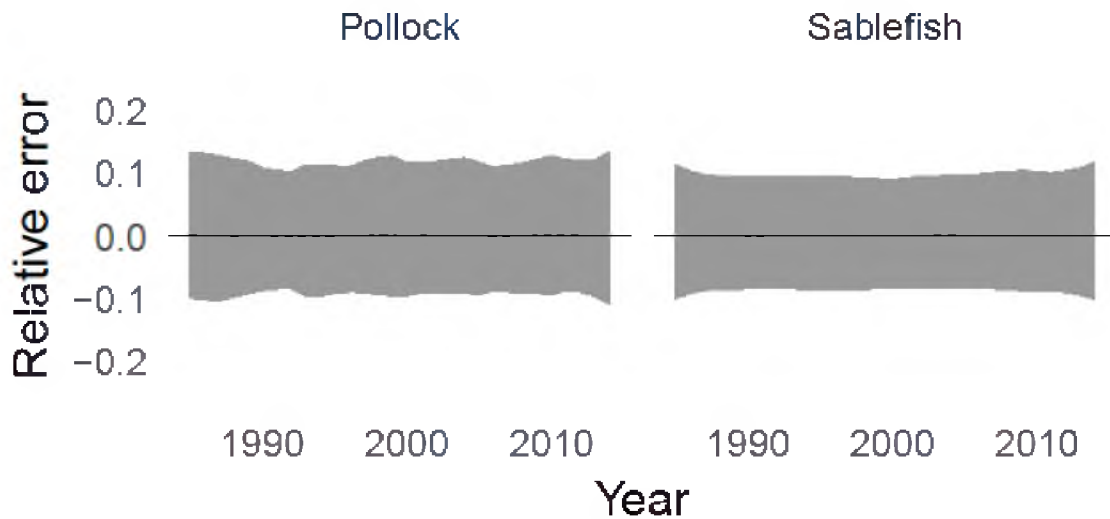


Figure 3.3 Relative errors (REs) in spawning stock biomass for pollock and sablefish. The dashed lines and light grey areas correspond to the median and interquartile range of REs, respectively.

Although results were similar on the level of estimating SSB, differences were observed in the accuracy and precision with which individual parameters were estimated for each species (Figure 3.4). Each species has its own most problematic parameter other than  $M_0$ . For walleye pollock, this is  $q$ . For sablefish, the problematic parameter other than  $M_0$  is  $\mu_R$ . Sablefish are characterized by highly variable recruitment, which is reflected by the value assumed for  $\sigma_R$  by the 2015 assessment, 1.20 (Hanselman et al. 2015). In this study, the value for  $\sigma_R$  used to simulate recruitment events is 1.36 for sablefish and 0.64 for pollock (Table 2.3), both of which correspond to the variance of estimated recruitment events since the regime shift in 1978. The variance of both the simulated and estimated recruitment deviations is therefore much higher for sablefish than pollock. REs for  $\mu_R$  tended to be of larger magnitude for sablefish compared to pollock (Figure 3.4), but much smaller in magnitude than the REs for pollock  $q$ .

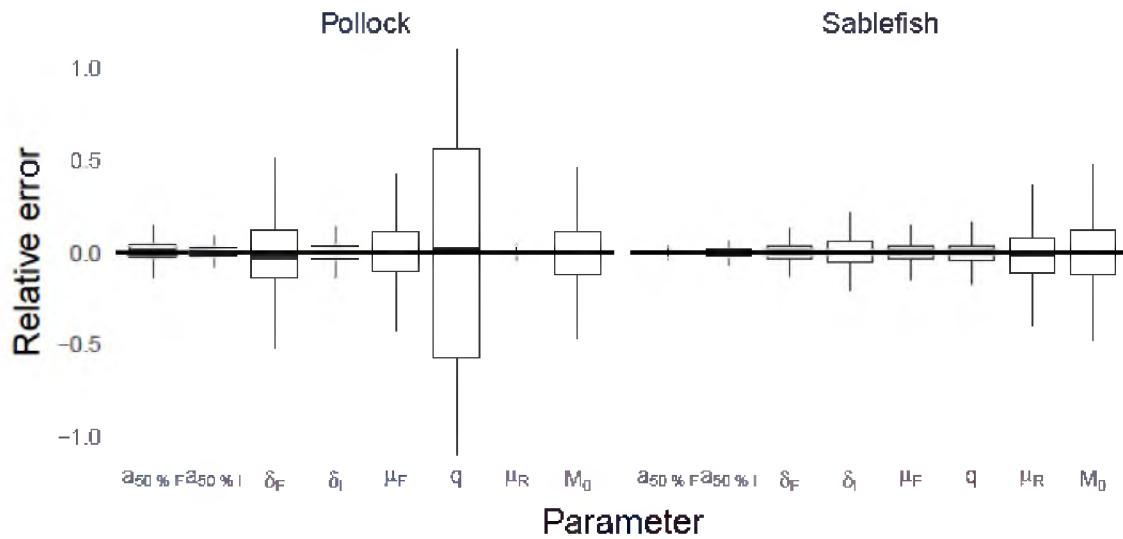


Figure 3.4 Relative errors of parameters for pollock and sablefish. Outliers are excluded.

### 3.2.2 Natural mortality trend

When considering the results of all estimation methods together, there is not much difference in the accuracy and precision of SSB estimates among trends in  $M$ . This is expected, as some estimation methods are not even estimating time-variable  $M$  and one estimation method could be responsible for the largest errors across trends. The performance of individual estimation methods is investigated in section 3.3, but some general takeaways can still be gained here regarding the effect that trend in  $M$  has on estimation. One result is that REs appear to be unbiased and approximately median-unbiased across trends in  $M$  (Figure 3.5). Another result is that errors differ between cases in which  $M$  is a linearly increasing function of time and cases in which  $M$  is a linear decreasing function of time. For pollock, the slight median bias changes direction between these two trends in  $M$  and is likely driven by the case in which  $M$  was fixed or estimated as a constant.



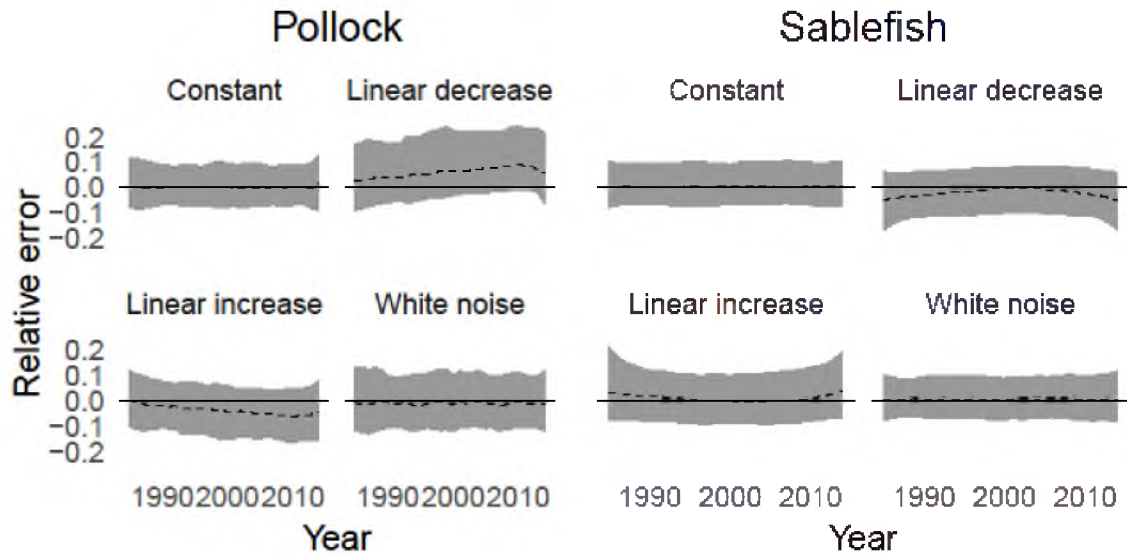


Figure 3.5 Relative errors (REs) in spawning stock biomass by species and natural mortality trend. The dashed lines and light grey areas correspond to the median and interquartile range of REs, respectively.

### 3.2.3 Survey biomass variability

As expected, a less precise survey resulted in more imprecise estimates of SSB (Figure 3.6). This result is similar between species. Bias was not observed under either of the survey CVs tested.

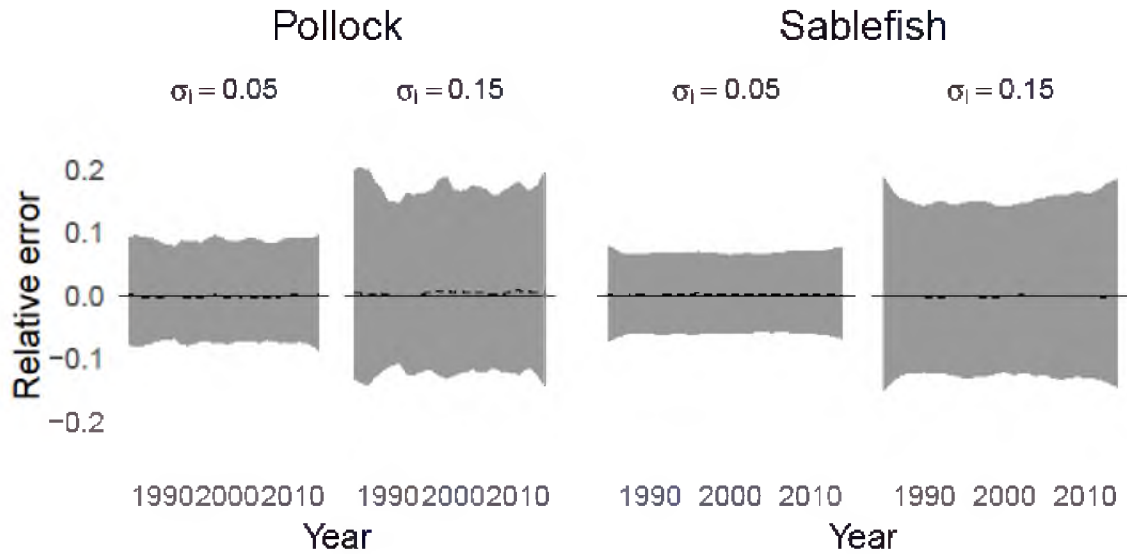


Figure 3.6 Relative errors (REs) in spawning stock biomass by species and survey precision. The dashed lines and light grey areas correspond to the median and inner 50% of REs, respectively.

### 3.2.4 Ratio of fishing to natural mortality

The ratio of fishing mortality to natural mortality influenced estimation considerably. Cases in which  $F/M = 2$  produced much more precise estimates of SSB than cases in which  $F/M = 1$  (Figure 3.7). This is intuitive, as fishery-dependent information becomes more informative as the ratio of fishing mortality increases relative to other sources of mortality. It should be noted that the  $\mu_F$  is adjusted in the OM in order to reach the appropriate ratio of either 1 or 2 and both ratios produce values of  $\mu_F$  that are higher than estimated for the wild populations of pollock and sablefish. These ratios were chosen for simplicity and to investigate the influence that the  $F/M$  ratio has on estimation. Again, I did not observe differences in estimation between species.

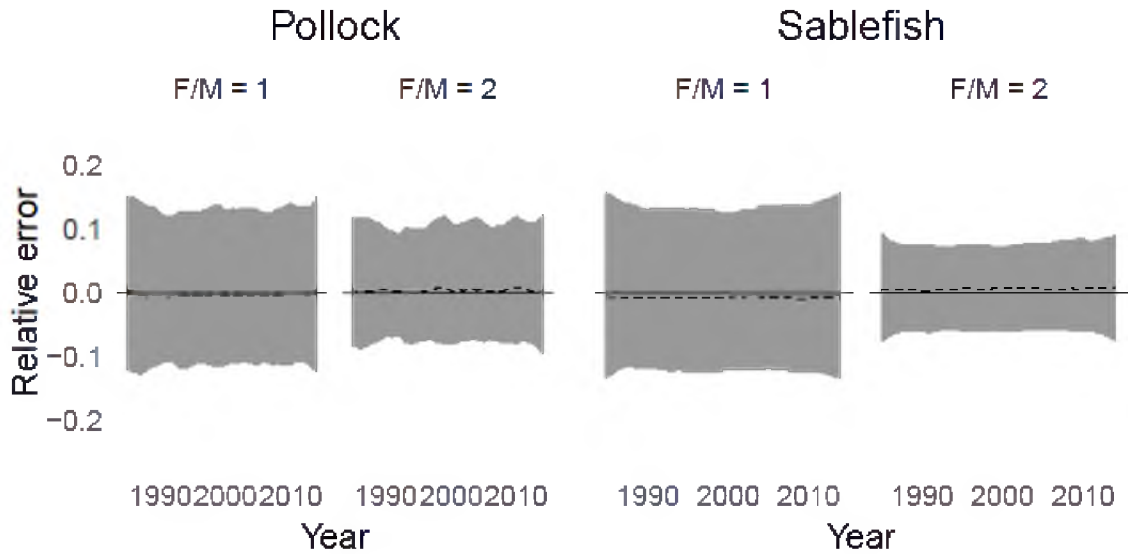


Figure 3.7 Relative errors (REs) in spawning stock biomass by species and ratio of fishing mortality to natural mortality. The dashed lines and light grey areas correspond to the median and interquartile range of REs, respectively.

### 3.3 Evaluating the performance of different estimation methods

#### 3.3.1 All estimation methods

Estimation models differed in the precision and accuracy with which they were able to estimate annual SSB. Each model's performance also differed between trends in  $M$ . For the case in which true  $M$  is constant over time (Figure 3.8), the EM in which  $M$  was fixed at that value acts more as a calibration case than a situation stock assessors are likely to find themselves, that is, knowing the true value of  $M$ . Fixing  $M$  at its true value in this case provided the extremely accurate and precise estimates expected. All other EMs produced slightly less precise estimates of SSB, but none were biased.

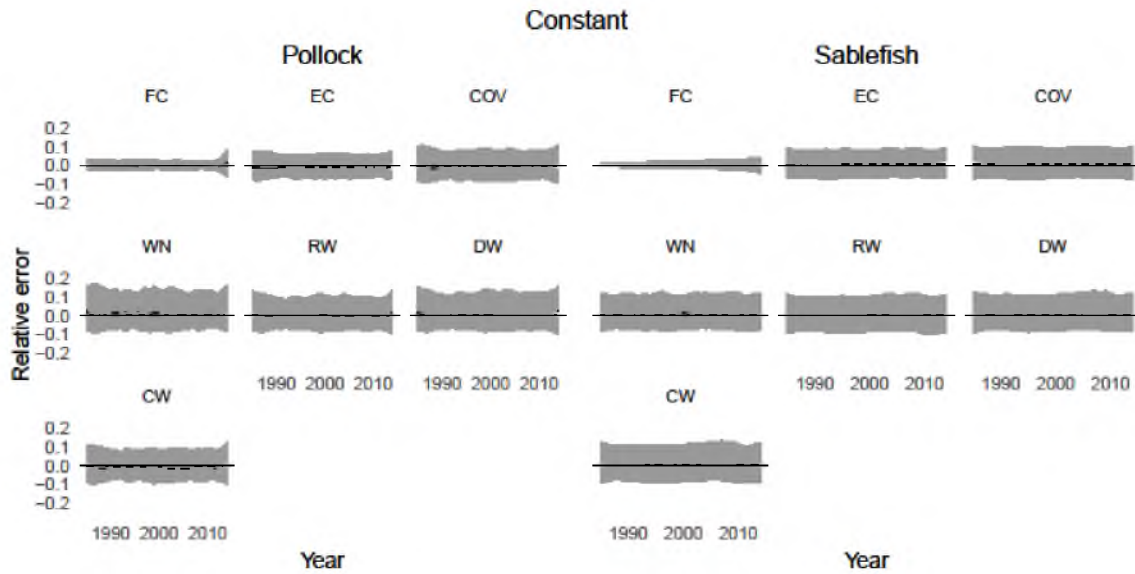


Figure 3.8 Relative errors (REs) in spawning stock biomass by species and estimation model for the case in which natural mortality is constant over time. The dashed lines and light grey areas correspond to the median and interquartile range of REs, respectively. Estimation model names are given in section 2.4.

The relative accuracy and precision between models changes substantially for cases in which true  $M$  is time-varying rather than constant. For the case in which  $M$  is simulated as a linear increase over time (Figure 3.9), the covariate approach still produces precise and unbiased estimates of SSB. For pollock, fixing a constant  $M$  resulted in biased estimates of SSB for much of the time-series. Estimating constant or white noise  $M$  resulted in biased estimates for the entirety of the time-series. These same EMs also produced biased estimates for sablefish, although results were mainly median-biased and concentrated at the very beginning and end of the time series. The random walk EM, with and without drift, resulted in unbiased estimates of SSB for each species. The results observed under a linear increase in  $M$  largely hold true for the case in which  $M$  is modeled as a linear decrease (Figure 3.10). The same models produce biased results, although the direction in bias shifts.

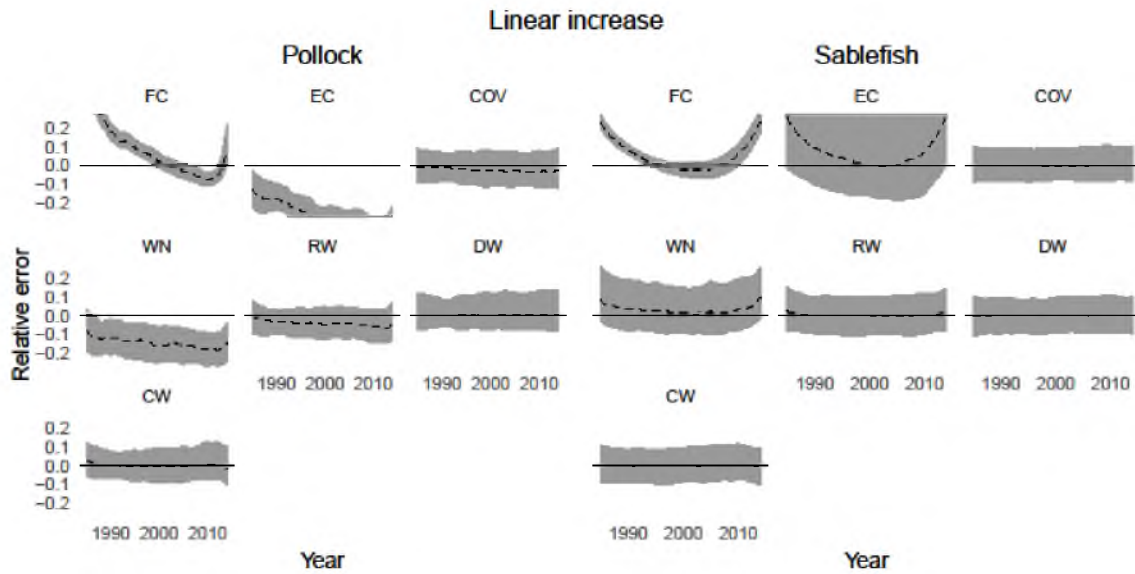


Figure 3.9 Relative errors (REs) in spawning stock biomass by species and estimation model for the case in which natural mortality increases linearly over time. The dashed lines and light grey areas correspond to the median and interquartile range of REs, respectively. Estimation model names are given in section 2.4.

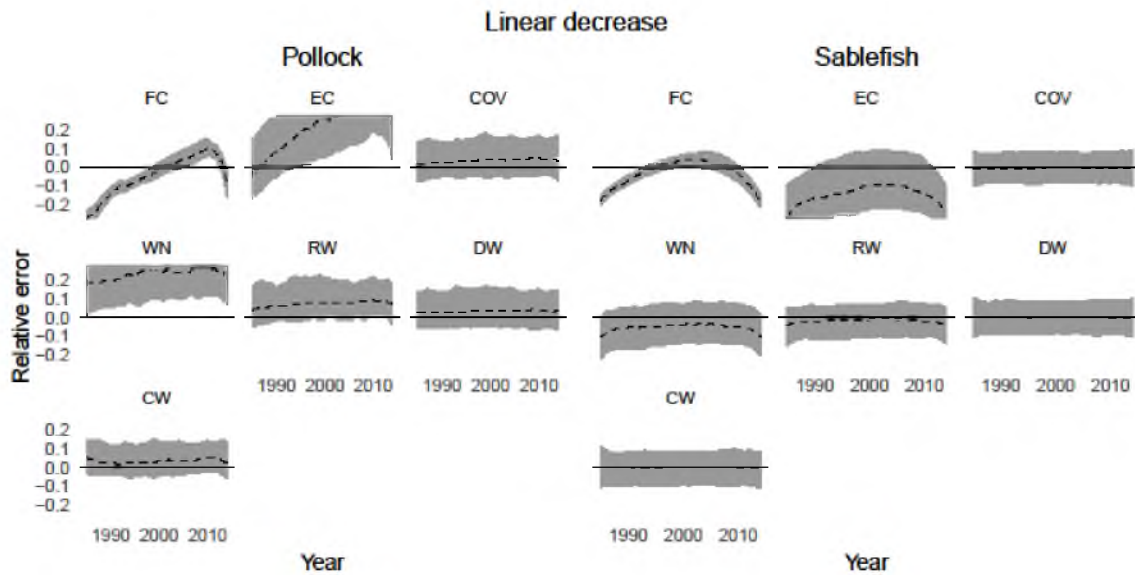


Figure 3.10 Relative errors (REs) in spawning stock biomass by species and estimation model for the case in which natural mortality decreases linearly over time. The dashed lines and light grey areas correspond to the median and interquartile range of REs, respectively. Estimation model names are given in section 2.4.

The case in which  $M$  is generated as white noise produces similar results (Figure 3.11) as the case in which true  $M$  is constant. All estimates of SSB are unbiased, although estimates are less precise than for true constant  $M$ . Hierarchical models, including that of white noise for pollock, appear to produce less precise estimates than non-hierarchical approaches.

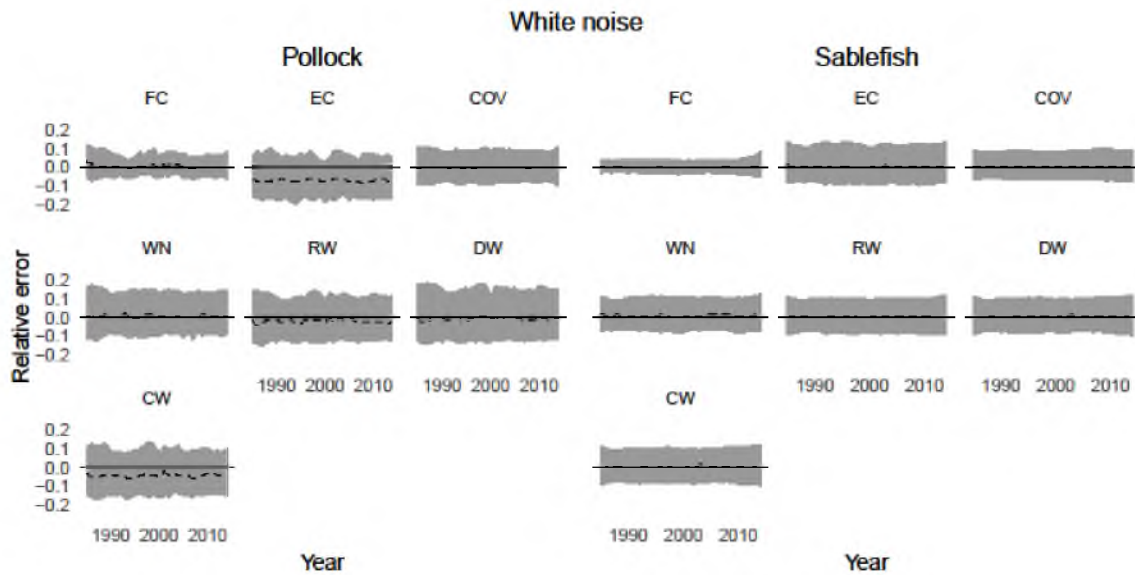


Figure 3.11 Relative errors (REs) in spawning stock biomass by species and estimation model for the case in which natural mortality has a normal random distribution. The dashed lines and light grey areas correspond to the median and interquartile range of REs, respectively. Estimation model names are given in section 2.4.

A model's ability to track abundance through time is important for assessing the assumptions made about the dynamics of a population. However, an important value for management is the estimate of abundance in the last year for which there are data. This terminal abundance provides the most recent estimate of where the population is in relation to the abundance that would support maximum sustainable yield. The model structures tested in this study differed in their ability to estimate terminal SSB (Figure 3.12). Estimates were most precise and accurate when the true underlying trend in  $M$  was constant. The precision of these estimates decreased for  $M$

generated as white noise, while retaining accuracy. I observed biased estimates for cases in which simulated  $M$  was either linearly increasing or decreasing, with estimated constant  $M$  producing the most biased results. Fixed and estimated constant models were biased in the opposite direction for pollock, but the same direction for sablefish.

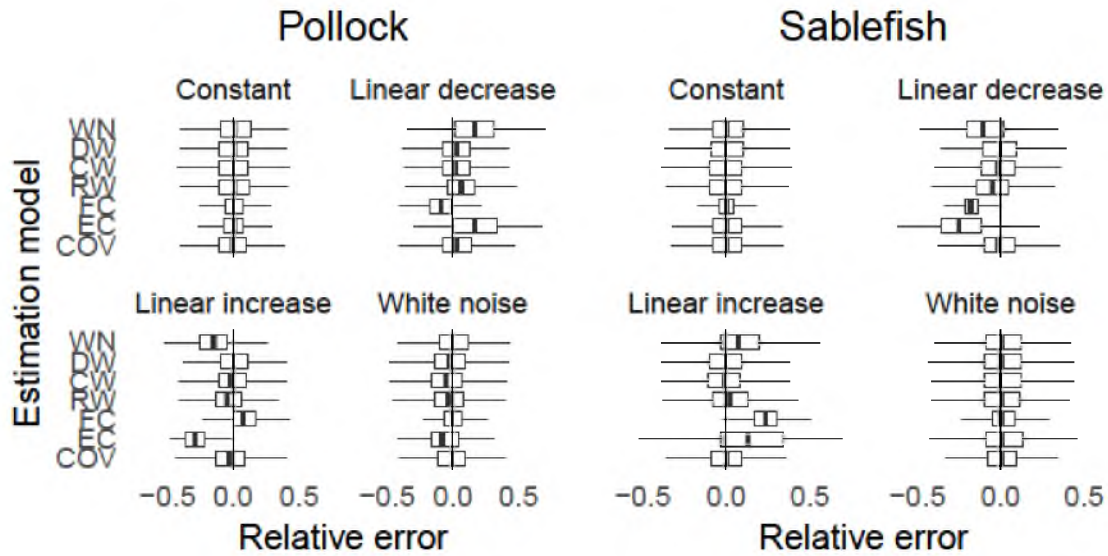


Figure 3.12 Relative errors in estimates of terminal year spawning stock biomass by species and natural mortality trend. Results are combined across the OM cases of species, survey precision, and fishing mortality. Estimation model names are given in section 2.4.

In attempting to parse out drivers of the biases observed in terminal year SSB, I examined the REs of individual parameters. Estimates of constant  $M$  in pollock were biased when true  $M$  was either linearly increasing or decreasing (Figure 3.13). Estimates of constant  $M$  in sablefish were highly imprecise when true  $M$  was changing linearly and median-biased when true  $M$  was decreasing.

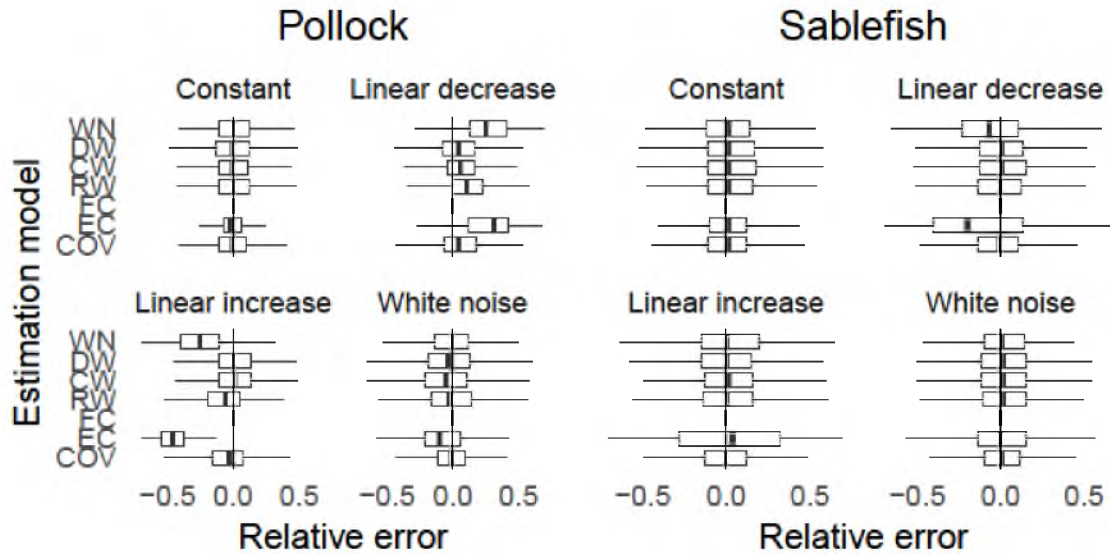


Figure 3.13 Relative errors in estimates of mean natural mortality by species and natural mortality trend. Results are combined across the OM cases of survey precision and fishing mortality. Estimation model names are given in section 2.4.

The biases observed in  $M_0$  were in the same direction as those in SSB. A low estimate of  $M_0$  should correspond to a high estimate of SSB, so errors in the same direction signify that biases in other parameters were driving the biases observed in terminal year SSB. The REs for  $q$  (Figure 3.14) suggest that this is one parameter counteracting the effects that biased  $M_0$  would have on terminal year SSB.



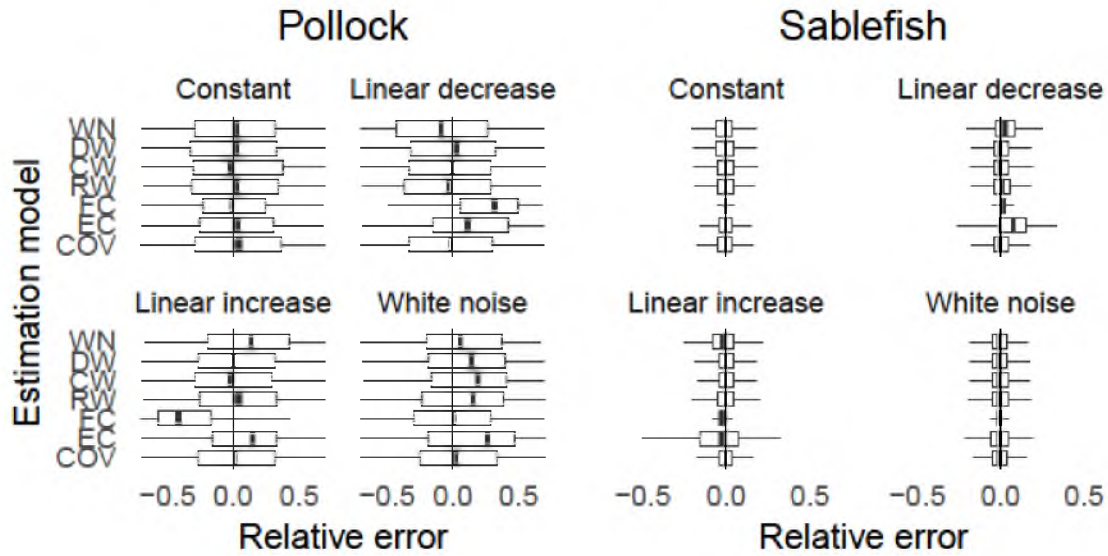


Figure 3.14 Relative errors in estimates of survey catchability by species and natural mortality trend. Results are combined across the OM cases of survey precision and fishing mortality. Estimation model names are given in section 2.4.

### 3.3.2 Sensitivity study

The results of the sensitivity study, in which I evaluated the performance of the covariate method at different levels of observation error variability ( $\sigma_\gamma$ ), suggest that observation error variability does not have a profound effect on the performance of this method. Relative errors in SSB are of similar magnitude and direction for cases in which  $\sigma_\gamma = 0.0$ ,  $\sigma_\gamma = 0.2$ ,  $\sigma_\gamma = 0.4$ , and  $\sigma_\gamma = 0.6$  (Figure 3.15).

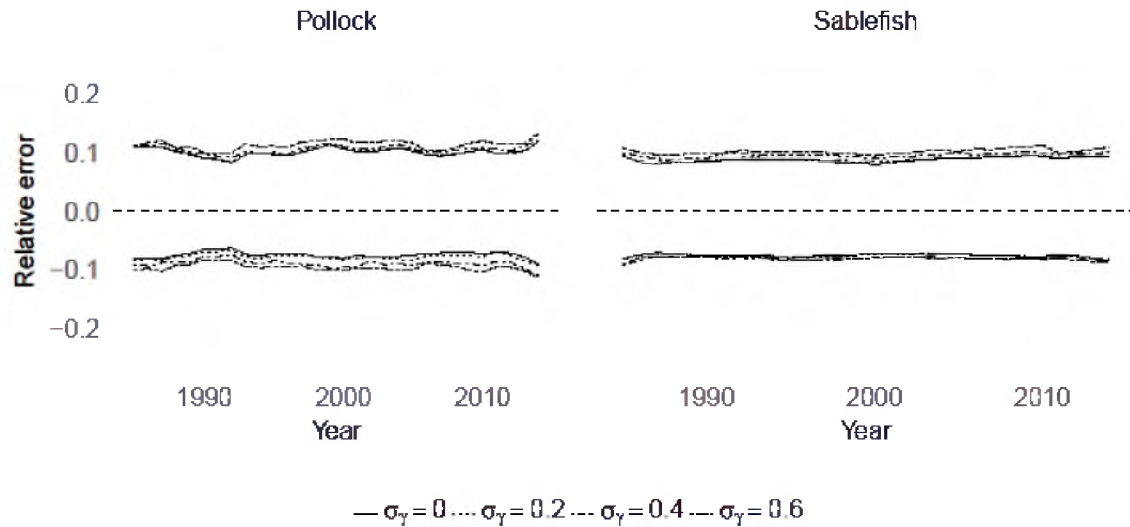


Figure 3.15 Relative errors (REs) in spawning stock biomass by species and covariate precision. Lines border the interquartile range of REs and line type corresponds to the variability of the error with which the covariate is observed.

### 3.3.3 Model comparison

AIC proved to be an effective tool for selecting the EM that matched the OM used to simulate data. For this analysis, I used data sets generated under low survey biomass variability (CV=0.05) and an  $F/M$  ratio of 1. In all scenarios, the model used to simulate data was the covariate model, as true  $M$  values were simulated from the covariate. The covariate model was selected by AIC as the best model on most occasions (Figure 3.16). AIC tended to select the model that fixed natural mortality at its true value for the case in which true  $M$  was constant. This is because a fixed constant  $M$  and constant  $M$  estimated with a covariate describe the correctly describe the same process, yet the fixed constant method does so with one less parameter than the covariate model. Again, the case in which constant natural mortality is correctly specified is a situation in which analysts are unlikely to find themselves, so it is worth

noting that the covariate model also performed well in the case in which true  $M$  is constant.

Curiously, the covariate model with no error was never selected when true  $M$  was constant.

Hierarchical models were never selected, even for the case in which  $M$  was both simulated and estimated as white noise. AIC tended to select the covariate EM that received values of the covariate without observation error ( $\sigma_\gamma = 0$ ). However, the covariate model with observation error ( $\sigma_\gamma = 0.2, \sigma_\gamma = 0.4, \sigma_\gamma = 0.6$ ) was sometimes selected.

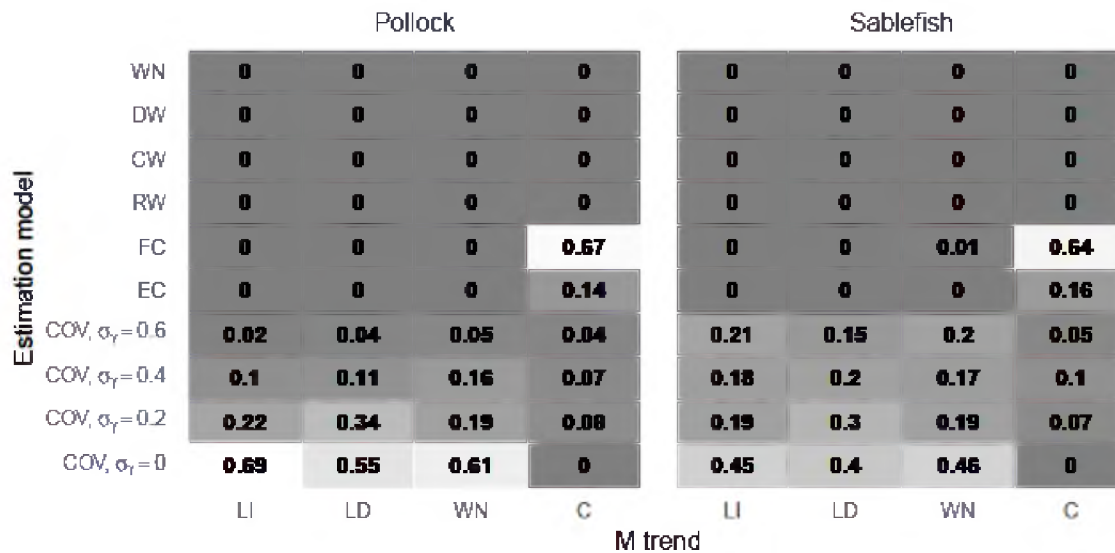


Figure 3.16 Proportion of instances in which AIC selected each estimation model by natural mortality trend. Lighter shading corresponds with higher proportion of runs chosen by AIC. Estimation model names are given in section 2.4.

In addition to examining the proportion of instances in which AIC correctly identified the true model, I calculated  $\Delta$ AIC values—the difference between each model’s AIC value and the lowest AIC value for each iteration. Median  $\Delta$ AIC values (Figure 3.17) illustrate the fit of each model relative to the best model, as defined by AIC. I used the median as opposed to the mean in order to avoid values being skewed by occasional high AIC values. Hierarchical models tended

to have higher  $\Delta AIC$  values than covariate approaches. Fixed and estimated constant EMs also produced high median  $\Delta AIC$  values for scenarios in which natural mortality was increasing linearly with time. For the case in which true  $M$  was constant, covariate EMs tended to have AIC values within a few points of the fixed constant EM.

		Pollock				Sablefish			
Estimation model	WN	35.7	35.6	37.15	36.4	43.85	45.35	45.9	45.65
	DW	51.75	51.05	76.25	53.55	54.75	55.15	60.4	57.3
	CW	57	57.05	95.25	58.75	55.1	55.15	60.7	57.35
	RW	50.15	49.8	74.3	51.6	53.05	53.65	58.15	55.4
	FC	91.35	79.1	53.8	0	88.4	73.8	14.6	0
	EC	55.1	66.7	48.95	1.65	67.05	57.65	12	1.7
	COV, $\sigma_\gamma=0.6$	14.35	13.45	12.15	2.9	3.3	4.05	2.65	2.75
	COV, $\sigma_\gamma=0.4$	7	7.1	6.2	2.75	1.8	2.5	1.8	2.8
	COV, $\sigma_\gamma=0.2$	2.25	1.5	2.25	2.65	1.2	0.65	0.9	2.8
	COV, $\sigma_\gamma=0$	0	0	0	3.65	0.15	0.4	0.35	3.7
			LI	LD	WN	C	LI	LD	WN
		M trend							

Figure 3.17 Median  $\Delta AIC$  values for each estimation model by natural mortality trend. Lighter shading corresponds with lower median  $\Delta AIC$  values. Estimation model names are given in section 2.4.

In order to compare the relationship of AIC model selection to the accuracy with which an EM is able to estimate a quantity relevant to management, I displayed the value of median REs for terminal year SSB in the same way (Figure 3.18). As expected, fixed and estimated constant EMs produced inaccurate estimates of terminal year SSB when  $M$  was either linearly increasing or decreasing. For pollock, the white noise EM also performed poorly in these scenarios. When  $M$  was either constant or undergoing white noise deviations over time, all models were able to estimate terminal year SSB with similar accuracy. The random walk, correlated walk, and

random walk with drift models were able to estimate terminal year SSB with approximately the same accuracy as models that relied on a covariate.

		Pollock				Sablefish			
Estimation model	WN	-0.17	0.24	0	0.03	0.04	-0.08	0.01	0
	DW	0	0.03	-0.03	0.01	-0.03	-0.01	0.02	-0.01
	CW	-0.01	0.03	-0.06	0.01	-0.03	-0.02	0	-0.02
	RW	-0.03	0.06	-0.04	0.01	-0.02	-0.04	0.01	-0.02
	FC	0.07	-0.08	0	0.01	0.28	-0.2	0	0
	EC	-0.33	0.24	-0.1	0.03	0.44	-0.3	-0.01	-0.01
	COV, $\sigma_\gamma=0.6$	-0.08	0.09	-0.01	0.02	-0.01	-0.02	0.01	0
	COV, $\sigma_\gamma=0.4$	-0.04	0.04	-0.01	0.02	-0.02	-0.02	0	0
	COV, $\sigma_\gamma=0.2$	0	0.02	0	0.01	-0.03	-0.02	0.02	0
	COV, $\sigma_\gamma=0$	0	0.02	0.01	0.03	-0.03	-0.01	0.01	-0.01
		LI	LD	WN	C	LI	LD	WN	C
		M trend							

Figure 3.18 Median relative errors (REs) in spawning stock biomass for each estimation model by natural mortality trend. Lighter shading corresponds with lower median REs. Estimation model names are given in section 2.4.

#### 4. Discussion

Including a covariate aided the estimability of time-varying  $M$ , regardless of the observation error associated with the covariate. Little degradation in estimation was observed, even when the standard deviation,  $\sigma_\gamma$ , of the observation error added to the covariate was as high as 0.6. The covariate aided in the estimation of time-varying  $M$  across trends in true mortality, resulting in median-unbiased estimates of terminal year SSB across trends in  $M$  and across species.

Estimating a constant value of  $M$  performed poorly when true  $M$  was time-varying. Biased estimates in terminal year SSB resulted when a constant value of  $M$  was estimated for linear increases or decreases in true  $M$ . The bias in  $M_0$  was in the same direction as the bias in terminal

year SSB, suggesting that bias in other model parameters were counteracting the effects on bias in  $M_0$ . Catchability appeared to be driving this counter-effect on SSB, particularly for pollock.

Interestingly, assuming  $M$  to be constant when it was undergoing a linear change resulted in quadratic trends in SSB relative errors. For the scenario in which  $M$  is fixed at its true mean but is actually increasing linearly, if parameters other than  $M$  were estimated exactly, one would expect to see an underestimate of SSB at the beginning of the time series followed by an overestimate of SSB in later years. The reason for this result remains unclear. Initial abundance in the first year of the time series was regularly overestimated for sablefish, but not for pollock, and I observed no trends over time in the errors associated with annual estimates of  $F$  or  $R$  (results not shown). This result illustrates the difficulty in assigning model results to any one parameter in a highly nonlinear model.

Although AIC proved to identify the true model effectively, there was some disparity between models chosen by AIC and models that minimized the median REs in terminal year SSB. Specifically, hierarchical models produced approximately median-unbiased estimates of terminal year SSB but were never selected by AIC. This suggests that the model chosen by AIC may not be the only model useful for estimating quantities of interest to management. Although comparing known values to the truth is not possible outside of simulation, analysts use metrics such as retrospective patterns to quantify the predictive power of a model (Mohn 1999).

The preference for fitting hierarchical models with an integrated, marginal likelihood, has been based on the observation that their penalized likelihood equivalents can produce biased estimates (De Valpine and Hilborn 2005) and on the inability to estimate the variance term associated with the hyper-distribution being used. However, hierarchical model structures fit by penalized

likelihood with a fixed variance term proved to be a robust approach to estimating time-varying  $M$  in cases in which auxiliary information, such as a covariate to  $M$ , is not available. Adding either a drift or correlation term to the basic random walk structure reduced the bias and increased the precision of terminal year SSB estimates. Because adding a term for drift is equivalent to estimating a linear function of time (if deviations go to 0), EMs that estimate  $M$  as a random walk with drift are specified very similarly to the OMs used to simulate constant and linear changes in  $M$  over time. The conclusions drawn from any simulation study are influenced by the amount of misspecification between the OM and EM (Francis 2012). However, in choosing between simplifying assumptions to make for a real system for which misspecification will be present but unknown, it may be beneficial to assume linear change rather than constant natural mortality, even in the case where true  $M$  is not changing linearly. The benefit of doing so can be judged in part by the effect on retrospective pattern.

Estimating time-varying natural mortality with hierarchical models proved untenable when attempting to integrate out the random effects in order to produce the marginal likelihood. However, this result has the potential to change with the software used to approximate the marginal likelihood. Although I found mixed effects models to be problematic using ADMB-RE, template model builder (TMB, Kristensen et al. 2016) is an ADMB-inspired software that shows promise in its ability to approximate the marginal likelihood in fitting random effects models (Albertsen et al. 2015). As the number of TMB users continues to grow, its performance in optimizing various non-linear, mixed effects models will continue to be evaluated. Such an evaluation and comparison to ADMB-RE is beyond the scope of this study, but further research should test the ability of TMB to fit mixed effects models for various applications.

Although I tested misspecifications in the way that models structured  $M$ , the assumptions made regarding other parameters, such as selectivity and catchability being time-invariant, are the same between the OM and EM. Time-varying selectivity and catchability also have great potential to influence the results of stock assessments, and this has been an active area of research (Wilberg and Bence 2006, Wilberg et al. 2010, Martell and Stewart 2014). In reality, which processes are time-varying and which are constant will change between stocks and time periods, leaving any individual model prone to misspecification over time. To reduce the risk of adopting a single, misspecified model, both model averaging (Ianneli et al. 2015a) and ensemble modeling (Stewart and Martell 2015) approaches have been applied in fisheries. Ensemble modeling techniques have begun to be used by fisheries scientists only recently, but the extent to which they can provide inference to a system with multiple potential time-varying processes should continue to be researched. However, the processes affecting populations are non-linear, so that simple averages over models result in inconsistent dynamics. This aspect of multi-model inference should be addressed as research continues.

## 5. Conclusions

The finding that estimating a constant value of  $M$  results in some of the most biased estimates when true  $M$  is time-varying differs from the results of Johnson et al. (2015), who found this to be a robust approach when true  $M$  underwent a sharp increase or decrease within the timeframe considered. The authors cited multiple examples of mass die-offs as rationale for simulating this sharp change. However, there is also evidence that populations may experience more gradual changes in  $M$  (Marty et al. 2010, Jiao et al. 2012) that are less easily detected. Differences in data quantity and model structure can affect results (Johnson et al. 2015), and this is evident in the results presented above. Therefore, despite attempts to identify a default approach to the



treatment of  $M$  within stock assessments (Johnson et al. 2015), I conclude that the most appropriate parameterization of  $M$  will likely depend on the dynamics, data quality, and data quantity associated with each population.

One caveat to the results of the covariate approach is that estimates would likely be influenced by covariate accuracy. In the sensitivity study, I evaluated the influences of covariate precision, but assumed the covariate to be accurate. At a certain degree of inaccuracy, a measurable variable can no longer be considered a covariate to a process. Further research is needed to identify the level of inaccuracy at which the covariate no longer aids in estimation.

When considering alternate model structures, it is important to keep in mind the ultimate goals of stock assessment. More valuable than a model's ability to estimate past and current population size is its ability to predict future population trends. While hierarchical models can quantify the amount of variability that can be expected in a process going forward, they cannot speak to the direction of change. In contrast, when a process is related to a covariate, predictions of the covariate inform predictions of the process. For instance, when related to the North Pacific Polar Front, projections of sablefish recruitment were shown to increase relative to the base model that did not include this relationship (Shotwell et al. 2012). After identifying a robust covariate to  $M$  in Pacific herring, that model was able to predict biomasses below the threshold needed to fish, preventing the fleet from mobilizing to the grounds in error (Marty et al. 2010). This increase in predictive power is not possible with hierarchical models, particularly those fit by penalized likelihood in which the amount of variance is chosen *a priori*. For these reasons, relating process variability to a predictable covariate should be preferred over describing that variability with hierarchical model structure. Efforts to identify covariates to time-varying processes should continue.

## 6. References

- Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle, In: Proceedings of the 2nd International Symposium on Information Theory. Academiai Kiado, Budapest. pp. 267–281.
- Albertsen, C.M., Whoriskey, K., Yurkowski, D., Nielsen, A., and Flemming, J.M. 2015. Fast fitting of non-Gaussian state-space models to animal movement data via Template Model Builder. *Ecology* 96: 2598–2604. doi:10.1890/14-2101.1.
- Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Chapman & Hall, London.
- Bolker, B.M., Gardner, B., Maunder, M., Berg, C.W., Brooks, M., Comita, L., Crone, E., Cubaynes, S., Davies, T., de Valpine, P., Ford, J., Gimenez, O., Kéry, M., Kim, E.J., Lennert-Cody, C., Magnusson, A., Martell, S., Nash, J., Nielsen, A., Regetz, J., Skaug, H., and Zipkin, E. 2013. Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. *Methods Ecol. Evol.* 4: 501–512. doi:10.1111/2041-210X.12044.
- Brodziak, J., Ianelli, J., Lorenzen, K., and Methot, R.D., Jr. 2011. Estimating natural mortality in stock assessment applications. NOAA Technical Memo. NMFS-F/SPO-119.
- Clark, W.G. 1999. Effects of an erroneous natural mortality rate on a simple age-structured stock assessment. *Can. J. Fish. Aquat. Sci.* 56: 1721–1731. doi:10.1139/f99-085.
- Clark, W.G., Hare, S.R., Parma, A.M., Sullivan, P.J., and Trumble, R.J. 1999. Decadal changes in growth and recruitment of Pacific halibut (*Hippoglossus stenolepis*). *Can. J. Fish. Aquat. Sci.* 56: 242–252. doi:10.1139/f98-163.
- Deriso, R.B., Maunder, M.N., and Pearson, W.H. 2008. Incorporating covariates into fisheries stock assessment models with application to Pacific herring. *Ecol. Appl.* 18: 1270–1286. doi:10.1890/07-0708.1.
- Deriso, R.B., Quinn, T.J., II, and Neal, P.R. 1985. Catch-age analysis with auxiliary information. *Can. J. Fish. Aquat. Sci.* 42: 815–824. doi:10.1139/f85-104.
- Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Methods Softw.* 27: 233–249. doi:10.1080/10556788.2011.597854.
- Fournier, D., and Archibald, C.P. 1982. A general theory for analyzing catch at age data. *Can. J. Fish. Aquat. Sci.* 39: 1195–1207. doi:10.1139/f82-157.
- Francis, R.I.C.C. 2012. The reliability of estimates of natural mortality from stock assessment models. *Fish. Res.* 119–120: 133–134. Elsevier B.V. doi:10.1016/j.fishres.2011.12.005.

- Fu, C., and Quinn, T.J., II. 2000. Estimability of natural mortality and other population parameters in a length-based model: *Pandalus borealis* in Kachemak Bay, Alaska. *Can. J. Fish. Aquat. Sci.* 57: 2420–2432. doi:10.1139/cjfas-57-12-2420.
- Gordoa, A., and Hightower, J.E. 1991. Changes in catchability in a bottom-trawl fishery for cape hake (*Merluccius capensis*). *Can. J. Fish. Aquat. Sci.* 48:1887–1895. doi:10.1139/f91-224.
- Hanselman, D.H., Lunsford, C.R., Fujioka, J.T., and Rodgveller, C.J. 2007. Alaska Sablefish Assessment for 2008 [online]. In Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions. North Pacific Fishery Management Council, Anchorage, Alaska. pp. 195-312. <http://www.afsc.noaa.gov/refm/docs/2007/sablefish.pdf>
- Hanselman, D.H., Lunsford, C.R., and Rodgveller, C.J. 2015. Assessment of the sablefish stock in Alaska [online]. In Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions. North Pacific Fishery Management Council, Anchorage, Alaska. pp. 297-414. Available from <http://www.afsc.noaa.gov/REFM/Docs/2015/BSAISablefish.pdf>
- Hilborn, R. 2003. The state of the art in stock assessment: where we are and where we are going. *Sci. Mar.* 67: 15–20. doi:10.3989/scimar.2003.67s115.
- Hilborn, R., and Walters, C. J. 1987. A general model for simulation of stock and fleet dynamics in spatially heterogeneous fisheries. *Can. J. Fish. Aquat. Sci.* 44: 1366-1369.
- Hulson, P.-J.F., Hanselman, D.H., and Quinn, T.J., II. 2012. Determining effective sample size in integrated age-structured assessment models. *ICES J. Mar. Sci.* 69: 281–292. doi:10.1093/icesjms/fsr189.
- Hulson, P.-J.F., Quinn, T.J., II, Hanselman, D.H., and Ianelli, J.N. 2013. Spatial modeling of Bering Sea walleye pollock with integrated age-structured assessment models in a changing environment. *Can. J. Fish. Aquat. Sci.* 70: 1402–1416. doi:10.1139/cjfas-2013-0020.
- Ianelli, J., Holsman, K.K., Punt, A.E., and Aydin, K. 2015a. Multi-model inference for incorporating trophic and climate uncertainty into stock assessments. *Deep Sea Res. II* 134: 379-389. doi:10.1016/j.dsr2.2015.04.002.
- Ianelli, J.N., Honkalehto, T., Barbeaux, S., and Kotwicki, S. 2015b. Assessment of the walleye pollock stock in the Eastern Bering Sea [online]. In Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions. North Pacific Fishery Management Council, Anchorage, Alaska. pp. 53–152. Available from <http://www.afsc.noaa.gov/REFM/Docs/2015/EBSpollock.pdf>.
- Jiao, Y., Smith, E.P., O'Reilly, R., and Orth, D.J. 2012. Modelling non-stationary natural mortality in catch-at-age models. *ICES J. Mar. Sci.* 69: 105–118. doi:10.1093/icesjms/fsr184.

- Johnson, K.F., Monnahan, C.C., McGilliard, C.R., Vert-pre, K.A., Anderson, S.C., Cunningham, C.J., Hurtado-Ferro, F., Licandeo, R.R., Muradian, M.L., Ono, K., Szuwalski, C.S., Valero, J.L., Whitten, A.R., and Punt, A.E. 2015. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. *ICES J. Mar. Sci.* 72: 137–150. doi:10.1093/icesjms/fsu055.
- Kimura, D.K., Shimada, A.M., and Shaw, F.R. 1998. Stock structure and movement of tagged sablefish, *Anoplopoma fimbria*, in offshore northeast Pacific waters and the effects of El Niño-Southern Oscillation on migration and growth. *Fish. Bull.* 96: 462–481.
- Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B. 2016. TMB: automatic differentiation and Laplace approximation. *J. Stat. Softw.* 70(5): 1–21. doi:10.18637/jss.v070.i05.
- Martell, S., and Stewart, I. 2014. Towards defining good practices for modeling time-varying selectivity. *Fish. Res.* 158: 84–95. Elsevier B.V. doi:10.1016/j.fishres.2013.11.001.
- Marty, G.D., Hulson, P.-J.F., Miller, S.E., Quinn, T.J., II, Moffitt, S.D., and Merizon, R. A. 2010. Failure of population recovery in relation to disease in Pacific herring. *Dis. Aquat. Organ.* 90: 1–14. doi:10.3354/dao02210.
- Maunder, M.N., and Watters, G.M. 2003. A general framework for integrating environmental time series into stock assessment models: model description, simulation testing, and example. *Fish. Bull.* 101: 89–99.
- Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. *ICES J. Mar. Sci.* 56: 473–488. doi:10.1006/jmsc.1999.0481.
- Munk, K.M. 2001. Maximum ages of groundfishes in waters off Alaska and British Columbia and considerations of age determination. *Alaska Fish. Res. Bull.* 8: 12–21.
- Myers, R.A. 1998. When do environment-recruitment correlations work? *Rev. Fish Biol. Fish.* 8: 285–305. doi:10.1023/A:1008828730759.
- Quinn, T.J., and Deriso, R.B. 1999. *Quantitative fish dynamics*. Oxford University Press, New York.
- R Development Core Team. 2016. *R: a language and environment for statistical computing* [online]. R Foundation for Statistical Computing, Vienna, Austria. Available from <http://www.R-project.org>.
- Sasaki, T. 1985. Studies on the sablefish resources in the North Pacific Ocean. *Bulletin 22, Far Seas Fishery Laboratory, Shimizu, Japan*.
- Schnute, J.T., and Richards, L.J. 1995. The influence of error on population estimates from catch-age models. *Can. J. Fish. Aquat. Sci.* 52: 2063–2077. doi:10.1139/f95-800.

- Shotwell, S.K., Hanselman, D.H., and Belkin, I.M. 2012. Toward biophysical synergy: investigating advection along the Polar Front to identify factors influencing Alaska sablefish recruitment. *Deep-Sea Res. II Top. Stud. Oceanogr.* 107: 40–53. doi:10.1016/j.dsr2.2012.08.024.
- Skaug, H.J., and Fournier, D.A. 2015. Random effects in AD Model Builder: ADMB-RE user guide.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and van der Linde, A. 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B.* 64: 583–639. doi:10.1111/1467-9868.00353.
- Stewart, I.J., and Martell, S.J.D. 2015. Reconciling stock assessment paradigms to better inform fisheries management. *ICES J. Mar. Sci.* 72: 2187–2196. doi:10.1093/icesjms/fsv061.
- Teerlink, S.F. 2011. Considering predation by humpback whales in the mystery of the failed recovery of Pacific herring in Prince William Sound, Alaska. M.S. thesis, School of Fisheries and Ocean Sciences. The University of Alaska Fairbanks, Juneau, AK.
- Thompson, G.G. 1994. Confounding of gear selectivity and the natural mortality rate in cases where the former is a nonmonotone function of age. *Can. J. Fish. Aquat. Sci.* 51: 2654–2664. doi:10.1139/f94-265.
- Thorson, J.T., Hicks, A.C., and Methot, R.D. 2015. Random effect estimation of time-varying factors in Stock Synthesis. *ICES J. Mar. Sci.* 72: 178–185. doi:10.1093/icesjms/fst211.
- Thorson, J.T., Johnson, K.F., Methot, R.D., and Taylor, I.G. 2017. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. *Fish. Res.* 192: 84-93. doi:10.1016/j.fishres.2016.06.005.
- Thorson, J.T., and Minto, C. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. *ICES J. Mar. Sci.* 72: 1245–1256. doi:10.1093/icesjms/fsu213.
- Trenkel, V.M., Bravington, M.V., and Lorange, P. 2012. A random effects population dynamics model based on proportions-at-age and removal data for estimating total mortality. *Can. J. Fish. Aquat. Sci.* 69: 1881–1893. doi:10.1139/f2012-103.
- Truesdell, S.B., Bence, J.R., Syslo, J.M., and Ebener, M.P. 2017. Estimating multinomial effective sample size in catch-at-age and catch-at-size models. *Fish. Res.* 192: 66-83. doi:10.1016/j.fishres.2016.11.003.
- De Valpine, P., and Hilborn, R. 2005. State-space likelihoods for nonlinear fisheries time-series. *Can. J. Fish. Aquat. Sci.* 62: 1937–1952. doi:10.1139/f05-116.
- Walters, C.J. 1969. A generalized computer simulation model for fish population studies. *Trans. Am. Fish. Soc.* 98: 505–512. doi:10.1577/1548-8659(1969)98.
- Wilberg, M.J., and Bence, J.R. 2006. Performance of time-varying catchability estimators in statistical catch-at-age analysis. *Can. J. Fish. Aquat. Sci.* 63: 2275–2285. doi:10.1139/f06-111.

Wilberg, M.J., Thorson, J.T., Linton, B.C., and Berkson, J.I.M. 2010. Incorporating time-varying catchability into population dynamic stock assessment models. 18: 7–24. doi:10.1080/10641260903294647.

Zheng, J., Murphy, M.C., and Kruse, G.H. 1995. A length-based population-model and stock-recruitment relationships for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. Can. J. Fish. Aquat. Sci. 52: 1229–1246. doi:10.1139/f95-120.