Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2023/022

Maritimes Region

A Framework for Conditioning Operating Models for the Southwest Nova Scotia/Bay of Fundy Spawning Component of 4VWX Herring

T.R. Carruthers[1], A.R. Hordyk[1], Q.C. Huynh[1], R. Singh[2], T.J. Barrett[2]

[1]Institute for the Oceans and Fisheries
University of British Columbia
Aquatic Ecosystems Research Laboratory, 2202 Main Mall
Vancouver, BC V6T 1Z4

[2]Population Ecology Division
Fisheries and Oceans Canada
St. Andrews Biological Station
125 Marine Science Drive
St. Andrews, NB E5B 0E4

June 2023

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044
ISBN 978-0-660-48729-8 Cat. No. Fs70-5/2023-022E-PDF

Correct citation for this publication:

Carruthers, T.R., Hordyk, A.R., Huynh, Q.C., Singh, R., and Barrett, T.J. 2023. A Framework for Conditioning Operating Models for the Southwest Nova Scotia/Bay of Fundy Spawning Component of 4VWX Herring. DFO Can. Sci. Advis. Sec. Res. Doc. 2023/022. v + 103 p.

Aussi disponible en français :

Carruthers, T.R., Hordyk, A.R., Huynh, Q.C., Singh, R., et Barrett, T.J. 2023. Cadre de conditionnement des modèles opérationnels pour la composante de reproducteurs du sud-ouest de la Nouvelle-Écosse et de la baie de Fundy du hareng des divisions 4VWX. Secr. can. de consult. sci. du MPO. Doc. de rech. 2023/022. vi + 111 p.

TABLE OF CONTENTS

ABSTRACT

An approach for conditioning operating models is described and demonstrated for use in a management strategy evaluation for the southwest Nova Scotia / Bay of Fundy spawning component of Atlantic Herring in NAFO area 4VWX. This document (1) provides a detailed account of operating model structure and estimation methods; (2) describes a reference case operating model; (3) investigates sensitivities of conditioning to central uncertainties identified for Atlantic Herring; (4) evaluates the impact of these sensitivities to determine a reference set grid of operating models spanning natural mortality rate, growth, resilience, and historical catch levels; (5) identifies those reference set operating models with the most contrasting implications for MSE projection for use in robustness operating models; (6) identifies outstanding sources of uncertainty and specifies robustness operating models that address these uncertainties; (7) references supporting documentation that allow for reproduction of all results. The reference set of operating models spanned a range of current stock status, magnitude of the current stock and the sustainable rate of exploitation. Additionally, 25 robustness operating models were specified that encompassed an additional eight sources of uncertainty in Herring fishery dynamics and can be used to further discriminate among candidate management procedures.

INTRODUCTION

SOUTHWEST NOVA SCOTIA / BAY OF FUNDY HERRING FISHERY AND MANAGEMENT

The Atlantic Herring, *Clupea harengus*, is a coastal pelagic species found on both sides of the North Atlantic Ocean. Herring are a schooling fish that form predictable aggregations for feeding, over-wintering, and spawning. The fishery for Atlantic Herring off southwest Nova Scotia and the Bay of Fundy, in NAFO area 4VWX, is one of the largest and oldest fisheries in the region. 4VWX Herring are caught by multiple gear types, including purse seine which accounts for 80–90% of the current total catches, with gillnet, weir, shutoff, and trap gears making up the remaining fraction of the catch (Singh et al. 2020). The majority of Herring in this region are fall spawners, forming large aggregations in several sites that are targeted by the purse seine fleet. Since 2002 there has been no directed winter fishery. Most Herring products from Atlantic Canada are exported to countries including Japan and the United States in fresh, frozen, smoked, and other forms.

The 4VWX Herring management unit contains a number of spawning areas, separated to various degrees in space and time. Spawning areas in close proximity with similar spawning times, and which share a larval distribution area, are considered part of the same component. Some spawning areas are large and offshore, whereas others are small and more localized, sometimes very near shore or in small embayments. The stock structure is complicated further as Herring migrate long distances and mix outside of the spawning period both with members considered part of the same component and with members of other components. For the purposes of evaluation and management, the 4VWX Herring fisheries are divided into four components:

1. Southwest Nova Scotia/Bay of Fundy (SWNS/BoF) spawning component;
2. Offshore Scotian Shelf banks spawning component;
3. Coastal (South Shore, Eastern Shore and Cape Breton) Nova Scotia spawning component;
4. Southwest New Brunswick (SWNB) migrant juveniles.

Each component, with the exception of SWNB migrant juveniles, has several spawning areas and there is mixing of fish among spawning components outside of the spawning period. The SWNB migrant juvenile weir fishery occurs within the spatial bounds of the SWNS/BoF area and are considered a separate component; however, weirs capture a mixture of Herring from different spawning areas.

The SWNS/BoF spawning component is the largest and the fishery on this component is presently managed by an annual Total Allowable Catch (TAC). The main fall spawning sites are German Bank, Scots Bay, and Trinity Ledge areas (Figure 1) but fishing also occurs on feeding aggregations outside of these spawning locations. In the past, additional weir fishing occurred in Nova Scotia along the Long Island shore within the Bay of Fundy; however, only a few weirs presently operate in the upper Bay of Fundy in Minas Basin. The gillnet fishery southeast of Yarmouth have fished in the SWNS/BoF stock quota area and hence landings are part of the quota. Gillnetting has taken place on Trinity Ledge and in the Spectacle Buoy area in the past (Power et al. 2013). More recently, gillnet fishing has also occurred on German Bank (since 2005) and in Scots Bay (since 2009; Singh et al. 2020).

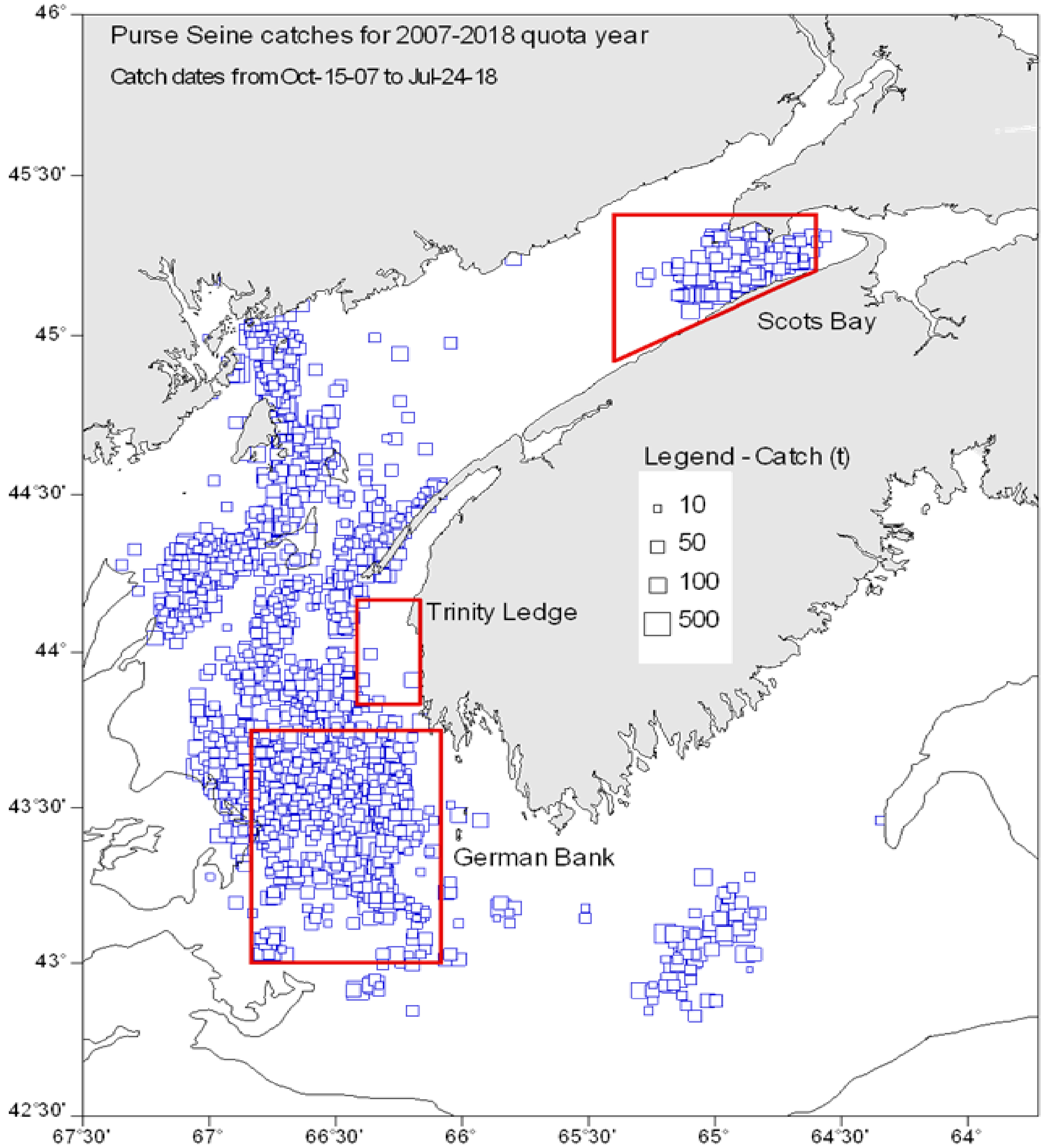


Figure 1. Map of Southwest Nova Scotia / Bay of Fundy region with total purse seine catches of Herring from 2007–2018 (blue squares) and the three spawning grounds indicated as red rectangles.

Landing records for the 4VWX commercial fishery extend back to the mid-1960s, when annual catches were greater than 140,000 tonnes (Figure 2). Catches tended to decline from this peak in the early years, with one of the lowest historical catch records occurring in 1978 (Figure 2). The apparent decline in resource abundance during the 1970s led to the introduction of annual TAC limits (Figure 2). Landings have tracked the TAC since 2002. The TAC was reduced to 50,000 tonnes (t) in 2005, followed by a reduction to 42,500 t in 2017 (Figure 2).

There have been apparent declines in abundance of spawning fish at some of the areas that were historically important to the fishery. For example, Trinity Ledge once supported a major portion of the overall catch but landings since the mid-2000s has been very low levels in response to the low estimate of spawning biomass from acoustic surveys in this region. The majority of the catch from the spawning grounds now comes from two areas: German Bank and Scots Bay (Figure 1).

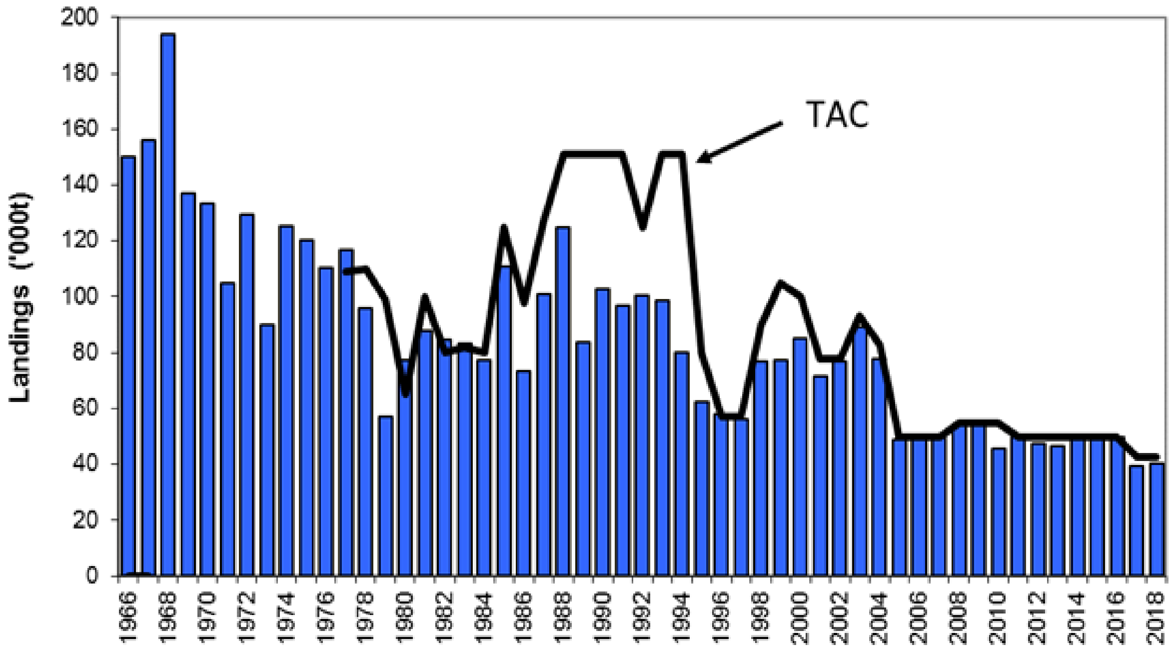


Figure 2. Time-series of the historical landings of Atlantic Herring in 4VWX (blue bars) and the Total Allowable Catch (TAC; black solid line). Years in the figure represent the "quota year" defined as Oct 15th of the previous year to Oct 14th of the quota year.

A previous investigation of stock assessment methods for the SWNS/BoF Herring stock revealed a number of concerns including conflicts between data (age composition and acoustic survey biomass), potentially misrepresentative age composition data (fishery composition are assumed to be representative of surveys) and a potentially misrepresentative acoustic survey due to incomplete spatio-temporal coverage (DFO 2011). Most of the assessment methods reviewed estimated a 5x lower stock size than was inferred by the acoustic survey (the explanation for this could not be determined at the meeting). DFO (2011) concluded that there were no reliable approaches available at that time to provide medium-term forecasts, and that short-term approaches to support management decision making should rely on "interpreting indicator levels and trends".

Subsequently, SWNS/BoF Herring has been managed with an informal harvest strategy relying on a variety of indicators such as observed survey biomass levels, fishermen input, age composition data, and trends in relative exploitation and total mortality rates. These data are interpreted in reference to management objectives that are stated in the fishery management plan (DFO 2003). Since there is no accepted analytical stock assessment, and central abundance indices only exist for 1999 onwards, conventional target and limit reference points relating to "unfished biomass" are not available. Instead, a limit reference point (LRP) was established as the mean 2005–2010 acoustic survey biomass (Clark et al. 2012).

ATLANTIC HERRING: UNCERTAINTIES

Recent assessment documents have outlined a number of central uncertainties which include the catchability of the acoustic survey, future recruitment strength (i.e., high uncertainty in short term abundance forecasts), the correct interpretation of catch composition data, and the possible continuation of observed historical changes in weight-at-age (DFO 2015, Melvin et al. 2014).

In other fisheries in the wider region, changing ocean conditions and increased numbers of marine predators have been linked to changes in distribution, growth, and survival (Neuenhoff et al. 2018).

The general areas of uncertainty were identified at a recent Atlantic Herring MSE working group meeting (Figure 3). These uncertainties and other example uncertainties that have been identified for this stock are presented in Table 1.

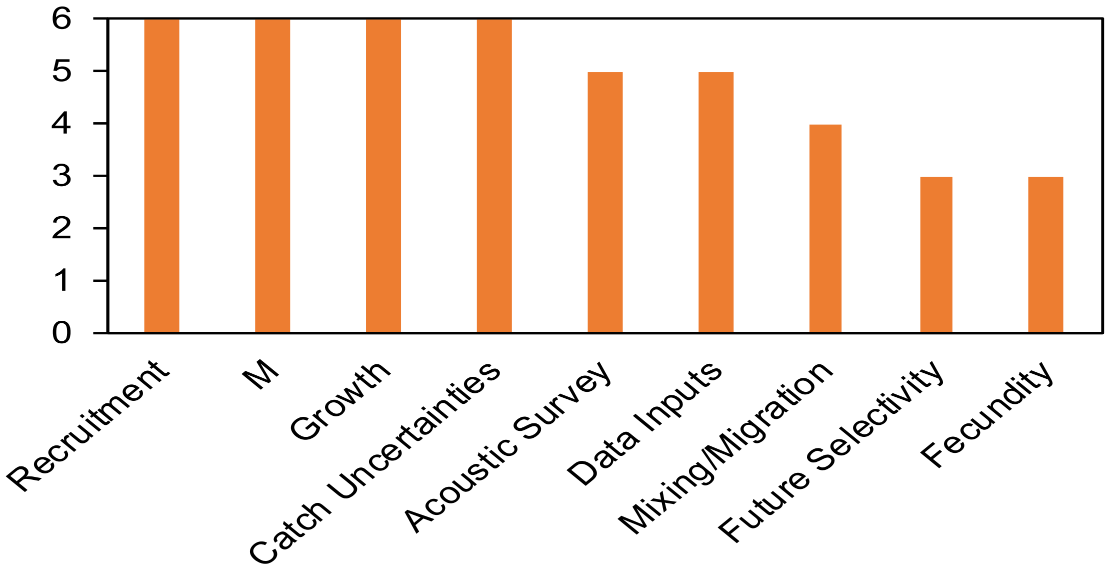


Figure 3. General areas of uncertainty in Atlantic Herring fishery dynamics identified by the Atlantic Herring Working Group in January 2020. The bars each represent a source of uncertainty in Atlantic Herring fishery dynamics. The height of the bar is the frequency with which the uncertainty was identified by each of six sub groups. For a description of the these sources of uncertainty see Table 1 below.

Table 1. Uncertainties identified by previous scientific processes, initial discussions with the Atlantic Herring working group, and from latest data sources.

Uncertainty	Data conflict
Recruitment	Includes uncertainty in variability and strength of recruitment in addition to resilience (e.g., steepness of the stock-recruitment relationship) and carrying capacity. Also includes different stock recruitment relationships and consideration of stochasticity. Potential drivers to changes in recruitment were identified by the Atlantic Herring Working Group in January 2020 to be changes in prey fields, phenology, temperature, ecosystem changes, egg predation, and larval condition. A relatively common axis of uncertainty in Management Strategy Evaluations, alternative plausible recruitment scenarios can impact estimated historical stock trajectory, current stock depletion, and future outcomes of alternative management procedures (e.g., rebuilding evaluations)
Natural mortality rate (M)	There is considerable uncertainty over temporal changes in natural mortality rate, hypothetically driven by changing ocean conditions, food availability, and the increasing presence of marine predators. There is additional uncertainty in natural mortality rate-at-age. Previous stock assessments have assumed a constant M of 0.2 (DFO 2004). Alternative age-variant M scenarios have been proposed for this stock based on predator consumption rates in a multispecies virtual population analysis (Guénette and Stephenson 2012) but have not been adopted in assessments. Herring assessments in the Gulf of Maine have used either a low age- and time-invariant instantaneous natural mortality rate (0.35) or a substantially high rate that declines with age (NEFSC 2012, Deroba 2015, Deroba 2017). Uncertainty in natural mortality rate is common in most fishery assessment settings due to a lack of independent data with which to reliably estimate it. Problematically, the typical level of uncertainty in M often corresponds with diverging conclusions about productivity, stock status, and resilience to current exploitation levels.
Growth	Unexplained changes in somatic growth (e.g., weight-at-age) have been observed over the period of fishery exploitation (DFO 2015). These changes could be due to changes in ocean temperatures, diet, density dependence, or productivity changes in the environment. If such changes continue in the future they have implications for both the effectiveness of proposed fishery management methods and fishery reference points relating to ‘unfished’ conditions.
Catches	A principal source of uncertainty relates to assigning catches correctly to management area. For example, it has been historically hypothesized that weir catches relate to stocks outside of Canadian fishery management areas. Additionally, a relatively small harvest of Herring still occurred prior to the 1960s and the onset of significant industrial catches and the collection of relative abundance indices. Assessment models assuming unfished conditions after these catches may not fully account for stock depletion impacting unfished stock size, current depletion, and future stock productivity.

Uncertainty	Data conflict
Acoustic Survey	There is some uncertainty about whether the annual acoustic surveys (conducted on the spawning stock of 4VWX Herring since 1999; Singh et al. 2016) should be considered as a relative (proportional, with constant of proportionality q, estimated) or an absolute (q = 1) index of vulnerable biomass. Based on the assumption that the spawning Herring turn over every 14 days, the acoustic surveys are separated by 10–14 days to account for double counting. The surveys use a target strength from the generic clupeid equation from Foote (1987), with estimates of the length-weight relationship estimated from sampling conducted throughout the survey period, to estimate the total spawning biomass of 4VWX Herring (Singh et al., 2016). The acoustic survey results have been used in previous assessments as absolute estimates of Spawning Stock Biomass (DFO, 2005). However, more recently, the biomass estimated by the acoustic survey has been considered as a relative index of abundance. Previous explorations of assessments (DFO 2011) found that composition data implied a substantially smaller absolute vulnerable biomass (2.5–7 times less) than that inferred by the Acoustic Survey.
Mixing and Migration	This includes uncertainties in the Canadian stock structure, migration within the management areas, and transboundary mixing into U.S. waters. Uncertainties over mixing and migration are closely related to correct assignment of historical **catches**. Tagging studies conducted in U.S. waters (Kanwit and Libby 2009) found substantial in-migration to Canadian waters. Depending on the relative size of Canadian and U.S. resources, the comparative recovery rates, and the directionality of migration, this could imply either catch over reporting of Canadian fish (erroneously high due to catch of U.S. fish) or under reporting of Canadian fish (due to exploitation of Canadian fish entering U.S. waters) since Canadian tagged fish has also been recovered in U.S. waters (Clark 2006).
Future Selectivity	This relates to the length-based selectivity of the fishing fleet driven by changes in fishing practices (e.g., towards a bait fishery), technical management measures or changing availability to fishing of fish of varying size.
Larval survey	The larval survey is intended for use as an index of Spawning Stock Biomass but measurement error relates only to density of larvae not the Spawning Stock Biomass from which larvae were produced. The relationship between the index and Spawning Stock Biomass is a key source of uncertainty and a strong determinant of early stock trends where only this index and no composition data were available.
Unreporting (bait)	Closely related to uncertainties relating to **catches** and **future selectivity**, catch reporting affects both the reconstruction of the historical fishery dynamics but also affects the success of management procedures if there is exploitation not accounted for in the simulated data. The uncertainty in unreporting is stemming from the recent shortage of bait in eastern Canada and the U.S.
Data Inputs	In addition to the **larval survey**, the available fishery data may be included in operating model conditioning in various ways. For example, operating model conditioning may lead to varying conclusions over stock status and projections given fits to additional length composition data or varying weight of relative abundance indices.

Uncertainty	Data conflict
Fecundity	Hypothesized changes in growth and mortality due to environmental conditions may also be linked to changes in condition factor and size-at-maturity which may impact the estimated stock productivity and future outcomes of candidate management procedures. There is also uncertainty in the effect of changes in weight-at-age on the reproductive capacity of the stock (e.g., changes in reproductive capacity may not be proportional to changes in Spawning Stock Biomass because fish are smaller).

MANAGEMENT STRATEGY EVALUATION

Management Strategy Evaluation (MSE) is an approach for establishing simple rules for managing a fishery by simulation testing their robustness to various hypothetical scenarios for fishery dynamics (Butterworth and Punt 1999, Cochrane et al. 1998) (Figure 4). Often referred to as Management Procedures (MPs, aka Harvest Strategies) these rules typically use streamlined data such as catches and a relative abundance index to generate management advice such as a Total Allowable Catch (TAC).

MSE differs substantially from conventional stock assessment in how models of fisheries dynamics (that approximate stock and exploitation dynamics) are used to derive management advice. In conventional stock assessment, fisheries dynamics models ('stock assessments') are used to directly derive management advice. For example, setting a TAC commensurate with fishing mortality rate at maximum sustainable yield (F_{MSY} levels). MSEs typically use a greater number of fitted fisheries dynamics models ('operating models') that span a much wider range of uncertainties in order to test the robustness of MPs. It follows that MSE allows managers and stakeholders to establish a comparatively simple management rule (the MP), understand its performance and have confidence that it can perform adequately even when uncertainty in dynamics may be high.

MSE involves closed-loop simulation whereby a candidate MP is applied recursively into the simulated future accounting for feedbacks with the operating model (E, Figure 4). Closed-loop simulation requires not only an operating model and MPs but an observation error model that can generate simulated data to be inputted to the MP, and an implementation error model that determines how well the management advice provided by the MP is adhered to in the simulation.

Punt et al. (2014) provide a comprehensive summary of the history of MSE implementations that, starting in the 1980s, covers more than 30 stocks including several short-lived pelagic species such as Herring. In the face of difficulties in establishing defensible stock assessment models (including retrospective biases, data conflicts, numerical convergence problems) and the recognition of uncertainties that cannot be accounted for the provision of advice within the assessment paradigm, MSE is increasingly adopted as a framework for the selection of management procedures for fisheries in Canada (Kronlund et al. 2013), the North Atlantic (NAFO 2010, 2018), and California (Hordyk et al. 2017, CDFW 2018). Additionally, once an MP is adopted, empirical data can be gathered and compared to those data predicted by the operating model, to indicate whether the operating models should be revised (aka exceptional circumstances, Carruthers and Hordyk 2018a).

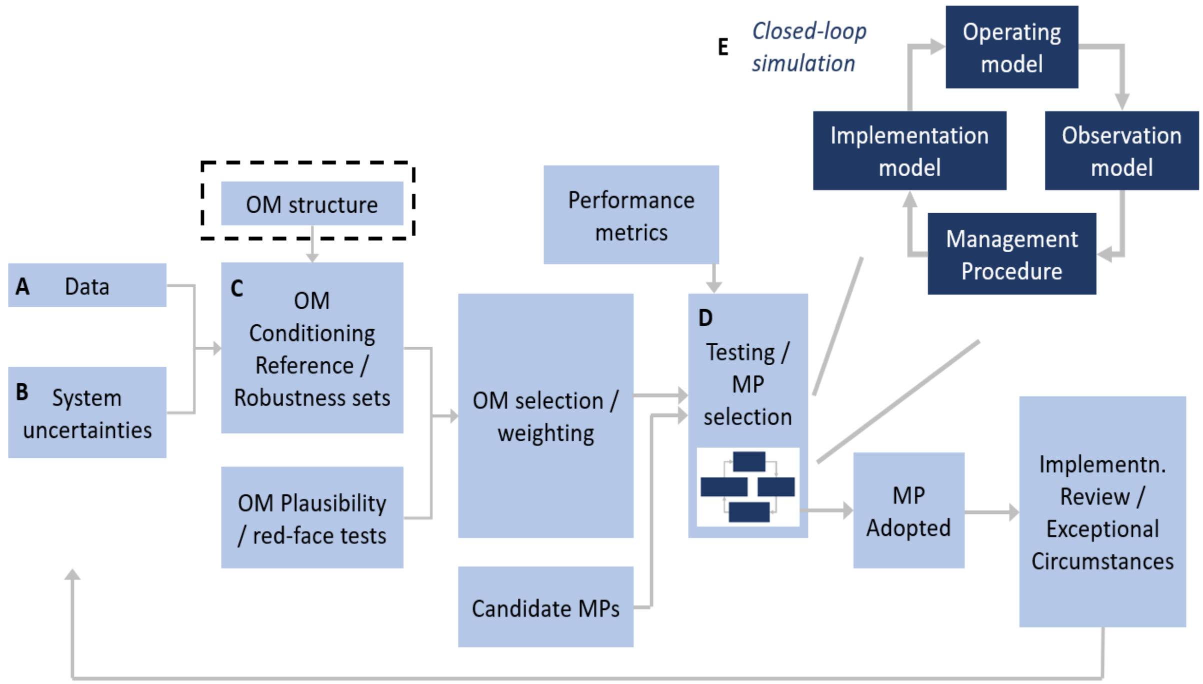


Figure 4. A Management Strategy Evaluation (MSE) process including closed-loop simulation. The focus of this document is the structure of the Operating Model (OM) (highlighted by a dashed box) and its capability to accept the data types currently available for Herring (A) across a range of system uncertainties (B) in a suitable model conditioning framework (C) that is compatible with existing software (MSEtool) for conducting closed-loop simulation testing of management procedures (D and E).

MSE IN THE CONTEXT OF ATLANTIC HERRING

For several reasons the Herring fishery in SWNS/BoF is an excellent candidate for MSE and the MP approach to fishery management. MSE was specifically designed to navigate situations where it is difficult to establish a defensible stock assessment for providing advice and uncertainties in extra-assessment dynamics are relatively high (Punt and Donovan 2007) (e.g., changing productivity of spawning grounds, scenarios for natural mortality rate). Operating models can also be used to establish MPs for differing inferences of the available data (e.g., southern bluefin tuna), offering a way forward in situations where there are prevailing data conflicts (e.g., the Acoustic survey and catch-at-age composition data).

MSE has the advantage that MP performance is evaluated against the operating model (the ‘true’ system) that has known biomass levels and reference points. Consequently, even in situations such as Herring where assessments, reference points, and hence *explicit* estimates of current stock status may be highly uncertain, the *implicit* performance of MPs can still be evaluated (exemplified by statements such as “MP X can be expected to exceed the limit reference point after 15 years for all operating models”).

Some fisheries, for example trans-boundary fisheries for highly migratory tuna species, are often poor candidates for MSE. These fisheries can have very high catch overages which disentangles management decisions from fishery performance and hence weakens the potential benefit of establishing an MP. In contrast, Herring has the significant advantage that management advice is generally well adhered to, and there are close links with a local fleet and organized industry groups.

Lastly, and not insignificantly, Herring is unusual in that the end product of an MSE - an (informal) MP linked to an index and index-based limit reference point - is already in use but whose theoretical performance is relatively unknown.

MSE has traditionally focused on the performance of MPs in closed-loop testing (E, Figure 4) but establishing a working set of operating models also allows for additional benefits to managers such as testing the theoretical cost-benefit of other data collection protocols (simulated in the observation model), enforcement strategies (simulated in the implementation model) and the identification of the most important knowledge gaps (which operating model uncertainties most strongly determine management performance).

THE MANAGEMENT STRATEGY EVALUATION TOOLKIT (MSETOOL)

Although MSE has many important advantages over conventional stock assessment-based management, historically MSE processes have tended to be relatively expensive, technically complex and time consuming.

However, since 2017, a DFO-UBC partnership agreement (DFO 2017) has supported the development of sophisticated open source R packages for MSE: the Data Limited Methods toolkit, (DLMtool; Carruthers and Hordyk 2018b,2019) and the Management Strategy Evaluation toolkit (MSEtool; Huynh et al. 2019). After years of development these are amongst the fastest most flexible and extensible open-source software packages for conducting MSE for fisheries in the full spectrum from data-poor (e.g., prescriptive management such as size limits and time-area closures) to data-rich (e.g., Statistical Catch-At-Age models linked with harvest control rules).

The MSEtool package contains computationally efficient functions for conditioning operating models on a wide range of fishery data types (greater detail on the stock reduction analysis used for conditioning Herring models is provided in the Methods and Appendix B). Also included in MSEtool are a wide range of data-rich MPs based on VPA, SRA, statistical catch-at-age, delay- difference and surplus production assessments. These provide a basis for evaluating the cost-benefit of using more complex approaches for provision of management advice. Additionally, MSEtool contains state-of-the-art exceptional circumstances protocols for empirically evaluating the suitability of an MP when it is in use (Carruthers and Hordyk 2018a).

An advantage of the MSEtool software is that it is computationally efficient (arguably the fastest MSE framework yet developed) and it has a high degree of flexibility in operating model structure, for example allowing for complex spatial dynamics, age-based movement, multi-stock dynamics and multi-fleet control rules. As such, the package can be used to investigate the robustness of candidate MPs to uncertainties that may be relevant to the Herring fishery (e.g., Table 1). Since MSEtool allows for tracking of a wide range of MSE data, it is straightforward to develop custom MPs that reflect the interests of various stakeholders and also reflect the differential value of fish of varying sizes (e.g., reflecting higher prices of smaller fish processed for bait or as canned sardines). MP development is further aided by more than 100 example MPs that are included in the R packages, from which tailor-made MPs can be adapted. Additionally, an online library of operating models is available that provides an extensive test-bed for a candidate management procedure.

MSEtool has extensive help documentation and user guides that allow DFO scientists and stakeholders to develop skills in the software and provide reference materials in support of future customization . Importantly MSEtool will remain open source and the operating model has published in the primary literature (MSEtool shares the same operating model as DLMtool; Carruthers and Hordyk 2018b,2019). All products developed using this software are made freely available, a comprehensive manual for the Herring MSE framework will be provided ensuring that it can be run by other analysts and readily adapted for future data, alternative performance metrics and revised MPs.

DLMtool and MSEtool are currently used by DFO, the California Department of Fish and Wildlife (Hordyk et al. 2017), the Marine Stewardship Council, the International Commission for the Conservation of Atlantic Tunas, and the US NOAA as MSE frameworks for the testing of management procedures, identifying data collection priorities and quantifying management reference points.

Through these collaborations, MSEtool and DLMtool have been subject to independent review by U.S. NOAA and peer review in the primary literature (Carruthers and Hordyk 2018b). These packages were used by the U.S. NOAA to establish operating models for six tropical reef fish in the Caribbean (SEDAR 2016a) and eight stocks in the Gulf of Mexico (SEDAR 2016b), the MP for lane snapper was used to set the catch limit currently implemented in that fishery. The U.S. Mid-Atlantic Fisheries Management Council also used DLMtool to identify acceptable biological catch limits (McNamee et al. 2016) and develop operating models to test MPs for black sea bass, Atlantic mackerel, and blueline tilefish (Miller 2016, Wiedenmann et al. 2019).

MSEtool was used recently as the MSE framework for San Francisco Bay Herring (A.R. Hordyk, UBC, pers. Comm.) from which an MP has now been adopted. Additionally, the California Department of Fish and Wildlife are in the process of establishing MPs for 9 other in-shore fisheries in California (A.R. Hordyk, UBC, pers. comm.). Ongoing MSEtool applications include Atlantic swordfish (International Commission for the Conservation of Atlantic Tunas) and B.C. yelloweye rockfish (DFO).

The applicability of MSEtool to short-lived forage fish such as Herring has been demonstrated in the specification of example operating models for two Herring stocks in Region 4T (Carruthers 2019a,b), Capelin in the Gulf of St Lawrence (Carruthers 2017) and Butterfish in the mid-Atlantic (Carruthers et al. 2014).

REQUIREMENTS OF OPERATING MODELS

No single operating model is a definitive representation of the real fishery system (i.e., a stock assessment), rather the range of operating models established should aim to encompass a plausible set of fishery scenarios. These scenarios serve as a test-bed for candidate MPs such that were an MP adopted there would be confidence over its robustness to uncertainties in the fishery system (e.g., Table 1).

MSE processes typically separate operating models into a **reference set** and a **robustness set**. The performance of candidate MPs across the reference set are the central focus for MP selection and represent core uncertainties. A single member of the reference set, referred to here as the **reference case** operating model, is identified that has a plausible combination of assumptions (a ‘Base Case’) and can serve as a basis for comparison with various other operating models. The robustness set operating models are used to consider scenarios for which there is weaker empirical evidence but could be useful in providing further discrimination among candidate MPs that perform similarly given reference set operating models.

Operating models may require greater flexibility in structure and assumptions than a typical assessment model since they must be able to recreate robustness scenarios that would not typically enter a conventional (‘best available science’) stock assessment process. It is therefore important that operating models have considerable flexibility over model structure, parameterization and data types for conditioning.

While the process of conditioning operating models is comparable to that of conditioning assessment models there are more specific considerations that relate to the use of data and the specification of MPs. For example, it is generally recognized in stock assessment that it is a priority to fit trends in relative abundance indices (Francis 2011), however, this may be more important for operating models that are intended to test index-based MPs. In the case of Herring, the only available index for future use in an MP is the acoustic survey, which is the principal input to the current index-based management approach including the derived limit reference point. It follows that it is implicit that these data are indicative of vulnerable biomass and therefore operating models should not show problematic misfit to these data. The definition of ‘problematic’ has not been clearly resolved in previous MSE processes; however, if it can be ensured that an index is positively correlated with vulnerable biomass, other statistical properties can still be estimated and preserved in future simulations such as error, temporal autocorrelation in errors, and hyperstability/hyperdepletion (an index that responds slower or faster than real vulnerable biomass trends, respectively).

In most MSEs, a dedicated estimation model is used for conditioning and the estimated fishery dynamics are then ‘copied’ into an MSE framework for closed-loop testing of MPs. This can be a non-trivial process and relies on exact parity of dynamics equations, including order of equations, and can be disrupted by phenomena that are easily overseen such as rounding of estimated parameters and internal parameter scaling of estimation software. It follows that when reviewing operating model conditioning for MSE purposes it can be valuable to demonstrate recreation of estimated fishery dynamics within the intended MSE framework.

Since MSEs are defined as open processes involving stakeholder engagement, it is critical that the technical aspects of the process (the operating models, MPs, performance definitions and results) are all available for review. The strength of a proposed approach for conditioning operating models should therefore also be evaluated according to its accessibility, transparency, ease of use and quality of documentation.

ORGANIZATION OF DOCUMENTS

The analyses of this document are presented in a series of steps:

1. Specifying and conditioning a reference case operating model
    a. the types and formatting of data for conditioning
    b. the assumed dynamics of the fishery system (equations for population and fishery dynamics) and the likelihood functions and numerical algorithms for conditioning operating models to data
2. Conducting sensitivity analyses in order to identify core uncertainties for defining a reference set of operating models
3. Specifying and conditioning a reference set of operating models
4. Accounting for other uncertainties in the specification and conditioning of robustness set operating models

Substantial written detail is required to ensure reproducibility of each of these steps that could obscure a clear and concise narrative of the approaches taken. To allow for both, we present an overview of each component in the main text, providing explanations for various decisions that were made (data formatting approaches, operating model structure, estimation software). This more concise narrative is supplemented by Appendix B that contains the mathematical descriptions of the modelling approach.

AIM AND OBJECTIVES OF THIS DOCUMENT

The aim of this document is to describe the identification and conditioning of a potential reference set and robustness set of operating models for Herring encompassing uncertainties that have been previously identified for Herring fishery dynamics (Table 1).

A list of key objectives is included in Table 2. Since these objectives are addressed in both descriptions of methodology and the results of operating model conditioning, Table 2 also lists the pertinent document sections.

To further ensure transparency and aid in review, R computer code was provided on a shared drive for stakeholders and reviewers for all methods and results that are presented in this document.

Table 2. Key objectives of this document and relevant material

Objective	Relevant Section
1. Provide a detailed account of operating model dynamics, fitting protocols, and open-source code to carry out fits for the purposes of review	**Overview:** Methods Sections 'Overview of Operating Model and Conditioning' & 'Developing a Reference Case Operating Model' **Detail:** Appendix B
2. Establish a reference case operating model that provides a suitable basis for conditioning given various data types (see Table 3) including: absolute indices of abundance (acoustic survey), relative indices of Spawning Stock Biomass (larval survey), estimation of selectivity for surveys using age composition data (acoustic survey), fit to multiple fleets with estimation of selectivity by length.	**Overview:** Results Section 'Reference Case Model Conditioning'
3. Investigate operating model sensitivity to establish a reference set of operating models	**Overview:** Methods Section 'Specification of Sensitivity Operating Models' Results Section 'Sensitivity Operating Models'
4. Specify and condition the reference set of operating models	**Overview:** Methods Section 'Establishing a reference grid of operating models' Results Section: 'Reference Set Operating Models'
5. Specify and condition the robustness set of operating models	**Overview:** Methods Section 'Robustness Operating Models' Results Section 'Robustness Set Operating Models'
6. Demonstrate that OM conditioning is sufficiently computationally efficient to allow for real-time exploration of alternative model configurations, input data, and weightings in a workshop setting, for example.	None

Objective	Relevant Section
7. Provide open source computer code for the reproduction of all results.	https://github.com/z5a1n/herring_OM_conditioning

OUTSIDE THE SCOPE OF THIS DOCUMENT

All operating model conditioning in this report is for the specification of alternative scenarios for fishery dynamics as a test-bed (contingency test) of candidate management procedures. This document should not be interpreted as an assessment of stock status.

This document focuses on historical reconstruction of fishery dynamics by conditioning on data. Aspects of MSE relating to forward projection and simulation testing of MPs will be evaluated in subsequent review processes, including OM plausibility, management performance metrics, and statistical models for future projection of data implementation models (parts D and E of Figure 4).

METHODS

OVERVIEW OF DATA

The fishery data used in operating model conditioning correspond with the Canadian portions of NAFO 4X (west of the Baccaro line in SWNS) and NAFO 5Yb (Figure 5).

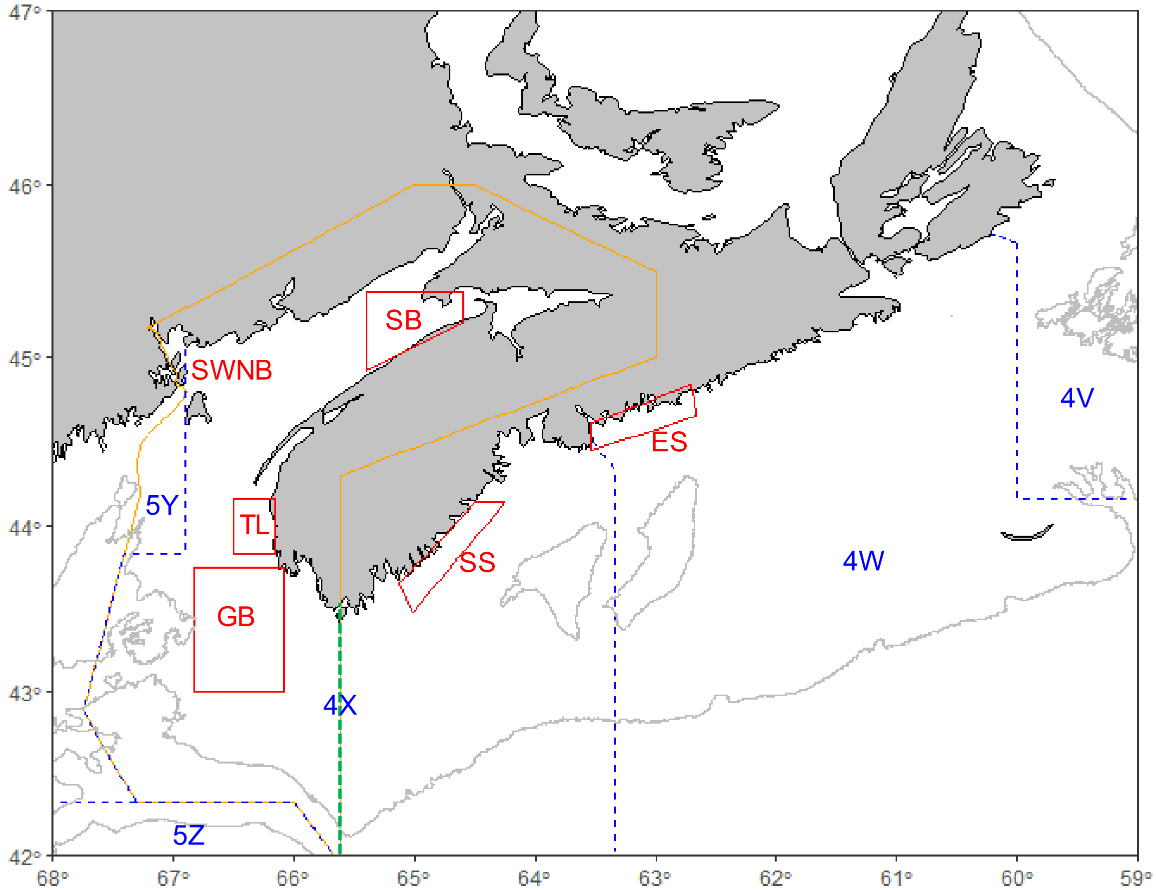


Figure 5. Map of Herring fishing areas in 4VWX. The Canadian area west of the green vertical (Baccaro) line (the orange polygon) is the SWNS/BoF Herring fishery area. SB = Scots Bay; GB = German Bank; TL = Trinity Ledge; SWNB = Southwest New Brunswick; SS = Coastal South Shore; ES = Coastal Eastern Shore.

For the time period 1968–2002, landings data originate from the COMLAND database. For 2003–2018 landings data come from the MARFIS database. Annual catch data included adjustment for unreported catches (Table A3), and were available for every year from 1968 to 2018 (Table 3; Figure 6). Landings were calculated by fleet using two different fleet structures (Table A9). The fleets were gillnet (GILL), purse seine (PS), weir and shutoff fishery (WEIR/SO) in New Brunswick, and OTHER which consists of Nova Scotia weir and all other gears types. The second fleet structure that was evaluated had 5 fleets where the PS fleet was divided into two fleets based on the size of fish landed in each fishing ground (Figure A17). The two purse seine fleets (PS_juv and PS_spa) were defined as purse seine landings that consisted of primarily juvenile and adult fish, respectively, based on catch length frequency distributions (Figure A16).

Length frequency data were available from 1968–2018 (Table 3; Figure 6). Catch-at-length composition data were converted from weight to numbers using a length-weight relationship, with parameters estimated by month, year, and fleet. A 1 cm length bin was used for the catch-at-length calculations and the catch-at-age was also evaluated using 0.5 cm length bins (Figures A10b and A10c). Effective sample sizes were calculated using the method from Pennington et al. (2002). When only one detailed sampling event was conducted for a year/fleet combination (occurred in some years for some gillnets), the effective sample size was estimated from the actual number of fish measured, adjusted using the average ratio of the actual number of fish caught to the effective sample size for years with multiple gillnet detailed sampling events.

Age samples were available from 1970–2018 so the catch-at-age time series began in 1970. Catch-at-age data were generated by converting numbers-at-length data to numbers-at-age using an Age-Length Key (ALK), defined by year and season. The seasons were defined to have sufficient age data for the ALKs, and were defined as:

1. Spring (January–May)
2. June
3. July
4. August
5. September/October
6. November/December

Missing ages in the ALK were estimated as follows in order in the list. If no ages for a length bin after step n then proceed to step n+1:

1. Combining data for the defined length from the seasons before and after
2. Fish < 10 cm are age 1 and fish > = 40 cm are 11+
3. Combining data for the given length from length groups before and after that length group
4. Manually adding the estimated ages

Effective sample sizes were calculated using the same methods described above for the catch-at-length data but were based on the number of fish aged. There were no detailed (age and weight) samples for some years for gillnets. In these years the detailed samples from the year prior were used. Catch-at-age was calculated and converted to proportions-at-age and then adjusted based on effective sample size.

An acoustic index is available from 1999–2018 (Table 3; Figure 6). The index was generated by summing the annual estimates of spawning biomass from the acoustic surveys in Scots Bay and German Bank. The acoustic index was also calculated as the sum of all acoustic surveys in 4X as well as using only the sum of the maximum survey in Scots Bay and German Bank in a given year. The implied depletion (estimated by comparing the mean biomass estimates from 1999–2001 to the mean estimates from 2016–2018 were similar to the current accepted index, therefore, these two alternative acoustic indices were not evaluated further. The acoustic survey has accompanying numbers-at-length data from the German Bank spawning box and the Scots Bay catch area. The windows for assigning biological data to the surveys for the purposes of estimating the numbers-at-age and biomass-at-age were +/- one day for length frequencies and +/- 5 days for detailed samples. When no samples were available within these windows then samples closest in date were used. Survey numbers-at-age were calculated from numbers-at-length using the same process described above for the fishery catch-at-age.

Empirical length-at-age and weight-at-age matrices were estimated by year (1970–2018) to be used by the model. Age-at-maturity and length-at-maturity were estimating using binary logistic regressions from year-specific estimates of age and length.

Table 3. Summary table of the available data. The column 'Used' indicates whether the data were used in the conditioning of one or more operating model scenarios. These data form the requirements for objective O2: a conditioning framework that can accept all or various combinations of these data types.

Data type	Description	Spatial range	Temporal range	Used*
Annual catches (by fleet)	Landings by gear and area. 4WX landings adjusted for unreported catches.	4VWX	1963–2018	1968–2018
Fishery length frequency data (by fleet)	Includes the survey grounds (Scots Bay and German Bank) and the areas outside of those grounds for the stock area.	SWNS/BoF	1965–2018	1968–2018
Fishery age composition data (by fleet)	Catch-At-Age (CAA) data (in numbers and weight) for the whole SWNS/BoF stock component by fleet.	SWNS/BoF	1965–2018	1970–2018
Survey age composition data	Breakdown of age by numbers and weight of sampled Herring found on spawning grounds during the acoustic surveys	Scots Bay and German Bank	1999–2018	1999–2018
Larval index	Annual plankton research surveys were completed in late October / early November in the SWNS/BoF area to determine larval Herring abundance. This series was ended for fiscal reasons and because the relationship between larval abundance and the spawning biomass was poor and of little predictive utility in the VPA (DFO 2007).	SWNS/BoF	1972–1998, 2009	1972–1998, 2009
Acoustic index	Structured acoustic surveys in specified spawning boxes (Scots Bay and German Bank). Sometimes transects are done just outside of those boxes. Biomass estimates are available for those areas within specified boxes and outside boxes. Data for Trinity Ledge also exist but this area is no longer a significant spawning area.	Scots Bay and German Bank	1999–2018	1999–2018
Tagging data	Several DFO reports (see reference list)	4VWX	1982–2018	

* raw data for landings and length frequency prior to 1968 were not in the database. Similarly, age data prior to 1970 were not in the database.

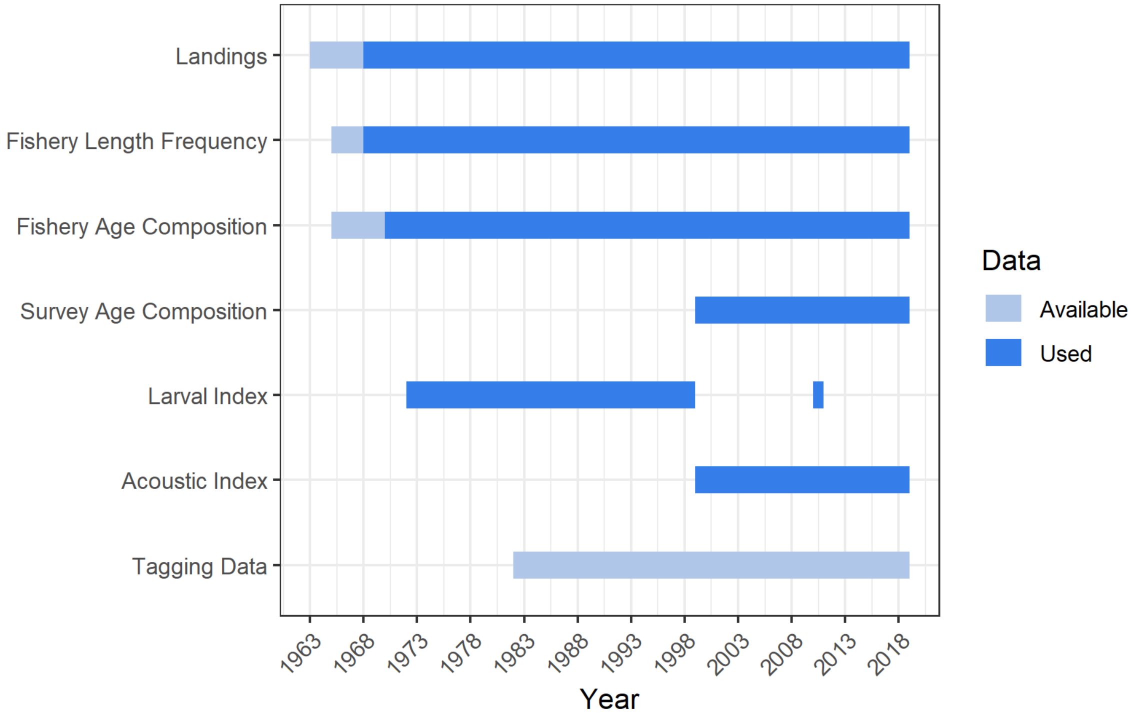


Figure 6. Data types and availability (light blue) and year ranges data were used for conditioning of operating models (dark blue).

OVERVIEW OF OPERATING MODEL AND CONDITIONING

A multi-fleet Stock Reduction Analysis (SRA; Walters et al. 2006) was developed in Template Model Builder (Kristensen et al. 2016) (see Appendix B for greater detail on model equations and conditioning). The SRA requires complete (all years, all fleets) catch data and any combination of other data types (that may be temporally patchy) including indices of abundance, catch-at-age composition data, catch-at-length composition data and mean length data. The model can interpret indices in various ways including relative/absolute measures of vulnerable/stock-wide, biomass/numbers/spawning biomass/biomass. Given age and/or length composition data, the model estimates time-invariant selectivities for fleets and surveys that are either logistic (asymptotic, 'flat-topped') or double-normal ('dome-shaped').

Indices and catches were fitted by log-normal likelihood functions while age composition and (optionally length composition) data were assumed to be distributed according to the multinomial (given the effective sample size correction of Pennington et al. 2002).

The SRA model applied here is comparable to previous Statistical Catch-At-Age (SCAA) models applied to Herring such as iSCAM (Martell 2017). As an SRA, the model assumes historical catches are known exactly. It is worth noting that commonly applied assessment frameworks such as Stock Synthesis (Methot and Wetzel 2013) are often cited as SCAs, however, in most applications, catches are fitted with such high precision that they function identically to their more computationally efficient counterparts – SRAs. For all operating models conditioned in these analyses the only substantive difference in fitting in SCA or SRA mode was a substantial improvement in stability and estimation time using the SRA formulation.

Given the substantial improvement in computational efficiency of TMB over previous estimation software such as ADMB (of iSCAM and SS3), the SRA model used in this conditioning is at least an order of magnitude faster than previous assessments, typically converging to a positive definite Hessian of model parameters (or not) in a few seconds on a modern laptop. This speed is important for developing operating models since it allows for in-meeting exploration of alternative operating model scenarios (all 58 operating models conditioned in these analyses can be refitted in under 5 minutes with a mobile computer).

An additional advantage of TMB is that it is native to the R statistical environment and hence linked directly to the MSEtool package and operating models proposed for conducting MSE analyses.

Standardized operating model conditioning reports are also available for all operating models described in this document that provide a summary of model estimates, fits to data and show the accuracy of conversion to the R MSE framework.

DEVELOPING A REFERENCE CASE OPERATING MODEL

A reference case model fit was established as the basis for exploring various sensitivities to data weighting, data types, and parameter values outlined in Table 1 in order to identify a reference set of operating models.

The reference case model was fitted to catch and age composition data for all four fleets (gillnet, purse-seine, weir and “other gears”) and the larval and acoustic surveys for the period 1968–2018. The model assumes virgin unfished stock conditions prior to the first model year (unfished in 1967, but see sensitivity OMs below.).

All previous assessments of this stock have excluded southwest New Brunswick weir catches from the SWNS/BoF landings under the assumption that weir catches are migrant juvenile fish of U.S. origin. A recommendation from the 2006 assessment framework (DFO 2007) was to change this assumption and include a fraction of the landings as part of the SWNS/BoF stock. This change was never implemented due to the subsequent large decline in weir catches. The Atlantic Herring Working Group discussed this issue at the January 2020 meeting and the group believes that a portion of weir catches are from the SWNS/BoF spawning component based on tagging data. The reference case was defined as zero weir catches included in the landings (status quo) with an extreme of 100% weir catches used to assess the influence of this uncertainty as a sensitivity OM.

The acoustic survey was assumed to be an index of vulnerable biomass (with selectivity estimated) whereas the larval survey was assumed to correspond with spawning stock numbers in the population. Although fitted to the catch-at-age composition data, the catch-at-length data were still submitted to the model (with zero weight) so that the implied fit could be evaluated (potentially to identify systematic pattern in length-to-age conversions).

Since it led to fewer data conflicts and greater model stability, the reference case model estimated catchability q for the acoustic survey (but see below for sensitivity OMs).

Previous assessments of this stock have used a constant M of 0.2 (DFO 2004). The Atlantic Herring Working Group agreed that M was most likely greater than 0.2 at the January 2020 meeting. A review was conducted to assess other M scenarios to consider. A constant M of 0.35 was used in the last US assessment for Atlantic Herring (NEFSC 2018).

Guénette and Stephenson (2012) conducted a MultiSpecies Virtual Population Analysis (MSVPA) to estimate predation mortality (M2) of Herring in southwest Nova Scotia/Bay of Fundy due to consumption by fish, birds, and marine mammals. Predator abundance was fixed and the number of Herring-at-age consumed was the result of biomass of prey available divided by suitable total biomass of prey plus total biomass of other prey. Major sources of uncertainty in the model were identified as residual mortality (M1), consumption rate and proportion of Herring eaten by predators, and the predators' abundance trends. Multiple scenarios were used to explore uncertainty about the percentage of Herring in diets. The base scenario was run with small residual mortality M1=0.01 (assuming most predators taken into account in the model). The juvenile (age 1–2) mortality rate from predation was estimated at 0.64 and the adult mortality rates at 0.37. As the estimate of Herring in fish diet varied from 20% to 130% for an average of 50% (Bundy et al. 2011), scenarios "plus50" and "minus50" assumed the consumption of Herring was being over or underestimated by 50% and this was used as a first estimate of uncertainty. The Herring consumption by the most important predators entered in the model were respectively increased and decreased by 50%. For juveniles, minus50 = 0.49 and plus50 = 0.72. For adults, minus50 = 0.26, plus50 = 0.45.

The U.S. MSE for Atlantic Herring in the Gulf of Maine conducted in 2017 (Deroba 2017), used a high and low age-varying M. Natural mortality varied among years and ages in previous US stock assessments with M being higher during 1996–2014 compared to previous years (NEFSC 2012, Deroba 2015). M was considered an uncertainty in the U.S. MSE and a high scenario was defined as age-specific M averages during 1996–2014 and a low scenario was defined as age-specific M averages during 2005–2014.

The high and low M scenarios defined by Deroba (2017) and Guénette and Stephenson (2012) are derived independently using different methods but are fairly consistent (Figure 7) and provide corroboration for their applicability for use for Herring. Three OM scenarios were selected for M for evaluation in sensitivity OMs: a constant 0.35 and an age-variant M with a high and low scenario using the values proposed by Guénette and Stephenson (2012).

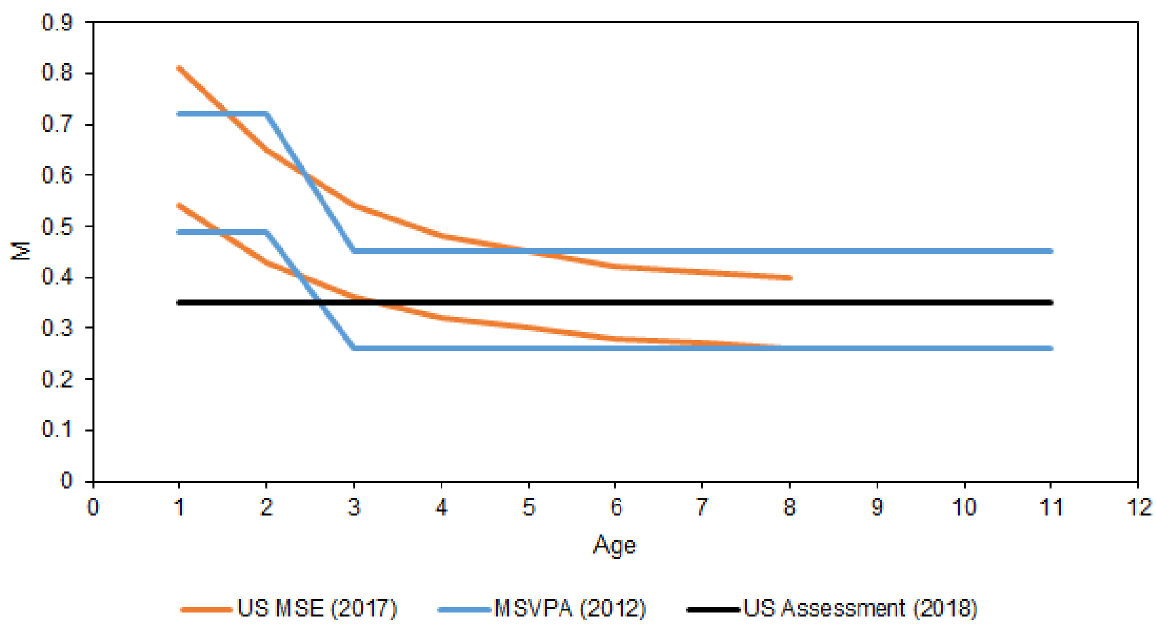


Figure 7. Natural mortality rates used in the U.S. MSE (Deroba 2017), a MSVPA (Guénette and Stephenson 2012), and U.S. Assessment (NEFSC 2018).

Growth for the reference case was estimated using the empirical weight-at-age for the time series (1970–2018) and the mean weight-at-age for the last 3 years of the time series for projections (Figure 8a) and the mean weight-at-age for the first 3 years for 1968 and 1970. An alternative growth scenario was considered in the sensitivity OM evaluation. This scenario is a continuation of the observed decline in weight-at-age for ages 4+ and a continuation of the observed increase in weight for ages 1 and 2 (Figure 8b). A linear regression of log_{10}(weight) on year was conducted separately for each age class and then slopes and intercepts were adjusted to correct for interactions among age classes that would have resulted in negative growth from one year class to the next in the projections.

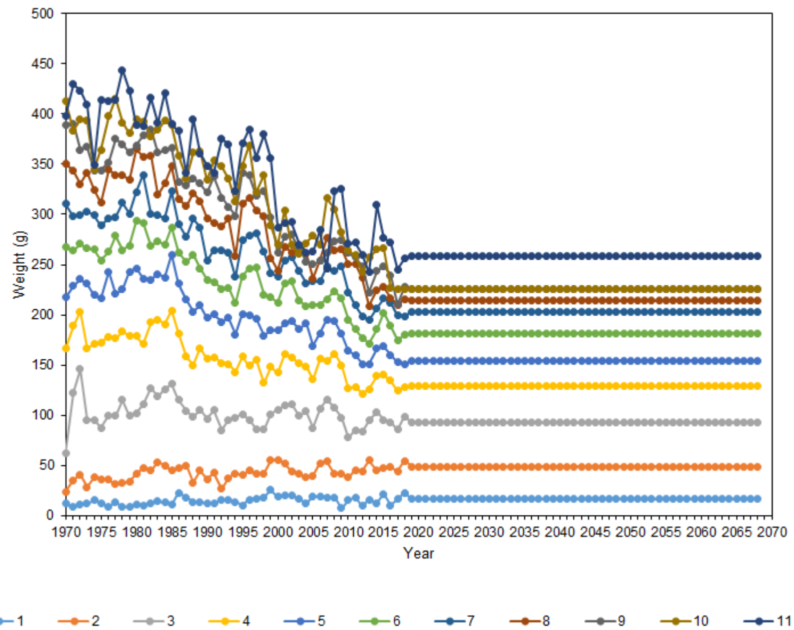


Figure 8a. Empirical weight-at-age 1970–2018 with projections to 2068 based on the mean weight-at-age for 2016–2018.

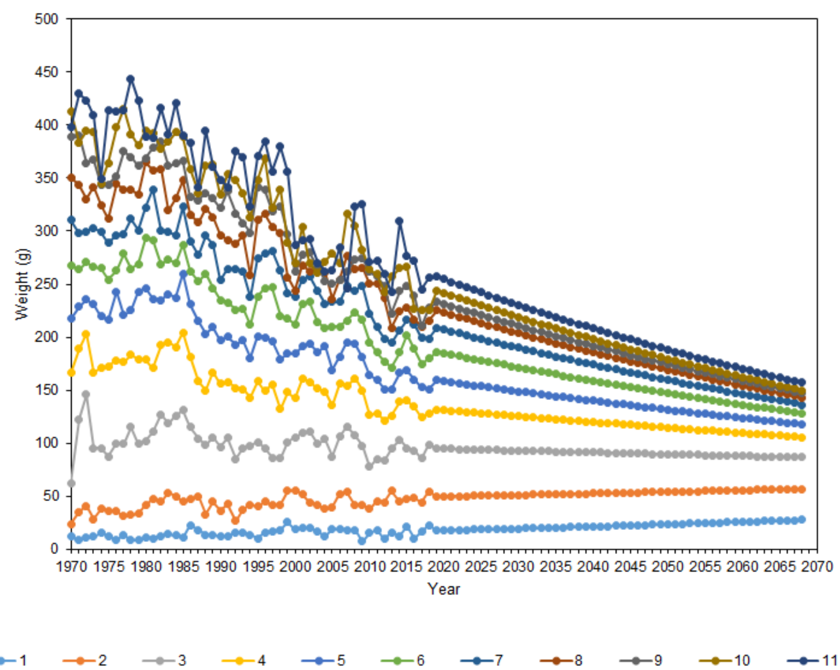


Figure 8b. Empirical weight-at-age 1970–2018 with projections to 2068 based on the regressions of $\log_{10}$(weight) on year by age using adjusted regression slopes and intercepts.

To fit the reference case operating model without ignoring recent trends in the acoustic survey index, it was necessary to down-weight the age-composition data. In the reference case OM, the effective sample size of all age composition data was reduced by a factor of 20 (the sensitivity to alternative scenarios is demonstrated in the sensitivity operating models below). This reduction in weighting did not appreciably reduce the fit to composition data but prevented negative correlation in the observed and predicted acoustic survey that would preclude its use as the primary input to an index-based management procedure.

All fleet and survey selectivities were estimated by length class. All but the purse seine fleet were assumed to follow a double-normal length selectivity function allowing for 'dome-shaped' reductions in selectivity for larger fish. The purse seine fleet was assumed to have a logistic ('flat-topped', asymptotic) selectivity in which selectivity increases with length and is 1 for large fish.

The sparsity of age-composition data for age classes 1 and 2 precludes the reliable estimation of annual recruitment in the last two years of the historical time period (2017 and 2018). For these years, recruitment is assumed to be equal to that predicted from the stock recruitment relationship without process error (recruitment deviations are zero). For future MSE projections, recruitment residuals from the final estimated year (2017 onwards) were estimated with statistical properties (variance and autocorrelation) determined by the estimates from 1968–2016.

When conditioning operating models, first the numerical optimization obtains the Maximum Likelihood Estimate (MLE) of model parameters. Then for each operating model, 48 stochastic simulations are generated by sampling from the joint normal distribution obtained from parameter covariance matrix (obtained by inverting the Hessian matrix). Detailed operating model reports are available for all operating models that show both the MLE fit and the range of outcomes for the stochastic simulations. A total of 48 simulations per operating model was used here for illustrative purposes. For following MSE analyses it is trivial to increase the number of stochastic simulations per operating model if required, however, given 24 reference set operating models, 48 simulations per operating model leads to 1,152 simulations in total. Typically MSE evaluations use only around 150 simulations to obtain stable performance ranking of candidate MPs and less than 300 simulations to obtain stable (precise) estimates of performance metrics such as long-term yield and probability of overfishing.

SPECIFICATION OF SENSITIVITY OPERATING MODELS

A total of 22 sensitivity scenarios were investigated to better understand reference case model behavior and identify consequential uncertainties for the formulation of the reference set operating models (for details of their specification see Table 4). Not all uncertainties were considered for investigation of sensitivity for including in the reference set of OMs because some uncertainties were determined *a priori* to be included as robustness OMs (e.g., some alternative catch scenarios).

Sensitivity was evaluated by examining model estimates of historical Spawning Stock Biomass in absolute terms (Figure 20a), spawning biomass relative to unfished conditions (Figure 20b) and relative to MSY levels (Figure 20c). MSY reference points were calculated based on current (2018) model parameters and aggregate fishery selectivity at age using the approach of Walters and Martell (2004).

Additionally, projection of the current fishing mortality rate at age was used to evaluate the impact of sensitivity assumptions on future Spawning Stock Biomass relative to MSY levels (Figure 21). For future years, MSY is recalculated each year to account for changing population parameters such as growth (i.e., sensitivity OM S4, 'G_ChangeGrowth'). The use of a fixed fishing mortality rate in projection is preferable to a fixed catch scenario because it provides compensatory dynamics that better mimic MPs that respond to index levels (that are the most likely type of MP to be applied in the case of Atlantic Herring).

Table 4. The sensitivity operating models developed for identifying a suitable reference set of operating models. All sensitivity operating models are a single-factor change from the reference case operating model.

OM #	Code	Description of single factor deviation from reference case
S1	RefCase	As described above.
S2	M_LowMv	Natural mortality rate varies with age: 0.49 (ages 1 and 2), 0.26 (ages 3+)
S3	M_HighMv	Natural mortality rate varies with age: 0.72 (ages 1 and 2), 0.45 (ages 3+)
S4	G_ChangeGrowth	Growth parameters are extrapolated by a linear model that assumes a continuation of the decline in weight-at-age for ages 4+ and an increase in weight-at-age for ages 1 and 2 (Figure 8b)
S5	R_Steep90	Lower steepness values (0.9–0.7) for the Beverton-Holt stock recruitment model were specified that imply a less resilient stock.
S6	R_Steep85	
S7	R_Steep80	
S8	R_Steep75	
S9	R_Steep70	
S10	C_WeirCat	Weir catches were included in the model fitting
S11	C_EqCat	Equilibrium catches are used to 'spool up' the model and are assumed to have occurred in all model years prior to the initial model year (1968). Equilibrium catches were calculated at the mean catches over the 30 years prior to the initial model year (1938–1967) for the purse seine (4,482t) and weir (33,999t) fleets.
S12	C_WeirEqCat	Both weir catches and equilibrium catches are included in the model
S13	q_1	Four sensitivity scenarios were considered that fixed q to 1 (acoustic survey is an index of absolute vulnerable biomass) and attempted to fit the acoustic survey index by prescribing progressively higher precisions in the index: 1x, 4x 16x and 36x increase.
S14	q_1x4	
S15	q_1x16	
S16	q_1x36	
S17	Comp_AgeLength	Although it could be considered as using the length composition twice (since it is used in the empirical age-length key), a sensitivity operating model was considered that was fitted to both length- and age-composition data (similarly to the age-composition data, length composition also was subject to a factor of 20 reduction in effective sample size).
S18	PS_Two	Two purse seine fleets were defined based on the spatial distribution of catches of juvenile fish (Figure A16 and A17). Fleet "PS_juv" = purse seine catches from fishing grounds where at last 50% of fish ≤ 23cm from length frequencies during the period 1995–2018. Fleet "PS_spa" = all other areas.
S19	CompWt_01	Four alternative weightings (reference case is 0.05) for the catch-at-age composition data were investigated including 0.01, 0.1, 0.2, 0.5.
S20	CompWt_1	
S21	CompWt_2	
S22	CompWt_5	

ESTABLISHING A REFERENCE GRID OF OPERATING MODELS

Four Central Sources of Uncertainty

The sensitivity operating models underlined the importance of four principal axes of uncertainty:

- the assumed natural mortality rate;
- future growth;
- stock resilience; and
- the correct assignment of historical catches to the stock.

While none of these four axes are particularly consequential in determining absolute spawning stock size (Figure 20a) or spawning stock size relative to unfished levels (Figure 20b), their impact on fishery reference points was substantial (Figure 20c) or strongly affected the expected outcomes of future projections (Figure 21; Table 10).

The reference case operating model steepness value of 0.95 offered comparable model fit to a much lower value of 0.75 (Figure 22). Although model conditioning revealed very weak information to discriminate among steepness values (compared with natural mortality for example, Figure 22), operating models would not reliably converge given a steepness value of 0.7 and therefore 0.75 was chosen as a lower value in contrast to the value of 0.95 of the reference case operating model.

Rather than consider multiple scenarios for historical catches, sensitivity analyses revealed that the combined scenario of equilibrium catches and inclusion of weir catches (S12) generally provided greater contrast in model estimates and projection outcomes in comparison to the reference case operating model.

Catchability of the Acoustic Survey to be Estimated in Reference Set Operating Models with q=1 Scenarios Included in the Robustness Set

The sensitivity scenarios in which the acoustic survey index catchability was fixed at 1 (q_1, q_1x4, q_1x16, q_1x36) provided estimates of absolute Spawning Stock Biomass between 100% and 180% higher than that of the reference case operating model (Figure 20a).

The fixed q=1 scenarios could only provide a positive correlation with the acoustic survey index when at least a 4x increase in the precision of the index was specified. For this reason, the assumption that the acoustic survey index is an absolute index of abundance (q=1) was not considered to be plausible, and was left out of the reference set operating models and instead moved to the robustness set of operating models.

Model Estimates and Projection Outcomes are Insensitive to Conditioning on Length Composition Data

Conditioning the model on both age and length composition data (Comp_AgeLength) did not provide a meaningful impact on either the estimates of absolute stock biomass (Figure 20a), relative stock size (Figure 20b) or stock level relative to MSY reference points (Figure 20c). This sensitivity also did not provide a substantial impact on the projected outcome of maintaining 2018 fishing levels (Figure 21) and was therefore not included in the reference set operating models.

Lack of Support for More Complex Purse Seine Structure

Splitting the purse seine fleet into two spatial fleets that represent catches of primarily juvenile fish and primarily adult fish (PS_Two) did not improve the conflict in inferred scale between the acoustic index and age composition data. Furthermore, the fit to the composition data did not substantially improve. This axis of uncertainty was not investigated further.

Inconsequential Alternative Weightings for Age Composition Data

Increasing in the weighting of age composition data (a higher weighting than the 1/20 weighting of the reference case operating model) led to flat or increasing trends in Spawning Stock Biomass (Figure 20a) and negative correlations with the acoustic survey index (CompWt_1, CompWt_2, CompWt_5). Reducing the weighting (CompWt_01) led to stronger decline in vulnerable stock biomass than inferred by the acoustic survey. In any case, the weighting of composition data did not have a substantial impact on projected outcomes of current exploitation rates (Figure 21) and was therefore left out of the reference set and moved to the robustness set of operating models.

Defining a Reference Grid of Operating Models

The proposed reference set of operating models includes the three levels of natural mortality rate (reference case, S2, S3) from the sensitivity analyses, two levels of future growth (reference case, S4), two levels of stock resilience (reference case, S8) and two levels for the historical catch levels (reference case, S12). Table 5 describes these levels in greater detail.

Often MSE results are summarized including multiple operating models. Similar to MSE frameworks established elsewhere, the reference set of operating models follows a full factorial cross of the four principal factors to avoid higher *a priori* weight for any factor level. The full cross leads to 24 reference set operating models (Table 6) (currently presented here with 48 simulations per operating model, but with the flexibility to change this if necessary for future MSE analyses).

Table 5. Reference set operating model factors and levels. (-) = not applicable.

Factor	Levels		
Factor 1: Natural Mortality	**Level 1**	**Level 2**	**Level 3**
	M = 0.35 (all ages)	M = 0.49 (ages 1 and 2)	M = 0.79 (ages 1 and 2)
	-	M = 0.26 (ages 3+)	M = 0.45 (ages 3+)
Factor 2: Growth	**Level A**	**Level B**	**-**
	Future growth is the average of the last three historical years (2016–2018)	Future growth is determined by a linear extrapolation of the temporal trend in weight-at-age	-
Factor 3: Resilience	**Level H**	**Level L**	**-**
	Steepness of the stock recruitment function is 0.95	Steepness of the stock recruitment curve is 0.75	-
Factor 4: Catch	**Level -**	**Level +**	**-**
	Weir catches are not included in model conditioning and the stock is assumed to be in virgin unfished conditioned in 1967	Weir catches are included in model conditioning and equilibrium catches for the 30 year period prior to 1968 were included	-

Table 6. The grid of reference set operating models: a factorial cross of the factors and levels identified in Table 5.

OM #	Natural Mortality	Growth	Resilience	Catch
1	1	A	H	-
2	2	A	H	-
3	3	A	H	-
4	1	B	H	-
5	2	B	H	-
6	3	B	H	-
7	1	A	L	-
8	2	A	L	-
9	3	A	L	-
10	1	B	L	-
11	2	B	L	-
12	3	B	L	-
13	1	A	H	+
14	2	A	H	+
15	3	A	H	+
16	1	B	H	+
17	2	B	H	+
18	3	B	H	+
19	1	A	L	+
20	2	A	L	+
21	3	A	L	+
22	1	B	L	+
23	2	B	L	+
24	3	B	L	+

ROBUSTNESS OPERATING MODELS

Identifying a Robustness Grid

Robustness operating models are intended to capture less plausible scenarios for fishery dynamics or alternatively, scenarios for which there is limited empirical evidence to evaluate plausibility. Robustness scenarios are typically single factor changes from other operating models. If robustness scenarios are only applied to the reference case operating model then outcomes are heavily constrained to a single relatively narrow set of simulations. Equally, if a robustness scenario is applied to the entire reference set grid the amount of MSE output becomes unmanageable given the very large number of operating models. The solution is to identify a small subset of reference set operating models (a ‘robustness grid’) that encompass a range of uncertainty and apply each single-factor robustness scenario to those operating models.

Examining the projected outcomes of current fishing mortality rate at age for the reference set reveals that some factors are generally attributable to more optimistic outcomes and others to more pessimistic outcomes, with the reference case providing an intermediate outcome (Figure 24). These projections were used to identify a pessimistic and optimistic scenario based on the various combinations of factor levels (Figure 26; Table 7).

The operating models identified in the robustness grid spanned all factor levels and a wide range of projected outcomes.

Table 7. Robustness Grid. Three reference set operating models with contrasting outcomes for specifying single factor robustness scenarios.

Code	Reference Set Operating Model	Description
R	**1AH-**	Reference case
O	**3BH+**	Optimistic High age-varying natural mortality rate Future growth is specified by the continuing temporal trend in weight-at-age High resilience (steepness = 0.95) Equilibrium catches and weir catches are included in conditioning
P	**2AL-**	Pessimistic Low age-varying natural mortality rate Future growth is the average of the last three years (2016–2018) Low resilience (steepness = 0.75) No equilibrium catches or weir catches are included in conditioning.

SPECIFYING THE ROBUSTNESS SET OF OPERATING MODELS

The robustness set of operating models encompassed 7 additional sources of uncertainty in fishery dynamics in addition to the four factors of the reference set:

- acoustic survey catchability;
- alternative weighting for the age-composition data;
- assuming a lower level of stock resilience;
- conditioning on length composition data;
- an alternative functional form for the stock-recruitment relationship;
- a rebuilding scenario in which Spawning Stock Biomass in 2018 is more depleted than the reference set of operating models (Table 8 provides a description of these additional factors); and
- alternative catch scenarios that include a proportion of U.S. landings and landings from the entire NAFO 4VWX area.

In some cases, combinations of robustness assumptions and the pessimistic configuration (P, 2AL-) would not converge to a positive definite Hessian and were therefore dropped from the robustness set.

Table 8. The robustness set of operating models. Codes including _R, _O, and _P refer to the same single factor robustness scenario applied to the three members of the robustness grid (Table 7): reference case, optimistic and pessimistic, respectively.

Number	Codes	Description
R1	**q1_R**	Acoustic Survey catchability is fixed to 1 and precision of the Acoustic survey is artificially increased by a factor of 16 in order to still obtain satisfactory fit to the acoustic survey index (e.g., sensitivity OM S15 'q_1x16').
R2	**q1_O**	
R3	**q1_P**	
R4	**AgeComp1_R**	Age composition data are down-weighted by 1/10 (e.g., sensitivity OM S20 'CompWt_1', rather than the 1/20 down-weighting of the reference case operating model.
R5	**AgeComp1_O**	
R6	**AgeComp1_P**	
R7	**Steep7_R**	A low resilience scenario. Steepness is assumed to be 0.7. This model did not converge for the pessimistic (P) model of the robustness grid.
R8	**Steep7_O**	
R9	**Length_R**	Models are fitted to both age and length composition data (e.g., Sensitivity OM #17, 'Comp_AgeLength').
R10	**Length_O**	
R11	**Length_P**	
R12	**Ricker_R**	The Ricker stock recruitment relationship is assumed. This model did not converge for the pessimistic (P) model of the robustness grid.
R13	**Ricker_O**	
R14	**Reb_R**	Previous conditioning exercises considering lower age-invariant natural mortality rate and estimated stock levels at around or slightly lower than half of BMSY. As an additional robustness test, a rebuilding scenario was assumed that starts each forward projection from half of BMSY levels given the same historical pattern in fishing mortality rates.
R15	**Reb_O**	
R16	**Reb_P**	
R17	**US20_R**	Historical catches include 20% of catches in adjacent U.S. waters (Kanwit and Libby 2009).
R18	**US20_O**	
R19	**US20_P**	
R20	**US40_R**	Historical catches include 40% of catches in adjacent U.S. waters (Stobo and Fowler 2009).
R21	**US40_O**	
R22	**US40_P**	
R23	**ALL_R**	All historical catches in area 4VWX are included in model conditioning.
R24	**All_O**	
R25	**All_P**	

RESULTS

REFERENCE CASE MODEL

Reference Case Model Fit to Indices

In general, the reference case model fit to the acoustic survey was good to excellent showing comparable trend without evidence for a pronounced pattern in residuals (Figure 9).

The larval survey has a much higher variance than the acoustic survey and for most years the estimated vulnerable stock numbers fell within the standard errors (Figure 10). The clear exceptions are 1974 and 1981 where the model underestimates then strongly overestimates the vulnerable stock numbers, respectively.

The sampling of parameter values from the parameter variance-covariance matrix, generally resulted in vulnerable stock size estimates with much lower variance than prescribed by the indices (Figures 8 and 9).

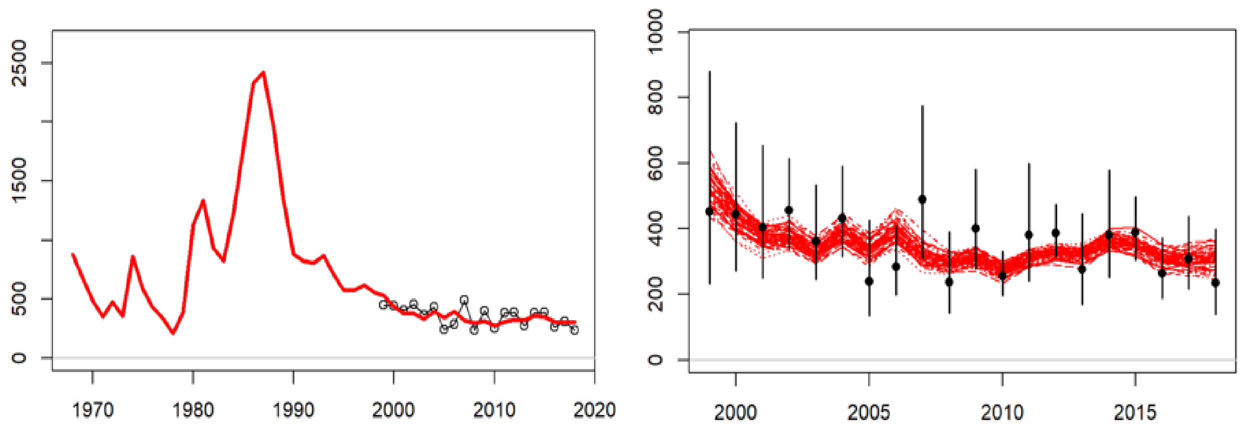


Figure 9. Base model fit to the acoustic survey index. Black points and line are observations, the red line is the model estimate (the left panel shows the maximum likelihood fit, the right panel shows the 48 stochastic simulations). Note that the observed acoustic survey is provided here in units of kilotons, the model fit is converted to the same scale as the observed index via the estimated catchability coefficient (q = 2.85)

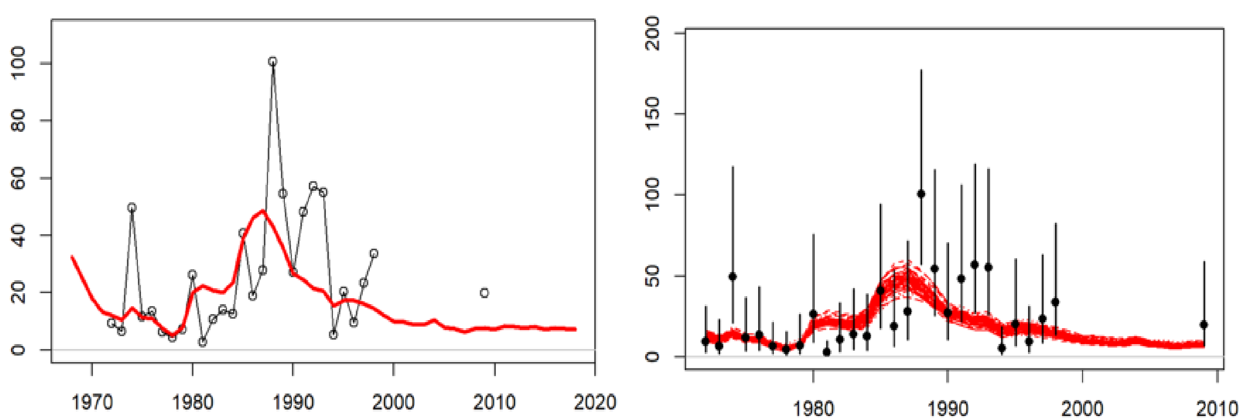


Figure 10. As Figure 9 but for the larval survey index.

Reference Case Model Fit to Acoustic Survey Age Composition Data

The overall fit to the acoustic survey age composition was very good (Figure 11) showing few problematic patterns in residuals or an inability to capture variable cohort strength (Figure 11).

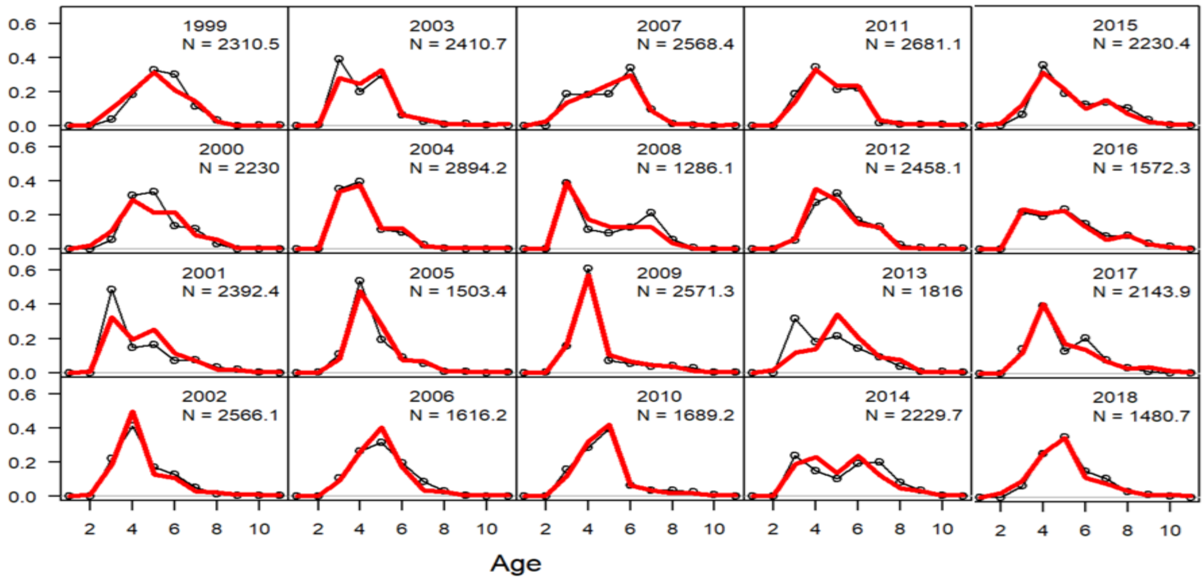


Figure 11. Base model fit to Acoustic survey age composition. Black lines and points are observed values, red lines are model predictions. The sample sizes (N) reported are the values used in the model after downweighting.

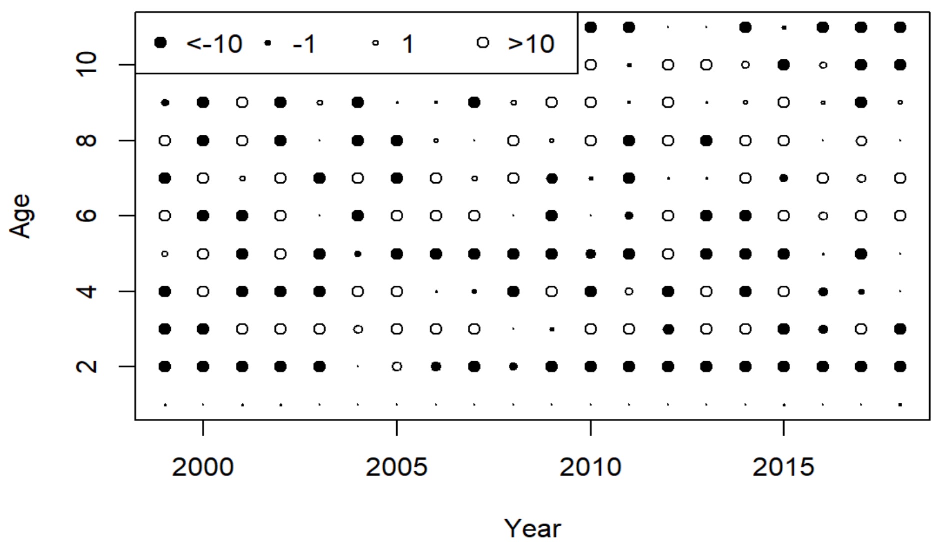


Figure 12. Standardized residuals for fits to acoustic survey age composition data.

Model Fit to Fleet Age Compositions

The reference case model fit to the age composition of the purse seine fleet (the most significant by catches) was good for recent years (2002–2018) but poorer for some of the earlier years (e.g., 1999 and 2000).

The fit to the age composition from fleets with smaller catches was substantially poorer, for example, the gillnet and ‘other’ fleets in 2017 and 2018. Since the mature age composition of the acoustic survey and principal fleet (the purse seine fleet is more than 95% of catches in 2018 excluding the weir fishery and around 75% of catches in 2018 including the weir fishery; Figure A1) are good these misfits indicate model misspecification, in which availability to the secondary fleets is more temporally (or spatially) variable.

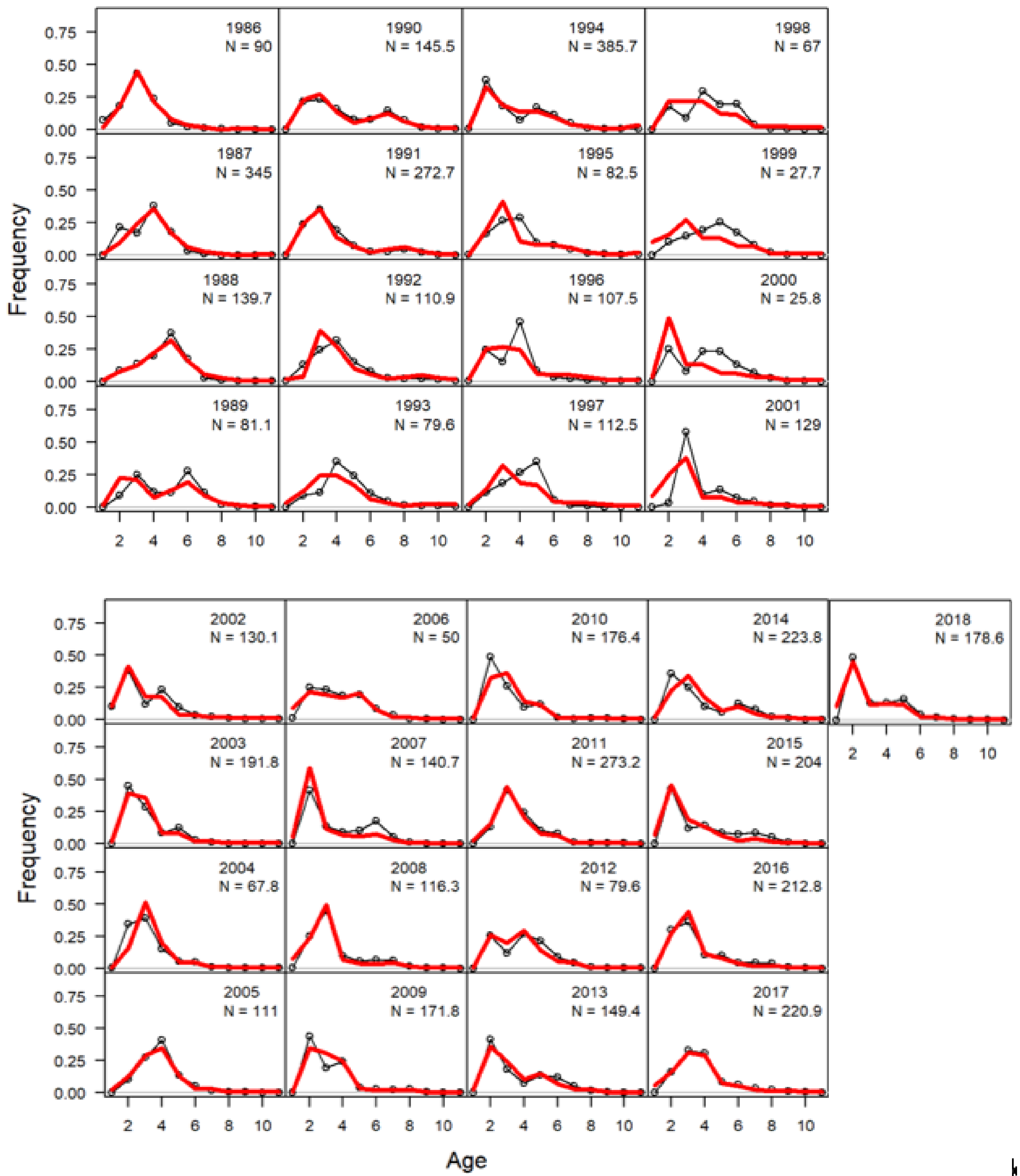


Figure 13. Fit to the purse seine fleet age composition data. Black lines and points are observed values, red lines are model estimates (fit of the maximum likelihood parameter estimates).

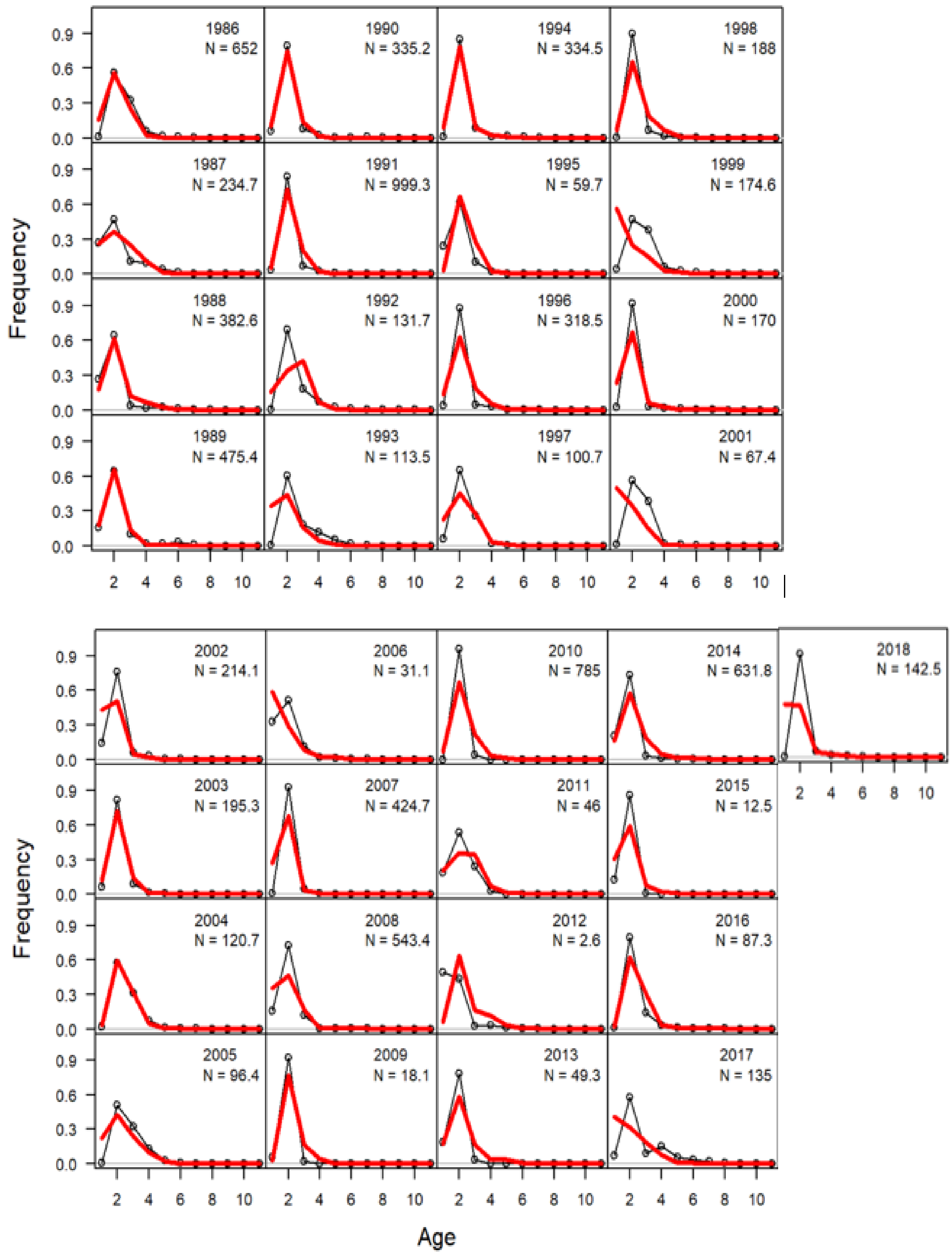


Figure 14. As Figure 13 but for the weir fleet.

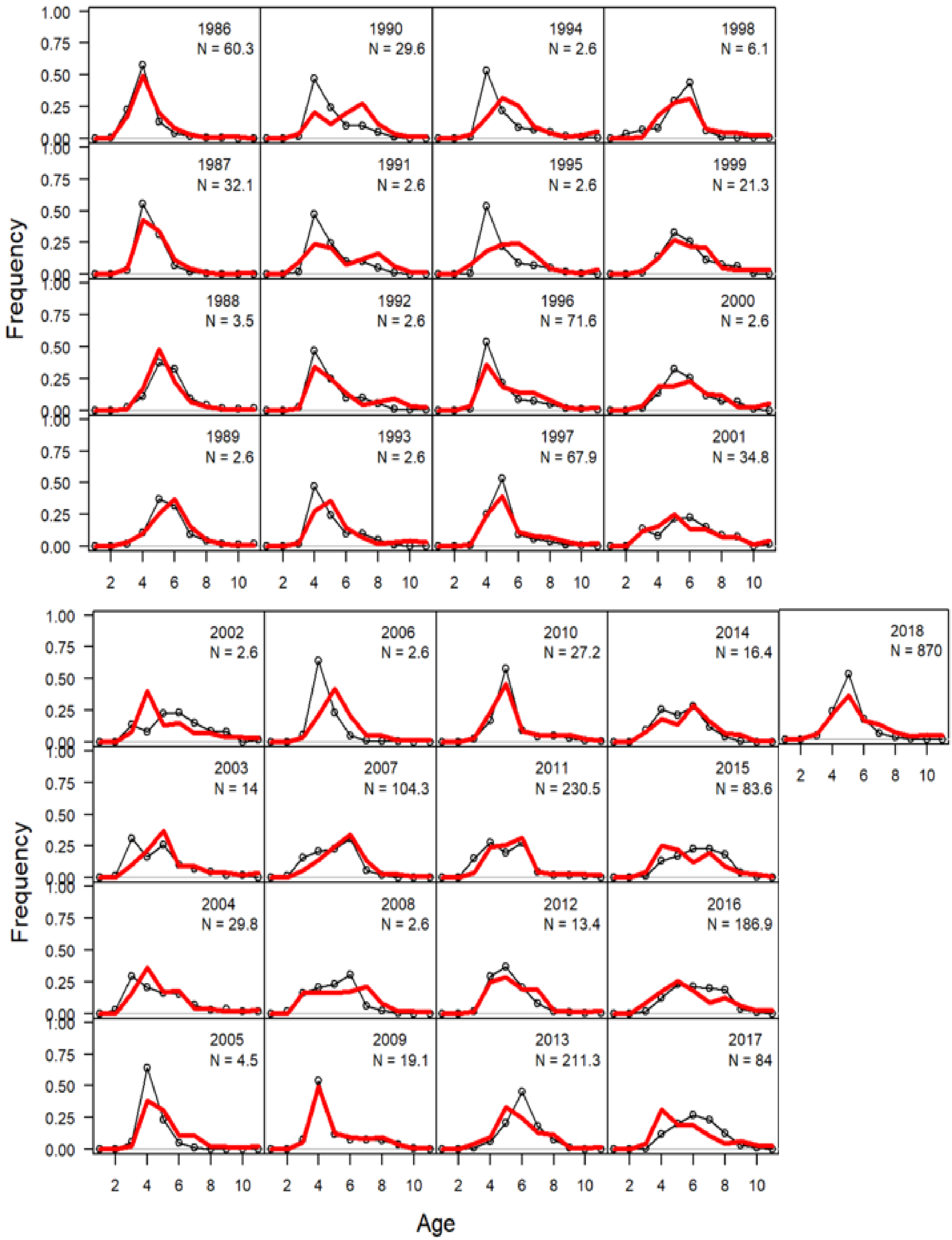


Figure 15. As Figure 13 but for the gillnet fleet.

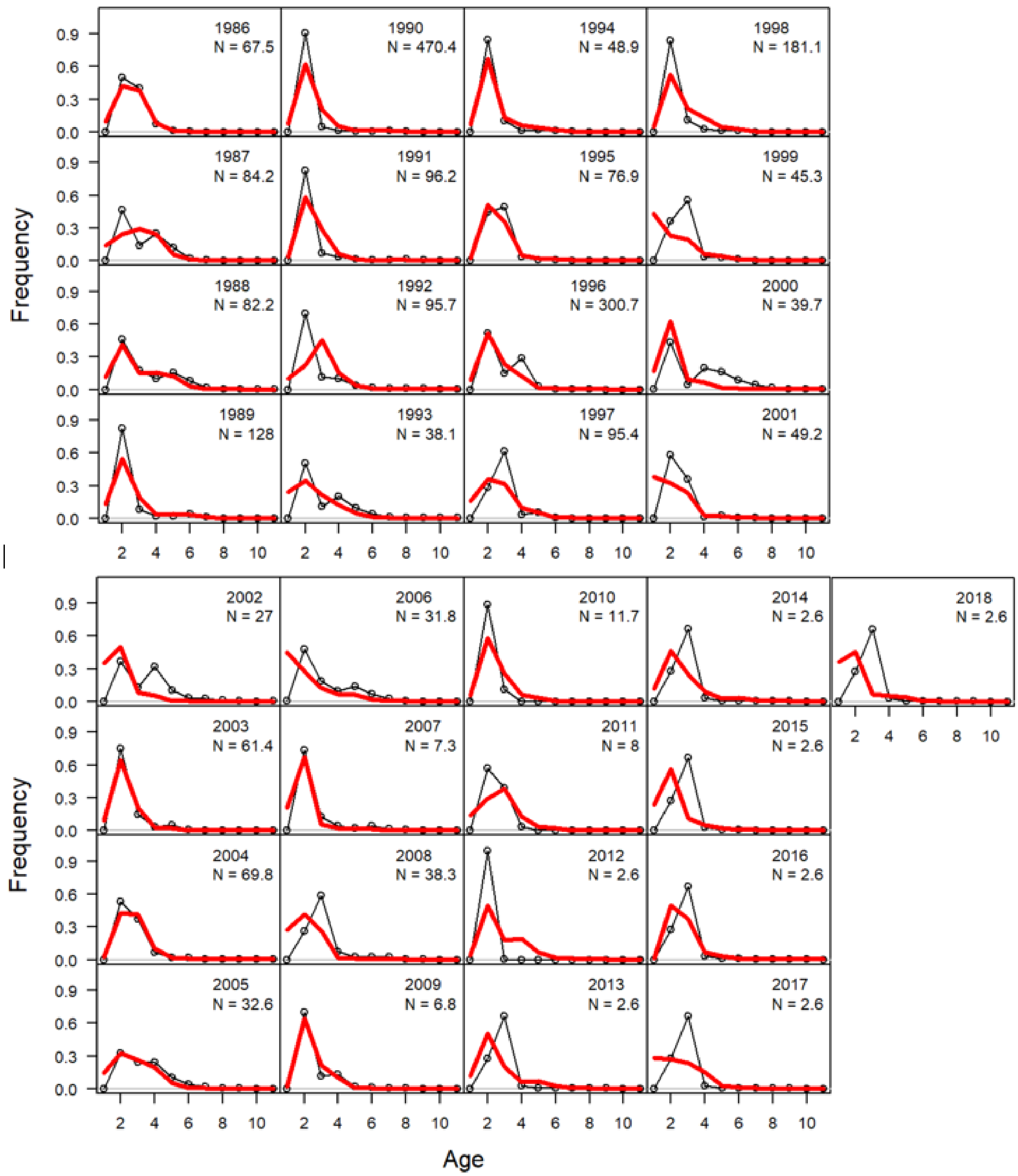


Figure 16. As Figure 13 but for the ‘other’ gears fleet.

Reference Case Model Estimates

The reference case model (and generally all models presented in this document) show a substantial decline in Spawning Stock Biomass from 1968 to 1978, after which there is a marked 2-phase recovery to well above initial levels in 1987, followed by a decelerating decline that levels off around 2003 to 2018 (Figure 17). The rapid increases in estimated spawning biomass around 1980 is caused by the estimation of particularly strong recruitment over proceeding years (Figure 18). Subsequent declines in the 1980s are driven by a resumption to mean recruitment levels and an initial spike in fishing mortality rates (Figure 19).

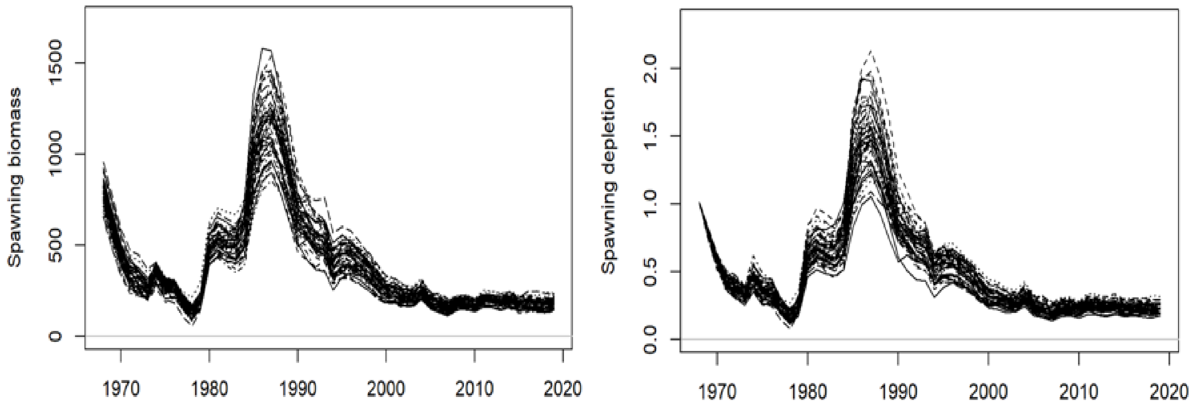


Figure 17. Model estimates of absolute Spawning Stock Biomass (kt) and Spawning Stock Biomass relative to unfished levels (Spawning depletion).

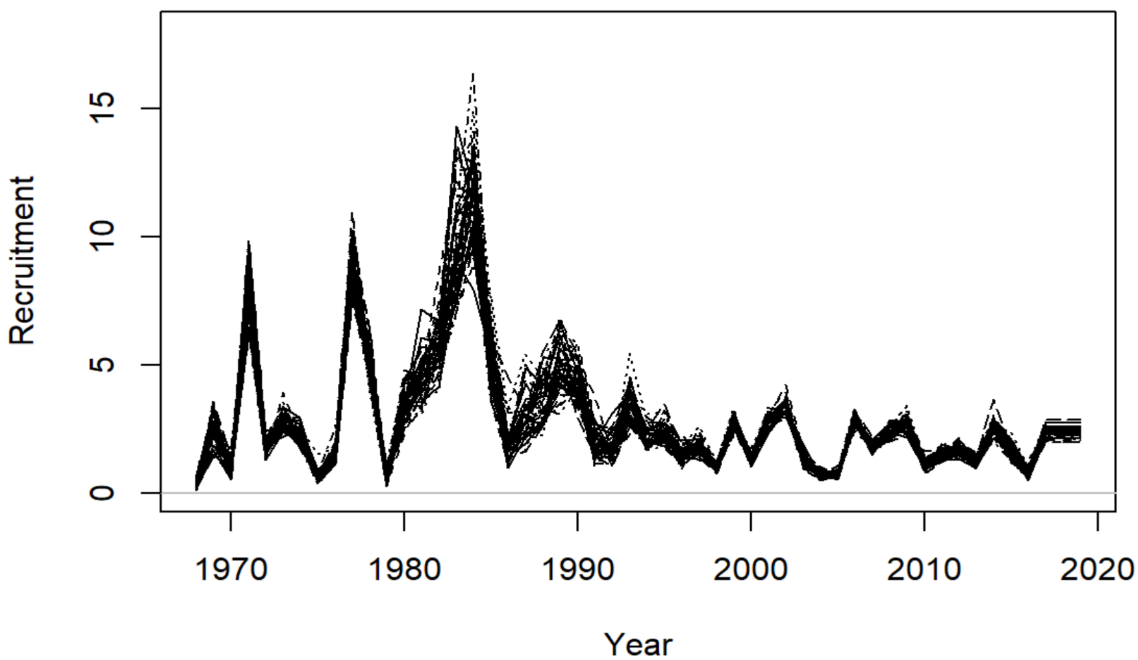


Figure 18. Estimated recruitment strength (unitless - rescaled numbers of fish entering age class 1).

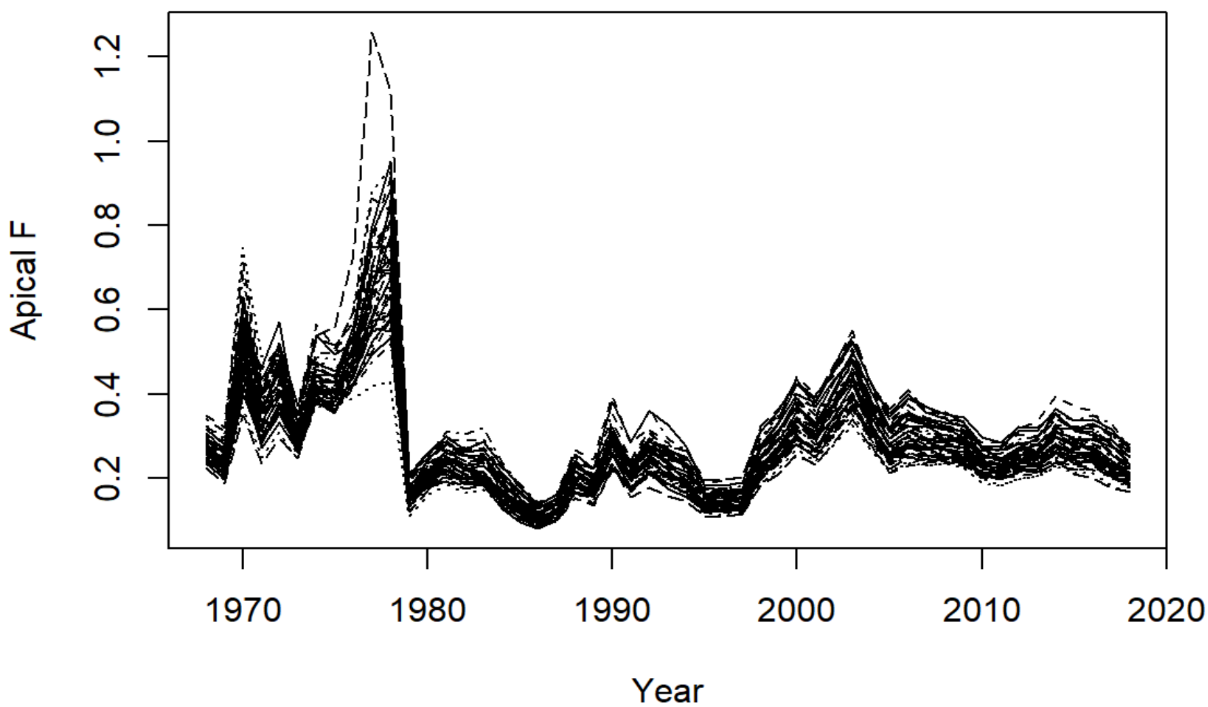


Figure 19. Model estimates of apical (maximum F among length classes) instantaneous fishing mortality rate (yr^{-1}).

SENSITIVITY OPERATING MODELS

The principal use of sensitivity operating model results was in the specification of the reference set operating models. It follows that the key results are presented in the methodological section "Establishing a Reference Grid of Operating Models" above. Here we present secondary results not related to the selection of the reference set.

Among the sensitivity analyses that did not reweight data or include additional data (S1–S8), overall fits summarized by the total negative log likelihood (nll; Table 9), were generally comparable (all close to the reference case OM negative log likelihood of 668.33). The exceptions were the low and high, age-varying natural mortality rate scenarios (S2 and S3) which showed worse (nll = 675) and much worse (nll = 681) overall fit, respectively.

Table 9. Negative log-likelihoods for maximum likelihood fitting of sensitivity operating models (lower values represent better fit). A dash (-) indicates no data.

OM	Code	Total	Equilibrium Catch		Age Composition (length composition, where applicable)					Indices		Index catchability	
			Purse seine	Weir	Gill	Other	Purse seine	Weir	Purse seine (juv)	Acoustic	Larval	q Acoustic	q Larval
S1	**RefCase**	668.33	-	-	58.38	100.33	142.56	154.43	-	-3.30	39.68	3.03	0.04
S2	**M_LowMv**	675.39	-	-	59.22	100.34	146.77	155.05	-	-3.23	38.00	3.48	0.04
S3	**M_HighMv**	680.78	-	-	58.82	100.82	146.11	158.33	-	-3.45	41.61	2.32	0.03
S4	**G_ChangeGrowth**	668.33	-	-	58.38	100.33	142.56	154.43	-	-3.30	39.68	3.03	0.04
S5	**R_Steep90**	668.23	-	-	58.41	100.27	142.60	154.42	-	-3.31	39.67	2.99	0.03
S6	**R_Steep85**	668.16	-	-	58.49	100.20	142.64	154.41	-	-3.34	39.66	2.93	0.03
S7	**R_Steep80**	668.15	-	-	58.61	100.12	142.67	154.38	-	-3.39	39.64	2.86	0.03
S8	**R_Steep75**	668.25	-	-	58.77	100.04	142.70	154.35	-	-3.45	39.60	2.78	0.03
S9	**R_Steep70**	668.50	-	-	58.96	99.95	142.75	154.30	-	-3.51	39.55	2.69	0.03
S10	**C_WeirCat**	663.69	-	-	56.36	99.79	144.05	152.10	-	-2.39	38.38	2.92	0.03
S11	**C_EqCat**	659.99	-3.67	-3.46	64.59	100.20	146.83	147.50	-	-2.13	38.13	3.16	0.04
S12	**C_WeirEqCat**	654.41	-3.69	-3.25	61.81	99.59	148.43	145.42	-	-1.20	36.72	3.00	0.04
S13	**q_1**	683.99	-	-	63.28	98.72	145.16	154.37	-	1.24	38.80	1	0.02
S14	**q_1x4**	701.20	-	-	64.60	98.69	145.58	154.96	-	10.99	38.88	1	0.02
S15	**q_1x16**	778.26	-	-	65.89	99.29	144.63	157.28	-	72.30	38.97	1	0.02
S16	**q_1x36**	888.74	-	-	67.27	99.81	145.79	161.21	-	151.45	39.13	1	0.03
S17	**Comp_AgeLength**	1,990.22	-	-	64.54 (107.60)	100.77 (199.66)	145.73 (404.64)	160.00 (469.07)	-	-2.50	39.33	3.55	0.05
S18	**PS_Two**	764.97	-	-	52.84	109.44	153.30	151.47	77.93	-0.76	36.12	2.88	0.05
S19	**CompWt_01**	260.92	-	-	16.13	23.74	37.03	38.80	-	-4.35	34.60	2.88	0.03
S20	**CompWt_1**	1,164.12	-	-	112.57	195.08	268.51	297.96	-	-2.14	42.95	2.98	0.04
S21	**CompWt_2**	2,148.86	-	-	223.39	384.72	518.75	584.98	-	-0.58	45.76	2.90	0.04
S22	**CompWt_5**	5,093.82	-	-	560.46	954.18	1,268.35	1,446.55	-	2.40	48.19	2.76	0.04

Table 10. Mean estimates derived from the sensitivity operating models. Maximum Sustainable Yield (MSY) quantities were calculated by the method of Walters and Martell (2004). FMSY is apical fishing mortality rate at MSY (maximum over length classes). SSBMSY = Spawning Stock Biomass at MSY. BMSY is total vulnerable biomass at MSY. UMSY is the fraction of vulnerable biomass caught at MSY (harvest rate). SSB0 is unfished Spawning Stock Biomass. RefY is the reference yield, the maximum yield obtainable by a fixed fishing rate given future conditions and current fishery selectivity. Blow is the biomass for which it would take 2 mean generation times to reach half of BMSY given current fishing and biological parameters. MGT is mean generation time calculated at the average age of a mature fish in the unfished population. SSB/SSBMSY is current spawning biomass relative to MSY levels. D is current stock depletion calculated as current Spawning Stock Biomass (SSB) divided by unfished Spawning Stock Biomass (SSB0) .

OM	Code	MSY (kt)	FMSY	SSBMSY (kt)	SSBMSY/ SSB0	BMSY (kt)	UMSY	SSB0 (kt)	RefY (kt)	Blow (kt)	MGT (yrs)	SSB/ SSBMSY	D
S1	**RefCase**	75.587	0.623	109.912	0.142	253.808	0.414	776.541	71.704	0.040	5.089	1.613	0.229
S2	**M_LowMv**	54.496	0.499	103.340	0.136	215.489	0.350	760.067	52.519	0.013	5.521	1.301	0.177
S3	**M_HighMv**	122.781	1.076	90.120	0.110	412.151	0.582	822.642	113.433	13.982	4.691	3.045	0.335
S4	**G_ChangeGrowth**	76.134	0.621	111.218	0.143	256.498	0.413	776.541	67.437	0.040	5.089	1.601	0.229
S5	**R_Steep90**	71.931	0.524	127.735	0.163	269.809	0.360	783.404	66.070	0.077	5.089	1.368	0.223
S6	**R_Steep85**	68.575	0.455	142.646	0.180	282.974	0.320	790.681	61.247	0.281	5.089	1.206	0.217
S7	**R_Steep80**	65.331	0.401	156.169	0.196	294.794	0.288	798.307	56.800	0.745	5.089	1.082	0.212
S8	**R_Steep75**	62.089	0.355	168.986	0.210	305.906	0.259	806.246	52.515	1.633	5.089	0.986	0.207
S9	**R_Steep70**	58.764	0.316	181.472	0.223	316.632	0.234	814.386	48.251	3.196	5.089	0.906	0.202
S10	**C_WeirCat**	89.531	0.531	138.897	0.143	311.423	0.360	972.187	82.472	0.012	5.089	1.406	0.201
S11	**C_EqCat**	94.633	0.627	137.526	0.142	317.657	0.416	971.446	87.846	0.014	5.089	1.359	0.193
S12	**C_WeirEqCat**	104.516	0.534	161.572	0.143	362.794	0.361	1,130.042	94.695	0.013	5.089	1.238	0.177
S13	**q_1**	94.072	0.566	139.029	0.141	319.108	0.380	986.390	89.628	27.687	5.089	3.525	0.497
S14	**q_1x4**	92.969	0.565	137.445	0.141	315.429	0.379	975.224	88.511	17.981	5.089	3.014	0.425
S15	**q_1x16**	88.222	0.574	129.871	0.141	298.564	0.384	921.246	84.047	10.255	5.089	2.596	0.366
S16	**q_1x36**	85.415	0.586	125.249	0.141	288.33	0.391	887.290	81.430	7.407	5.089	2.404	0.339
S17	**Comp_AgeLength**	62.589	0.841	93.524	0.151	211.12	0.546	618.802	60.032	1.026	5.089	1.742	0.263
S18	**PS_Two**	86.560	0.746	120.070	0.135	281.334	0.480	891.608	86.033	0.154	5.089	1.844	0.249
S19	**CompWt_01**	91.240	0.633	133.148	0.142	306.596	0.420	935.005	87.023	0.215	5.089	1.345	0.191
S20	**CompWt_1**	70.384	0.616	102.448	0.141	236.545	0.410	724.765	66.725	0.034	5.089	1.865	0.263
S21	**CompWt_2**	67.275	0.611	98.019	0.141	226.275	0.407	694.128	63.755	0.034	5.089	2.163	0.305
S22	**CompWt_5**	65.422	0.606	95.404	0.141	220.193	0.404	676.201	61.990	0.043	5.089	2.624	0.370

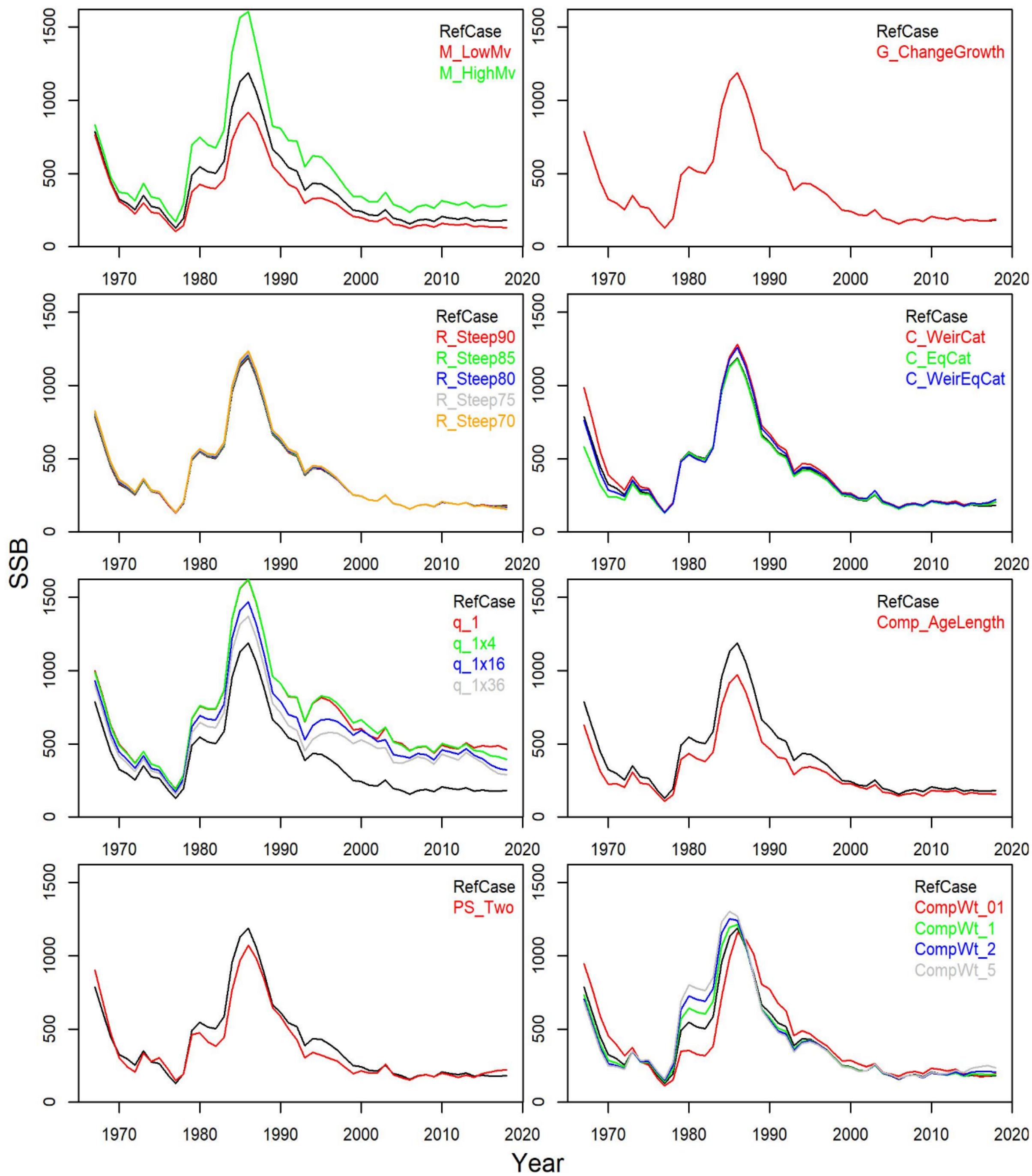


Figure 20a. Mean estimated Spawning Stock Biomass (SSB) (kt) for the various sensitivity operating models of Table 4.

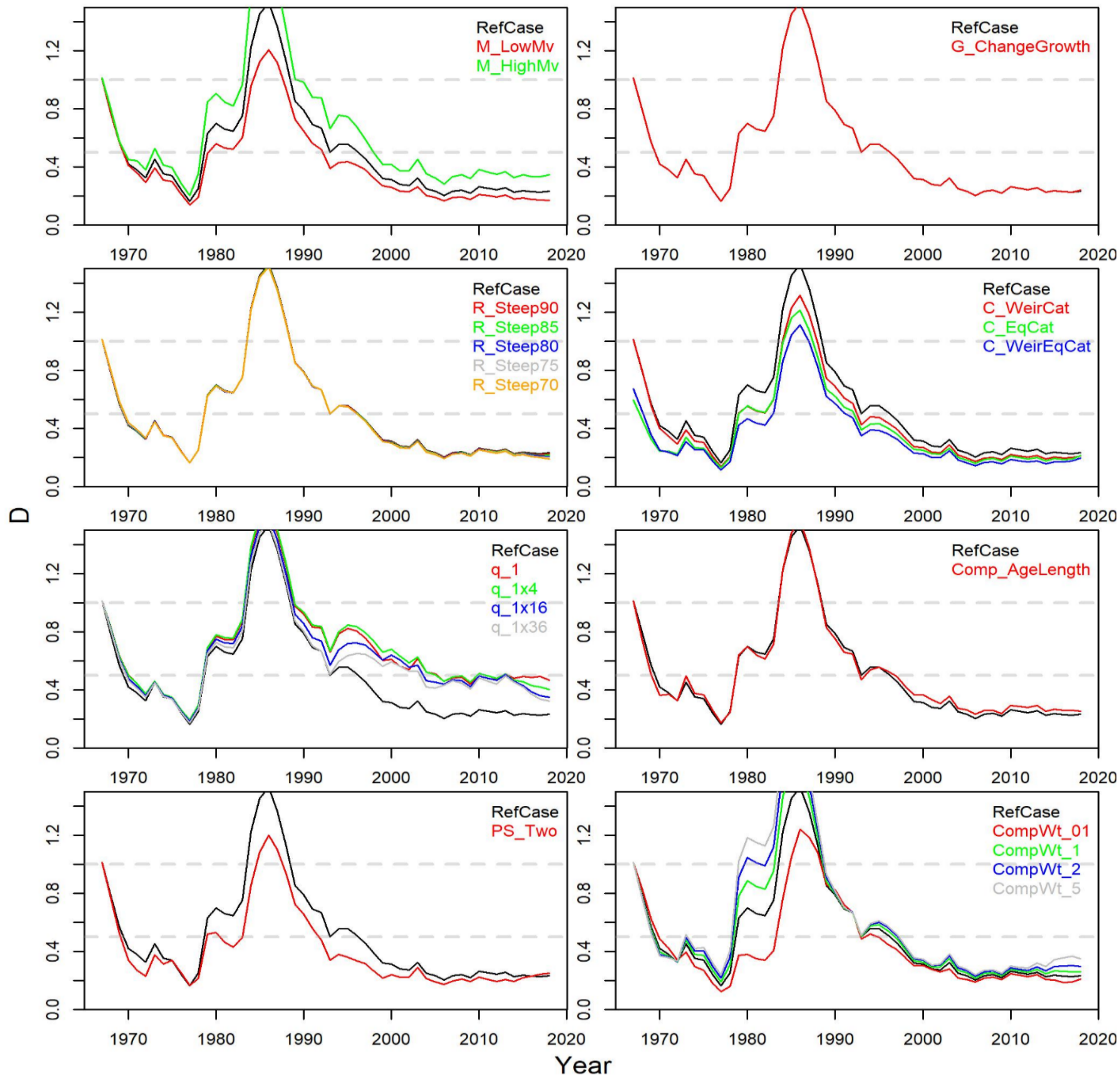


Figure 20b. Mean estimates of depletion (D) for the various sensitivity operating models described in Table 4.

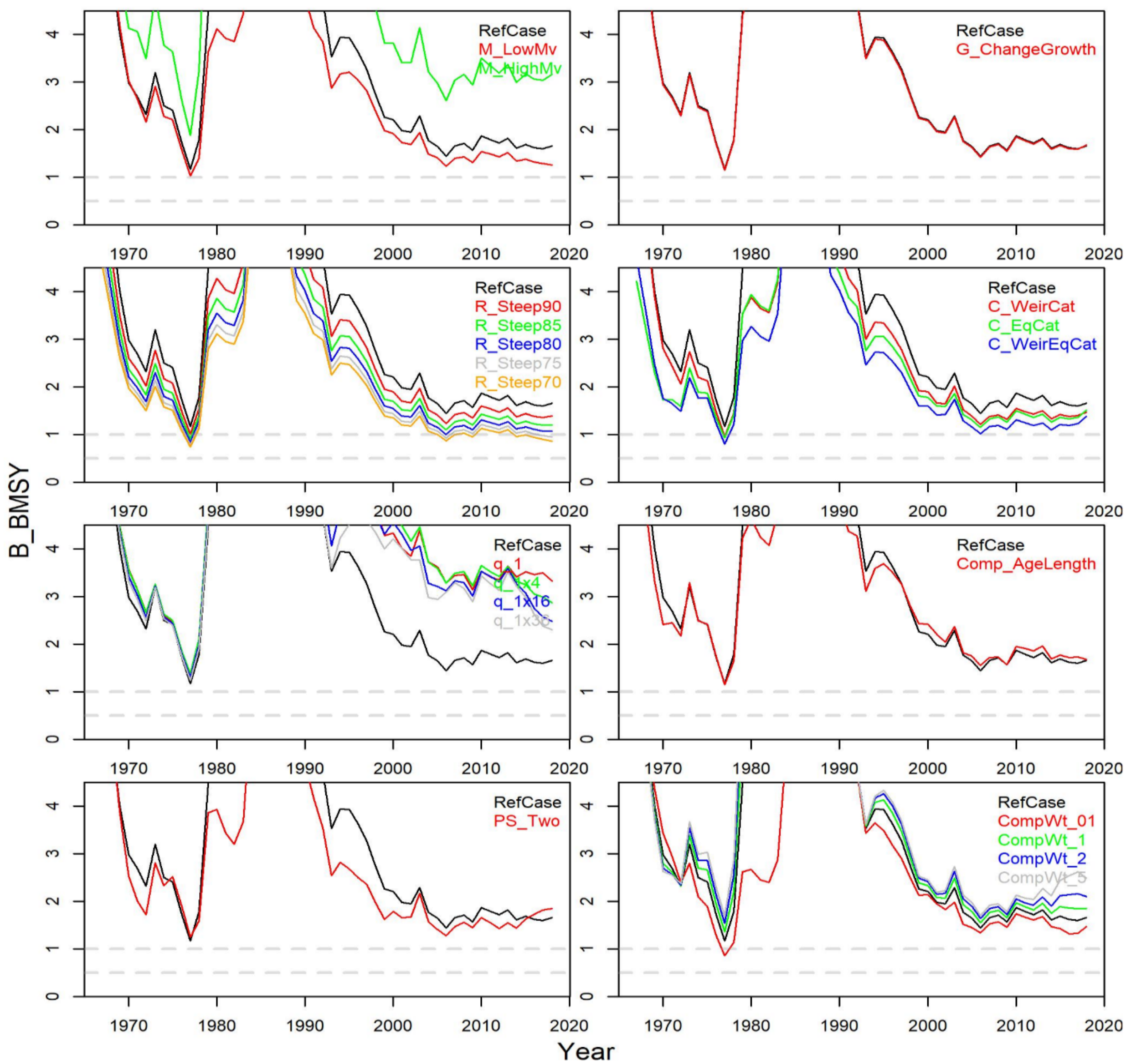


Figure 20c. Mean estimated Spawning Stock Biomass (SSB) relative to MSY levels (B_BMSY) for the various sensitivity operating models described in Table 4.

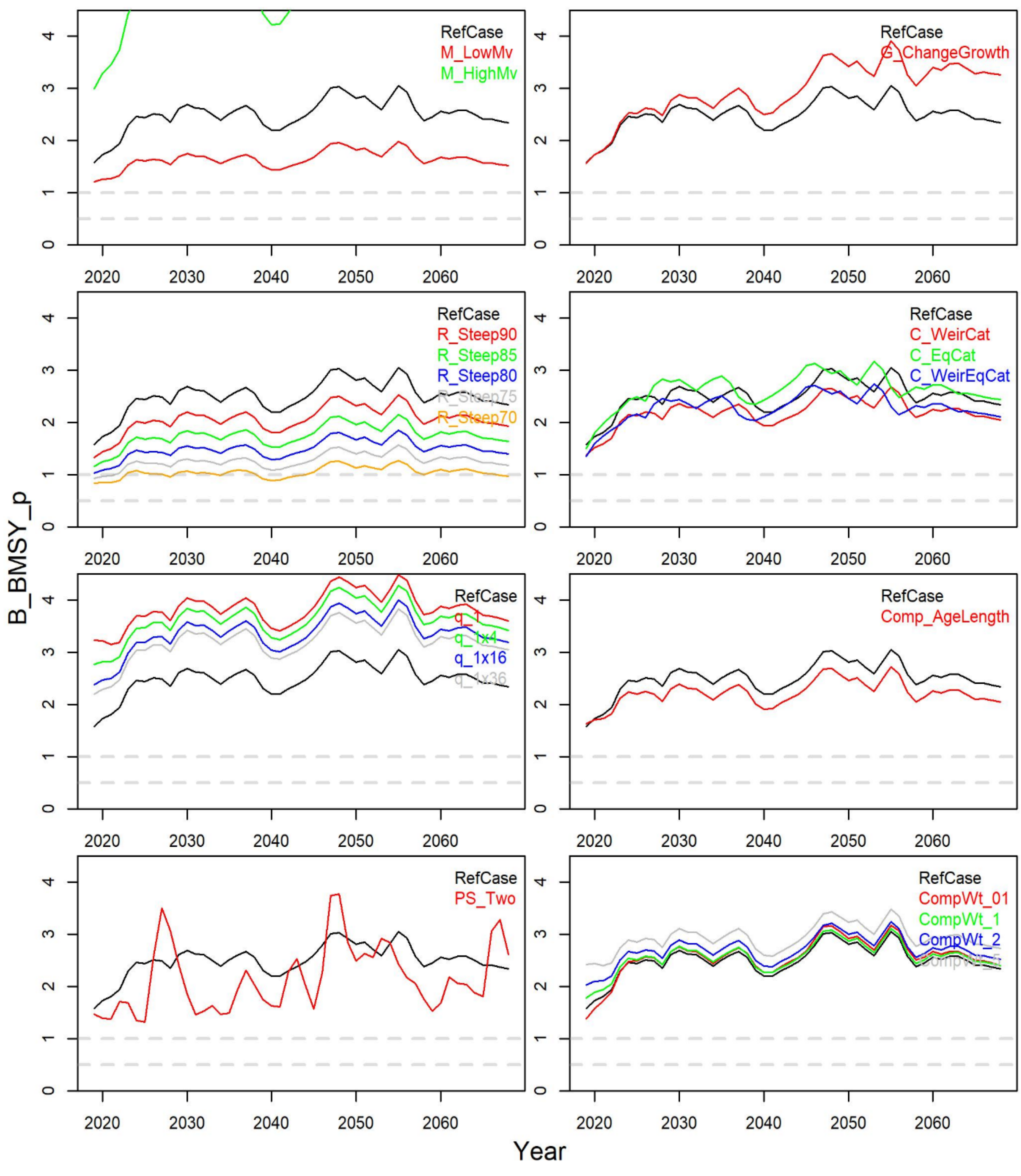


Figure 21. Mean estimates of projected Spawning Stock Biomass (B_BMSY_p) relative to MSY levels for a 50-year projection of a constant current fishing mortality rate scenario (a projection of status quo fishing). Note MSY is recalculated in each future year based on the corresponding growth in that year.

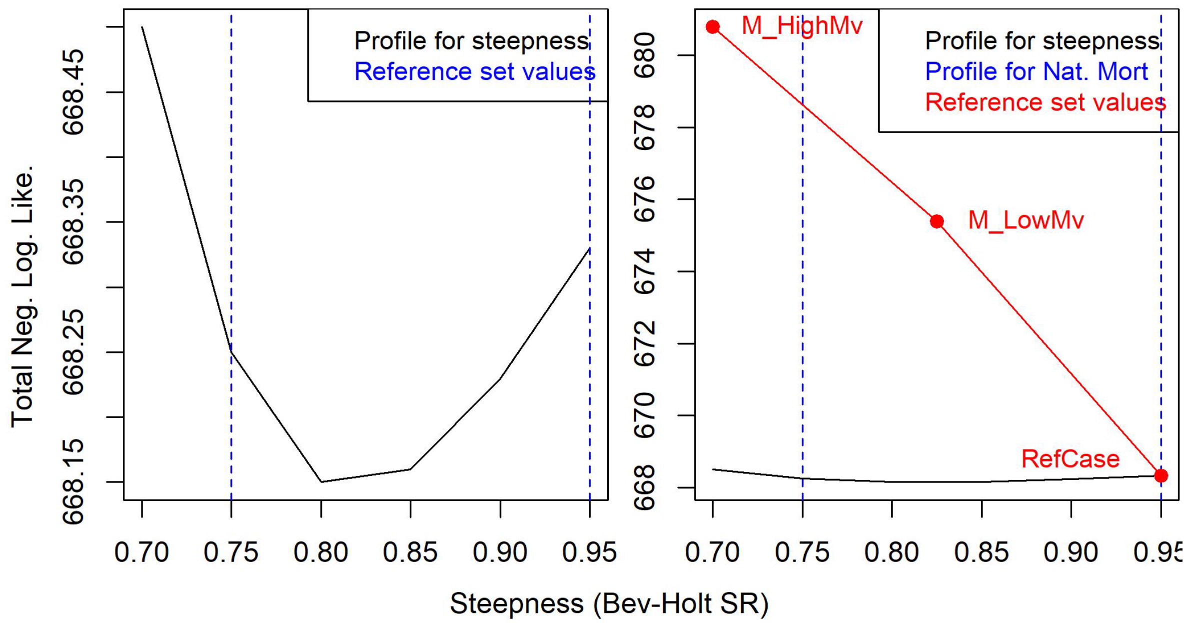


Figure 22. Negative log-likelihood with respect to specified steepness of the stock-recruitment function. Lower values indicate better model fit. The black line in each plot is identical and show the global model objective function value across the range in specified steepness from 0.7 to 0.95. The vertical dashed blue lines show the proposed values for the reference set (0.7 and 0.95; 0.95 is the reference case assumption). The plots differ in the scaling of the y-axis with the righthand plot rescaled over the range in global objective function obtained across the various sensitivity analyses for natural mortality rate, M (M_HighMv is high M that is age varying, M_LowMv is low M that is age varying and the reference case is a fixed age invariant M of 0.35). Note that the red line and points that indicate the various M scenarios are superimposed to demonstrate scale and are not linked to steepness denoted by the x-axis.

REFERENCE SET OPERATING MODELS

Similar to the sensitivity operating models, the biggest determinant of model fit was the specification of natural mortality rate. The constant M = 0.35 assumption of the reference case model provided the best overall fit to the data followed by the low age-varying M scenario (Factor level 2) and then the high age-varying M scenario (Factor level 3) (Table 11).

The reference set operating models span a relatively large range in current stock status (SSB relative to MSY levels between 0.75 and 3), sustainable harvest rate (UMSY in the range 20%–60% of vulnerable biomass) and asymptotic maximum sustainable yield (50–150kt). These constitute a reasonably challenging test bed for management procedures given that no single calibration to the Acoustic index or target harvest rate will necessarily perform well among all reference set operating models.

Estimates of current stock status (SSB relative to MSY levels) are more optimistic than previous model fits primarily due to the specification of natural mortality rates that are higher in all three levels (a more resilient stock) than values previously assumed.

Variance in model estimates was higher among operating model configurations than within the simulations of an individual operating model (see Figure 25). For example, current stock status is not appreciably wider when presenting model estimates by simulation (Figure 25) compared with the range among mean (averaged over simulations) estimates across alternative models (Figure 23c).

Table 11. As Table 9 but for the reference set operating models. A dash (-) indicates no data.

OM	Code	Total	Equilibrium Catch		Age Composition				Indices		Index catchability	
			Purse seine	Weir	Gill	Other	Purse seine	Weir	Acoustic	Larval	q Acoustic	q Larval
1	**1 A H -**	668.33	-	-	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
2	**2 A H -**	675.39	-	-	59.22	100.34	146.77	155.05	-3.23	38.00	3.48	0.04
3	**3 A H -**	680.78	-	-	58.82	100.82	146.11	158.33	-3.45	41.61	2.32	0.03
4	**1 B H -**	668.33	-	-	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
5	**2 B H -**	675.39	-	-	59.22	100.34	146.77	155.05	-3.23	38.00	3.48	0.04
6	**3 B H -**	680.78	-	-	58.82	100.82	146.11	158.33	-3.45	41.61	2.32	0.03
7	**1 A L -**	668.25	-	-	58.77	100.04	142.70	154.35	-3.45	39.60	2.78	0.03
8	**2 A L -**	679.19	-	-	61.18	99.99	147.50	155.10	-3.89	37.52	2.99	0.04
9	**3 A L -**	680.42	-	-	58.66	100.83	146.14	158.28	-3.39	41.75	2.29	0.03
10	**1 B L -**	668.25	-	-	58.77	100.04	142.70	154.35	-3.45	39.60	2.78	0.03
11	**2 B L -**	679.19	-	-	61.18	99.99	147.50	155.10	-3.89	37.52	2.99	0.04
12	**3 B L -**	680.42	-	-	58.66	100.83	146.14	158.28	-3.39	41.75	2.29	0.03
13	**1 A H +**	654.41	-3.69	-3.25	61.81	99.59	148.43	145.42	-1.20	36.72	3.00	0.04
14	**2 A H +**	659.32	-3.67	-2.68	63.64	99.66	149.52	146.07	-1.70	36.21	3.22	0.05
15	**3 A H +**	670.05	-3.69	-3.36	62.62	99.88	153.92	149.83	-1.21	37.89	2.38	0.03
16	**1 B H +**	654.41	-3.69	-3.25	61.81	99.59	148.43	145.42	-1.20	36.72	3.00	0.04
17	**2 B H +**	659.32	-3.67	-2.68	63.64	99.66	149.52	146.07	-1.70	36.21	3.22	0.05
18	**3 B H +**	670.05	-3.69	-3.36	62.62	99.88	153.92	149.83	-1.21	37.89	2.38	0.03
19	**1 A L +**	653.56	-3.67	-3.01	62.38	99.08	147.40	145.91	-2.06	36.60	2.58	0.04
20	**2 A L +**	690.22	-3.55	-2.55	71.16	98.81	151.83	147.64	4.95	35.97	2.08	0.04
21	**3 A L +**	669.67	-3.69	-3.36	61.96	99.85	154.51	149.93	-1.59	37.33	2.49	0.03
22	**1 B L +**	653.56	-3.67	-3.01	62.38	99.08	147.40	145.91	-2.06	36.60	2.58	0.04
23	**2 B L +**	690.22	-3.55	-2.55	71.16	98.81	151.83	147.64	4.95	35.97	2.08	0.04
24	**3 B L +**	669.67	-3.69	-3.36	61.96	99.85	154.51	149.93	-1.59	37.33	2.49	0.03

Table 12. As Table 10 but for the reference set operating models.

OM	Code	MSY (kt)	FMSY	SSBMSY (kt)	SSBMSY/ SSB0	BMSY (kt)	VBMSY	UMSY	SSB0 (kt)	RefY (kt)	Blow (kt)	MGT (yrs)	SSB/ SSBMSY	D
1	1 A H -	75.587	0.623	109.912	0.142	253.808	182.764	0.414	776.541	71.704	0.040	5.089	1.613	0.229
2	2 A H -	54.496	0.499	103.340	0.136	215.489	155.627	0.350	760.067	52.519	0.013	5.521	1.301	0.177
3	3 A H -	122.781	1.076	90.120	0.110	412.151	211.220	0.582	822.642	113.433	13.982	4.691	3.045	0.335
4	1 B H -	76.134	0.621	111.218	0.143	256.498	184.563	0.413	776.541	67.437	0.040	5.089	1.601	0.229
5	2 B H -	54.921	0.498	104.662	0.138	217.934	157.309	0.349	760.067	47.145	0.013	5.521	1.283	0.177
6	3 B H -	123.591	1.074	91.113	0.111	416.624	213.042	0.581	822.642	114.936	13.982	4.691	3.018	0.335
7	1 A L -	62.089	0.355	168.986	0.210	305.906	239.636	0.259	806.246	52.515	1.633	5.089	0.986	0.207
8	2 A L -	44.562	0.281	156.658	0.200	261.079	207.044	0.215	784.036	38.103	2.219	5.521	0.950	0.190
9	3 A L -	95.262	0.509	171.647	0.203	477.972	290.812	0.328	847.383	79.099	1.317	4.691	1.488	0.302
10	1 B L -	62.571	0.355	170.800	0.212	309.021	241.927	0.259	806.246	42.993	1.655	5.089	0.976	0.207
11	2 B L -	44.944	0.281	158.333	0.202	263.790	209.065	0.215	784.036	29.348	2.250	5.521	0.941	0.190
12	3 B L -	95.929	0.508	173.420	0.205	482.994	293.392	0.327	847.383	69.483	1.334	4.691	1.473	0.302
13	1 A H +	104.516	0.534	161.572	0.143	362.794	289.771	0.361	1130.042	94.695	0.013	5.089	1.238	0.177
14	2 A H +	86.497	0.445	170.391	0.136	349.514	274.411	0.315	1254.164	79.679	0.016	5.521	0.926	0.126
15	3 A H +	150.166	0.872	114.101	0.110	503.533	311.525	0.483	1034.140	134.365	1.282	4.691	2.709	0.298
16	1 B H +	105.302	0.533	163.301	0.145	366.438	292.455	0.360	1130.042	91.664	0.013	5.089	1.221	0.177
17	2 B H +	87.186	0.445	172.348	0.137	353.252	277.143	0.315	1254.164	73.713	0.017	5.521	0.920	0.126
18	3 B H +	151.272	0.872	115.080	0.111	508.635	314.057	0.483	1034.140	141.384	1.282	4.691	2.685	0.298
19	1 A L +	87.775	0.317	252.739	0.210	451.023	376.212	0.233	1203.773	71.044	2.472	5.089	0.757	0.159
20	2 A L +	72.743	0.262	269.116	0.199	443.743	361.302	0.201	1351.531	58.736	3.682	5.521	0.889	0.177
21	3 A L +	117.835	0.437	217.560	0.202	597.332	417.343	0.283	1078.546	94.585	1.714	4.691	1.282	0.259
22	1 B L +	88.463	0.316	255.331	0.212	455.479	379.666	0.233	1203.773	59.838	2.505	5.089	0.750	0.159
23	2 B L +	73.357	0.262	271.822	0.201	448.158	364.597	0.201	1351.531	46.859	3.731	5.521	0.881	0.177
24	3 B L +	118.707	0.437	219.679	0.204	603.448	421.040	0.282	1078.546	85.891	1.735	4.691	1.270	0.259

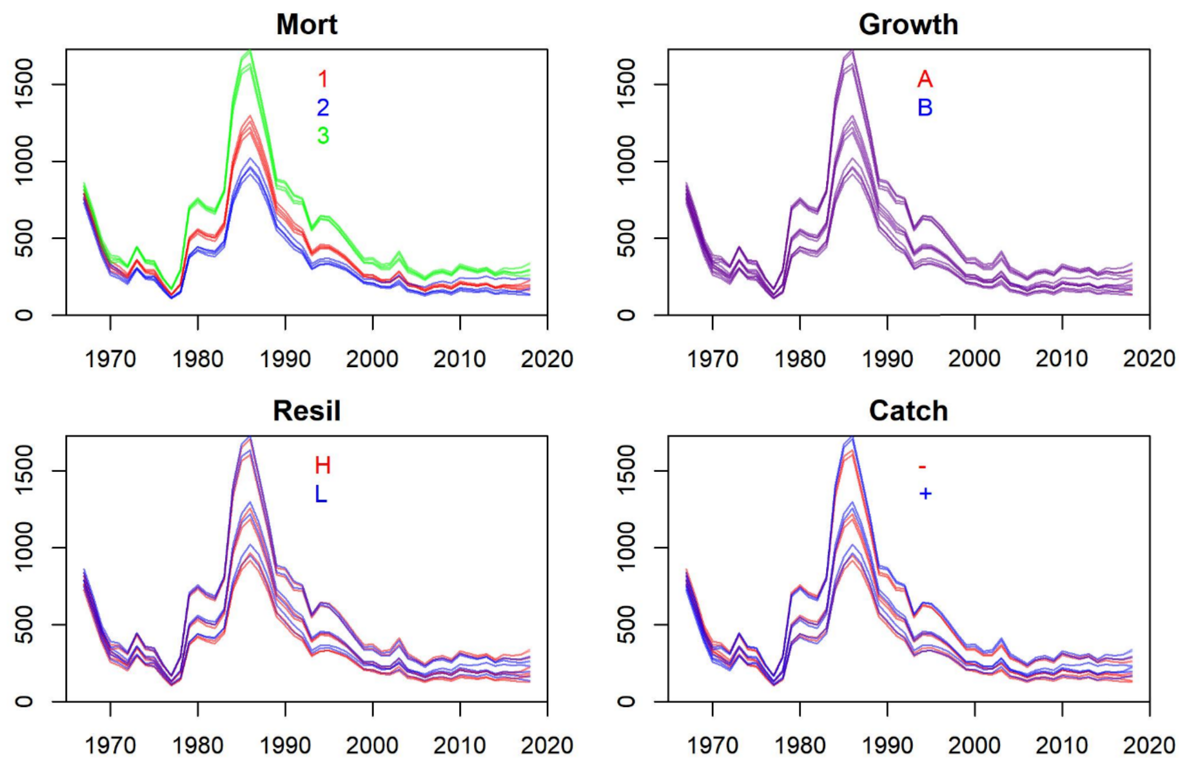


Figure 23a. Estimates of Spawning Stock Biomass (kt) of the various factors of the reference set grid.

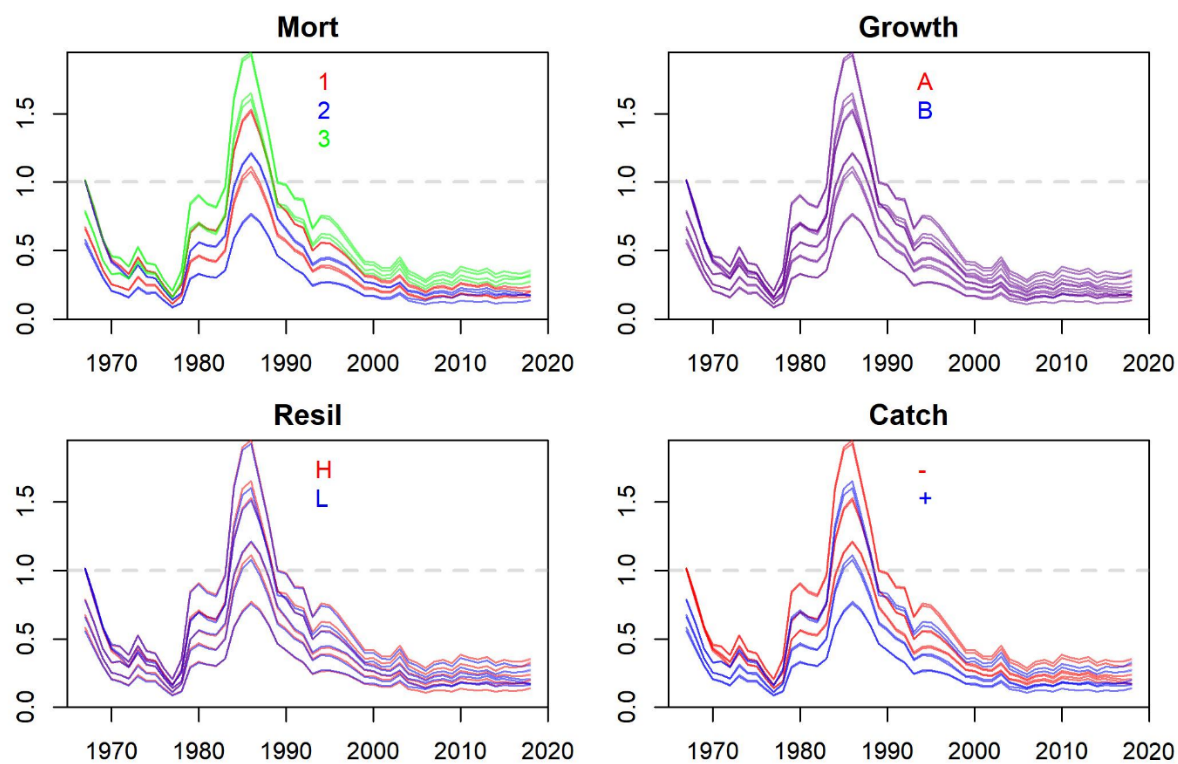


Figure 23b. Estimates of spawning stock depletion for the various factors of the reference set grid.

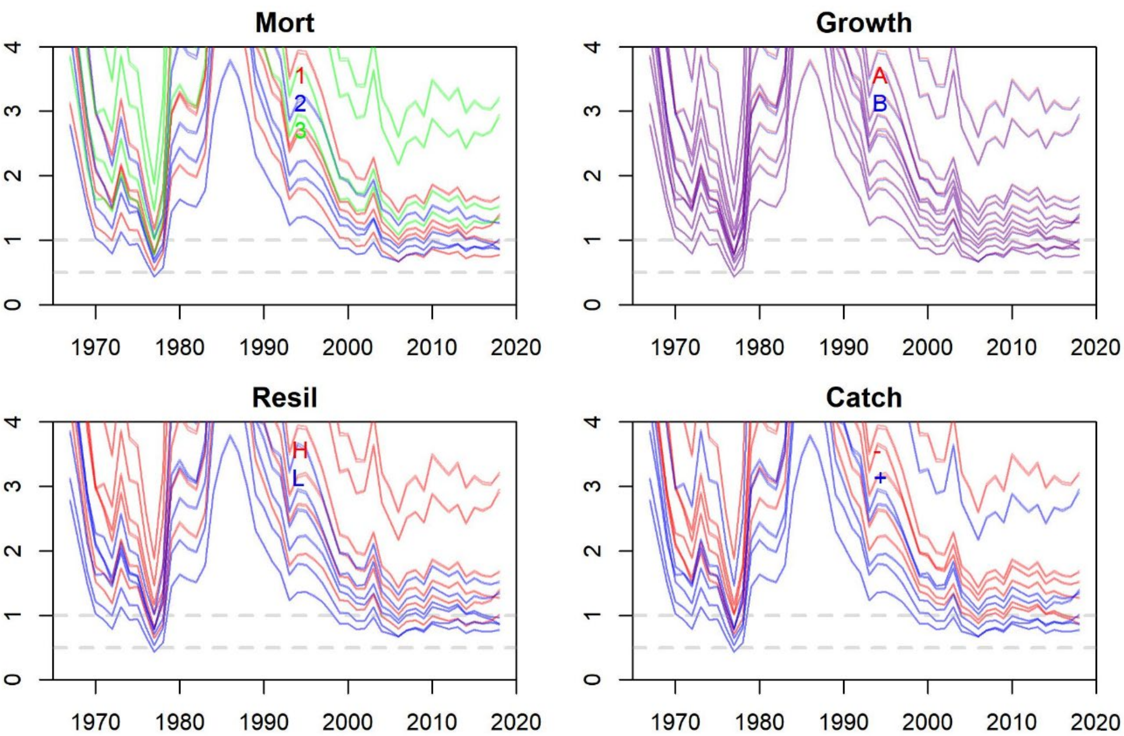


Figure 23c. Estimates of Spawning Stock Biomass relative to MSY levels among the various factors of the reference set operating models.

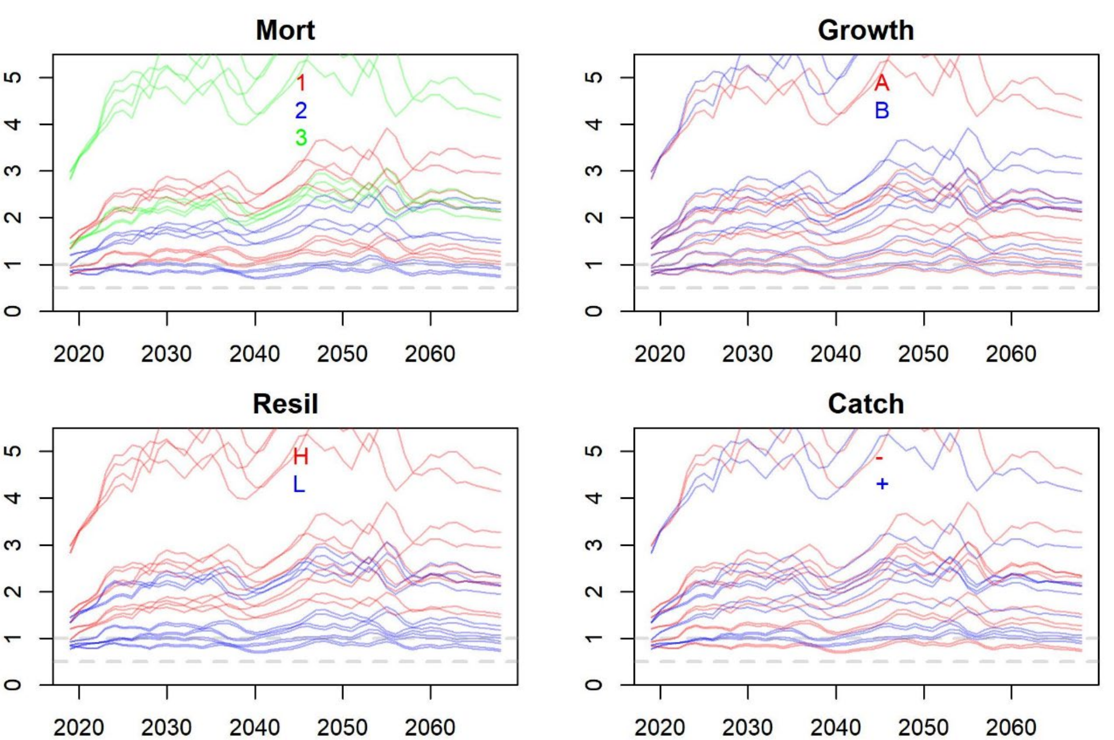


Figure 24. Projected Spawning Stock Biomass relative to MSY levels for current (2018) estimated fishing mortality rate at age.

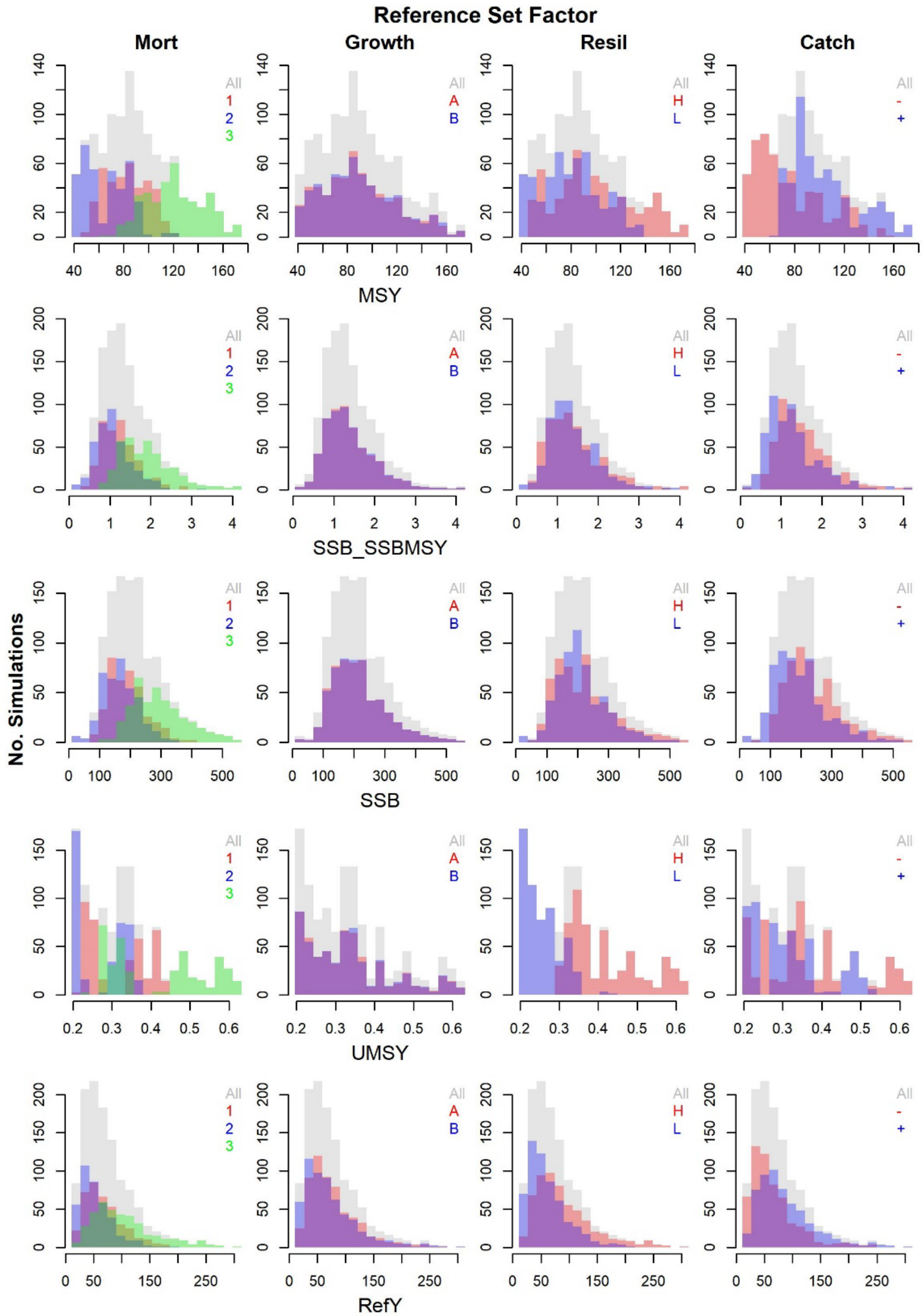


Figure 25. Stochastic model estimates (by simulation, 48 simulations per operating model) aggregated over various factors of the Reference grid of operating models. Maximum Sustainable Yield (MSY), Spawning Stock Biomass (SSB) and Reference Yield (RefY) are in units of kilotons. SSB_SSBMSY is Spawning Stock Biomass in 2018 relative to MSY levels. SSB is current (2018) Spawning Stock Biomass. UMSY is the fraction of vulnerable biomass by weight that was caught in 2018. RefY is the maximum yield that can be obtained from a fixed fishing mortality rate given 2018 selectivity and future fishery conditions (growth and recruitment).

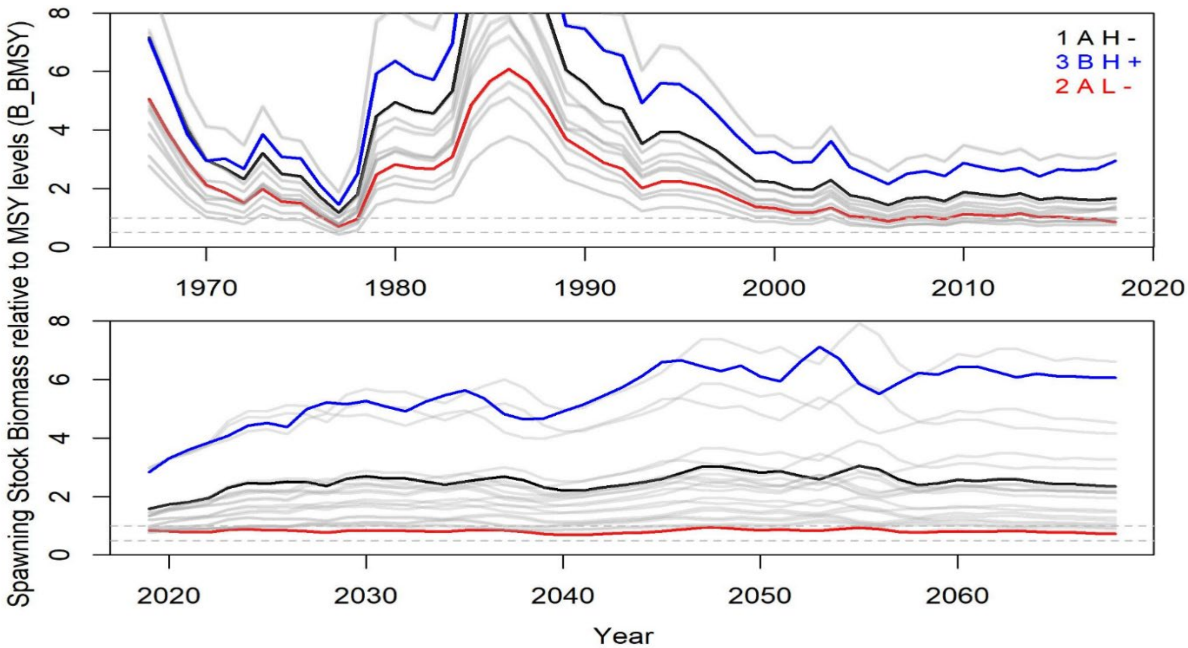


Figure 26. Projections of Spawning Stock Biomass relative to MSY levels given current fishing mortality rate at age for the reference set of operating models (grey lines). Highlighted in black is the reference case (R) operating models. The blue and red lines represent reference set operating models that span a wide range of uncertainty in outcomes and have relatively optimistic (O) and pessimistic (P) outcomes. These three scenarios, R (reference case), O (optimistic, reference set OM #18, '3BH+') and P (pessimistic, reference set OM #8, '2AL+') form the robustness grid that is replicated over factor levels for the robustness set.

ROBUSTNESS SET OPERATING MODELS

The robustness scenarios that provided the greatest contrast to those of the reference set were the alternative scenarios for the stock recruitment relationship (R12 and R13, assuming a Ricker model) and rebuilding (R14-R16) that simulated current stock levels starting from around 50% BMSY levels (Table 14; Figure 27c). R13 (Ricker_O) and R16 (Reb_P) were the only two robustness operating models that would not lead to increasing stock levels given current (2018) fishing mortality rate-at-age (Figure 28). The optimistic rebuilding scenario (R15, Reb_O) led to increasing spawning biomass trends in recent years in contradiction to the acoustic survey index, that may be considered a substantial enough misfit to exclude this operating model from further analyses.

The Ricker stock recruitment function (Ricker_R) provided fractionally better fit to the data (nll = 666.59) compared to the Beverton-Holt stock recruitment function of the reference case operating model (nll = 668.33).

In general, the alternative catch scenarios (R17–R25) all led to more pessimistic outcomes but were not as impactful on model estimates and projections as other axes of uncertainty (e.g., comparing red and black lines in Figure 28). Scenarios where all of the 4VWX catches were included in historical fitting had the biggest impact of all of the catch scenarios.

Generally, the robustness set operating models provided a similar range of UMSY, MSY and SSBMSY as the reference set operating models (Table 14).

Table 13. As Table 9 but for the robustness set operating models.

OM	Code	Total	Equilibrium Catch		Age Composition				Indices		Index catchability	
			Purse seine	Weir	Gill	Other	Purse seine	Weir	Acoustic	Larval	q Acoustic	q Larval
R1	**q1_R**	778.26	0	0	65.89	99.29	144.63	157.28	72.30	38.97	1	0.02
R2	**q1_O**	755.10	-3.63	-2.99	66.88	99.25	150.70	155.35	55.23	38.24	1	0.03
R3	**q1_P**	829.05	0	0	73.91	98.90	161.18	157.23	85.58	37.46	1	0.03
R4	**AgeComp1_R**	1164.12	0	0	112.57	195.08	268.51	297.96	-2.14	42.95	2.98	0.04
R5	**AgeComp1_O**	1172.95	-3.69	-3.46	117.97	193.92	284.28	293.30	1.08	42.15	2.11	0.03
R6	**AgeComp1_P**	1178.82	0	0	116.69	195.07	274.32	299.38	-3.03	40.31	3.28	0.05
R7	**Steep7_R**	668.50	0	0	58.96	99.95	142.75	154.30	-3.51	39.55	2.69	0.03
R8	**Steep7_O**	669.53	-3.69	-3.36	61.79	99.82	154.40	149.99	-1.58	37.27	2.50	0.03
R9	**Length_R**	1990.22	0	0	64.54	100.77	145.73	160.00	-2.50	39.33	3.55	0.05
R10	**Length_O**	1985.99	-3.61	-3.29	68.44	100.89	148.09	156.09	-1.23	37.45	2.86	0.04
R11	**Length_P**	2017.68	0	0	71.02	100.30	149.11	161.27	-3.98	37.70	3.39	0.05
R12	**Ricker_R**	666.59	0	0	58.37	100.24	140.96	155.06	-3.60	42.69	2.83	0.04
R13	**Ricker_O**	681.41	0	0	58.23	100.10	146.73	153.77	0.89	39.82	3.70	0.04
R14	**Reb_R**	668.33	0	0	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
R15	**Reb_O**	670.05	-3.69	-3.36	62.62	99.88	153.92	149.83	-1.21	37.89	2.38	0.03
R16	**Reb_P**	679.19	0	0	61.18	99.99	147.50	155.10	-3.89	37.52	2.99	0.04
R17	**US20_R**	662.26	0	0	56.36	99.77	143.62	152.05	-2.37	37.97	2.68	0.03
R18	**US20_O**	669.21	-3.69	-3.36	62.58	99.85	153.78	149.74	-1.20	37.68	2.20	0.03
R19	**US20_P**	679.49	0	0	61.04	99.24	150.31	152.61	-2.94	35.89	2.43	0.04
R20	**US40_R**	661.10	0	0	56.37	99.76	143.25	152.03	-2.34	37.63	2.48	0.03
R21	**US40_O**	668.60	-3.69	-3.35	62.54	99.84	153.71	149.66	-1.18	37.51	2.06	0.03
R22	**US40_P**	677.60	0	0	60.87	99.22	149.64	152.59	-2.91	35.59	2.25	0.03
R23	**ALL_R**	662.03	0	0	54.90	99.69	143.62	152.26	-2.52	37.78	2.43	0.03
R24	**All_O**	669.08	-3.69	-3.36	61.50	99.76	153.56	150.03	-1.34	37.75	2.00	0.03
R25	**All_P**	679.05	0	0	59.53	99.27	150.46	152.69	-3.09	35.93	2.29	0.03

Table 14. As Table 10 but for the robustness set operating models

OM	Code	MSY (kt)	FMSY	SSBMSY (kt)	SSBMSY/ SSB0	BMSY (kt)	UMSY	SSB0 (kt)	RefY (kt)	Blow (kt)	MGT (yrs)	SSB/ SSBMSY	D
1	**RefCase**	75.587	0.623	109.912	0.142	253.808	0.414	776.541	71.704	0.040	5.089	1.613	0.229
R1	**q1_R**	88.222	0.574	129.871	0.141	298.564	0.384	921.246	84.047	10.255	5.089	2.596	0.366
R2	**q1_O**	158.785	0.849	120.618	0.111	533.058	0.472	1,088.918	148.774	8.607	4.691	2.757	0.306
R3	**q1_P**	51.727	0.258	185.969	0.200	308.640	0.198	931.460	44.386	2.628	5.521	2.075	0.415
R4	**AgeComp1_R**	70.384	0.616	102.448	0.141	236.545	0.410	724.765	66.725	0.034	5.089	1.865	0.263
R5	**AgeComp1_O**	147.305	0.853	111.790	0.111	494.537	0.473	1,009.464	137.737	2.246	4.691	3.099	0.344
R6	**AgeComp1_P**	40.463	0.281	142.312	0.200	237.158	0.215	712.171	34.613	2.059	5.521	0.980	0.196
R7	**Steep7_R**	58.764	0.316	181.472	0.223	316.632	0.234	814.386	48.251	3.196	5.089	0.906	0.202
R8	**Steep7_O**	112.005	0.384	240.746	0.220	622.041	0.254	1,093.969	76.177	3.405	4.691	1.127	0.248
R9	**Length_R**	62.589	0.841	93.524	0.151	211.120	0.546	618.802	60.032	1.026	5.089	1.742	0.263
R10	**Length_O**	122.703	1.261	99.287	0.120	420.665	0.688	828.804	112.457	1.621	4.691	2.717	0.326
R11	**Length_P**	37.510	0.340	127.888	0.202	213.514	0.259	633.785	32.137	1.915	5.521	0.970	0.196
R12	**Ricker_R**	96.090	0.399	235.338	0.303	440.512	0.287	776.802	76.657	6.504	5.089	0.630	0.191
R13	**Ricker_O**	55.147	0.241	228.739	0.300	364.909	0.188	762.374	41.663	17.733	5.521	0.650	0.195
R14	**Reb_R**	75.587	0.623	109.912	0.142	253.808	0.414	776.541	71.704	0.040	5.089	0.500	0.071
R15	**Reb_O**	151.272	0.872	115.080	0.111	508.635	0.483	1,034.140	141.384	1.282	4.691	0.505	0.056
R16	**Reb_P**	44.562	0.281	156.658	0.200	261.079	0.215	784.036	38.093	2.219	5.521	0.500	0.100
R17	**US20_R**	96.199	0.533	148.882	0.143	334.163	0.361	1,042.174	88.833	0.013	5.089	1.427	0.204
R18	**US20_O**	162.246	0.878	123.329	0.111	545.706	0.486	1,107.803	151.545	1.748	4.691	2.703	0.300
R19	**US20_P**	58.512	0.261	216.667	0.199	357.210	0.201	1,088.258	48.800	3.328	5.521	1.005	0.200
R20	**US40_R**	102.951	0.534	158.973	0.143	357.167	0.361	1,112.981	95.277	0.014	5.089	1.448	0.207
R21	**US40_O**	173.250	0.884	131.588	0.111	582.869	0.489	1,181.622	161.730	2.362	4.691	2.712	0.301
R22	**US40_P**	62.537	0.260	231.118	0.199	381.183	0.200	1,160.586	52.289	3.519	5.521	1.015	0.202
R23	**All_R**	99.808	0.699	159.721	0.151	351.698	0.472	1,059.433	93.578	0.014	5.089	1.391	0.210
R24	**All_O**	164.918	1.283	128.871	0.115	560.000	0.713	1,119.614	153.243	3.391	4.691	2.687	0.309
R25	**All_P**	61.248	0.293	224.501	0.204	369.159	0.224	1,102.664	51.719	3.551	5.521	0.936	0.191

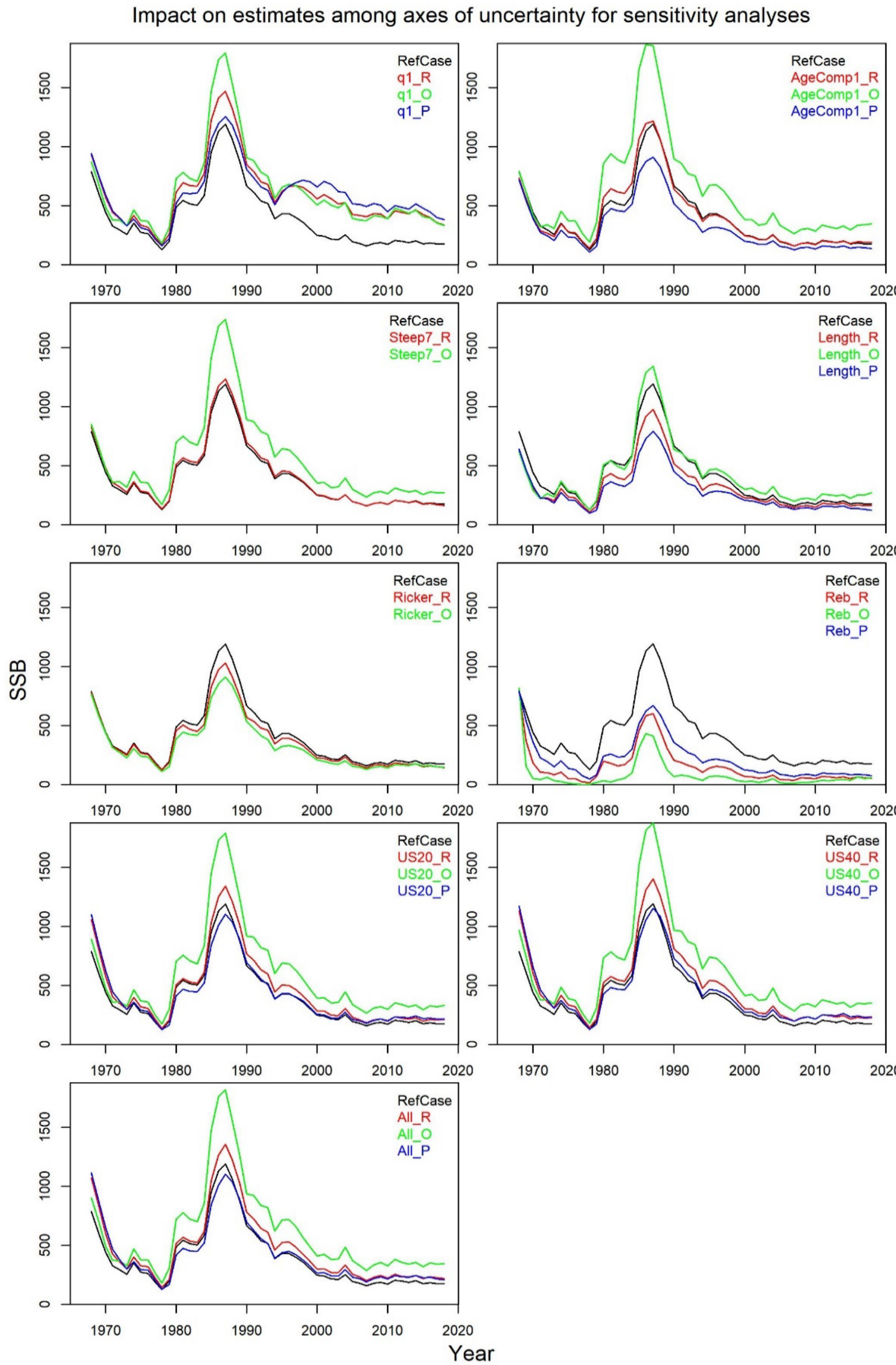


Figure 27a. Robustness operating model mean estimates of Spawning Stock Biomass (SSB)(kt).

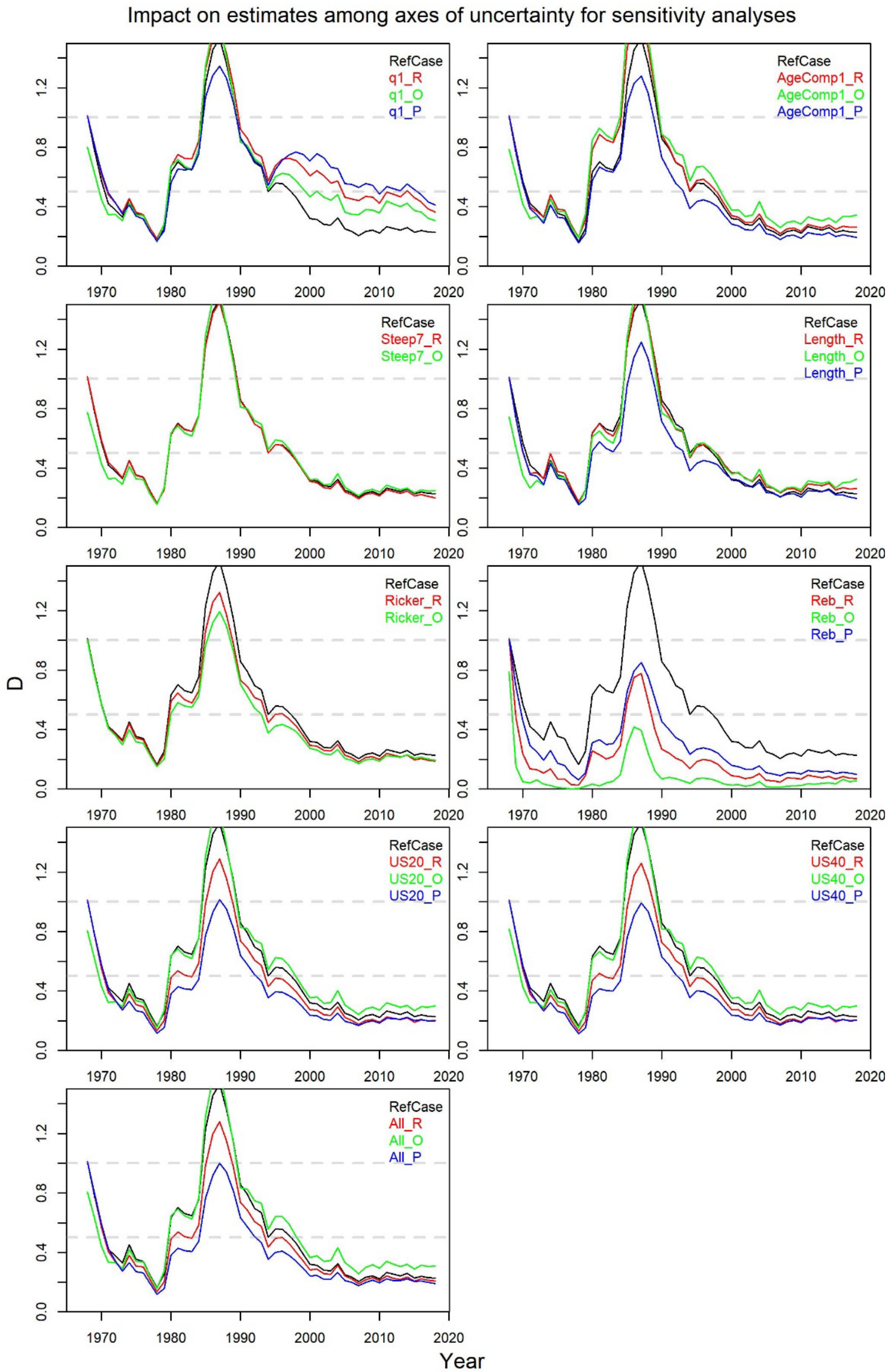


Figure 27b. Robustness operating model mean estimates of stock depletion (D, Spawning Stock Biomass relative to unfished levels).

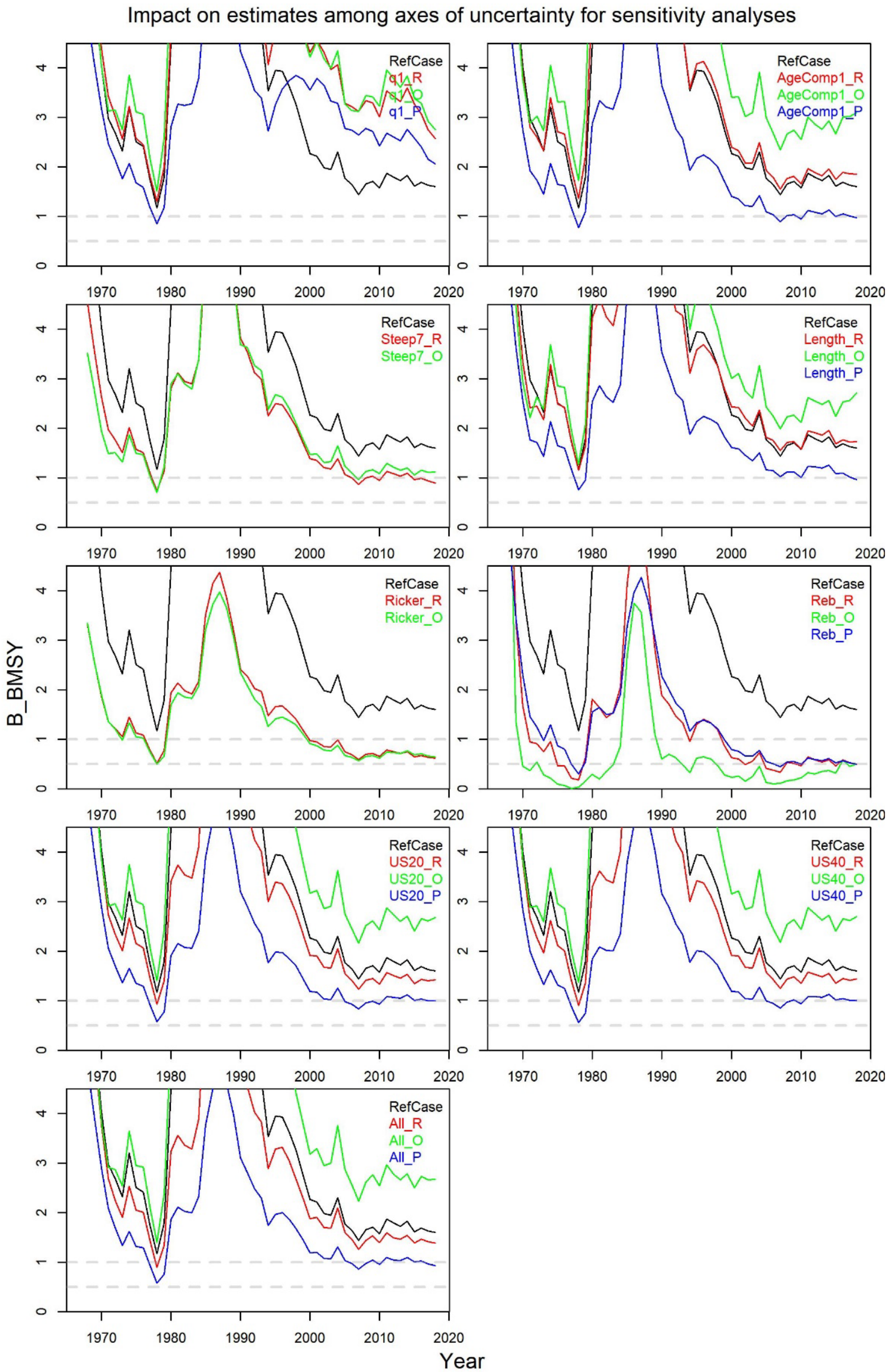


Figure 27c. Robustness operating model estimates of Spawning Stock Biomass relative to MSY levels (B_BMSY).

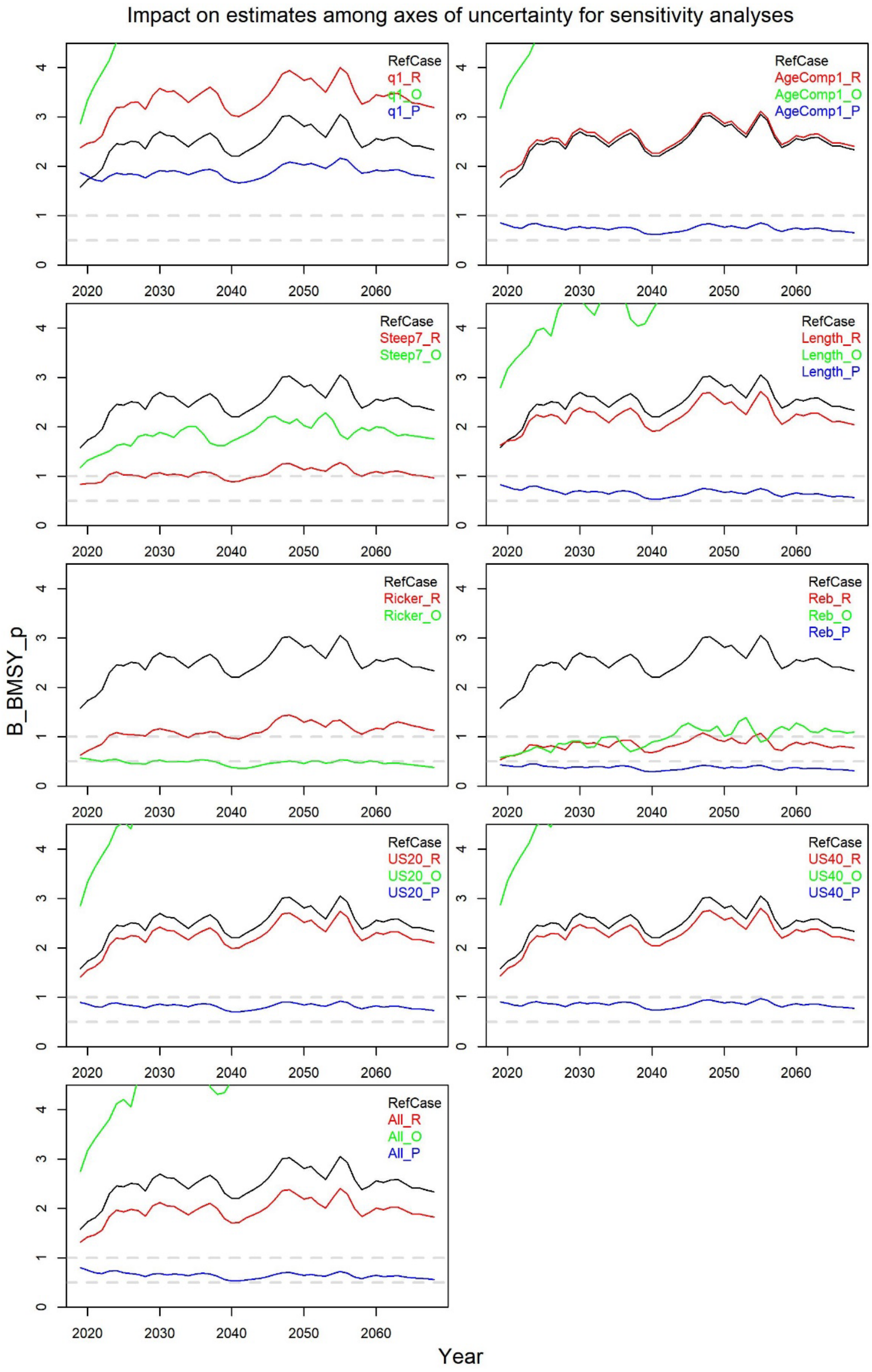


Figure 28. Projected mean Spawning Stock Biomass (SSB) (kt) given current fishing mortality rate at age for the robustness operating models.

DISCUSSION

FINDINGS

The operating model conditioning approach applied in these analyses was remarkably stable over a very wide range of scenarios for fishery dynamics and input data. The model could be fitted very rapidly using a maximum likelihood approach. Drawing stochastic samples from the parameter variance-covariance matrix provided a means of characterising within-model parameter uncertainty with evidence for very few spurious or implausible draws. These stochastic simulations were relatively precise, often falling well within the observed data. Nonetheless the current approach goes further than other MSE frameworks such as those applied to Southern and Atlantic bluefin tunas that do not include model parameter uncertainty and only include observation and process errors for MSE projections.

In general, the variance in model estimates among OM types (i.e., MLE fits among operating models) was substantially higher than variance among stochastic simulations within operating models. This suggests that there may be little advantage to sampling a very large number of simulations other than the calculation of performance metrics of higher precision.

In general, the estimates of current stock status among the reference grid of operating models was somewhat more optimistic than previous analyses presented to the Atlantic Herring Working Group that assumed a much lower rate of natural mortality (0.2). The robustness set of operating models did, however, include at least 5 scenarios that would test the ability of a Candidate Management Procedure (CMP) to recover from stock levels substantially below target MSY levels.

The base model fit revealed a similar conflict between the scale (fishing mortality rate) inferred by the composition data and the acoustic survey. In this case, the base model estimated vulnerable biomass that was around three times smaller than that of the acoustic survey (survey q = 3.03). This is less severe than found in a previous investigation of assessment models (average discrepancy across assessments was around 5, DFO 2011).

While the fits to the age composition of the survey and the primary fleet (purse seine) were very good, the fit to fleet age composition data could be poor for other fleets in certain years, pointing to model misspecification and the inability to approximate processes relating to either variable selectivity or availability. Previously it has been hypothesized that splitting the purse seine fishery into juvenile and spawning fleets might address both this misspecification and perhaps also the conflict in inferred scale among data types. The corresponding sensitivity operating model proved this to be wrong.

The reference set operating model scenarios identified here vary substantially in their estimates of current stock status (which fall just under and above SSB_{MSY}), have differing sustainable harvest rates, absolute current stock sizes and varying calibration to the available acoustic survey index (Tables 11 and 12). As a reference set of operating models, these provide a reasonably challenging test of a CMP. A CMP would need to manage exploitation for long-term yield maximization in some operating models whilst maintaining or increasing exploitation rate in others. A CMP assuming a fixed proportion of the acoustic survey index may also struggle due to variable harvest rate and q among operating models. It is likely that given a comparable reference set of operating models there are performance opportunities for CMPs that are either model-based (involve some type of estimation of status) or are adaptive, and do not rely on simple fixed interpretations of the survey index for providing TAC advice.

STRENGTHS OF PROPOSED APPROACH

MSEtool (Huynh et al. 2019 and its operating model conditioning approach that was used here, were developed to address many of the principal criticisms of other MSE frameworks including computation inefficiency (time taken to get results, investigate other model scenarios), lack of transparency (the problem of 'black box' software in processes intended to be open), difficulty in customizing OMs, MPs and performance metrics, inaccessibility and lack of documentation and insufficient flexibility to investigate a wide range of operating model scenarios. It is hoped that the methods and provision of code for these analyses demonstrates the advances made in the development of MSEtool.

WEAKNESSES OF PROPOSED APPROACH

The operating model approach still could not reconcile the data conflicts discussed above (i.e., q for the acoustic index > 1) that previously precluded the use of assessment models to provide management advice. There is general agreement among the Atlantic Herring Working Group that q could be greater than 1 due to uncertainty in turnover on the spawning grounds and the target strength that is used to estimate Herring biomass from acoustic backscatter.

Due to time constraints it was not possible to provide more in-depth evaluation of the reference case operating model in this report. Additional diagnostics such as retrospective analyses and simulation testing may be presented later.

The SRA approach does not account for catch observation error and therefore compresses uncertainty in parameter estimates (and management reference points) for individual model fits. This is however less important in MSE where the uncertainties in fishery dynamics estimates among operating models typically dwarf those estimated within any given operating model (as was also the case here).

By aggregating two distinct spawning areas into a single population these models may fail to account for regional changes in productivity or range shifts. While it may be difficult to estimate such changes empirically in model conditioning, MSEtool is inherently a spatial model and such scenarios can be specified theoretically and investigated as robustness operating models.

The reference set of operating models was not appreciably narrower in model estimates and projection outcomes than the robustness set, with the possible exception of the Ricker stock recruitment and rebuilding operating models that had much lower current stock status. It follows that the reference set may be a relatively challenging basis for testing a CMP relative to the robustness set. This is generally not the intended situation for these two sets of scenarios.

UNCERTAINTIES NOT DESCRIBED IN THE CURRENT REFERENCE AND ROBUSTNESS SETS OF OPERATING MODELS

Some scenarios for fishery dynamics were discussed at the MSE workshop on uncertainties in January 2020 but were not presented in the current reference and robustness sets of OMs. Some of these uncertainties were explored prior to OM development and considered to be encompassed in the existing set of OMs. Additional sensitivities analyses were requested during and directly following the meetings on May 23th and May 24th 2020 and these results are provided in Appendix D.

The relationship between growth and ocean temperature was explored using surface and bottom temperatures from the AZMP Prince 5 station and predicting weights based on projected changes in ocean temperatures for the Prince 5 stations from the BIO North Atlantic Model (BNAM) (Brickman et al. 2016). The BNAM model considers four different ocean temperature projections (two different projections for CO_2 emissions and two time periods (2055 and 2075).

Temperature data for Prince 5 were obtained from the AZMP program and the DFO Oceanographic Database for 1970–2018. Linear regressions were conducted between weight (log_{10}-transformed) and temperature separately by age. The predictor variables that were explored were:

- mean annual temperature (surface and bottom)
- mean winter temperature (months 1 to 3)
- mean spring temperature (months 4 to 6)
- mean summer temperature (months 7 to 9)
- mean fall temperature (months 10 to 12)

Mean temperatures for the time series 1970–2018 were calculated as the mean of the least squares means from an ANOVA of temperature vs. month (categorical factor) and year (categorical factor), in order to control for missing data in a particular month. The fall temperature predictor variable explained the most variability in weight (selected based on the mean r^2 across ages) was selected as the variable for evaluation. Temperature projections for the highest CO2 emissions in 2055 and 2075 were selected as the two future climate scenarios for evaluation. Temperature deviations from the baseline climatology (defined in the model as the mean temperature from 1986–2005) for the area that contains the Prince 5 station were provided by David Brickman (D. Brickman, DFO, pers. comm.). These deviations were added to the baseline climatology for Prince 5 and used as future temperature projections for 2055 and 2075. The projected mean temperatures actually fell within the range of temperatures observed in the 1970–2018 time series. The predicted mean weights-at-age based on the temperature projections for Herring ages 3+ were greater than the two scenarios included in the reference OM set. An additional scenario for projections of weight-at-age was therefore not generated.

The influence of changes in seal abundance in relation to changes in natural mortality was explored. Guénette and Stephenson (2012) estimated the biomass of Herring consumed from 1970 to 2006) for the primary predator groups (Figure 29). In 2006, gray seals comprised of approximately 15% of the biomass of Herring consumed by all predators. By 2017, gray seal abundance had increased 68% (Hammill et al. 2017) relative to 2006. Assuming consumption rates remain unchanged and that the biomass consumed is proportional to the mortality rate from 2006 to 2017, the increase in gray seal abundance would result in approximately a 10% increase in natural mortality rate. A time-varying M scenario that addresses the change in gray seal abundance could not be developed at this time because the contribution of gray seals to the total biomass is relatively small and a scenario needs to consider the contribution of other species. Further research would need to be conducted to properly estimate a time-varying M for Herring.

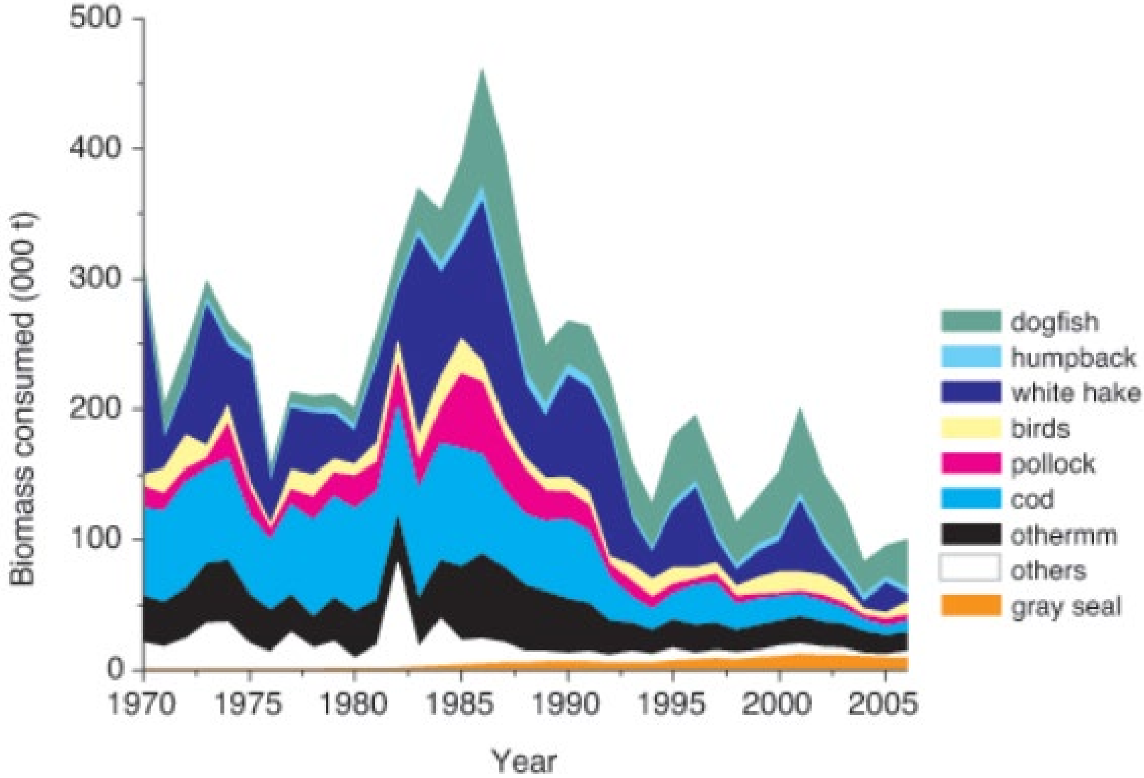


Figure 29. Herring biomass consumed by predators in the base scenario (figure from Guénette and Stephenson 2012).

Other specific uncertainties discussed at the at the workshop on uncertainties in January 2020 included climate change, changes in ocean conditions, diet, phenology, larval condition, adult condition, fecundity, and sub-stock structure (e.g., return of spawning to particular areas). Although these uncertainties have not been directly included in the reference and robustness OMs, some of the influences of changes in these uncertainties would result in changes in natural mortality, recruitment, and growth which have been captured with upper and lower limits in the reference OM set.

An additional uncertainty that has been explored is ageing error. Ageing error matrices have been defined for the period of 2010–2018 for comparisons between multiple reads of the same otolith. Melvin and Campana (2010) present a more serious aging error of underestimating the true age of Herring aged 5 and older. A robustness OM will be explored in the future to assess the influence of this uncertainty.

Although not fully demonstrated here, it worth listing some features of the MSEtool operating model that are potentially relevant to possible scenarios for Herring operating models. The underlying operating model includes specified arrays for movement that are specific to area, age and year allowing for complex shifts in distribution that vary among age classes. This allows for robustness testing of shifts in spawning location for example. The model also accepts arrays for recruitment strength allowing for scoping of future productivity scenarios. Additionally, it is possible to specify time varying natural mortality-at-age in order to account for, for example, changing number of marine predators in future robustness operating models.

ATLANTIC HERRING MSE PRIORITIES

Preliminary Performance Objectives and Metrics

Ideally, stakeholders and managers would have achievable fishery management objectives explicitly stated ahead of technical development of an MSE framework (Punt et al. 2014), however, experience with MSE suggests that this is unrealistic and that the process is necessarily iterative providing opportunities for stakeholders and managers to see what outcomes are possible and how best to operationalize broader objectives (Nakatsuka 2017). Typically, stakeholders initially identify a large number of objectives (e.g., high yields) and corresponding metrics (e.g., mean yield 2026–2030), most of which are later found to be redundant (collinear with other performance metrics) when considered together in a working MSE framework. At least one iteration is required to visualize metrics (quantitative expressions of objectives) and another to understand the relationships among proposed metrics in order to simplify results to a more concise and digestible format.

An agreed set of performance metrics is required for the full process of CMP development as developers typically use metrics to tune their CMPs for comparability and optimal performance.

Operating Model ‘red face’ Tests and Plausibility

The presentation of operating models in this document provides a basis to start considering when an operating model should not be considered as a plausible representation of the system or inappropriate for CMP testing. For example, a negative correlation of model predicted vulnerable biomass and the corresponding acoustic survey (e.g., scenario q1). This is an example of a so-called “red face test”: a qualitative statement to exclude an operating model from consideration. Establishing a range of red face tests is the first priority. Later, it may be necessary to consider a quantitative system of evaluating operating model plausibility if equal weighting of operating models is considered unacceptable.

CONCLUSIONS

An approach for conditioning operating models was described and demonstrated for use in management strategy evaluation for Atlantic Herring off southwest Nova Scotia and the Bay of Fundy, in NAFO area 4VWX, meeting objectives 1–7 (Table 2).

The estimation framework demonstrated sufficient flexibility to construct operating models that can span the major uncertainties for the fishery, and computationally efficient enough to be investigated in real-time during workshops.

A plausible reference set of 24 operating models was specified and fitted which spanned a range of scenarios for current stock status, magnitude of the current stock and the sustainable rate of exploitation. Additionally, 25 robustness operating models were specified that can be used to further discriminate among candidate management procedures.

ACKNOWLEDGEMENTS

While the authors are responsible for any errors or inaccuracies in this document, we wish to thank the valuable input of the Atlantic Herring MSE Working Group. Many thanks to Gary Melvin, Michelle Greenlaw, Rob Stephenson, Christa Waters, Jenna Munden, and Tara McIntyre for their guidance and support of this work. MSEtool and DLMtool development has been funded by the Natural Resources Defense Council, a DFO partnership agreement

(PAC2016.12), the Marine Stewardship Council, the David & Lucile Packard Foundation and the United Nations Food and Agricultural Organization.

REFERENCES CITED

Brickman, D., Wang, Z., and Detracey, B. 2016. High resolution future climate ocean model simulations for the Northwest Atlantic Shelf region. Can. Tech. Rep. Hydrogr. Ocean Sci. 315: 1488–5417. [accessed July 2020]

Bundy, A., Link, J.S., Smith, B.E., and Cook, A.M. 2011. You are what you eat, whenever or wherever you eat it: An integrative analysis of fish food habits in Canadian and U.S.A. waters. Journal of Fish Biology. 78: 514–539.

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

Carruthers, T.R. 2017. Capelin in the Gulf of St Lawrence: a DFO Case Study Operating model. [accessed July 2020]

Carruthers, T.R. 2019a. Fall Spawning Atlantic Herring in NAFO Region 4T: a DFO Case Study Operating Model.

Carruthers, T.R. 2019b. Spring Spawning Atlantic Herring in NAFO Region 4T: a DFO case study operating model.

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

Carruthers, T.R., and Hordyk, A.R. 2019. Using management strategy evaluation to establish indicators of changing fisheries. Can. J. Fish. Aquat. Sci. 76: 1653–1668.

Carruthers, T.R. and Hordyk, A.R. 2018a. Using management strategy evaluation to establish indicators of changing fisheries. Can. J. Fish. Aquat. Sci. 76(9): 1653–1668.

Carruthers, T.R. and Hordyk, A.R. 2018b. The Data-Limited Methods Toolkit (DLMtool): An R package for informing management of data-limited populations. Methods in Ecology and Evolution. 9: 2307–2413.

Carruthers, T.R. and Hordyk, A.R. 2019. DLMtool: Data-limited methods toolkit. R package version 5.4. [accessed July 2020]

CDFW. 2018. 2018 Master plan for fisheries: A guide for implementation of the Marine Life Management Act. California Department of Fish & Wildlife. [accessed July 2020]

Clark, K.J. 2006. An examination of turnover rate of herring on the spawning grounds of Scots Bay and German Bank using tagging data. DFO Can. Sci. Advis. Sec. Res. Doc. 2006/47.

Clark, D.S., Clark, K.J, Claytor, R., Leslie, S., Melvin, G.D., Porter, J.M., Power, M.J., Stone, H.H., and Waters, C. 2012. Limit Reference Point for Southwest Nova Scotia / Bay of Fundy spawning component of Atlantic Herring, *Clupea harengus* (German Bank and Scots Bay). DFO Can. Sci. Advis. Sec. Res. Doc. 2012/025. iii + 14 p.

Cochrane, K L., Butterworth, D.S., De Oliveira, J.A.A., and Roel, B.A. 1998. Management procedures in a fishery based on highly variable stocks and with conflicting objectives: experiences in the South African pelagic fishery. Rev. Fish. Biol. Fisher. 8, 177–214.

Deroba, J.J. 2015. Atlantic Herring operational assessment report 2015. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 15–16; 30 p.

Deroba, J.J. 2017. Technical details of operating models and harvest control rules used in the Georges Bank/Gulf of Maine Atlantic Herring management strategy evaluation. p.1–34.

DFO. 2003. 2003–2006 Scotia-Fundy Fisheries Integrated Herring Management Plan, NAFO subdivisions 4WX, 4Vn and 5Z. Department of Fisheries and Oceans.

DFO. 2004. 4VWX Herring. Can. Science Advis. Sec. Stock Status Rep. 2004/034.

DFO. 2005. Stock Assessment Report on 4VWX Herring. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2005/033 (Revised).

DFO. 2007. Proceedings of the Maritimes Provinces Regional Advisory Process on the Assessment Framework for 4VWX Herring; 31 October–1 November 2006 and 9–11 January 2007. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2007/002.

DFO. 2011. Proceedings of the Maritimes Region Science Advisory Process on the Assessment Framework for Southwest Nova Scotia/Bay of Fundy Herring; 24–28 January 2011. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2011/031: iv + 28p.

DFO. 2015. 2015 Assessment of 4VWX Herring. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/040.

DFO. 2017. DLMtool Phase II: Developing a Management Strategy Evaluation Package for Advancing the Science and Management of Data-limited and At-risk Canadian Fish Stocks. Project Number: PAC2016.12. [accessed July 2020]

Foote, K.G. 1987. Fish target strengths for use in echo integrator surveys. The Journal of the Acoustical Society of America. 82: 981–987.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68(6): 1124–1138.

Guénette, S. and Stephenson, R.L. 2012. Accounting for Predators in Ecosystem-Based Management of Herring Fisheries of the Western Scotian Shelf, Canada. *In*: G.H. Kruse, H.I. Browman, K.L. Cochrane, D. Evans, G.S. Jamieson, P.A. Livingston, D. Woodby, and C.I. Zhang (eds.), Global Progress in Ecosystem-Based Fisheries Management. Alaska Sea Grant, University of Alaska Fairbanks.

Hammill, M.O., den Heyer, C.E., Bowen, W.D., and Lang, S.L.C. 2017. Grey seal population trends in Canadian waters, 1960–2016 and harvest advice. DFO Can. Sci. Advis. Sec. Res. Doc. 2017/052. v + 30 p.

Hordyk, A.R, Newman, D., Carruthers, T., and Suatoni, L. 2017. Applying Management Strategy Evaluation to California Fisheries: Case Studies and Recommendations. [accessed July 2020]

Huynh, Q., Hordyk, A., and Carruthers, T. 2019. MSEtool: Management Strategy Evaluation Toolkit. R package version 1.4. [accessed July 2020]

Kanwit, J.K. and Libby, D.A. 2009. Seasonal movements of Atlantic Herring (*Clupea harengus*): results from a four year tagging study conducted in the Gulf of Maine and Southern new England. J. Northw. Atl. Fish. Sci. 40: 29–39.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B.M. 2016. TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software. 70(5): 1–21.

Kronlund, A.R., Holt, K.R., Cleary, J.S., and Shelton, P.A. 2013. Current approaches for the provision of scientific advice on the precautionary approach for Canadian fish stocks: Sections 8 – Management Strategy Evaluation. DFO Can. Sci. Advis. Sec. Res. Doc. 2013/081. v + 26 p. [accessed July 2019]

Mace, P.M. 1985. Catch rates and total removals in the 4WX herring purse seine fisheries. Canadian Atlantic Fisheries Scientific Advisory Committee Research Document 85/74. [accessed July 2020]

Martell, S. 2017. Integrated Statistical Catch-at-Age Model. The iSCAM project. [accessed July 2020]

Melvin, G.D. and Campana, S.E. 2010. High resolution bomb dating for testing the accuracy of age interpretations for a short-lived pelagic fish, the Atlantic Herring. Environmental Biology of Fishes. 89: 297–311.

Melvin, G.D., Martin, R., and Power, M.J. 2014. Estimating German Bank and Scots Bay Herring Spawning Ground Turnover Rates from Tag Returns. DFO Can. Sci. Advis. Sec. Res. Doc. 2014/068: iv + 22 p.

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

Miller, T. 2016. Memorandum. Proposed blueline tilefish subcommittee report. [accessed July 2020]

McNamee, J., Fay, G., and Cadrin, S. 2016. Data Limited Techniques for Tier 4 Stocks: An alternative approach to setting harvest control rules using closed loop simulations for management strategy evaluation. SEDAR46-RD-08. [accessed July 2020]

NAFO. 2010. Northwest Atlantic Fisheries Organization. Report of the Working Group on Greenland Halibut Management Strategy Evaluation (WGMSE). NAFO/FC Doc. 11/8. [accessed July 2020]

NAFO. 2018. Northwest Atlantic Fisheries Organization. Management Strategy Evaluation – Harvest Control Rules. [accessed July 2020]

Nakatsuka, S. 2017. Management strategy evaluation in regional fishery management organizations – How to promote robust fisheries management in international settings. Fish. Res. 187:127–138.

NEFSC. 2012. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Report. US Department of Commerce, Northeast Fisheries Science Center Reference Document. 12-18; 600 p [accessed July 2020]

NEFSC. 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. US Department Commerce, Northeast Fisheries Science Center Reference Document. 15-24; 251 p. [accessed July 2020]

NEFSC. 2018. 65th Northeast Regional Stock Assessment Workshop (65th SAW) Assessment Summary Report. US National Marine Fisheries Services. Northeast Fisheries Science Center Reference Document. 18-08: 38 p.

Neuenhoff, R., Swain, D. P., Cox, S. P., McAllister, M.K., Trites, A.W., Walters, C.J., and Hammill, M.O. 2018. Continued decline of a collapsed population of Atlantic cod (*Gadus morhua*) due to predation-driven Allee effects. Can. J. Fish. Aqua. Sci. 76 (1): 168–184.

Pennington, M., Burmeister, L.-M., and Hjellvik, V. 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. Fish. Bull. 100:74–80.

Power, M.J., Knox, D., MacIntyre, A., Melvin, G.D., and Singh, R. 2013. 2011 Evaluation of 4VWX Herring. DFO Can. Sci. Advis. Sec. Res. Doc. 2012/085. iv + 85 p.

Punt, A.E. and Donovan, G.P. 2007. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. ICES J. Mar. Sci. 64 603–612.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., and Haddon, M. 2014. Management strategy evaluation: best practices. Fish and Fisheries. 17(2): 303–334.

SEDAR. 2016a. Southeast Data, Assessment and Review. SEDAR 46: Caribbean Data-Limited Species. US National Oceanic and Atmospheric Administration. [accessed July 2020]

SEDAR. 2016b. Southeast Data, Assessment and Review. SEDAR 49: Stock Assessment Report Gulf of Mexico Data-Limited Species. US National Oceanic and Atmospheric Administration. [accessed July 2020]

Singh, R., Dalton, A, Clay, A., and Melvin, G.D. 2016. Summary of 2013 and 2014 Herring Acoustic Surveys in Northwest Atlantic Fisheries Organization (NAFO) Divisions 4VWX. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/005. v + 203 p.

Singh, R., Knox, D., and MacIntyre, A. 2020. 2019 Southwest Nova Scotia/Bay of Fundy Atlantic Herring Framework: Data Inputs. DFO Can. Sci. Advis. Sec. Res. Doc. 2020/028. v + 123 p.

Stephenson, R.L., Lane, D.E., Aldous, D.G., and Nowak, R. 1993. Management of the 4WX Atlantic Herring (*Clupea harengus*) fishery: an evaluation of recent events. Can. J. Fish. Aquat. Sci. 50: 2742–2757.

Stephenson, R.L., Power, M.J., Laffan, S.W., and Suthers, I.M. 2015. Tests of larval retention in a tidally energetic environment reveal the complexity of the spatial structure in herring populations. Fisheries Oceanography. 24: 553–570.

Stobo, W.T. and Fowler, G.M. 2009. Herring tagging in the vicinity of the Scotian shelf and Gulf of St. Lawrence by the Maritimes region, 1973–1982. Can. Tech. Rep. Fish. Aquat. Sci. 2851: 69 p.

Walters, C.J. and Martell, S.J.D. 2004. Fisheries ecology and management. Princeton University Press. 448 pp.

Walters, C.J., Martell, S.J.D., and Korman, J. 2006. A stochastic approach to stock reduction analysis. Can. J. Fish. Aquat. Sci. 63(1): 212–223.

Wiedenmann, J., Free, C.M., and Jensen, O.P. 2019. Evaluating the performance of data-limited methods for setting catch targets through application to data-rich stocks: A case study using Northeast U.S. fish stocks. Fish. Res. 209: 129–142.

APPENDICES

APPENDIX A. FISHERY DATA AND BIOLOGICAL PARAMETERS

Table A1. Historical landings (t) by fleet for the SWNS/BoF stock area (GILL, OTHER, PS) and for the N.B. migrant juvenile stock (WEIR/SO), 1968–2018, based on calendar year with adjustments for misreporting of purse seine landings from 1968–1993.

	SWNS/BoF				NB Mig Juv
Year	GILL	OTHER	PS	Total	WEIR/SO
1968	361	8,108	171,872	180,341	40,174
1969	417	5,413	117,534	123,364	33,748
1970	1,195	7,034	152,516	160,745	22,285
1971	2,753	4,601	96,306	103,660	16,243
1972	2,344	4,058	122,757	129,159	38,326
1973	4,157	7,664	90,781	102,602	25,269
1974	2,214	2,834	144,462	149,510	24,495
1975	4,453	519	110,510	115,481	38,266
1976	6,918	8,223	104,505	119,645	29,294
1977	12,176	755	81,334	94,265	29,952
1978	4,566	9,145	95,210	108,922	39,298
1979	4,615	9,262	39,864	53,740	37,254
1980	14,769	4,420	82,569	101,758	13,544
1981	9,967	3,288	110,576	123,831	21,059
1982	5,400	2,203	106,400	114,003	24,737
1983	6,597	2,257	126,557	135,411	13,371
1984	4,199	3,246	127,181	134,627	8,676
1985	5,474	5,298	135,642	146,414	27,862
1986	3,532	2,238	131,087	136,858	27,886
1987	2,290	7,114	134,034	143,438	27,319
1988	675	9,048	165,156	174,879	39,058
1989	100	4,193	140,687	144,980	44,103
1990	239	5,296	172,907	178,442	40,479
1991	147	1,555	132,356	134,058	24,582
1992	157	3,053	130,555	133,766	32,090
1993	138	3,013	100,254	103,406	31,540
1994	106	2,051	75,652	77,809	22,247
1995	71	3,248	56,338	59,656	18,278
1996	6,402	3,666	47,330	57,398	15,913
1997	6,781	4,209	42,976	53,966	20,618
1998	2,328	3,670	71,444	77,442	21,117
1999	1,693	4,613	73,482	79,789	19,589
2000	882	175	83,159	84,216	17,742
2001	1,932	1,544	66,005	69,481	22,583
2002	386	322	77,511	78,219	12,705
2003	725	15	85,763	86,504	10,861
2004	465	112	72,538	73,115	25,287
2005	1,127	185	44,160	45,471	15,153
2006	1,177	1,503	46,221	48,901	13,886
2007	1,466	11	48,380	49,857	32,063
2008	139	740	52,539	53,417	8,314
2009	287	264	52,693	53,245	4,374
2010	459	155	43,766	44,381	12,414
2011	778	63	47,404	48,244	4,399
2012	562	15	46,256	46,833	541
2013	1,352	233	46,304	47,889	6,243
2014	2,208	210	47,813	50,230	1,987
2015	1,944	2	47,462	49,408	146
2016	1,614	1	48,178	49,793	4,132
2017	797	1	39,199	39,997	2,133
2018	1,182	1	39,665	40,849	12,458

Table A2. Historical landings (t) by fleet for the SWNS/BoF stock area (GILL, OTHER, PS_juv, PS_spa) and for the N.B. migrant juvenile stock (WEIR/SO), 1968–2018, based on calendar year with adjustments for misreporting of purse seine landings from 1968–1993.

	SWNS/BoF					NB Mig Juv
Year	**GILL**	**OTHER**	**PS_juv**	**PS_spa**	**Total**	**WEIR/SO**
1968	361	8,108	50,183	121,689	180,341	40,174
1969	417	5,413	22,751	94,782	123,364	33,748
1970	1,195	7,034	41,062	111,454	160,745	22,285
1971	2,753	4,601	10,727	85,579	103,660	16,243
1972	2,344	4,058	37,301	85,456	129,159	38,326
1973	4,157	7,664	4,337	86,444	102,602	25,269
1974	2,214	2,834	20,583	123,879	149,510	24,495
1975	4,453	519	23,801	86,709	115,481	38,266
1976	6,918	8,223	11,875	92,630	119,645	29,294
1977	12,176	755	3,506	77,828	94,265	29,952
1978	4,566	9,145	11,720	83,491	108,922	39,298
1979	4,615	9,262	10,863	29,000	53,740	37,254
1980	14,769	4,420	19,674	62,895	101,758	13,544
1981	9,967	3,288	65,178	45,398	123,831	21,059
1982	5,400	2,203	82,686	23,715	114,003	24,737
1983	6,597	2,257	81,433	45,124	135,411	13,371
1984	4,199	3,246	86,767	40,414	134,627	8,676
1985	5,474	5,298	99,331	36,311	146,414	27,862
1986	3,532	2,238	96,273	34,814	136,858	27,886
1987	2,290	7,114	108,326	25,708	143,438	27,319
1988	675	9,048	151,481	13,675	174,879	39,058
1989	100	4,193	116,991	23,696	144,980	44,103
1990	239	5,296	135,172	37,735	178,442	40,479
1991	147	1,555	105,868	26,489	134,058	24,582
1992	157	3,053	56,121	74,435	133,766	32,090
1993	138	3,013	57,030	43,224	103,406	31,540
1994	106	2,051	51,986	23,666	77,809	22,247
1995	71	3,248	20,169	36,168	59,656	18,278
1996	6,402	3,666	13,830	33,500	57,398	15,913
1997	6,781	4,209	16,644	26,332	53,966	20,618
1998	2,328	3,670	32,284	39,159	77,442	21,117
1999	1,693	4,613	29,498	43,984	79,789	19,589
2000	882	175	29,840	53,320	84,216	17,742
2001	1,932	1,544	24,820	41,185	69,481	22,583
2002	386	322	28,724	48,787	78,219	12,705
2003	725	15	34,970	50,793	86,504	10,861
2004	465	112	21,880	50,658	73,115	25,287
2005	1,127	185	16,282	27,877	45,471	15,153
2006	1,177	1,503	14,662	31,559	48,901	13,886
2007	1,466	11	15,924	32,456	49,857	32,063
2008	139	740	16,062	36,477	53,417	8,314
2009	287	264	20,116	32,578	53,245	4,374
2010	459	155	20,738	23,029	44,381	12,414
2011	778	63	16,576	30,827	48,244	4,399
2012	562	15	5,448	40,808	46,833	541
2013	1,352	233	19,435	26,869	47,889	6,243
2014	2,208	210	13,850	33,963	50,230	1,987
2015	1,944	2	11,953	35,510	49,408	146
2016	1,614	1	17,807	30,371	49,793	4,132
2017	797	1	13,436	25,763	39,997	2,133
2018	1,182	1	17,720	21,945	40,849	12,458

Table A3. Multipliers for purse seine landings to adjust for misreporting for 1968–1993.

Year	Multiplier
1968	1.01
1969	1.11
1970	1.49
1971	1.26
1972	1.20
1973	1.35
1974	1.30
1975	1.35
1976	1.20
1977	1.15
1978	1.15
1979	1.20
1980	1.45
1981	1.55
1982	1.55
1983	1.63
1984	1.77
1985	1.37
1986	1.88
1987	1.49
1988	1.46
1989	1.61
1990	1.67
1991	1.49
1992	1.38
1993	1.08

Multipliers for 1973–1984 from Mace (1985)
Multipliers for 1985–1991 reported in Stephenson (1993)
Multipliers for 1968–1972 and 1992–1993 estimated as (T_{adj} – Gill – NS weir)/(PS) where T_{adj} = 4WX stock adjusted landings, and Gill, NS weir, and PS, are the nominal landings for 4WX gillnet, Nova Scotia weir, and purse seine gear components as reported in Singh et al. (2020).

Table A4. Acoustic index of Spawning Stock Biomass (t) as the sum of German Bank and Scots Bay surveys, adjusted for turnover as outlined in Melvin et al. (2014).

Year	Total Biomass (t)	COV (%)
1999	452,197	34.0
2000	443,637	24.9
2001	404,176	24.6
2002	455,631	15.3
2003	362,626	19.8
2004	431,807	16.0
2005	239,385	29.3
2006	284,839	18.4
2007	489,628	23.5
2008	236,892	25.5
2009	401,707	18.8
2010	254,625	13.4
2011	380,676	23.2
2012	386,759	10.4
2013	275,009	24.7
2014	380,885	21.3
2015	389,523	12.5
2016	264,147	17.4
2017	307,759	18.0
2018	234,520	27.1

Table A5. Mean autumn larval density (number/m² to bottom) 1972–1998, and 2009. Coefficient Of Variation (COV) adjusted based on the regression of VPA SSB on larval density in Figure A4.

Year	Density (#/m²)	COV (%)	COV adj (%)
1972	9.4	19.1	61.3
1973	6.6	19.7	63.6
1974	49.5	22.0	44.1
1975	11.7	12.8	58.0
1976	13.5	21.5	59.3
1977	6.3	15.9	62.7
1978	4.5	11.1	63.2
1979	7.1	29.6	66.9
1980	26.2	25.6	54.2
1981	2.7	11.1	64.8
1982	10.6	11.3	58.5
1983	13.9	11.5	56.2
1984	12.7	11.0	56.9
1985	40.8	11.3	42.8
1986	18.9	11.1	53.0
1987	27.9	11.5	48.3
1988	100.7	11.4	28.9
1989	54.5	11.2	38.4
1990	27.2	11.4	48.6
1991	48.2	11.4	40.3
1992	57	11.2	37.7
1993	55	11.3	38.2
1994	5.4	13.0	62.8
1995	20.3	22.7	55.8
1996	9.5	16.8	60.6
1997	23.3	11.6	50.7
1998	33.6	11.3	45.7
2009	19.9	21.1	55.4

Table A6. Regression coefficients for the binary logistic regression of maturity (1 = mature; 0 = immature) vs. total length (cm) and age (years) all fish collected in SWNS/BoF by year (1970–2018).

Year	Length b_0	Length b_1	Age b_0	Age b_1
1970	-21.803	0.80163	-3.5451	1.0259
1971	-21.994	0.81280	-4.7747	1.4913
1972	-21.408	0.79251	-5.1490	1.6038
1973	-27.262	1.0724	-8.2715	2.4651
1974	-24.644	0.95193	-12.072	3.4138
1975	-32.400	1.2096	-11.337	3.0653
1976	-26.731	1.0062	-8.9728	2.5563
1977	-30.485	1.1470	-10.548	2.8349
1978	-26.463	0.99777	-9.4820	2.6658
1979	-25.892	1.0024	-11.493	3.5233
1980	-26.446	1.0629	-11.618	3.4564
1981	-26.044	1.0168	-9.2417	2.6286
1982	-23.314	0.91934	-7.6284	2.4289
1983	-24.640	0.94462	-8.4333	2.5493
1984	-25.460	0.97982	-8.3763	2.5411
1985	-22.954	0.91121	-8.9318	2.9729
1986	-24.744	0.96808	-9.2282	2.9366
1987	-31.417	1.2337	-10.646	3.1732
1988	-29.020	1.1133	-9.9650	2.6615
1989	-31.334	1.1841	-8.7828	2.4352
1990	-28.495	1.0749	-9.4001	2.5504
1991	-22.601	0.90070	-9.0342	2.7303
1992	-22.525	0.90416	-9.2688	2.6349
1993	-32.857	1.3161	-15.103	4.5314
1994	-25.975	1.0314	-9.5092	2.7671
1995	-19.457	0.82321	-9.0469	2.9954
1996	-31.804	1.3449	-12.590	4.0398
1997	-31.911	1.3124	-14.188	4.3164
1998	-34.053	1.4325	-13.738	4.1790
1999	-29.313	1.2022	-11.268	3.4596
2000	-30.845	1.2875	-9.4275	3.1525
2001	-25.680	1.0997	-8.7408	3.0914
2002	-27.914	1.1710	-11.073	3.7817
2003	-40.069	1.7333	-14.875	5.2294
2004	-28.993	1.2781	-13.905	5.0082
2005	-37.599	1.6345	-13.654	4.4881
2006	-27.310	1.1427	-12.376	4.0577
2007	-35.023	1.4870	-14.656	5.3526
2008	-25.328	1.0937	-11.705	4.0680
2009	-27.647	1.1946	-10.907	3.6154
2010	-30.268	1.3261	-11.684	3.7900
2011	-31.997	1.4036	-11.123	3.5858
2012	-30.448	1.3260	-7.4981	2.3669
2013	-35.988	1.5548	-10.749	3.6438
2014	-36.479	1.5736	-14.856	5.1180
2015	-37.506	1.6071	-11.866	3.9797
2016	-28.076	1.2291	-10.285	3.4968
2017	-34.810	1.5270	-10.368	3.4678
2018	-39.000	1.7101	-11.711	4.2708

Table A7. Empirical weights (g) at age for all fish sampled in SWNS/BoF by year (1970–2018).

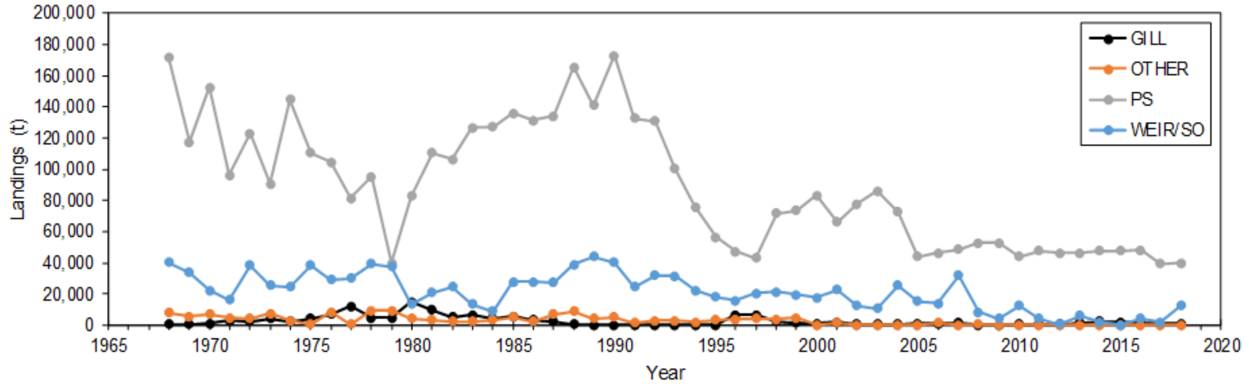
	Age										
Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1970	12.3	24.0	62.6	167	217	267	310	351	389	413	398
1971	8.59	35.1	123	189	228	264	298	343	390	384	429
1972	11.0	41.0	146	203	236	271	300	330	364	395	423
1973	12.1	28.5	95.7	167	231	266	303	342	367	393	409
1974	15.2	38.2	95.5	171	220	265	299	324	344	345	349
1975	12.4	36.2	87.6	173	216	254	289	312	344	364	414
1976	8.46	35.8	100	178	243	263	297	345	352	398	413
1977	13.6	31.3	99.4	177	221	279	297	339	375	415	414
1978	8.71	32.5	115	183	225	264	312	339	370	391	443
1979	8.57	33.7	99.3	179	243	269	301	334	361	381	423
1980	11.1	42.2	102	179	246	293	323	365	369	395	389
1981	10.3	47.3	111	171	236	292	339	357	379	393	388
1982	12.6	44.9	127	192	235	268	300	359	385	378	416
1983	14.3	53.4	118	195	241	274	299	319	362	384	391
1984	13.3	50.0	126	191	237	270	297	331	364	393	421
1985	10.8	44.9	131	204	260	287	323	348	366	389	390
1986	22.9	47.2	115	181	231	262	290	316	332	358	383
1987	18.2	50.2	104	158	215	252	278	308	329	336	341
1988	13.5	33.0	98.1	150	203	260	295	321	336	361	395
1989	13.6	44.9	106	167	209	246	287	313	331	363	361
1990	12.2	36.5	96.5	156	197	235	254	296	322	334	348
1991	12.0	42.9	105	158	201	232	264	292	338	353	342
1992	15.7	26.7	85.3	152	193	226	264	289	316	348	375
1993	15.6	37.5	95.2	151	197	227	261	296	307	336	370
1994	13.9	41.8	97.6	143	180	212	238	258	299	313	324
1995	10.2	40.2	101	158	201	239	275	311	342	348	370
1996	15.8	44.8	95.4	149	200	246	278	316	339	369	384
1997	16.9	41.3	86.3	156	197	247	281	304	319	322	357
1998	18.2	42.0	85.9	132	179	220	263	298	324	339	380
1999	25.6	55.3	101	148	185	218	242	257	298	289	356
2000	19.1	55.7	105	143	184	212	238	244	262	269	287
2001	20.0	51.6	110	161	191	232	254	268	278	304	292
2002	19.8	44.5	111	158	194	234	257	262	280	270	292
2003	17.4	41.2	100	152	186	214	243	262	266	261	270
2004	11.8	38.2	104	149	192	209	232	261	253	271	262
2005	19.2	39.7	87.7	136	169	210	233	236	250	279	263
2006	18.8	52.3	106	156	182	210	233	254	254	269	284
2007	18.2	53.9	116	154	195	215	248	276	262	316	246
2008	18.1	41.4	107	161	194	224	244	264	273	305	323
2009	7.45	42.2	97.4	150	181	216	249	266	275	282	326
2010	15.4	38.1	78.0	127	164	195	222	250	262	264	271
2011	18.3	45.3	85.2	128	160	186	210	251	257	259	272
2012	9.61	44.5	83.6	121	151	177	199	237	248	243	260
2013	16.0	55.0	94.6	126	151	172	195	209	223	257	243
2014	12.7	44.8	103	139	166	186	207	225	243	265	309
2015	21.4	46.9	95.2	140	169	202	216	228	248	266	277
2016	10.6	48.3	93.4	134	160	189	212	217	240	227	273
2017	16.6	44.1	85.9	124	153	175	199	210	210	225	245
2018	23.0	54.7	99.0	128	151	180	199	216	227	225	256

Table A8. Empirical lengths (cm) at age for all fish sampled in SWNS/BoF by year (1970–2018).

	Age										
Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1970	12.3	14.4	20.7	28.2	30.6	32.4	33.8	35.2	36.3	36.5	35.8
1971	10.6	16.4	25.1	29.0	30.9	32.4	33.6	35.0	36.2	37.0	37.7
1972	11.9	17.7	26.5	29.7	31.1	32.7	33.9	34.9	35.9	36.8	37.5
1973	12.4	16.2	23.6	28.1	31.0	32.5	33.8	35.0	35.8	36.7	37.1
1974	13.5	17.2	23.4	28.2	30.7	32.5	33.6	34.7	35.5	35.8	36.0
1975	12.3	17.1	23.0	28.6	30.7	32.4	33.9	35.2	35.9	36.3	36.8
1976	11.0	17.5	24.1	28.8	31.7	32.7	34.3	35.4	36.3	36.8	37.8
1977	12.5	15.8	23.0	28.2	30.4	32.6	33.5	35.0	36.2	37.4	37.6
1978	11.4	16.3	24.3	28.4	30.4	32.0	33.7	34.7	35.8	36.8	37.9
1979	11.4	17.3	24.1	28.5	31.2	32.3	33.4	34.6	35.5	36.4	37.7
1980	12.1	18.1	23.4	27.6	30.7	32.6	33.7	35.1	35.8	36.3	36.5
1981	11.8	18.6	24.4	27.7	30.4	32.6	34.7	35.5	36.3	36.7	37.4
1982	12.5	18.2	24.9	28.5	30.3	31.6	33.0	35.3	36.3	36.4	36.8
1983	12.9	19.3	24.8	28.8	31.0	32.1	33.0	34.4	35.7	37.4	37.0
1984	12.5	18.9	25.6	29.0	30.9	32.3	33.3	34.0	35.2	36.8	37.5
1985	11.8	18.1	25.3	29.1	31.3	32.5	33.6	34.2	34.8	36.5	36.8
1986	14.7	18.7	24.9	28.7	31.0	32.3	33.4	34.5	35.1	35.9	37.1
1987	13.7	19.3	24.4	27.5	30.2	32.0	33.2	34.3	35.3	35.9	36.4
1988	12.6	16.3	23.7	27.0	29.5	31.8	33.3	34.4	35.3	35.9	37.3
1989	12.5	18.5	24.1	27.8	29.7	31.2	32.9	34.0	34.8	35.7	36.1
1990	12.3	17.3	23.8	27.6	29.6	31.2	32.1	33.6	34.6	35.5	35.9
1991	12.0	18.1	24.2	27.3	29.6	31.1	32.3	33.1	34.5	35.3	35.7
1992	13.3	15.1	22.4	26.9	29.1	30.7	32.2	33.4	34.1	35.2	36.1
1993	13.4	17.7	23.6	27.0	29.2	30.7	32.1	33.7	34.3	35.0	35.9
1994	12.4	18.1	23.8	26.7	28.7	30.3	31.4	32.6	33.7	34.6	34.9
1995	10.8	17.9	23.8	27.3	29.4	31.0	32.3	34.1	35.0	35.6	36.3
1996	13.3	18.5	23.4	26.9	29.4	31.2	32.4	33.8	35.0	35.7	36.4
1997	13.6	17.7	22.5	27.0	28.7	31.0	32.6	33.6	35.1	35.6	35.5
1998	12.5	18.3	22.7	26.0	28.5	30.2	32.0	33.4	34.5	35.1	36.5
1999	14.8	18.9	23.1	26.1	27.9	29.5	30.7	31.5	34.1	34.0	34.8
2000	14.2	19.8	24.1	26.5	28.5	29.8	31.0	31.6	32.5	32.2	33.0
2001	14.4	19.2	24.2	27.2	28.8	30.4	31.3	32.1	32.3	33.7	33.2
2002	14.5	18.7	24.4	27.2	29.0	30.6	31.5	31.8	32.3	32.3	32.8
2003	13.6	18.0	23.6	26.9	28.5	30.0	31.3	32.1	32.3	32.2	31.8
2004	11.7	17.7	24.1	26.8	28.8	29.5	30.6	31.7	32.2	32.0	31.8
2005	14.1	17.7	22.8	26.1	27.8	29.6	30.4	30.8	31.7	32.4	32.2
2006	13.8	19.4	23.9	26.7	28.1	29.3	30.3	30.9	31.6	30.9	32.4
2007	14.2	19.4	24.5	26.7	28.7	29.6	30.7	31.8	32.0	32.8	31.4
2008	14.1	17.7	23.9	27.1	28.5	29.7	30.5	31.3	31.7	32.4	33.5
2009	10.4	18.1	23.3	26.6	28.3	29.7	30.8	31.5	31.8	32.1	33.7
2010	13.0	17.9	22.2	25.8	27.6	29.1	30.2	31.4	31.8	31.9	32.1
2011	13.6	18.4	22.4	25.3	27.2	28.5	29.6	31.3	31.8	31.8	32.3
2012	11.4	18.6	22.7	25.2	26.7	28.1	29.1	31.1	31.7	32.2	32.4
2013	13.4	19.6	23.2	25.2	26.7	27.8	29.0	29.6	30.7	32.2	32.7
2014	12.0	18.2	23.9	25.9	27.2	28.1	29.0	30.0	30.5	31.0	32.2
2015	15.0	19.0	23.5	26.4	27.7	28.9	29.5	29.8	30.5	31.6	32.5
2016	11.0	18.7	23.2	25.9	27.4	28.7	29.6	29.8	30.6	30.5	31.8
2017	13.7	18.6	23.1	25.7	27.5	28.7	29.7	30.1	30.2	30.7	31.5
2018	15.2	19.9	23.8	25.8	27.1	28.7	29.6	30.2	30.7	30.8	31.6

Table A9. Classification of fleets by fishing ground and gear.

<table>
<tr><th>Fleet Structure</th><th>Fleet(s)</th><th>Fishing Grounds</th><th>Gear(s)</th></tr>
<tr><td rowspan="5">A</td><td>WEIR/SO</td><td>Grand Manan, Grand Manan Banks, NB Coastal</td><td>Weir, shutoff</td></tr>
<tr><td>PS</td><td>All</td><td>Purse seine</td></tr>
<tr><td>GILL</td><td>All</td><td>Gill net</td></tr>
<tr><td rowspan="2">OTHER</td><td>Gannet Dry Ledge, Long Island, Scots Bay</td><td>Weir</td></tr>
<tr><td>All</td><td>Gears other than weir, shutoff, purse seine, gill net</td></tr>
<tr><td rowspan="3">B</td><td>WEIR/SO, GILL, OTHER</td><td colspan="2">As described for Fleet Structure A</td></tr>
<tr><td>PS_juv</td><td>Grand Manan, Grand Manan Banks, Long Island, NB Coastal, Trinity, Yankee Bank</td><td>Purse seine</td></tr>
<tr><td>PS_spa</td><td>4W winter, Browns Bank, Gannet Dry Ledge, Lurcher, Scots Bay, Seal Island, SW Grounds</td><td>Purse seine</td></tr>
</table>



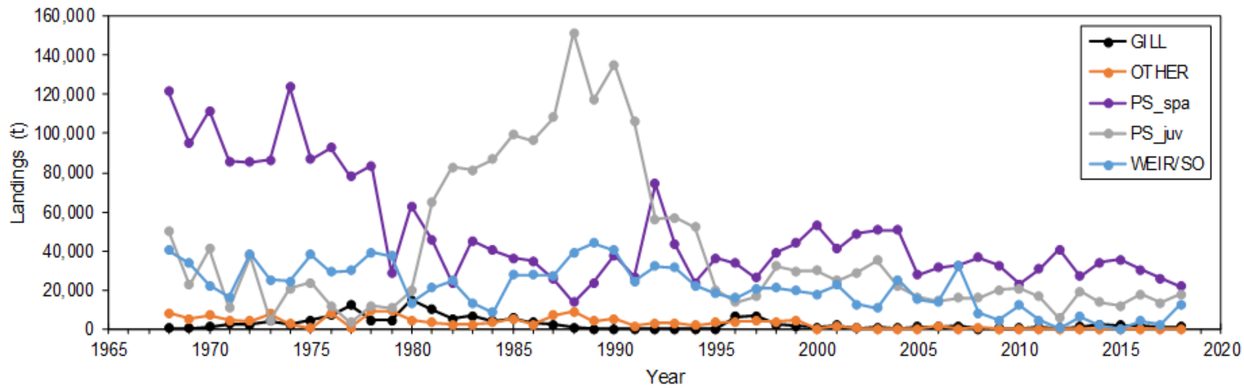


Figure A1. Landings by fleet 1968–2018 based on data in Table A1 and A2.

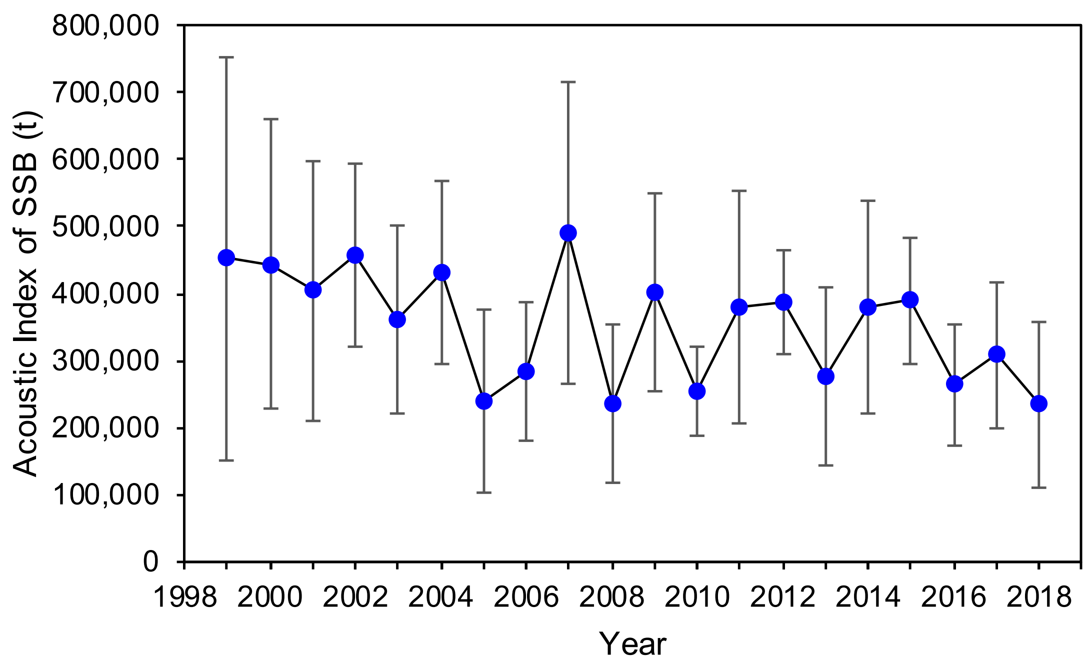


Figure A2. Acoustic index of Spawning Stock Biomass (t) as the sum of German Bank and Scots Bay surveys, adjusted for turnover as outlined in Melvin et al. (2014). Error bars represent 95% confidence intervals.

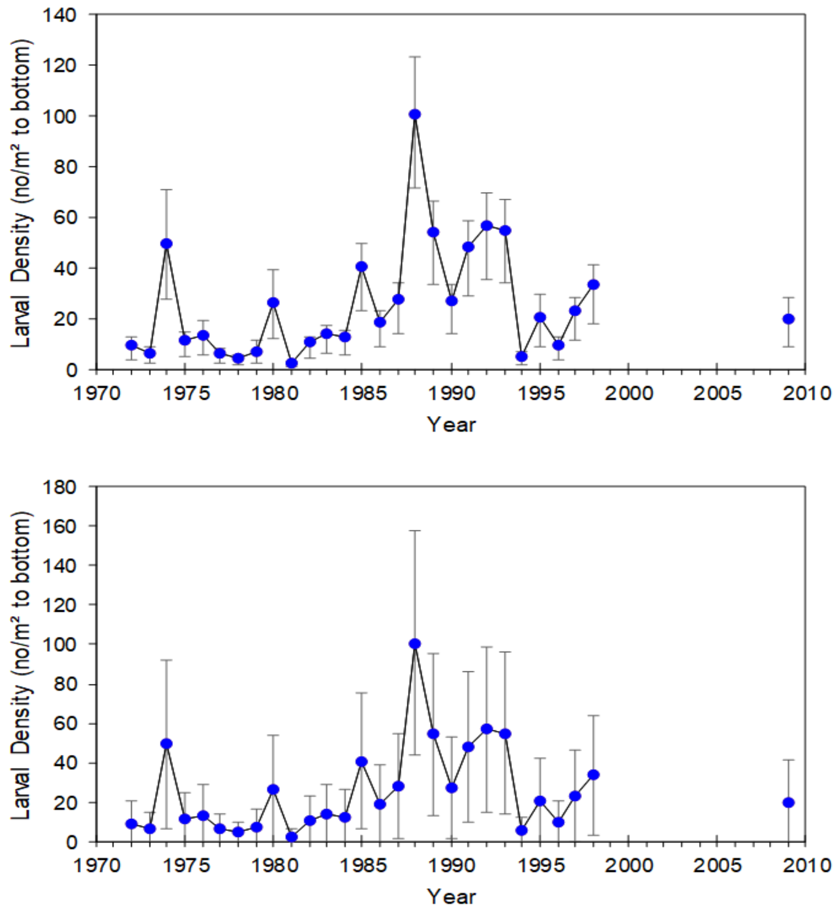


Figure A3. Mean autumn larval density (number/m^2 to bottom) 1972–1998, and 2009. Error bars represent 95% confidence intervals for the mean density (top) and 95% confidence intervals for the mean density after adding the variability in the relationship between SSB from the VPA and Larval Density (Figure A4).

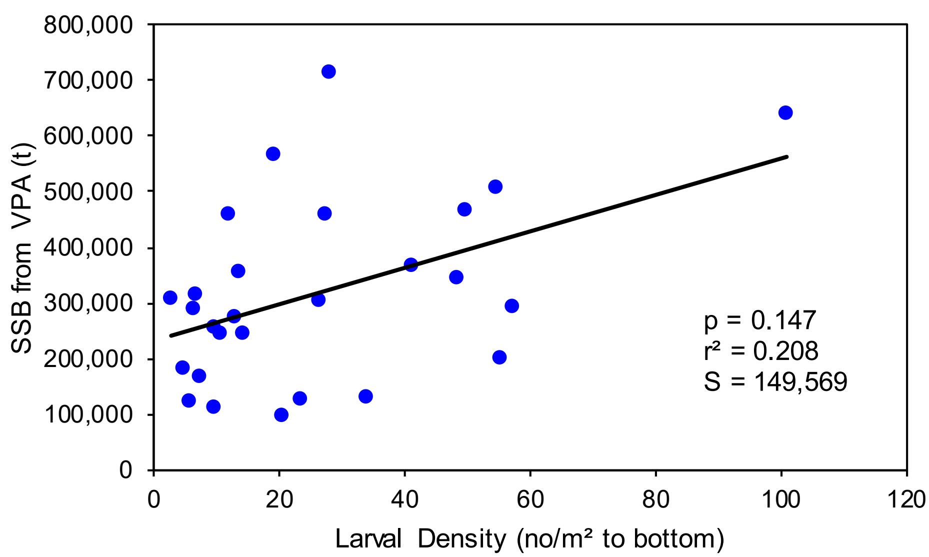


Figure A4. Scatterplot and linear regression of SSB from the VPA vs. mean autumn larval density 1972–1998, and 2009. S = square root of the mean squared error.

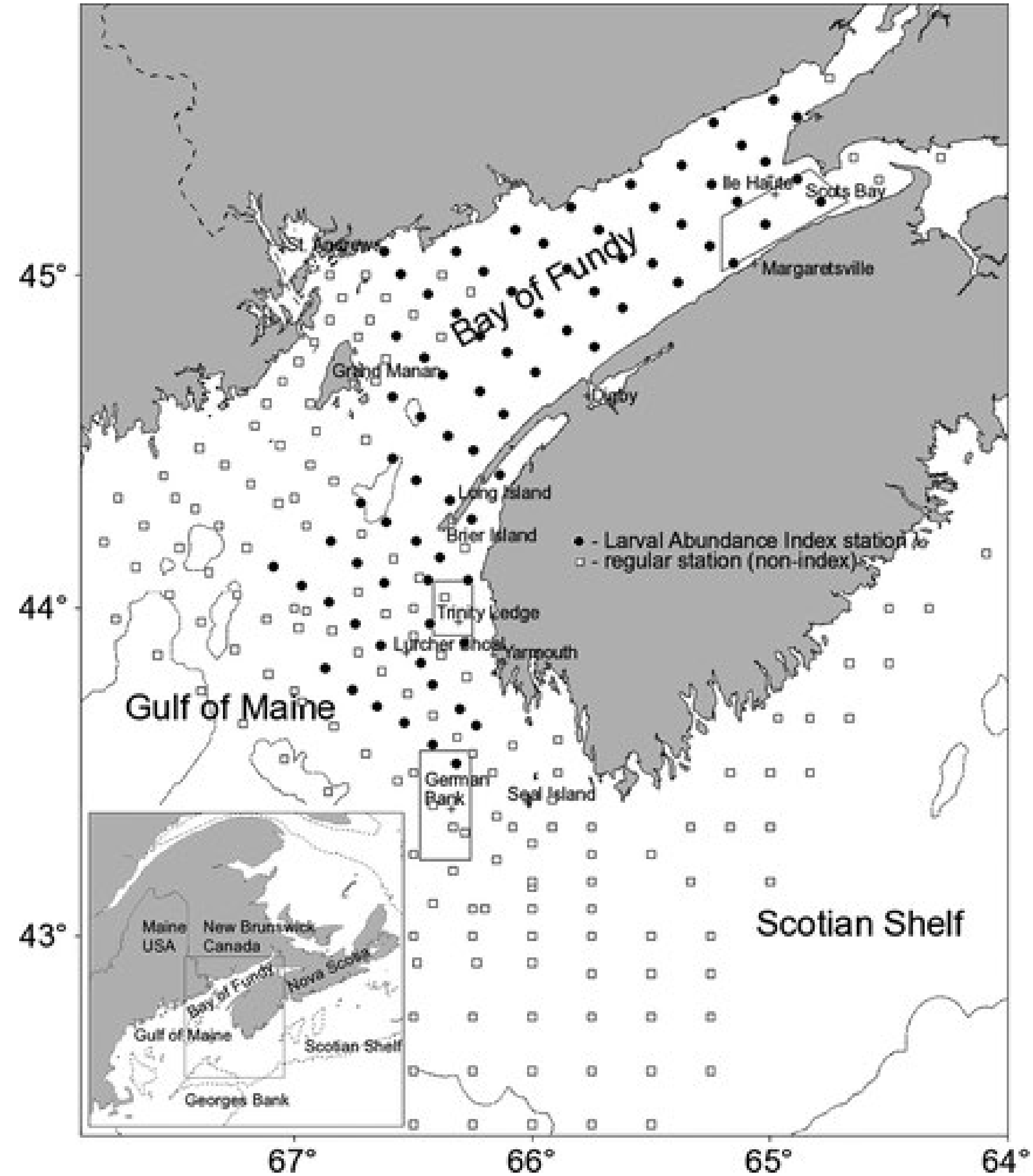


Figure A5. Bay of Fundy larval Herring survey stations (● n = 79 standard survey stations; □ other stations surveyed between 1972 and 1998). Figure from Stephenson et al. (2015).

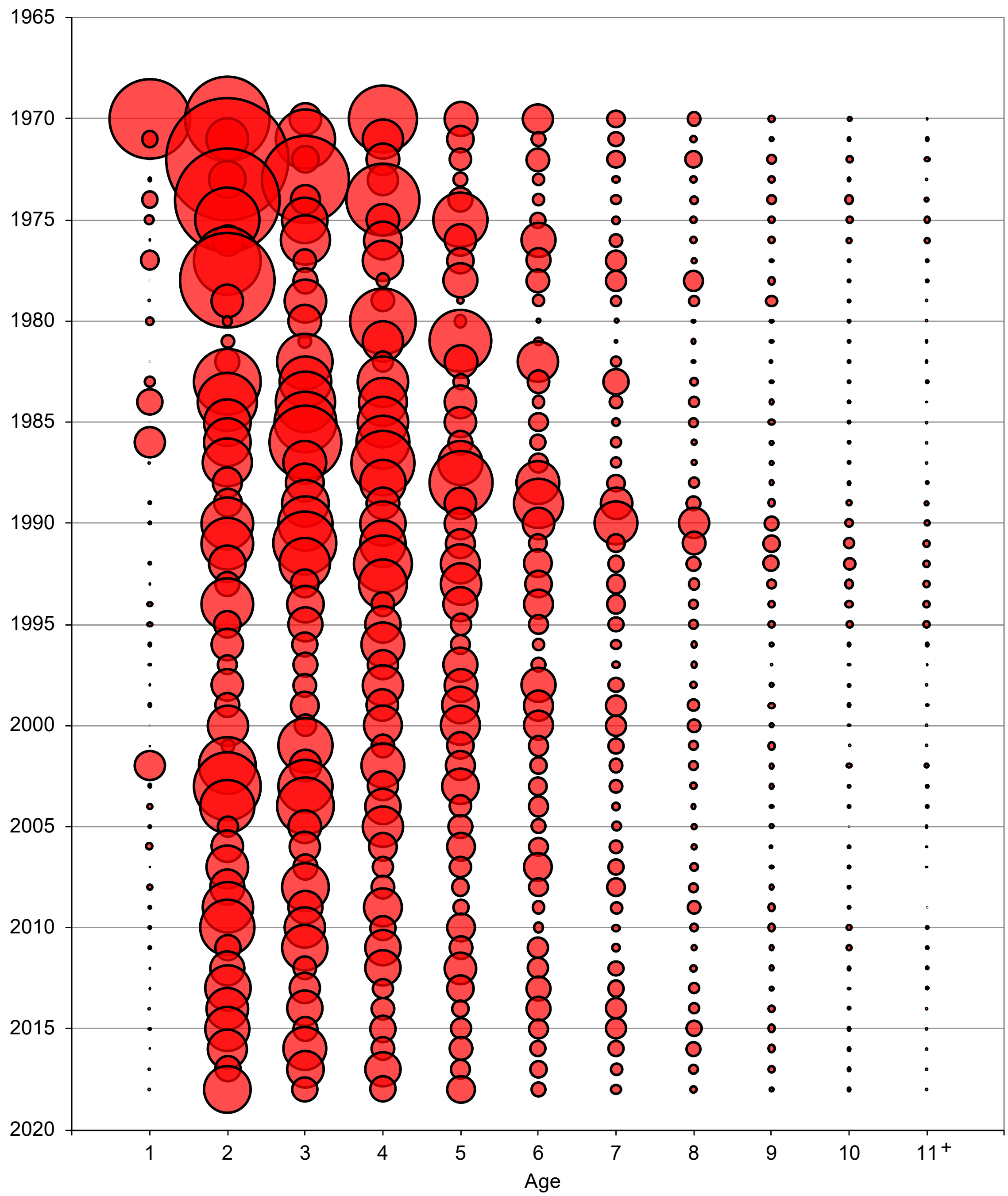


Figure A6. Relative catch-at-age (numbers) for the SWNS/BoF purse seine (PS) fleet, 1970–2018.

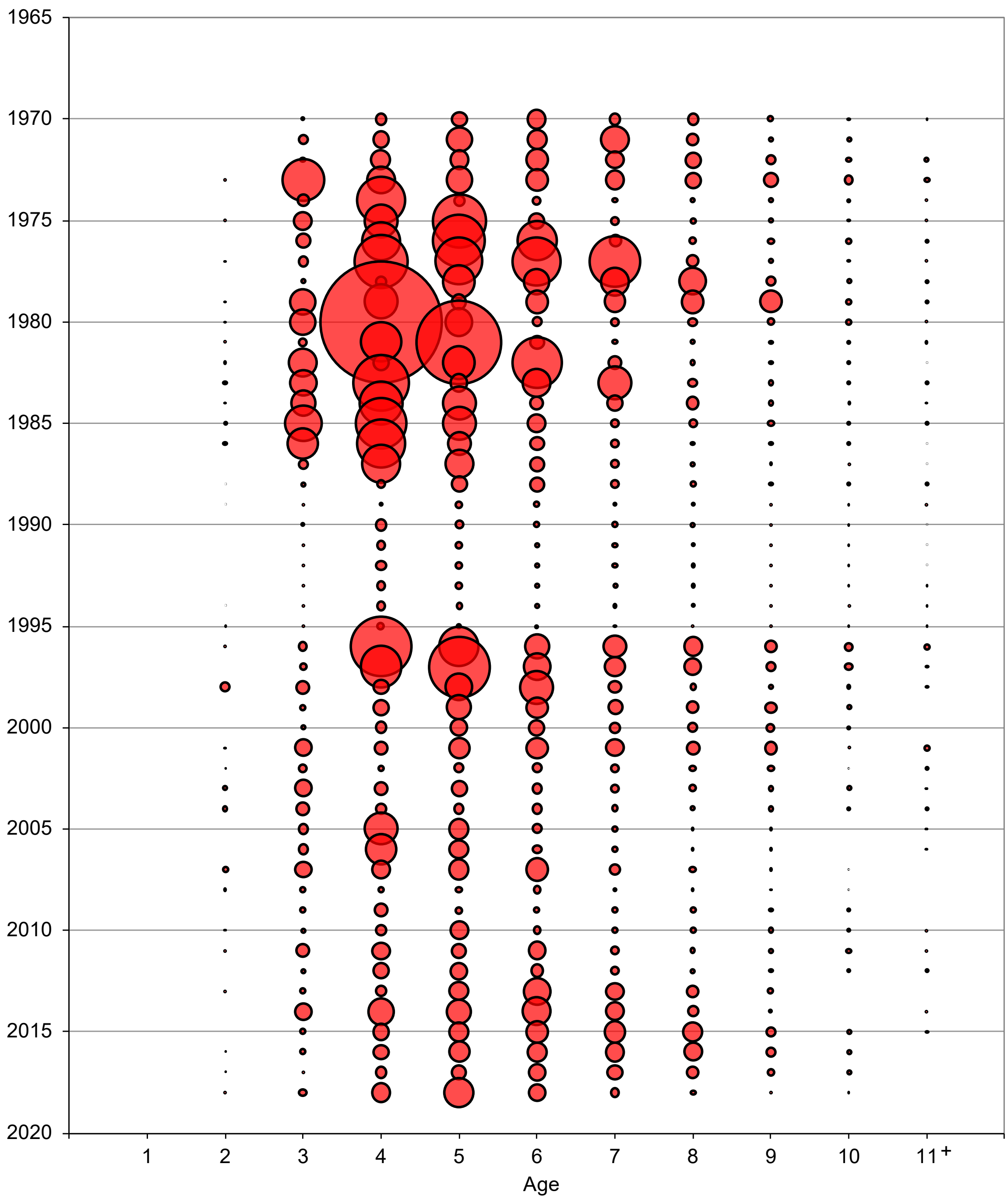


Figure A7. Relative catch-at-age (numbers) for the SWNS/BoF gill net (GILL) fleet, 1970–2018.

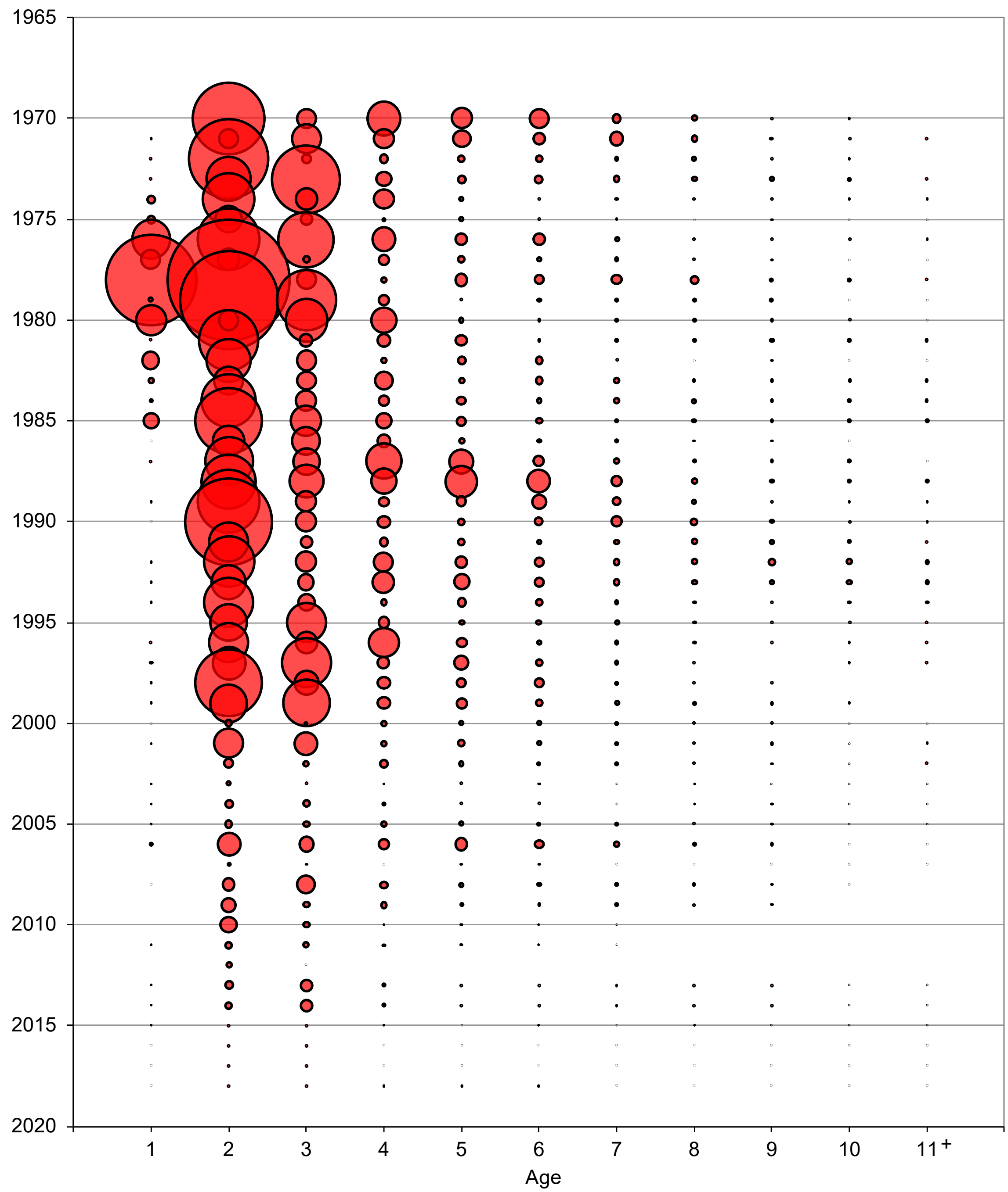


Figure A8. Relative catch-at-age (numbers) for the SWNS/BoF “OTHER” fleet, 1970–2018.

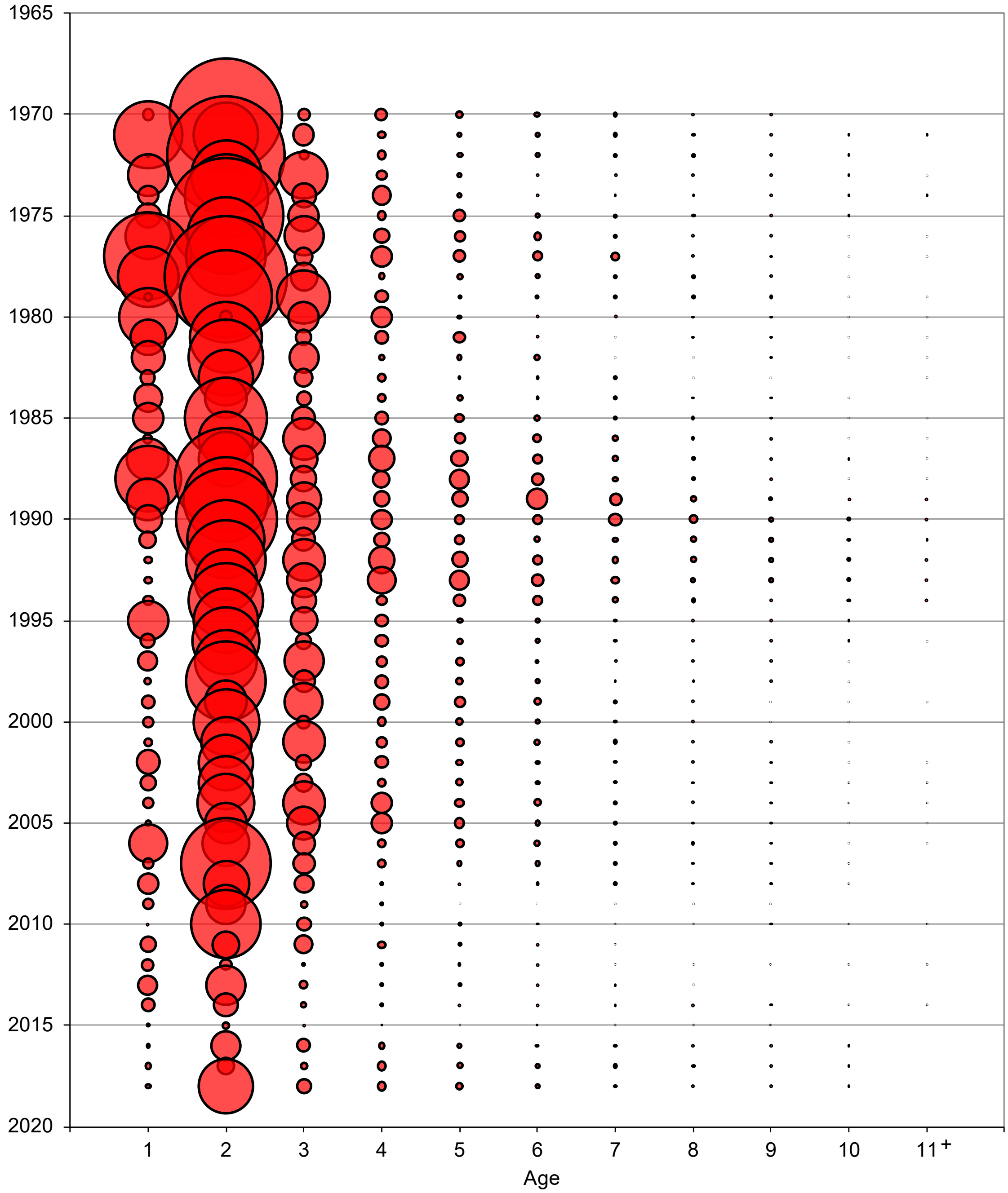


Figure A9. Relative catch-at-age (numbers) for the SWNS/BoF weir and shut off (WEIR/SO) fleet (New Brunswick catches only), 1970–2018.

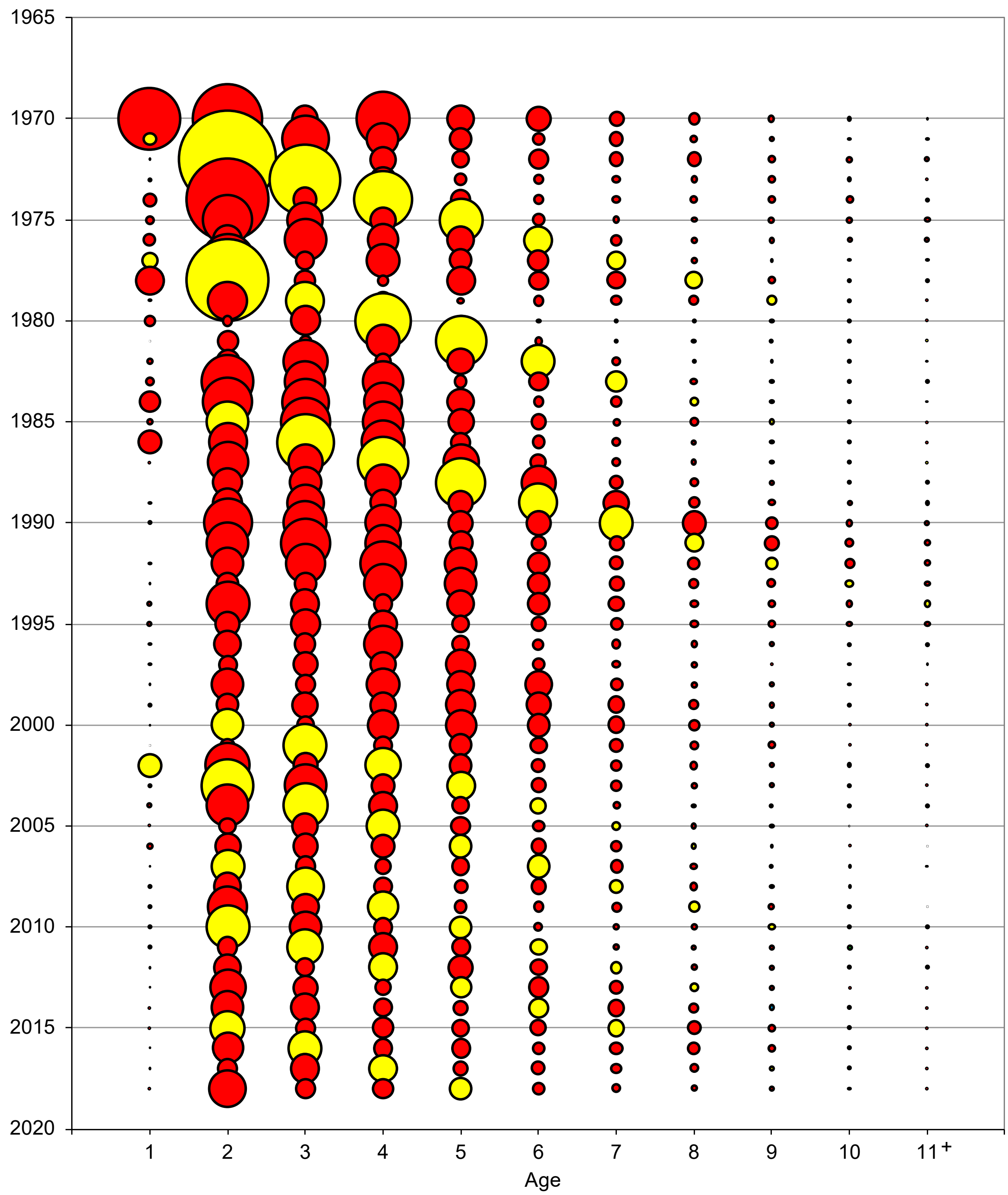


Figure A10a. Relative catch-at-age (numbers) for the SWNS/BoF purse seine (PS), gill net (GILL), and other (OTHER) fleets (i.e., total quota catches). Some stronger year classes shown in yellow (1970, 1976, 1983, 1998, 2001, 2005, 2008, and 2013). This figure was used for comparison back to historical catch-at-age.

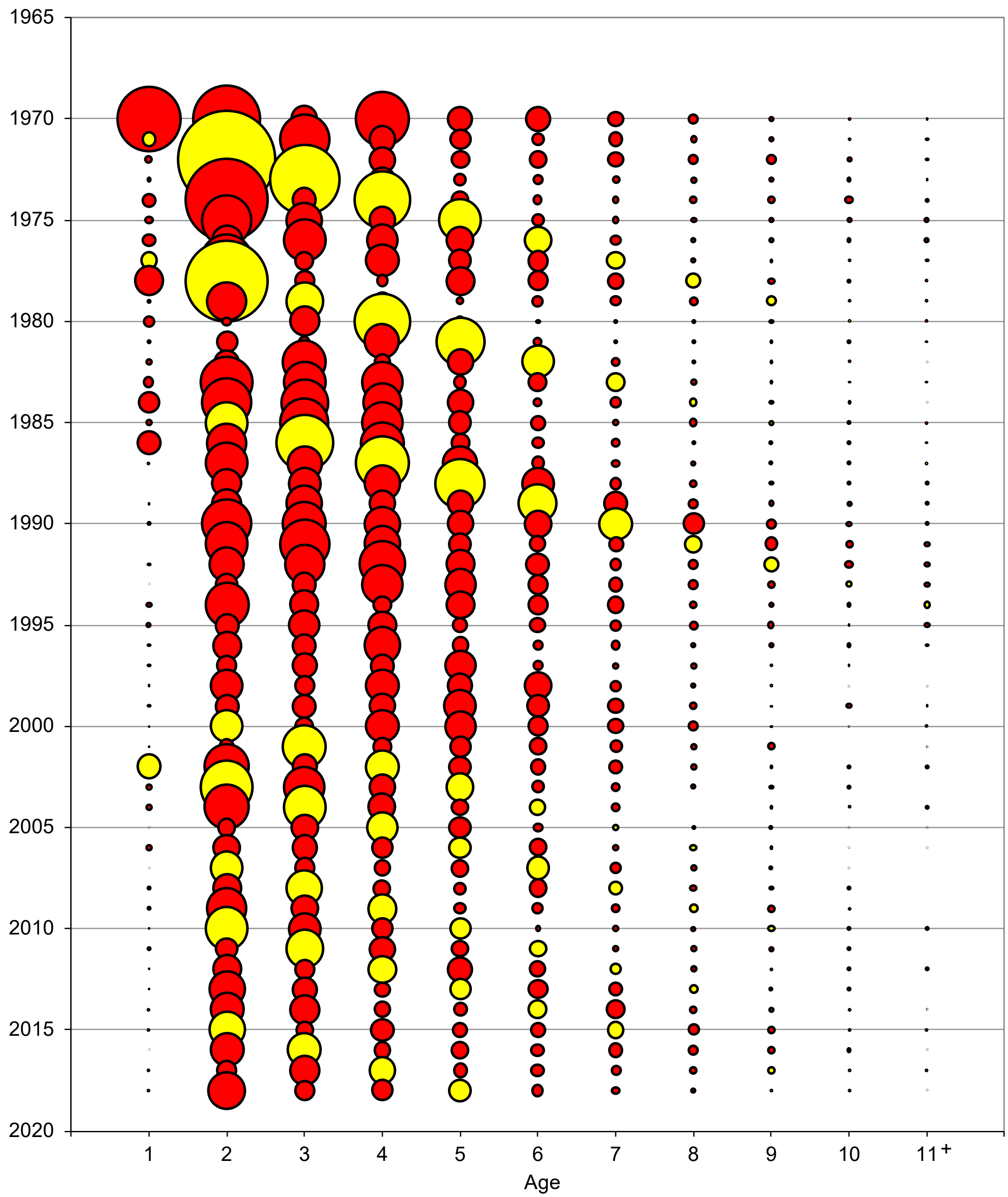


Figure A10b. Relative catch-at-age (numbers) for the SWNS/BoF purse seine (PS), gill net (GILL), and other (OTHER) fleets (i.e., total quota catches) using length bins of 0.5 cm in the age-length keys. Some stronger year classes shown in yellow (1970, 1976, 1983, 1998, 2001, 2005, 2008, and 2013). Contrast with Figure A10a that uses 1 cm bins.

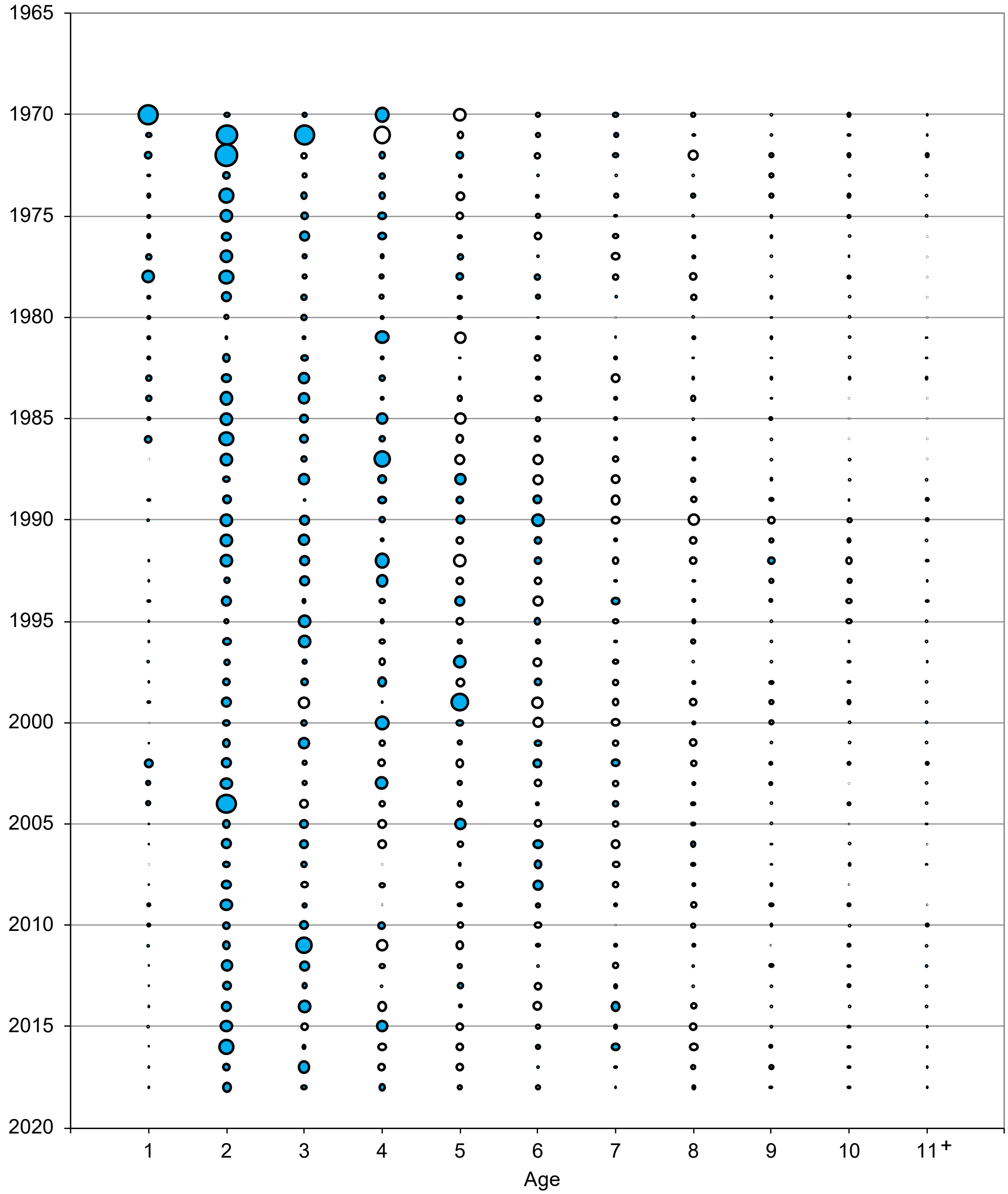


Figure A10c. Differences in catch-at-age (calculated as numbers-at-age using 0.5 cm length bins minus numbers-at-age using 1 cm length bins; blue = positive difference and white = negative difference) for the SWNS/BoF purse seine (PS), gill net (GILL), and other (OTHER) fleets (i.e., total quota catches). Scale is the same as the relative catch-at-age in Figure A10b.

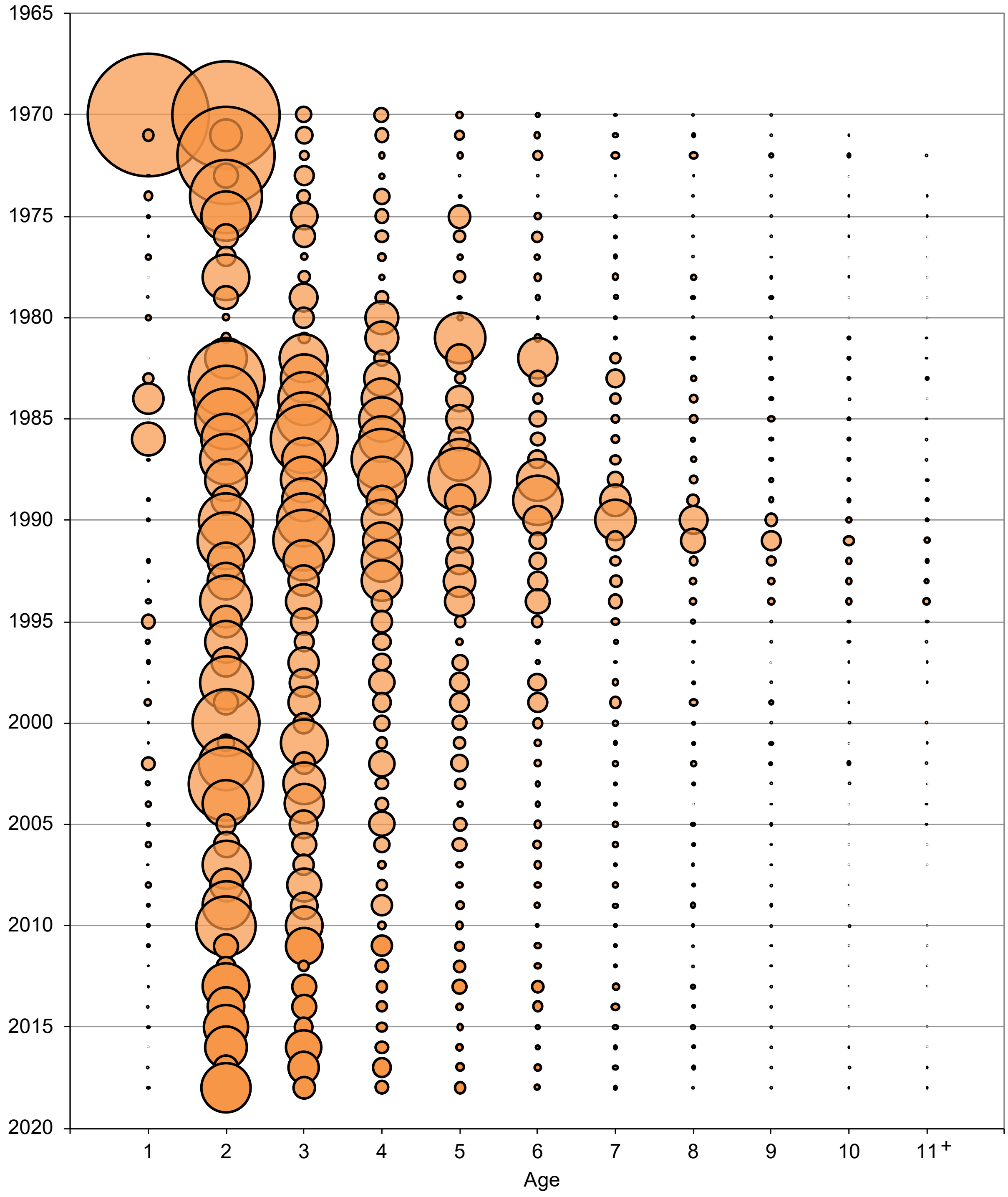


Figure A11. Relative Catch-At-Age (numbers) for the SWNS/BoF purse seine (PS_juv) fleet, 1970–2018.

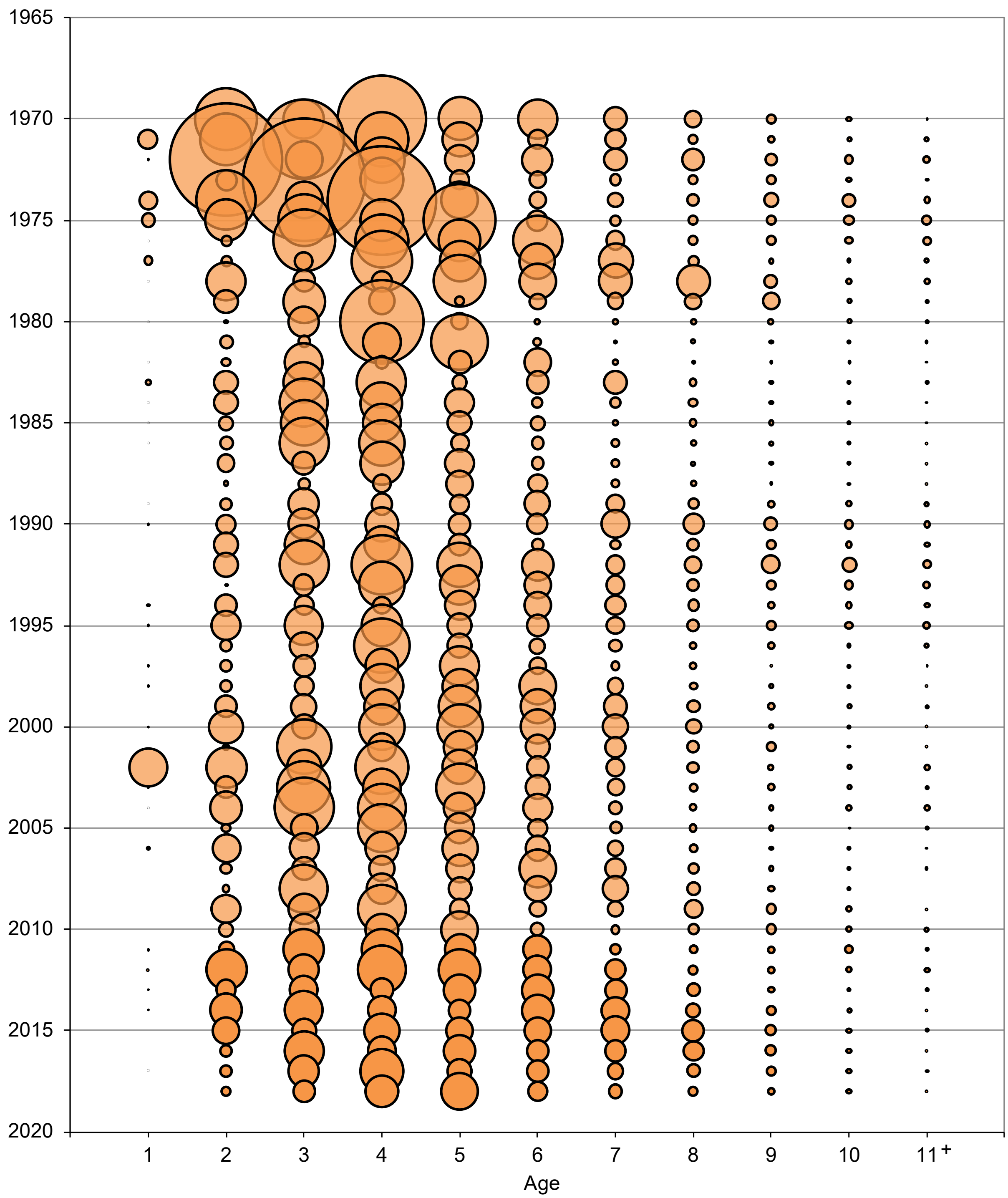


Figure A12. Relative Catch-At-Age (numbers) for the SWNS/BoF purse seine (PS_spa) fleet, 1970–2018.

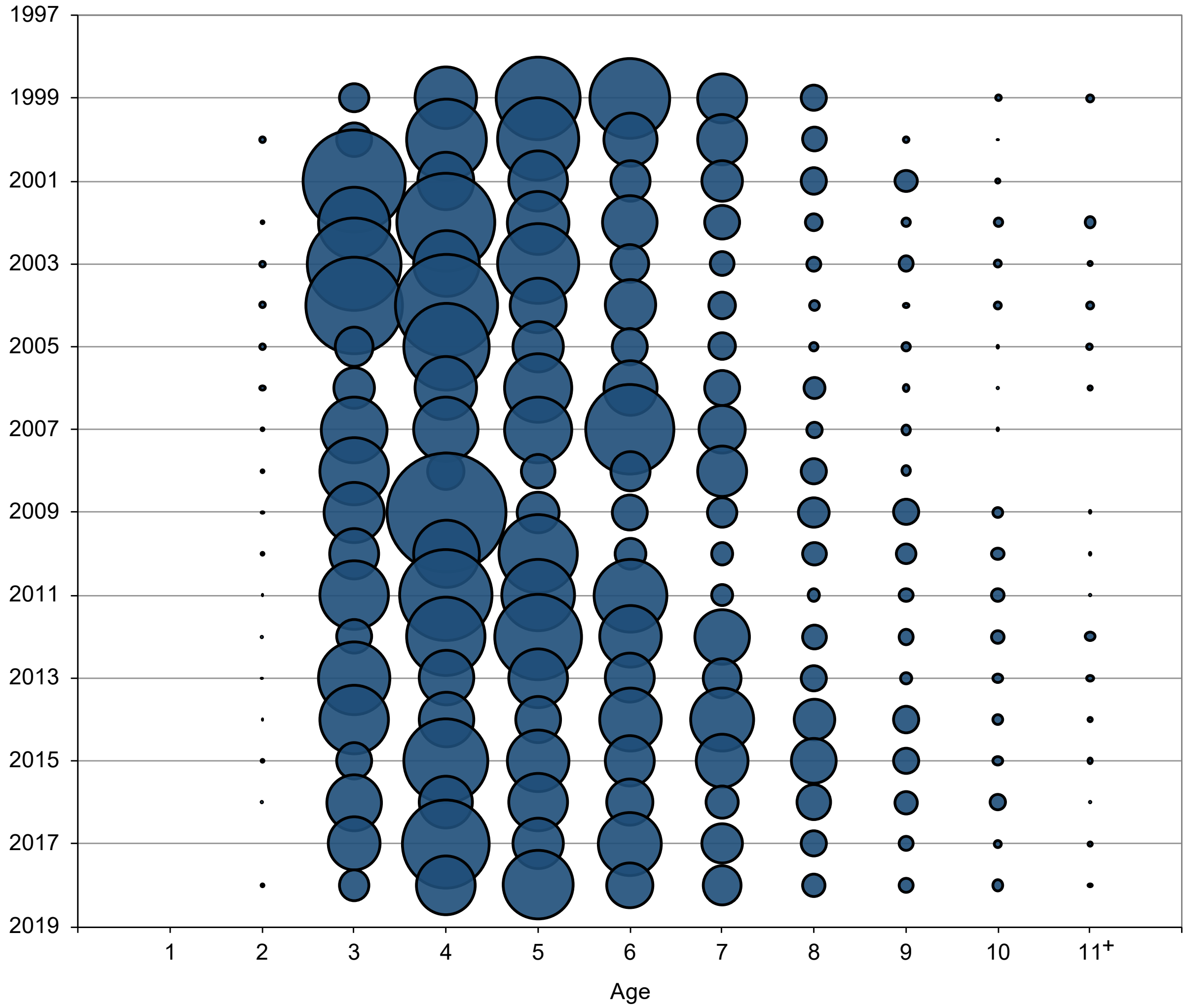


Figure A13. Relative numbers of sexually mature fish at age estimated from the acoustic surveys (sum of German Bank and Scots Bay) for the overall SWNS/BoF stock component, 1999–2018. Notes: Numbers of fish determined using target strength for a fish by length bin and total backscatter. Age 2 and 3 fish: removed per survey using proportion mature at age 2 or 3 from the detailed samples used for the individual survey. If age 2 fish were present, but no detailed samples for maturity for a particular survey, then a percent mature of 48% was used (overall % mature of all age 2 fish from all detailed samples used for the surveys).

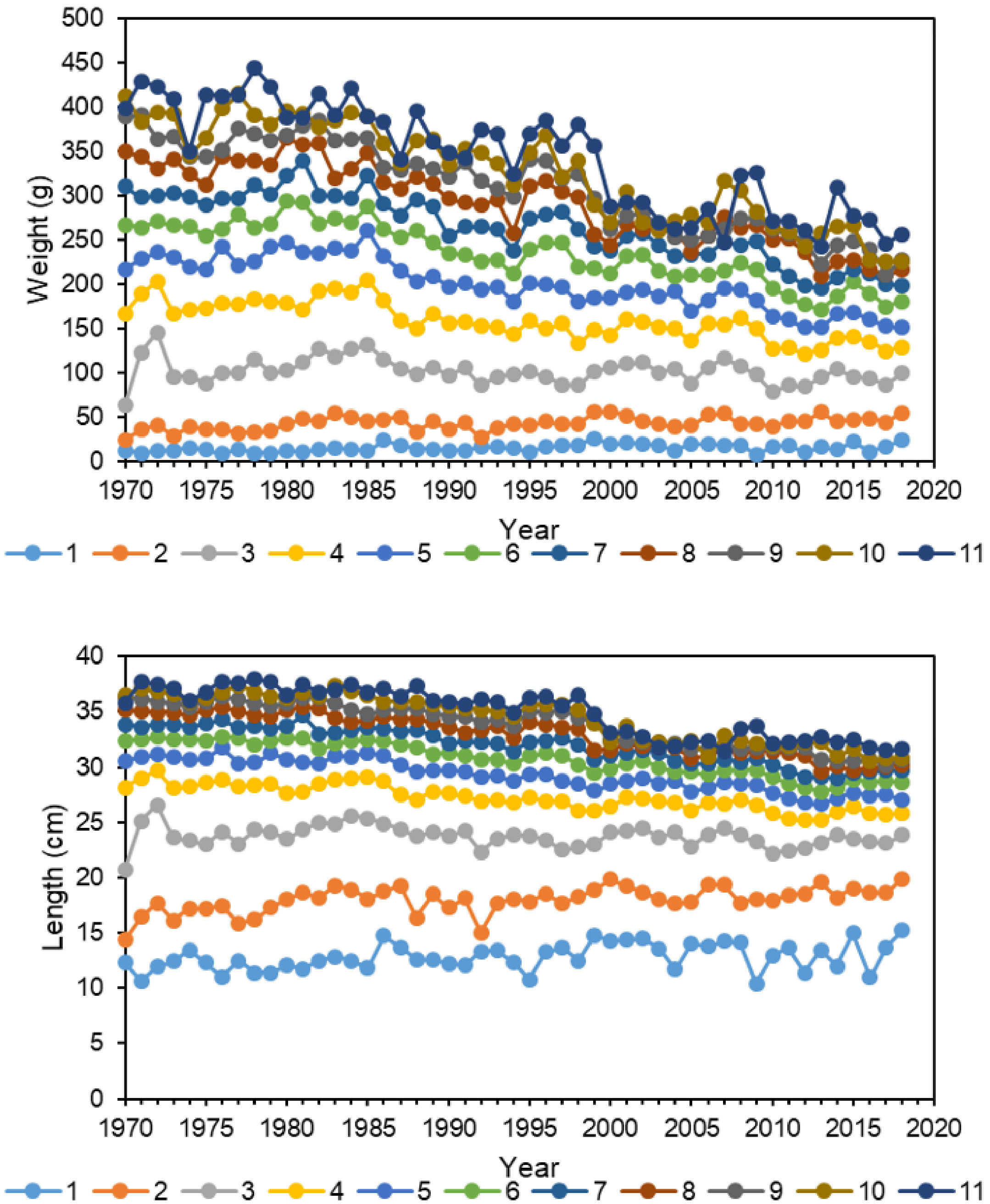
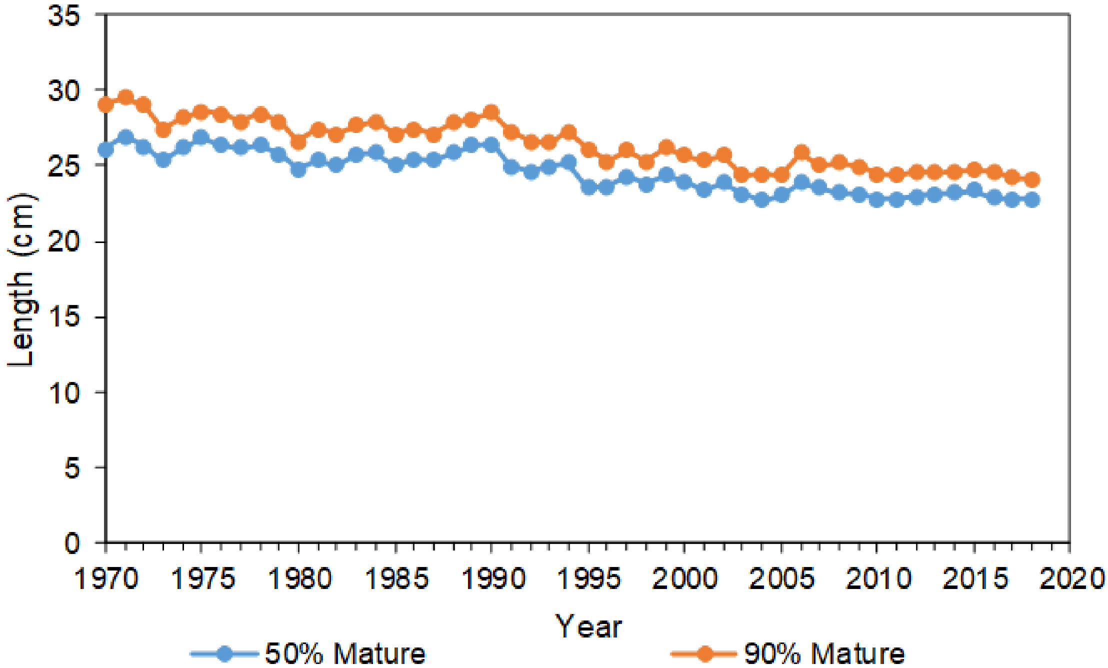


Figure A14. Empirical weight (g) and length (cm) at age for all fish sampled in SWNS/BoF by year (1970–2018).



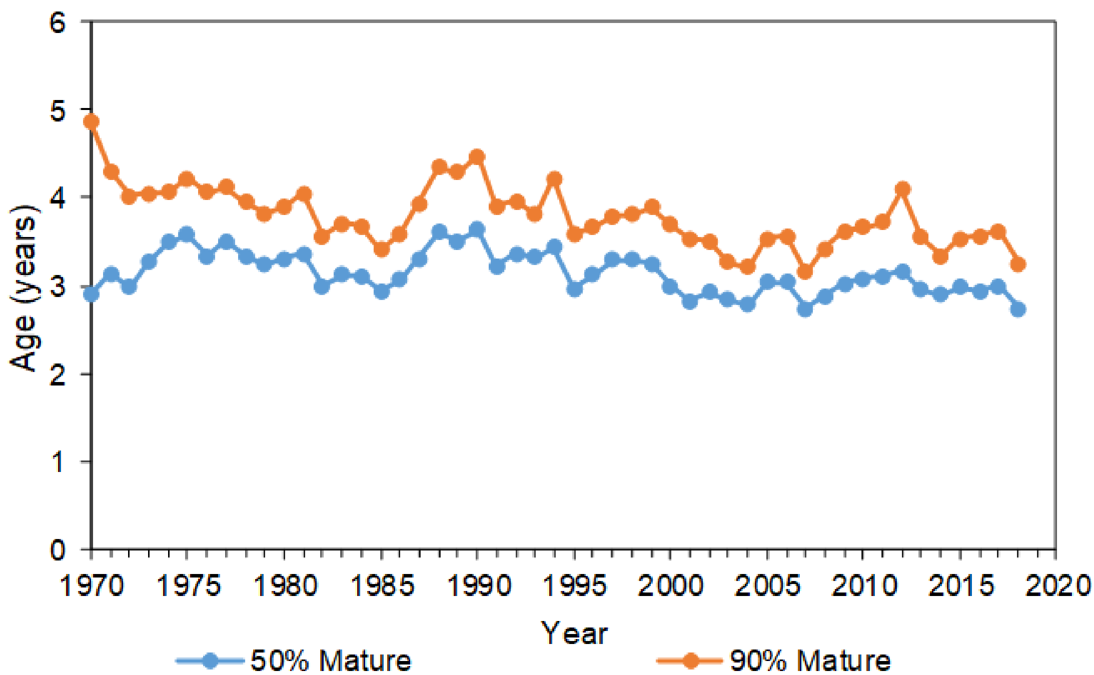


Figure A15. Total length (cm) and age (years) at 50% and 90% maturity for all fish collected in SWNS/BoF by year (1970–2018).

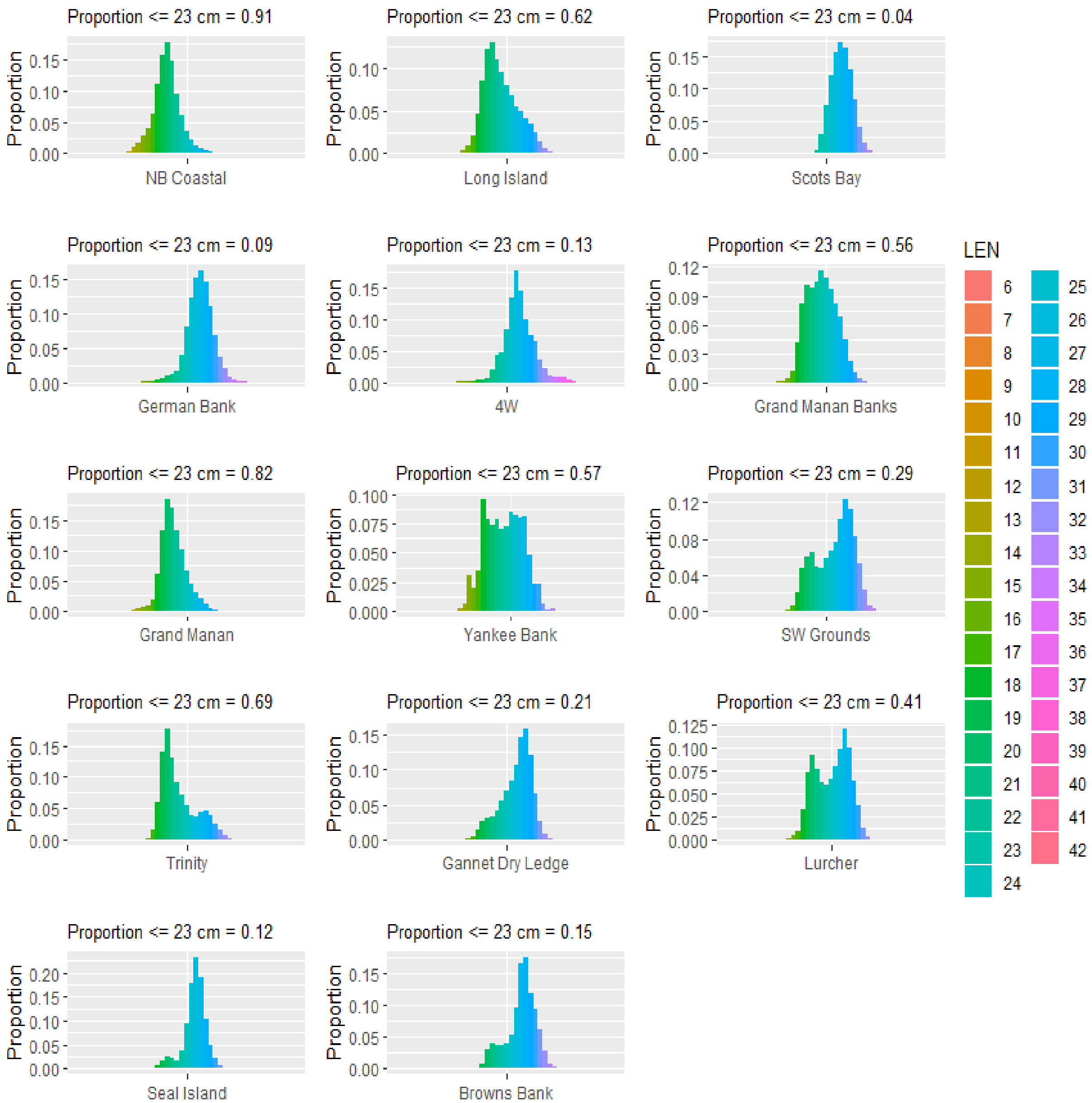


Figure A16. Length frequency distributions for purse seine catches by fishing ground for 1995–2018.

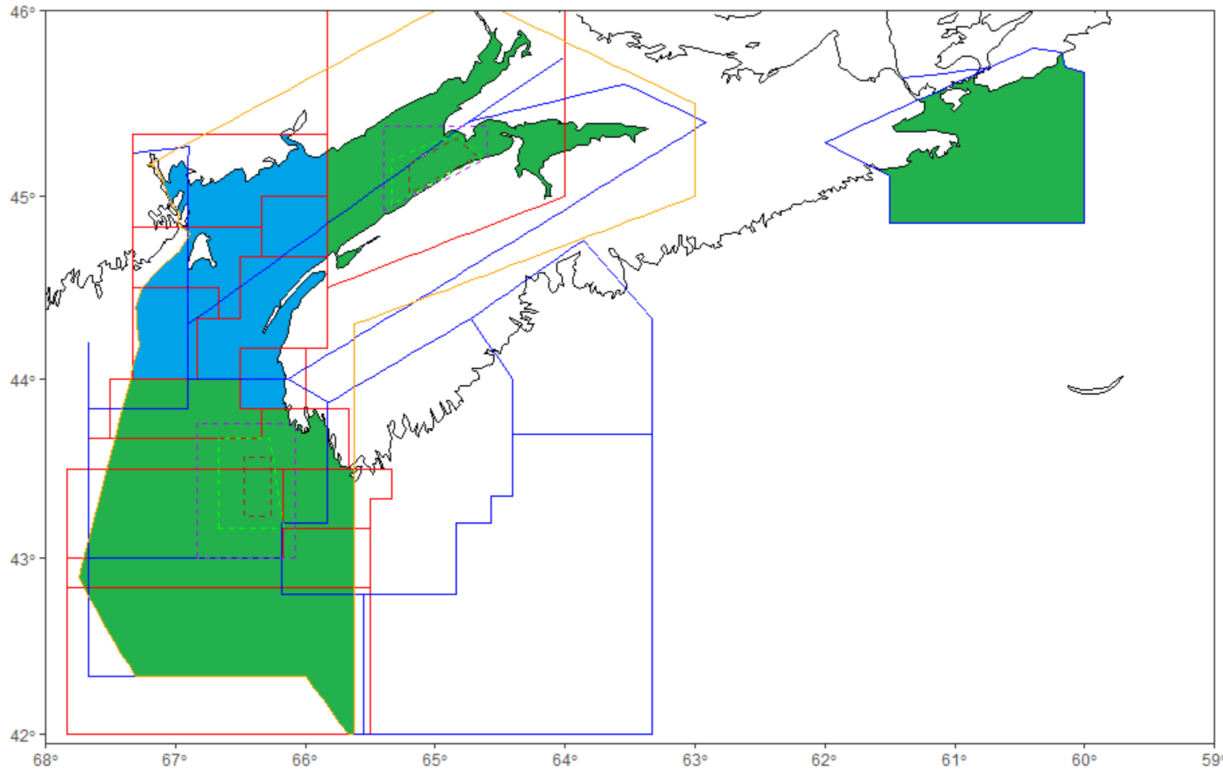


Figure A17. Spatial separation of fishing grounds based on the proportion of fish ≤ 23 cm from purse seine catches (see Figure A16). Blue = spatial area for fleet PS_juv. Green = spatial area for fleet PS_spa.

APPENDIX B. EQUATIONS OF THE POPULATION AND EXPLOITATION DYNAMICS MODEL INCLUDING LIKELIHOOD FUNCTIONS AND NUMERICAL ALGORITHMS FOR CONDITIONING OPERATING MODELS

Mathematical Description Of The SRA Model

Selectivity and Mortality

Selectivity v is length-based and modeled as a double-exponential function (using base 2). For fleet f with flat-topped selectivity, two parameters are used, the length of 5% selectivity (L_f^5) and the length of full selectivity L_f^{FS}. For dome selectivity, a third parameter, the selectivity at L_∞, $V_f^{L\infty}$ is also used. Length-based selectivity is converted to age-based selectivity in the age-structured model as:

$$v_{y,a,f} = \begin{cases} 2^{-[(L_{y,a}-L_f^{\text{FS}})/(\sigma_f^{\text{asc}})]^2} & \text{if } L_{y,a} < L_f^{\text{FS}} \\ 1 & \text{if logistic and } L_{y,a} \geq L_f^{\text{FS}}, \\ 2^{-[(L_{y,a}-L_f^{\text{FS}})/(\sigma_f^{\text{des}})]^2} & \text{if dome shaped and } L_{y,a} \geq L_f^{\text{FS}} \end{cases} \tag{B1}$$

where $L_{y,a}$ is the mean length-at-age in year y, and $\sigma_f^{\text{asc}} = (L_f^5 - L_f^{\text{FS}})/\sqrt{-\log_2(0.05)}$ and $\sigma_f^{\text{des}} = (L_\infty - L_f^{\text{FS}})/\sqrt{-\log_2(V^{L\infty})}$ control the shape of the ascending and descending limbs, respectively, of the selectivity function. In this parameterization, length-based selectivity is constant over time. The corresponding age-based selectivity is constant over time if growth is not time-varying.

Total mortality Z in year y and for age a is the sum of fishing mortality F from all fleets and natural mortality M:

$$Z_{y,a} = M_{y,a} + \Sigma_f v_{y,a,f} F_{y,f} \tag{B2}$$

Initial Population Distribution

The population age distribution in the first year of the model $y = 1$ is in equilibrium where

$$N_{y=1,a} = \begin{cases} R^{\text{eq}} \exp\left(-\Sigma_{i=1}^{a-1} Z_i^{\text{eq}}\right) & a = 1, \dots, A-1 \\ \dfrac{R^{\text{eq}} \exp\left(-\Sigma_{i=1}^{a-1} Z_i^{\text{eq}}\right)}{1 - \exp\left(-Z_A^{\text{eq}}\right)} & a = A, \end{cases} \tag{B3}$$

where the R^{eq} is the equilibrium recruitment and $Z_a^{\text{eq}} = M_{1,a} + \Sigma_f v_{1,a,f} F_f^{\text{eq}}$ is the equilibrium total mortality rate. Unfished conditions are modeled by setting $F_f^{\text{eq}} = 0$. To estimate F_f^{eq}, the corresponding equilibrium catch in weight $\tilde{C}_f^{\text{eq}}$ prior to the first year of the model should be provided. In the equilibrium yield curve, F_f^{eq} would be the fishing mortality corresponding to fishing at F_f^{eq}. Once Z_a^{eq} is obtained, then the equilibrium recruitment using a Beverton-Holt stock recruitment relationship is calculated as:

$$R^{\text{eq}} = \frac{\alpha^{\text{BH}} \phi^{\text{eq}} - 1}{\beta^{\text{BH}} \phi^{\text{eq}}}, \tag{B4}$$

where ϕ^{eq} is the spawners-per-recruit at the initial equilibrium mortality rate. From steepness h, $\alpha^{\text{BH}} = \frac{4h}{(1-h)\phi_0}$, $\beta^{\text{BH}} = \frac{5h-1}{(1-h)B_0^S}$, where ϕ_0 and B_0^S are unfished spawners-per-recruit and unfished spawning biomass, respectively.

For the Ricker stock-recruit relationship, the equilibrium recruitment is

$$R^{\text{eq}} = \frac{\log(\alpha^{\text{R}}\phi^{\text{eq}})}{\beta^{\text{R}}\phi^{\text{eq}}}, \tag{B5}$$

where $\alpha^{\text{R}} = \frac{(5h)^{1.25}}{\phi_0}$ and $\beta^{\text{R}} = \frac{\log(5h)}{B_0^S}$.

Dynamics Equations

After setting the equilibrium population age distribution in the first year of the model, the population abundance $N_{y,a}$ in subsequent years is:

$$N_{y,a} = \begin{cases} R_y & a = 1 \\ N_{y-1,a-1}\exp(-Z_{y-1,a-1}) & a = 2, \ldots, A-1, \\ N_{y-1,a-1}\exp(-Z_{y-1,a-1}) + N_{y-1,a}\exp(-Z_{y-1,a}) & a = A \end{cases} \tag{B6}$$

where R_y is the recruitment and A is the maximum-age as the plus-group. Recruitment is modelled as:

$$R_y = \frac{\alpha^{\text{BH}} B_{y-1}^S}{1 + \beta^{\text{BH}} B_{y-1}^S}\exp(\delta_y - 0.5\tau^2), \tag{B7}$$

where δ_y are recruitment deviates and τ is the standard deviation of the deviates.

The spawning biomass is B_y^S is:

$$B_y^S = \sum_a w_{y,a}\, m_{y,a} N_{y,a}, \tag{B8}$$

where $m_{y,a}$ and $w_{y,a}$ are the maturity at age and weight at age, respectively.

The catch (in numbers) C^N at age for fleet f is:

$$C_{y,a,f}^N = \frac{v_{y,a,f}F_{y,f}}{Z_{y,a}} N_{y,a}(1 - \exp[-Z_{y,a}]) \tag{B9}$$

If the model is conditioned on catch with zero equilibrium catches prior to the first year of the model, then $F_{y,f}$ are solved such that the annual predicted catches (in weight) match the observed values. Otherwise, $F_{y,f}$ are estimated parameters.

The catch-at-length is calculated assuming a normally distributed length-at-age $P(\ell, a)$, where:

$$C_{y,\ell,f}^N = \sum_a C_{y,a,f}^N\, P(\ell|a) \tag{B10}$$

and

$$P(\ell|a) = \begin{cases} \phi(L'_{\ell+1}) & \ell = 1 \\ \phi(L'_{\ell+1}) - \phi(L'_{\ell}) & \ell = 2, \ldots, L-1, \\ 1 - \phi(L'_{\ell}) & \ell = L \end{cases} \tag{B11}$$

with L'_ℓ as the length at the lower boundary of length bin ℓ and $\phi(L'_\ell)$ as the cumulative distribution function of a normal variable with mean $\tilde{L}_{y,a}$ (the expected mean length at age a) and standard deviation $\tilde{L}_{y,a} \times CV^L$ (CV^L is the coefficient of variation in mean length at age).

The catch in weight $\tilde{C}$ is:

$$\tilde{C}_{y,f} = \sum_a C^N_{y,a,f}\, w_{y,a}. \tag{B12}$$

The proportion of the catch-at-age is:

$$p_{y,a,f} = \frac{C^N_{y,a,f}}{\sum_a C^N_{y,a,f}} \tag{B13}$$

The proportion of the catch-at-length is:

$$p_{y,\ell,f} = \frac{C^N_{y,\ell,f}}{\sum_\ell C^N_{y,\ell,f}} \tag{B14}$$

If the s^{th} survey is biomass-based, then the survey value $I_{y,s}$ is calculated as:

$$I_{y,s} = q_s \sum_a v_{y,a,s}\, N_{y,a} w_{y,a} \tag{B15}$$

where q is the scaling coefficient and s indexes survey.

If the survey is abundance-based, then:

$$I_{y,s} = q_s \sum_a v_{y,a,s}\, N_{y,a} \tag{B16}$$

The predicted proportions-at-age vulnerable to the survey is

$$p_{y,a,s} = \frac{v_{a,s}\, N_{y,a}}{\sum_a v_{a,s}\, N_{y,a}} \tag{B17}$$

Likelihoods

If the model is conditioned on catch and fishing mortality rates are estimated parameters, then the log-likelihood component Λ_1 of the catch is:

$$\Lambda_1 = \sum_f \left[\lambda^{\tilde{C}}_f \sum_y \left(-\log(0.01) - \frac{[\log(\tilde{C}^{\text{obs}}_{y,f}) - \log(\tilde{C}^{\text{pred}}_{y,f})]^2}{2 \times 0.01^2} \right) \right] \tag{B18}$$

where obs and pred indicate observed and predicted quantities, respectively, and λ are likelihood weights. With a small standard deviation for the catch likelihood relative to the variance in other likelihood components, the predicted catch should match the observed catch.

The log-likelihood component Λ_2 of survey data is:

$$\Lambda_2 = \sum_s \left[\lambda^I_s \sum_y \left(-\log(\sigma_{y\,s}) - \frac{[\log(I^{\text{obs}}_{y\,s}) - \log(I^{\text{pred}}_{y\,s})]^2}{2\sigma^2_{y\,s}} \right) \right] \tag{B19}$$

The log-likelihood component Λ_3 of catch-at-age data is:

$$\Lambda_3 = \sum_f \lambda_f^A \left[\sum_y O_{y\,f}^A \sum_a p_{y\,a\,f}^{\text{obs}} \log\left(p_{y\,a\,f}^{\text{pred}}\right) \right] \tag{B20}$$

where O^A is the annual sample sizes for the age compositions.

The log-likelihood component Λ_4 of catch-at-length data is:

$$\Lambda_4 = \sum_f \lambda_f^L \left[\sum_y O_{y\,f}^L \sum_a p_{y\,\ell\,f}^{\text{obs}} \log\left(p_{y\,\ell\,f}^{\text{pred}}\right) \right] \tag{B21}$$

where O^L is the annual sample size for the length compositions.

The log-likelihood component Λ_5 of the survey proportions-at-age is:

$$\Lambda_5 = \sum_s \lambda_s^{\text{IA}} \left[\sum_y \left(O_{y,s}^{IA} \sum_a p_{y,a,s}^{\text{obs}} \log\left(p_{y\,a\,s}^{\text{pred}}\right) \right) \right] \tag{B22}$$

where O^{IA} is the annual sample size of the survey age compositions.

The log-likelihood component Λ_6 of annual estimated recruitment deviates δ_y in log space is:

$$\Lambda_6 = \sum_y \left(-\log(\tau) - \frac{\delta_y^2}{2\tau^2} \right) \tag{B23}$$

where τ is the standard deviation of recruitment deviates.

The log-likelihood component Λ_7 of the equilibrium catch is:

$$\Lambda_7 = \sum_f \lambda_f^{\tilde{C}} \left(-\log(0.01) - \frac{[\log(\tilde{C}_f^{\text{eq,obs}}) - \log(\tilde{C}_f^{\text{eq,pred}})]^2}{2 \times 0.01^2} \right) \tag{B24}$$

The total log-likelihood LL to be maximized is:

$$\text{LL} = \sum_{i=1}^{7} \Lambda_i \tag{B25}$$

APPENDIX C. ADDITIONAL SENSITIVITY OPERATING MODELS

Additional sensitivity OMs were proposed during and following the May 26 and 27, 2020 meetings. The results of these sensitivity OM fits are capture here for consideration by the Atlantic Herring Working Group for changes to the reference and robust set of OMs.

Table C1. The sensitivity operating models developed for identifying a suitable reference set of operating models. All sensitivity operating models are a single-factor change from the reference case operating model.

OM #	Code	Description of single factor deviation from reference case
Add1	RefCase	As described above
Add2	M_oldM	Old natural mortality rate of 0.2
Add3	M_oldM-L	Old natural mortality rate of 0.2 with steepness 0.75
Add4	LS_none	Larval survey weight = 0
Add5	LS_none1	Larval survey weight = 0; index value = 1 for all year (check that weight of zero works)
Add6	R_post1994	Shift in recruitment after 1994. Recruitment for projections estimated from 1994 forward.
Add7	R_pre1994	Shift in recruitment after 1994. Recruitment for projections estimated from 1968–1993.
Add8	R_post2010	Shift in recruitment after 2010. Recruitment for projections estimated from 2010 forward.
Add9	G_0.5	Future growth is a linear increase of 0.5% of the mean 2016–2018 weights-at-age per year
Add10	G_1	Future growth is a linear increase of 1% of the mean 2016–2018 weights-at-age per year
Add11	Land_up25	Mixing and migration – landings are 25% higher
Add12	Land_down25	Mixing and migration – landings are 25% lower
Add13	Lognorm_comp	Composition distributions are lognormal instead of multinomial

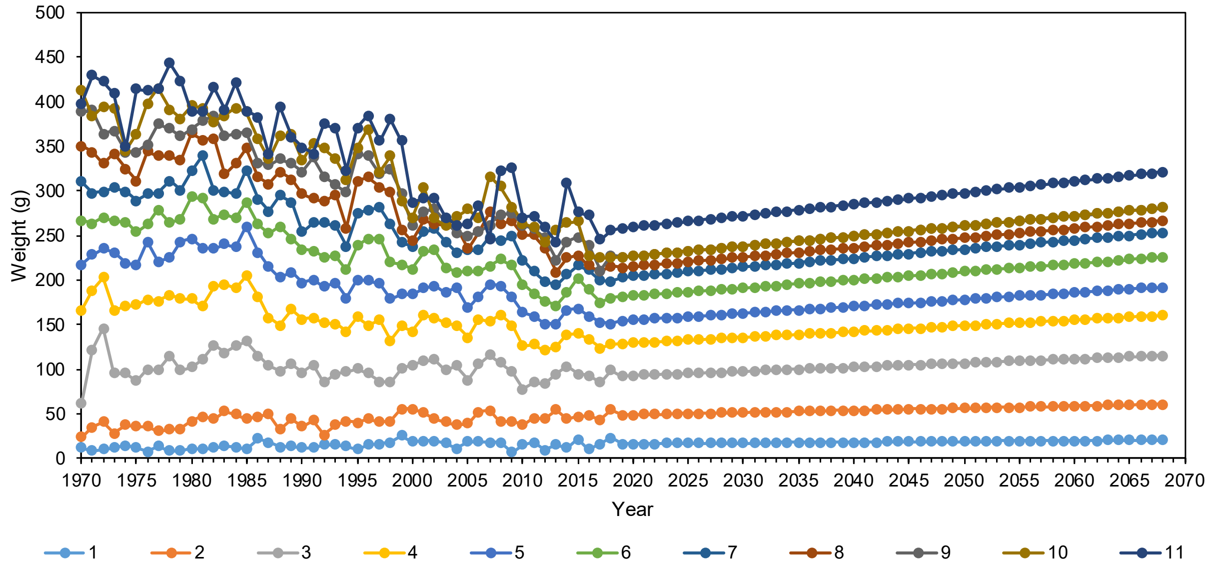


Figure C1a. Empirical weight-at-age 1970–2018 with projections to 2068 based on a 0.5% increase in the mean weight-at-age for 2016–2018.

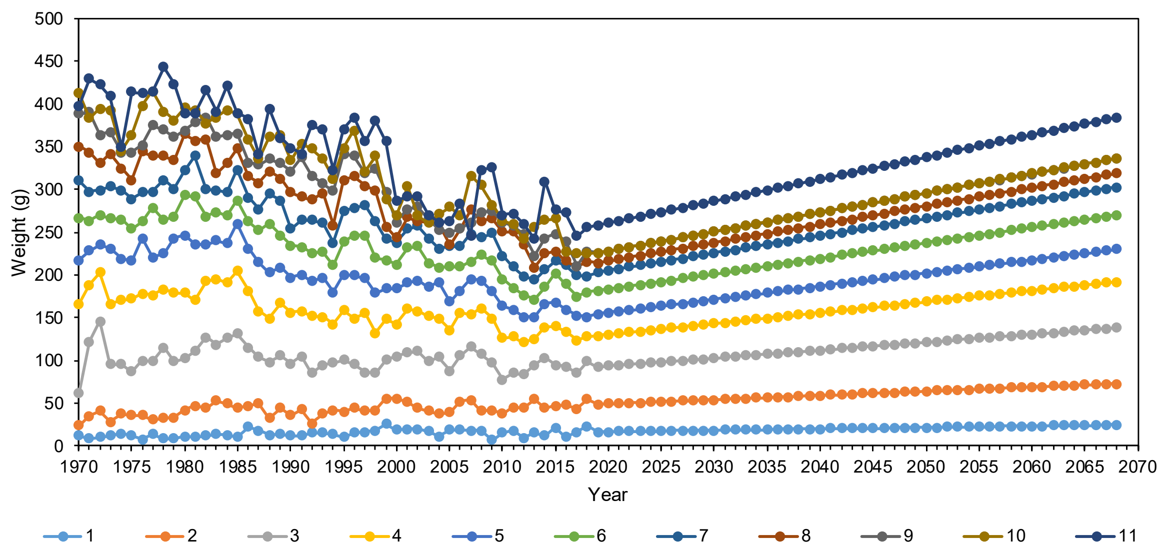


Figure C1b. Empirical weight-at-age 1970–2018 with projections to 2068 based on a 1% increase in the mean weight-at-age for 2016–2018.

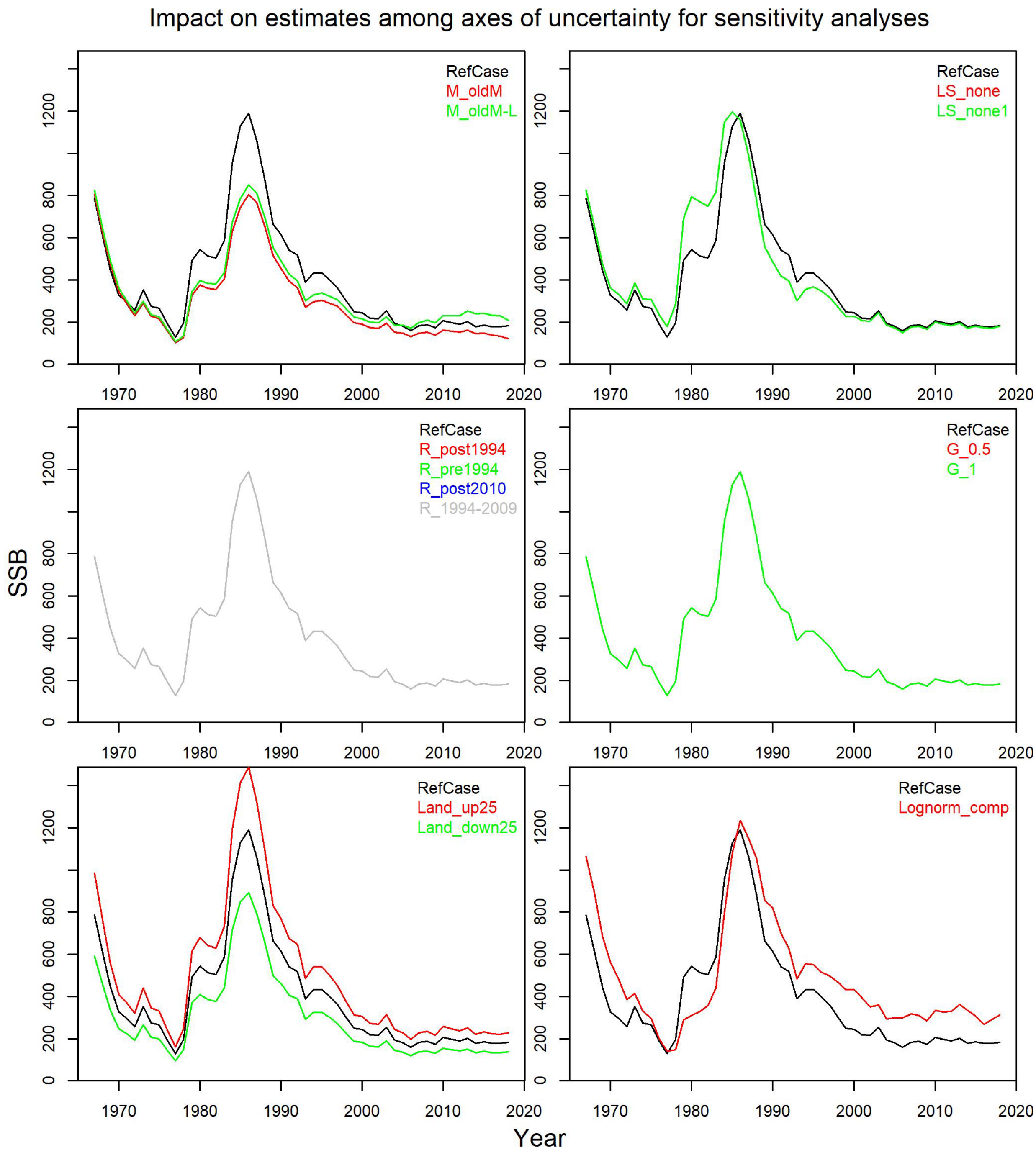


Figure C2a. Mean estimated Spawning Stock Biomass (kt) for the various sensitivity operating models of Table C1.

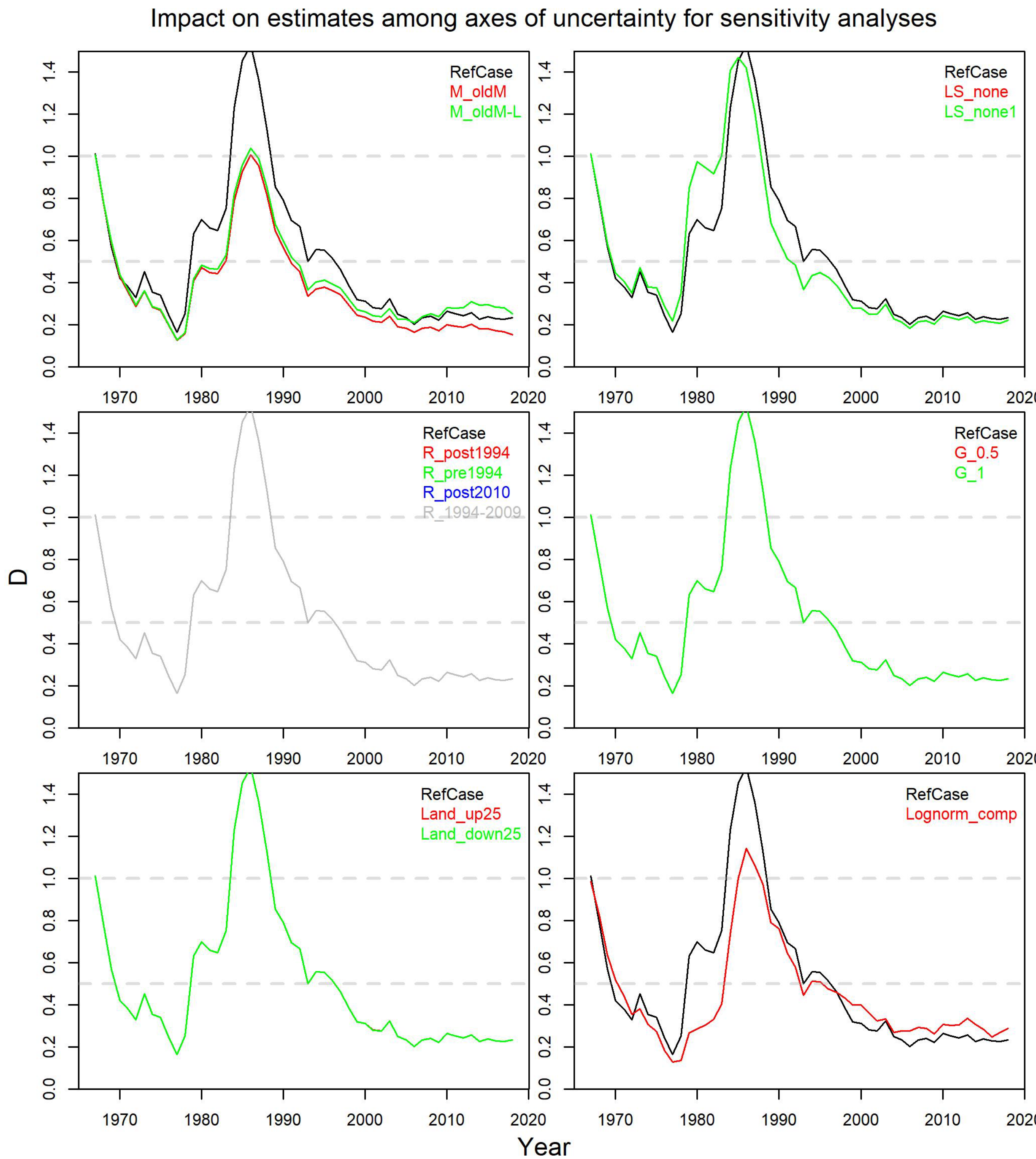


Figure C2b. Mean estimates of depletion (D) for the various sensitivity operating model described in Table C1.

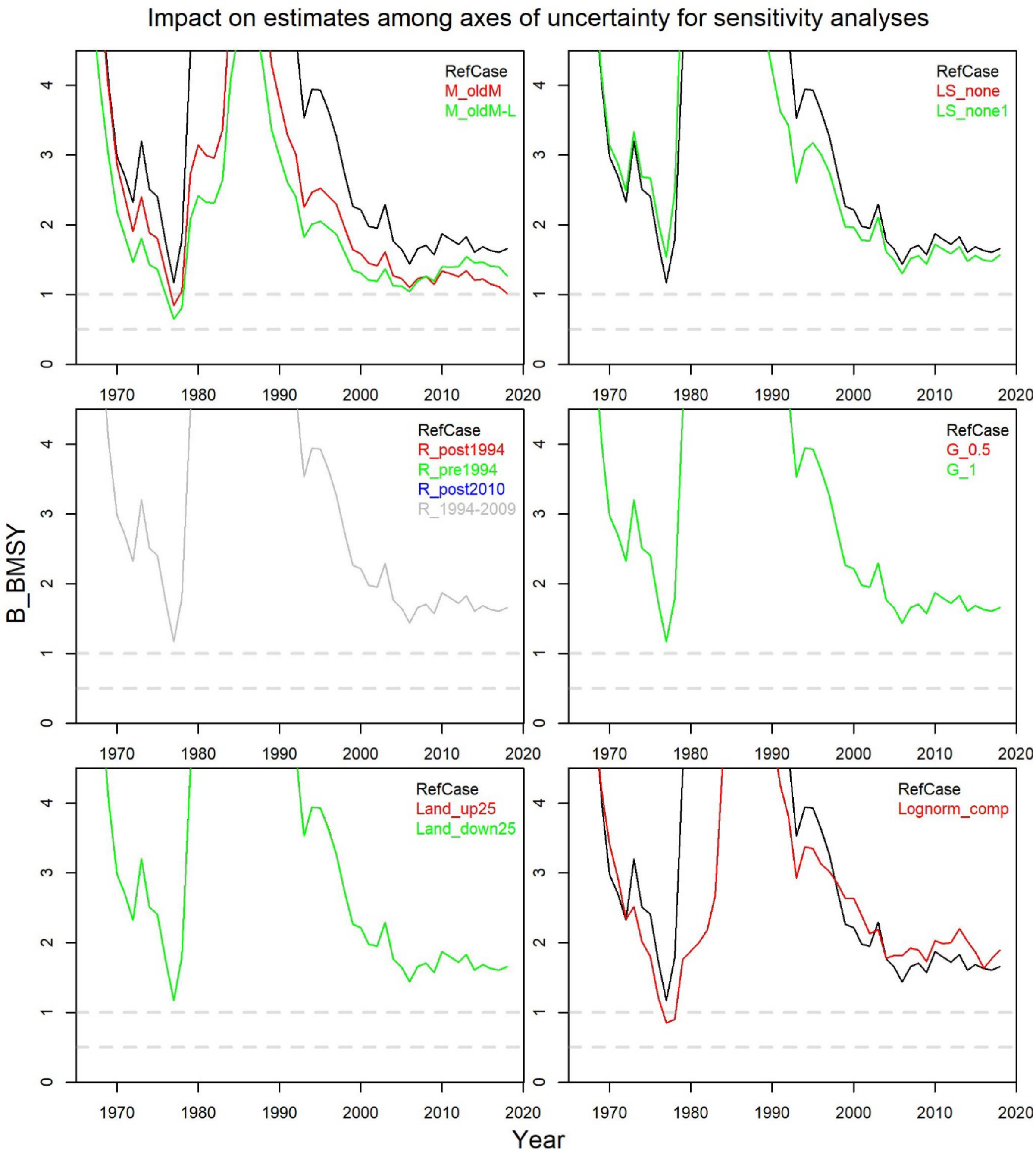


Figure C2c. Mean estimated Spawning Stock Biomass relative to MSY levels (B_BMSY) for the various sensitivity operating model described in Table C1.

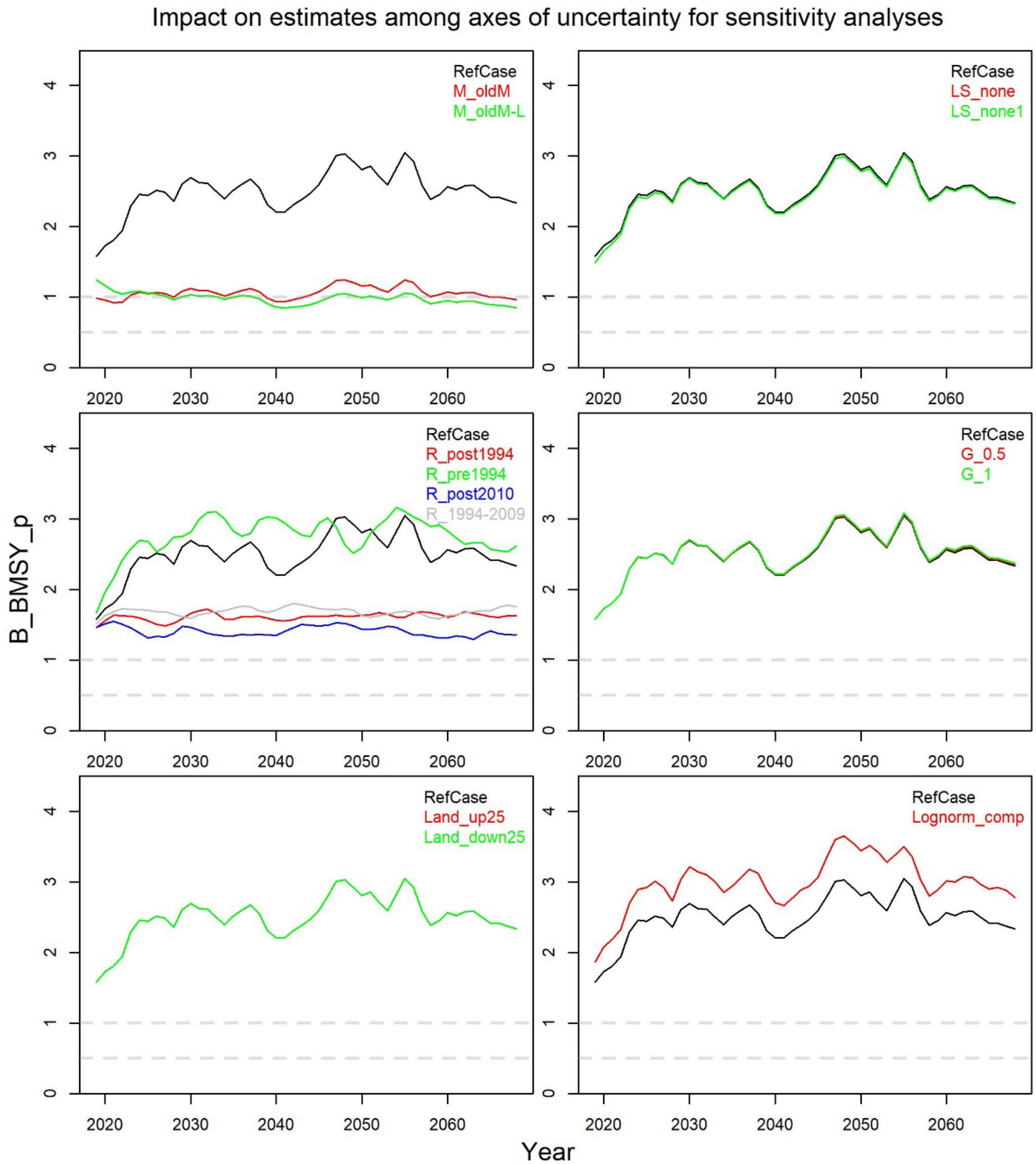


Figure C3. Mean estimates of projected Spawning Stock Biomass (B_BMSY_p) relative to MSY levels for a 50-year projection of a constant current fishing mortality rate scenario (a projection of status quo fishing). Note MSY is recalculated in each future year based on the corresponding growth in that year.

Table C2. Negative log-likelihoods for maximum likelihood fitting of sensitivity operating models (lower values represent better fit).

OM	Code	Total	Age Composition (length composition, where applicable)				Indices		Index catchability	
			Gill	Other	Purse seine	Weir	Acoustic	Larval	q Acoustic	q Larval
Add1	RefCase	668.33	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
Add2	M_oldM	685.54	63.51	100.01	150.05	153.18	-3.62	36.35	3.22	0.05
Add3	M_oldM-L	715.86	70.43	99.63	154.42	154.26	2.28	35.54	2.4	0.04
Add4	LS_none	623.71	57.82	99.35	140.85	154.71	-3.48	0	3.18	0.03
Add5	LS_none1	623.71	57.82	99.35	140.85	154.71	-3.48	0	3.18	0
Add6	R_post1994	668.33	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
Add7	R_pre1994	668.33	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
Add8	R_post2010	668.33	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
Add9	R_1994-2009	668.33	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
Add10	G_0.5	668.33	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
Add11	G_1	668.33	58.38	100.33	142.56	154.43	-3.30	39.68	3.03	0.04
Add12	Land_up25	668.33	58.38	100.33	142.56	154.44	-3.30	39.68	2.43	0.03
Add13	Land_down25	668.33	58.38	100.33	142.56	154.43	-3.30	39.68	4.04	0.05
Add14	Lognorm_comp	508.85	92.57	108.49	59.01	95.89	-3.80	34.22	1.99	0.03

Table C3. Mean estimates derived from the sensitivity Operating Models (OM). Maximum Sustainable Yield (MSY) quantities were calculated by the method of Walters and Martell (2004). FMSY is apical fishing mortality rate at MSY (maximum over length classes). SSBMSY = Spawning Stock Biomass at MSY. BMSY is total vulnerable biomass at MSY. UMSY is the fraction of vulnerable biomass caught at MSY (harvest rate). SSB0 is unfished Spawning Stock Biomass. RefY is the Reference Yield, the maximum yield obtainable by a fixed fishing rate given future conditions and current fishery selectivity. Blow is the biomass for which it would take two mean generation times to reach half of BMSY given current fishing and biological parameters. MGT is Mean Generation Time calculated at the average age of a mature fish in the unfished population. SSB/SSBMSY is current spawning biomass relative to MSY levels. D is current stock depletion calculated as current Spawning Stock Biomass divided by unfished Spawning Stock Biomass.

OM	Code	MSY (kt)	FMSY	SSBMSY (kt)	SSBMSY/ SSB0	BMSY (kt)	UMSY	SSB0 (kt)	RefY (kt)	Blow (kt)	MGT (yrs)	SSB/ SSBMSY	D
Add1	RefCase	75.587	0.623	109.912	0.142	253.808	0.414	776.541	71.704	0.040	5.089	1.613	0.229
Add2	M_oldM	41.496	0.343	119.947	0.150	181.949	0.267	799.766	40.781	0.038	5.848	1.120	0.168
Add3	M_oldM-L	34.341	0.210	164.300	0.201	221.541	0.173	818.268	29.630	4.473	5.848	1.398	0.281
Add4	LS_none	78.957	0.613	115.290	0.141	265.802	0.408	814.774	74.905	0.013	5.089	1.489	0.210
Add5	LS_none1	78.957	0.613	115.290	0.141	265.802	0.408	814.774	74.905	0.013	5.089	1.489	0.210
Add6	R_post1994	75.587	0.623	109.912	0.142	253.808	0.414	776.541	47.264	0.039	5.089	1.613	0.229
Add7	R_pre1994	75.587	0.623	109.912	0.142	253.808	0.414	776.541	78.370	0.046	5.089	1.613	0.229
Add8	R_post2010	75.587	0.623	109.912	0.142	253.808	0.414	776.541	39.753	0.039	5.089	1.613	0.229
Add9	R_1994–2009	75.587	0.623	109.912	0.142	253.808	0.414	776.541	51.185	0.042	5.089	1.613	0.229
Add10	G_0.5	75.587	0.623	109.912	0.142	253.808	0.414	776.541	89.857	0.040	5.089	1.613	0.229
Add11	G_1	75.587	0.623	109.912	0.142	253.808	0.414	776.541	108.138	0.040	5.089	1.613	0.229
Add12	Land_up25	94.480	0.623	137.382	0.142	317.246	0.414	970.628	89.628	0.050	5.089	1.613	0.229
Add13	Land_down25	56.693	0.623	82.442	0.142	190.370	0.414	582.453	53.780	0.029	5.089	1.613	0.229
Add14	Lognorm_comp	98.251	0.532	164.315	0.152	353.855	0.366	1,079.303	93.953	2.109	5.089	1.783	0.271

APPENDIX D. TERMINOLOGY AND ACRONYMS

Table D1. Terminology and Acronyms.

Term	Meaning
MSE	Management Strategy Evaluation. A process for identifying robust management procedures that can meet management performance objectives by simulation testing against a range of system uncertainties.
MP	Management Procedure (aka Harvest Strategy). A rule (algorithm) that provides fishery management advice from fishery data (typically a more streamlined data set than a conventional stock assessment, such as catch data and an index of relative abundance).
CMP	Candidate MP. One of multiple options for providing management advice that are to be comparatively evaluated using closed-loop simulation.
OM	Operating Model. The simulated ('true') stock and exploitation dynamics for testing performance of management procedures. Often similar to the structure of a conventional stock assessment but are sufficiently flexible to span a range of scenarios that cover fishery system uncertainties.
Harvest Control Rule	An function providing additional rules for target exploitation level typically applied to the outputs of a data-rich stock assessment in order to provide management advice (for example a '40-10' rule that prescribes FMSY fishing when stocks are estimated to be above 40% of unfished biomass and imposes a linear decline to zero exploitation between 40 and 10% of unfished stock biomass).
Closed-loop simulation	A control-systems simulation approach that models the feedback between a control rule (MP) and a system (operating model) accounting for feedbacks.
Observation error model	Simulates data collection for use in closed-loop testing of management procedures including statistical properties based on fit of operating model to data (e.g., bias, imprecision, hyperstability in indices, etc.).
Implementation error model	Simulates how well management advice provided by an MP is followed in closed-loop simulation (e.g., underages / overages of TACs)
Depletion	Spawning Stock Biomass in a given year relative to asymptotic (expected, or mean) 'unfished' (zero exploitation) Spawning Stock Biomass calculated from the growth, survival and resilience in the same year.
MSY	Maximum Sustainable Yield: maximum equilibrium (long term) yield achieved by the aggregate fishery selectivity, growth, resilience in a given year. Here MSY and related quantities are calculated by the numerical approach of Walters and Martell 2004.
F_{MSY}	Apical (maximum of all age classes) instantaneous fishing mortality rate commensurate with MSY.
TAC	Total Allowable Catch (in the case of Herring an annual tonnage)

Term	Meaning
SRA	Stock Reduction Analysis. A fishery dynamics model that assumes reported catches were taken exactly (requires complete catches) and can use age composition, length composition and relative abundance index data (or a combination thereof).
MSEtool	Management Strategy Evaluation toolkit (Huynh et al. 2019): an open source R package for developing and testing of candidate management procedures conducting and MSE for data-rich fisheries.
DLMtool	Data Limited Methods toolkit (Hordyk and Carruthers 2019: an open source R package for developing and testing of candidate management procedures conducting and MSE for data moderate (e.g., a relative abundance index or composition data) and data-limited (e.g., only catch data) fisheries.
SRA_scope	An MSEtool function that conditions an operating model to fishery data using Template Model Builder.
Catchability, *q*	The constant of proportionality of an abundance index (e.g., vulnerable biomass = *q* x index). When q = 1 the index is an *absolute* index of vulnerable biomass.
SSB	Spawning Stock Biomass. The sum product of numbers at age, maturity at age and weight at age for a given model year.
TMB	Template Model Builder (Kristensen et al. 2016): An R library for numerical parameter estimation
VPA	Virtual Population Analysis: a stock assessment modelling approach that assumes total catches at age are known, and which back-constructs cohorts subtracting catch-at-age and accounting for natural and fishing mortality rate.