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 2025/064

Maritimes Region

Framework Review for 4X5Y Haddock: Operating Model Specification, Projections, Simulation Approach, and Reference Points

Barrett, T.J., Barrett, M.A., and Hebert, N.E.

Saint Andrews Biological Station
Fisheries and Oceans Canada
125 Marine Science Drive
St. Andrews, New Brunswick, E5B 0E4

October 2025

Canada

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/
DFO.CSAS-SCAS.MPO@dfo-mpo.gc.ca

ISSN 1919-5044
ISBN 978-0-660-78887-6 Cat. No. Fs70-5/2025-064E-PDF

Correct citation for this publication:

Barrett, T.J., Barrett, M.A., and Hebert, N.E. 2025. Framework Review for 4X5Y Haddock: Operating Model Specification, Projections, Simulation Approach and Reference Points. DFO Can. Sci. Advis. Sec. Res. Doc. 2025/064. v + 63 p.

Aussi disponible en français :

Barrett, T.J., Barrett, M.A. et Hebert, N.E. 2025. Examen du cadre pour l'aiglefin des divisions 4X5Y : spécification des modèles opérationnels, projections, approche de simulation et points de référence. Secr. can. des avis sci. du MPO. Doc. de rech. 2025/064. v + 70 p.

TABLE OF CONTENTS

ABSTRACT

The 4X5Y Haddock stock assessment model was rejected in 2018 due to large retrospective patterns and a mismatch between model-estimated biomass and survey biomass. Catch advice has since been provided qualitatively based on the DFO ecosystem survey index. In 2023, DFO science began a framework for a new assessment method for the 4X5Y Haddock stock, which included a review of the data inputs and the development of assessment models. Three assessment models were identified and are presented in this document with different hypotheses to explain the mismatch between fishery and survey data inputs: density dependent natural mortality (Models 1 and 2) and constant natural mortality with random effects on the stock numbers-at-age transitions (Model 3). Reference points were defined for each model and weighted to provide an overall stock status. Two sets of operating models were defined based on these three models to account for uncertainty in future stock productivity. One set assumed status quo productivity with no large recruitment events, and the other assumed productivity consistent with the last 25 historical years and includes a large recruitment event consistent with the 2013 cohort. A management procedure (MP) evaluation framework using closed-loop simulation is presented in this document and was used for evaluating the performance of MPs relative to management objectives. Example management procedures were used in this document to demonstrate how the MP evaluation framework will be applied. When an MP meets the minimum performance standard defined for conservation objectives, tradeoffs among MPs for all objectives can be evaluated. This MP evaluation framework can be used in the future to select an MP to provide catch advice for the 4X5Y Haddock fishery. Exceptional circumstances are defined to address situations outside the range of scenarios for which the MP was simulation tested or when the data required to apply the MP are not available.

1. INTRODUCTION

Haddock (*Melanogrammus aeglefinus*) are caught as part of a multi-species groundfish fishery concentrated on the western Scotian Shelf (SS) and in the Bay of Fundy (BoF) in the Northwest Atlantic Fisheries Organization (NAFO) Divisions 4X5Y. In 2017, a virtual population analysis (VPA) model with fixed natural mortality time blocks on older fish, was chosen to provide catch advice (Wang et al. 2017). In 2018, projections from the VPA model showed large retrospective patterns and there was a mismatch between the model-predicted biomass and the survey biomass (Finley et al. 2018). The VPA model has therefore not been used to provide catch advice or estimate stock status since 2018, and stock status updates have been provided qualitatively by comparing the survey biomass index to 40% and 80% of the 1985–2020 time series median (DFO 2023a).

The last data review for 4X5Y Haddock was completed in October 2023 (Barrett and Barrett 2025). The data review was the first meeting of a series of peer-review meetings for the 4X5Y Haddock framework review, initiated by Fisheries and Oceans (DFO) Maritimes Science. The second meeting was a review of the structural assumptions and fits of population dynamics models, which will form the basis of a set of operating models (OMs) that capture the main uncertainties in the fishery and population dynamics for 4X5Y Haddock. The final decisions from that meeting are included in this document.

The objective of the third (and final) peer-review meeting was to review the projection scenarios and simulation approach using the set of models that were chosen to capture uncertainty in the historical fishery dynamics at the second peer-review meeting. These models and projection scenarios were used to define OMs to be used in the framework to evaluate the performance of management procedures (MPs) for the 4X5Y Haddock fishery. Reference points were also defined to inform on stock status and to be used in performance metrics in the MP evaluation framework.

1.1 OBJECTIVES

The specific objectives of the third peer-review meeting were to review the scientific components of the MP evaluation framework:

- The set of OMs that capture uncertainty in the historical and future stock dynamics.
    - Uncertainty in the historical dynamics was captured by the models from the second peer-review meeting.
    - Uncertainty in the future dynamics was captured by defining projection scenarios.
- The projection scenarios for weight-at-age, maturity-at-age, selectivity, natural mortality rate (M), and recruitment for each OM.
- A biomass limit reference point (LRP) and fishing mortality reference point (F_{ref}) and one or more candidate biomass upper stock reference point(s) (USR) for each OM.
- An approach to estimate overall stock status from the set of OMs.
- The simulation approach to test MPs.
    - Observation model assumptions.
    - Weighting of OMs.
- Exceptional circumstances that would indicate assumptions of the OMs are no longer valid.

- The proposed frequency and timing of interim-year updates to be provided between full peer-reviewed frameworks, and the recommended timing of the next framework.

2. MANAGEMENT OBJECTIVES

The MP evaluation framework is a simulation-based, analytical framework used to evaluate the performance of candidate MPs relative to pre-specified management objectives. In September 2023, management objectives for the 4X5Y Haddock fishery were drafted to meet both the conservation objectives of DFO and the economic considerations for Industry. The management objectives were defined by DFO in consultation with stakeholders during a meeting of the Scotia-Fundy Groundfish Advisory Committee. Management objectives are expressed in terms of performance metrics with associated probabilities and timeframes (Table 1). The first four management objectives (Table 1) are consistent with DFO's Precautionary Approach (PA) policy (DFO 2009) and the recently defined guidelines for implementing the fish stocks provisions (FSPs) in the Fisheries Act (DFO 2022). These first four management objectives have a minimum performance standard (a probability threshold) that must be met by an MP. The fishery objectives related to maximizing the total allowable catch (TAC) and avoiding inter-annual variability in TAC do not have probability thresholds, and performance relative to these objectives is evaluated through tradeoffs among MPs.

Table 1. Management objectives and performance metrics developed for the 4X5Y Haddock stock.

Management Objective	Performance metric	Probability threshold	Timeframe
1. Maintain stock above LRP	P(SSB > LRP)	75-95%[1]	Average across two generations (10 years) AND average across years in the 25-year projection period OR (if stock is < LRP): After the rebuilding[2] timeframe (e.g., in projection year 10) AND average across years in projection period after the rebuilding timeframe (e.g., projection year 10 to end of 25-year projection period)
2. Maintain stock above USR	P(SSB > USR)	≥50%	Average across two generations (10 years) AND average across years in the 25-year projection period
3. Maintain fishing mortality below F_{ref}	$P(F < F_{ref})$	≥50%	Every year in the 25-year projection period
4. Promote stock growth when stock is below USR	$P(SSB_{y+5} > SSB_y)$[3]	≥95% if stock is < LRP ≥75% if stock is < USR and recent trajectory is a decline[4] ≥50% if stock is < USR and recent trajectory is not a decline[4]	One generation (5 years)
5. Maximize TAC	Average catch	–	5, 10, 15 years
6. Avoid large inter-annual changes in TAC	Number of years that interannual change in catch exceeds 15% in short, medium, and long terms	–	5, 10, 15 years

[1] Final threshold to be determined by Resource Management and will be used to tune MPs
[2] Reasonable would be defined to be consistent with the DFO PA Policy and DFO rebuilding guidelines (e.g., maximum of 1.5–2 generations or 2 – $3T_{min}$, where T_{min} is the time the stock would take to rebuild to that target in the absence of fishing).

[3] The probability here is the probability of preventable decline. Should SSB decrease under F = 0 then this should be P(SSB_{y+5} > $SSB_{y+5 \text{ with } F = 0}$).
[4] "recent trajectory" has not yet been defined.
Dash indicates not applicable

3. MODEL INPUTS

A review of the data inputs for this framework, including stock structure, biological parameters, ecosystem considerations, fishery catch, indices of abundance, and catch and survey age-composition was conducted in a CSAS peer-review meeting on October 17–18, 2023 (Barrett and Barrett 2025).

3.1 FISHERY

Haddock are landed using both otter trawl and fixed (e.g., long line, gill net, hand line) gears, with otter trawl catches predominant over the last two decades. The fishery data (catch and catch composition) were summarized separately for Bay of Fundy (BoF, DFO ecosystem survey strata 482–495) and the western Scotian Shelf (SS, DFO ecosystem survey strata 470–481) based on differences in growth. Each area and gear type were combined into one fleet in the model but were weighted by region (SS and BoF).

Length frequency (LF) distributions from the fishery were estimated from port sampling of the catch, while the ages were estimated from samples taken from both the port and observer sampling programs. The estimated weights-at-length for each LF distribution were estimated using the DFO ecosystem survey weight-length relationships. Fish aged 1–12+ are included in both the catch-at-age (CAA) and weight-at-age (WAA) inputs. Both inputs were calculated separately by region (BoF and SS) and then combined together to create a single input matrix for the stock.

3.2 SURVEY

The DFO summer ecosystem survey index (1970–2022) and survey age-composition data were used as model inputs. Two methods for calculating the DFO summer ecosystem survey index were explored; 1) the *status quo* (annual mean biomass per tow as a weighted mean proportional to size of the stratum and number of tows), and 2) a weighted mean assuming a delta-lognormal distribution (Pennington 1996). Although the indices estimated using the two different methods were similar, the delta-lognormal index reduced the influence of extreme tows and had a more stable coefficient of variation over time.

Model inputs from the DFO ecosystem survey included survey weight-at-age (WAA), spawning stock biomass (SSB) WAA and the January 1st stock WAA by adjusting the survey WAA to the spawning time using the Rivard method (Rivard 1982). Maturity-at-age inputs were based on the mean values from the DFO summer ecosystem survey (1986–2022), weighted by the survey numbers-at-age for each region to account for differences in maturity between regions.

4. POPULATION DYNAMICS MODELS

The Woods Hole assessment model (WHAM, Stock and Miller 2021) was selected as the modeling platform for the 4X5Y Haddock stock. WHAM is an open-source R package (https://github.com/timjmiller/wham) that has the ability to fit state-space models that use random effects to model population parameters, resulting in fewer parameters being estimated compared to traditional statistical catch-at-age models. WHAM was specifically chosen for 4X5Y Haddock given its flexibility to include random effects (process error) for natural mortality rate (M), numbers-at-age (NAA), selectivity, and catchability (q), and its ability to include environmental covariates in the estimation of model parameters (M, recruitment, and q).

4.1 MODEL CONFIGURATIONS AND EXPLORATION

An initial model (*m0*; Table A1) was defined and used as a starting point to identify candidate models with acceptable diagnostics for which to explore alternative model configurations. The primary structural assumptions influencing model fits were *M*, specification of fishery and survey selectivity, choice of likelihood distribution for age composition data, and data sources (e.g., including the ITQ survey or catches in survey strata 482 and 483 in the south of NAFO area 4Xp). Model exploration began with assuming a Dirichlet distribution for age composition data because it is self-weighting. The multinomial distribution was also explored to allow for flexibility in the relative weights of different data sources.

All models estimated recruitment as random about a mean value and the age-at-recruitment used in WHAM is age-1. Haddock recruitment in adjacent stocks is episodic (NFSC In Prep[1], Wang et al. 2022) with occasional strong cohorts and has not been modeled using a stock-recruitment relationship. Based on the survey age-composition data (Barrett and Barrett 2025), the 4X5Y Haddock stock had a large recruitment event in 2013 that has a strong contribution to the survey and the fishery age-composition data. 4X5Y Haddock lacks significant contrast in abundance over time so the assumption of deviations from mean recruitment (i.e., a null stock-recruitment relationship) is consistent with pragmatic recommendations from a recent review paper (Brooks 2024) for modeling recruitment of Haddock.

The initial model (*m0*; Table A1) exhibited retrospective patterns in spawning stock biomass (SSB), fishing mortality rate (*F*), and recruitment, and had a poor fit to the DFO survey index with a long series of positive residuals around 1990–2010 (Table A3). Model exploration was focused on improving diagnostics and improving the fit to the DFO survey index (mismatch in the DFO survey index and the fishery and survey age composition data). The main focus of the exploration was on alternative structural assumptions: higher *M* (i.e., > 0.2) and/or time-varying *M,* alternative forms of fishery and survey selectivity, and including random effects on the NAA transitions. A summary of all model configurations explored and reviewed in the second peer-review meeting is provided in Appendix A. Model diagnostics were assessed through visual inspections of fit to the data, convergence checks (self-test, jitter analysis of model initial values), examination of raw and one-step ahead (OSA) residuals, and summaries of the mean Mohn's rho over a 7-year peel for SSB, *F*, and recruitment (NAA-1). Akaike's Information Criteria (AIC) was used to inform parsimonious models with a better fit (for models with the same data inputs).

4.2 FINAL MODEL SELECTION

Three models were accepted at the conclusion of the second peer-review meeting (March 20–21, 2024). These models accounted for different hypotheses to explain the mismatch between fishery and survey data:

> **Model 1**: used DFO survey NAA1+ as a covariate for *M* and assumed logistic annual selectivity for the fishery (model *m0_Me1+_dln_sig0.05_L1*; Table 2; Appendix B).
> **Model 2**: used DFO survey NAA1+ as a covariate for *M* and used random effects (2dar1 correlation structure) on fishery selectivity parameters (model *m0_Me1+_dln_sig0.05_L1_V2dar1;* Table 2; Appendix B).

[1] Northeast Fisheries Science Center (NFSC), In Prep. 2022 Management Track Assessment Report for Georges Bank Haddock. US Dept Commer. Northeast Fish Sci Cent Ref Doc.

Model 3: assumed a constant *M* = 0.2 and used random effects (2dar1 correlation structure; recruitment decoupled) on NAA, with no bias-correction on the mean for the NAA transitions (*m0_M0.2_NAARE_2dar1_DFO0.5cv_noBC*; Table 2; Appendix B).

The specific structural assumptions of these model configurations are summarized in Table 2. Models 1 and 2 assume a time-varying *M*, to account for the hypothesis of density-dependent natural mortality. Although a model (*m0_M2b_2010*; Table A1) with *M* estimated in two time blocks (1979–2009 and 2010–2022) had lower retrospective patterns, Models 1 and 2 were preferred because they have a mechanism to drop *M* from the high value caused by the strong 2013 cohort and they do not rely on the subjectivity of choice for the time blocking. The additional uncertainty captured in Models 1 and 2 is the structure of the time-varying selectivity. Model 1 estimated annual fishery selectivity parameters. Model 2 estimated a single set of selectivity parameters, and annual selectivity was determined by estimating random effects on the selectivity parameters, assuming a 2dar1 correlation structure. Model 3 was fit with random effects on NAA assuming a 2dar1 correlation structure across years and ages as an alternative hypothesis to time-varying *M*. This model assumed a time- and age-invariant *M* of 0.2, consistent with previous assessments for Haddock.

Diagnostics for model configurations in Appendix A were presented in the second peer-review meeting. In this paper, the detailed diagnostics for Models 1–3 are presented in Appendix B.

Table 2. Configuration of population dynamics models 1–3.

Model feature	*Model 1*	*Model 2*	*Model 3*
Model years	1970 – 2022	1970 – 2022	1970 – 2022
Modeled age classes	1 – 12+	1 – 12+	1 – 12+
Catches	All[1]	All[1]	All[1]
Fishery Fleet structure	Single fleet	Single fleet	Single fleet
Fishery selectivity	Logistic, annual blocks	Logistic, 2dar1 correlation random effects	Logistic, annual blocks
Survey	DFO survey [1970–2022 without 2021]	DFO survey [1970–2022 without 2021]	DFO survey [1970–2022 without 2021]
Survey selectivity	Logistic, single block	Logistic, single block	Logistic, single block
Stock recruitment model	Mean recruitment with log deviations estimated as fixed effects	Mean recruitment with log deviations estimated as fixed effects	Mean recruitment with log deviations estimated as random effects
Natural mortality rate	Time-varying based on survey NAA-1, with one year time lag	Time-varying based on survey NAA-1, with one year time lag	M = 0.2
Likelihood function for fishery catch and survey index data	Lognormal	Lognormal	Lognormal
Likelihood function for catch and survey age-composition data	Dirichlet, pooling zeros	Dirichlet, pooling zeros	Dirichlet, pooling zeros
Random effects (process error)	None	Fishery selectivity parameters	NAA (2dar1 correlation), recruitment decoupled

[1]Some models explored excluded catches from the south of NAFO subdivision 4Xp.

5. OPERATING MODELS

A set of OMs was defined to capture uncertainty in the historical and future stock dynamics. The major sources of uncertainty for the historical dynamics were captured using three models fit in WHAM with different structural assumptions (Section 4.2). The WHAM models were converted to OMs in MSEtool (Hordyk et al. 2022) using a modified version of the MSEtool function *WHAM2OM*. The modified *WHAM2OM* function generates a set of simulations by sampling from a multivariate normal distribution defined by the WHAM model parameter estimates and their variance-covariance matrix. This results in matrices for stock numbers-at-age (NAA), fishing mortality-at-age (FAA), and for models with random effects on NAA, the process error terms in NAA ($\varepsilon_{a,y}$). For each model, 100 simulations were generated to capture uncertainty in the corresponding WHAM model's parameter estimates. Matrices for spawning stock biomass weight-at-age (SSB WAA), catch WAA, maturity-at-age, and natural mortality-at-age (MAA), which were used as time-varying data inputs to the WHAM models, were provided to the modified *WHAM2OM* function and remained constant across simulations. R_0 was simulation specific, defined as the median NAA-1 from the historical period.

Future recruitment was modeled in the OM as the median recruitment (i.e., the R_0 parameter) times an annual recruitment deviation [$\exp(\varepsilon_{0,y})$; Eqn 2]. The MSEtool OM structure uses an age of recruitment of zero, so the *WHAM2OM* function defines the OM by shifting the NAA-1 for year y to NAA-0 in year y – 1 and defines *M* of 1×10^{-6} (i.e., approximately 0) for age zero. The initial year NAA values are model-estimated parameters in the WHAM models. In MSEtool, the initial year NAA values are estimated using a_{max} = 12 (i.e., the number of age classes in the model), recruitment deviations prior to the first historical year (1970), and using the initial year MAA to establish the initial year NAA. The OM stock dynamics follow the equations:

$$N_{a,y} = \begin{cases} R_0 \exp(\varepsilon_{0,y}) & a = 0 \\ N_{a-1,y-1} \exp(-F_{a-1,y-1} - M_{a-1,y-1}) & a \in (1, a_{max} - 1) \\ N_{a-1,y-1} \exp(-F_{a-1,y-1} - M_{a-1,y-1}) + N_{a,y-1} \exp(-F_{a,y-1} - M_{a,y-1}) & a = a_{max} \end{cases}$$

(Eqn 1)

For Model 3 with random effects on the NAA transitions, the NAA process error term was combined with the *M* term in the specification of the OM dynamics. i.e., the equation:

$$N_{a,y} = N_{a-1,y-1} \exp\left(-F_{a-1,y-1} - M_{a-1,y-1} + \varepsilon_{a,y}\right) \quad \text{(Eqn 2)}$$

was captured as

$$N_{a,y} = N_{a-1,y-1} \exp\left(-F_{a-1,y-1} - M'_{a-1,y-1}\right) \quad \text{(Eqn 3)}$$

where

$$M'_{a-1,y-1} = M_{a-1,y-1} - \varepsilon_{a,y} \quad \text{(Eqn 4)}$$

Modifications to the *WHAM2OM* function, *Simulate* function, *Project* function, and other functions called within these functions in the MSEtool package were made to accommodate negative values of *M*. Note that M' was only used in the NAA transitions (Eqn 1) and was not used in the estimation of catch where catch ($C_{a,y}$) in numbers was calculated using *M* and not M' as:

$$C_{a,y} = N_{a,y} \left(\frac{F_{a,y}}{F_{a,y} + M_{a,y}}\right) \left(1 - \exp(-F_{a,y} - M_{a,y})\right) \quad \text{(Eqn 5)}$$

Comparisons were conducted between the WHAM models and the median across 100 simulations from the OMs for historical SSB, fully-selected F, and recruitment to confirm that the OM dynamics truly capture the WHAM model dynamics (Figure 1). The full matrix of NAA from WHAM Model 1 and the median across 100 simulations from OM1 is shown in Figure 2.

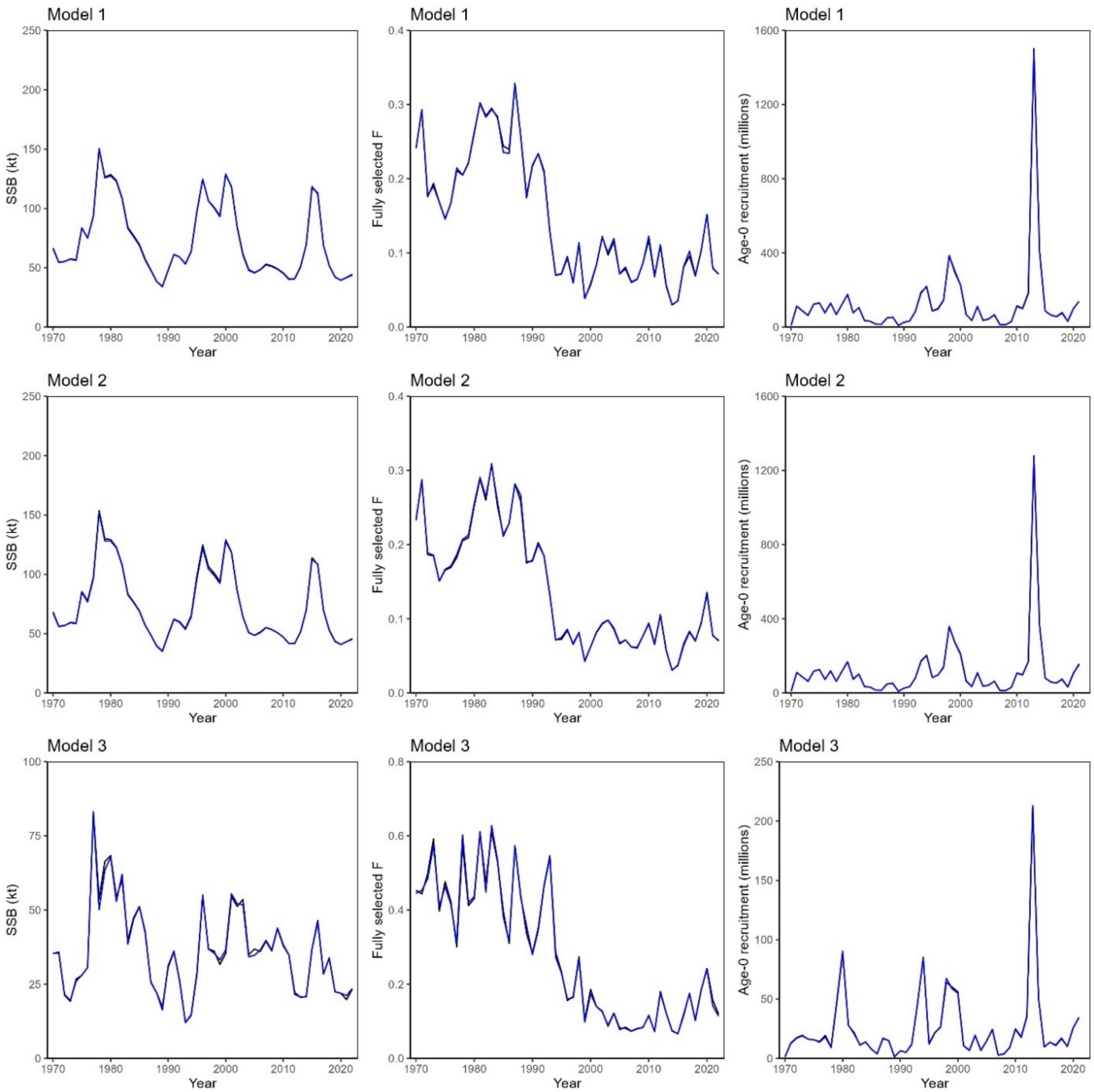

Figure 1. Comparison of historical spawning stock biomass (SSB), fully-selected fishing mortality (F), and recruitment for the WHAM model (blue) to the median of 100 simulations from the operating model (black) for Model 1 (top row), Model 2 (middle row), and Model 3 (bottom row).

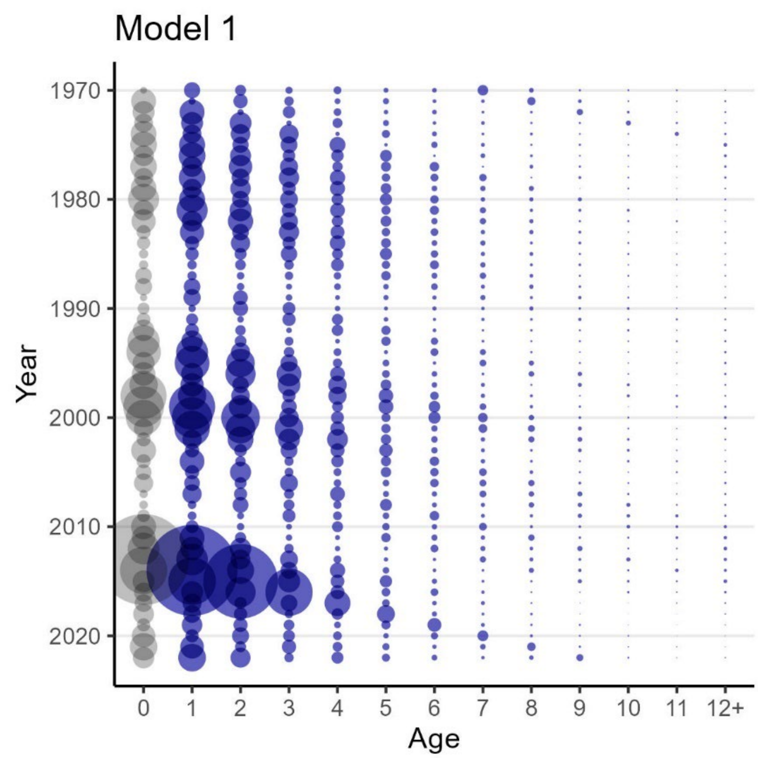

Figure 2. Comparison of the historical stock numbers-at-age-matrix for the WHAM model (blue) to the median of 100 simulations from the operating model (OM) (black) for Model 1. Points are plotted with 50% transparency and a natural mortality rate (M) of 1×10^{-6} is assumed for age 0 in the OM.

Two OMs were defined for each WHAM model to account for uncertainty in the future stock dynamics (Table 4). One set of OMs (OM1, OM2, OM3) was defined with future recruitment randomly sampled from historical recruitments (excluding the recruitment for the 2013 cohort). The objective here was to simulate conditions consistent with average recruitment without large cohorts. The other projection assumptions for these OMs are described in the following sections and in Table 4.

The second set of OMs (OM1b, OM2b, OM3b) was defined to have the same recruitment deviations, NAA deviations (OM3b), growth, selectivity, and *M* in the 25-year projection period as was observed in the last 25 historical years (Table 4), although growth was modelled into the projection period to allow a transition from terminal year WAA (Section 5.5). The objective here was to simulate conditions consistent with what was observed in the past, specifically, accounting for any potential density-dependent effects on *M* and WAA as well as simulating a high recruitment event. The progression of the strong cohort over time creates an additional productivity scenario over which the performance of MPs can be evaluated.

5.1 FUTURE PROCESS ERROR FOR NAA DEVIATIONS

The process error deviations for the NAA transitions for OM3 were projected following the continuation of the AR1 process without bias correction, such that the deviation (ε) for age a (a > 1) in projection year y was:

$$\varepsilon_{a,y} = \rho_{year}\varepsilon_{a,y-1} \quad \text{(Eqn 6)}$$

The process error deviations for the NAA transitions for OM3b were projected to be the simulation-specific historical deviations from years 1998–2022.

5.2 FUTURE SELECTIVITY

Time-varying selectivity was assumed in the WHAM models, either as annual selectivity (Model 1 and Model 3), or as selectivity with process error on the selectivity parameters (Model 2). One challenge with the time-varying selectivity is deciding how to project future selectivity. One hypothesis for the time-varying selectivity is that when there is a strong recruitment event, selectivity shifts to higher selectivity of smaller fish due to the presence of a strong cohort. This assumption was tested by exploring the relationship between the selectivity parameters (k and a_{50}) and two measures of cohort strength: 1) model estimated NAA-1, and 2) the mean of the model estimated NAA-1 and NAA-2 for the cohort, each evaluated using time lags of 0, 1, and 2 years. A significant linear relationship was detected for k and a_{50} selectivity parameters vs. model estimated NAA-1 for Model 1 and Model 3 which had annual selectivity (Figure 3). Only the k parameter had a significant linear relationship with NAA-1 for Model 2, which had random effects on the selectivity parameters (Figure 3). The regressions were conducted over the time period 1990–2022 where there appeared to be a linear relationship (unique relative to the entire time series) based on inspection of scatterplots. The predictor variables with a time lag of 0-years for NAA-1 had the best fit (i.e., lowest p-value and greatest coefficient of determination) and the strengths of the regressions using the predictor variable NAA-1 were essentially indistinguishable from those using the mean of the model estimated NAA-1 and NAA-2 for the cohort.

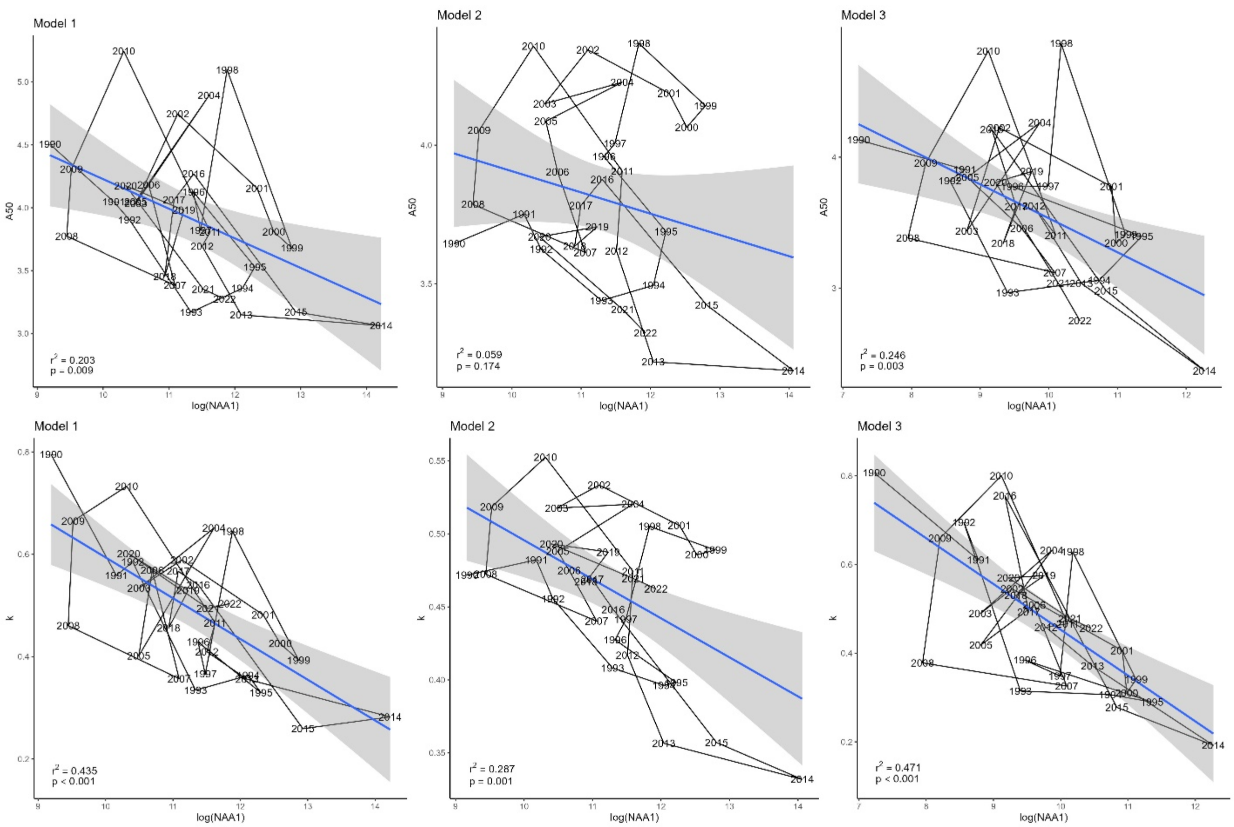

Figure 3. Scatterplot and linear regression of estimates of the selectivity parameters (A_{50} and k) vs log(NAA1) for each model from 1990–2022. The grey shading represents the 95% confidence interval.

5.3 FUTURE NATURAL MORTALITY RATE (*M*)

Future *M* values for OM1 and OM2 were projected using the model-estimated relationship between *M* and the DFO survey NAA1+ (Figure 4), lagged by one year, with the projected DFO survey NAA1+ estimated with observation error as described in Section 7.2. Future *M* for OM3 was assumed fixed at 0.2.

The future *M* values for OMs 1b, 2b, and 3b were projected to be the simulation-specific historical *M* values from years 1998–2022.

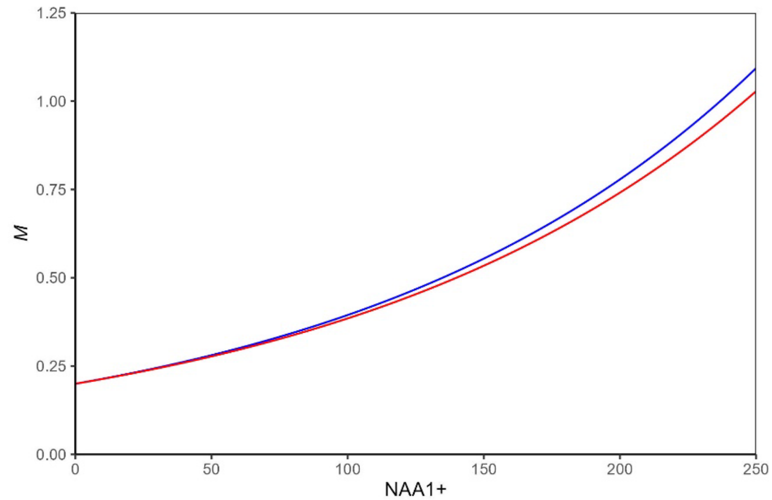

Figure 4. Plot of the model-estimated relationship between natural mortality rate (M) and the DFO survey numbers-at-age 1+ (NAA1+), lagged by one year for Model 1 (blue) and Model 2 (red).

5.4 FUTURE RECRUITMENT

Future recruitment (age of recruitment in the OMs is age 0) was projected for OMs 1–3 by resampling the recruitment deviations from the historical period, excluding the recruitment from 2013. To account for uncertainty in recruitment dynamics below the minimum observed SSB, a hockey stick (HS) stock recruitment relationship (SRR) was assumed with a change point at the minimum observed SSB for each simulation and the median historical recruitment for that simulation. The blade of the hockey stick was a line segment joining the change point and the origin (Figure 5). This SRR was projected by multiplying the recruitment deviation for year y by a factor of $\frac{SSB_y}{SSB_{min}}$ when the current year SSB (SSB_y) was below the minimum historical SSB (SSB_{min}).

Future recruitment for OMs 1–3 was therefore based on the assumption that future recruitment will be consistent with historical recruitment, with the exception of exceptionally large recruitments (like 2013). These large recruitment events were considered in the second set of OMs. The assumption of recruitment declining linearly from median recruitment below the minimum observed SSB to the origin is consistent with the assumption used for ICES stock type 5 (i.e., stocks showing no evidence of impaired recruitment with no clear relation between stock and recruitment) (ICES 2021).

Future recruitment for OMs 1b, 2b, and 3b was projected to be the simulation-specific historical recruitments from years 1998–2022 (Table 4).

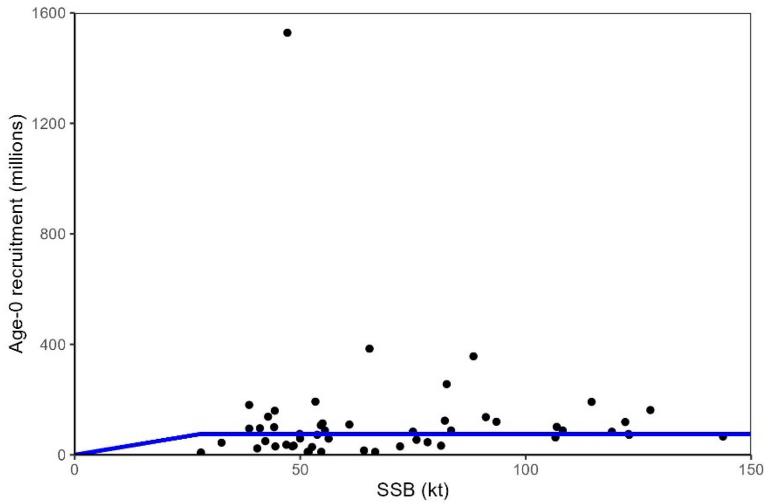

Figure 5. The hockey stick stock-recruitment relationship (blue line) for one simulation for OM1, with the change point at the minimum historical spawning stock biomass (SSB) and median historical recruitment.

5.5 FUTURE GROWTH

An exploration of relationships between von Bertalanffy growth parameters (estimated by cohort) and measures of cohort strength were explored similar to the exploration for selectivity (Section 5.2) but no significant relationships were detected.

For OMs 1–3, future stock WAA was projected as a model-estimated WAA by region (BoF vs. SS), weighted by survey catches. This was estimated specifically by modeling LAA using a two-parameter von Bertalanffy growth model from mean LAA values from 2018–2022 (Figure 6). The LAA was converted to WAA (Figure 7) using the mean weight-length relationship parameters from the DFO survey. The overall stock WAA (Figure 7) was estimated as the weighted mean WAA by region, using weights proportional to the mean NAA by region from 2018–2022.

For OMs 1b–3b, stock WAA was projected using a model-estimated LAA for cohorts up to 2022 (Section 5.5.1; Figure 8; Figure 9) converted to stock WAA using the DFO survey weight-length relationship, then reverting to historical stock WAA over the last 25 year period to fill in the remaining WAA values (Figure 10). Using the model-estimated LAA for recent cohorts provided a smoother transition from terminal year WAA into the projected stock WAA and allowed the continuation of slower growth in recent cohorts to continue into the projection period for OMs 1b–3b (Figure 10). The projected catch WAA for all OMs was estimated as the average ratio of catch WAA to stock WAA over the last 25 years (Figure 11) multiplied by the model-estimated stock WAA for OMs 1–3 and for OMs 1b–3b for cohorts up to 2022. The remaining catch WAA values for OMs 1b–3b were filled in using the historical catch WAA over the last 25 year period (Figure 11).

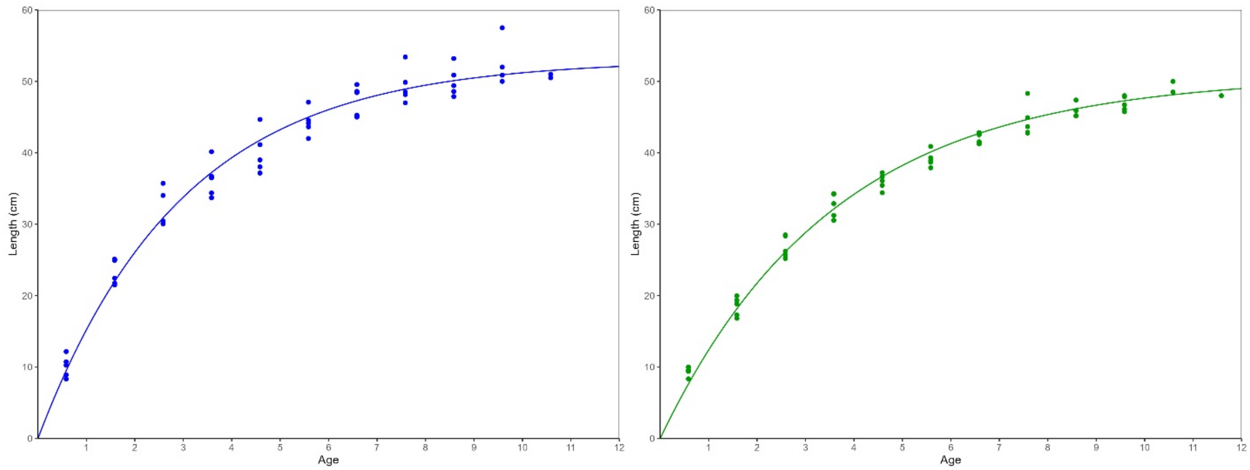

Figure 6. Mean DFO survey length-at-age (adjusted for survey timing in month 7) with a two-parameter von Bertalanffy growth model for Bay of Fundy (blue; left panel) and Scotian Shelf (green; right panel) using data from 2018–2022.

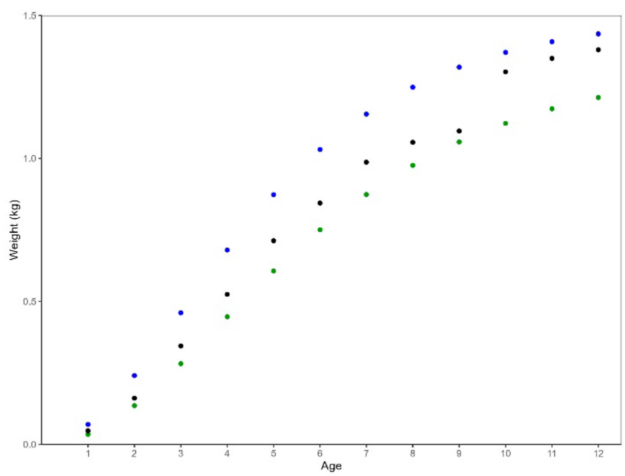

Figure 7. Estimated stock weight-at-age (WAA) for Bay of Fundy (blue) and Scotian Shelf (green) estimated using length-at-age models from Figure 6 and the survey weight-length relationships. The stock WAA (black) is estimated as the weighted average WAA, with weights proportional to the survey NAA by region.

5.5.1 Length-at-age (LAA) model

Future stock LAA for cohorts with at least one observation in the historical data were predicted using a state-space model to account for initial growth for recent cohorts following a similar approach to Gulf of Maine Haddock (C. Perretti, NFSC, pers. comm.). The model predicts annual LAA for stock (DFO survey) data. Length-at-age for cohort c ($L_{a,c}$) was estimated using a two parameter von Bertalanffy growth model:

$$\log\left(L_{a,c}\right) = \log\left(L_{\infty,c}\right) + \log\left(1 - e^{-k_c a}\right) + \varepsilon_{a,c} \; , \; \varepsilon_{a,c} \sim N(0, \sigma_L^2) \qquad \text{(Eqn 7)}$$

with parameters L_∞ and k and error term $\varepsilon_{a,c}$. An AR1 correlation process model was assumed for the von Bertalanffy growth parameters:

$$k_c | k_{c-1} \sim N(\varphi_k, k_{c-1}, \sigma_k^2) \qquad \text{(Eqn 8)}$$

$$L_{\infty,c}|L_{\infty,c-1} \sim N(\varphi_{L_\infty}, L_{\infty,c-1}, \sigma^2_{L_\infty}) \quad \text{(Eqn 9)}$$

Observations with missing lengths were added to the matrix of lengths by age and year for the projection years that contain ages in the historical cohorts (e.g., ages 1 to 12+ in projection year 1 correspond to cohorts 2022–2011, and ages 2 to 12+ in projection year 2 correspond to cohorts 2022–2012). Estimates of lengths for these missing values were estimated by the model by treating the missing values as random effects (Figure 8). Models were fit using RTMB (Kristensen 2024).

The model-estimated LAA values by cohort (Figure 9) were converted to estimates of WAA using annual weight-length relationships from the DFO survey (Barrett and Barrett 2025). The weight-length relationship for 2022 was used to estimate WAA for future years.

Additional models were explored but did not converge. These models included estimating the weight-length relationships in the same LAA model, and estimating a catch LAA with a shared L_∞ parameter with the survey LAA model.

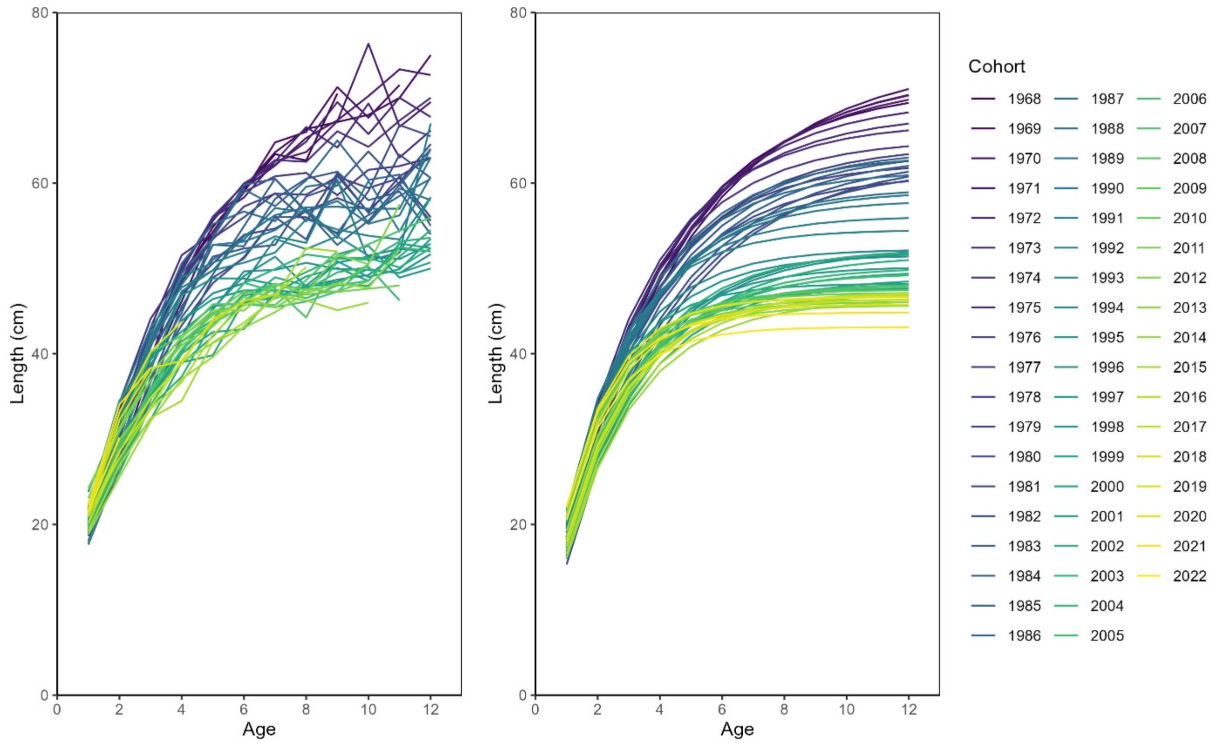

Figure 8. Mean empirical (left panel) and model-estimated (right panel) length-at-age by cohort for 4X5Y Haddock.

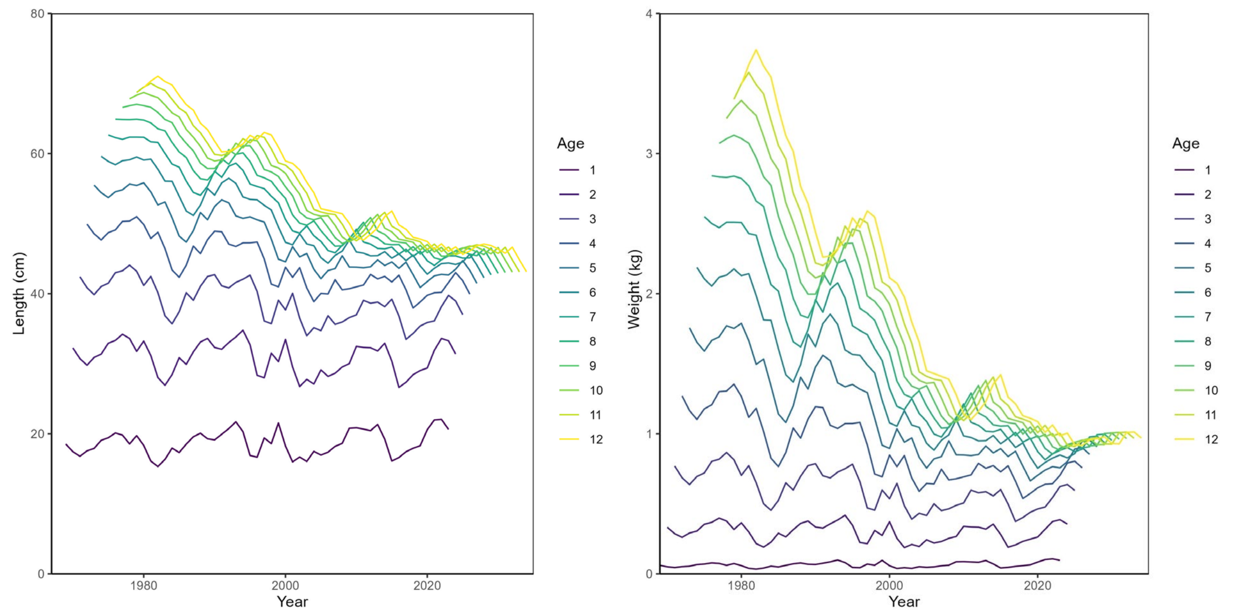

Figure 9. Model-estimated length-at-age (LAA; left panel) and weight-at-age (converted from LAA and survey weight-length relationship; right panel), plotted by year to illustrate the projected LAA and WAA for cohorts up to 2022.

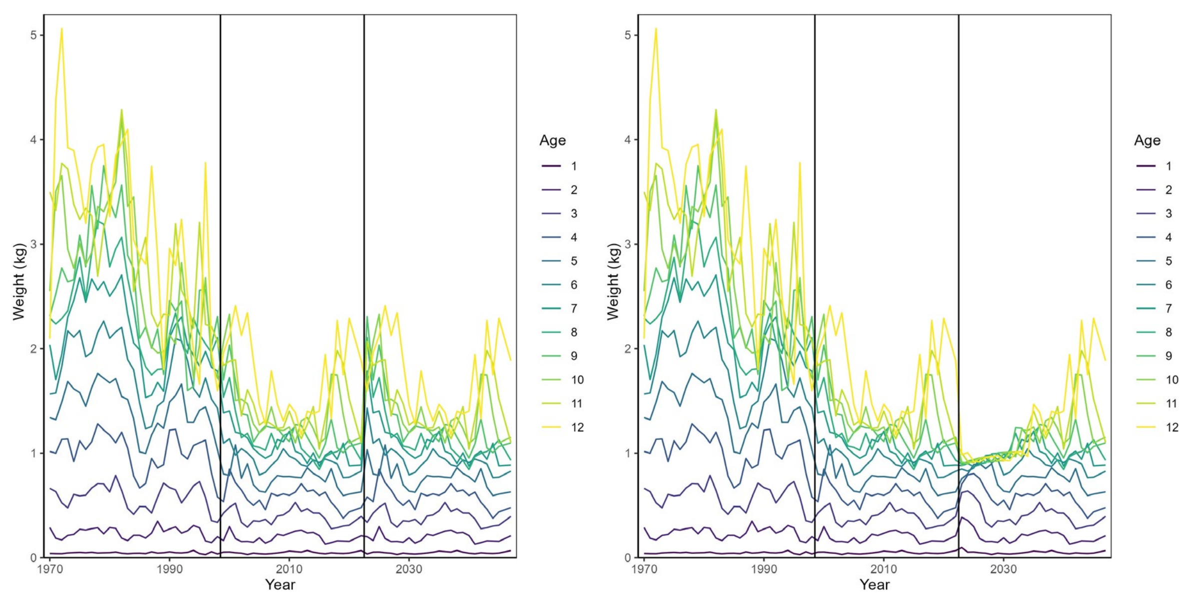

Figure 10. Empirical historical weight-at-age (WAA) and projected WAA assuming 1998–2022 WAA (left panel) and projected WAA using model-estimated LAA for cohorts up to 2022 then reverting to 1998–2022 WAA to fill in remaining values (right panel).

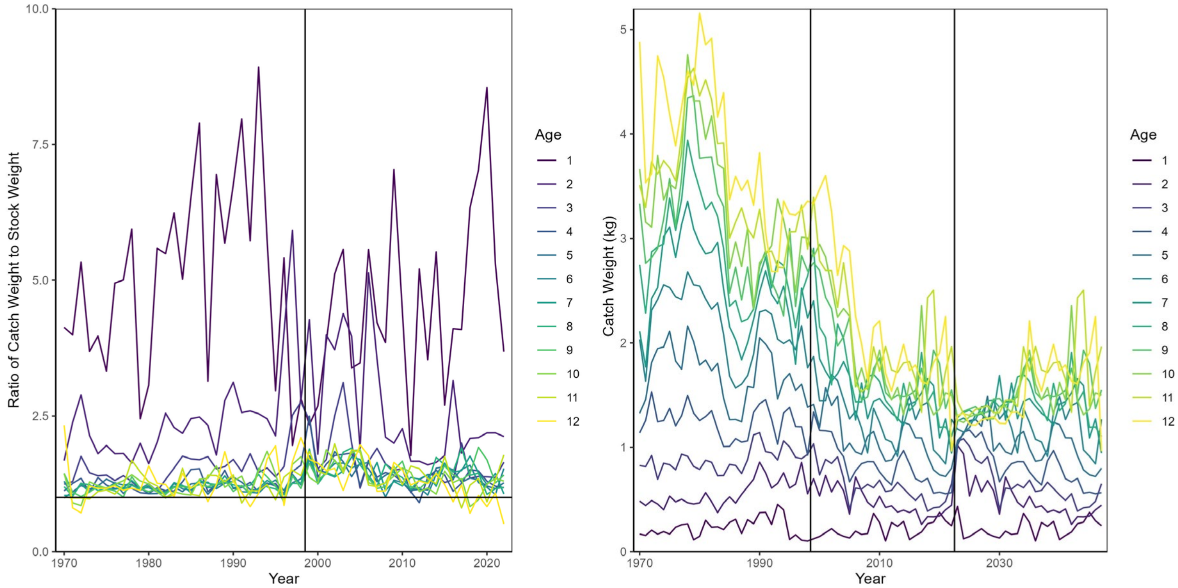

Figure 11. Ratio of empirical historical catch weight-at-age (WAA) to stock WAA (left panel) and projected catch WAA assuming catch WAA is the projected stock WAA (scaled using the mean ratio of catch WAA to stock WAA) for cohorts up to 2022 then reverting to 1998–2022 catch WAA to fill in remaining values (right panel).

5.6 FUTURE MATURITY

Limited maturity data were available from the DFO survey since 1985 (Barrett and Barrett 2025). The input data for each model for the 1986–2022 period were the mean maturity-at-age, weighted by the landings by region. Projected maturity was therefore assumed to be the mean maturity over the last 5 historical years for OMs 1–3 (Table 3).

Future maturity-at-age for OMs 1b, 2b, and 3b was projected to be the simulation-specific historical maturity-at-age from years 1998–2022 (i.e., very similar to the last 5 historical years).

Table 3. Future maturity-at-age of 4X5Y Haddock for operating models 1–3.

Age	0	1	2	3	4	5	6	7+
Proportion mature	0	0.061	0.320	0.795	0.966	0.996	0.999	1.000

5.7 WEIGHTING OF OPERATING MODELS

The primary uncertainty among the set of three WHAM models is whether the population dynamics are characterized by time-varying *M*, using the DFO survey NAA1+ as a covariate (OM1 and OM2) or characterized using random effects on the NAA transitions (OM3). For the evaluation of the performance of MPs, equal weight was placed on each uncertainty. Equal weight was similarly placed on the uncertainty in the parameterization on selectivity (annual fishery selectivity or random effects on fishery selectivity parameters) for OM1 and OM2, providing an OM weighting of 0.25, 0.25, and 0.5 for OMs 1, 2, and 3, respectively.

For the evaluation of the performance of MPs across OMs 1b, 2b, and 3b, the same OM weighting of 0.25, 0.25, and 0.5, respectively, was used.

Table 4. Projection assumptions for each operating model (OM).

OM Name	WHAM model for historical dynamics	Future process error for NAA deviations	Future selectivity	Future *M*	Future recruitment deviations	Future weight-at-age	Future maturity-at-age
OM1	Model 1	–	Selectivity/recruitment relationship[1]	Estimated[2]	Resampled[3]	Estimated using last 5 years[5]	Mean of last 5 years
OM2	Model 2	–	Continuation of AR1 process correlation	Estimated[2]	Resampled[3]	Estimated using last 5 years[5]	Mean of last 5 years
OM3	Model 3	Continuation of AR1 process correlation	Selectivity/recruitment relationship[1]	0.2	Resampled[3]	Estimated using last 5 years[5]	Mean of last 5 years
OM1b	Model 1	–	Historical selectivity from 1998–2022	Historical *M* from 1998–2022	Historical recruitment deviations[4]	Estimated for cohorts up to 2022[6]	Historical maturity-at-age[7]
OM2b	Model 2	–	Historical selectivity from 1998–2022	Historical *M* from 1998–2022	Historical recruitment deviations[4]	Estimated for cohorts up to 2022[6]	Historical maturity-at-age[7]
OM3b	Model 3	Historical deviations from 1998–2022	Historical selectivity from 1998–2022	Historical *M* from 1998–2022	Historical recruitment deviations[4]	Estimated for cohorts up to 2022[6]	Historical maturity-at-age[7]

[1]Estimated from relationship between selectivity parameters and recruitment from 1990–2022 (see Figure 3).
[2]Estimated from WHAM model relationship between *M* and survey NAA1+
[3] Resampled from historical without 2013 age-0 recruitment, adjusted using a hockey stick SRR
[4] Historical recruitment deviations from 1998–2022
[5] Model-estimated using LAA and weight-length relationships from last 5 years
[6] Model-estimated for cohorts up to 2022 and historical weight-at-age from 1998–2022
[7] Historical maturity-at-age from 1998–2022
Dash indicates not applicable for this model; NAA = numbers-at-age; AR1 = first order autoregression; SRR = stock-recruitment relationship; LAA = length-at-age; *M*= natural mortality rate; WHAM= Woods Hole assessment model

5.8. CONSIDERATION OF ADDITIONAL OPERATING MODELS

An additional set of OMs was explored at the request of the reviewers during the peer-review meeting to capture a scenario with low productivity (average recruitment with no large recruitment event similar to the 2013 cohort), but while still maintaining a chance to observe high *M*. The OMs were defined as OMs 1c and 2c with the same projection assumptions as OMs 1 and 2, respectively, but with *M* randomly sampled from historical *M* values. Based on this exploration, it was determined that the higher *M* values (not dependent on density) have a large impact on catch advice and while the productivity scenario could potentially occur in the future, it is not consistent with what has been observed in the past so these OMs were not included in the MP evaluation. Detection of a change in productivity, consistent with low recruitment and high *M*, may take several years to detect and this situation would be captured under an exceptional circumstance (see Section 12).

6. REFERENCE POINTS

The selection of reference points for 4X5Y Haddock was consistent with DFO's fishery decision-making framework incorporating the precautionary approach (PA policy) (DFO 2009) and the recently defined guidelines for implementing the fish stocks provisions (FSPs) in the *Fisheries Act* (DFO 2022). The primary components of the PA policy are reference points and stock status zones (healthy, cautious, and critical), harvest strategy and harvest decision rules, and the need to take into account uncertainty and risk when developing reference points and developing and implementing decision rules (DFO 2009). The FSPs establish legally binding obligations on DFO to:

- manage prescribed major fish stocks at levels necessary to promote sustainability (section 6.1);
- develop and implement rebuilding plans for prescribed major fish stocks that have declined to or below their LRP to grow the stock above that point (section 6.2); and
- prescribe in regulation the fish stocks to which the provisions will apply (section 6.3).

A clear definition of the LRP and its role is important because the LRP is the trigger for a rebuilding plan under the FSP. The LRP represents the upper bound of stock states that should be avoided in order to prevent serious harm to the stock and is the boundary between the Critical and Cautious zones of the PA policy (DFO 2023b). The USR marks the boundary between the PA policy's Cautious and Healthy stock status zones. When biomass falls below the USR, the removal rate at which fish are harvested should be reduced in order to avoid the LRP (DFO 2009). A removal reference (RR) and an optional target reference point (that can differ from the USR) are also reference points defined in the PA policy (DFO 2009). It is the role of DFO Science to define the LRP. The USR is defined by DFO Resource Management, in consultation with stakeholders. As such, DFO Science will provide advice on one or more candidate USR(s), to be considered during a meeting of the Scotia-Fundy Groundfish Advisory Committee. The RR in the PA policy is defined as the maximum acceptable removal rate for the stock (DFO 2009). The RR can therefore be interpreted as the removal rate defined by the MP that will be selected using the MP evaluation framework, since this MP will specify a removal rate that meets the conservation objectives of the PA policy and the FSPs.

Two fishing mortality (*F*) reference points were defined for the stock to support the investigation of MPs. These were $F_{lim} = F_{crash}$, defined as a constant *F* that would lead the stock to crash in the long-term (see section 6.3) and $F_{ref} = F_{PA}$, a constant *F* that results in maximum yield (consistent with the concept of *F* at maximum sustainable yield) while maintaining a high probability of avoiding the LRP in the long-term. This is consistent with the ICES (2021) definition of a F_{PA}, which is an F “below which exploitation is considered to be sustainable, having accounted for estimation uncertainty”.

6.1 ROLES OF REFERENCE POINTS

Reference points can serve multiple roles, and the role of the LRP can get confounded in a framework with multiple models and where the LRP may also appear in performance metrics and in HCRs. We outline three roles of reference points in this framework, following the DFO guidance for stock status and LRPs under the FSPs (Barrett et al. 2024).

Role 1: Component of Management Objectives and Performance Metrics

The reference points can serve a role in management objectives. For example, the objective “Maintain the stock above the LRP” (Table 1) includes the LRP. The LRP in this context is defined as the OM- and simulation-specific LRP. There are multiple ways to quantify the probability of dropping below the LRP in a performance metric (ICES 2013). In this framework, the probability of SSB falling below the LRP, was defined as the average annual P(SSB < LRP) over the specified projection period. Performance metrics were calculated across all weighted OMs (see weighting approach in Section 5.7), consistent with the Commission for the Conservation of Southern Bluefin Tuna approach defined in ICES (2020).

Role 2: Metric of Stock Status

The historical dynamics of the fishery were characterized by three stock assessment models: Model 1, Model 2, and Model 3. Each model has its own set of reference points (i.e., LRP, USR). The stock status for the purpose of implementing the FSPs is estimated using an ensemble approach, using the model weights outlined in Section 5.7, namely, 0.25, 0.25, and 0.5 for Model 1, Model 2, and Model 3, respectively. This approach is consistent with that used for Sablefish *Anoplopoma fibria* (DFO 2020) and consistent with DFO guidance for stock status and LRPs under the FSPs (DFO 2023b, Barrett et al. 2024). The indicator of stock status is the model-weighted ratio of the SSB to the LRP across simulations. The probability of falling below the LRP is estimated as the model-weighted proportion of simulations with SSB below the LRP. A breach of the LRP is defined when 50% of more of the simulations have an SSB < LRP, following the guidelines for implementing the FSPs in the Fisheries Act (DFO 2022).

The objective of this framework is to identify an MP to provide future catch advice. The data source in the future that is used to inform stock size (and therefore the catch advice using the MP) is the DFO survey index. While the MP is used to provide future catch advice, a stock assessment model is not fit each year in the future. A stock status monitoring trigger is therefore needed in interim years between modeling frameworks to inform whether the LRP may have been breached. The indicator is the three-year moving average of the DFO survey index with a trigger for an assessment being a value of the three-year moving average below the weighted average (across OMs) and q-adjusted LRP. This trigger is outlined in the exceptional circumstances (Section 12) for this MP evaluation framework.

Role 3: Operational Control Points

In the DFO PA policy, the LRP and USR serve as operational control points in the default policy harvest control rule (HCR). Although the model-estimated LRP and USR can be used to inform control points of a HCR to be tested in this framework, there is no requirement to have control points of a HCR based on the LRP and USR. To avoid confusion, the terms “upper control point” and “lower control point” will be used to describe the points in an HCR, instead of using any mention to the reference points. This concept is discussed further in Section 9.

6.2 LIMIT REFERENCE POINT (LRP)

Indicator

SSB was selected as the indicator to represent the reproductive capacity of the stock. The models estimate annual SSB (at a proportion of year = 0.4, to approximate the timing of spawning) and uncertainty in SSB is captured by the 100 simulations. Annual fecundity data are not available for the stock; however, fecundity data were available for the neighboring eastern Georges Bank (EGB) Haddock stock in 2019 (Acheson 2020) and these data were used in this document to evaluate the assumption of isometry (i.e., fecundity is proportional to body weight; Marshall et al. 2021). The slope of the regression of log(fecundity) on log(body weight) was 1.21 with a 95% confidence interval that does not overlap zero (Figure 12), indicating that fecundity is not proportional to body weight.

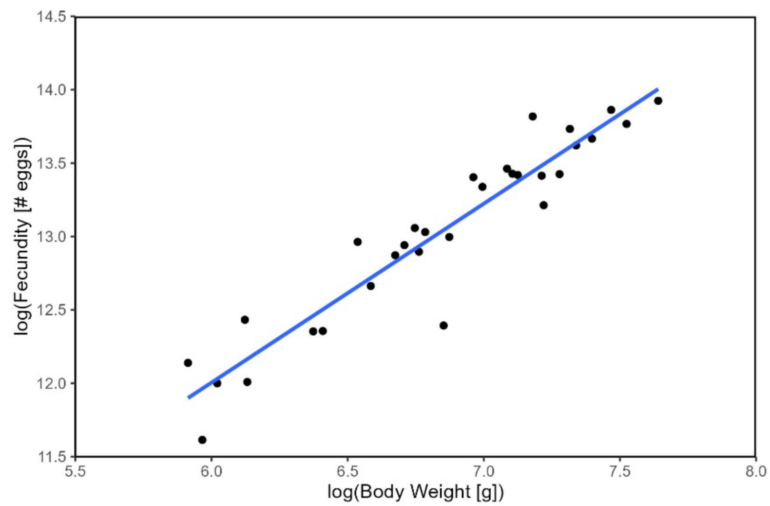

Figure 12. Linear regression of log(fecundity) on log(body weight) for eastern Georges Bank Haddock collected in 2019 with slope of 1.22 with a 95% confidence interval of (1.06, 1.38).

Methods

LRPs are commonly defined in terms of a proportion of the biomass at maximum sustainable yield (B_{MSY}) or a proportion of unfished biomass (B_0) or a proxy for these (Sainsbury 2008, DFO 2023b). Estimates of B_{MSY} are based on a SRR, and in the absence of a SRR, a proxy of the equilibrium biomass from fishing at $F_{X\%}$ (i.e., the F that results in a spawning potential ratio (SPR) of X%) may be used. For example, for medium productivity stocks (MF 2011), a B_{MSY} proxy of the equilibrium biomass from fishing at $F_{40\%}$ is commonly used (e.g., Georges Bank Haddock, NFSC In Prep[1]; Gulf of Maine Haddock, NFSC In Prep[2]). The SPR (Gabriel et al. 1989) is defined as the SSB-per-recruit (φ) at a given constant, long-term F divided by the φ at long-term F = 0 (i.e., SPR = ${}^{\varphi_F}/_{\varphi_0}$).

Challenges with defining equilibrium biomass-based reference points for 4X5Y Haddock are that the stock has experienced changes in productivity over time (e.g., declining growth (Figure 9), temporal changes in M (Model 1 and 2, Figure 13) or temporal changes in survival (i.e., NAA transition) deviations (Model 3), and spasmodic recruitment (Figure 5)). Equilibrium reference

[2] Northeast Fisheries Science Center (NFSC), In Prep. 2022 Management Track Assessment Report for Gulf of Maine Haddock. US Dept Commer. Northeast Fish Sci Cent Ref Doc.

points were therefore estimated under the productivity scenarios assumed in the long-term projections for OMs 1, 2, and 3 (i.e., productivity defined by the projected weight-at-age, selectivity-at-age, maturity-at-age, *M*-at-age, and in the case of OM3, survival deviations) with median historical recruitment (R_0) but assuming a HS SRR that reduces recruitment below the minimum observed SSB (Section 5.4). Given the variability in annual *M* obtained in long-term projections (Figure 13), a long-term equilibrium *M* for OMs 1 and 2 was estimated as the mean *M* over projection years 26–50 for each OM (Figure 13).

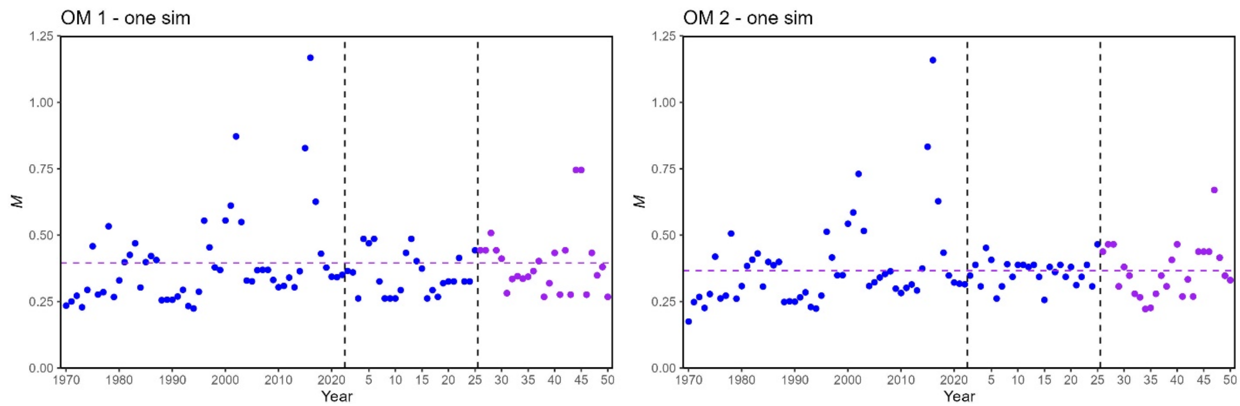

Figure 13. Historical natural mortality rate (M) for the times series (years 1970–2022) and for 50-year projections (years 1–50 on x-axis) under F = 0 for one simulation for OMs 1 and 2. The mean value over projection years 26–50 (purple) is plotted as a horizontal line in each panel.

The ICES guidance for reference point estimation for a stock with occasional large year-classes like Haddock (ICES stock type 1; ICES 2021) is to define the LRP as B_{lim} which is equal to the lowest SSB where large recruitment is observed. The stock-recruitment pattern for Haddock is also consistent with ICES stock type 5: “stocks showing no evidence of impaired recruitment or with no clear relation between stock and recruitment”, where LRP is the lowest observed SSB (B_{loss}). Consistent with these approaches, a common approach used by DFO is B_{recover} defined as the minimum biomass that produced the recruitment that led to stock recovery (Barrett et al. 2024).

Candidate references points that were explored for estimation of an LRP were:

- a proportion (0.2) of B_0
- a proportion (0.4) of the equilibrium biomass from fishing at $F_{40\%}$ ($B_{F40\%}$, a B_{MSY} proxy)
- the lowest SSB where large recruitment is observed (B_{lim})
- the lowest observed SSB (B_{loss})
- the minimum biomass that produced the recruitment that lead to stock recovery (B_{recover}).

Under equilibrium assumptions, equilibrium biomass values were estimated as the product of φ_F and R_0, where φ_F is defined as:

$$\varphi_F = \sum_{a=a_{rec}}^{a=a_{max}} (l_a w_a m_a) \quad \text{(Eqn 10)}$$

where a_{rec} is the age at recruitment (age 1) and where a_{max} is the maximum age (plus group of 12), and w_a is the weight-at-age a, m_a is the maturity-at-age a, l_a is the survivorship-at-age a defined as:

$$l_a = \begin{cases} 1, & a = a_{rec} \\ l_{a-1}e^{-(M_{a-1}+Fv_{a-1})}, & a_{rec} < a < a_{max} \\ \frac{l_{a-1}e^{-(M_{a-1}+Fv_{a-1})}}{1-e^{-(M_{a-1}+Fv_{a-1})}}, & a = a_{max} \end{cases} \quad \text{(Eqn 11)}$$

For OM1 and OM2, *M* was not in equilibrium in the projection period so l_a was estimated using the mean *M* over projection years 26–50 for the projection of the specified *F* (e.g., see Figure 13).

SPR values were estimated over a range of *F* values in increments of 0.05 over [0, 1]. $F_{40\%}$ was estimated to the nearest hundredth using linear interpolation between these *F* values to obtain the approximate SPR of 40% (Figure 14). An $F_{40\%}$ was estimated to be 0.30 for OM3, but could not be estimated for OM1 and OM2 (i.e., the SPR at F_{lim} was > 0.4) (Figure 14).

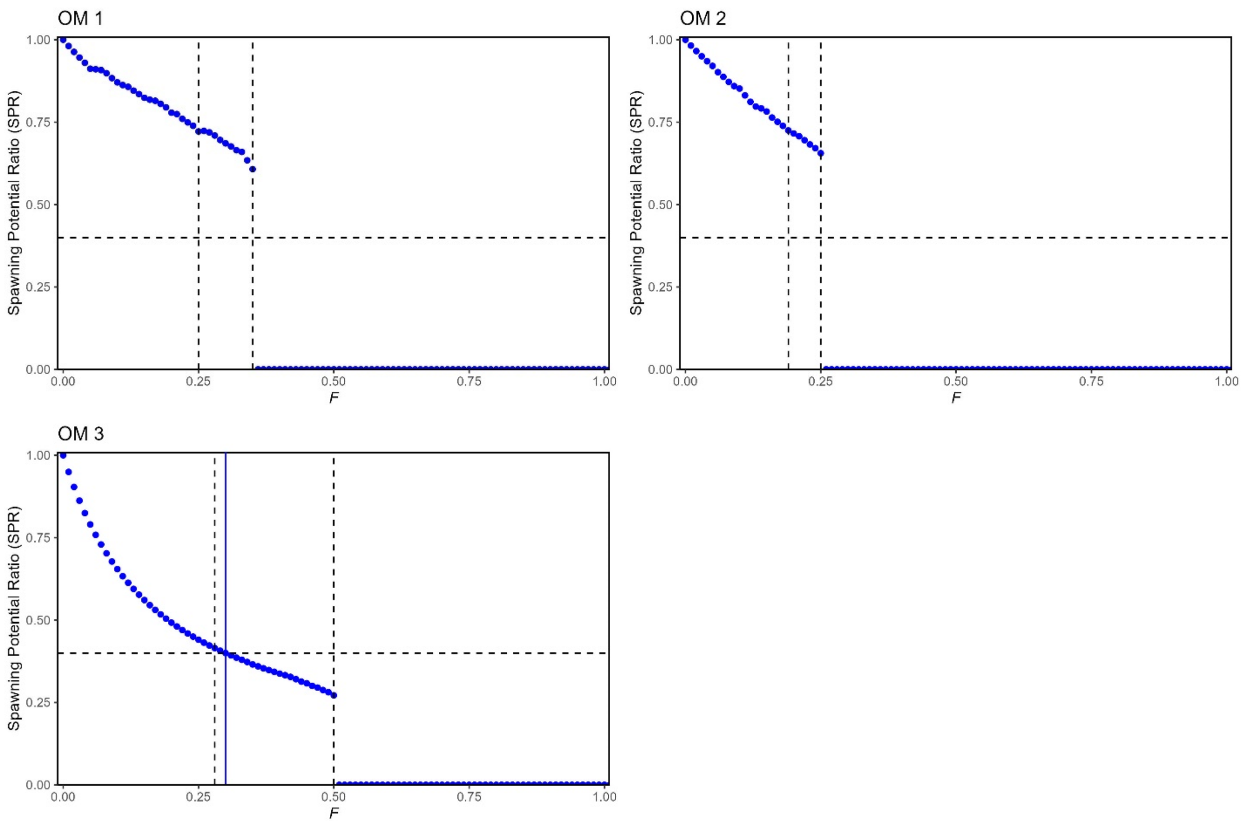

Figure 14. Plot of spawning potential ratio (SPR) vs. fishing mortality rate (F) over a range of F-values for operating models (OM) 1–3. The horizontal reference line is SPR = 0.4 and the vertical dashed reference lines are at the median F_{lim} and at F_{PA}, while the vertical blue reference line for OM3 is $F_{40\%}$.

Two candidate B_{lim} options were explored; the SSB that produced the largest recruitment (2013 for all three OMs, (Figure 15, Table 5), and the minimum SSB associated with the largest 10% of historical recruitments (2013 for OM1 and OM2, 1994 for OM3, Figure 15, Table 5). The SSB that produced the recruitment that led to the most recent increase in SSB was also the 2013 recruitment event so B_{recover} was similarly defined as the SSB in 2013 for all OMs. B_{loss} was the SSB in year 1989 for OM1 and OM2 and the SSB in 1993 for OM3 (Figure 15, Table 5).

To account for the allometric relationship between body weight and fecundity, the SSB required in 2022 to produce the number of eggs produced by B_{loss} in year 1989 (OM1 and OM2) or year

1993 (OM3) was estimated using the model-estimated NAA in each year and the assumed WAA in each year. The ratios of the eggs per unit of biomass in the year associated with B_{loss} to eggs per unit of biomass in 2022 was 2.66, 2.69, and 3.94, for OMs 1, 2, and 3, respectively. This resulted in B_{loss} estimates, adjusting for eggs (B_{loss_eggs}) as outlined in Table 5.

Table 5. Candidate limit reference points (LRPs; medians across 100 simulations) by operating model (OM).

OM	0.2 B_0 (kt)	0.4 $B_{F_{40\%}}$ (kt)	B_{lim} = $B_{recover}$ = (kt)	B_{loss} (kt)	B_{loss_eggs} (kt)
OM1	10.3	-	50.9 [year 2013]	34.2 [year 1989]	91.0
OM2	10.3	-	51.1 [year 2013]	35.3 [year 1989]	95.1
OM3	8.22	6.62	20.6 [year 2013] 14.5 [year 1994]	12.1 [year 1993]	47.7

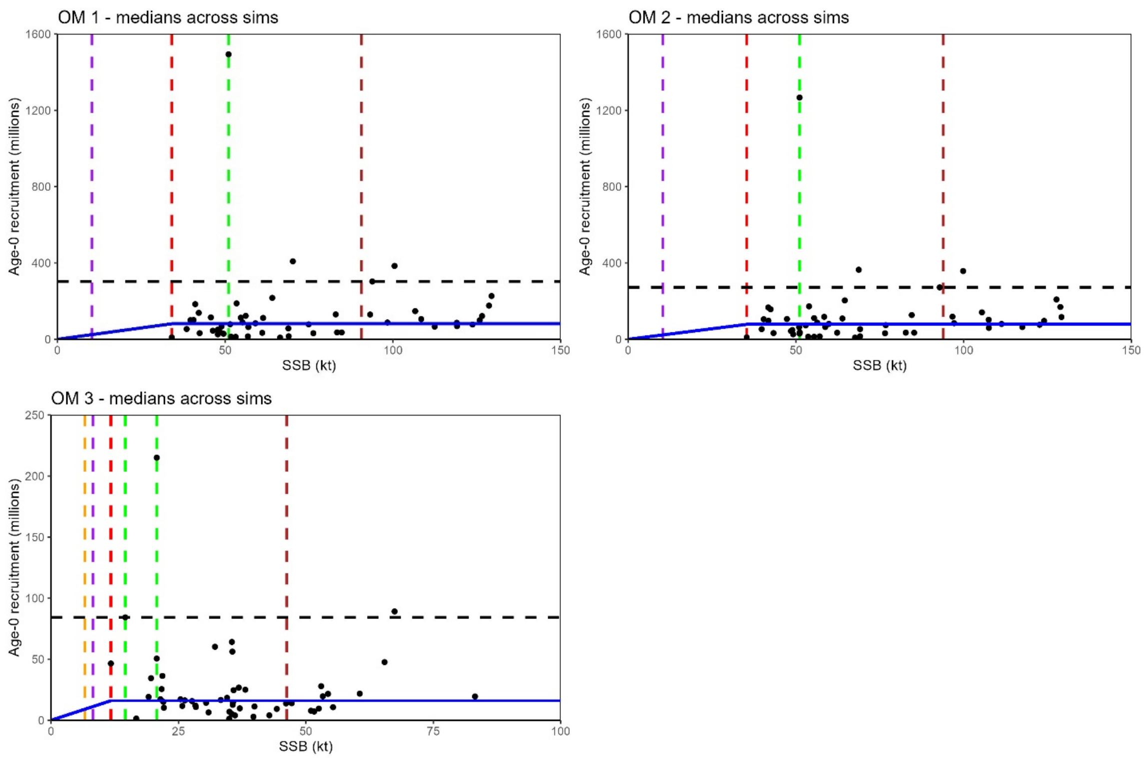

Figure 15. Plots of historical model-estimated recruitment vs. model-estimated SSB (medians across simulations; 1970–2021) for each OM. The hockey stick SRR is plotted in blue. The black horizontal reference line is the 90th percentile of historical recruitments. The vertical reference lines are 0.2 B_0 (purple), B_{loss} (red), B_{lim} = $B_{recover}$ (green; two lines for OM3), 0.4 $B_{F40\%}$ (orange; OM3 only), and B_{loss_eggs} (brown).

The estimate of B_0 has greater uncertainty in OMs 1 and 2, given the variability in the projected *M* (Figure 13), and the estimate falls below the minimum observed SSB, where recruitment dynamics are uncertain, despite the assumptions of the HS SRR used in the projections. The B_{lim} was the SSB associated with high recruitment, where maximum recruitment occurred the same year as the minimum SSB of the highest 10% of historical recruitments (i.e., SSB in

2013). Given there are multiple observations of recruitment at SSB values below SSB in 2013, this B_{lim} estimate appears overly conservative. B_{loss} was therefore selected to be used as the LRP for the models, and this choice is supported by the ICES guidance for reference points (2021) and the DFO guidance for LRPs under the FSPs (DFO 2023b). The concept of B_{loss} is linked to serious harm to the productivity of the stock by representing a proxy for recruitment overfishing, reflecting the uncertainty in population dynamics (e.g., recruitment, *M*, survival deviations) at stock sizes below the minimum observed. The estimates of B_{loss_eggs} were based on a fecundity/body weight relationship for Eastern Georges Bank Haddock and were relatively high. Further research would be needed on the fecundity, egg size, and survival of eggs in order to use fecundity as the indicator of reproductive potential for 4X5Y Haddock.

6.3 FISHING MORTALITY REFERENCE POINTS

A limit on fishing mortality (F_{lim}) was defined as the equilibrium *F* that gives φ_F of B_{loss}/R_0. This F_{lim} is consistent with F_{crash} (Shelton and Rice 2002), the equilibrium fishing mortality rate that would lead the stock to crash in the long-term since an $F > F_{lim} = F_{crash}$ would have a φ_F replacement line that doesn't intersect the hockey stick SRR (Figure 16). The F_{lim} estimates are simulation-specific and the median F_{lim} values are provided in Table 6.

Given the nature of the hockey stick SRR, there is no compensation in the SRR at low SSB. Therefore, a B_{MSY} and F_{MSY}, cannot be estimated for practical purposes.

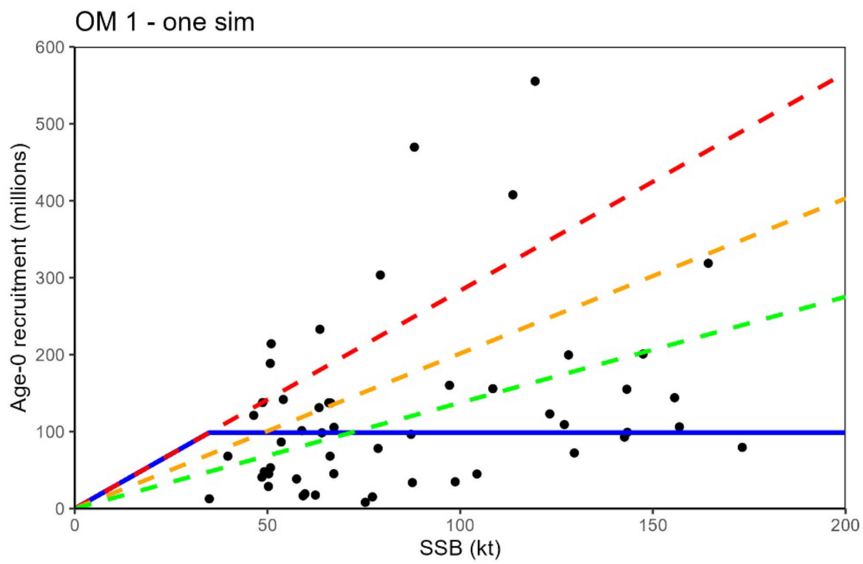

Figure 16. Scatterplot of model-estimated recruitment vs. model-estimated spawning stock biomass (SSB) (1970–2021) for one simulation for OM1 with the hockey stick stock recruitment relationship (blue), SSB-per-recruit (φ_F) replacements lines for $F_{lim} = F_{crash}$ (red), F_{PA} (orange), and F=0 (green). The y-axis is truncated at 600 such that the observation for 2013 is not shown.

Long-term projections over a range of constant *F*-values were conducted using the assumptions described in Section 6.2 for LRPs. The relationship between the SPR and equilibrium *F* are shown for OMs 1–3 in Figure 14. With the higher (i.e., > 0.2) values of *M* in Models 1 and 2, an SPR of 40% cannot be attained before *F* reaches F_{lim} (Figure 14); however, a median $F_{40\%}$ across simulations is estimated for Model 3 as 0.30 (Figure 14).

F_{PA} is defined by ICES as the equilibrium *F* that would lead to SSB > LRP with a 95% probability, while accounting for the expected stochastic variability in the biology and the fishery (ICES 2021). F_{PA} was therefore estimated as the equilibrium *F* that resulted in a probability (combined across 100 simulations and projection years 26–50) of being above the simulation-specific LRP. This probability was estimated based on a linear interpolation of probabilities for the long-term projection *F*-values over the range [0, 1] in increments of 0.05 (Figure 17).

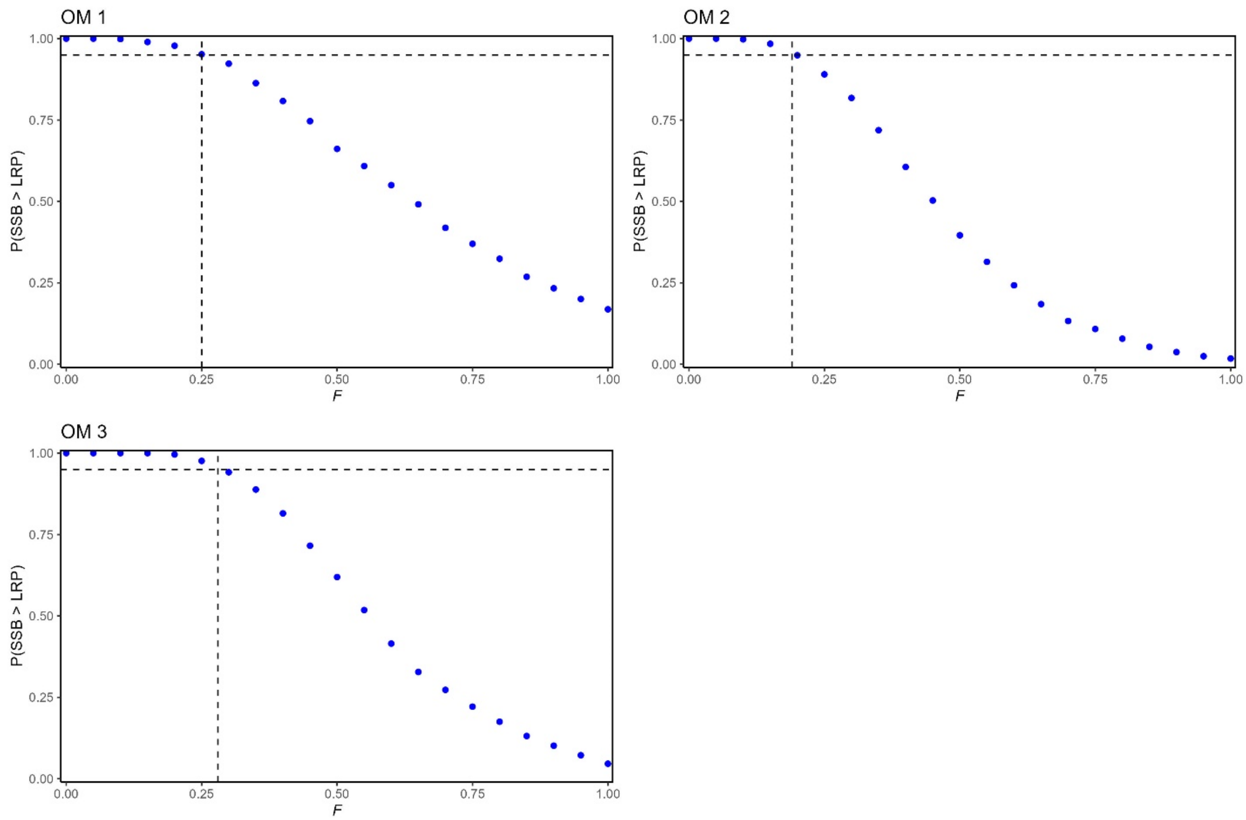

Figure 17. Plots (by OM) of the probability of spawning stock biomass exceeding the limit reference point, P(SSB > LRP), in projection years 26–50 over a range of long-term projection F values [0, 1] in increments of 0.05 for 100 simulations. The probabilities represent the proportion of year/simulation SSB values that exceed the simulation specific LRP.

6.4 PROPOSED UPPER STOCK REFERENCE POINT (USR)

The proposed USR was defined as B_{PA}, the stock size above which the stock is considered to have full reproductive capacity, having accounted for estimation uncertainty (ICES 2021). B_{PA} was estimated for each simulation for each OM as the equilibrium biomass from fishing at F_{PA}, specifically as the product of $\varphi_{F_{PA}}$ and R_0 which is equivalent to the intersection of the $\varphi_{F_{PA}}$ replacement line and the SRR (e.g., Figure 14).

Table 6. Limit reference points (LRPs), proposed upper stock reference points (USRs), fishing mortality reference points (F_{lim}, F_{ref}, and $F_{40\%}$), and probability of terminal year SSB (SSB_{2022}) exceeding the LRP and USR by operating model (OM).

OM	LRP = B_{loss} (kt)	USR = B_{PA} (kt)	F_{lim} = F_{crash}	F_{ref} = F_{PA}	$F_{40\%}$	P(SSB_{2022} > LRP)	P(SSB_{2022} > proposed USR)
OM1	34.2	52.2	0.36	0.25	–	0.94	0.10
OM2	35.3	54.4	0.25	0.19	–	0.97	0.15
OM3	12.1	23.4	0.50	0.28	0.30	1.00	0.56
Weighted	23.4	38.3	0.40	0.25	–	0.98	0.34

Notes: Dash indicates not estimated.

6.5 ESTIMATION OF STOCK STATUS

The indicator of stock status was defined as the model-weighted ratio of the SSB to the LRP and USR across simulations (Table 6). The probability the stock is above the LRP is estimated as the model-weighted proportion of simulations with SSB above the LRP as 0.98, such that the stock is not in the critical zone of the DFO PA policy framework (DFO 2009). The probability the stock is above the proposed USR is estimated as 0.34, placing the stock in the cautious zone of the PA policy framework.

7. OBSERVATION MODEL

An observation model was used to simulate the future data collection used in the closed-loop simulation testing of MPs. The two data sources needed in the simulation testing were:

1. Future DFO survey index (to inform on stock size, used in MPs)
2. Future DFO survey NAA1+ (to inform future *M* for OM1 and OM2)

7.1 DFO SURVEY INDEX

The DFO survey index was simulated as the model-estimated stock biomass-at-age, multiplied by the survey selectivity to obtain the model-estimated survey biomass-at-age. This was then q-adjusted by the model-estimated catchability coefficient (q), and observation error was added by randomly sampling an observation error from the distribution of historical observation errors between the annual observed DFO survey index and the model-estimated survey biomass. The same estimated annual observation errors were applied across each OM to allow consistent comparisons across OMs (e.g., if the simulated observation error was 50% for simulation s in year y in OM1, an observation error of 50% was also applied in simulation s in year y for all other OMs).

7.2 DFO SURVEY NUMBERS-AT-AGE (NAA)

The DFO survey NAA were simulated as the model-estimated stock NAA, multiplied by the survey selectivity to obtain the model-estimated survey NAA. To account for observation error, the relationship between the historical annual observed DFO survey NAA1+ ($SNAA_{1+O}$) and the model-estimated survey NAA1+ ($SNAA_{1+E}$) was explored (Figure 18). The relationship was approximately linear on the log-scale so a linear regression with random sampling of an observation from this regression was explored; however, the random samples led to cases of extreme $SNAA_{1+O}$ estimates that exceeded the historical data and lead to extrapolation of the *M* vs. $SNAA_{1+O}$ relationship that was estimated in the WHAM models (see Figure 4), and resulted in extreme estimates of *M*. Extrapolating beyond the range of the *M* vs. $SNAA_{1+O}$ relationship that was estimated in the WHAM models was undesirable, so a resampling approach was taken to obtain a simulated (projected) $SNAA_{1+O}$ as follows:

- If simulated $SNAA_{1+E}$ ≥ 85th percentile of historical $SNAA_{1+E}$, resample historical $SNAA_{1+O}$ values from 3 nearest historical ($SNAA_{1+E}, SNAA_{1+O}$) pairs to obtain a simulated $SNAA_{1+O}$.
- If simulated $SNAA_{1+E}$ < 85th percentile of historical $SNAA_{1+E}$, resample historical $SNAA_{1+O}$ values from 7 nearest historical ($SNAA_{1+E}, SNAA_{1+O}$) pairs to obtain a simulated $SNAA_{1+O}$.
- If simulated $SNAA_{1+E}$ < historical $SNAA_{1+E}$, estimate $SNAA_{1+O}$ from a linear interpolation between the origin and the mean of the ($SNAA_{1+E}, SNAA_{1+O}$) pairs with the lowest $SNAA_{1+E}$ values.

The 85th percentile was chosen as a cutoff based on visual inspection of the $SNAA_{1+O}$ vs. $SNAA_{1+E}$ plot (Figure 18), where the trend and variability in the relationship appear to change. One observation from 1977 (the year with one survey tow that was large) was excluded from the $SNAA_{1+O}$ vs. $SNAA_{1+E}$ relationship for the resampling approach.

By using this resampling approach and restricting the range of $SNAA_{1+O}$ to 0 and the maximum value observed in the historical period, the bounds of the projected *M* values for OM1 and OM2 were 0.2 and the maximum observed *M* in the historical period.

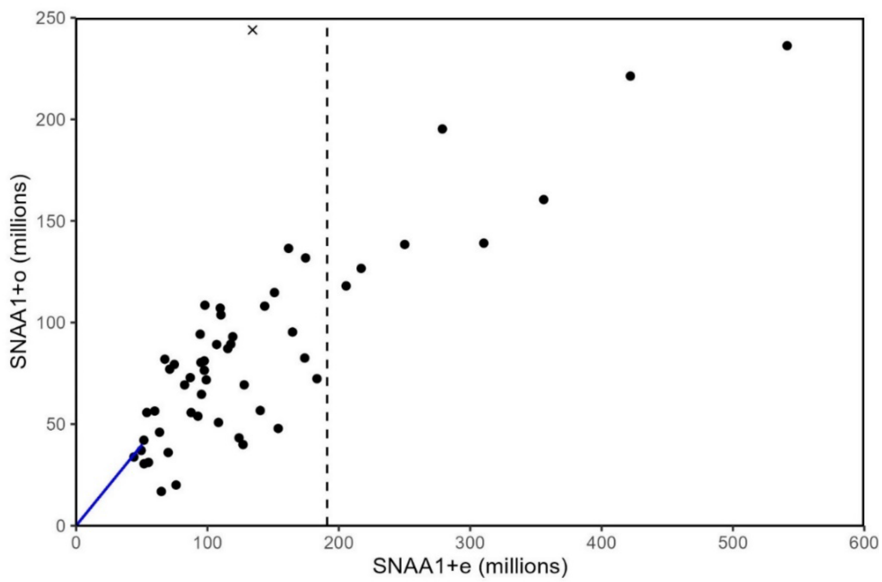

Figure 18. Scatterplot of the historical annual observed DFO survey numbers-at-age (NAA) 1+ ($SNAA_{1+O}$) vs. the model-estimated survey NAA1+ ($SNAA_{1+E}$) for simulation 1 from Model 1. Vertical dashed line is the 85th percentile. The blue line segment is a line joining the origin and the mean paired values of the smallest 5 $SNAA_{1+E}$ values. The outlier from 1977 (driven by a single large survey tow) that was excluded from the resampling approach is plotted as ×.

8. CLOSED-LOOP SIMULATIONS

The MP evaluation framework uses closed-loop simulation in the MSEtool package (Hordyk et al. 2022) in R (R Core Team 2021). The simulations were conducted on each OM/MP combination and involved simulating fishery data for replicate (n = 100) simulations from the

OM. Each simulation varied in the estimated model parameters (sampled from the model parameter variance-covariance matrix), future recruitment (described in Table 4), and future *M* (for OM1 and OM2; described in Table 4). Simulated fishery data were generated for each simulation by applying the observation model with observation error and uncertainty designed to replicate the data collection for the projected DFO survey index and DFO survey index NAA. The simulated survey NAA1+ is used to inform the *M* in the OM for the following year.

A candidate MP is applied to the simulated data (step 2 in Figure 19) and the MP generates a TAC recommendation. The TAC is implemented (step 3 in Figure 19) without implementation error (i.e., the MP recommended catch is removed without error) and the final step of the simulation loop (step 4 in Figure 19) is to remove the catch from the stock (using the projected selectivity-schedule) by updating the OM. This process is repeated until the end of the 25-year projection period (approximately five generations for 4X5Y Haddock, defined as the mean age of sexually mature fish in the terminal model year). At the end of the projection period, the performance of the MPs can be evaluated.

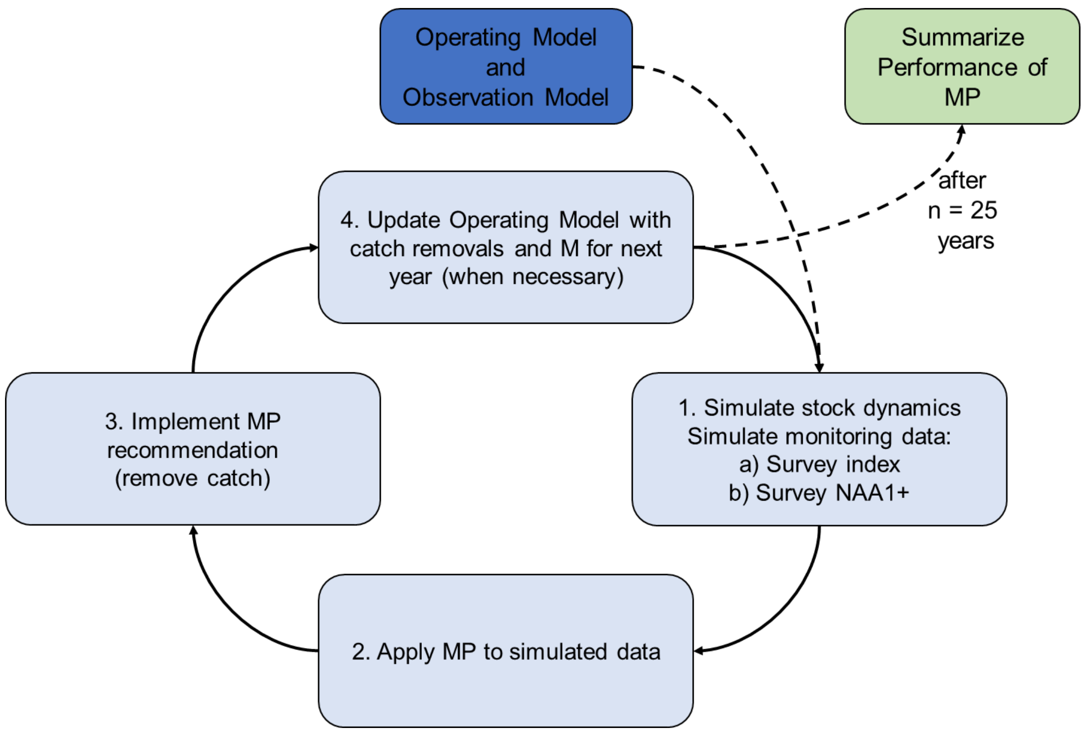

Figure 19. Illustration of the closed-loop simulations. For each operating model (OM) and management procedure (MP) combination, simulated stock dynamics are generated by projecting the population dynamics model and applying an observation model to generate the simulated monitoring data (index and index NAA) (step 1). The MP is applied to the simulated data (step 2) and the MP generates a total allowable catch (TAC). The TAC is implemented (step 3) without implementation error. The final step of the simulation loop (step 4) is to remove the catch from the population and update the M for the next year (when necessary) by updating the OM. This process (steps 1 to 4) is repeated until the end of the 25-year projection period.

9. EXAMPLE MANAGEMENT PROCEDURES (MPs)

A set of candidate MPs is defined in this document and used to illustrate the application of the MP evaluation framework. All MPs were defined as rules to set the TAC (in kt) as a function of the DFO survey index (Figure 20; Table 7). These MPs are examples that are used only to illustrate how the MPs are evaluated. MP development will occur at a later date.

The first example MP, *PA_0_8.6* was defined to be consistent with the DFO PA policy provisional harvest rule (DFO 2009), using the q-adjusted and model-weighted LRP (15.2 kg/tow) and proposed USR (25.9 kg/tow) as the lower and upper control points for the index. Several upper control points for TAC were evaluated to meet the performance metric that keeps F below F_{ref} with at least 50% probability in each projection year, and a TAC of 8.6 kt was identified for an initial MP. The lower control point of 0 kt is based on the PA policy. The relative harvest rate (ratio of TAC to the index) declines linearly from the upper control point to the lower control point and is constant above the upper control point (Figure 20).

Given that Haddock are collected as part of a multi-species groundfish fishery, a TAC of zero may not be practical if the stock is below the LRP. Another example MP was therefore defined as a modification to *PA_0_8.6*, with a catch below the LRP set to 1 kt to give *PA_1_8.6* (Figure 20, Table 7).

The other example MPs were modifications to *PA_1_8.6* such that the MP provides two-year catch advice (*PA_1_8.6_2yr*), limits the annual change in TAC to 15% (*PA_1_8.6_TAC15*), and constant TACs of 6.198 kt (the constant TAC used for the 4X5Y Haddock fishery over the last three years), and 10 kt (Figure 20, Table 7).

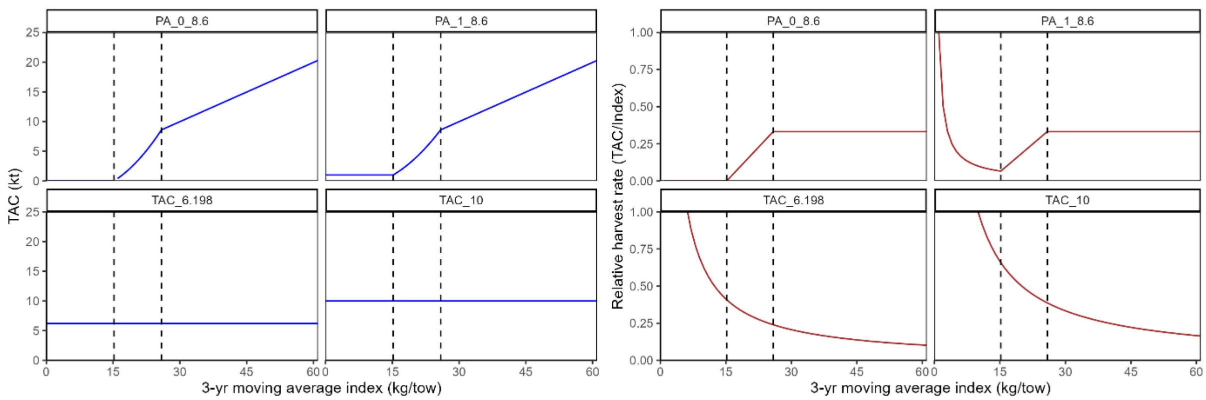

Figure 20. Plot of recommended total allowable catch (TAC) (left panel) and relative harvest rate (TAC/index) (right panel) vs. the three-year moving average DFO survey index for four example management procedures (MPs) evaluated in this document.

Table 7. Description of example management procedures (MPs) evaluated in this document.

MP name	MP Description
PA_0_8.6	DFO PA Policy provisional harvest rule to set TAC based on three-year moving average Index LCP = 15.2 kg/tow; UCP = 25.9 kg/tow TAC = 0 kt when Index < LCP; TAC = 8.6 kt when Index = UCP Harvest rate declines linearly from UCP to LCP Harvest rate above UCP is the harvest rate at the UCP
PA_1_8.6	Same as *PA_0_8.6* except TAC = 1 kt when Index < LCP; TAC = 8.6 kt when Index = UCP
PA_1_8.6_2yr	Same as *PA_1 _8.6* except Based on 2-year catch advice
PA_1_8.6_TAC15	Same as *PA_1 _8.6* except Limits inter-annual changes in TAC by 15%
TAC_6.198	Constant TAC of 6.198 kt
TAC_10	Constant TAC of 10 kt

PA=precautionary approach, TAC=total allowable catch, LCP=lower control point, UCP=upper control point.

10. EXAMPLE MANAGEMENT PROCEDURE EVALUATION

The performance of the example MPs defined in Section 9 was evaluated relative to the management objectives defined in Section 2. The first step (Figure 21) of the MP evaluation is to average (using the OM weights) performance across OMs 1–3 (Table 8). The MP must meet the minimum performance standards for the first three management objectives in Table 1:

1. Maintain stock above the LRP:
   a. P(SSB > LRP) > 75–95% over 10 years
   b. P(SSB > LRP) > 75–95% over 25 years
2. Maintain stock above the USR:
   a. P(SSB > USR) ≥ 50% over 10 years
   b. P(SSB > USR) ≥ 50% over 25 years
3. Maintain fishing mortality below F_{ref}:
   a. P(F < F_{ref}) ≥ 50% every year in projection period
4. Promote stock growth when stock is below USR
   a. P(SSB_{y+5} > SSB_y) ≥ 50%, since recent trajectory is not a decline.

Management objectives 5 and 6 relate to maximizing TAC and avoiding large inter-annual changes in TAC (Table 1), and MP performance for these objectives are reported as summary statistics for which tradeoffs among MPs can be evaluated (Table 8).

The second step (Figure 21) of the MP evaluation is to average performance (using the OM weights) across OMs 1b–3b for the first two management objectives (Table 1), over the entire 25-year projection period (Table 9). Productivity assumed for the projection period for OMs 1b–3b is the historical productivity from 1998–2022, that includes the “2013” cohort. The objective of the evaluation over these OMs, is to ensure the MP meets the minimum performance standard for the LRP and USR objectives over the 25-year projection period. The projection period is divided into two additional time periods for the evaluation of trade-offs in TAC and variability in TAC: the first 17 years (before the “2013” cohort enters the SSB) and years 18–25 (progression of the “2013” cohort) (Table 10). The F_{ref} management objective is not applied to the evaluation across OMs 1b–3b because F_{ref} is specific to the productivity period assumed in OMs 1–3.

If an MP meets the minimum performance standard for the first four objectives for the first set of OMs (OMs 1–3), and meets the minimum performance standard for the first two objectives over the 25-year projection period for the second set of OMs (OMs 1b–3b) then it is considered to have met the conservation objectives (Table 10).

The annual projected SSB, catch, *F*, P(SSB > LRP), DFO survey index, and annual variability in catch were plotted by MP for OMs 1–3 (Figure 22, Figure 23) and OMs 1b–3b (Figure 24, Figure 25). The distribution (across years and simulations) of annual variability in catch $\left(\frac{|C_{y+1}+C_y|}{C_y} \times 100\%\right)$ was plotted by MP for OMs 1–3 (Figure 26) and OMs 1b–3b (Figure 27). The magnitude of the annual variability in catch was capped at a maximum of 100% to accommodate catches near zero which inflate the annual variability calculation.

MPs that meet the minimum performance standard for the conservation objectives (i.e., all except TAC_10 and PA_1_8.6_2yr, Table 10) can be carried forward to evaluate the tradeoffs among MPs.

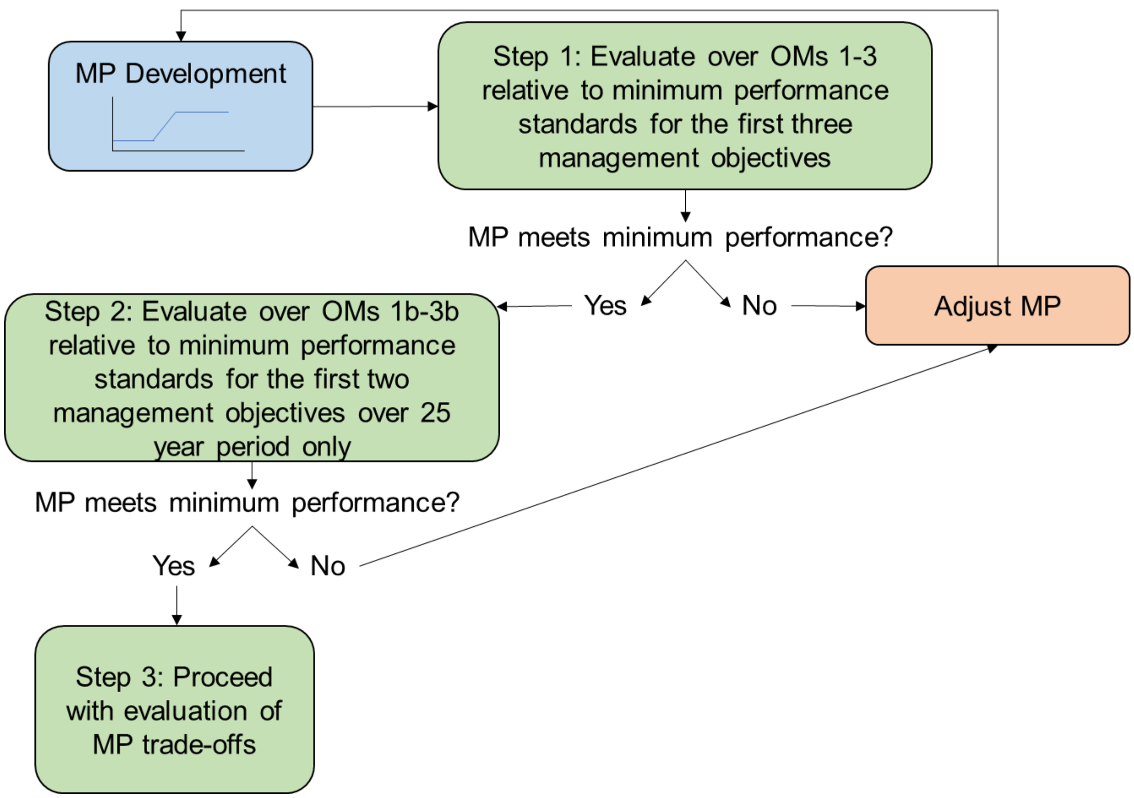

Figure 21. Steps taken for the management procedure (MP) evaluation.

Table 8. Summary of performance metrics for the six management procedures (MPs) based on the weighting of Operating Models 1–3 where SSB = spawning stock biomass, LRP = limit reference point, USR = proposed upper stock reference point, Fref = fishing mortality reference, and y= year. Shaded cells indicate that the probability (P) exceeded (red) or did not exceed (green) the threshold set for the performance metric.

Management Objectives	1a	1b	2a	2b	3	4	5			6		
	P(SSB>LRP)		P(SSB>USR)		Proportion of years $P(F<F_{ref}) >$ 50%	$P(SSB_{y+5} > SSB_y)$	Median Catch (kt)			Annual Variability in Catch (%)		
MP	10 yr	25 yr	10 yr	25 yr	25 yr	5 yr	5 yr	10 yr	15 yr	5 yr	10 yr	15 yr
PA_0_8.6	0.97	0.96	0.71	0.65	1	0.88	7.06	9.17	9.21	23.96	19.37	19.20
PA_1_8.6	0.97	0.96	0.69	0.63	1	0.86	7.05	9.05	9.10	21.22	18.17	17.55
PA_1_8.6_2yr	0.97	0.94	0.71	0.63	0.96	0.92	5.16	9.03	9.07	0.01	0	0.01
PA_1_8.6_TAC15	0.99	0.92	0.81	0.66	1	0.95	4.58	5.90	7.04	13.04	13.04	13.04
TAC_10	0.86	0.72	0.56	0.46	0.32	0.67	10	10	10	0	0	0
TAC_6.198	0.97	0.94	0.78	0.75	1	0.90	6.198	6.198	6.198	0	0	0

Table 9. Summary of performance metrics for the six management procedures (MPs) based on the weighting of Operating Models 1b–3b where SSB = spawning stock biomass, LRP = limit reference point, USR = proposed upper stock reference point, and AV = annual variability. Shaded cells indicate that the probability (P) exceeded (red) or did not exceed (green) the threshold set for the performance metric.

	Historical Productivity (1998–2014)				Historical Productivity (2015–2022) with large “2013” cohort				Historical Productivity (1998–2022)			
MP	P(SSB>LRP)	P(SSB>USR)	Median Catch (kt)	AV in Catch (%)	P(SSB>LRP)	P(SSB>USR)	Median Catch (kt)	AV in Catch (%)	P(SSB>LRP)	P(SSB>USR)	Median Catch (kt)	AV in Catch (%)
PA_0_8.6	0.93	0.57	9.05	26.2	0.73	0.41	9.23	43.5	0.86	0.52	9.08	30.6
PA_1_8.6	0.92	0.56	8.93	23.9	0.72	0.39	9.23	34.1	0.86	0.51	9.02	26.3
PA_1_8.6_2yr	0.90	0.53	8.82	0.01	0.69	0.40	9.16	10.9	0.83	0.49	8.90	2.73
PA_1_8.6_TAC15	0.90	0.61	7.43	13.0	0.87	0.45	6.25	13.0	0.88	0.56	6.99	13.1
TAC_10	0.82	0.54	10	0	0.50	0.30	10	0	0.72	0.46	10	0
TAC_6.198	0.98	0.73	6.198	0	0.93	0.53	6.198	0	0.96	0.66	6.198	0

Table 10. Summary of the MP evaluation over operating models (OMs) 1–3 and OMs 1b–3b.

MP	Summary of MP evaluation for objectives 1a, 1b, 2a, 2b, and 3 for OMs 1–3	Summary of MP evaluation for objectives 1 and 2 for OMs 1b–3b	MP meets conservation objectives
PA_0_8.6	Pass	Pass	Yes
PA_1_8.6	Pass	Pass	Yes
PA_1_8.6_2yr	**Fail**	**Fail**	No
PA_1_8.6_TAC15	Pass	Pass	Yes
TAC_10	**Fail**	**Fail**	No
TAC_6.198	Pass	Pass	Yes

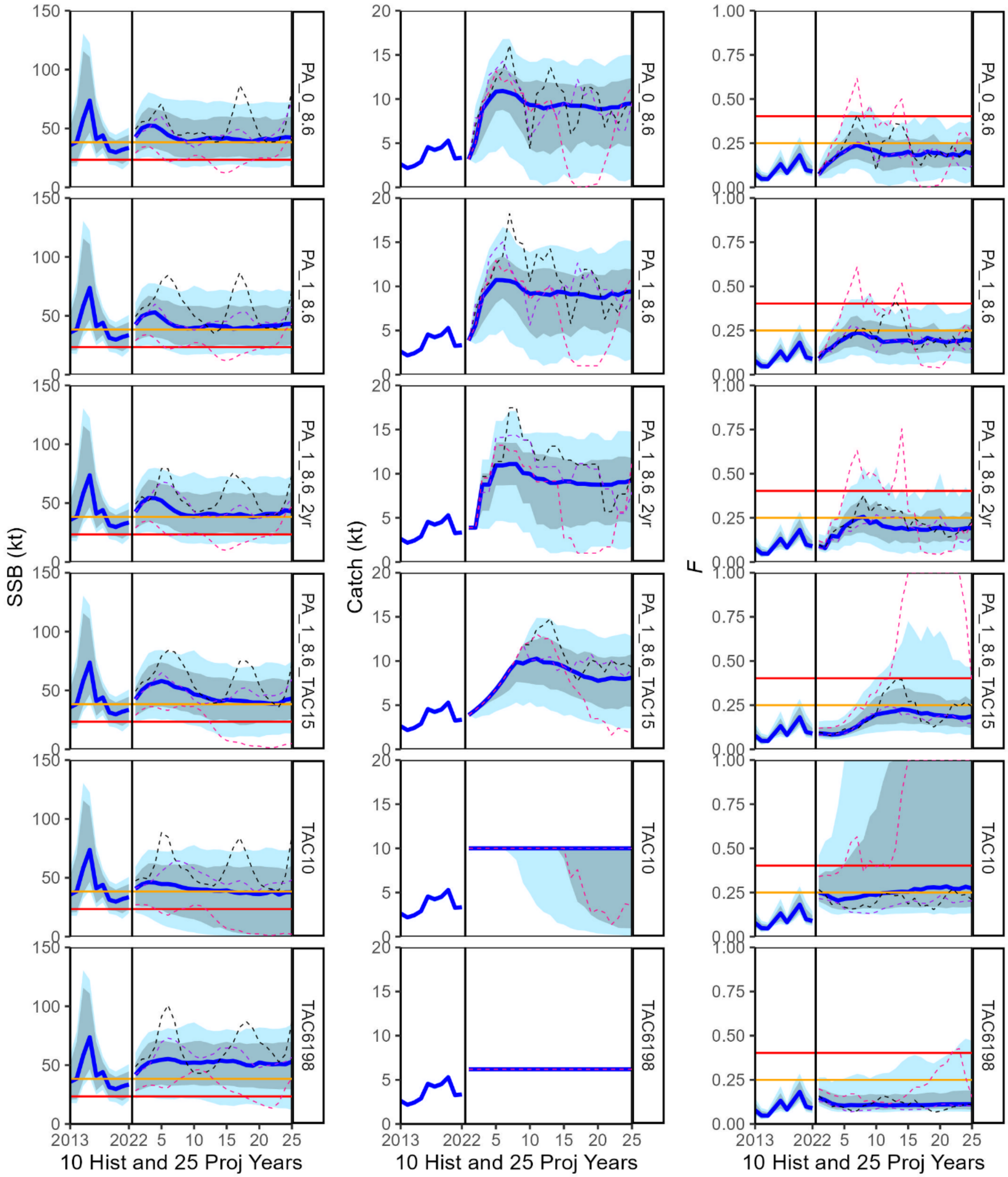

Figure 22. Historical and projected median (blue line), 25th and 75th percentiles (dark shading), 10th and 90th percentiles (light shading), for spawning stock biomass (SSB), catch, and fishing mortality rate (F), weighted across OMs 1–3 for 100 simulations. Trajectories for three randomly selected projected simulations are plotted as dashed lines. Horizontal reference lines for SSB are the OM weighted limit reference point (LRP; red) and proposed upper stock reference point (USR; orange). Horizontal reference lines for F are the OM weighted fishing mortality reference and limit (F_{ref}; orange and F_{lim}; red). Each row of figures are for a different management procedure.

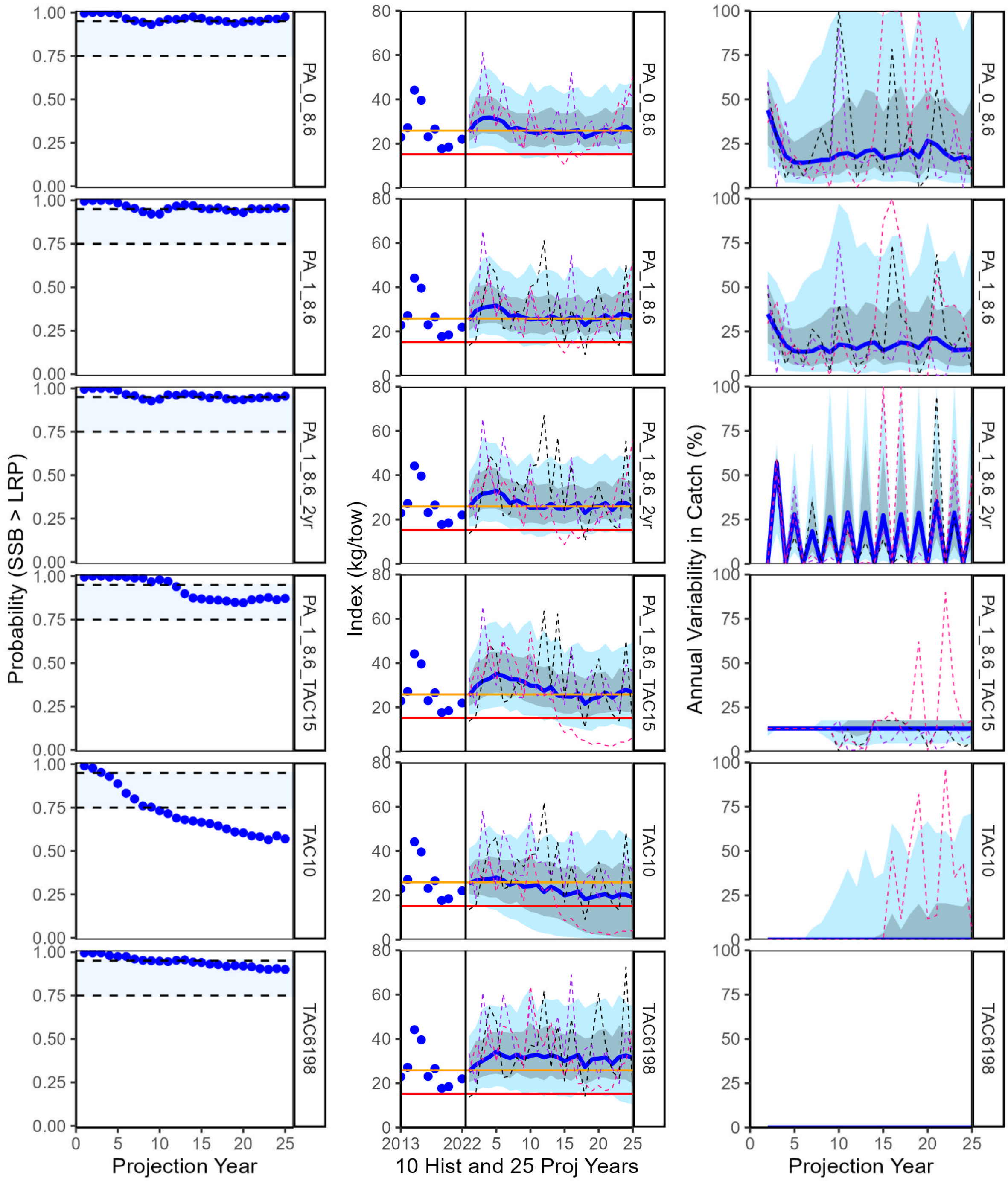

Figure 23. The projected probability spawning stock biomass exceeds the limit reference point P(SSB > LRP), weighted by operating model (OM) for OMs 1–3 (left panel). The blue shading indicates the probability range of 75–95%. Historical (blue circles) and projected median (blue line), 25th and 75th percentiles (dark shading), 10th and 90th percentiles (light shading), for the DFO survey index (middle panel), trajectories for three randomly selected projected simulations are plotted as dashed lines. Annual variability in catch (right panel), weighted across OMs 1–3 for 100 simulations. Horizontal reference lines for the index are the OM weighted and q-adjusted limit reference point (red) and proposed USR (orange). The annual variability in catch was capped at 100% for plotting purposes. Each row of figures are for a different management procedure.

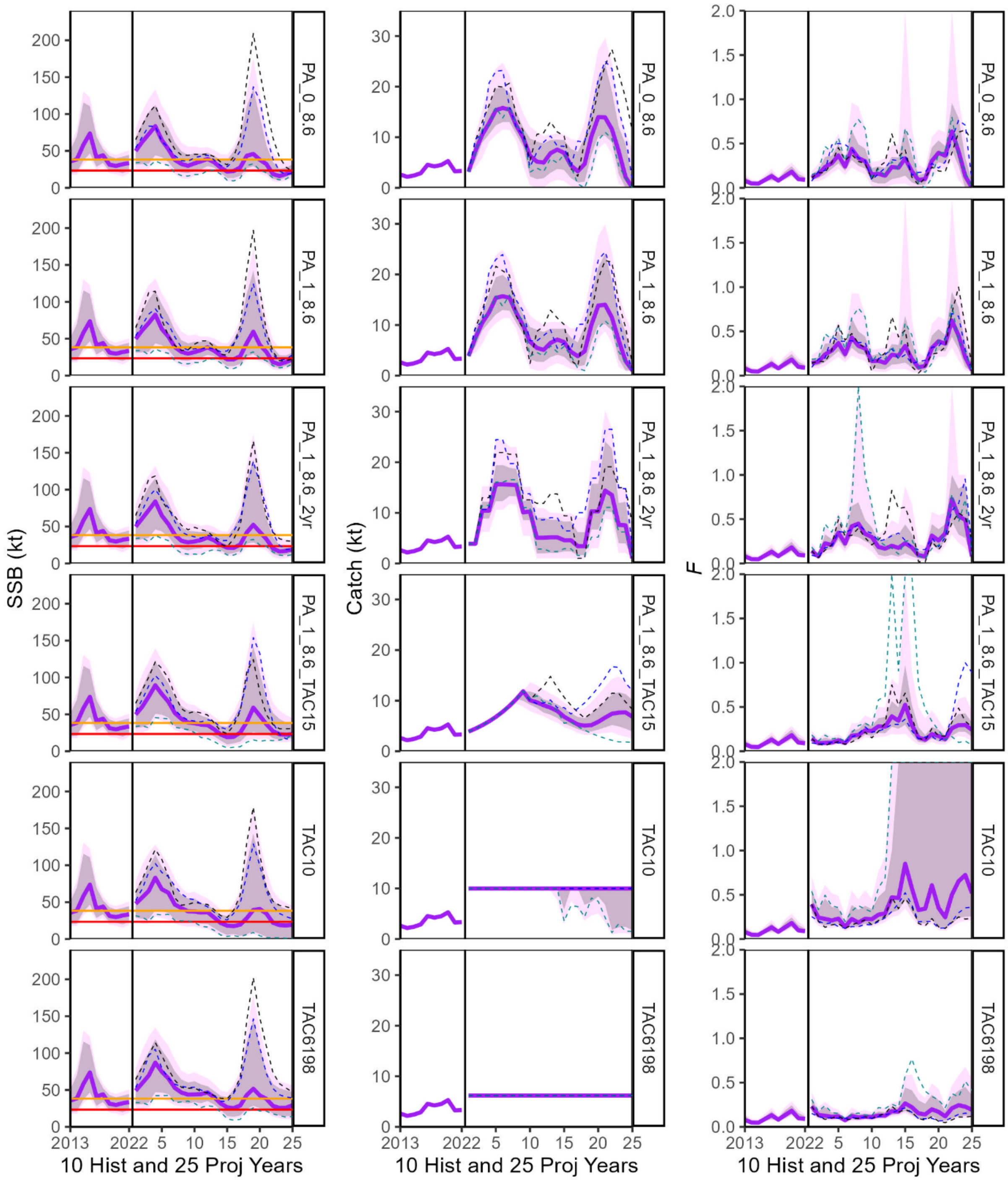

Figure 24. Historical and projected median (purple line), 25th and 75th percentiles (dark shading), 10th and 90th percentiles (light shading), for spawning stock biomass (SSB), catch, and fishing mortality rate (F), weighted across OMs 1b–3b for 100 simulations. Trajectories for three randomly selected projected simulations are plotted as dashed lines. Horizontal reference lines for SSB are the OM weighted limit reference point (LRP; red) and proposed upper stock reference point (USR; orange). Each row of figures represent a different management procedure.

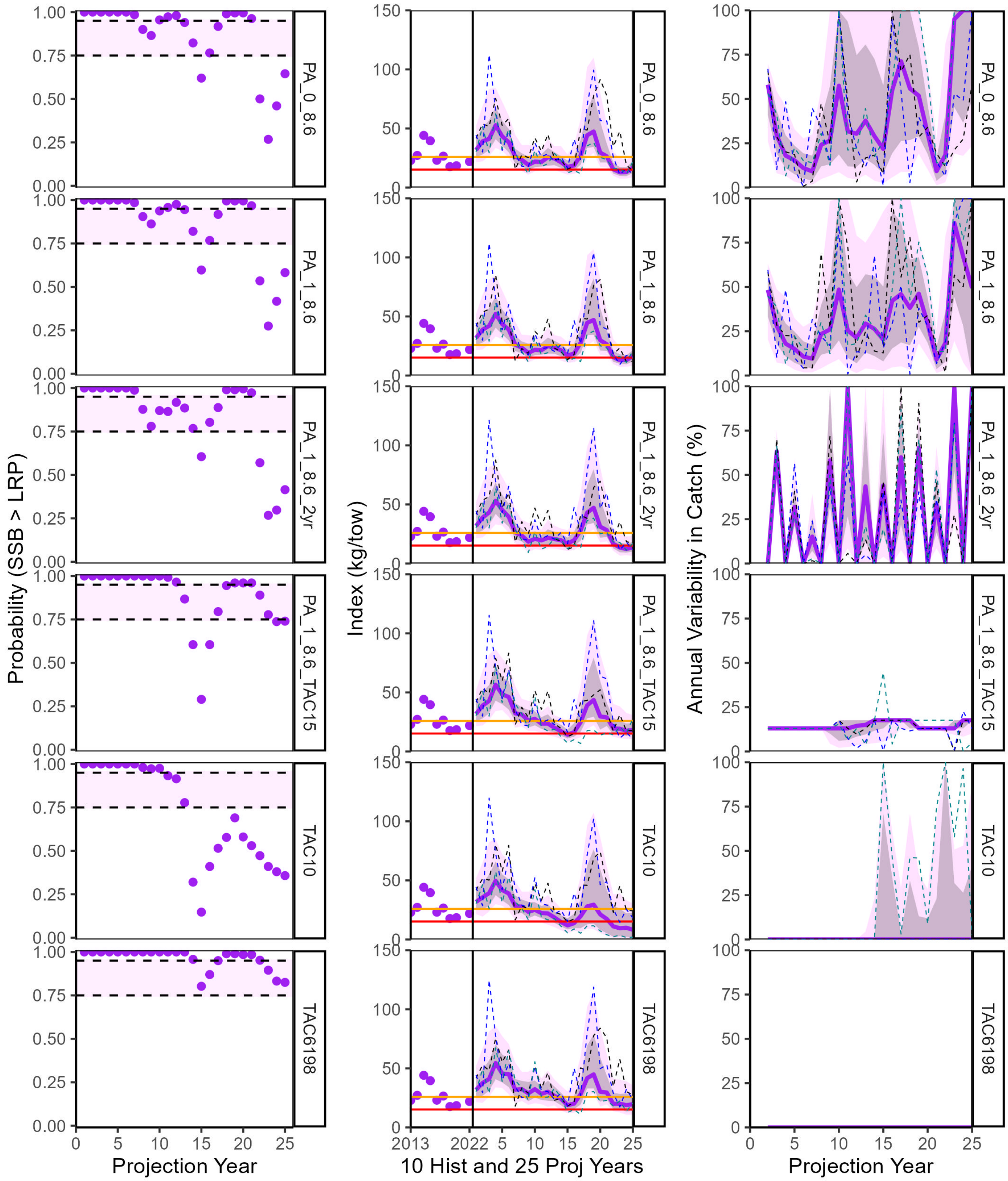

Figure 25. The projected probability spawning stock biomass exceeds the limit reference points P(SSB > LRP), weighted by operating model (OM) for OMs 1b–3b (left panel). Historical (purple circles) and projected median (purple line), 25th and 75th percentiles (dark shading), 10th and 90th percentiles (light shading), for the DFO survey index (middle panel), trajectories for three randomly selected projected simulations are plotted as dashed lines. Annual variability in catch (right panel), weighted across OMs 1–3 for 100 simulations. Horizontal reference lines for the index are the OM weighted and q-adjusted limit reference point (red) and proposed upper stock reference point (orange). The annual variability in catch was capped at 100% for plotting purposes. Each row of figures represent a different management procedure.

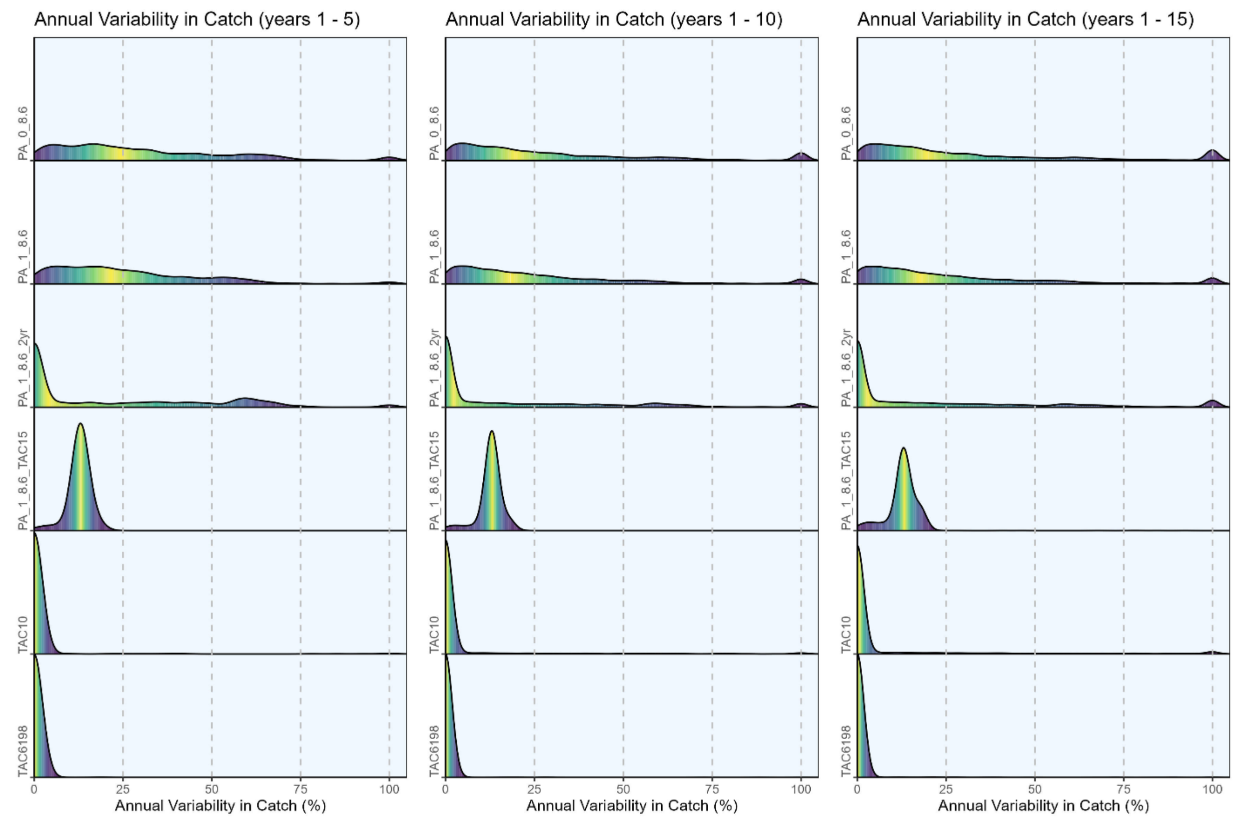

Figure 26. Density plots for the annual variability in catch by management procedure (MP) for OMs 1–3 for three time periods (5 years, left panel; 10 years, middle panel; 15 years, right panel). Note that the annual variability in catch was capped at 100% for plotting purposes and the viridis colour scale is applied over quantiles (e.g., median is yellow).

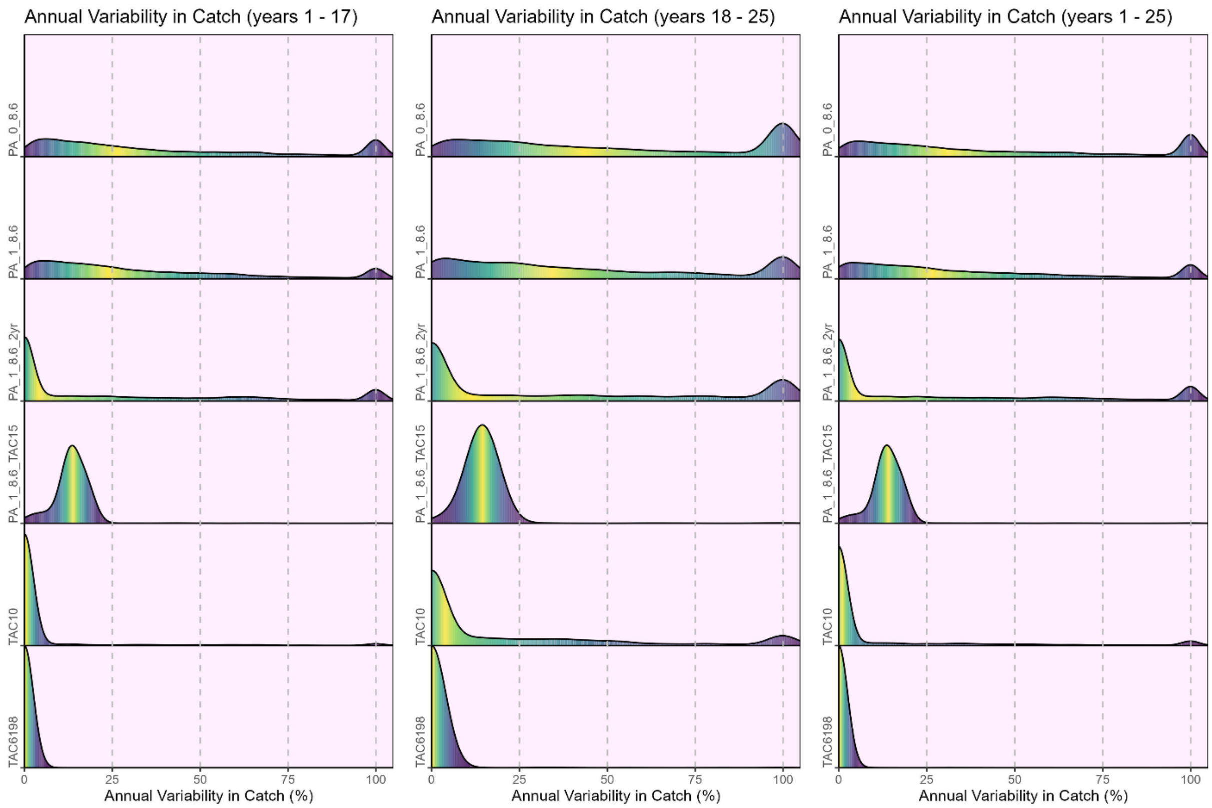

Figure 27. Density plots for the annual variability in catch by management procedure (MP) for OMs 1b–3b for three time periods (1–17 years, left panel; 18–25 years, middle panel; 1–25 years, right panel). Note that the annual variability in catch was capped at 100% for plotting purposes and the viridis colour scale is applied over quantiles (e.g., median is yellow).

11. MANAGEMENT PROCEDURE SELECTION

All MPs presented in this document are for demonstration purposes and only the simulation approach used to evaluate the performance of the MPs was reviewed in this third CSAS peer-review meeting. This MP evaluation framework will be applied after the meeting to test other MPs submitted to DFO Science. The performance of each MP will be provided along with an evaluation of trade-offs among the management objectives to RM for consultation with stakeholders. After this consultation, one MP will be chosen for providing future catch advice for the 4X5Y Haddock fishery.

12. EXCEPTIONAL CIRCUMSTANCES AND ASSESSMENT TIMING

Once an MP is adopted, it will be used to provide catch advice. Exceptional circumstances are commonly defined when an MP is adopted for a fishery to address situations outside the range of scenarios for which the MP was simulation tested or when the data required to apply the MP are not available. The primary reasons identified for triggering an exceptional circumstance for this MP evaluation framework for 4X5Y Haddock are:

A. There is evidence that the assumed stock dynamics in OMs 1–3 have changed or are no longer appropriate.

B. There are substantial changes to the fishery or fleet dynamics.

C. The data required to apply the MP are no longer available or no longer appropriate.

The expectation is that catch advice will be provided by evaluating new data (i.e., the DFO survey index) relative to the MP and the exceptional circumstances would be evaluated in an update meeting. The frequency of the meetings (e.g., annually or every two years) will be dependent on the prioritization by the regional CSAS office. Specific actions to be taken when each exceptional circumstance is triggered are provided generally; however, specific actions will be evaluated on a case by case basis and may involve discussions with Resource Management and the Scotia Fundy Groundfish Advisory Committee. If exceptional circumstances are not triggered (Table 11), the MP is proposed to be used for 5 to 7 years, upon which the MP will be reviewed to determine next steps.

Additional indicators, outside of the exceptional circumstances, were compiled based on discussions during the peer-review meeting (Table 12). These indicators will be evaluated regularly, and accompany the typical assessment indicators for the fishery and survey (e.g., catch, catch-at-age, catch-at-length), as well as the evaluation of the exceptional circumstances.

Table 11. Indicators, evaluation criteria, and science considerations for the evaluation for the exceptional circumstances. OM= operating model; M= natural mortality; MP= management procedure

Trigger	Indicator	Evaluation Criteria	Science Considerations
A1	DFO survey index	Observed index is below 15.2 kg/tow in two consecutive years (Figure 28).	Trigger a review of stock status
A2	Data inputs	DFO Science identifies new data to suggest that data inputs or model assumptions (e.g., *M*) are no longer valid (e.g., Higher *M* under low productivity state).	Findings could trigger a new framework, revision of OMs, or re-evaluation of MPs
B2	Fishery selectivity	Evidence that fishery selectivity may differ substantially from selectivity assumed in the OMs.	Findings could trigger a new framework, revision of OMs, or re-evaluation of MPs
C1	DFO survey index	The index is not available to apply the MP.	Roll over of an annual TAC if missing for one year. To be determined if missing for more than one year.

Table 12. Additional indicators to be evaluated regularly for updates. OM= operating model; TAC= total allowable catch.

Indicator	Evaluation	Science Considerations
DFO survey index	Observed index is outside the 80% prediction interval (10^{th} and 90^{th} percentiles) in the same direction for OMs 1–3 in two consecutive years. See Figure 28 for an example	Low values addressed through exceptional circumstances
Weight-at-age (WAA)	Compare observed DFO survey WAA to assumed projected WAA	Differences in WAA not expected to impact MP performance
Landings	Compare landings to TAC	Discussion with Resource Management
Relative harvest rate by region (SS and BoF)	Compare relative harvest rate by region to a proxy for F_{crash}	Avoid local depletion

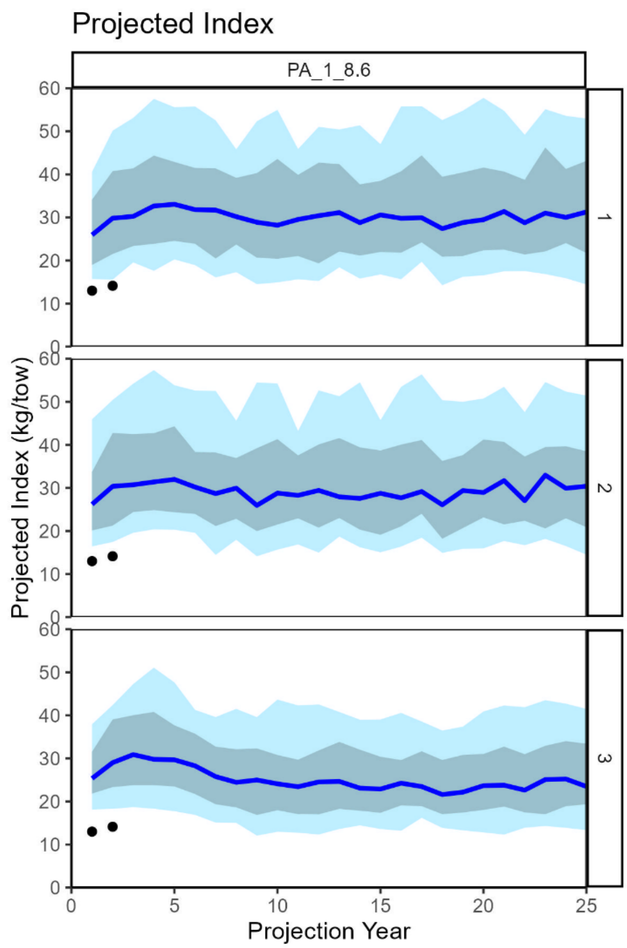

Figure 28. Projected median (blue line), 25th and 75th percentiles (dark shading), and 10th and 90th percentiles (light shading) for the DFO survey index by operating model (OM) (1, 2, and 3) for management procedure (MP) PA_1_8.6, showing two future observed index values (projection year 1 and 2) falling outside of the 80% prediction limits for the index for all three OMs.

REFERENCES CITED

Acheson, C. 2020. Fecundity of Haddock (*Melanogrammus aeglefinus*) of Eastern Georges Bank. Thesis (B.Sc.) University of New Brunswick, Fredericton, N.B., 32 p.

Barrett, T.J., Marentette, J.R., Forrest, R.E., Anderson, S.C., Holt, C.A., Ings, D.W., and Thiess, M.E. 2024. Technical Considerations for Stock Status and Limit Reference Points under the Fish Stocks Provisions. DFO Can. Sci. Advis. Sec. Res. Doc. 2024/029. v + 57 p.

Barrett, T.J., and Barrett, M.A. 2025. 2024 4X5Y Haddock Framework: Data Inputs. DFO Can. Sci. Advis. Sec. Res. Doc. 2025/020 iv+85 p

Brooks, E.N. 2024. Pragmatic approaches to modeling recruitment in fisheries stock assessment: A perspective. Fisheries Research 270: 106896.

Chen, S., and S. Watanabe. 1989. Age dependence of natural mortality coefficient in fish population dynamics. Bull. Jap. Soc. Sci. Fish. 55: 205–208.

Cope, J.M., and Hamel, O.S. 2022. Upgrading from M version 0.2: An application-based method for practical estimation, evaluation and uncertainty characterization of natural mortality. Fish. Res. 256:106493.

Crone, P., Maunder, M., Valero, J., McDaniel, J., and Semmens, B. 2013. Selectivity: theory, estimation, and application in fishery stock assessment models. CAPAM Workshop Series Report 1. 50 p.

DFO. 2009. A fishery decision-making approach incorporating the Precautionary Approach. Fisheries and Oceans Canada. Accessed September 2024.

DFO. 2020. Evaluating the robustness of candidate management procedures in the BC sablefish (*Anoplopoma fibria*) fishery for 2019-2020. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2020/025.

DFO. 2022. Guidelines for implementing the fish stocks provisions in the Fisheries Act. Fisheries and Oceans Canada. Accessed September 2024.

DFO. 2023a. Stock Status Update of Haddock (*Melanogrammus aeglefinus*) in NAFO Divisions 4X5Y for 2022. DFO Can. Sci. Advis. Sec. Sci. Resp. 2023/018.

DFO. 2023b. Science Advice on Guidance for Limit Reference Points under the Fish Stocks Provisions. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2023/009.

Finley, M., Wang, Y., and Stone, H.H. 2018. Assessment of 4X5Y Haddock (*Melanogrammus aeglefinus*) in 2016. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/041. iv + 54 p.

Fu, C., and Quinn II., T.J., 2000. Estimability of natural mortality and other population parameters in a length-based model: *Pandalus borealis* in Kachemak Bay, Alaska. Can. J. Fish. Aquat. Sci. 57: 2420–2432.

Gabriel, W.L., Sissenwine, M.P., and Overholtz, W.J. 1989. Analysis of spawning stock biomass per recruit: an example for Georges Bank Haddock. N. Am. J. Fish. Manag. 9: 383-391.

Hamel, O.S., 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES J. Mar. Sci. 72: 62–69.

Hamel, O.S., Cope, J.M., 2022. Development and considerations for application of a longevity-based prior for the natural mortality rate. Fish. Res. 256:106477.

Hamel, O.S., Punt, A.E., Kapur, M.S., and Maunder, M.N. 2023. Natural mortality: theory, estimation and application in fishery stock assessment models. NOAA Processed Report NMFS-NWFSC-PR-2023-02.

Hordyk, A., Huynh, Q., Carruthers, T., and Grandin, C. 2022. Package “MSEtool”: Stock Assessment Methods Toolkit. Version 1.2.5.

ICES. 2013. Report of the Workshop on guidelines for management strategy evaluations (WKGMSE). 21-23 January 2013, Copenhagen, Denmark. ICES CM 2013/ACOM:39. 122 p.

ICES. 2020. The third workshop on guidelines for management strategy evaluations (WKGMSE3). ICES Scientific Reports 2:116. 112 p.

ICES. 2021. ICES fisheries management reference points for category 1 and 2 stocks. ICES technical guidelines.

Kristensen K (2024). RTMBode: Solving ODEs with 'deSolve' and 'RTMB'. R package version 1.0.

MacCall, A., 2013. Use of the delta method to evaluate the precision of assessments that fix parameter values. Fish. Res. 142, 56–60.

Marshall, D.J., Bode, M., Mangel, M., Arlinghaus, R., and J. Dick. 2021. Reproductive hyperallometry and managing the world’s fisheries. PNAS 118(34): e2100695118.

Maunder, M.N., Hamel, O.S., Less, H., Piner, K.R., Cope, J.M., Punt, A.E., Ianelli, J.N., Castillo-Jordan, C., and Kapur, M.S. 2023. A review of estimation methods for natural mortality and their performance in the context of fishery stock assessment. Fish. Res. 257: 106480.

MF (New Zealand Government, Ministry of Fisheries). 2008. Harvest Strategy Standard for New Zealand Fisheries. October 2008. 30 p.

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

Pennington, M. 1996. Estimating the mean and variance from highly skewed marine data. Fish. Bull. 94: 1996.

Pope, J.G., Gislason, H., Rice, J.C., and Daan, N. 2021. Scrabbling around for understanding of natural mortality. Fish. Res. 240: 105952.

Punt, A.E., Hurtado-Ferro, F., and Whitten, A.R., 2014. Model selection for selectivity in fisheries stock assessment. Fish. Res. 158: 124–134.

Punt, A.E., Castillo-Jordan, C., Hamel, O.S., Cope, J.M., Maunder, M.N., and Ianelli, J.N. 2021. Consequences of error in natural mortality and its estimation in stock assessment models. Fish. Res. 233:105759.

Rivard, D. 1982. APL programs for stock assessment (revised). Canadian Technical Report of Fisheries and Aquatic Sciences. No. 1091.

Sainsbury, K. 2008. Best practices for reference points for Australian fisheries. Australian Fisheries Management Authority Report R2001/0999.

Sampson, D.B., and Scott, R.D. 2012. An exploration of the shape and stability of population-selection curves. Fish and Fisheries 13: 89–104.

Shelton, P.A., and J.C., Rice. 2002. Limits to overfishing: reference points in the context of the Canadian perspective on the precautionary approach. DFO Can. Sci. Advis. Sec. Res. Doc. 2002/084. iv + 29 p.

Stock, B.C., and Miller, T.J. 2021. The Woods Hole assessment model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fish. Res. 240:105967.

Szuwalski, C.S., 2022. Estimating time-variation in confounded processes in population dynamics modeling: A case study for snow crab in the eastern Bering Sea. Fish. Res. 251:106298.

Szuwalski, C.S., Ianelli, J.N., Punt, A.E., 2018. Reducing retrospective patterns in stock assessment and impacts on management performance. ICES J. Mar. Sci. 75:596–609.

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

Then, A.Y., Hoenig, J.M., Hall, N.G., Hewitt, D.A., 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES J. Mar. Sci. 72, 82–92.

Wang, Y., Stone, H.H., and Finley, M. 2017. 4X5Y Haddock 2016 framework assessment: Modelling and reference points. DFO Can. Sci. Advis. Sec. Res. Doc. 2017/026.

Wang Y., Regnier-McKellar, C., and K. Kraska. 2022. Assessment of Haddock on Eastern Georges Bank for 2022. TRAC Reference Document 2022/02.

APPENDIX A

A1. MODEL CONFIGUATIONS

A1.1 Natural Mortality Rate (*M*)

M is often considered to be one of the most important parameters in a fish stock assessment and affects productivity estimates for the population (Punt et al. 2021). *M* is defined as all sources of mortality (e.g., predation, disease) not attributed to fishing and can include unreported fishery catch and movement of fish outside a management area (Maunder et al. 2023). For fish with strong cohorts, similar to Haddock, *M* may be elevated for younger fish based on limited resources and reduced growth (Hamel et al 2023). The estimation of *M* within a stock assessment model is often confounded by selectivity, growth, and the steepness of the stock-recruitment relationship and may vary by sex, age, size, and time-period (Punt et al. 2021, Hamel et al. 2023). A low time-invariant *M* (e.g., *M* = 0.2, Pope et al. 2021), has traditionally been assumed for many stock assessments when *M* is unknown, partly because of the difficulty of estimating *M* and because a lower *M* may lead to more precautionary advice (Hamel et al 2023). Estimating *M* may be better than fixing *M* at an incorrect value (Fu and Quinn 2000); however, standard practice should include constructing likelihood profiles for *M* or examining sensitivities with alternative values of *M* (MacCall 2013, Punt et al. 2021) to avoid misspecification of the model (Szuwalski et al. 2018). The use of management strategy evaluation (MSE) allows for OMs to be specified to capture the uncertainty in *M*, leading to harvest control rules (HCRs) that are robust to these uncertainties (Hamel et al. 2023).

A review by Hamel and Cope (2022), recommended that even when direct data that inform on *M* are incorporated into a stock assessment model, other methods that estimate *M* outside of the model should be explored and considered for use as a prior or as a diagnostic tool. The Natural Mortality Tool is an R application that was developed to encompass more than 20 empirical estimators of *M* using life history parameters related to growth, maturity, and longevity (Cope and Hamel 2022). These estimates can then be used individually or combined in a weighted-average across methods. Based on the general recommendations from Cope and Hamel, three of the estimators in the tool were used to estimate *M* empirically for 4X5Y Haddock. These approaches were based on longevity (Hamel 2015), the von Bertalanffy growth parameter *k* (Hamel 2015), and both *k* and asymptotic length (L_{inf}) (Then et al. 2015). The average of these approaches was used to estimate a constant (*M* = 0.412) and a time-varying *M* that were evaluated for 4X5Y Haddock. An additional method proposed by Chen and Watanabe (1989) was used to estimate a time-varying age-specific *M*.

The various *M* scenarios explored are summarized in Table A2. A time- and age-invariant *M* was estimated outside the model using life history parameters and provided a mean estimate of 0.412 (model *m0_M0.412*). A time- and age-invariant *M* was model-estimated (*m0_Me*) to be 0.552. The higher *M* did improve the fit to the DFO survey index (e.g., *m0_Me*) but the model *m0_Me* generally underestimated the DFO survey index from 1995 to 2015. *M* was estimated in two time blocks (1970–2009; 2010–2022) based on residual patterns and the blocking period is consistent with that used in the eastern Georges Bank Haddock assessment model (Wang et al. 2022). For model *M0_M2b0.2_2010*, *M* was fixed in the first time block to 0.2 and for model *m0_M2b_2004*, *M* was estimated separately in each time block (Table A2).

A likelihood profile across values of *M* was attempted. The negative log-likelihood decreased as *M* increased from 0.2 to 0.5 in increments of 0.05 but a model with *M* = 0.55 did not converge, thus no minimum was identified in the profile.

Some age-varying *M* scenarios were explored using blocks of ages 1–4 and 5+ with different values of *M* (estimated, 0.2, or 0.412; *m0_a*, *m0_a2*). A default method coded in WHAM based on weight-at-age estimated an age- and time-varying *M* (Table A2; m0_Mwaa). The DFO survey numbers-at-age (NAA) 1+ was used as a covariate for estimating *M* under a few different assumptions on transformation and CV. A model with a CV on the NAA 1+ of 0.05 (model *m0_Me1+_sig0.05*) with *M* estimates ranging from 0.23 to 1.21 had relatively good diagnostics (Table A3). From the evaluation of models with different assumptions on *M*, models *m0_M2b_2010* and *m0_Me1+_sig0.05* were identified as candidate models that capture the assumption of higher (i.e., > 0.2) *M* (Table A3). Model *m0_Me1+_sig0.05* was selected as the best candidate due to its ability to adjust *M* down in the projection period, and the model was revised (after the second peer-review meeting) to use the DFO survey estimated assuming a delta-lognormal distribution (*dln*) and using a one-year lag *(L1)* on the relationship between the DFO survey NAA1+ and *M*, to allow for a time-varying M scenario to be projected forward using the previous year DFO survey NAA1+ (model: *m0_Me1+_dln_sig0.05_L1*; Section 4.2).

A1.2 Selectivity and Catchability

Fishery selectivity represents the combined factors that affect fish vulnerability which includes contact sensitivity (probability of being encountered by a gear) and availability (probability that fish are in the area where the fishery occurs) (Crone et al. 2013). The spatial area of the 4X5Y Haddock management unit is relatively large, and fishing generally occurs in concentrated areas and not uniformly over time within a year. Although the bottom trawl gear may be assumed to have asymptotic selectivity, dome-shaped selectivity may occur due to spatial and temporal availability of Haddock and may be influenced by the presence of strong cohorts if Haddock aggregate by size.

At the population level, selectivity is unlikely to be homogeneous over space and time (Crone et al. 2013) suggesting that selectivity should be time-varying for most fisheries. Failing to account for time-varying selectivity can lead to biased estimates of key management quantities (Sampson and Scott 2012, Punt et al. 2014) while varying selectivity too much over time can introduce variance by not filtering out observation noise (Nielsen and Berg 2014). One challenge in allowing time-varying selectivity is the increase in the number of parameters to be estimated which can lead to issues of overparameterization. Nielsen and Berg (2014) demonstrated the ability of state-space models to efficiently estimate time-varying selectivity with the addition of only one parameter (a smoothing parameter) that allows selectivity to vary over time under different assumptions of correlation structure.

WHAM uses the following parameterization for double-logistic selectivity-at-age (S_a) which is the product of a logistic function and a negative logistic function:

$$S_a = \left(\frac{1}{1+\exp\left(\frac{-(a-a_{50\_1})}{k_1}\right)}\right)\left(\frac{1}{1+\exp\left(\frac{a-a_{50\_2}}{k_2}\right)}\right) \quad \text{(Eqn A1)}$$

where the a_{50_1} and a_{50_2} parameters are the age at 50% selectivity for each logistic curve and k_1 and $-k_2$ are the inverse of the slope of the curve at the respective a_{50}. The four selectivity parameters are estimated in WHAM as positive values between 0 and the number of age classes in the model. While a_{50} must be bounded between 0 and the number of age classes, WHAM restricts the range of the slopes of the selectivity curves to 1/12 and -1/12 for the logistic and negative logistic portions of the function, respectively, where 12 is the number of age classes in the 4X5Y Haddock models.

There is no logical explanation for a change in survey catchability in the time series, other than uncertainty in the constant 1.2 conversion factor used for the DFO survey index from 1970–1981 (Barrett and Barrett 2025). The only catchability scenario considered for explanation for 4X5Y Haddock is a separate q time block of 1970–1981 for the DFO survey index.

Model *m0* assumed logistic selectivity for the fishery and both surveys. As an alternative to the high *M* (i.e., *M* > 0.2) scenarios that had acceptable diagnostics, double-logistic (i.e., dome-shaped) selectivity scenarios were explored (Table A4). Model *m0_Vd1* with time-invariant double-logistic selectivity and model *m0_Vd53* with annual double-logistic selectivity (4 parameters × 53 years) did not converge (Table A5). To reduce the number of selectivity parameters to estimate, blocks of 10 years were used (one block of 3 years and 5 blocks of 10 years) but the model (*m0_Vd6*) did not converge and parameter estimates for the negative logistic portion of the selectivity curve for the first 5 blocks hit the upper parameter bound of 12.

Time-varying double-logistic selectivity options were explored by using a single fishery selectivity block with random effects. The four different correlation structures were explored (*iid*, *ar1_y, ar1_a, 2dar1*). These models either did not converge or parameter estimates for the negative logistic portion of the selectivity curve hit the upper parameter bound and diagnostics were poor (Table A5). Two potential hypotheses for the lack of convergence and selectivity parameters hitting bounds were:

1. the parameter bounds of 12 for the k parameters in Equation A1 were too low, restricting the range of selectivity curves that could be explored, or
2. there is little support for dome-shaped selectivity in the data.

In theory, the k_2 parameter can exceed the maximum number of age classes and as it increases, the double-logistic function approaches a logistic function. To evaluate potential support for a dome-shaped selectivity, selectivity was estimated using WHAM's age-specific selectivity pattern. To reduce the number of selectivity parameters to estimate, time-blocks of 10 years were used with 3 years in the sixth block (model *m0_Vas6*; Table A4, Table A5). The selectivity-at-age 6 was fixed to 1 in each time block. The selectivity-at-age for each time block was approximately logistic, with the exception of the last block (2020–2022) which appeared dome-shaped and the 2010–2019 block had selectivity-at-age 12+ of about 0.5. To assess the support for the dome-shaped selectivity at the end of the time-series, the selectivity blocks were set in blocks of 10 years with 3 years in the first block and 10 years in the sixth block (model *m0_Vas6a*, Table A4). The selectivity-at-age pattern for the 2013–2022 block was dome-shaped and all other blocks were approximately logistic (e.g., considering a smoother through the estimates of selectivity-at-age). The *m0_Vas6* and *m0_Vas6a* models did not converge (Table A5). Using these age-specific selectivity curves, and the *m0_Vas6a* model time block structure, model *m0_Vl5d1* was fit with logistic selectivity in all blocks with the exception of the 2013–2022 block (double-logistic). Model *m0_Vl5d1* converged but the model did not fit the DFO survey index well (fit comparable to model *m0*).

In an attempt to improve model fit to the DFO survey index, *q* was estimated in two time blocks (1970–1981 and 1982–2022, *m0_Vl5d1_q82*). The blocking for selectivity for this model was adjusted so that selectivity blocks didn't overlap the two *q* blocks. The model did not converge and the double-logistic selectivity parameters hit bounds. The double-logistic selectivity was changed to age-specific and ages 5 to 10 had to be fixed to one to get the model (*m0_Vl5as1_q82*) to converge. One more change to the selectivity blocks was made in an attempt to improve the residuals for the fishery age-composition data by allowing for lower selectivity for the 12+ age group in block 5 (*m0_Vl4as2_q82*); however, this did not improve overall model diagnostics (Table A5).

The shape of the DFO survey index selectivity was also explored as an additional option (without changing the assumption on *M*) to improve the fit to the index. Model *m0_Vl5d1_dDFO* assumed double-logistic selectivity for the DFO survey index and selectivity only dropped below 0.8 for age 12+ and provided little improvement to model diagnostics (Table A5).

A model (*m0_Vl5d1_asDFO_q82*) that assumed a dome-shaped selectivity for the fishery and DFO survey and had *q* estimated in two time blocks was fit. Double-logistic selectivity parameters for the DFO survey hit bounds so the selectivity was assumed age-specific and selectivity-at-age 5 to 10 had to be fixed at one for the model to converge. The fit to the DFO survey index was still not improved after 1982. One final model was fit to explore the influence of the changes in selectivity and *q* assumptions on *M*. Model *m0_Vl5d1_asDFO_q82_Me* estimated *M* as 0.471, a relatively small decrease (e.g., considering an *M* = 0.2 for model *m0*) from model *m0_Me* where *M* was estimated to be 0.552.

None of the models in Table A5 based on *m0* with different assumptions on selectivity were identified as candidate models for the stock; however, given that a model with annual selectivity has 106 (53 x 2) fishery selectivity parameters, it was recommended during the second peer-review meeting to explore a model with one set of fish selectivity parameters with annual selectivity estimated using random effects on the selectivity parameters. This model (*m0_Me1+_dln_sig0.05_L1_V2dar1*, Section 4.2) was identified as a candidate model to capture this uncertainty.

A1.3 Multinomial Age-Composition Likelihood Distribution

The Dirichlet was chosen as the age-composition distribution for model *m0* because it is self-weighting and model fits do not depend on potentially arbitrary weights put on different data sources used in the likelihood function. The influence of this decision on model fit was explored by fitting models with different age-composition distributions. For the multinomial distribution, the weight of the age-composition data sources are influenced by the input effective sample size (ESS) and the weight of the catch and indices is influenced by the coefficient of variation (CV) on these data sources. The multinomial distribution was chosen for exploration specifically so that the weights of different data sources could be adjusted in order to evaluate different hypotheses to explain conflicts in the survey index and age-composition data.

The initial model *m1* (Table A6) assumed an annual ESS equal to the number of tows for each survey, and an annual ESS equal to the number of unique fishing trips with age samples for the catch-composition. The CVs for the indices were estimated from the data and the CVs for catch were fixed at 0.05.

A key challenge with the diagnostics for models assuming a Dirichlet distribution was the fit to the DFO survey index (Table A6). There appears to be a mismatch between the index and the age-composition data for the fishery and DFO survey index, for which a high *M* can improve model diagnostics. Not all strong cohorts in the age-composition data for the catch appear in the survey composition data and the strong 2013 cohort that appears in the fishery and survey age-composition data does not contribute to very high biomass in the DFO survey index. While increasing *M* above 0.2 (e.g., model *m0_M2b_2010*) improved model diagnostics and may explain the mismatch, there are some other hypotheses such as decreased selectivity or catchability, or a decrease in growth that could explain this mismatch (e.g., Szuwalski 2022). Given that the model inputs address temporal changes in growth, the focus of alternative model configurations for *m1* was on changes in selectivity (i.e., dome-shaped selectivity for either the fishery or survey). The alternative hypothesis for this exploration of alternative forms of selectivity is a “hide them” scenario instead of a “kill them” scenario (e.g., Taylor and Methot 2013).

The focus on the model exploration for *m1* was using the multinomial distribution for the age-composition data so that the relative weights of different data sources could be adjusted in an attempt to get a model to fit the DFO survey index with *M* = 0.2. To reduce the residual pattern in the fishery catch age-composition data caused by the 2003 cohort, models were fit with catches in the south of NAFO area 4Xp (survey strata 482 and 483, see Barrett and Barrett

2025) removed based on the hypothesis that these catches include Haddock from EGB. Selectivity could not be estimated for the ITQ survey (similar to model *m0*) so it was removed as a data source. Model *m1* was therefore defined as model *m0* with catches in the south of 4Xp removed and a multinomial distribution for the age composition data with initial ESS as described for model *m0_AC_mult_DFOcv0.1*. Model *m1* did not fit the DFO survey index so the weight of the index was increased by reducing the CV by half (model *m1_DFO0.5cv*) and also by setting the CV to 0.1 for all years except 1977 where the index is driven by one large tow (model *m1_DFOcv0.1;* Table A6). The decrease in CV to 0.1 did improve the fit of the survey, but resulted in strong residual patterns in the fishery and survey age-composition data (e.g., a block of negative residuals for the 12+ age group). A double-logistic selectivity was assumed for the fishery over the entire time series (*m1_DFO_cv0.1_Vd1*) and for a time block of 2003–2022 (*m1_DFOcv0.1_Vl33d1*). Model *m1_DFOcv0.1_Vl33d1* had improved fishery age-composition residuals and improved fit to the DFO survey index but poor residuals for the survey age-composition data (Table A6). The DFO survey selectivity was changed to double-logistic (model *m1_DFOcv0.1_Vl33d1_dDFO*) and age-specific and ESSs for the survey age-composition data were increased. Model diagnostics were improved with the double-logistic selectivity but increases in the ESS forced the DFO survey index fit to fall apart (similar fit to the initial model *m0* fit). A model with age-based selectivity for the DFO survey estimated selectivity-at-age was very similar to the double-logistic selectivity model and wasn't considered further. Residual patterns (blocks of negative residuals for age 12+) in the fishery age-composition data suggested that a double-logistic selectivity estimated in age blocks may improve the fit. Model *m1_DFOcv0.1_Vd5_dDFO* had acceptable diagnostics (Table A6) with the fishery selectivity estimated as double-logistic in 5 time blocks (one block of 13 years and 4 blocks of 10 years).

Changes to *M* (increasing to 0.25 and trying to estimate within the model) resulted in failure to converge, caused by the double-logistic parameters hitting bounds. As a final exploration for model *m1*, random effects on fishery selectivity were explored for model *m1_DFOcv0.1_Vd1* and random effects on the DFO survey selectivity were explored for model *m1_DFOcv0.1_Vd5_dDFO*. Random effects for *m1_DFOcv0.1_Vd1* improved the retrospective patterns and AIC but model *m1_DFOcv0.1_Vd5_dDFO* (selectivity in blocks without random effects) had a lower AIC (Table A6). Model *m1_DFOcv0.1_Vd5_dDFO* with random effects did not converge.

From the exploration of model configurations for *m1*, model *m1_DFOcv0.1_Vd5_dDFO* had the best fit; however the assumption of dome-shaped selectivity for the fishery and for the survey was deemed inappropriate and no model based on the *m1* configuration was identified as a candidate model for the stock.

A1.4 Other Age-Composition Likelihood Distributions

Model *m0_AC_mult* (Table A7) assumed a multinomial distribution using ESSs and CVs as described above for model *m1*. Model *m0_AC_mult* did not converge so age-composition ESSs were manually adjusted to 70 for the catch and 40 (*m0_ AC_mult_70_40*) for the index but did not converge (values arbitrarily chosen based on values used for GB haddock (NFSC In Prep[1]). Model *m0_ AC_mult_200_40* converged but the diagnostics were poor and similar to *m0* (Table A7). In an attempt to improve the model fit to the DFO survey index, the CV on the index was decreased by 50% and also decreased to 0.1 (models *m0_ AC_mult_DFO0.5cv* and *m0_ AC_mult_DFOcv0.1*). Although fits to the DFO survey index were improved, the model diagnostics were poor. Temporal blocking was identified in the one-step ahead residuals for the catch and DFO survey age composition data from model *m0_ AC_mult_DFOcv0.1* where blocks of negative residuals were observed beginning around 2004 for the catch and 2010 for the DFO survey index. Models with logistic-normal and Dirichlet-multinomial distributions did not converge. The decision to pool zero observations (composition data) with adjacent age classes for the Dirichlet likelihood was arbitrary but a model was fit treating zeros as missing as a sensitivity and this decision appeared inconsequential with only a slight difference in *F* in the mid-2000s.

A1.5 Data Inputs

The influence on model fits of different data sources was explored. These data sources were:

- the ITQ survey index and age-composition data,
- catch and survey age-composition data estimated using age 9+ as the plus group,
- catch in the south of NAFO area 4Xp (survey strata 482 and 483),
- data inputs before 1985 excluded (based on uncertainty in ages before 1985, Wang et al. 2017), and
- DFO survey index (estimated assuming a delta-lognormal distribution).

Sensitivity model runs for these uncertainties were conducted and presented in the second peer-review meeting for each of the three candidate models (model *m0_M2b_2010,* model *m0_Me1+_sig0.05*, and model *m1_DFOcv0.1_Vd5_dDFO*).

Sensitivities for model *m0_M2b_2010* had little influence on diagnostics for model estimates of *F*, recruitment, or SSB. One exception was the *q* estimate for the model with 9+ as the plus group which resulted in a difference in scale for *F*, recruitment, and SSB, as well as a lower CV for these metrics.

Sensitivities for model *m0_Me1+_sig0.05* had little influence on diagnostics for model estimates of *F*, recruitment, or SSB other than differences in scale related to different estimates of *q*.

Sensitivities for model *m1_DFOcv0.1_Vd5_dDFO* had little influence on diagnostics or model estimates of *F*, recruitment, or SSB; however, for the sensitivity for dropping data pre-1985, the 5 blocks for selectivity were changed to 4 blocks (two 9-yr and two 10-yr blocks). The first two blocks needed to be set to logistic for the model to converge (double-logistic parameters hit bounds). This model had a much lower CV and different *q* (although similar trends in *F*, recruitment, and SSB).

A1.6 Random Effects on NAA

Model *m0* was fit with random effects on NAA to explore alternative hypotheses to an increase in *M* to improve model diagnostics. The correlation structure for age-1 (age of recruitment) was decoupled from that for ages 2+. Model *m0_M0.2_NAARE_iid* was fit with random effects on NAA (age 1 separate from ages 2+) assuming deviations by age and year but uncorrelated (Table A8). Models with NAA deviations correlated by age (*m0_M0.2_NAARE_ar1_a*), year (*m0_M0.2_NAARE_ar1_y*), and age/year (*m0_M0.2_NAARE_2dar1*) were also fit. The model with the best fit was model *m0_M0.2_NAARE_2dar1* (Table A8) and this model was selected as the candidate model for the stock as an alternative hypothesis to the high *M*. The model was revised (during the second peer-review meeting) using the Multinomial distribution and decreasing the CV on the DFO survey index by 50%. The model was revised again (after the second peer-review meeting) to turn off the bias-correction (*BC*) in the estimation of the mean NAA in the projection of the process-error. The final model used was *m0_M0.2_NAARE_2dar1_DFO0.5cv_noBC* (Section 4.2).

Table A1. Configuration of model m0, m0_M2b_2010, m0_Me1+_sig0.05, m1, m1_DFOcv0.1_Vd5_dDFO.

Model feature	***m0***	***m0_M2b_2010***	***m0_Me1+_sig0.05***	***m1***	***m1_DFOcv0.1_Vd5_dDFO***
Model years	1970 – 2022	1970 – 2022	1970 – 2022	1970 – 2022	1970 – 2022
Modeled age classes	1 – 12+	1 – 12+	1 – 12+	1 – 12+	1 – 12+
Catches	All	All	All	Excludes catches in survey strata 482-483	Excludes catches in survey strata 482-483
Fishery Fleet structure	Single fleet	Single fleet	Single fleet	Single fleet	Single fleet
Fishery selectivity	Logistic, annual blocks	Logistic, annual blocks	Logistic, annual blocks	Logistic, annual blocks	Double-logistic with one 13-yr and five 10-yr blocks
Survey	DFO survey [1970-2022 without 2021]	DFO survey [1970-2022 without 2021]	DFO survey [1970-2022 without 2021]	DFO survey [1970-2022 without 2021]	DFO survey [1970-2022 without 2021]
Survey selectivity	Single block, time-invariant, logistic	Single block, time-invariant, logistic	Single block, time-invariant, logistic	Single block, time-invariant, logistic	Single block, time-invariant, double-logistic
Stock recruitment model	Mean recruitment with log deviations estimated as fixed effects with sigma = 1	Mean recruitment with log deviations estimated as fixed effects with sigma = 1	Mean recruitment with log deviations estimated as fixed effects with sigma = 1	Mean recruitment with log deviations estimated as fixed effects with sigma = 1	Mean recruitment with log deviations estimated as fixed effects with sigma = 1
Natural mortality rate	0.2	0.484 [1970-2009] 0.621 [2010-2022]	0.227 – 1.21 [Time-varying based on survey AAA-1]	0.2	0.2
Likelihood function for fishery catch and survey index data	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal
Likelihood function for catch and survey age-composition data	Dirichlet, pooling zeros	Dirichlet, pooling zeros	Dirichlet, pooling zeros	Multinomial	Multinomial with CV on DFO survey index as 0.1
Random effects (process errors)	None	None	None	None	None

Table A2. *Model configurations explored with alternative M scenarios*

Model	*M*	Description
m0	0.2	Initial assumption
m0_M0.412	0.412	Mean estimate from life history parameters
m0_Me	0.552	1 block: model-estimated
m0_Mgrowth	Time-varying	Based on annual von Bertalanffy growth estimates
m0_Matagegrowth	Age- and time-varying	Based on annual von Bertalanffy growth estimates
m0_M2b_2004	0.466 and 0.559	2 blocks: 1970–2003 and 2004–2022 (both model-estimated)
m0_M2b0.2_2004	0.2 and 0.422	2 blocks: 1970–2003 (fixed) 2004–2022 (model-estimated)
m0_M2b0.2_2004_q2b	0.2 and 0.376	2 blocks: 1970–2003 (fixed) 2004–2022 (model-estimated); *q* estimated in 2 blocks [1970–1981;1982–2022]
m0_M2b_2010	0.479 and 0.614	2 blocks: 1970–2009 and 2010–2022 (both model-estimated)
m0_M2b0.2_2010	0.2 and 0.437	2 blocks: 1970–2009 (fixed) 2010–2022 (model-estimated)
m0_M2b0.2_2010_q2b	0.2 and 0.295	2 blocks: 1970–2009 (fixed) 2010–2022 (model-estimated); *q* estimated in 2 blocks [1970–1981;1982–2022]
m0_Me1+	Time-varying	Model-estimated using DFO survey abundance (ages 1+) as a covariate
m0_Me1+_sig0.05	Time-varying	*m0_Me1+* with index cv = 0.05
m0_Me1+_sig0.1	Time-varying	*m0_Me1+* with index cv = 0.1
m0_Me1+_log	Time-varying	Model-estimated using DFO survey log(abundance) (ages 1+) as a covariate
m0_Me1+_logsig0.1	Time-varying	*m0_Me1+_log* with index cv = 0.1
m0_Me1+_ logsig0.05	Time-varying	*m0_Me1+_log* with index cv = 0.05
m0_Mwaa	Age- and time-varying	Default WHAM weight-at-age method
M0_a	Age-varying	0.412 (ages 1–4); model-estimated age-specific (age 5+)
M0_a2	Age-varying	model-estimated age-specific (ages 1–4); 0.2 (age 5+)

Table A3. Summary of diagnostics for model m0 and alternative M scenarios (that converged) based on model m0 structure.

Model	*M*	N parameters	AIC	ΔAIC	ρSSB	ρ*F*	ρRec	DFO Survey Residual Trend[1]
m0	0.2	228	-6243.7	240.7	1.135	-0.453	0.977	Yes
m0_M0.412	0.412	228	-6321.4	163.0	0.923	-0.422	0.972	Yes
m0_Me	0.552	229	-6394.3	90.1	-0.690	3.887	-0.715	No
m0_Mgrowth	0.285 – 0.553	228	-6386.4	98.0	0.287	-0.249	0.241	No
m0_Matagegrowth	0.181 – 1.23	228	-6249.1	235.3	0.695	-0.375	0.569	Yes
m0_M2b_2004	0.466 and 0.559	232	-6469.0	15.4	0.472	-0.230	0.405	No
m0_M2b0.2_2004	0.2 and 0.422	231	-6400.1	84.3	-0.010	-0.085	-0.169	Yes
m0_M2b0.2_2004_q2b	0.2 and 0.376	234	-6474.9	11.5	0.139	-0.144	0.040	Yes
m0_M2b_2010	0.479 and 0.614	232	-6484.4	0	-0.072	0.036	-0.166	No
m0_M2b0.2_2010	0.2 and 0.437	231	-6348.7	135.7	-0.062	0.029	-0.174	Yes
m0_M2b0.2_2010_q2b	0.2 and 0.295	234	-6459.8	24.6	0.109	-0.039	0.045	Yes
m0_Me1+	0.227 – 1.29	231	-5792.3	-	0.042	-0.166	-0.036	No
m0_Me1+_logsig0.05	0.382 – 0.669	231	-6297.2	-	-0.063	-0.031	-0.107	No
m0_Me1+_logsig0.1	0.439 – 0.626	231	-6294.0	-	-0.054	-0.042	-0.106	No
m0_Me1+_sig0.05	0.228 – 1.21	231	-5790.6	-	0.084	-0.192	0.021	No
m0_Me1+_sig0.1	0.228 – 1.14	231	-5802.1	-	0.073	-0.181	0.016	No
m0_Mwaa	0.183 – 0.258	229	-6242.4	-	1.110	-0.488	0.840	Yes
m0_Ma	0.334 – 0.765	236	-6453.4	31.0	-0.443	0.809	-0.393	No

Notes: ΔAIC only reported for models with identical data inputs, ρ = average Mohn's rho for a 7-yr peel, green shading indicates desirable model diagnostics.

[1] "DFO Survey Residual Trend" = trend over time in the DFO survey index residuals, caused by a long series of positive residuals around 1990-2010. "No trend" = 95% confidence bands for regression of residuals on year overlaps zero at all ages.

Table A4. Model configurations explored with alternative selectivity scenarios.

Model	Fishery	DFO Survey	Notes
m0	L [53b]	L	-
m0_Vd53	DL [53b]	L	-
m0_Vd1	DL [1b]	L	-
m0_Vd1_iid	DL [1b] RE	L	-
m0_Vd1_ar1_a	DL [1b] RE	L	-
m0_Vd1_ar1_y	DL [1b] RE	L	-
m0_Vd1_2dar1	DL [1b] RE	L	-
m0_Vd6	DL [6b]	L	five 10-yr blocks, one 3-yr block
m0_Vd6a	DL [6b]	L	one 3-yr block, five 10-yr blocks
m0_Vas6	AS [6b]	L	five 10-yr blocks, one 3-yr block, SAA 6 fixed at 1
m0_Vas6a	AS [6b]	L	one 3-yr block, five 10-yr blocks, SAA 6 fixed at one
m0_Vas6a_fix5-7	AS [6b]	L	one 3-yr block, five 10-yr blocks, SAA 5-7 fixed at 1
m0_Vl4as2_q82	L [4B] AS [2b]	L	year blocks [6,6,11,12,9,9], SAA 5-10 fixed at 1
m0_Vl5as1_q82	L [5B] AS [1b]	L	year blocks [6,6,11,11,10,10], SAA 5-10 fixed at 1, *q* estimated in 2 blocks [1970-1981; 1982-2022]
m0_Vl5d1	L [5B] DL [1b]	L	one 3-yr block, five 10-yr blocks
m0_Vl5d1_asDFO_q82	L [5B] DL [1b]	AS	year blocks [6,6,11,11,10,10], SAA 5-10 fixed at 1, *q* estimated in 2 blocks [1970-1981; 1982-2022]
m0_Vl5d1_asDFO_q82_ Me [M= 0.471]	L [5B] DL [1b]	AS	year blocks [6,6,11,11,10,10], SAA 5-10 fixed at 1, *q* estimated in 2 blocks [1970-1981; 1982-2022]. *M* is model-estimated
m0_Vl5d1_dDFO_q82	L [5B] DL [1b]	DL	year blocks [6,6,11,11,10,10] , *q* estimated in 2 blocks [1970-1981; 1982-2022]
m0_Vl5d1_dDFO	L [5B] DL [1b]	DL	one 3-yr block, five 10-yr blocks
m0_Vl5d1_q82	L [5B] DL [1b]	L	year blocks [6,6,11,11,10,10]

Notes: "L" = logistic; "DL" = double-logistic, "AS" = age-specific; "RE" = random effects, "b" = time block, "SAA" = selectivity-at-age

Table A5. Summary of diagnostics for model m0 and alternative fishery and survey selectivity scenarios based on model m0 structure.

Model	Fishery Selectivity	Survey Selectivity	N parameters	AIC	ΔAIC	ρSSB	ρ*F*	ρRec	Converged	DFO Survey Residual Trend[1]
m0	L [53b]	L	228	-6243.7	181.8	1.135	-0.453	0.977	Yes	Yes
m0_Vd53	DL [53b]	L	334	9285.7	NR	1.586	-0.485	1.550	No	Yes
m0_Vd1	DL [1b]	L	126	-6067.1	358.4	1.456	-0.457	1.409	No	Yes
m0_Vd1_iid	DL [1b] RE	L	127	-6081.3	NR	1.474	-0.471	1.345	No	Yes
m0_Vd1_ar1_a	DL [1b] RE	L	128	-6063.1	362.4	1.456	-0.457	1.409	Yes	Yes
m0_Vd1_ar1_y	DL [1b] RE	L	128	-6094.0	NR	1.480	-0.462	1.275	No	Yes
m0_Vd1_2dar1	DL [1b] RE	L	129	-6133.6	291.9	1.478	-0.475	1.484	Yes	Yes
m0_Vd6	DL [6b]	L	146	-6123.8	NR	1.412	-0.449	1.485	No	Yes
m0_Vd6a	DL [6b]	L	146	-6136.7	NR	1.429	-0.470	1.767	No	Yes
m0_Vas6	AS [6b]	L	188	-6156.1	NR	1.570	-0.490	1.674	No	Yes
m0_Vas6a	AS [6b]	L	188	-6172.7	NR	1.676	-0.494	2.183	No	Yes
m0_Vas6a_fix5-7	AS [6b]	L	176	-6158.6	NR	1.691	-0.512	2.208	No	Yes
m0_Vl4as2_q82	L [4b] AS [2b]	L	145	-6335.9	89.6	1.205	-0.448	1.699	Yes	Yes
m0_Vl5as1_q82	L [5b] AS [1b]	L	141	-6354.9	70.6	0.773	-0.366	1.148	Yes	Yes
m0_Vl5d1	L [5B] DL [1b]	L	136	-6193.1	232.4	1.277	-0.455	1.588	Yes	Yes
m0_Vl5d1_asDFO_q82	L [5B] DL [1b]	AS	143	-6349.3	76.2	0.769	-0.344	1.114	Yes	Yes
m0_Vl5d1_asDFO_q82_Me [M = 0.471]	L [5B] DL [1b]	AS	144	-6425.5	0	-0.256	0.344	0.011	Yes	No
m0_Vl5d1_dDFO	L [5B] DL [1b]	DL	138	-6217.2	208.3	1.290	-0.449	1.635	Yes	Yes
m0_Vl5d1_dDFO_q82	L [5B] DL [1b]	DL	141	-6353.8	NR	0.797	-0.351	1.091	No	Yes
m0_Vl5d1_q82	L [5B] DL [1b]	L	139	-6342.7	NR	0.718	-0.345	0.935	No	Yes

Notes: "L" = logistic; "DL" = double-logistic, "AS" = age-specific; "RE" = random effects, "b" = time block, "NR" = not reported for models that did not converge "NS" = not significant (95% confidence bands for regression of residuals on year overlaps zero at all ages), ρ = average Mohn's rho for a 7-yr peel, green shading indicates desirable model diagnostics.

[1] "DFO Survey Residual Trend" = trend over time in the DFO survey index residuals, caused by a long series of positive residuals around 1990-2010. "No trend" = 95% confidence bands for regression of residuals on year overlaps zero at all ages.

Table A6. *Summary of diagnostics for model m1 and alternative fishery and survey selectivity scenarios based on model m0 structure.*

Model	Fishery Selectivity	Survey Sel	N param	AIC	ΔAIC	ρSSB	ρ*F*	ρRec	Converged	DFO Survey Residual Trend[1]
m1	L [53b]	L	228	-6288.1	-	1.108	-0.466	0.895	Yes	Yes
m1_DFO0.5cv	L [53b]	L	226	3781.9	-	0.654	-0.340	0.717	Yes	Yes
m1_DFOcv0.1	L [53b]	L	226	4049.8	503	0.449	-0.292	0.735	Yes	No
m1_DFOcv0.1_Vd1	DL [1b]	L	124	4038.8	492	0.472	-0.324	0.352	Yes	No
m1_DFOcv0.1_Vd1_iid_dDFO	DL [1b]	L	127	3643.7	96.9	-0.028	-0.023	0.440	Yes	No
m1_DFOcv0.1_Vd1_2dar1_dDFO	DL [1b]	L	129	3590.6	43.8	-0.093	0.018	0.428	Yes	No
m1_DFOcv0.1_Vl33d1	L [33b] DL [1b]	L	190	3856.4	309.6	0.391	-0.295	0.802	Yes	No
m1_DFOcv0.1_Vl33d1_asDFO	L [33b] DL [1b]	AS	196	3635.4	88.6	0.185	-0.187	0.688	Yes	No
m1_DFOcv0.1_Vl33d1_dDFO	L [33b] DL [1b]	DL	192	3643.0	96.2	0.158	-0.171	0.682	Yes	No
m1_DFOcv0.1_Vl33d1_essDFO100	L [33b] DL [1b]	DL	190	4287.7	-	0.432	-0.313	0.875	Yes	No
m1_DFOcv0.1_Vl33d1_essDFO400	L [33b] DL [1b]	DL	190	6769.1	-	1.072	-0.490	1.207	Yes	Yes
m1_DFOcv0.1_Vd5	DL [5b]	L	140	3833.5	NR	0.323	-0.252	0.681	No	No
m1_DFOcv0.1_Vd5_dDFO	DL [5b]	DL	142	3555.4	8.6	-0.116	0.065	0.279	Yes	No
m1_DFOcv0.1_Vd6_dDFO	DL [6b]	DL	146	3546.8	0	-0.072	0.020	0.334	Yes	No
m1_DFOcv0.1_Vd6_dDFO_M0.25	DL [6b]	DL	147	3537.4	NR	-0.033	0.026	0.146	No	No
m1_DFOcv0.1_Vd6_dDFO_Me [M = 0.339]	DL [6b]	DL	147	3537.4	NR	-0.033	0.026	0.146	No	No

Notes: "L" = logistic; "DL" = double-logistic, "AS" = age-specific, "RE" = random effects, "b" = time block, "NR" = not reported for models that did not converge, ΔAIC only reported for models with identical data inputs (DFO survey index CV = 0.1), "NS" = not significant (95% confidence bands for regression of residuals on year overlaps zero at all ages), ρ = average Mohn's rho for a 7-yr peel, green shading indicates desirable model diagnostics.
1 "DFO Survey Residual Trend" = trend over time in the DFO survey index residuals, caused by a long series of positive residuals around 1990-2010. "No trend" = 95% confidence bands for regression of residuals on year overlaps zero at all ages.

Table A7. *Summary of diagnostics for model m0 and alternative age-composition likelihood distribution scenarios based on model m0 structure.*

Model	**Age-composition Distribution**	**N parameters**	**AIC**	**ΔAIC**	**ρSSB**	**ρ*F***	**ρRec**	**Converged**
m0	Dirichlet (pooled)	228	-6243.7	13.6	1.135	-0.453	0.977	Yes
m0_AC_dir_miss0	Dirichlet (missing)	228	-6257.3	0	1.110	-0.450	0.944	Yes
m0_AC_dirmult	Dirichlet-multinomial	228	3516.0	–	1.033	-0.428	0.841	No
m0_AC_dirmult_200_40	Dirichlet-multinomial	228	4434.8	–	1.109	-0.456	0.660	No
m0_AC_dirmult_70_40	Dirichlet-multinomial	228	3462.6	–	1.129	-0.448	0.831	No
m0_AC_lnorm_miss0	Lognormal (missing)	228	-5461.6	–	0.815	-0.288	1.183	No
m0_AC_lnorm_pool0	Lognormal (pooled)	228	-5467.3	–	0.719	-0.314	0.746	No
m0_AC_mult	Multinomial	226	3512.0	–	1.158	-0.450	0.905	Yes
m0_AC_mult_200_40	Multinomial	226	4446.0	–	1.180	-0.462	0.662	Yes
m0_AC_mult_70_40	Multinomial	226	3458.6	–	1.219	-0.464	0.857	Yes
m0_AC_mult_DFO0.5cv	Multinomial	226	3804.1	–	0.652	-0.343	0.718	Yes
m0_AC_mult_DFOcv0.1	Multinomial	226	4074.3	–	0.445	-0.289	0.732	Yes

Notes: ΔAIC only reported for models with identical data inputs, ρ = average Mohn's rho for a 7-yr peel, green shading indicates desirable model diagnostics. Dash indicates not applicable.

Table A8. Summary of models with random effects on numbers-at-age.

Model	AIC	ΔAIC	ρSSB	ρ*F*	ρRec	Converged
m0	-6243.7	250.9	1.135	-0.453	0.977	Yes
m0_M0.2_NAARE_iid	-6382.2	112.4	0.849	-0.412	0.737	Yes
m0_M0.2_NAARE_ar1_a	-6436.9	57.7	0.429	-0.293	0.398	Yes
m0_M0.2_NAARE_ar1_y	-6488.1	6.5	0.316	-0.186	0.276	Yes
m0_M0.2_NAARE_2dar1	-6494.6	0	0.304	-0.195	0.255	Yes

Notes: ρ = average Mohn's rho for a 7-yr peel, AIC=Akaike information criterion, SSB=spawning stock biomass, F=fishing mortality, Rec=Recruitment, green shading indicates desirable model diagnostics.

APPENDIX B

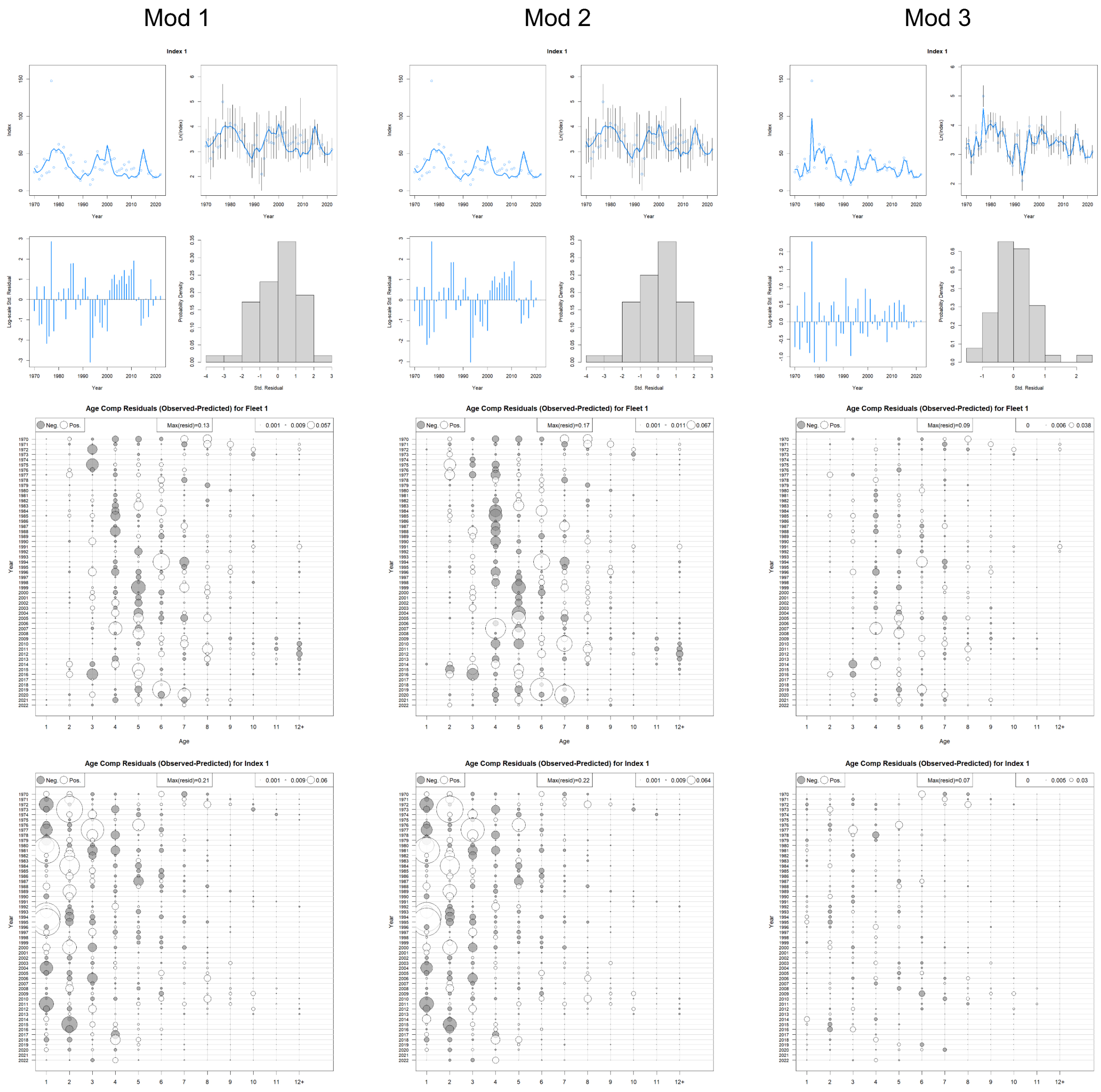

Figure B1. Diagnostics for the DFO survey index (top), age composition residuals for the fishery catch (middle), and age composition residuals for the survey index (bottom) for the 3 models.

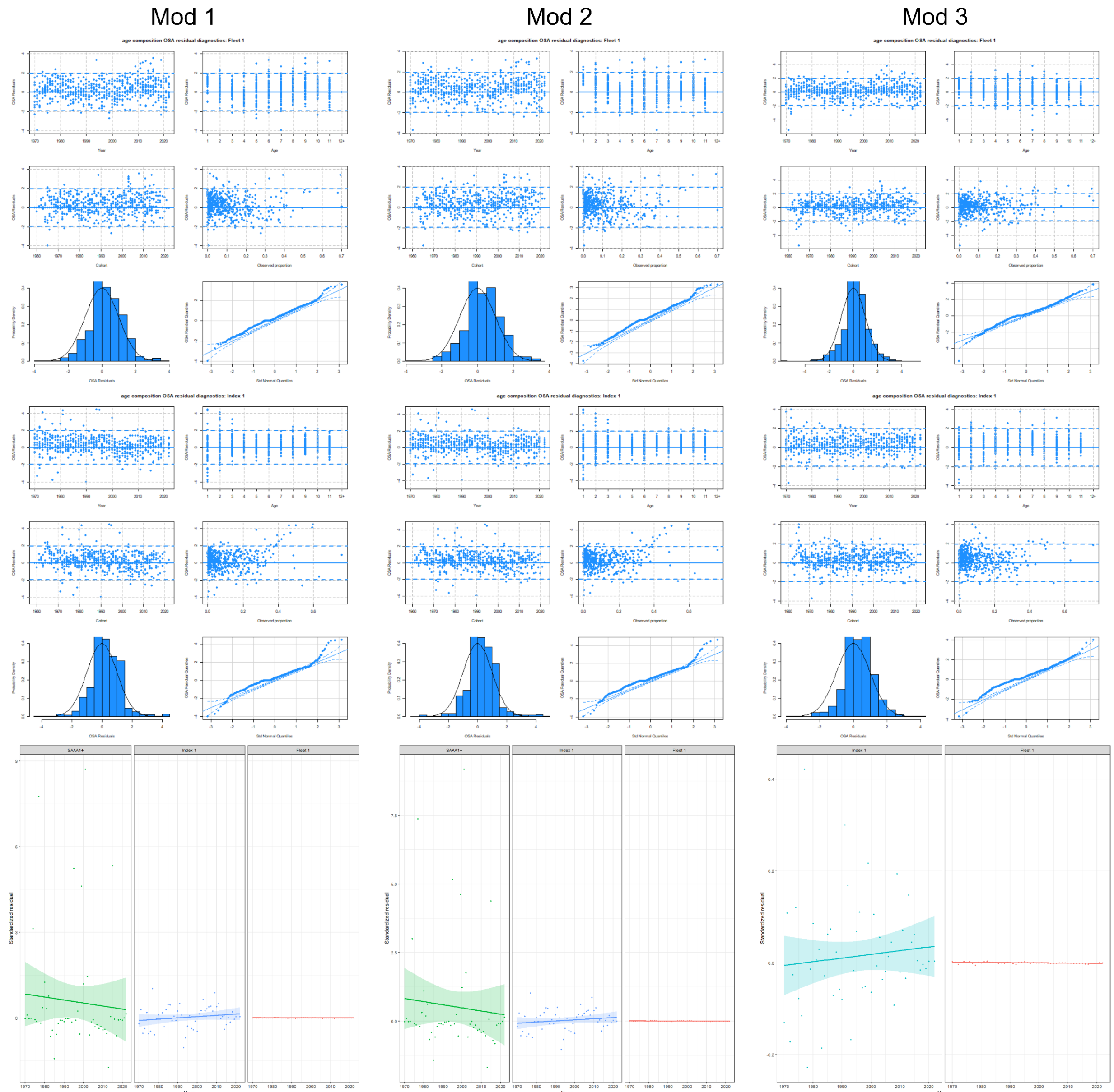

Figure B2. One step ahead residual diagnostics for the fishery catch (top) and the DFO survey index (middle), and standardized residual diagnostics for the fishery and survey index (bottom) for the 3 models.

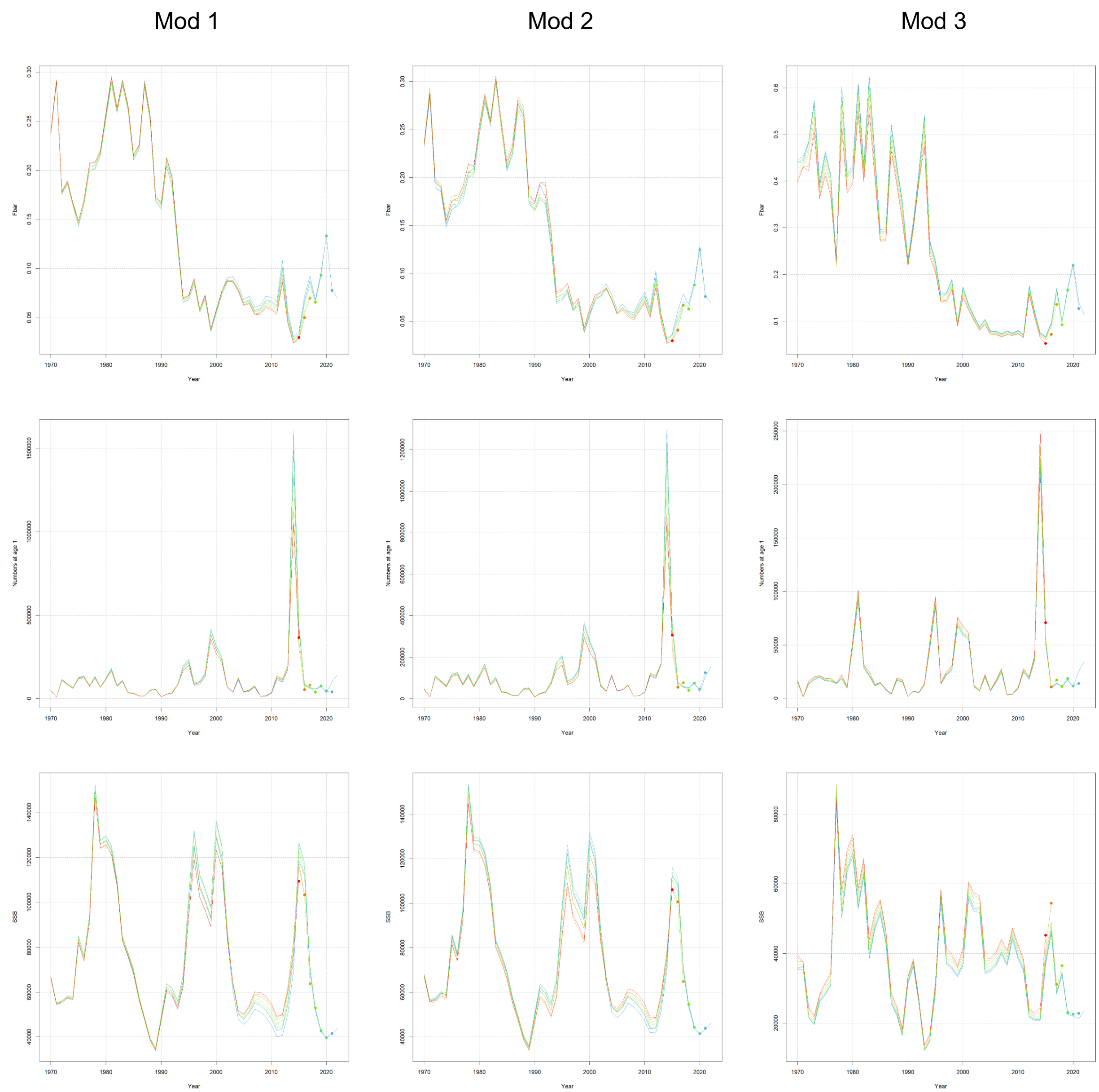

Figure B3. Retrospective analysis of mean fishing mortality rate (F_{bar}), recruitment (Numbers at age 1), and spawning stock biomass (SSB) for the 3 models.

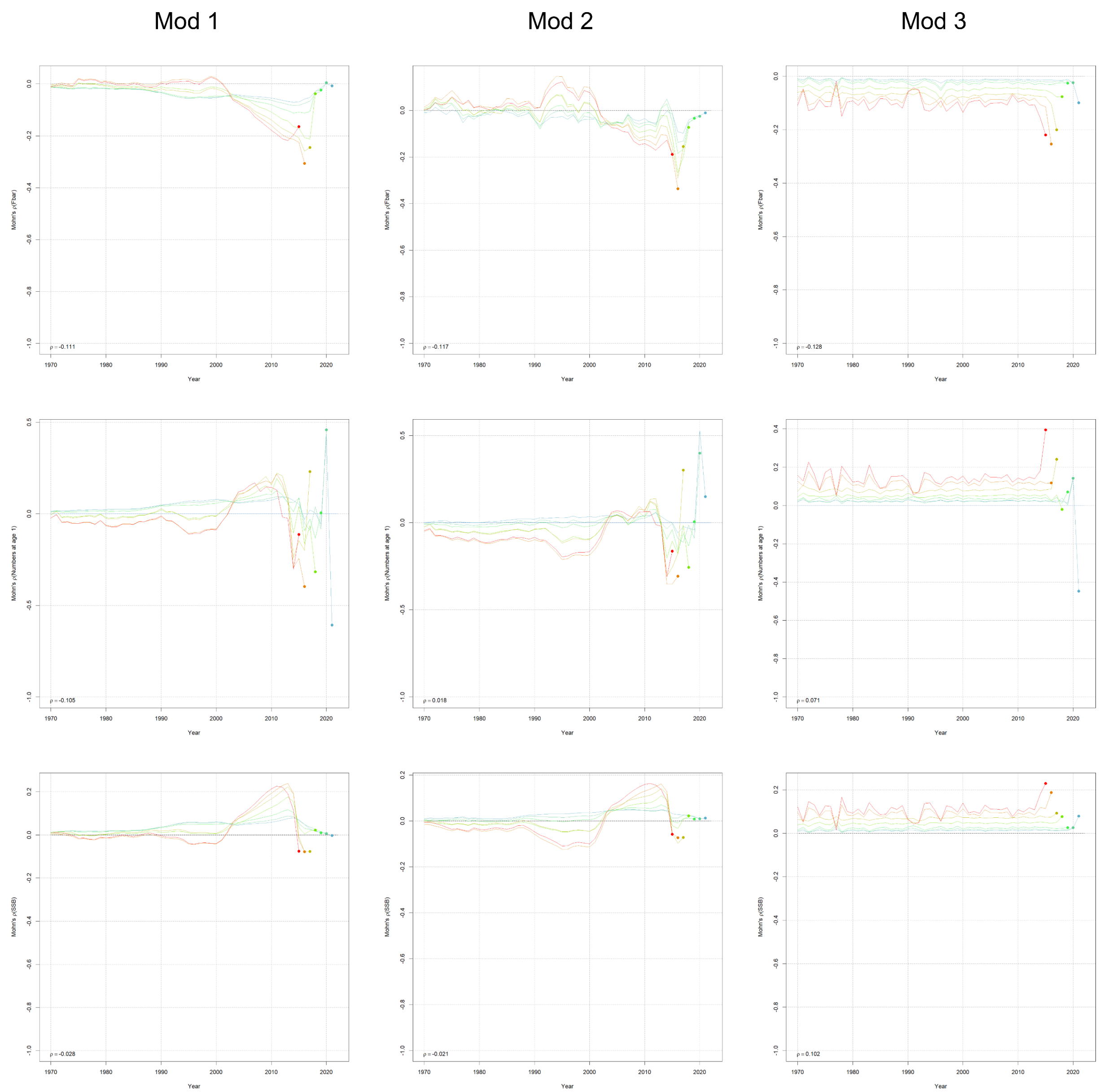

Figure B4. Retrospective analysis of mean fishing mortality rate (F_{bar}), recruitment (Numbers at age 1), and spawning stock biomass (SSB) using Mohn's rho (ρ) for the 3 models.

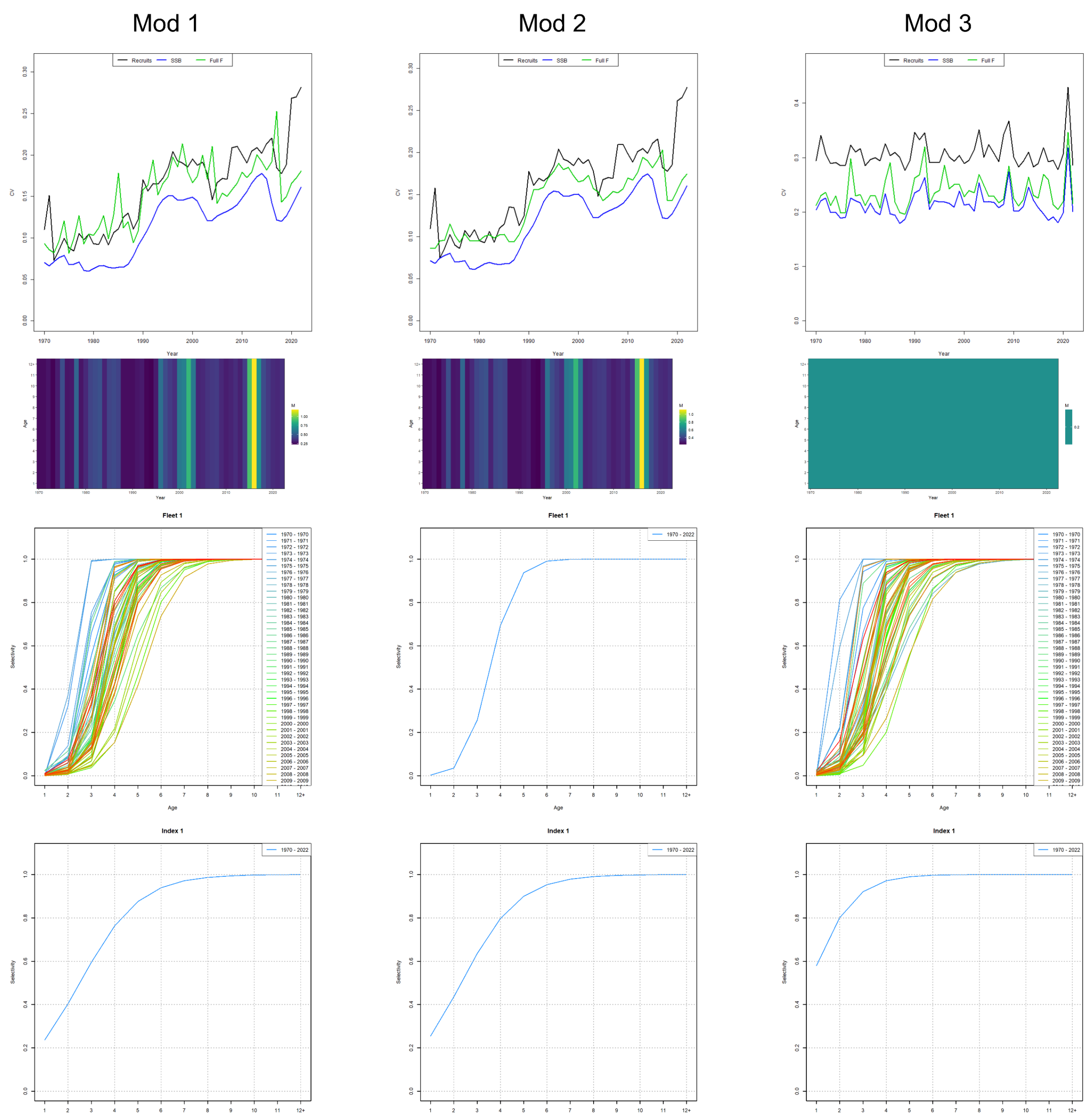

Figure B5. Coefficients of variation (CV) for spawning stock biomass (SSB), fishing mortality (F) and recruitment (Numbers at age 1, top), natural mortality (M) at age (middle), and fishery selectivity and survey selectivity (bottom two panels) for the 3 models.

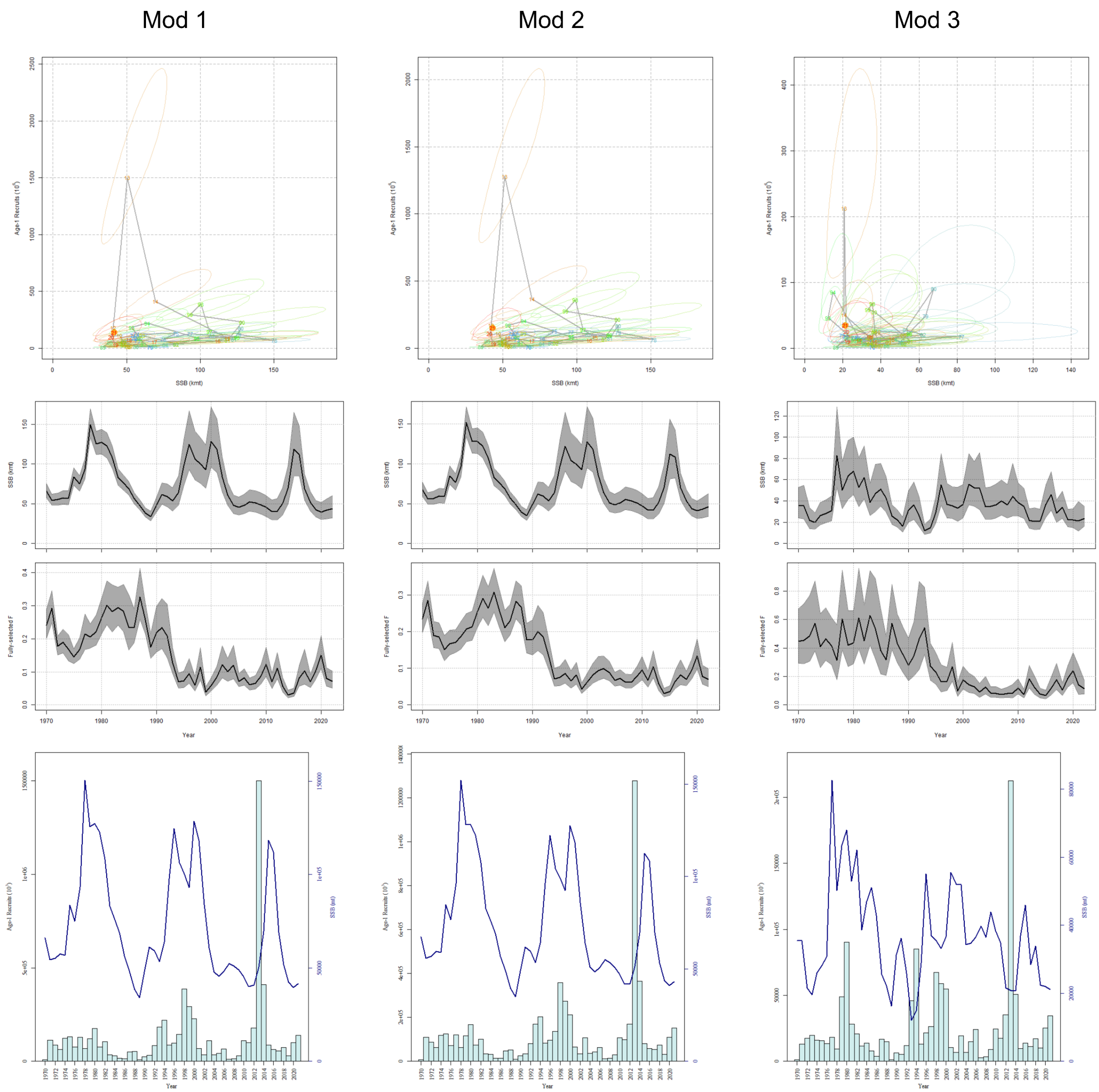

Figure B6. Relationship between recruitment at age 1 and spawning stock biomass (SSB) (top), model estimated SSB and model estimated fully selected fishing mortality (F) (middle), and model estimated SSB (line) and recruitment (bars, bottom) for the 3 models. Grey shading represents the 95% confidence intervals.