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

**Canadian Science Advisory Secretariat (CSAS)**

**Research Document 2023/039**

**Arctic Region and Ontario and Prairie Region**

# A modeling framework for stock assessment and harvest strategy evaluation for the Northwest Atlantic Fisheries Organisation Subarea 0+1 (offshore) Greenland Halibut (*Reinhardtius hippoglossoides*) fishery

Samuel D. N. Johnson and Sean P. Cox

Landmark Fisheries Research
213-2414 St Johns St
Port Moody, BC, V3H 2B1

May 2023

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/
csas-sccs@dfo-mpo.gc.ca


ISSN 1919-5044
ISBN 978-0-660-48810-3 Cat. No. Fs70-5/2023-039E-PDF

**Correct citation for this publication:**

Johnson, S.D.N, and Cox, S.P. 2023. A modeling framework for stock assessment and harvest strategy evaluation for the NAFO 0+1 (offshore) Greenland Halibut (*Reinhardtius hippoglossoides*) fishery. DFO Can. Sci. Advis. Sec. Res. Doc. 2023/039. iv + 88 p.

***Aussi disponible en français :***

*Johnson, S.D.N, and Cox, S.P. 2023. Cadre de modélisation pour l'évaluation du stock de flétan du Groenland (*Reinhardtius hippoglossoides*) de la sous-zone 0+1 (au large des côtes) de l'Organisation des pêches de l'Atlantique Nord-Ouest et des stratégies de pêche connexes. Secr. can. des avis sci. du MPO. Doc. de rech. 2023/039. iv + 94 p.*

## TABLE OF CONTENTS

## ABSTRACT

This paper presents an assessment modeling framework that integrates all available fishery monitoring and survey data into a Spatially Implicit Statistical Catch-At-Length (SISCAL) operating model for Northwest Atlantic Fishery Organization (NAFO) Subareas 0+1 (offshore) Greenland Halibut (GH-0+1). The model is subsequently used to provide both an assessment of stock status and productivity as well as a closed loop simulation framework for evaluating GH-0+1 feedback harvest strategies. The SISCAL model fit reasonably well to GH-0+1 data, as determined by standard goodness of fit metrics, although some sensitivities and data issues were noted. The retrospective behaviour of the model was also within reason. Simulation-evaluation self-tests also showed that the model was unlikely to be biased over a large number of simulated data sets. After model testing, SISCAL was used to condition a closed-loop simulation framework for testing GH-0+1 management procedures against performance metrics based on NAFO precautionary approach fishery management policy. As an illustrative example, we defined an adaptive model/index-based management procedure that set total allowable catches on a biennial basis, using decision rule parameters that were updated via simulated SISCAL stock assessments every 6 years. Simulated SISCAL stock assessments were fit to historical and simulated catch and biological data from six commercial fleets, differing by nation and gear type, and stock indices and length compositions from three fishery independent surveys. The three surveys included two existing offshore (NAFO Divisions 0A1CD) and inshore (NAFO Divisions 1A to F) research vessel surveys, as well as an additional proposed survey that will begin in 2022 and was assumed to encounter small fish in inshore waters of Divisions 0A and 0B. For comparison, a non-adaptive index-based method was also tested, where decision rule parameters are based off the initial SISCAL model for the entire simulation. The adaptive procedure performed well, keeping biomass above the limit reference point of $B_{lim} = 0.3B_{MSY}$ in all simulations, and avoiding the limit fishing mortality rate $F_{lim} = F_{MSY}$ with high probability. In contrast, the non-adaptive procedure ended up slightly overfishing the GH-0+1 stock, with biomass appearing to continue declining past the end of the simulation. Moreover, the non-adaptive procedure had a roughly neutral probability of exceeding $F_{lim}$, making it unacceptable under NAFO policy. We close with recommendations for future work to expand this framework to a full management strategy evaluation, enabling the development of a full harvest strategy for the GH-0+1 fishery consistent with fishery management policy.

# INTRODUCTION

The Greenland halibut (*Reinhardtius hippoglossides*) stock in North Atlantic Fisheries Organisation (NAFO) Subareas 0+1 is exploited in a transboundary fishery that has been expanding over the past several decades. Local traditional and subsistence fishing by Greenlanders has been recorded since as early as 1852 (Smidt 1969, Bowering and Brodie 1995) and, while commercial fishing has occurred since the early 1900s, catches have increased significantly since 1968 (Bowering and Brodie 1995, Treble and Nogueira 2020). Historically, from 1976 to 1992, both offshore and inshore fisheries were managed under a single total allowable catch (TAC); however, since 1993, the contemporary offshore NAFO 0+1 Greenland halibut fishery (referred to hereafter as GH-0+1) has been managed with its own TAC. Offshore TACs have ranged from around 11 kt (1995 - 2001) to as high as 36 kt in 2018 (Treble and Nogueira 2020).

There are currently limited information feedbacks available to adjust TACs in response to changes in Greenland Halibut stock status. Since 2016, TACs have been set using a moving average biomass index (Treble and Nogueira 2020), which was produced via the offshore research vessel (RV) survey in NAFO Divisions 0A1CD. While the survey was conducted and an index produced for 2019, significant gear and vessel differences affected gear performance, reducing confidence in the results (Nogueira and Treble 2020, Treble and Nogueira 2020); therefore, the fishery has been under a constant total allowable catch (TAC) since 2018. While there are other available data sources, such as the inshore survey of small fish in Divisions 1A-F, fishery catch-per-unit-effort (CPUE), and length composition, there are also well documented uncertainties that may weaken feedbacks based on all of them. For example, the inshore survey appears to target smaller fish, which are largely distinct from the larger fish targeted by commercial fisheries. Further, portions of these inshore areas are linked to inshore stocks that are considered sink populations, which if true may further erode information linkages between TAC decisions and the index over the whole region. Finally, without the aid of a model, it is generally difficult to determine year-class strength from length composition data collected over different surveys and fisheries that each have somewhat unique sampling designs and protocols.

In this paper, we present an example end-to-end (data-to-advice) assessment modelling framework in three parts:

a) estimating GH-0+1 stock status and biological reference points from fishery and survey data;

b) simulation-testing the assessment model to better understand the range of estimation performance (i.e., bias and precision of estimates) given the available data types, quality, and quantity; and

c) conditioning operating models for simulation-testing precautionary harvest strategies.

We use this framework to develop and test an example precautionary harvest strategy involving an adaptive model-/index-based management procedure (MP) for setting GH-0+1 TACs as future information accumulates from existing and planned new surveys.

The modelling framework in Part A develops a spatially integrated statistical catch-at-length (SISCAL) approach that integrates GH-0+1 data sources via interactions between stock dynamics and data-generating mechanisms. Throughout this document, we refer to this model as SISCAL-GH. SISCAL-GH produces Bayes posterior probability distributions for key biological and management parameters (e.g., $B_{MSY}$, $MSY$, $F_{MSY}$). In Part B, we develop and apply a simulation self-testing framework for SISCAL-GH in which bias and precision of assessment

estimates are evaluated based on simulated data sets. Finally, in Part C, we embed SISCAL-GH in a closed-loop simulation algorithm to generate new data, conduct an assessment, and set TACs on a predetermined temporal cycle into the future. We then apply an example MP assuming the assessment is updated every 6 years and sets TACs annually based on an index-based method similar to the existing approach. The index-based method uses survey catchability estimated from the assessment to estimate exploitable biomass. In the simulations, we provide realistic performance expectations of this approach since the assumed catchability of the index is updated in each assessment cycle, rather than incorrectly assuming that survey catchability is known perfectly over a 47-year simulation horizon. This adaptive approach is specifically designed to allow scientists and managers to design surveys, sampling designs, and protocols that optimize fishery performance (i.e., completing the "data-to-advice" loop) to the extent possible.

In Part A, we also conduct sensitivity analyses on several SISCAL-GH model assumptions, including larger GH-0+1 size-at-age, time-varying catchability for fishery CPUE, form of fishery selectivity, and temporal variability in natural mortality and bottom trawl catchability. We also discuss how selected sensitivities can be used as alternative operating model hypotheses in harvest strategy evaluation.

# METHODS

## SISCAL-GH: A TWO SEX STATISTICAL CATCH-AT-LENGTH OPERATING MODEL

The SISCAL-GH model is an age- and sex-structured population dynamics model fit to fishery-independent biomass indices and length compositions, and fishery-dependent landings, CPUE, and catch-at-length data (Figure 1). Fishery landings and length compositions are split by gear type and, where possible, nation, but aggregated over Subareas 0 and 1. The SISCAL-GH model is an implicitly spatial or 'areas-as-fleets' model, as only Canadian vessels fish in Subarea 0, and all other vessels fish in Subarea 1. Model notation is given in Table 1 and population dynamics and statistical model equations in Tables 2 and 3, respectively. Commercial fishing fleet definitions are given in Table 4. Below, we provide a narrative description of the model using these tables for equation references.

SISCAL-GH partitions the base model parameters into four subsets (Table 2, P.1 - P.4), consisting of:

- estimated parameters ($\theta^{est}$),
- nuisance catchability and variance parameters estimated conditionally on the leading parameters ($\theta^{cond}$),
- fixed parameters for growth, maturity-at-length, recruitment standard deviations, and fixed model variance for time-varying parameters ($\theta^{fixed}$), and
- parameters specifying the prior distributions ($\theta^{prior}$) for other model parameters.

Parameter membership in the fixed and estimated sets differ from the base model for some of the model variations considered in the sensitivity analysis.

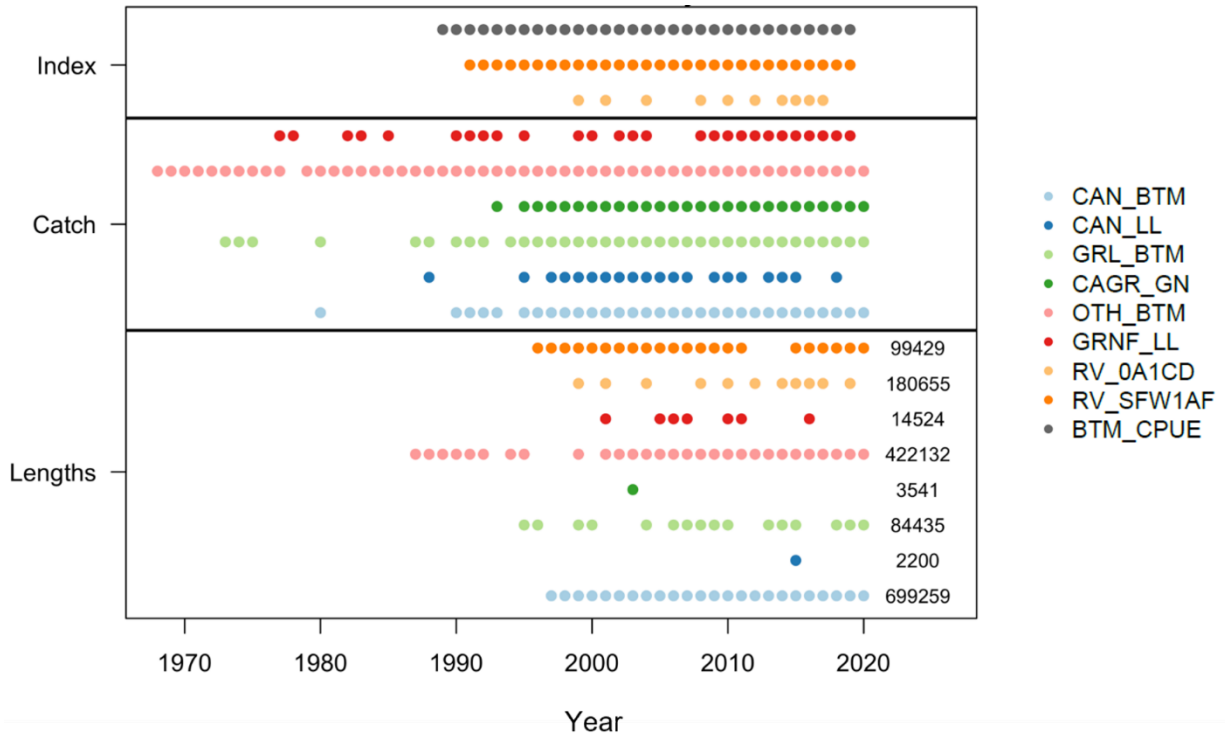


*Figure 1. Summary of available index, catch, and length-composition data for the SISCAL-GH stock assessment model as of 2020. Points indicate the presence of data in each year (x-axis) and fleet (colours explained in figure legend). For length composition data, the total sample size over all years is shown at the right hand end of the panel. CAN_BTM = Canada Bottom Trawl; CAN_LL = Canada Longline; GRL_BTM = Greenland Bottom Trawl; CAGR_GN = Canada Greenland Gillnet ; OTH_BTM = Other Bottom Trawl; GRNF_LL = Greenland Norway Faroe Islands Longline; RV_0A1CD = Research Vessel 0A1CD Survey; RV_SFW1AF = Research Vessel Shrimp and Fish West Greenland 1AF; BTM_CPUE = Bottom Trawl Catch per Unit Effort.*

*Table 1. Notation used in the specification of the SISCAL-GH stock assessment model, along with a description of each variable, and possible fixed values.*

| Symbol | Value | Description |
|---|---|---|
| $T$ | 53 | Total number of time steps 1965 - 2021 |
| $A$ | 35 | Plus group age-class |
| $L$ | 32 | Number of length bins (4cm width) |
| $t$ | $1,2,\ldots,T$ | Time step |
| $a$ | $1,2,\ldots,A$ | Age-class index |
| $l$ | $2,6,\ldots,112$ | Length-bin mid-points ($L = 32$ total length bins) |
| $g$ | $1,\ldots,9$ | Gear index as described in Table 4 ($g = 9$ only used in projections) |
| $x$ | 1,2 | Sex index for males ($x = 1$) and females ($x = 2$) |
| $B_0$ | - | Unfished female spawning stock biomass |
| $h$ | - | Beverton-Holt stock-recruitment steepness |
| $R_0$ | - | Unfished equilibrium age-1 recruitment |
| $S_{a,x}$ | - | Unfished equilibrium survivorship-at-age and sex |
| $\phi_0$ | - | Unfished equilibrium spawning biomass per recruit |

| Symbol | Value | Description |
|---|---|---|
| $\beta_1, \beta_2$ | $28, 10$ | Beta prior parameters for steepness (mean of .73, sd of 0.07) |
| $\omega_t$ | - | Annual recruitment processs error log-deviations |
| $\sigma_R$ | $1$ | Standard error of $\omega_t$ recruitment deviations |
| $q_g$ | - | Catchability coefficient for RV surveys ($g = 7{,}8{,}9$) |
| $q_t$ | - | Time-varying catchability coefficient for commercial CPUE index |
| $M_{x,t}$ | - | Time-varying natural mortality rate for males and females |
| $M_{0,x}$ | - | Time-averaged natural mortality rate for males and females |
| $\mu_M$ | $0.14$ | Natural mortality prior mean for males and females |
| $\sigma_M$ | $0.05$ | Natural mortality prior standard deviation for males and females |
| $\epsilon_{M,t}$ | - | Time-varying natural mortality random walk log-deviations |
| $L_{\infty,x}$ | $68, 72$ | Asymptotic length (cm) for males and females |
| $\sigma_{L,x}$ | $0.08, 0.11$ | CV in length-at-age distribution |
| $K_x$ | $0.10, 0.559$ | von Bertalanffy growth constant for males and females |
| $L_{1,x}$ | $16.26, 16.53$ | Length-at-age 1 for males and females |
| $c_1, c_2$ | $3.86E^{-6}, 3.22$ | Allometric length-weight transformation coefficients |
| $l_{50}, l_{95}$ | $67{,}78$ | Length-at-50% and -95% maturity |
| $L_{a,x}$ | - | Mean length-at-age (cm) for males and females |
| $w_{a,x}$ | - | Mean weight-at-age (cm) for males and females |
| $m_l$ | - | Proportion females mature-at-length |
| $m_a$ | - | Proportion females mature-at-age |
| $s_{l,g}$ | - | Mean selectivity-at-length $l$ for gear $g$ |
| $s_{a,x,g}$ | - | Mean selectivity-at-age $a$ for gear $g$ and sex $x$ |
| $L^A_{50,g}$ | - | Length-at-50% selectivity for ascending limb |
| $L^A_{95,g}$ | - | Length-at-95% selectivity for ascending limb |
| $L^D_{50,g}$ | - | Length-at-50% selectivity for descending limb |
| $L^D_{95,g}$ | - | Length-at-95% selectivity for descending limb |
| $N_{a,x,t}$ | - | Numbers-at-age $a$ for sex $x$ in year $t$ |
| $B_t$ | - | Female spawning biomass in year $t$ |
| $C_{g,t}$ | - | Observed total catch (biomass units) for gear $g$ in year $t$ |
| $C_{a,x,g,t}$ | - | Expected catch-at-age (numbers) $a$ and sex $x$ by gear $g$ in year $t$ |
| $C'_{a,x,g,t}$ | - | Expected catch-at-age (biomass units) $a$ and sex $x$ by gear $g$ in year $t$ |
| $B_t^{exp}$ | - | Allocation weighted average exploitable biomass in year $t$ |
| $U_{g,t}$ | - | Exploitation rate by gear $g$ in year $t$ |
| $U_t$ | - | Total exploitation rate in year $t$ |
| $I_{g,t}$ | - | Observed biomass/abundance index for gear $g \in \{7{,}8\}$ in year $t$ |
| $\hat{I}_{g,t}$ | - | Expected biomass/abundance index for gear $g \in \{7{,}8\}$ in year $t$ |
| $\tau_g$ | - | Standard deviation of biomass index observation log-residuals |
| $u_{l,x,g,t}$ | - | Observed length composition data for sex $x$ in gear $g$ at time $t$ |

| Symbol | Value | Description |
|---|---|---|
| $\hat{u}_{l,x,g,t}$ | - | Expected length composition data for sex $x$ in gear $g$ at time $t$ |
| $\hat{\tau}_{x,g}^{len}$ | - | Conditional MPDE of length composition sampling error, by sex and gear |
| $w_{B_0}$ | 10 | Weighting on unfished biomass improper Jeffreys prior |

*Table 2. Process and observation model equations for the SISCAL-GH stock assessment model. Equation number (No.) prefixes help readers partition equations into subsets, defined as P = parameters; G = growth and maturity; M = Mortality, S = Selectivity; EQ = Equilibrium characteristics; C = Catch equations; N = Numbers-at-age equations.*

| No. | Equation |
|---|---|
| (P.1) | $\Theta^{lead} = \left(B_0, \{\omega_t\}_{t\in 1:T}, h, M_x, \{\epsilon_t\}_{t\in 1:T}, \{L_{50,g}^A, L_{95,g}^A, L_{50,g}^D, L_{95,g}^D\}_{g\in 1:8}\right)$ |
| (P.2) | $\Theta^{cond} = \left(\{\log q_g\}_{g\in\{7,8\}}, \{\tau_{x,g}^{len}\}_{g\in 1:8}\right)$ |
| (P.3) | $\Theta^{fixed} = \left(l_{mat,50}, l_{mat,95}, \sigma_R, \sigma_M\right)$ |
| (P.4) | $\Theta^{priors} = \left(m_M, s_M, \{m_{L50,A,g}, m_{L95,A,g}, m_{L50,D,g}, m_{L95,D,g}, \sigma_g^{Sel}\}_{g\in 1:8}\right)$ |
| (G.1) | $l_{a,x} = L_1 + \left(L_1 - L_{\infty,x}\right)\cdot e^{-k_x(a-1)}$ |
| (G.2) | $D(l \mid a,x) = e^{-\left(frac\left(l-l_{a,x}\right)2\cdot\sigma_L\cdot l_{a,x}\right)^2}$ |
| (G.3) | $P(l \mid a,x) = \frac{D(l \mid a,x)}{\sum_{l'} D\,(l' \mid a,x)}$ |
| (G.4) | $w_{a,x} = c_1 l_{a,x}^{c_2}$ |
| (G.5) | $m_l = \left(1 + e^{-\log 19\frac{l-l_{mat,50}}{l_{mat,95}-l_{mat,50}}}\right)^{-1}$ |
| (G.6) | $m_a = \sum_l P\,(l \mid a, x=2) m_l$ |
| (M.1) | $M_{x,t} = \begin{cases} M_x & t = 1 \\ M_{x,t-1}\cdot e^{\sigma_M epsilon_t} & t > 1 \end{cases}$ |
| (M.2) | $M_{0,x} = \frac{1}{T}\sum_t M_{x,t}$ |
| (S.1) | $s_{l,g}^X = \left(1 + e^{-\log 19\frac{l-L_{50,g}^X}{L_{95,g}^X - L_{50,g}^X}}\right)^{-1}$ |
| (S.2) | $s_{l,g} = s_{l,g}^A \cdot s_{l,g}^D$ |
| (S.3) | $s_{a,x,g} = \sum_l P\,(l \mid a,x) s_{l,g}$ |
| (EQ.1) | $S_{a,x} = \begin{cases} 0.5 & a = 1, \\ S_{a-1,x}e^{-M_{0,x}} & 1 < a < A \\ S_{a-1,x}e^{-M_{0,x}}/(1-e^{-M_{0,x}}) & a = A. \end{cases}$ |
| (EQ.2) | $\phi = e^{-M_{0,x=2}}\cdot\sum_a S_{a,x=2}\cdot\bar{w}_{a,x=2}\cdot m_a$ |
| (EQ.3) | $R_0 = B_0/\phi$ |
| (EQ.4) | $N_{a,x}^{eq} = R_0\cdot S_{a,x}$ |

| No. | Equation |
|---|---|
| (C.1) | $N_{a,x,t+\delta_g} = N_{a,x,t+\delta_{g-1}} \cdot e^{-1\cdot(\delta_g-\delta_{g-1})M_{x,t}}$ |
| (C.2) | $N_{a,x,g,t} = N_{a,x,t+\delta_g} \cdot s_{a,x,g}$ |
| (C.3) | $B_{a,x,g,t} = N_{a,x,g,t} \cdot w_{a,x}$ |
| (C.4) | $B_{g,t} = \sum_{a,x} B_{a,x,g,t}$ |
| (C.5) | $B_t^{exp} = \sum_g \rho_g B_{g,t}$ |
| (C.6) | $C'_{a,x,g,t} = C_{g,t} \cdot \frac{B_{a,x,g,t}}{\sum_{a'} B_{a',x,g,t}}$ |
| (C.7) | $C_{a,x,g,t} = C'_{a,x,g,t}/w_{a,x}$ |
| (C.8) | $N_{a,x,t+\delta_g} = e^{-(\delta_g-\delta_{g-1})\cdot M_{x,t}} \cdot N_{a,x,t+\delta_{g-1}} - C_{a,x,g,t}$ |
| (C.9) | $U_{g,t} = C_{g,t}/B_t^{exp}$ |
| (N.1) | $SB_t = \sum_{a,x=2} m_a B_{a,x=2,t}$ |
| (N.2) | $R_{t+1} = \frac{a_p SB_t}{1+b_p SB_t} \cdot e^{\sigma_R \omega_{R,t}}$ |
| (N.3) | $N_{a,x,t+1} = \begin{cases} 0.5R_{t+1} & a = 1 \\ e^{-(1-\delta_G)M_{x,t}} \cdot N_{a-1,x,t+\delta_G} & 2 \le a \le A-1 \\ e^{-(1-\delta_G)M_{x,t}} \cdot \left(N_{a-1,xs,t+\delta_G} + N_{a,x,t+\delta_G}\right) & a = A. \end{cases}$ |

*Table 3. Statistical model prior and likelihood functions for the SISCAL-GH stock assessment model. The function $1(X)$ is the indicator function, taking value 1 when $X$ is true, and 0 when $X$ is false. Equation number (No.) prefixes help readers partition equations into subsets, defined as O = observation model; NLL = Negative log-likelihood for index data; LL = Length data Likelihood, Pr = Priors; OF = Objective Function.*

| No. | Equation |
|---|---|
| (O.1) | $q_{g,t} = \begin{cases} q_g & t = t_{g,1} \\ q_{g,t-1} e^{\epsilon_{q,g,t}} & t_{g,1} < t < t_{g,n_g} \end{cases}$ |
| (O.2) | $\hat{I}_{g,t} = q_{g,t} B_{g,t}$ |
| (O.3) | $\hat{u}_{l,x,g,t} = \frac{\sum_a P(l \mid a,x) s_{a,x,g} N_{a,x,g} e^{-f_g Z_{a,x,t}}}{\sum_{l'} \sum_{a'} P(l \mid a',x) s_{a',x,g} N_{a',x,g} e^{-f_g Z_{a',x,t}}}$ |
| (NLL.1) | $n_g = \sum_{t=1}^{T} 1\left(I_{g,t} > 0\right)$ |
| (NLL.2) | $z_{g,t} = \begin{cases} \log \frac{I_{g,t}}{B_{g,t}} & g = 5 \\ \log \frac{I_{g,t}}{q_{g,t} B_{g,t}} & g = 1,4 \end{cases}$ |
| (NLL.3) | $\hat{q}_g = \frac{1}{n_g} z_{g,t}, \quad g = 4$ |

| No. | Equation |
|---|---|
| (NLL.4) | $\hat{\tau}_g^2 = \begin{cases} \frac{1}{n_g}\sum_t 1\left(I_{g,t} > 0\right) \cdot \left(z_{g,t} - \hat{q}_g\right)^2 & g = 5 \\ \frac{1}{n_g}\sum_t 1\left(I_{g,t} > 0\right) \cdot \left(z_{g,t}\right)^2 & g = 1{,}4 \end{cases}$ |
| (NLL.5) | $l_{g,1} = \frac{1}{2}\left(n_g \log\hat{\tau}^2 + n_g\right)$ |
| (LL.1) | $n_{x,g,t}^{len} = \sum_a 1\left(u_{l,x,g,t} > 0\right)$ |
| (LL.2) | $\eta_{l,x,g,t} = \log u_{l,x,g,t} - \log\hat{u}_{l,x,g,t}$ |
| (LL.3) | $Z_{x,g} = \sum_t \sum_a \left(\eta_{l,x,g,t} - \frac{1}{n_{x,g,t}^{len}} \sum_{l\prime} \eta_{l\prime,x,g,t}\right)$ |
| (LL.4) | $\hat{\tau}_{len,g,x}^2 = \frac{1}{\sum_t n_{x,g,t}^{len}} Z_{x,g}$ |
| (LL.5) | $l_{g,2} = \sum_x \left(\frac{1}{2}\sum_t n_{x,g,t}^{len} \cdot \log\hat{\tau}_{age,g,x}^2\right)$ |
| (Pr.1) | $p_h = -[(\beta_1 - 1)\log h + (\beta_2 - 1)\log(1-h)]$ |
| (Pr.2) | $p_M = \frac{M_m - \mu_m}{2\sigma_M^2} + \frac{M_f - \mu_f}{2\sigma_M^2}$ |
| (Pr.3) | $p_s = \sum_g \left(\frac{\alpha_g - \mu_{\alpha_g}}{2\sigma_{sel,g}^2} + \frac{\beta_g - \mu_{\beta,g}}{2\sigma_{sel,g}^2}\right)$ |
| (Pr.4) | $p_R = \sum_{t=2}^{T} \omega_t^2$ |
| (OF.1) | $f = \sum_g \left(w_g^{idx} l_{g,1} + w_g^{len} l_{g,2}\right) + p_h + p_M + p_s + p_R + w_{B_0} \log B_0$ |

*Table 4. GH-0+1 commercial fleet code definitions (Fleet), which indicate Nation and Gear, as well as average catch between 1968 and 2020 (Mean Historical Catch), and allocation based on average catch share during 2011 – 2020 (10-yr Allocation).*

| Fleet | Nation | Gear | Mean Historical Catch | 10-yr Allocation |
|---|---|---|---|---|
| CAN_BTM | Canada | Bottom Trawl | 3.285 | 0.3480722 |
| CAN_LL | Canada | Longline | 0.166 | 0.0059228 |
| GR_BTM | Greenland | Bottom Trawl | 3.221 | 0.3666405 |
| CAGR_GN | Canada and Greenland | Gillnet | 1.759 | 0.2094441 |
| OTH_BTM | Russian, Norway, Japan, Germany, Spain, Faroe Islands, Estonia, Latvia | Bottom Trawl | 1.564 | 0.0631904 |
| GRNF_LL | Greenland, Russia, Norway, Faroe Islands | Longline | 0.148 | 0.0067300 |

## SISCAL-GH model state dynamics

The model parameters, equilibrium states, and full state dynamics for SISCAL-GH are given in Table 2. Unfished equilibrium recruitment (EQ.3) and numbers-at-age (EQ.4) are derived via spawning biomass per recruit (EQ.2), which is itself a function of female natural mortality (M.2), weight-at-age, and unfished equilibrium survivorship-at-age (EQ.1), as well as female maturity-at-age integrated over uncertainty in the length-at-age distribution (G.1 – G.6, Figure 2). The Von Bertalanffy length-at-age model was estimated from 315 fish sampled during the 2017 offshore survey cruise in Division 0A, while the allometric weight-at-length model was estimated from observations of 9057 fish, sampled in the same area between 2001 and 2019. Values for length-at-50% and -95% maturity were taken from studies of Division 0A fish (Morgan and Treble 2006, Harris et al. 2009), and converted to maturity-at-age by integrating over the length-at-age distribution.

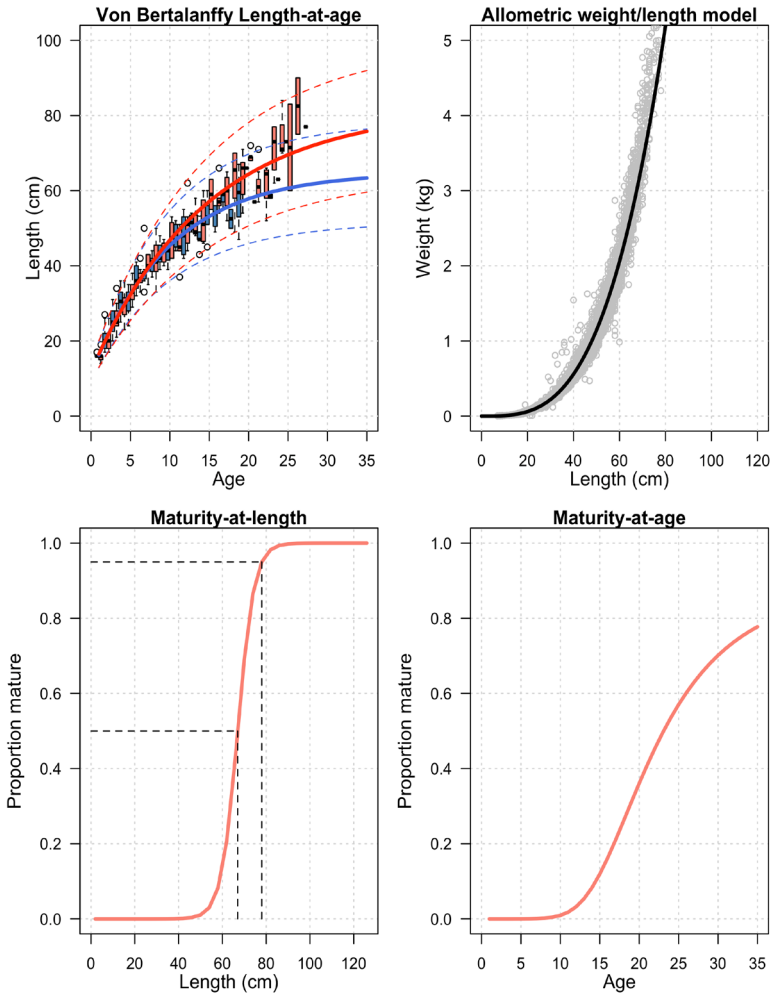


*Figure 2. GH-0+1 growth (top) and maturity (bottom row) models. Von Bertalanffy length-at-age models (top left panel) are fit to data for males (blue) and females (red) sampled by the 2017 offshore research vessel survey cruise in Division 0A, while the allometric length-weight conversion is fit to length/weight data collected between 2001 and 2019 by the same survey and aggregated over sexes (grey points). Maturity-at-length is an input to the model, based on length-at-50% and length-at-95% values taken from studies of Greenland Halibut in NAFO Division 0A, while maturity-at-age is derived by integrating the maturity-at-length over the female length-at-age distribution.*

A model year in SISCAL-GH begins on January 1 and ends on December 31. Age- and sex-class abundances are initialised in an unfished state in 1968, which seems reasonable given the trace levels of catch up to that time. Spawning biomass in 1968 is generated by summing weight-at-age over an unfished equilibrium abundance-at-age structure (A.1). Annual recruitment (occurring on the first day of the year) is assumed to follow a Beverton-Holt stock-recruitment function parameterized via stock-recruitment steepness $h$, unfished spawning stock biomass $B_0$, and recruitment process error deviations $\omega_t$ (A.2). The female spawning stock biomass is calculated from numbers-at-age based on the maturity-at-age ogive and mean weight-at-age (Figure 2). The population dynamics progressed from year-to-year by advancing each age-class after applying mortality due to natural causes $M_{x,t}$ and commercial fishing (explained below) and adding age-1 recruitment (A.3). In cases where natural mortality is modeled as time-varying, the $M_{x,t}$ series proceeds as a simple random walk with annual deviations $\epsilon_t$ (M.1).

Selectivity-at-length was asymptotic (logistic; Table 2, S.1) with length-at-age for all fleets except the Canadian Bottom Trawl fleet, and both the offshore and inshore research vessel (RV) surveys, in Divisions 0A1CD and 1A-F, respectively. Those three fleets had dome shaped (double logistic) selectivity, which was taken as a product of an ascending limb $s^A_{l,g}$ and descending limb $s^D_{l,g}$ (S.2). All fleets use the same parameters for the ascending limb: length-at-50% selectivity $L^A_{50,g}$ and length-at-95% selectivity $L^A_{95,g}$. For the fleets with dome-shaped selectivity, there are additional descending limb length-at-95% selectivity $L^D_{95,g}$ and length-at-50% $L^D_{50,g}$ selectivity parameters. Apart from the ascending limb $L^A_{50,g}$ parameter, all other selectivity model parameters are modeled as increases from the previous parameter on the left (i.e., $L^D_{50} > L^D_{95} > L^A_{95}$), thereby ensuring the descending limb is always to the right of the ascending limb, which reduces the sharpness of the selectivity peak. Selectivity-at-length was converted to sex-specific selectivity-at-age via the sexually dimorphic von Bertalanffy growth model (S.3, Figure 2). Sensitivity to the assumption of dome shaped or asymptotic selectivity for trawl fleets is considered in Appendix A.

Removals by all commercial fisheries in both Subareas 0 and 1 were represented as discrete fisheries occurring around halfway through the year at a fractional time step $\delta_g$ (Table 2, C.1 - C.9), where $0.45 < \delta_g < 0.55$. Fish were removed from the population by estimating the catch-at-age in biomass for each fishery (C.6), which was then converted to total caught numbers-at-age via mean weight-at-age (C.7), and then subtracted from population numbers-at-age (C.8). Annual exploitation rates for total landings $U_t$ and each commercial fleet $U_{g,t}$ are calculated as the ratio of landed catch to an allocation-weighted exploitable biomass, which weights the selected biomass of the commercial fleets by the average proportion of landed catch from each over the 2011-2020 period (Table 2, C.5, C.9; Table 4 for allocation).

## SISCAL-GH Observation Models

Temporal variation in GH-0+1 abundance and population composition are monitored via an offshore research vessel trawl survey of the southern part of Division 0A and Division 1CD (RV_0A1CD), and an inshore small mesh survey in Divisions 1A to F (RV_SFW1AF). Length compositions are also collected by both surveys and all commercial fishing fleets to varying degrees. Additionally, there is a bottom trawl catch per unit effort (CPUE) series that combines catch and effort data across several nations via a catch-weighted average of selected biomass for each trawl fleet (CAN, GRL, and OTH, using mean catches in Table 4).

Survey indices and CPUE are assumed to be linear (i.e., no hyperstability or hyperdepletion) in the population state that they are indexing, which is total survey biomass for the offshore RV

survey, total survey abundance for the inshore RV survey, and weighted average vulnerable biomass for the combined trawl fishery CPUE (Table 3, O.1; Table 4 for average catch weights). Catchability parameters for the surveys, and model residual standard deviations for all three indices, are estimated via conditional maximum likelihood (NLL.1 – NLL.5). Given the possible spatial expansion of fisheries comprising the trawl CPUE time series, we estimated a time-varying catchability parameter for that series beginning in 2014, which is meant to track possible catchability changes as the fishery expanded (O.1).

Proportion-at-length (i.e., length composition) observations are modeled in 4 cm length bins via a logistic-normal likelihood function (Schnute and Haigh 2007, Francis 2014), with expected values equal to the proportions of the vulnerable numbers-at-length (Table 3, O.3). We used a tail-compression procedure that combined data from length bins with less than 0.1% of the samples with neighbouring length bins that were above that threshold. This procedure eliminates zeroes or very small proportions in the length composition data, but also creates a variable number of bins at each time step (LL.1). Fleet and sex specific length sampling error variances were estimated via conditional maximum likelihood as nuisance parameters (LL.4).

## Objective function and optimization

The SISCAL-GH objective function (Table 3, OF.1) is proportional to the negative log Bayes joint posterior density function. The negative log posterior is defined as the sum of the negative log likelihood functions for observed data (NLL.5, L.8 and F.4), negative log prior densities for recruitment (Pr.1) and natural mortality (Pr.2) process errors, and other priors on leading parameters (Pr.3 - Pr.4).

The SISCAL-GH objective function is implemented in Template Model Builder (TMB), and optimised via the nlminb() function in the R statistical package (Kristensen et al. 2016, R Core Team 2015). Model parameters are considered converged when the maximum gradient component of the log-posterior surface had absolute value less then $10^{-2}$ and the model Hessian matrix (i.e., the inverse of the covariance matrix of leading parameters) was positive definite.

Bayes posterior distributions are generated via four independent chains of 1000 samples each using Hamiltonian Monte-Carlo (Monnahan and Kristensen 2018). Chain starting values were overdispersed by sampling from a normal distribution with mean values at the maximum posterior density estimates (MPDEs) and standard deviations equal to three times the standard error of each parameter. Hamiltonian Monte-Carlo differs from Markov-Chain Monte-Carlo by minimising the auto-correlation between successive posterior samples, thereby producing a mixed model posterior sample with lower absolute sample sizes and little or no thinning (Monnahan et al. 2017).

## Goodness of fit

Goodness of fit is reflected in two ways. First, residual standard errors and likelihood function values are shown for all datasets. Second, the posterior chain characteristics are examined, including within-chain autocorrelation, correlation among leading parameters, potential scale-reduction factors (i.e., $\hat{R}$ values less than 1.01), and effective sample size (ESS; at least 100 per chain).

## MSY based reference points

Biological maximum sustainable yield (MSY-based) reference points were estimated from SISCAL-GH estimates of GH-0+1 unfished biomass and productivity parameters, fleet selectivity, and the allocation of catch among fleets based on the catches from 2011 - 2020

(Table 4). The resulting reference points provide estimates of optimal spawning biomass $B_{MSY}$, exploitable biomass $B_{MSY}^{exp}$, yield $MSY$, and optimal exploitation rate $U_{MSY} = MSY/B_{MSY}^{exp}$. While an optimal fishing mortality rate ($F_{MSY}$) is estimated as part of the reference point derivation, it is a grid search parameter used to scale fleet specific fishing mortality rates based on the allocation, from which equilibrium yield is calculated. Therefore, the $F_{MSY}$ value is not directly interpretable as a total fishing mortality rate, given the range of fleet selectivities. Instead, the optimal exploitation rate $U_{MSY}$ is used to define MPs and performance metrics in the closed loop simulation procedure used in Part C.

## Sensitivity analysis

This section describes sensitivity tests of several model assumptions, including the asymptotic length, time-varying versus constant parameters for aggregated trawl CPUE catchability and natural mortality, and the minimum sample size and tail compression settings for length composition likelihood calculations. Additional tests of SISCAL-GH performance relative to changes in data sets and structural assumptions were conducted as part of the peer review process, which are summarised in Appendix A.

### Asymptotic Length

Preliminary SISCAL-GH fits suggested that the growth model may be estimating a negatively biased mean length-at-age for the GH-0+1 stock. Evidence for this bias is based on the observed maturity-at-length proportions for areas 0A and 0B (Cooper et al. 2007, Harris et al. 2009) showing that fully mature fish are in the upper tails of estimated length-at-age distribution for age 35+ fish, which results in less than 100% maturity-at-age in the age 35+ group (Figure 2). Additionally, there was trouble fitting to the gillnet fishery length compositions, which had a modal length in the upper tails of the length-at-age-35. We tested sensitivity to $L_{\infty,x}$ for both sexes by re-estimating the von Bertalanffy growth model and fixing $L_{\infty,x} = \alpha \cdot \hat{L}_{\infty,x}$, where $\hat{L}_{\infty,x}$ was the original value freely estimated from the 2017 offshore survey age/length data. The remaining parameters for growth rate $K_x$, length-at-age coefficient of variation (CV) $c_{L,x}$, and theoretical age-at-length-0 $t_{0,x}$ were re-estimated (Figure 3) and used as inputs when re-fitting SISCAL-GH to the full GH-0+1 assessment data.

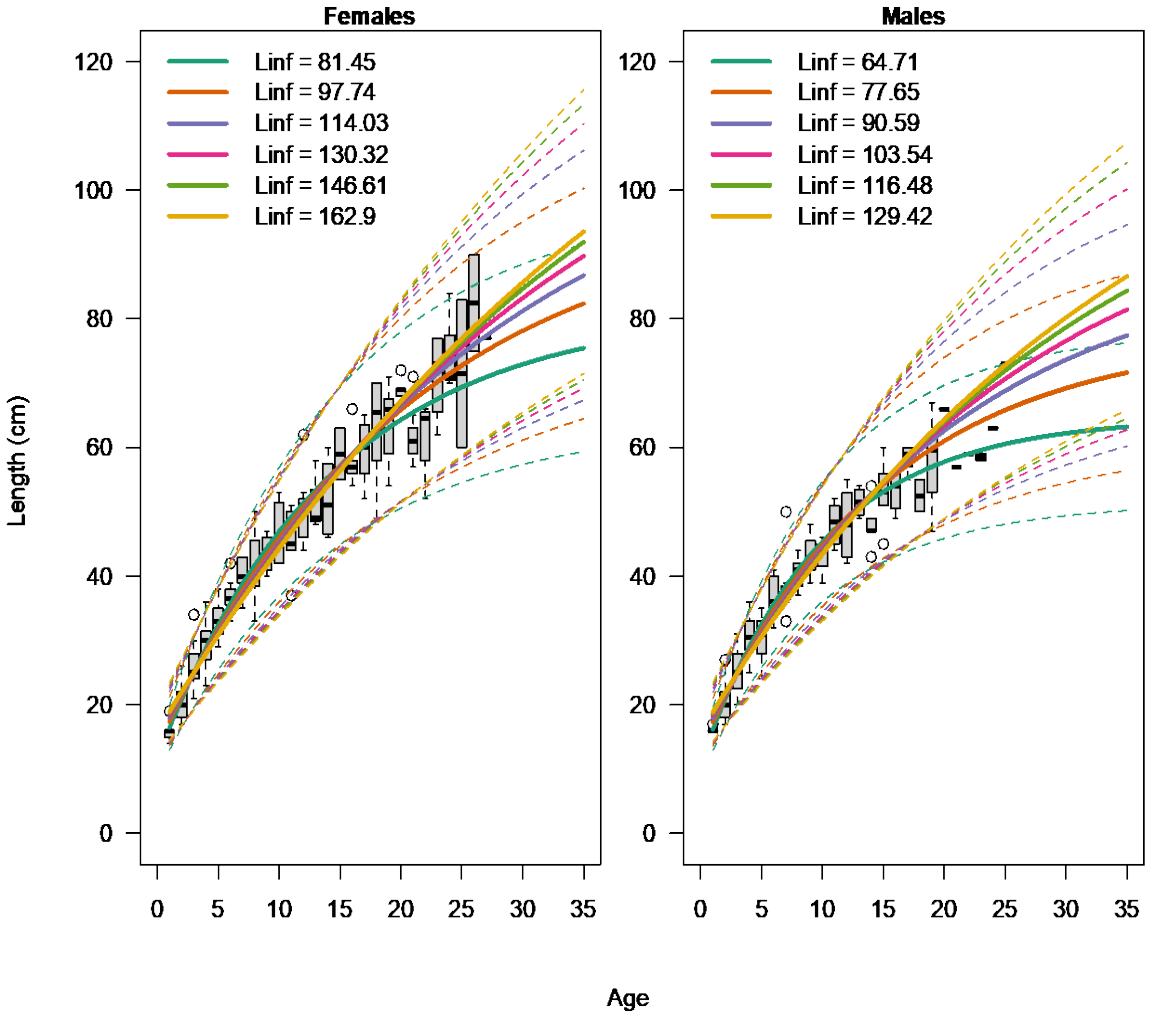


*Figure 3. Refit growth models used for the $L_\infty$ sensitivity analysis. Solid lines show the maximum likelihood estimate of the mean length-at-age, while dashed lines show the central 95% of the length-at-age distribution for each asymptotic length (colours in legend).*

### Time varying parameters

Sensitivity was tested to both time-varying or constant catchability and time-varying natural mortality as applied in the base SISCAL-GH model hypothesis. Time-varying catchability was applied post 2014 to improve the fit to recent high CPUE data points possibly associated with increasing spatio-temporal range of the fishery. The time-varying mortality was applied to match a depletion signal in CPUE data from 1989 - 2000, which may be associated with mortality that is otherwise unaccounted for in SISCAL-GH (e.g., possible bycatch from cod fishing effort, unreported catch by the cod fleet following the northern cod collapse, or fish landed in other areas). All four combinations of both parameters varying, one parameter varying, or both constant, are tested.

### Length sample size and tail compression

We also tested sensitivity to length-composition data treatments for likelihood function calculations. Both the minimum sample size for inclusion of length composition data in the data set and the minimum proportion used for tail compression are adjusted to demonstrate model

sensitivity, as well as changes in the residual standard error for some fleets across those settings.

### Retrospective analyses

A retrospective analysis was performed by fitting to successive 'peels' of data going from 2010 to 2020. Each peel is compared via MPDEs of spawning biomass time series, unfished biomass, unfished recruitment, natural mortality, and stock-recruit steepness. Additionally, MPDEs of MSY-based reference points are calculated for each data peel to understand the effect of new data on model equilibria and the associated management targets.

### Simulation self-testing

The ability of SISCAL-GH to estimate key model parameters was tested via a simulation self-test in which the 'true' operating model was based on MPDEs of SISCAL-GH parameters ($\Theta^*$). Parameters were estimated under the base hypothesis with time-varying selectivity for the inshore survey, time-varying catchability for trawl CPUE, and time-varying natural mortality. All recruitment and natural mortality process errors were held fixed at their MPDEs, but new observation errors (i.e., data) were simulated for all data series.

New data were simulated in three scenarios, where simulated data replaced the observed data originally used to estimate the base model parameters. The three scenarios were:

1. **SimLengths**: simulated length composition data and true (original) observed biomass/abundance index data;
2. **SimIdx**: true observed length composition data and simulated biomass/abundance index data;
3. **SimAll**: All length and index data simulated.

Under all scenarios, simulated data were generated using the standard errors estimated for the base SISCAL-GH model for each data series.

SISCAL-GH self-test performance was measured via relative errors of selected parameters summarised as median relative errors (MREs) and median absolute relative errors (MAREs) of key parameters. The MRE measures the bias of a model estimate, and the MARE can be interpreted as a measure of precision relative to the MRE. Both $MRE(\Theta)$ and $MARE(\Theta)$ summarise the distribution of relative errors $RE_i(\Theta)$, which is defined as:

$$RE_i(\Theta) = \frac{\hat{\Theta}_i - \Theta^*}{\Theta^*}$$

where $i$ indexes the simulation replicate, $\hat{\Theta}_i$ is the estimate of SISCAL-GH model parameters from simulated data in replicate $i$, and $\Theta^*$ is the MPDE of $\Theta$ from the true data. For both metrics, values closer to zero indicate better model estimates.

An additional simulation test was conducted during peer review to test the sensitivity of SISCAL-GH estimates to the assumption of deterministic recruitment from 1968 to 1988 (Appendix B).

## CLOSED-LOOP FEEDBACK SIMULATION FRAMEWORK

We conditioned an operating model with SISCAL-GH model estimates, using the same dynamics in Table 2. The model, which we refer to as MS3-GH, simulates the GH-0+1 fishery management system forward in time from 2021, generating stock responses to fishery removals of TACs, which are set based on simulated MPs. Simulated dynamics are then used to calculate

fishery performance metrics based on (hypothetical) fishery objectives derived from current NAFO precautionary approach policy (PA) (Brodie et al. 2013).

The MS3-GH operating model provides a practical and realistic representation of GH-0+1 stock dynamics, fishery harvesting processes, and fishery monitoring data so that non-linear feedbacks and data uncertainties can be accounted for in annual TAC advice. These processes interact to determine short- and long-term performance of fishery harvest strategies with respect to (hypothetical) fishery objectives. In the absence of fishery objectives, we have defined standard conservation and yield metrics based on the NAFO PA (Brodie et al. 2013). Performance metrics then provide a quantitative approach to assessing fishery compliance with international fishery policy.

The closed loop simulations include an additional RV survey to reflect the new survey design currently being initialised for the GH-0+1 stock, which includes the previous offshore RV survey design, plus some additional inshore sets in Division 0A and an expansion to include sets in Division 0B. The new survey design was simulated in projections as two separate indices. The first continues the existing RV survey, with the same selectivity and catchability. The second is a simulated **new RV** survey as a 9th fleet ($g = 9$), which has similar selectivity to the existing inshore survey in Subarea 1 and a simulated catchability of $q_9 = 1$. While the simulated selectivity and catchability values are conjecture at this time given a lack of any existing data for this survey, there is some basis for the choice of values. Selectivity is similar to the current inshore survey to simulate the capture of small fish in shallower waters, as there was no catch-at-length data available at the time of writing and additional sets are planned in shallower waters. Catchability was set to 1 (i.e., an absolute survey) as without data this is an arbitrary choice, which should not affect the model performance in closed loop simulations, as long as simulated assessments estimate $q_9$. The precision of the newRV survey is the basis of two alternative operating model hypotheses for the closed loop simulations, explained below.

## Adaptive model-/index-based management procedure

The GH-0+1 adaptive model-/index-based MP attempts to capture the main elements of a full fishery stock assessment and management framework cycle. In management frameworks for similar species (e.g., Atlantic Halibut) a full age-structured stock assessment is performed every 5 - 8 years, while in the intervening years a simpler index-based procedure is used to set TACs on an annual basis. Control parameters (e.g., catchability and target exploitation rate) for the index-based procedure are re-estimated at full stock assessment updates. Updated parameters are then usually re-tested in closed-loop simulations to ensure expected procedure performance is commensurate with fishery objectives under contemporary estimates of stock-status, biomass, and productivity, but we skipped this step for computational reasons.

For the adaptive model-/index-based MP, the full stock assessment model update is scheduled every six years and TAC revisions happen every two years with both beginning in 2024. We chose 2024 because it was the first year where the simulated new RV survey would have enough data for conditional estimates of both catchability and residual standard errors, and will also be the next time that a TAC update is scheduled for the real GH-0+1 stock.

In full assessment years, only the index-based stock assessment parameters are updated (i.e., target harvest rates and MSY-based reference points remain constant at their current estimates). At each full assessment, we update the offshore RV survey catchability $q_{RV,t}$ estimate, as well as an exploitable biomass adjustment factor

$$\rho_t = e^{\left(\frac{1}{6}\sum_{t\prime=t-5}^{t}\log\left(B_{t\prime}^{exp}/B_{t\prime}^{RV}\right)\right)},$$

where $B_t^{exp}$ and $B_t^{RV}$ are the estimates of exploitable and RV survey biomass over the most recent 6 years. The biomass adjustment factor attempts to track differences between the exploitable and survey biomasses resulting from variation in recruitment and the underlying age-structure. This approach is similar to the JABBA select model (Winker et al. 2020), which scales spawning biomass to exploitable biomass using model equilibria. The main difference here is that we used an age-structured model to estimate both RV and exploitable biomass, which were compared directly and updated over time.

The index-based component of the MP sets TACs according to the following steps at each time step:

1. Compute the 3-year average of the offshore **RV_0A1CD** survey index

$$\bar{I}_t = e^{\left(\frac{1}{3}\sum_{t'=t-2}^{t} \log I_{RV,t'}\right)};$$

2. Compute the estimated exploitable biomass from the 3-year average RV survey index, the latest catchability estimate, and the biomass adjustment factor via

$$\hat{B}_t^{exp} = \rho_t \cdot \frac{\bar{I}_t}{\hat{q}_{RV,t}},$$

where $\hat{q}_{RV,t}$ is the conditional MPDE of the RV_0A1CD survey index catchability parameter estimated in the most recent assessment, and $\rho_t$ is the exploitable biomass adjustment factor. If $\rho_t$ and $q_{RV,t}$ are not estimated for time $t$, then use the values from time $t-1$.

3. Calculate the target fishing mortality rate $F_t$ from the harvest control rule

$$F_t = \begin{cases} 0 & \hat{B}_t^{exp} \le 0.3B_{MSY}^{exp} \\ F_{buf} \cdot \dfrac{\hat{B}_t^{exp} - 0.3\hat{B}_{MSY}^{exp}}{0.7\hat{B}_{MSY}^{exp} - 0.3\hat{B}_{MSY}^{exp}} & 0.3\hat{B}_{MSY}^{exp} < \hat{B}_t^{exp} \le 0.7\hat{B}_{MSY}^{exp}, \\ F_{buf} & 0.7\hat{B}_{MSY}^{exp} < \hat{B}_t^{exp} \end{cases}$$

where the maximum harvest rate is $F_{buf} = 0.8F_{MSY}^{exp} = 0.099$, the upper control point is $0.7B_{MSY}^{exp} = B_{buf}^{exp} = 95.6$ kt, and the lower control point, where harvest rates drop to 0, is $0.3B_{MSY}^{exp} = 41$ kt in this example. The biomass $B_{MSY}^{exp}$ is the equilibrium exploitable biomass producing $MSY$ (see results for values used in the simulations). Note that this rule uses a simple hockey-stick functional form common to Canadian fisheries policy (DFO 2006), which is effectively 'mapped onto' the NAFO PA policy by the choice of control/reference points (Figure 4; Brodie et al. 2013). All control/reference point values such as $F_{buf}$, $F_{lim}$, and $B_{buf}$ may be adjusted based on fishery manager preferences, and as tuning parameters, to bring fishery performance metrics in line with NAFO PA framework requirements under projected operating model conditions (Brodie et al. 2013).

4. Compute the proposed $TAC'_t = (1 - e^{-F_t})\hat{B}_t^{exp}$;

5. Apply any interannual TAC change limit, here set as $\Delta_t = 1.0$, i.e.,

   a. If $TAC'_t > (1 + \Delta_t)TAC_{t-1}$, then $TAC_t = (1 + \Delta_t)TAC_{t-1}$;

   b. Otherwise, $TAC_t = TAC'_t$

   We chose $\Delta_t = 1.0$ so that TACs could be dropped to 0 if exploitable biomass was estimated below $0.3B_{MSY}^{exp}$, however this change limit can be adjusted based on manager/stakeholder preferences to reduce interannual TAC variation.

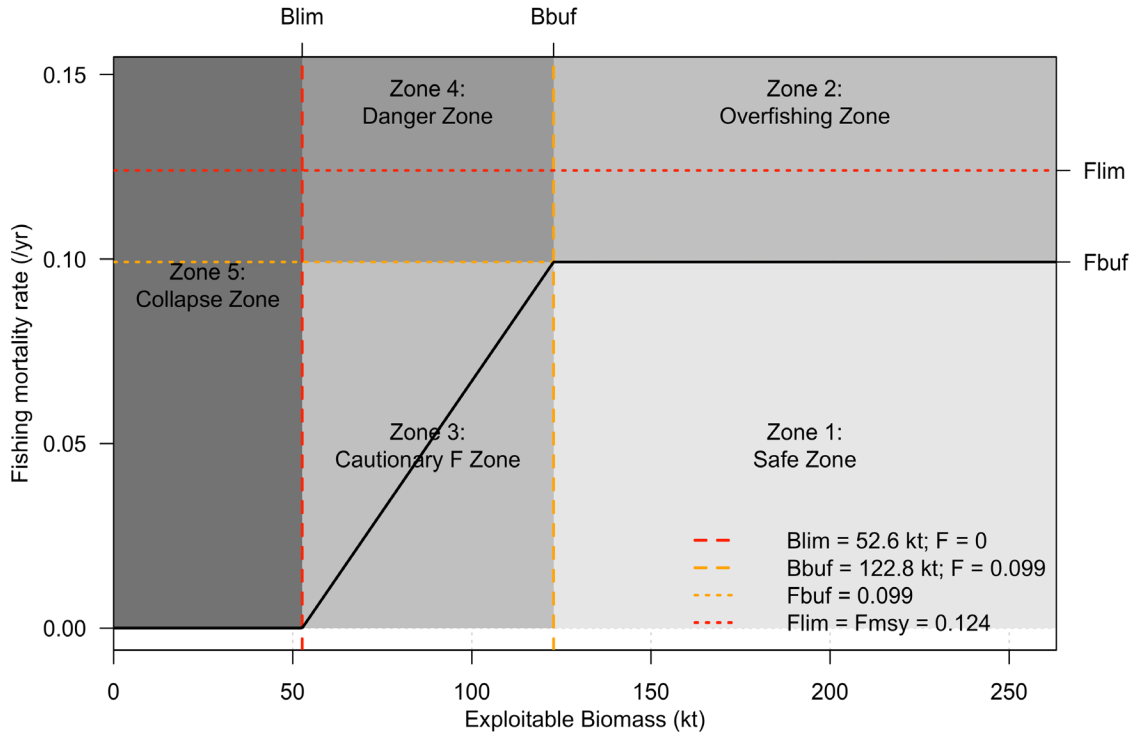


*Figure 4. A 'hockey-stick' harvest control rule for GH-0+1, setting target fishing mortality rates (y-axis) as a function of exploitable biomass (x-axis) as estimated by the index-based management procedure (both adaptive and fixed $q$ versions). The hockey-stock rule is overlaid on the five zones defined by the NAFO precautionary approach (PA) policy to show how Canadian precautionary fishery management policy effectively maps on to the NAFO PA policy. For this example, $F_{MSY} = F_{lim}$ is set as a limit mortality rate, and $F_{buf} = 0.8F_{MSY}$ as the target fishing mortality rate.*

The final $TAC_t$ for each year is then allocated among the 6 commercial fleets according to allocated proportions (Table 4). Catches are removed from the population as in the SISCAL-GH model, where fleets are assumed to take catch in one discrete pulse in the middle of the fishing year, before and after which half of the natural mortality is applied. For years where the MP is not formally completed (e.g., under a biennial TAC schedule as in the GH-0+1 fishery) the procedure carries forward the previous year's TAC.

The adaptive approach, which we called **qAdapt**, used here to update the RV survey $q_{RV}$ parameter and biomass adjustment factor $\rho$ at longer time intervals than the TAC, serves two purposes. First, it simulates future uncertainty in the $q_{RV}$ parameter rather than unrealistically assuming a perfectly known value that is close to an optimal choice. Secondly, while the offshore RV survey has the highest spatial overlap with the exploited portion of the stock, the RV survey selects for a higher proportion of small fish than any fishery, meaning that as natural and fishing mortality alter the age structure of the stock, the survey biomass and exploitable biomass will diverge. Therefore, the updated $q_{RV,t}$ and $\rho_t$ values are required to bring the TAC decision closer to the target fishing mortality rate, as the proportional difference between exploitable and offshore survey biomass will vary over time and among posterior draws used to condition simulation replicates.

For comparison purposes, we also evaluated a non-adaptive fixed parameter procedure, called **qFixed**, with no simulated model updates. Under the qFixed procedure, we fixed $\rho_t = 0.82$ and

$q_{RV,t} = 0.38$ for all time steps $t$ at their Bayes posterior medians estimated by the SISCAL-GH model fit to data up to 2020.

## Closed-loop feedback simulation algorithm for evaluating harvest strategies

We use the following closed-loop simulation algorithm to apply the simulated model-/index-based MP (Walters 1986, de la Mare 1998, Cooke 1999, Punt and Smith 1999, Sainsbury et al. 2000, Butterworth 2007):

1. Initialize a pre-conditioned operating model for the historical period (1968 – 2020) based on the SISCAL-GH model;
2. Project the operating model population and fishery one time step into the future for 47 years (corresponding to two GH-0+1 generations), starting in 2021. At each time step apply the following:
    i. Generate the catch and survey data available for stock assessment;
    ii. If a full stock assessment is scheduled, estimate SISCAL-GH states and parameters from the simulated data, and generate new $q_{RV,t}$ and $\rho_t$ values for the index-based procedure;
    iii. If a TAC update is scheduled, apply the harvest control rule to generate a new TAC, otherwise bring the previous time step's TAC forward;
    iv. Update the MS3-GH operating model population dynamics given the total natural mortality, removals due to fishing allocated among fleets, and new recruitment;
    v. Repeat steps 2.i - 2.iv until the projection period ends.
3. Repeat step 2 for 100 replicates
4. Calculate quantitative performance statistics across all 100 replicates;

In the real GH-0+1 management system, the current and new survey index and catch-at-length data will not be collected before the 2022 summer season, therefore, our simulations held the TAC at the 2020 catch of 32 kt until 2025, following the first simulated full stock assessment in 2024, after which TACs will be updated on a biennial cycle and stock assessments ($q_{RV,t}$ and $\rho_t$ values) will be updated on a 6-year cycle. While the true TACs may in fact be lower (but probably not higher), the performance under the assumption of a constant 32 kt TAC until 2024 can be considered a conservative under-estimate of true fishery performance, as biomass would be higher with lower TACs in the short term.

The adaptive model-/index-based procedure is evaluated under two operating model scenarios, which differ by precision used for simulated newRV survey data:

1. **isPrec**: precision is taken from the existing inshore RV survey;
2. **osPrec**: precision is taken from the existing offshore RV survey.

The qFixed MP was evaluated only under one scenario, as it did not rely on the newRV survey data at all so results are identical for both.

## Performance Metrics

We calculated the following example performance metrics for evaluating the adaptive GH-0+1 MP in closed loop simulation. These are chosen to exhibit a balance of conservation and yield metrics in line with NAFO PA policy.

1. Probability that female spawning biomass is below the limit reference point (LRP) $B_{lim} = 0.3B_{MSY}$ (pLRP),
2. Probability that the exploitation rate exceeds $U_{lim} = U_{MSY}$ (pOverfish)
3. Mean exploitation rate above $U_{MSY}$ (mUoverfish)
4. Median average catch over 10 years (avgC)
5. Average annual variation in yield over 10 years (AAV)

The limit reference point $B_{lim}$ is defined as 30% of the female spawning biomass that produces MSY, and the limit exploitation rate is the optimal exploitation rate $U_{MSY}$ (a proxy for $F_{lim}$ in NAFO PA policy) (Brodie et al. 2013).

For the adaptive model-/index-based procedure, we also calculated and compared relative errors in key model parameters for the simulated stock assessments in 2024 and 2054, similar to the simulation self-test procedure outlined earlier.

# RESULTS

## SISCAL-GH BASE MODEL ESTIMATES

### Fits to GH-0+1 data

The observation model residual standard deviations for each data set (Table 5) are good indicators of the quality of SISCAL-GH fits shown in Figures 5 and 6. In general, these standard errors are in log-space, so they can be interpreted more or less as coefficients of variation (CV). For example, the offshore RV survey had a residual standard error of $\tau_{RV0A1CD} = 0.15$ (i.e., 15% CV), indicating that the model and RV index have strong agreement not only on the trend, but also on the interannual variation (Figure 5, upper panel), despite the somewhat large residual in 2016. Occasional large deviations are not necessarily unexpected for a log-normal distribution, even one with CV ~ 10%.

*Table 5. Estimated standard errors for length composition data (first three columns), proportion female in each length bin (fourth column), and stock indices (last column) for Canadian trawl (CANBTM), Canadian longline (CANLL), Greenland trawl (GRBTM), Canadian and Greenland gillnet (CAGRGN), all other trawl (OTHBTM), Greenland/Russian/Norway/Faroe Isl. longline (GRNFLL), offshore survey (RV0A1CD) and inshore survey (RV_SFW1AF).*

| | Male | Female | Combined | Prop Female | Indices |
|---|---|---|---|---|---|
| $\tau_{CANBTM}$ | - | - | 0.431 | - | - |
| $\tau_{CANLL}$ | - | - | 0.762 | - | - |
| $\tau_{GRBTM}$ | - | - | 0.842 | - | - |
| $\tau_{CAGRGN}$ | - | - | 0.604 | - | - |
| $\tau_{OTHBTM}$ | - | - | 0.827 | - | - |
| $\tau_{GRNFLL}$ | - | - | 1.144 | - | - |
| $\tau_{RV\_0A1CD}$ | - | - | 0.387 | - | 0.15 |
| $\tau_{RV\_SFW1AF}$ | 0.71 | 0.565 | | 0.392 | 0.38 |

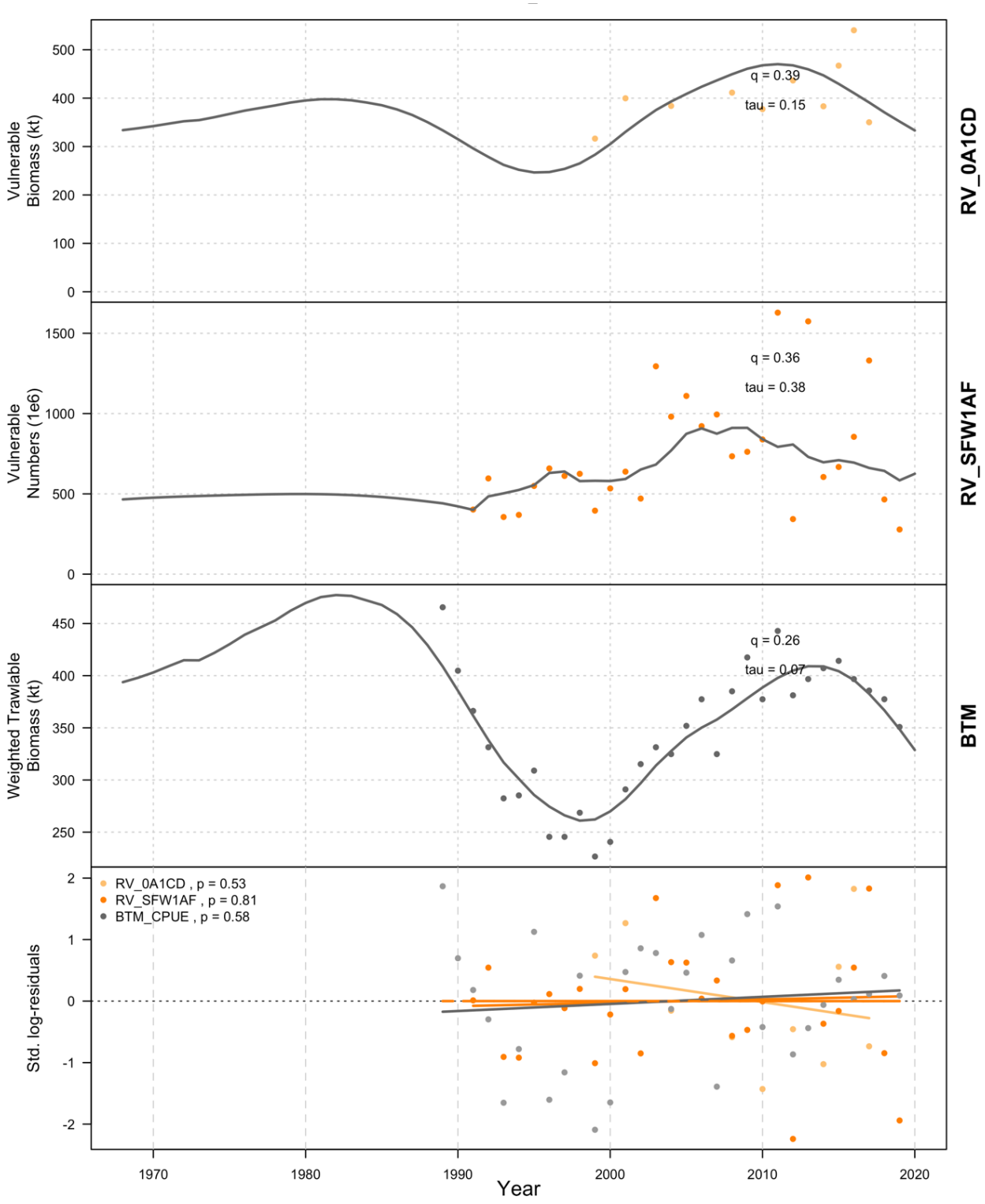


*Figure 5. SISCAL-GH fits (lines) to abundance and biomass indices (points), with standardised log-residuals for each index (bottom panel). Trawl CPUE indices are scaled by time-varying catchability after 2014, which results in a downward trend of the data to meet the stock.*

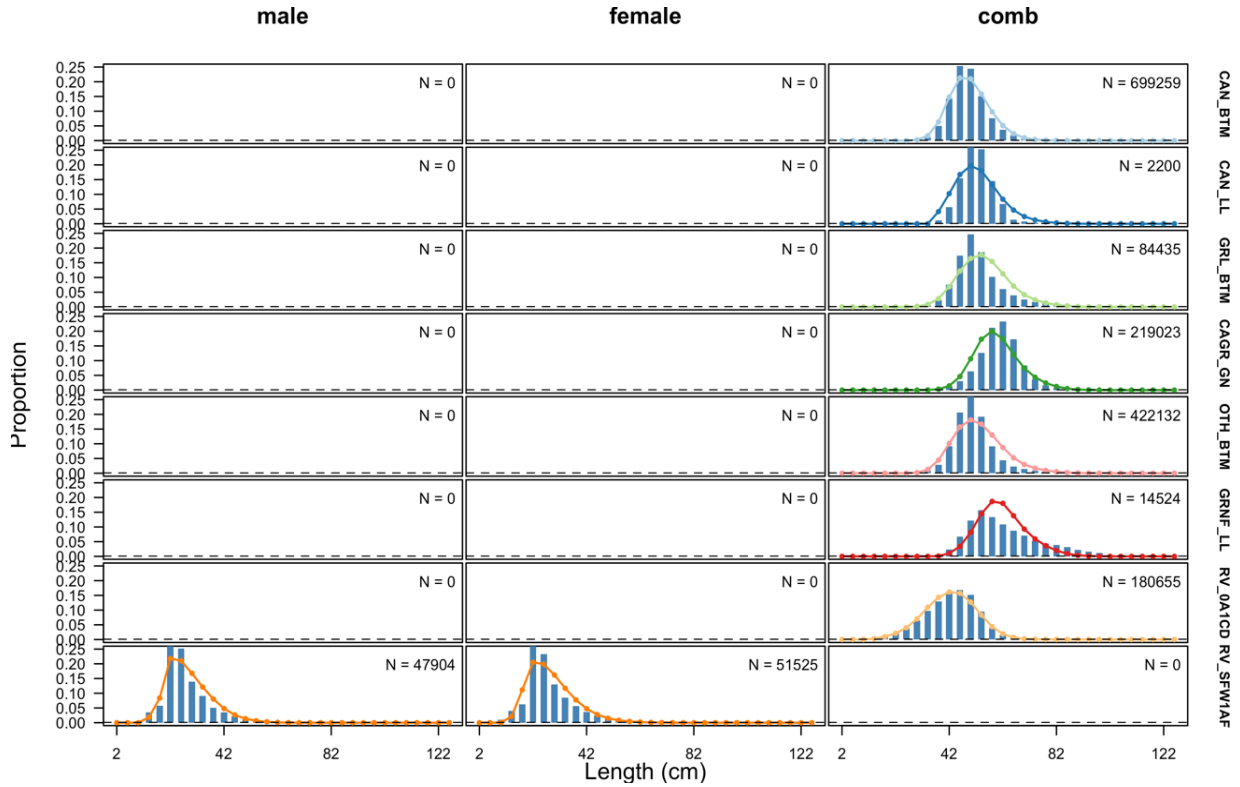


*Figure 6. Time-averaged SISCAL-GH fits (lines and points) to length composition data (blue bars) for males (left), females (middle), and combined sexes (right), sampled by each fishery and survey (labeled on the right). Total sample size N over all years is given in the top-right hand corner of each panel.*

The inshore RV survey had a higher residual standard error at $\tau_{RVSFW1AF} = 0.38$ indicating a marginally acceptable fit with several large residuals occurring after 2004 (Figure 5, second panel), which coincide with some variation in execution of this survey (Treble and Nogueira 2020). While comparative towing can align the mean catchability of two surveys, it cannot account for any process variance that stems from different gear or a variable number of sets. On the bright side, variation in survey designs/protocols are fixable in the future and could lead to lower residual standard deviations in future assessments.

The combined bottom trawl CPUE had a very low residual standard error of $\tau_{BTM} = 0.073$ (Figure 5, third panel), which is not totally indicative of quality since the model had extra flexibility in the form of time-varying mortality and time-varying catchability to help fit that series. In such cases, one needs to also consider the variances used in those random-walk components in assessing fit quality, which was $0.1$ on the log scale, or roughly a 10% CV. None of the biomass index series had a significant residual trend (Figure 5, bottom panel, $p > 0.05$ for all indices) or an indication of bias.

Time-averaged fits to length compositions are given in Figure 6, with annual fits provided in Appendix D. The time-averaged fits are adequate to show where the model consistently misses certain aspects of the data. In general, residual standard errors of 0.3 - 0.5 are considered very good for compositional data, 0.5 - 0.8 are good, 0.8 - 1.0 are usually adequate, and above 1.0 indicates that there may be model mis-specification that fails to sufficiently capture the data collection process. Here, the Canadian bottom trawl and longline fleets, mixed gillnet fleet, and two fishery independent surveys fall into the very good and good ranges, while the remaining commercial fleets fall into adequate (2 fleets) and likely mis-specified (1) categories (Table 5). Note, however, that the term ‘mis-specified’ here is used generally and may include not properly

capturing inadequate or biased sampling processes; that is, the model assumes ideal sampling designs, which is almost surely incorrect, especially for fishery-dependent data. Nevertheless, model fits and variance indicators are helpful in identifying where standardizing sampling designs and protocols could improve overall assessment performance. For example, the GR_BTM (Greenland) and the OTH_BTM (not Greenland or Canada) fleets both had reasonable fits to length compositions in many years (Figures D.3 and D.5); however, there were also several years where the expected composition was translated either to the right or left, leading to higher residuals despite matching the overall shape well. These residuals could be reduced by using a time-varying selectivity, which would allow some flexibility to shift the expected values to meet the observations, or by adjusting the tail compression settings (explored in sensitivity analysis below). But, more importantly, they point to possible issues in the underlying data that should be examined further and corrected, if possible.

The GRNF_LL (Greenland, Russia, Norway, and Faroe Islands) fleet had high proportions of length observations in the upper tail of the length-at-age distributions for older fish that were not present for any other fleets (Figure D.6). Assuming this is probably not a sampling bias or error (although it could be), there may be a size-at-age mismatch between the portion of the stock generating age/length samples for the growth model and the portion of the GH-0+1 stock fished by these fleets. A growth model misspecification means that adjustments to the selectivity function will not completely improve the fit to the length composition data and may just mask the problem. In sensitivity analyses below, we investigate the potential implications of this mismatch via alternative asymptotic length hypotheses.

The residual standard error for the proportion female data from the inshore RV survey (Table 5) is similar to length composition data for that survey. Most large residuals were observed in length bins towards either end of the range, which is expected since outer bins typically have smaller sample sizes and higher individual variances than those in the centre of the composition (Figure 7).

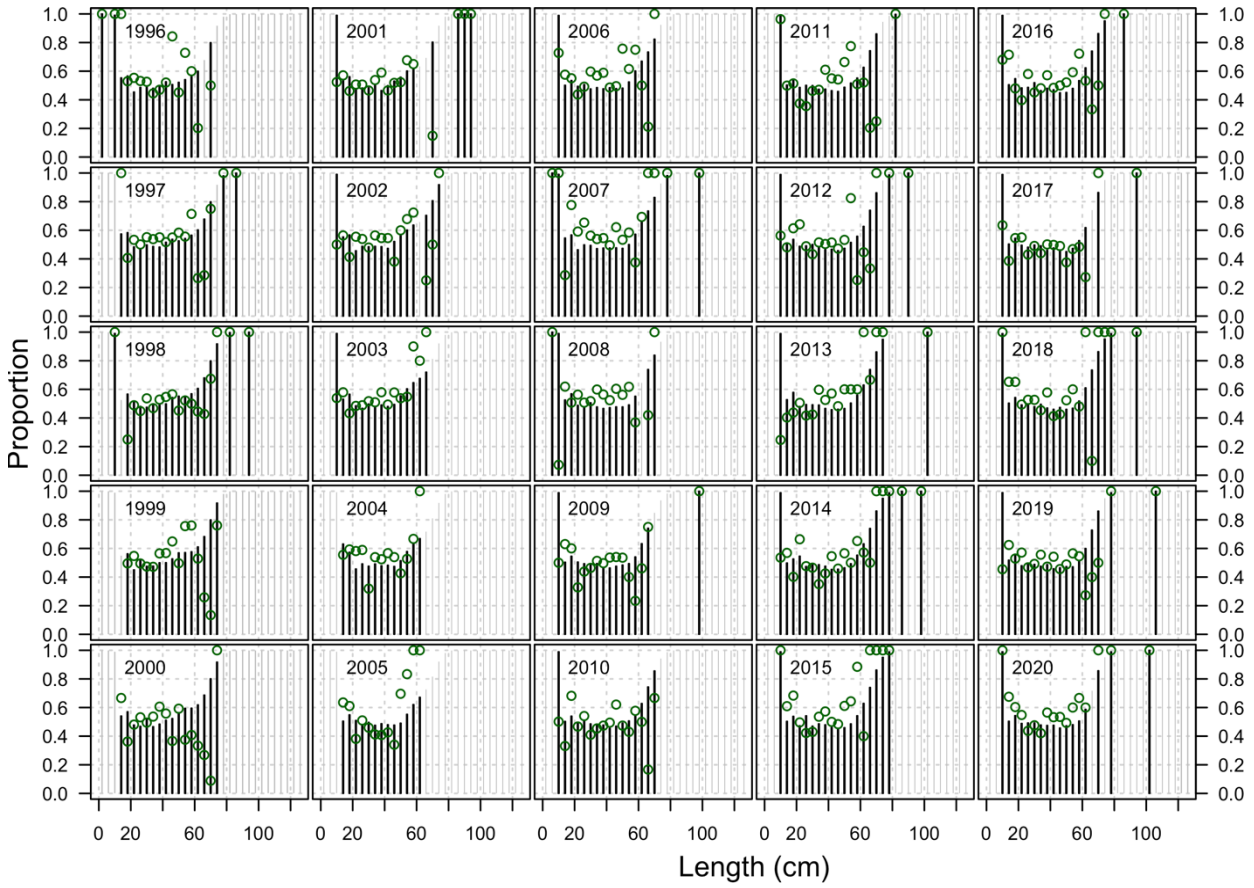


*Figure 7. SISCAL-GH model fits to proportion of females observed in inshore survey length compositions by length bin. Green circles show data, while vertical line segments show expected values from the model for bins with (black) and without (light grey) data.*

Standard Bayesian posterior chain diagnostics suggest that chains converged (Figure 8) with most leading parameters having potential scale reduction factor $\hat{R}$ values less than 1.01 (i.e., that posterior standard deviations within each chain are within 1% of each other across chains) (Betancourt and Girolami 2015). There were 10 marginal parameters where $1.02 \geq \hat{R} > 1.01$ and bulk or tail effective sample sizes were below 400 (i.e., 100 effective samples per individual chain), namely the length-at-50% selectivity for the Canadian longline fleet, four time-varying selectivity deviations, recruitment deviations in 1996, 2006, and 2009, and time-varying natural mortality deviations for 1969 and 2019. Low effective sample sizes and corresponding $\hat{R} > 1.01$ indicates that there was limited information for that parameter in the data, so a more informative prior or improved/increased sampling is warranted in future revisions; indeed, all marginal parameters had $\hat{R} \leq 1.02$ and visual inspections of ESS versus chain length (not shown) indicate that increased sampling should eliminate all marginal convergence indicators. Moreover, there is limited available length composition for the Canadian Longline fleet, so it is unsurprising that selectivity parameters have quite high uncertainty (Figure 8, CV(theta)). Similarly, time-varying mortality, selectivity, and catchability parameters are all expected to be quite uncertain, even at convergence. Finally, high uncertainty in some annual recruitment deviations is also not surprising when fitting to length data compared to age composition directly. Parameter lag-1 autocorrelation within chains ranged between -0.5 and 0.6, but most of the density was between -0.4 and 0.4, with an average around 0 as expected for Hamiltonian Monte Carlo.

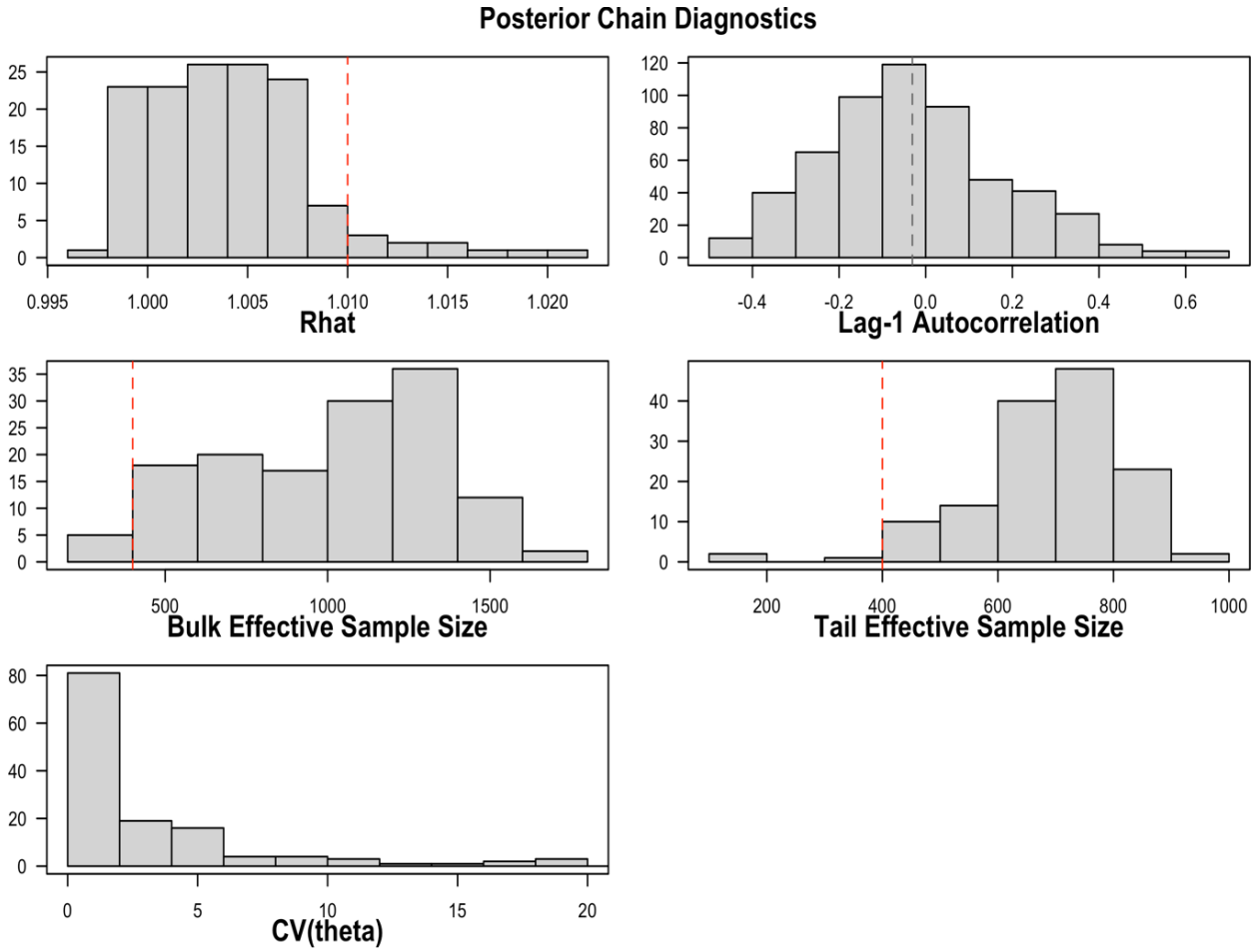


*Figure 8. Distributions of Hamiltonian Monte Carlo convergence diagnostics over the set of leading parameters, Θ, for the SISCAL-GH model, showing scale reduction factors (Rhat), Lag-1 autocorrelation, Bulk and tail effective sample size, and coefficient of variation (CV (theta)). Red vertical lines show minimum threshold for effective sample size, and the grey vertical line shows mean lag-1 autocorrelation.*

## Implications for stock dynamics

SISCAL-GH estimates of female spawning biomass, recruitment, exploitation rate, and natural mortality rates suggest that the GH-0+1 stock is currently above $B_{MSY}$ and, therefore, most likely in a healthy state according to Canadian fisheries policy (Figure 9) (DFO 2006). Posterior distributions for female spawning biomass show a slight increase from unfished equilibrium in 1968 to around 110 kt in 1985. In the early 1980s, natural mortality starts to increase, driving GH-0+1 spawning biomass back down to below unfished levels. The estimated spawning biomass then appears to stabilize around 75 kt by the year 2000, around which time the natural mortality series returns close to the time-averaged $M_0$. When compared to MSY-based reference points (Figure 10), the biomass was almost certainly above $B_{MSY}$ in 2020, but the posterior distribution estimates a 50% probability that the exploitation rates were above $U_{MSY}$ (Figure 11, grey crosshair in 2020). Put another way, the GH-0+1 stock is almost certainly not overfished according to this SISCAL-GH model, but there is a 50% chance that current TACs are overfishing the stock.

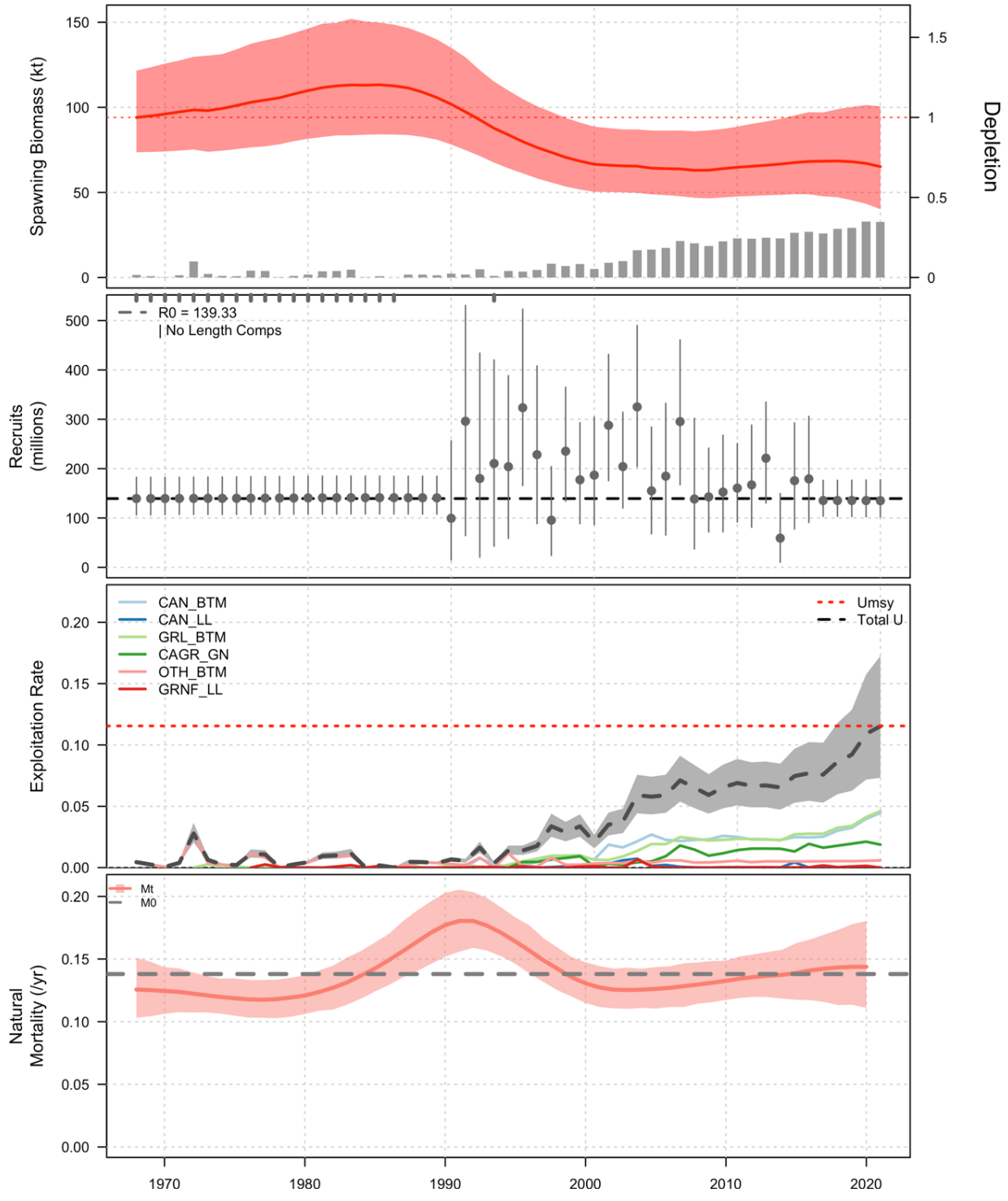


*Figure 9. SISCAL-GH estimates of historical GH-0+1 population dynamics showing (top panel) credibility intervals (red shaded region) and posterior median (red line) estimates of GH-0+1 female spawning biomass and total landings (grey bars), posterior 95% credibility intervals and median recruitment estimates (second row), posterior median harvest rates by fishery and in total (third row), and 95% credibility intervals and posterior median estimates of annual natural mortality rates averaged between males and females (bottom row).*

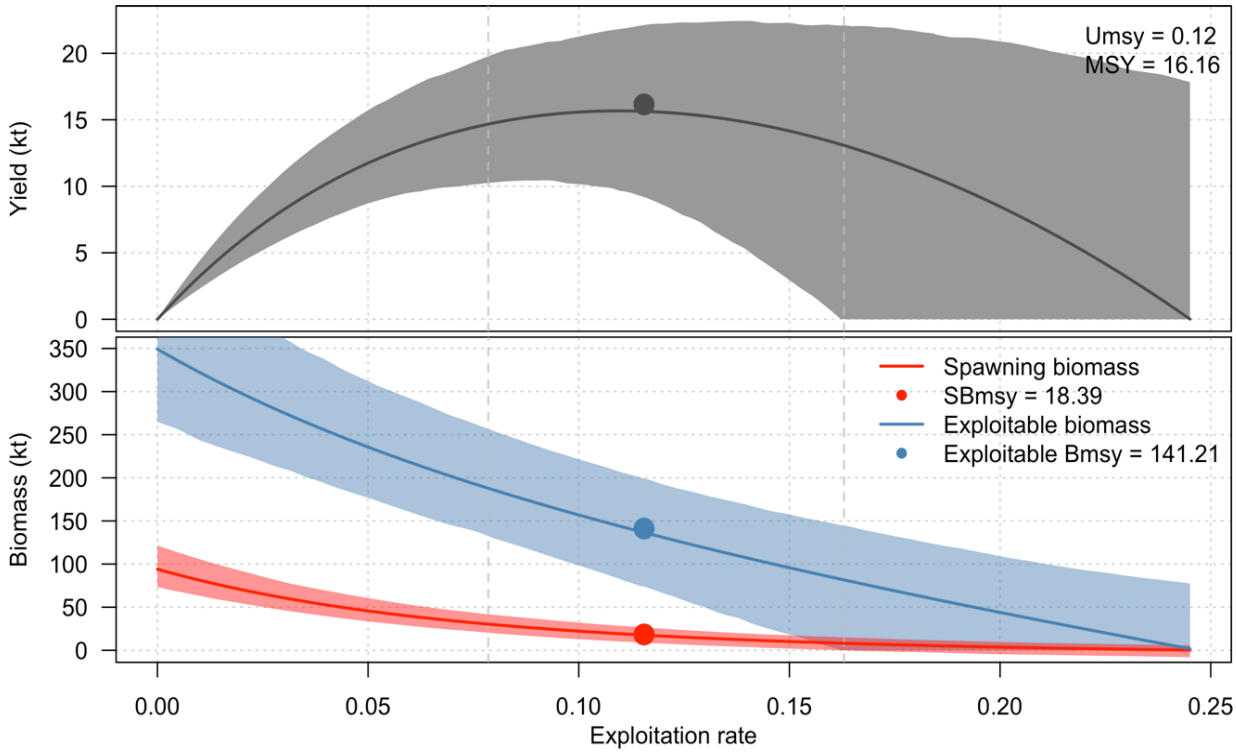


*Figure 10. Equilibrium GH-0+1 yield (top) and biomass (bottom) curves as a function of total exploitation rate. $U_{MSY}$ reference points are shown as closed circles on each line. Posterior 95% credibility intervals in yield and biomass are shown as envelopes, while the 95% credibility interval for $U_{MSY}$ is shown by the vertical dashed lines.*

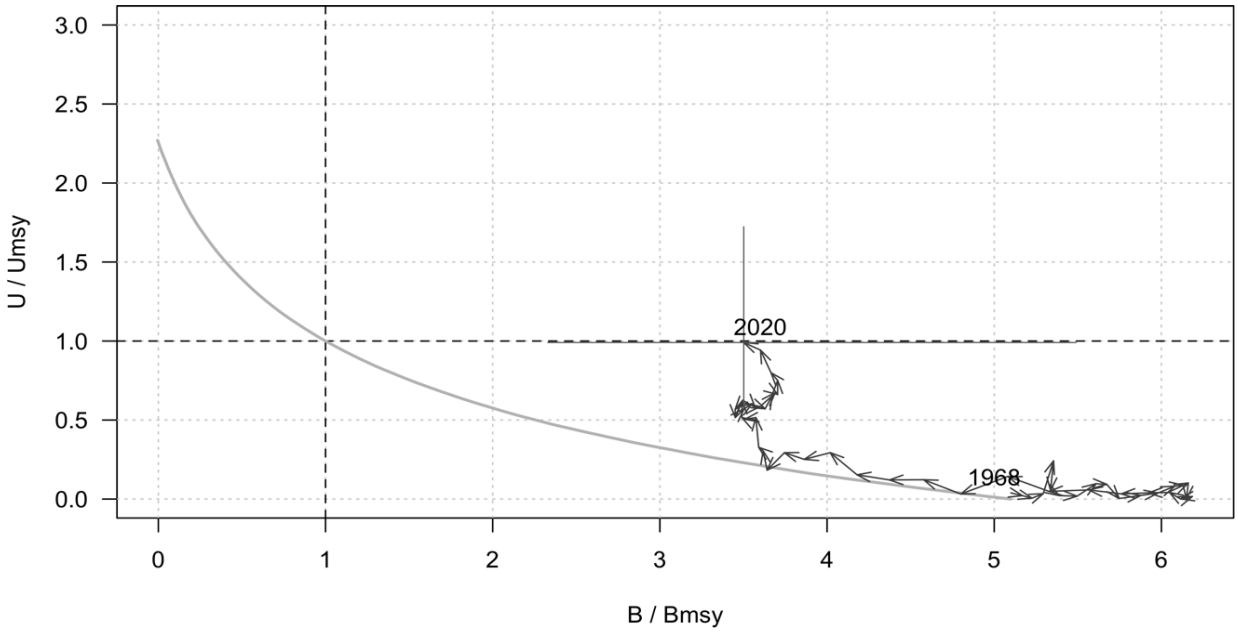


*Figure 11. Phase plot showing posterior median GH-0+1 female spawning biomass (vertical axis) and total exploitation rate (horizontal axis) relative to $U_{MSY}$ reference points. Arrows show the direction of time, beginning in 1968 and ending in 2020. Equilibrium spawning biomass is shown as a faint grey curve in the background of the plot. The distribution of stock status and exploitation rates relative to $MSY$ based reference points in 2020 is shown as a grey crosshair.*

It is expected that the increase in natural mortality estimated by the model during the 1980s was driven by fitting to combined trawl fishery CPUE data. Trawl CPUE declined between the late 1980s and late 1990s, presumably as trawl landings increased while fished area remained static (Treble and Nogueira 2020). While there are well documented issues with non-linearity in fishery CPUE data (i.e., hyper-depletion or hyper-stability), the trawl CPUE data was the only biomass index time series indicating any historical depletion signal for the GH-0+1 stock, which we considered an important signal to include. The increase in natural mortality estimated by the model around the early 1990s is consistent with temporal patterns estimated for other groundfishes such as Atlantic Halibut and Northern Cod (Johnson et al. In prep[1], Cadigan 2016). It is also concurrent with an estimated decrease in natural mortality for Greenland Halibut in NAFO Subarea 2 and Divisions 3KLMNO (Regular 2020); negatively correlated mortality patterns between neighbouring management units of GH may indicate that migration between those areas is being misrepresented as mortality.

Beginning around the year 2000, the increasing combined trawl CPUE was driven by a combination of high recruitment and, possibly, a spatio-temporal expansion of the fishery to year-round timing and areas further north (Treble and Nogueira 2020). In the current SISCAL-GH formulation and data, there was not enough information to distinguish between either factor contributing to the CPUE increase. However, assuming that combined trawl catchability was time-varying after 2014 effectively assigns the increase to spatio-temporal expansion of fishing after that time by allowing catchability to increase instead of forcing the biomass to continue growing. The increase from 2000 - 2014 is then explained by the model with higher recruitment, creating a positive mean recruitment residual around 0.18 standard deviations (Figure 12, lower panel).

[1] Johnson, S., Hubley, B., Cox, S.P., den Heyer, C.E., and Li, L. In prep. Framework assessment of Atlantic Halibut on the Scotian shelf and Southern Grand Banks (NAFO Divs. 3NOPs4VWX5Zc) model update. DFO Can. Sci. Advis. Sec. Res. Doc. In preparation.

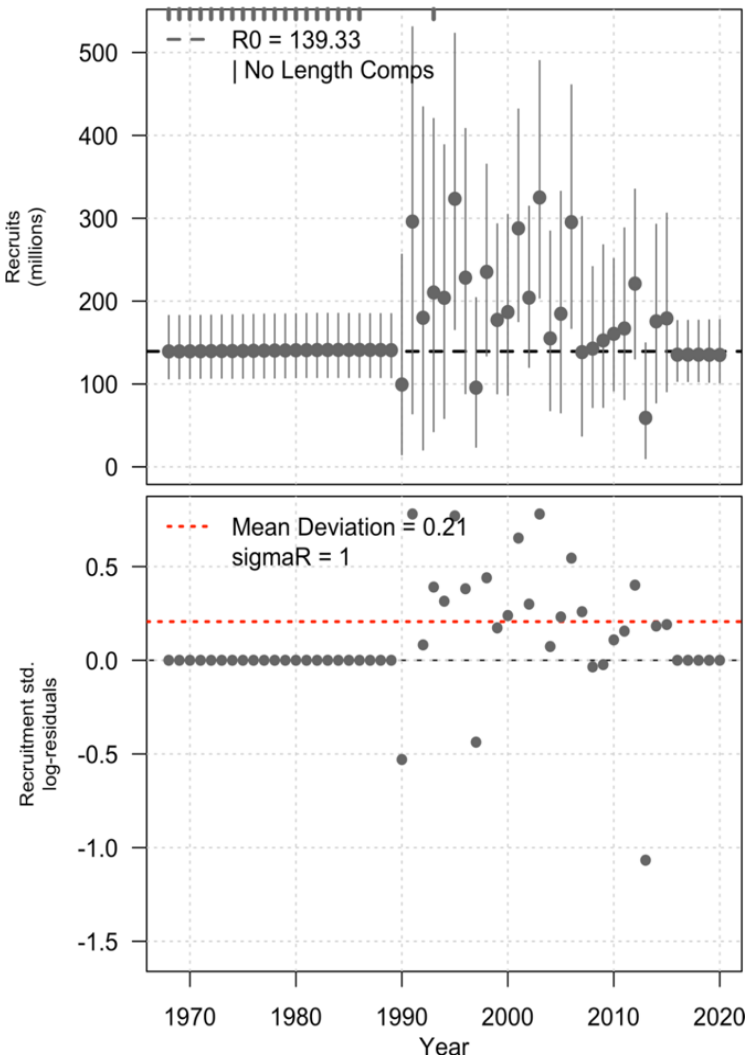


*Figure 12. Time series of age-1 GH-0+1 recruitments (top) and standardised recruitment log-residuals (bottom). Absolute recruitments show equilibrium unfished recruitment $R_0$ (horizontal dashed line), and 95% credibility intervals (vertical line segments), and residuals are plotted with the average estimated residual (horizontal red dashed line).*

Despite a generally healthy picture of the GH-0+1 stock, there is considerable uncertainty in current biomass, and as a result, MSY-based biological reference points. For example, the 95% credible interval (CI) for Bayes posterior female spawning biomass ranges between 60 kt and 160 kt in 2020, and there are similar relative variations in other biomass estimates as well (Figure 13). As a result, the 95% CI for exploitation rates ranges between 0.05 and 0.13 (Figure 9, lower panel), leading to the aforementioned low chance that overfishing is occurring (Figure 11). Similarly, there is high uncertainty in the $MSY$-based reference points with optimal exploitation rate $U_{MSY}$ ranging between 0.07 and 0.16 (Figure 10, upper panel), and similar relative uncertainty in female spawning biomass at MSY (14 - 36 kt) and exploitable biomass (132 - 256 kt) (Figure 10, lower panel).

Stability in spawning biomass since the early 2000s suggests that removals may have been balancing production since that time (Figure 13). That balance has been sustained by the generally high recent recruitments, which may not continue if future recruitments are closer to

the expected stock-recruitment relationship (Figure 14). If that were to happen, then the declines estimated in exploitable, total, and RV biomasses (Figure 13) after 2010 will materialise in the spawning biomass, which, if TACs remain at the current levels, may result in an overfished stock.

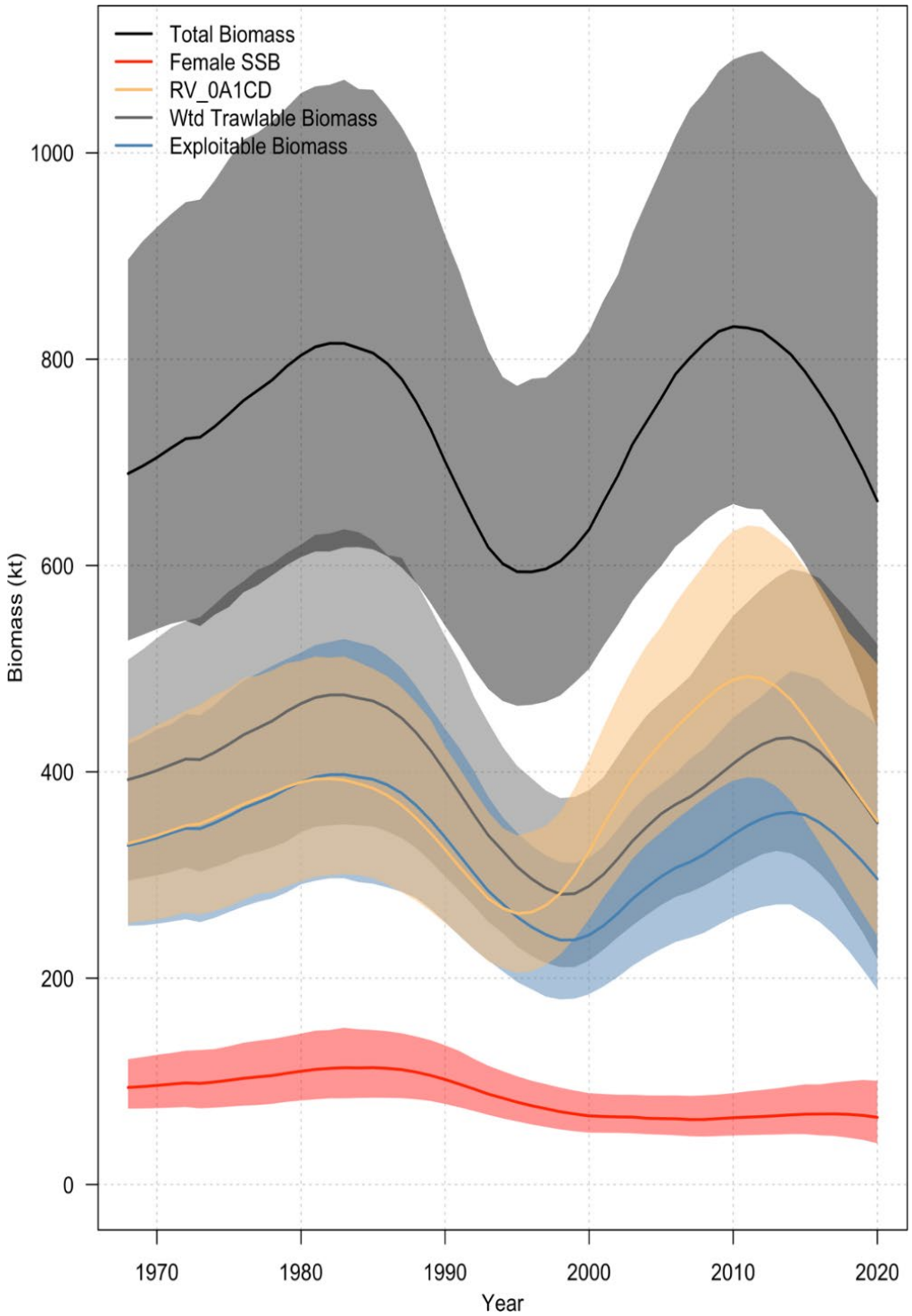


*Figure 13. SISCAL-GH posterior mean and 95% credibility intervals of exploitable (dark blue), total (black), female spawning (red), offshore RV survey (orange), and weighted trawlable (grey) biomass.*

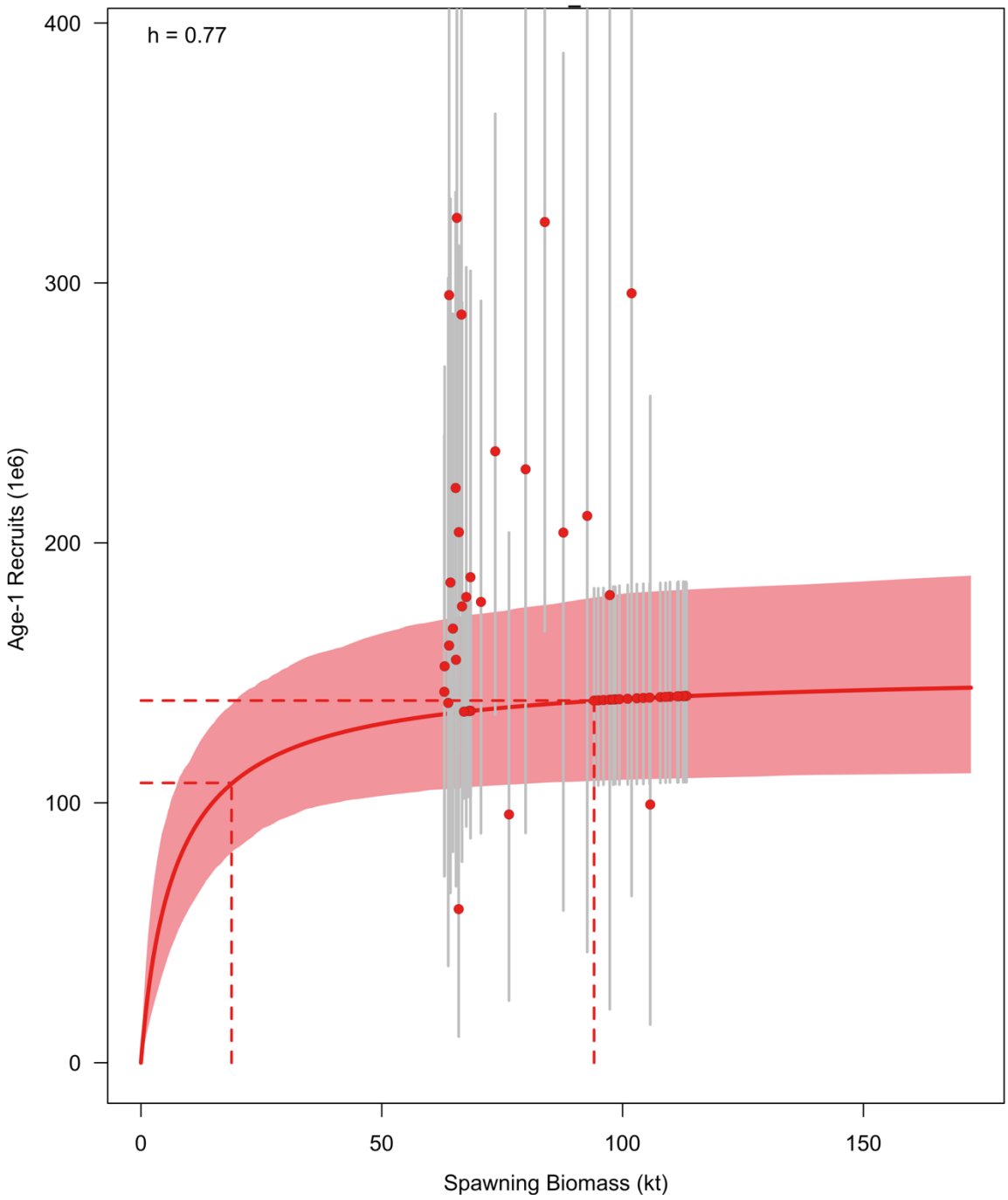


*Figure 14. Bayes posterior stock-recruit relationship showing the posterior mean (solid line), central 95% credibility region (red envelope), as well as estimated recruitments (points) and their 95% credibility intervals.*

## Sensitivity analyses

### Asymptotic Length

In total, there were six alternative GH-0+1 growth models fit to the 2017 age/length samples from the offshore RV survey. All six models were largely the same for ages 1 - 15, but begin diverging for males at age 15, and females around age 16 (Figure 3).

The estimated unfished biomass ($B_0$) decreased as $L_\infty$ increased from 1.0 to 2.0 times the values estimated from the data. The largest percentage drop in unfished biomass occured between 1.0 and 1.2 times the estimated $L_\infty$, where $B_0$ dropped by roughly 25%. After that the decreases in unfished biomass were much smaller, dropping by around 8% each time from 1.2 to 1.6 times the estimated values (Table 6). Unfished recruitment $R_0$, $B_{MSY}$, and MSY all decreased with unfished biomass while natural mortality and the mean recruitment deviation $\bar{\omega}_t$ increased, and optimal exploitation rates $U_{MSY}$ showed no obvious pattern.

*Table 6. Asymptotic length ($L_\infty$) sensitivity analysis results showing SISCAL-GH maximum posterior density estimates of unfished biomass ($B_0$), unfished recruitment ($R_0$), time-averaged natural mortality ($M_0$), MSY-based optimal biomass $B_{MSY}$, optimal harvest rate $U_{MSY}$, and maximum sustainable yield MSY, mean recruitment deviation $\omega_t$, and a convergence indicator (pdHess).*

| multiplier | $L_\infty$ | $B_0$ | $R_0$ | $M_0$ | $B_{MSY}$ | $U_{MSY}$ | $MSY$ | $\bar{\omega}_t$ | pdHess |
|---|---|---|---|---|---|---|---|---|---|
| 1.0 | 64.71, 81.45 | 95.210 | 141.587 | 0.125 | 18.491 | 0.116 | 16.420 | 0.21 | TRUE |
| 1.2 | 77.65, 97.74 | 71.587 | 75.155 | 0.127 | 14.604 | 0.113 | 9.264 | 0.30 | TRUE |
| 1.4 | 90.6, 114.03 | 66.114 | 60.935 | 0.160 | 13.525 | 0.115 | 7.744 | 0.22 | TRUE |
| 1.6 | 103.53, 130.32 | 61.779 | 51.622 | 0.217 | 12.283 | 0.121 | 6.899 | 0.15 | TRUE |
| 1.8 | 116.48, 146.61 | 64.667 | 60.947 | 0.182 | 13.217 | 0.115 | 7.480 | 0.17 | TRUE |
| 2.0 | 129.42, 162.9 | 60.654 | 53.196 | 0.227 | 12.016 | 0.122 | 6.699 | 0.14 | TRUE |

As unfished biomass decreased, the temporal patterns in spawning biomass changed significantly (Figure 15). The initial increase in spawning biomass flattens out and changes concavity as the $L_\infty$ multiplier increases, eventually becoming a monotonically decreasing "one-way trip" (Hilborn and Walters 1992) under the highest $L_\infty$ values. This reduction in informative contrast is mirrored in the weighted trawlable biomass series that is used to generate expected values for the combined trawl CPUE data series. As a result, models with higher $L_\infty$ have a lower negative log likelihood for CPUE data (Table 7), but no longer track the initial depletion signal between 1989 and 1999 in that index (not shown).

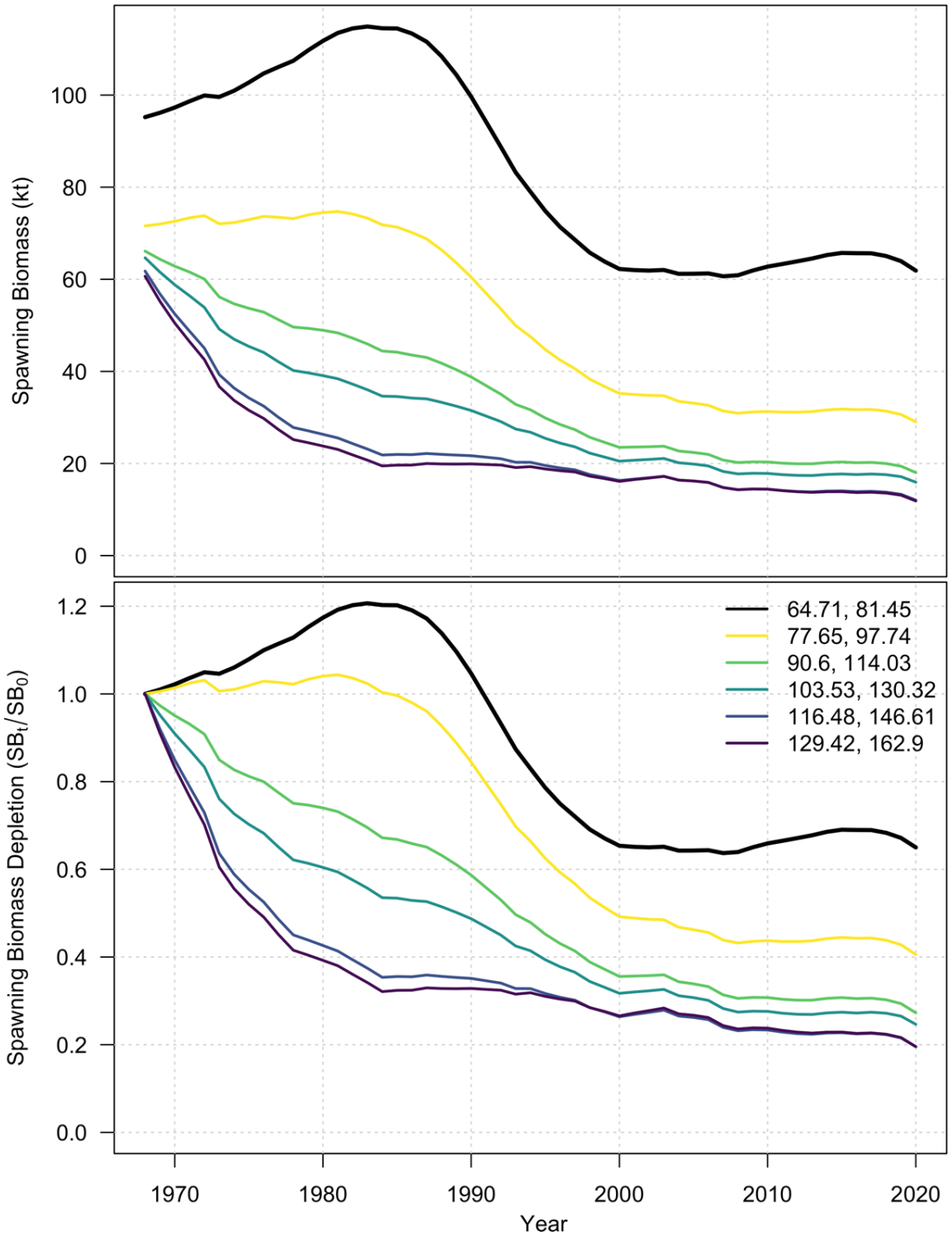


*Figure 15. Asymptotic length ($L_\infty$) sensitivity analysis estimates of female spawning biomass (top) and spawning biomass depletion (bottom).*

*Table 7. Asymptotic length sensitivity analysis results, showing absolute ($L_\infty$) asymptotic length, as well as their value relative to the base model (Multiplier), and negative log likelihood function values for all data (TotLike) and contributions from biomass/abundance index components for data from the offshore survey (idxRV0A1CD), inshore survey (idxRVSFW1AF), and trawl fishery CPUE (CPUE_BTM).*

| Multiplier | Linf | TotLike | idxRV0A1CD | idxRVSFW1AF | CPUE_BTM |
|---|---|---|---|---|---|
| 1.0 | 64.71, 81.45 | 153.42 | -13.94 | -13.4 | -67.09 |
| 1.2 | 77.65, 97.74 | 172.5 | -13.48 | -13.33 | -67.98 |
| 1.4 | 90.6, 114.03 | 139.16 | -13.48 | -12.21 | -65.2 |
| 1.6 | 103.53, 130.32 | 105.65 | -14.12 | -10.7 | -54.87 |
| 1.8 | 116.48, 146.61 | 134.35 | -13.73 | -11.31 | -61.86 |
| 2.0 | 129.42, 162.9 | 105.57 | -14.3 | -10.43 | -51.31 |

Based on negative log-likelihood values (Table 8), the best candidate $L_\infty$ multipliers for improving fits to the GRNF_LL and CAGR_GN length compositions are 1.4 and 1.6, which we recommend for deeper analysis in future, and as a possible candidate for an alternative operating model hypothesis.

*Table 8. Asymptotic length sensitivity analysis results, showing absolute ($L_\infty$) asymptotic length, as well as their value relative to the base model (Multiplier), and negative log likelihood function values for all data (TotLike) and contributions from the inshore proportion female data (propFSFW), and length composition data from the Canadian trawl (lenCANBTM), Canadian longline (lenCANLL), Greenland trawl (lenGRLBTM), combined gillnet (lenCAGRGN), all other trawl (lenOTHBTM), Greenland/Russia/Norway/Faroe Islands longline (lenGRNFLL), and offshore (lenRV0A1CD) and inshore (lenRVSFW1AF) surveys.*

| Multiplier | Linf | TotLike | propFSFW | lenCANBTM | lenCANLL | lenGRLBTM | lenCAGRGN | lenOTHBTM | lenGRNFLL | lenRV0A1CD | lenRVSFW1AF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.0 | 64.71, 81.45 | 153.42 | -152.45 | -57.75 | 2.98 | 121.28 | 7.84 | 185.06 | 69.87 | -54.27 | 125.28 |
| 1.2 | 77.65, 97.74 | 172.5 | -174.55 | -34.29 | 4.99 | 103.51 | 17.87 | 220.78 | 63.5 | -42.7 | 108.17 |
| 1.4 | 90.6, 114.03 | 139.16 | -181.83 | -37.56 | 5.04 | 89.27 | 21.92 | 214.52 | 58.88 | -43.42 | 103.23 |
| 1.6 | 103.53, 130.32 | 105.65 | -179.95 | -35.38 | 5.07 | 84.23 | 25.6 | 175.2 | 54.6 | -43.8 | 99.76 |
| 1.8 | 116.48, 146.61 | 134.35 | -180.06 | -38.66 | 5.14 | 86.62 | 23.82 | 212.64 | 56.15 | -44.1 | 99.7 |
| 2.0 | 129.42, 162.9 | 105.57 | -179.9 | -35.25 | 4.98 | 83.32 | 26.23 | 172.53 | 55.05 | -43.89 | 98.53 |

### Time-varying vs constant parameters

Similar to the $L_\infty$ sensitivity, switching between time-varying and constant $M$ and $q$ for the CPUE index had implications for both unfished biomass and fitting the CPUE index data. When keeping $q$ constant but allowing $M$ to vary, the unfished biomass increased slightly from 90 kt at unfished up to 126 kt (Table 9). With the removal of the time-varying catchability parameter the model decreased recent $M$ and increased recruitment to fit to the higher CPUE towards the end of the data series. At the same time, the mean recruitment deviation increased to 0.27 standard deviations (Table 9, $\bar{\omega}_t$) and brought spawning biomass back up towards unfished in 2020 (Figure 16).

When $M$ was constant, almost all temporal variability was removed from the spawning biomass time series. In addition, unfished biomass dropped significantly for both $q$ hypotheses, but less so for the constant $q$ hypothesis (90.4 kt at unfished versus 82.9 kt; Table 9). Similar to the growth model sensitivity, the drop in unfished biomass and flatter trend led to poorer qualitative fits to the early depletion in the CPUE data (not shown).

*Table 9. Time-varying CPUE catchability and natural mortality sensitivity analysis results showing SISCAL-GH maximum posterior density estimates of unfished biomass ($B_0$), unfished recruitment ($R_0$), time-averaged natural mortality ($M_0$), MSY-based optimal biomass $B_{MSY}$, optimal harvest rate $U_{MSY}$, and maximum sustainable yield MSY, mean recruitment deviation $\omega_t$, and a convergence indicator (pdHess).*

| Sensitivity | M | $B_0$ | $R_0$ | $M_0$ | $B_{MSY}$ | $U_{MSY}$ | $MSY$ | $\bar{\omega}_t$ | pdHess |
|---|---|---|---|---|---|---|---|---|---|
| con q | conM | 90.440 | 134.364 | 0.138 | 19.192 | 0.107 | 14.495 | 0.45 | TRUE |
| con q | tvM | 126.130 | 187.787 | 0.127 | 24.741 | 0.115 | 21.487 | 0.27 | TRUE |
| tv q | conM | 82.902 | 122.797 | 0.138 | 17.635 | 0.107 | 13.244 | 0.39 | TRUE |
| tv q | tvM | 95.210 | 141.587 | 0.125 | 18.491 | 0.116 | 16.420 | 0.21 | TRUE |

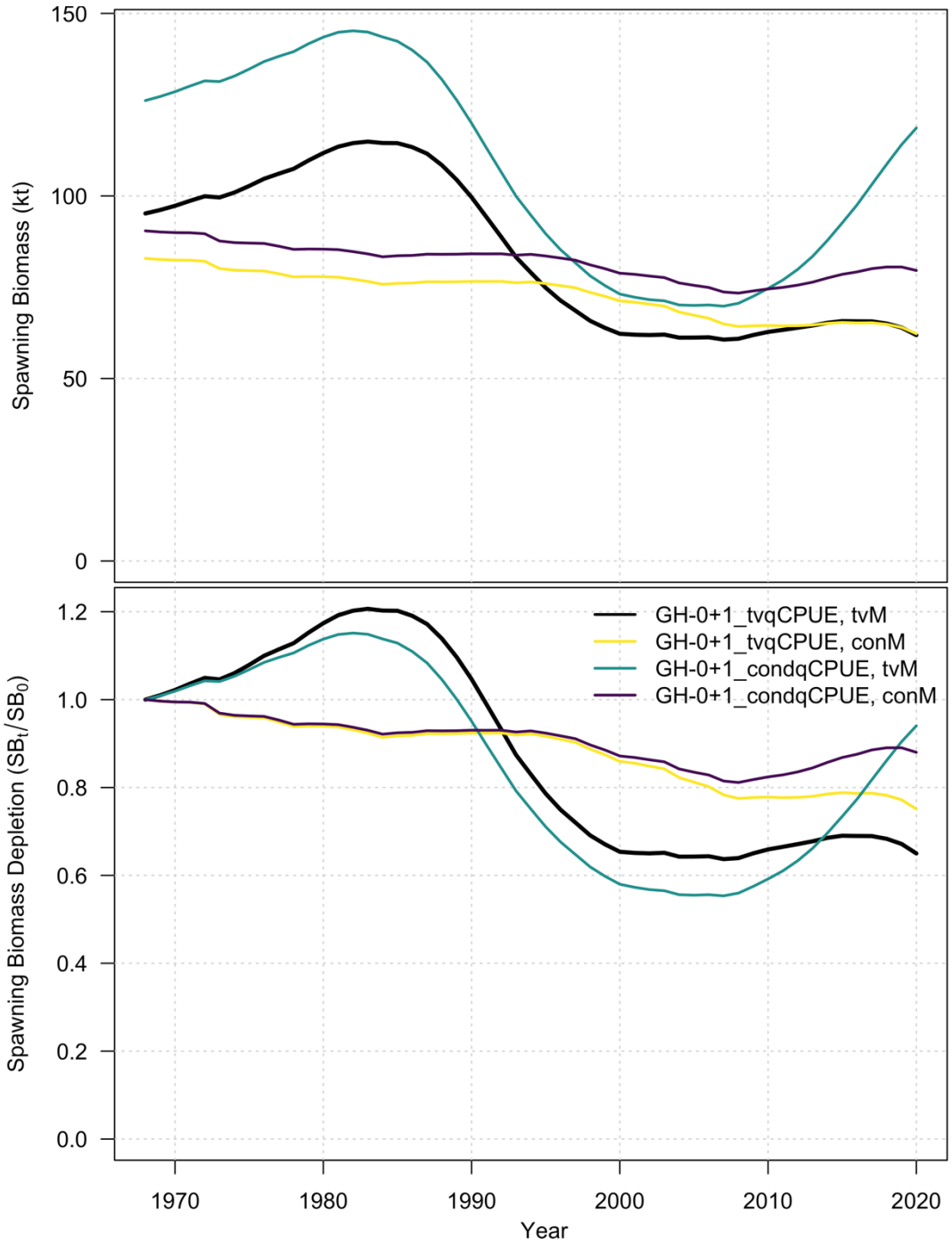


*Figure 16. Time-varying catchability and mortality sensitivity analysis estimates of female spawning biomass (top) and spawning biomass depletion (bottom).*

### Length composition sample size and tail compression

SISCAL-GH is sensitive to the way that length composition data are filtered for including within the logistic normal likelihood function. Raising the tail compression parameter (i.e., minimum threshold on the proportion of samples under which observations and predictions are aggregated with neighbouring bins) causes a steep drop in unfished biomass. Unfished biomass $B_0$, and reference points $B_{MSY}$ and MSY all drop by close to 50%, with a jump in mean

recruitment deviation by about 61% (Table 10, minProp hypotheses). This suggests that the information in the tails of the length composition data drive a lot of the model estimates and is related to the difficulty fitting data in the low probability length bins (i.e., GRNF longline fisheries). While the absolute biomass scale of the stock changed a lot under the minimum bin proportion, the productivity appeared to be insensitive, with only minor differences in $M_0$ and $U_{MSY}$.

In contrast, SISCAL-GH model was only slightly sensitive to the minimum sample size for inclusion of length composition data (Table 10, minSampSize hypotheses). As minimum sample sizes were increased from 100 to 800 per year, there was a relatively small drop in unfished biomass, going from 101 kt to 95 kt, close to the base model estimate where a minimum sample size of 700 is used. As with the minProp sensitivity, the productivity of the stock appeared to be insensitive to the minimum sample size, with $M_0$ and $U_{MSY}$ differing on the order of $10^{-3}$.

*Table 10. Length composition minimum sample and minimum bin proportion threshold sensitivity analysis results showing SISCAL-GH maximum posterior density estimates of unfished biomass ($B_0$), unfished recruitment ($R_0$), time-averaged natural mortality ($M_0$), MSY-based optimal biomass $B_{MSY}$, optimal harvest rate $U_{MSY}$, and maximum sustainable yield $MSY$, mean recruitment deviation $\omega_t$, and a convergence indicator (pdHess).*

| modelHyp | $B_0$ | $R_0$ | $M_0$ | $B_{MSY}$ | $U_{MSY}$ | $MSY$ | $\bar{\omega}_t$ | pdHess |
|---|---|---|---|---|---|---|---|---|
| minProp.001 | 95.210 | 141.587 | 0.125 | 18.491 | 0.116 | 16.420 | 0.21 | TRUE |
| minProp.01 | 57.084 | 90.058 | 0.133 | 11.116 | 0.110 | 9.987 | 0.34 | TRUE |
| minProp.02 | 50.320 | 80.092 | 0.133 | 9.624 | 0.113 | 8.827 | 0.30 | TRUE |
| minSampSize100 | 101.481 | 149.703 | 0.125 | 19.783 | 0.115 | 17.414 | 0.23 | TRUE |
| minSampSize400 | 100.076 | 148.188 | 0.125 | 19.529 | 0.115 | 17.195 | 0.23 | TRUE |
| minSampSize600 | 96.878 | 143.526 | 0.125 | 18.767 | 0.116 | 16.707 | 0.19 | TRUE |
| minSampSize800 | 95.278 | 141.803 | 0.125 | 18.534 | 0.116 | 16.423 | 0.20 | TRUE |

## Retrospective analysis

Retrospective model fits show a shift in unfished biomass and abundance ($B_0, R_0$) over time (Table 11). Since 2010, model estimates of unfished biomass ($B_0$), optimal biomass ($B_{MSY}$) and yield (MSY) have increased by around 45%. On the other hand, spawning biomass depletion has remained fairly consistent over the 2010 - 2020 period (Figure 17), staying within 10 percentage points for most of the years with overlap between peels.

*Table 11. Retrospective analysis results showing SISCAL-GH maximum posterior density estimates of unfished biomass ($B_0$), unfished recruitment ($R_0$), time-averaged natural mortality ($M_0$), MSY-based optimal biomass ($B_{MSY}$), optimal harvest rate ($U_{MSY}$), and maximum sustainable yield ($MSY$), mean recruitment deviation($\omega_t$), and a convergence indicator (pdHess).*

| Peel | $B_0$ | $R_0$ | $M_0$ | $B_{MSY}$ | $U_{MSY}$ | $MSY$ | $\bar{\omega}_t$ | pdHess |
|---|---|---|---|---|---|---|---|---|
| 2010 | 66.472 | 98.800 | 0.123 | 13.053 | 0.118 | 11.636 | 0.30 | TRUE |
| 2011 | 81.479 | 121.027 | 0.126 | 16.231 | 0.116 | 13.875 | 0.37 | TRUE |
| 2012 | 90.386 | 133.654 | 0.125 | 17.709 | 0.117 | 15.518 | 0.18 | TRUE |
| 2013 | 91.511 | 134.220 | 0.126 | 18.038 | 0.112 | 15.121 | 0.43 | TRUE |
| 2014 | 92.465 | 135.628 | 0.125 | 18.154 | 0.111 | 15.290 | 0.32 | TRUE |
| 2015 | 88.639 | 130.150 | 0.125 | 17.421 | 0.112 | 14.641 | 0.37 | TRUE |
| 2016 | 97.918 | 143.663 | 0.125 | 19.402 | 0.113 | 16.361 | 0.29 | TRUE |

| Peel | $B_0$ | $R_0$ | $M_0$ | $B_{MSY}$ | $U_{MSY}$ | $MSY$ | $\bar{\omega}_t$ | pdHess |
|---|---|---|---|---|---|---|---|---|
| 2017 | 100.324 | 148.091 | 0.124 | 19.502 | 0.116 | 17.165 | 0.13 | TRUE |
| 2018 | 97.584 | 144.129 | 0.125 | 19.003 | 0.114 | 16.696 | 0.17 | TRUE |
| 2019 | 99.420 | 147.232 | 0.125 | 19.446 | 0.113 | 17.061 | 0.19 | TRUE |
| 2020 | 95.210 | 141.587 | 0.125 | 18.491 | 0.116 | 16.420 | 0.21 | TRUE |

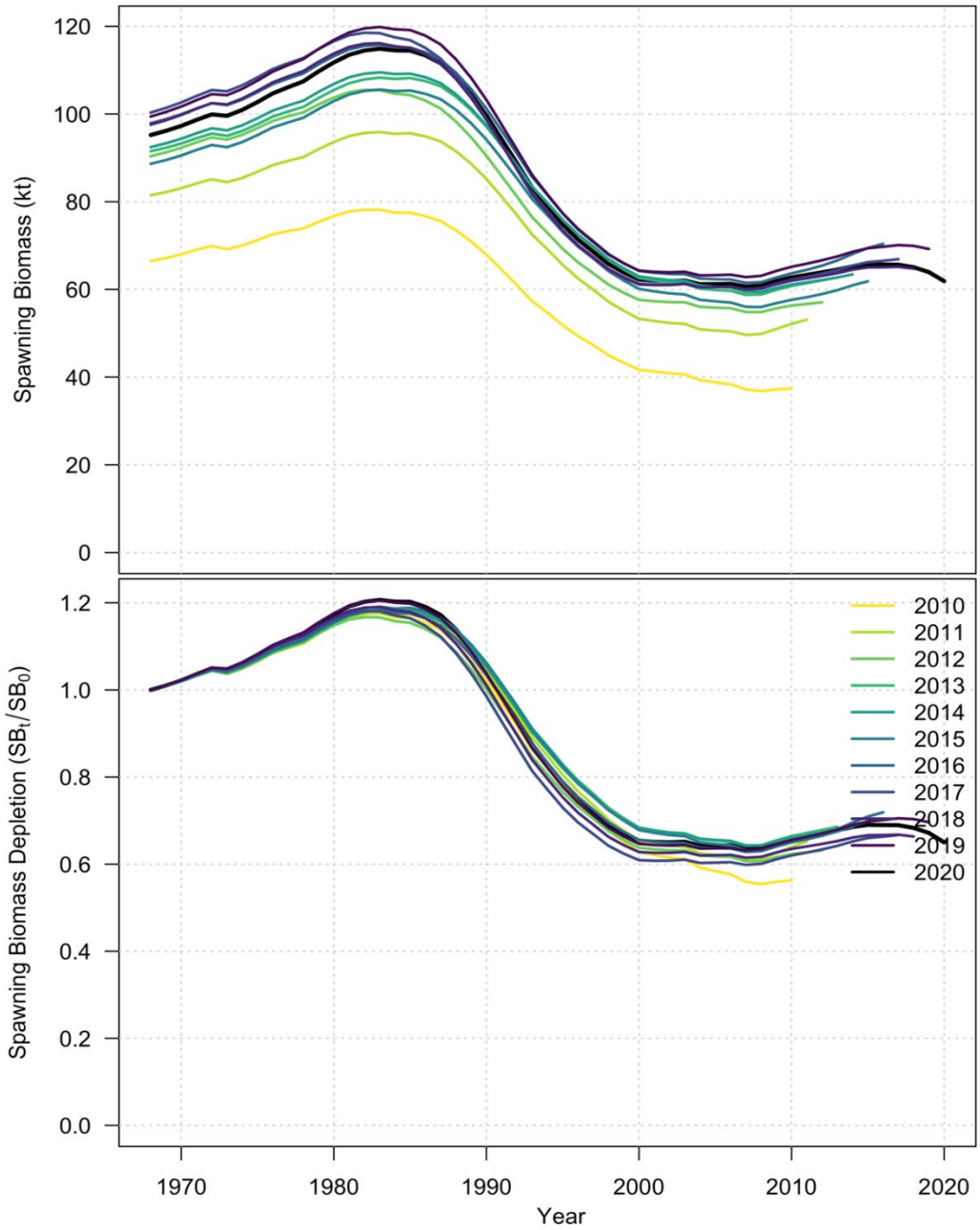


*Figure 17. SISCAL-GH retrospective estimates of female spawning biomass from 2010 to 2020.*

The increase in unfished biomass is largely driven by the increasing catch, and somewhat flat (or missing) survey indices. Since 2017, all peels have had estimates of unfished and terminal

biomass varying around similar values. The year 2017 is also the last usable offshore RV survey data point, and TACs have been fairly stable since this time. Stable TACs, flat survey data series, and limited recruitment information mean that there has been little added information about the scale of the GH-0+1 stock in intervening years.

## Simulation self-tests

The simulation self-test suggests that SISCAL-GH is reasonably good at recapturing estimates of leading model parameters and survey biomass time series. MREs and MAREs for selected parameters are given for each data scenario in Table 12, while relative error distributions are shown for some of the same parameters in Figure 18 and for spawning and vulnerable biomass time series in Figure 19.

*Table 82. Median relative errors (MREs) and median absolute relative errors (MAREs) for selected SISCAL-GH parameters under the simulation-evaluation self-test.*

| Variable | MRE | MARE | MRE | MARE | MRE | MARE |
|---|---|---|---|---|---|---|
| B0 | -0.01 | 0.06 | -0.02 | 0.04 | -0.03 | 0.09 |
| H | -0.03 | 0.03 | 0.00 | 0.00 | -0.03 | 0.03 |
| M m | 0.04 | 0.04 | 0.00 | 0.00 | 0.04 | 0.04 |
| M f | 0.06 | 0.06 | 0.00 | 0.00 | 0.06 | 0.06 |
| L50A CAN BTM | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.01 |
| L95A CAN BTM | 0.02 | 0.02 | 0.00 | 0.00 | 0.01 | 0.01 |
| L50D CAN BTM | 0.02 | 0.02 | 0.00 | 0.00 | 0.02 | 0.02 |
| L95D CAN BTM | 0.05 | 0.05 | 0.00 | 0.00 | 0.04 | 0.04 |
| L50A CAN LL | -0.02 | 0.04 | 0.00 | 0.00 | -0.02 | 0.05 |
| L95A CAN LL | -0.01 | 0.04 | 0.00 | 0.00 | -0.01 | 0.04 |
| L50A GRL BTM | 0.03 | 0.03 | 0.00 | 0.00 | 0.03 | 0.03 |
| L95A GRL BTM | 0.05 | 0.05 | 0.00 | 0.00 | 0.05 | 0.05 |
| L50A OTH BTM | 0.03 | 0.03 | 0.00 | 0.00 | 0.03 | 0.03 |
| L95A OTH BTM | 0.06 | 0.06 | 0.00 | 0.00 | 0.06 | 0.06 |
| L50A GRNF LL | -0.01 | 0.01 | 0.00 | 0.00 | -0.01 | 0.01 |
| L95A GRNF LL | -0.01 | 0.01 | 0.00 | 0.00 | -0.01 | 0.01 |
| q RV 0A1CD | -0.07 | 0.10 | 0.02 | 0.05 | -0.06 | 0.09 |
| tau RV 0A1CD | -0.04 | 0.04 | -0.10 | 0.17 | -0.10 | 0.17 |
| L50A RV 0A1CD | -0.02 | 0.02 | 0.00 | 0.00 | -0.02 | 0.02 |
| L95A RV 0A1CD | -0.02 | 0.02 | 0.00 | 0.00 | -0.02 | 0.02 |
| L50D RV 0A1CD | 0.01 | 0.01 | 0.00 | 0.00 | 0.01 | 0.01 |
| L95D RV 0A1CD | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.01 |
| q RV SFW1AF | -0.06 | 0.10 | -0.02 | 0.07 | -0.06 | 0.12 |
| tau RV SFW1AF | 0.03 | 0.03 | -0.06 | 0.12 | -0.07 | 0.13 |
| L50A RV SFW1AF | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| L95A RV SFW1AF | 0.01 | 0.01 | 0.00 | 0.00 | 0.01 | 0.01 |
| L50D RV SFW1AF | 0.03 | 0.03 | 0.00 | 0.00 | 0.03 | 0.03 |
| L95D RV SFW1AF | 0.04 | 0.04 | 0.00 | 0.00 | 0.04 | 0.04 |

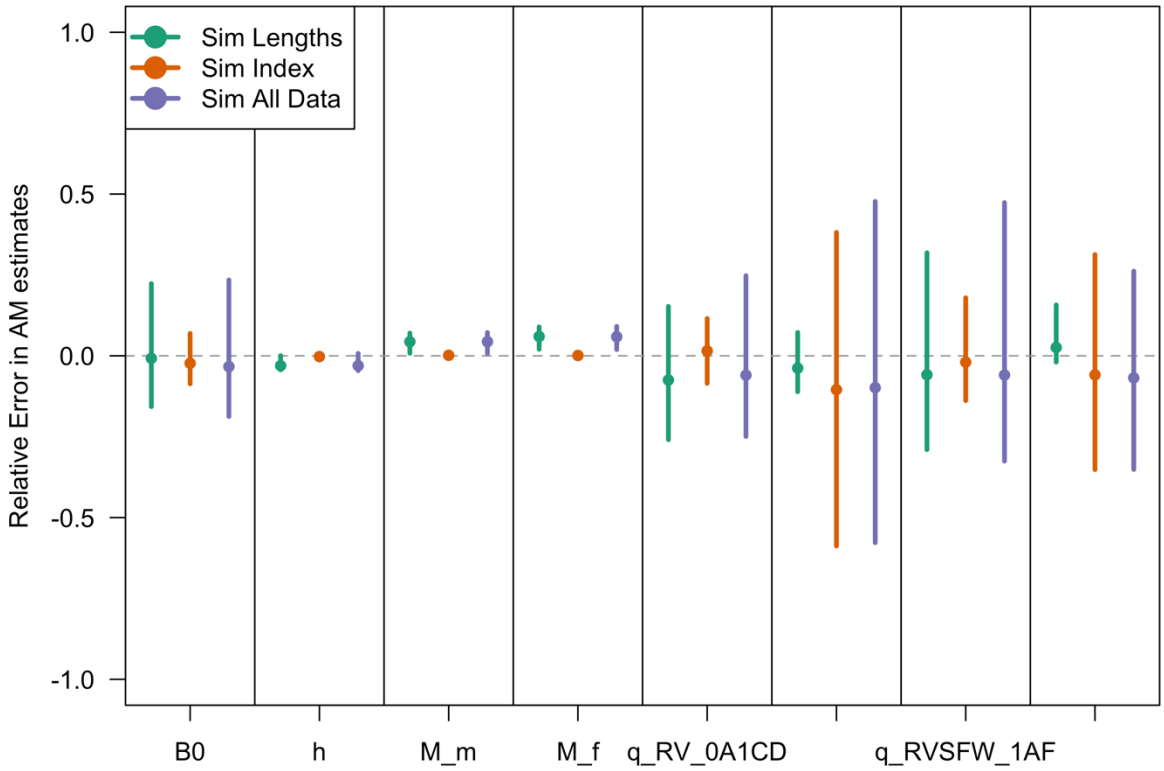


*Figure 18. Median (points) and central 95% (segments) of distributions of relative errors for SISCAL-GH Assessment Model (AM) estimates of leading parameter under the three simulation self-test scenarios (colours).*

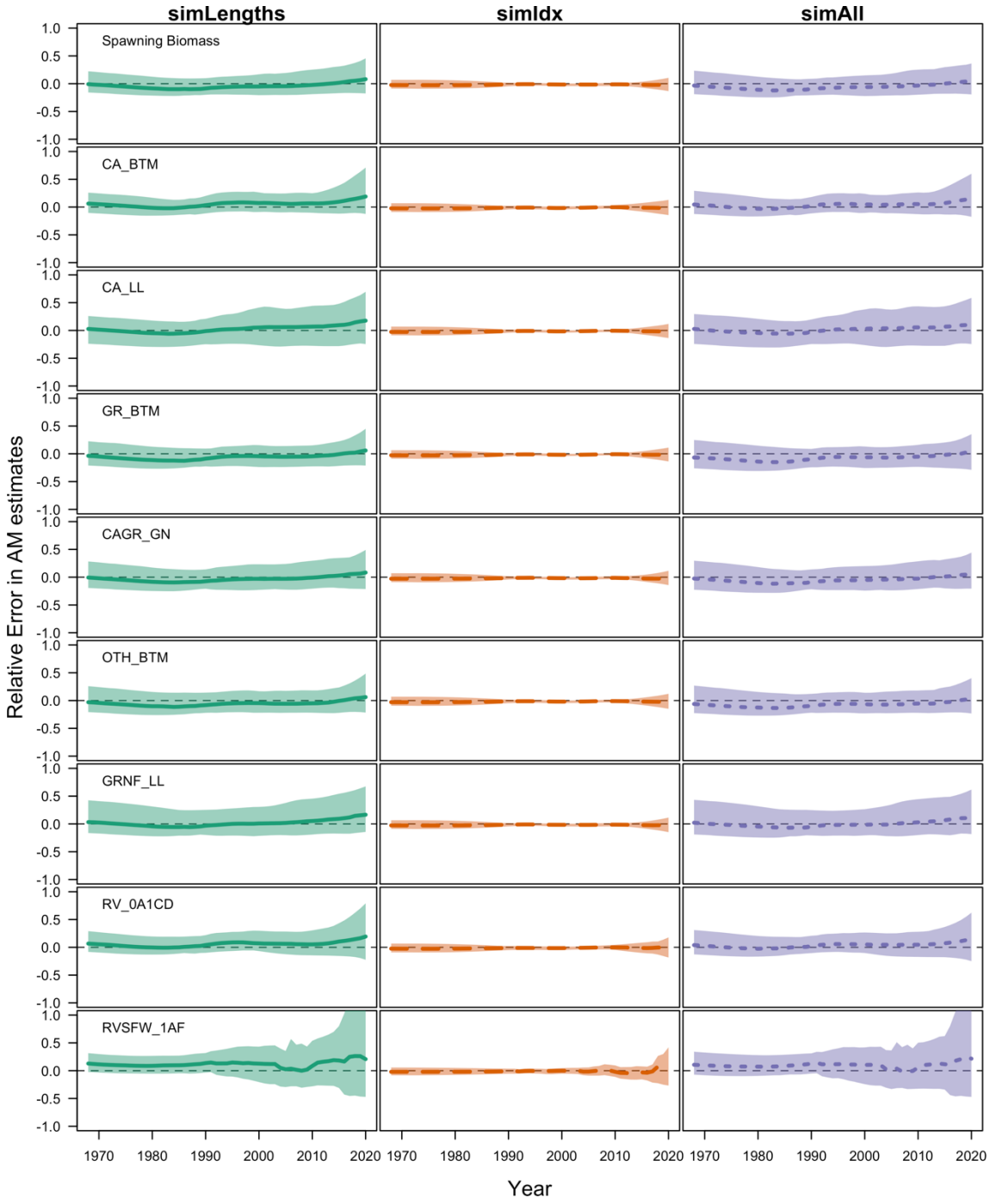


*Figure 19. Distributions of relative errors in SISCAL-GH estimates of female spawning biomass (top), and vulnerable biomass time series for each modeled fleet (remaining rows) under the three simulation self-test scenarios (columns). Heavy lines show median relative errors (i.e., median bias), while shaded regions show the central 95% (i.e., variation) of errors in each year.*

On average, SISCAL-GH is more sensitive to simulated length composition data than to simulated biomass and abundance index data, which is consistent with other length- and/or age-structured stock assessments since there are simply more data points to fit. MREs under the simulated index data only scenario (simIdx) are generally much lower than the scenarios including simulated length composition (simLengths) for unfished biomass $(B_0)$, stock-recruit

steepness ($h$), male/female natural mortality ($M_m/M_f$), and catchability ($q$) for offshore and inshore surveys. Similarly, selectivity parameters were also sensitive to the simulated lengths, but not index data. In contrast, observation error standard deviations ($\tau$) were often more biased and less precise under the simulated index scenario (Table 12, Figure 18). This dichotomous behaviour is expected, since length compositions inform selectivity and mortality parameters, which affect catchability and biological parameter estimates, while (unbiased) simulated indices affect model estimates of precision (i.e., $\tau$) more, as reflected in the time-series of biomass relative errors (Figure 19).

Overall, SISCAL-GH estimates of spawning biomass and vulnerable biomass for the commercial fleets were slightly negatively biased under the simLengths scenario, but not significantly so, and close to unbiased under the simIdx and simAll scenarios. Under the simLengths scenario, median relative errors for spawning and vulnerable biomass for fisheries were around -0.1, but relative error envelopes always contained 0 (Figure 19). By extension, estimates of exploitable biomass would also be slightly negatively biased (not shown). In contrast, relative error distributions for the survey biomass estimates are slightly positively biased under the simLengths and simAll scenario, with median errors around 0.1 on average. Most time series had very low and stable variation in relative errors under each scenario, with the exception of the inshore RV survey where variability increased towards the end of the time series due to time-varying selectivity.

## ADAPTIVE MODEL-/INDEX-BASED MANAGEMENT PROCEDURE EVALUATION FOR GH-0+1

The adaptive model-/index-based MP performs relatively well based on the metrics we defined, which were chosen to roughly fit into NAFO PA policy guidelines (Table 13). Under either precision scenario, there was zero probability that female spawning biomass dropped below $B_{lim} = 0.3B_{MSY}$ over the 47 year simulation of the GH-0+1 stock. Moreover, the probability of being above $B_{MSY}$ was around 82% for the whole projection period, and around 71% during the last 10 years of the projection, which is well above a probability of 50% usually associated with target biomass levels (DFO 2006). The probability of overfishing is around 22% for both scenarios, which is a little higher than the NAFO PA policy limit of 20% (Brodie et al. 2013), but the difference is small enough that the procedure can be tuned via a slightly lower $F_{buf}$ value. Despite the higher-than-acceptable probability of overfishing in the earlier period, the expected exploitation rate above $U_{MSY}$ was only 0.13, or 8% higher than the $U_{MSY}$ limit.

*Table 93. Example performance metrics for the GH-0+1 stock under two management procedures and two newRV survey precision scenarios.*

| Scenario | MP | pLRP | pBtGtBmsy | pBtGtBmsyEnd | pOverfish | mUoverfish | avgCatch10 | catchAAV10 |
|---|---|---|---|---|---|---|---|---|
| ***Adaptive model/index-based procedure*** | | | | | | | | |
| isPrec | qAdapt | 0.000 | 0.821 | 0.710 | 0.213 | 0.13 | 25.1 | 0.06 |
| osPrec | qAdapt | 0.000 | 0.821 | 0.707 | 0.219 | 0.13 | 25.1 | 0.06 |
| ***Non-adaptive index-based procedure*** | | | | | | | | |
| osPrec | qFixed | 0.016 | 0.667 | 0.443 | 0.423 | 0.14 | 25.1 | 0.06 |

Catch statistics in the first 10 years are heavily weighted by the initial constant TAC period from 2021 - 2024 (Table 13). Average catch over the 2021 - 2030 period is 25.1 kt, which is only slightly lower than the current TAC of around 32 kt, and the AAV is 6%, again reflecting the early constant TACs and a relatively slow decline of TACs towards the MSY level given that they are set on a biennial frequency.

Given the very similar performance under both precision scenarios, here we present detailed results for the isPrec scenario only. The same plots for the osPrec scenario are given in Appendix E (Figures E.1 - E.3). To compare the simulated variation among scenarios, Figures E.4 and E.5 show a single replicate of simulated abundance indices for the isPrec and osPrec scenarios, respectively. We also discuss the non-adaptive procedure at the end of this section in contrast to the adaptive procedure, with plots also given in Appendix E (Figures E.6 - E.8). Similarly, the bias and precision of simulated SISCAL-GH assessments was very similar under both isPrec and osPrec scenarios (Appendix F).

The adaptive procedure does a good job of managing the development phase of the fishery, fishing down the standing biomass in a controlled manner across all the posterior draws, i.e., the range of parameter uncertainty (Figure 20), while also allowing the TACs to change in response to variation in survey biomass indices (Figure 21). Final median female spawning biomass is slightly above the posterior median $B_{MSY}$ value (Figure 20), which is expected given the target maximum fishing mortality rate of $F_{buf} = 0.8F_{MSY}$, or exploitation rate of around $0.8U_{MSY}$ (Figure 22, compare solid median line to horizontal black dashed line). Year-to year, the probability of exceeding $U_{MSY}$ starts out above 20%, based on the constant TACs for the first 4 projection years, but settles below 20% after about 25 years as the procedure adapts (Figure 22, dark grey envelope).

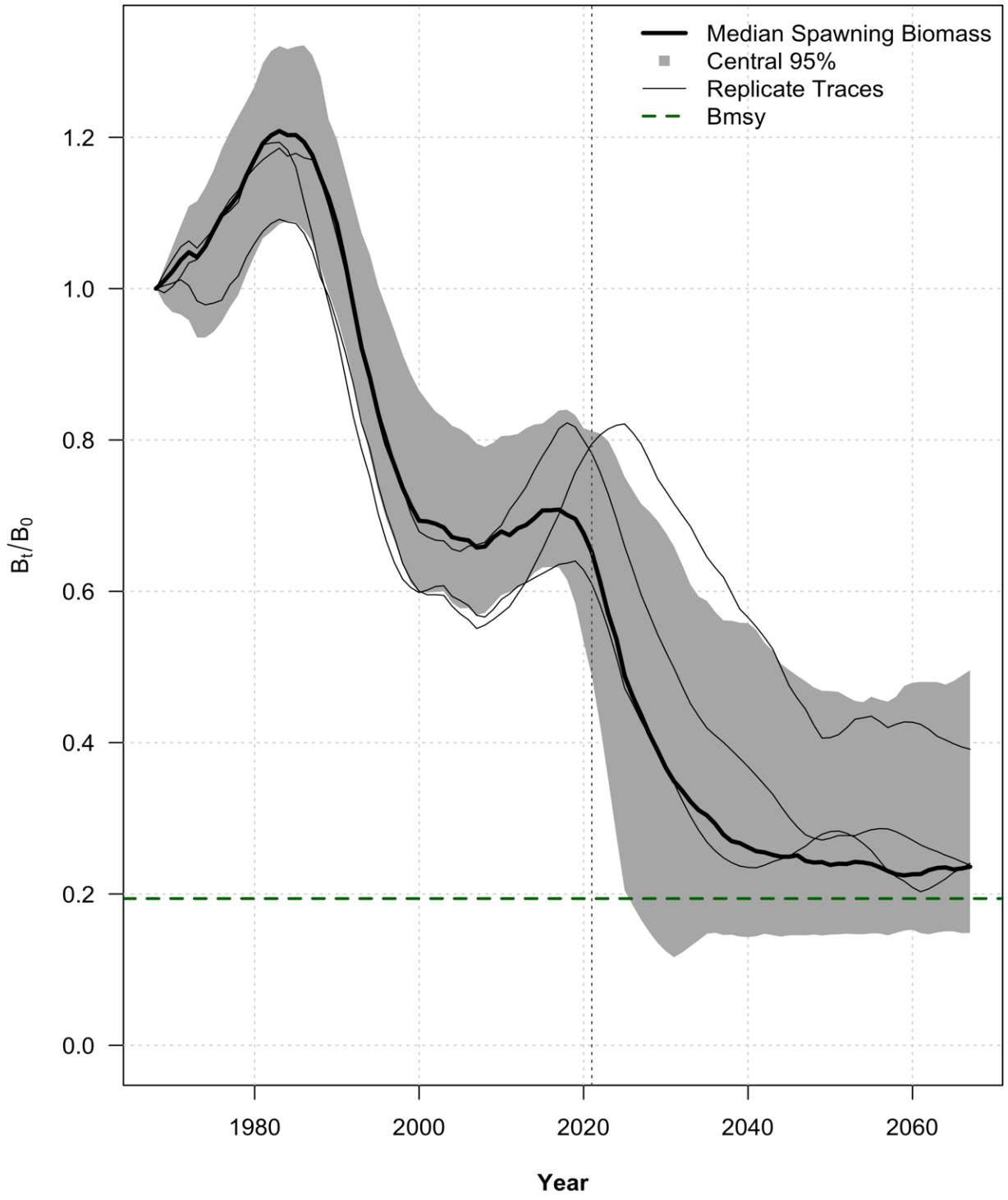


*Figure 20. Simulated GH-0+1 female spawning biomass depletion under the adaptive model/index-based MP and the isPrec survey precision scenario. Grey simulation envelopes show the central 95% (grey region), median (heavy black line), and 3 random simulation replicates (thin black lines) of spawning biomass depletion, the horizontal green line shows spawning biomass producing MSY, and the vertical dashed line indicates the start of the simulated management procedure in 2021.*

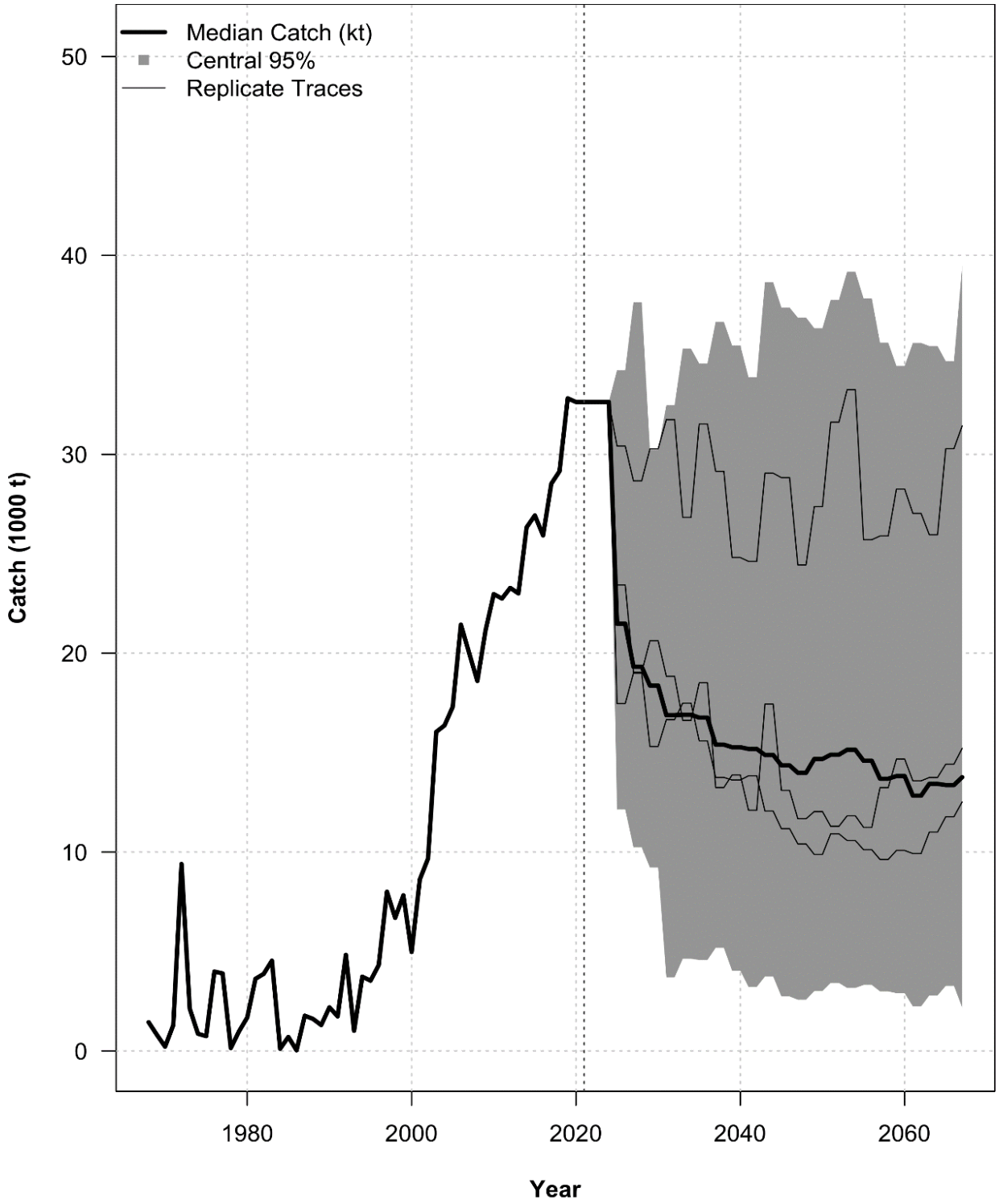


*Figure 21. Simulated total GH-0+1 catches under the adaptive model/index-based management procedure (MP) and the isPrec survey precision scenario. Grey simulation envelopes show the central 95% (grey region), median (heavy black line), and 3 random simulation replicates (thin black lines) of spawning total catch across all fisheries. The beginning of the MP is shown as a vertical dashed line in 2021.*

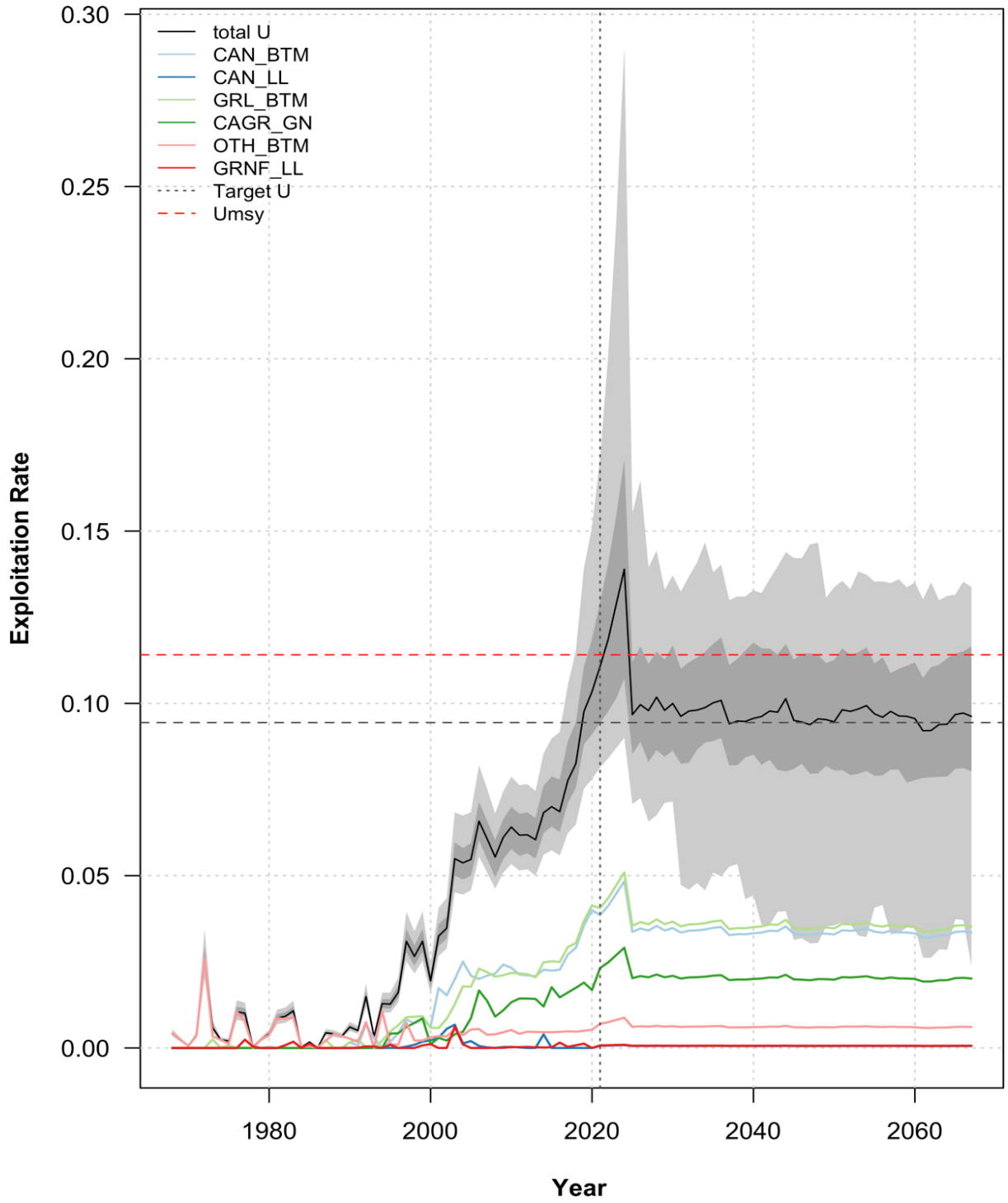


*Figure 22. Simulated GH-0+1 exploitation rates under the adaptive model/index-based management procedure (MP) and the isPrec survey precision scenario. Lines show median total (black) and fleet specific (colours) exploitation rates, with the simulation envelope showing the central 60% (dark grey) and central 95% (light grey) of total exploitation rates over the 50 replicates. Target and optimal exploitation rates are shown as horizontal dashed lines, and the beginning of the MP is shown as a vertical dashed line in 2021.*

A single simulation replicate shows the ability of the MP to capture the GH-0+1 simulated dynamics, and bring biomass towards $B_{MSY}$ via its adaptive nature (Figure 23). For this particular posterior draw, the spawning biomass estimates from SISCAL-GH are slightly positively biased, but the bias appears to decrease over the projection period as more data is generated for model fitting (Figure 23, grey lines). There was also a slight positive bias in exploitable biomass estimates from the index-based procedure early on, partly based on the bias in SISCAL-GH biomass estimates (Figure 23, blue circles). However, that positive bias in the biomass estimates is reduced as updated assessments revise the $rho_t$ values. As a result, spawning biomass is brought almost exactly to its MSY-level (horizontal red dashed line), and although exploitable biomass levels out a little below the MSY level (horizontal blue dashed line), the stock is stable and on average biomass estimates are unbiased.

Adaptation of the model-/index-based procedure can be seen in the trends in exploitable biomass estimates over time (Figure 23). For the chosen replicate, the trends in exploitable biomass and RV survey biomass diverge around 2023, with exploitable biomass decreasing further while RV biomass levels off after a short increase. The diverging biomasses, along with the observation errors in the survey indices, translate into a short period of positively biased exploitable biomass estimates between 2024 and 2028 (Figure 23, blue circles). The bias was eliminated after the next full SISCAL-GH simulated assessment, adapting the ratio $\rho_t$ and survey catchability $q_{RV,t}$ control parameters for the index-based method, bringing the exploitable biomass estimates much closer to the true exploitable biomass value, and long term spawning biomass close to the optimal $B_{MSY}$ level.

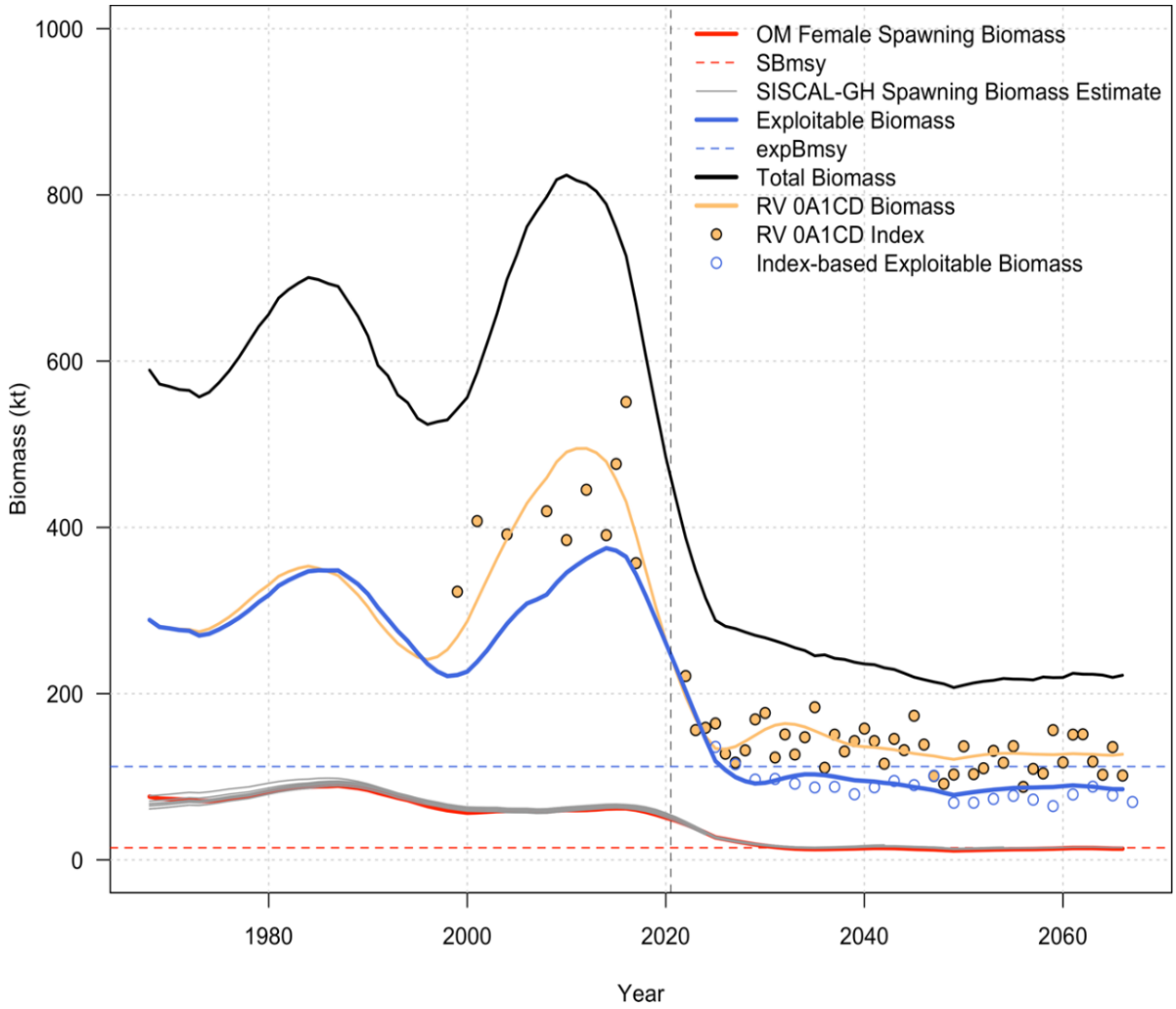


*Figure 23. A single simulation replicate from MS3-GH under the isPrec scenario and adaptive management procedure (MP) showing the true operating model female spawning biomass (red line), exploitable biomass (blue line), total biomass (black line), and RV survey biomass (orange line), as well as simulated RV survey biomass indices (orange points). Also shown are MP estimates of female spawning biomass (grey lines) produced by the simulated SISCAL-GH model update, and estimates of exploitable biomass from the index-based procedure (blue points).*

In contrast, the non-adaptive MP performs poorly relative to the chosen performance metrics. While there is still a very low probability of biomass dropping below the LRP, this is somewhat biased by the health of the stock in early years during the development phase (Figure B.7). The fixed $\rho$ and $q_{RV}$ values do not adapt to changes in the ratio of exploitable and survey biomass, leading to overestimates of biomass. As a result, TACs are too high (Figure B.8), leading to median harvest rates around $U_{MSY}$ (Figure B.9), with a 42% probability of exceeding the limit (Figure B.9, Table 13). As a result, there is a 56% probability of female spawning biomass being below $B_{MSY}$ at the end of the projections (Table 13). Furthermore, the median spawning biomass does not appear to level off, indicating that it may not have equilibrated to the higher exploitation rate by the end of the projection (Figure B.7)

Finally, simulated assessments appear to perform similarly in projections as under the simulation-estimation procedure (Tables C.1 and C.2).

## DISCUSSION

In this paper, we demonstrated an example end-to-end (data-to-advice) assessment modelling framework in three parts:

A. estimating GH-0+1 stock status and biological reference points from fishery and survey data;

B. simulation-testing the assessment model to better understand the range of estimation performance (i.e., bias and precision of estimates) given the available data types, quality, and quantity;

C. conditioning operating models and simulation-testing precautionary harvest strategies.

According to SISCAL-GH estimates, the GH-0+1 stock is likely in a development phase and not overfished, but there is a chance (50% probability) that the recent increases in exploitation rates are overfishing the stock. The SISCAL-GH model presented for Part A showed that the GH-0+1 stock has sufficient data to produce a Bayesian stock assessment, estimating biological reference points and their associated uncertainty. It should be noted that there is considerable uncertainty about the scale of the stock, largely due to low statistical power data such as flat and noisy survey indices, and missing observations of older fish. Sensitivities presented here show a range of unfished female spawning biomass between 30 kt and 250 kt, with total biomass ranging much higher. While absolute scale is very uncertain, all models that fit the data well told a similar story (Appendix A): the GH-0+1 stock is likely to be above $B_{MSY}$, placing it in a relatively healthy state. Moreover, given the 1-way trip dynamics and catch history, commercial exploitation is likely to have been fishing down a standing stock in a fishery development phase. On the other hand, despite little evidence to suggest the stock is overfished, several barriers continue to prevent more certain determination of stock scale, and without scale information it is hard to determine whether current TACs are overfishing GH-0+1, and if so by how much (Appendix A). For example, it is difficult to observe larger fish, making growth estimates uncertain, and the portion of the stock that is in un-surveyed water deeper than 1500 m is not known—all of which suggest that now is an opportune time to develop and apply a simulation-tested feedback harvest strategy approach for the GH-0+1 stock, tested for robustness against the uncertainty in the absolute biomass of the stock. Without such information feedbacks, it will be impossible to detect signs of overfishing and adjust management actions accordingly.

There were some differences between the GH-0+1 model and the state-space model used to manage Greenland Halibut in NAFO Subarea 2 and Divisions 3KLMNO (referred to here as GH-2+3 for short) (Rademeyer and Butterworth 2017, Regular et al. 2017, Regular 2020, Varkey et al. 2020). First, there is a large difference between the number of modeled age-classes for GH-0+1 (plus group for age 35+) and GH-2+3 (plus group for age 10+). There is also a considerable difference in the weight-at-age for those ages that overlap, with age 10+ GH-2+3 fish being about the same weight as fish age 20 and higher in the GH-0+1 model. Finally, the GH-2+3 model is not sex-structured, where SISCAL-GH is a sex-structured model. These differences are all related to the availability and quality of ageing data for each stock. There is substantially more ageing data available for GH-2+3 stock, with annual survey and commercial catch-at-age observations. However, there is considerable uncertainty in the age readings for fish above 10 years old in the GH-2+3 stock (P. Regular, DFO, pers. comm.), and therefore the plus group age is set to that threshold. Given that there is a large portion of the stock aggregated into the plus-group for GH-2+3, there is considerable variation in size-at-age from year-to-year. Therefore, annual empirical weight-at-age is estimated from biological samples, leading to a higher weight-at-age assigned to age 10+ fish for the GH-2+3 stock assessment.

In contrast to the GH-2+3 stock, the GH-0+1 stock has age/length observations for the 2017 offshore survey in Division 0A only. The limited data set means that a parametric growth model is necessary to convert numbers-at-age to stock biomass within the SISCAL-GH model and is also required to provide expected numbers-at-length for fitting to length composition data. While there is a suggestion of ageing error given large variation and overlap in length-at-age distributions for successive age classes above age 10 (Figure 2), there is also a positive trend in median length-at-age in the data. That positive trend is also reflected in estimated von Bertalanffy models that suggest that age observations are not completely unusable above age 10, and that both sexes keep growing beyond the observed data. Therefore, despite differences to the GH-2+3 model, as well as identified limitations in how biology is modeled for GH-0+1—such as the low maturity-at-age and possible bias in length-at-age given dome-shaped selectivity for the survey—we are confident that the plus-group age and growth models for GH-0+1 are appropriate.

It is not clear that there is the same level of ageing uncertainty for GH-0+1 as in the GH-2+3 stock. As mentioned above, there is considerable overlap between length distributions for successive age-classes, which may suggest ageing error (Heifetz et al. 1999, Hanselman et al. 2012). On the other hand, the sample size is very low, with about 175 fish for each sex, which if randomly assigned to age-classes would give an average of 7 individuals per age-class, and less for older ages given survivorship. Therefore, it is difficult to assign variation in length-at-age to low sample size, or a systemic difficulty in reading ages from Greenland Halibut otoliths.

One notable similarity between the GH-2+3 stock assessment and the GH-0+1 stock assessment is the high mortality in the late 1980s and early 1990s. For GH-0+1, we have assigned it to time-varying natural mortality, given that the model is conditioned on catch. For GH-2+3, it is produced as elevated fishing mortality required to fit a large spike in catch from 1990 to 1993. The GH-2+3 State Space Model (SSM) fits the catch reasonably well, but misses some of the points by around 5 kt (Regular et al. 2017, Regular 2020). Coincidentally, the GH-2+3 SSM has a run of positive process errors in years corresponding to the elevated $M$ for GH-0+1, which may be interpreted as fish moving between Subareas 0+1 and 2+3.

Retrospective analysis of the SISCAL-GH model shows that the dynamics are relatively stable under the addition of new data. While the increased TACs over the years have pushed estimates of the absolute size (scale) of the stock up, there has been little variation in depletion estimates among data peels. This stability indicates that while estimates of stock scale have been increasing, there has been stability in estimates of the productivity of the stock (as reflected in several of the sensitivity analyses). On the other hand, that stability is partially due to informative priors on steepness and natural mortality, which compensate for low statistical power (i.e., flat or noisy) indices and length compositions, which are an imperfect substitute for improved age data.

Overall, the SISCAL-GH model shows acceptable bias and precision in simulation self-tests. The SISCAL-GH model base hypothesis was used to simulate alternative data histories (based on the same population dynamics), either simulating stock indices, length composition data, or both. Each simulated data scenario tested SISCAL-GH for its bias and precision in selected biological and observation model parameters, as well as spawning and fishery/survey vulnerable biomass time series. Model estimates of biomass time series, biological parameters, and catchability were most sensitive to whether length composition data was simulated, while stock index residual standard errors were more sensitive to simulated index data. While MRE and MARE values all indicated low bias and moderate precision of parameters, there was moderately persistent negative bias of around -20% in spawning and vulnerable biomass for commercial fisheries. However, these biases are often ultimately irreducible, as there is always uncertainty in the data and the underlying population dynamics, which is something that closed-

loop feedback simulations are designed to address. When fit to data generated by the same life-history parameters but simulating random recruitments for the 1968 to 1988 period, there was more bias but the central 95% of distributions of biomass contained the true values (Appendix B).

We closed the data-to-advice loop by demonstrating a closed-loop feedback simulation framework for testing potential harvest strategies. MS3-GH was conditioned on the estimates under the base SISCAL-GH hypothesis and used to simulation test an adaptive model-/index-based MP also based on the SISCAL-GH stock assessment model. Using the current harvest control rule parameters (e.g., $F_{buf} = 0.8F_{MSY}$ and $B_{buf} = 0.7B_{MSY}$) the MP tends to meet management targets implied by those parameters. Moreover, the MP does a good job of guiding the GH-0+1 fishery through the development phase, fishing the stock down so spawning biomass ends up close to the optimal $B_{MSY}$ level without any large probabilities of overfishing. The speed with which the MP adapts to changes in stock status could be increased by narrowing the time window used to update the full SISCAL-GH model or by updating target fishing mortality rates along with the catchability and biomass adjustment factor parameters. By using a shorter window (e.g., 3-5 years instead of 6), the $\rho_t$ scalar would be more responsive to changes in the ratio of exploitable to RV survey biomass. On the other hand, higher responsiveness in the index-based method would likely come at the cost of increased catch variability due, in part, to increased sensitivity to SISCAL-GH assessment errors to survey observation errors.

The SISCAL-GH stock assessment model performed as expected (i.e., similar to self-tests) in projections as well. Bias and precision was similar to the simulation self-test across newRV survey precision scenarios and projection years (Appendix C). Predictably, the newRV survey parameter estimates improved with additional simulated data.

As mentioned in the results around length composition uncertainties, a modelling framework like the one proposed here can be used for more than just TAC decision making. At the assessment level, there is also the ability to identify data sources that have the largest impact on uncertainty as identified by higher residual standard errors and/or negative log-likelihood function values. Once identified, problematic data sources can be investigated further, and potentially remedied. If there are deeper issues that require significant investment, such as standardisation of data collection processes, then the simulation framework can help to identify the data that will provide the largest return on investment (e.g., more stable TACs through smaller simulated assessment errors when future length data are simulated with, say, lower standard errors).

The end-to-end (data-to-advice) framework presented in this paper can be readily extended to include new data sources and additional complexities. For example, different types of tagging data can be incorporated in the SISCAL-GH model. Tagging data is often integrated into assessments to inform movement among areas (for spatially explicit models), mortality rates (Cox et al. 2016), or selectivity estimates (Cox et al. 2019). There is also an increasing prevalence of genetic mark-recapture methods for conservation and management of fisheries including Greenland Halibut (Carrier et al. 2020), some of which may draw on SISCAL-GH estimates of cohort strength. Additionally, the framework can also be used to estimate the value of information from a new survey, such as a longline survey in deeper waters. Initial exploratory survey sets could determine the rough size structure of the portion of the GH-0+1 stock below 1400 m, and be used to help design a new survey series that could be included in the closed loop simulations for MP testing.

## FUTURE WORK FOR A FULL PEER-REVIEWED MSE

This paper shows one example realisation of a closed-loop data-to-advice framework. As a result, there are several sources of uncertainty that we did not fully explore. A full, peer-reviewed Management Strategy Evaluation process could expand upon this work in several ways, including, but not limited to:

i. Additional MS3-GH scenarios for closed-loop simulations, using the same adaptive MP that fits SISCAL-GH under the base hypothesis for simulated assessments. This would test the MP under a model mis-specification scenario (sometimes called robustness tests). We suggest an operating model that fits to GRNF longline fishery length compositions better, perhaps via a larger $L_\infty$ value and alternative selectivity functional forms.

ii. Additional model sensitivities to (i) selected parameter prior distributions, and (ii) the initial year of time-varying catchability for the CPUE series, as well as any others suggested by reviewers.

iii. Fit to any available data from the newRV survey to estimate selectivity and (in time) catchability, incorporating additional uncertainty in parameter values associated with its short time series of data.

iv. Design a longline survey, perhaps in partnership with industry, with sets below 1400 m to identify the size/age structure and abundance of the unobserved portion of the GH-0+1 stock (Cox et al. 2018).

## ACKNOWLEDGEMENT

The authors wish to thank Margaret Treble and Kevin Hedges from Fisheries and Oceans Canada, and Adriana Nogueira from the Greenland Institute of Natural Resources for providing Greenland halibut data and context from surveys conducted in Northwest Atlantic Fisheries Organization Divisions 0AB and 1CD. We also acknowledge the participants of the Canadian Science Advisory Secretariat peer-review meeting, whose comments substantially improved this research document.

## REFERENCES CITED

Betancourt, M., and Girolami, M. 2015. Hamiltonian Monte Carlo for hierarchical models. *In* Current trends in Bayesian methodology with applications. Edited by S.K. Upadhyay, U. Singh, D.K. Dey, and A. Loganathan. CRC Press, Boca Raton. pp. 80–101.

Bowering, W.R., and Brodie, W.B. 1995. Greenland halibut (*Reinhardtius hippoglossoides*). A review of the dynamics of its distribution and fisheries off eastern Canada and Greenland. *In* Deep-water fisheries of the North Atlantic Oceanic slope. Edited by A.G. Hopper. Springer, Dordrecht. pp. 113–160.

Brodie, W., Shelton, P.A., Couture, E., and Dwyer, K. 2013. A discussion of the NAFO precautionary approach framework. NAFO SCR Doc: 13/024(N6177). 14 p.

Butterworth, D.S. 2007. Why a management procedure approach? Some positives and negatives. ICES J. Mar. Sci. 64(4): 613–617.

Cadigan, N.G. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Can. J. Fish. Aquat. Sci. 73(2): 296–308.

Carrier, E., Ferchaud, A.-L., Normandeau, E., Sirois, P., and Bernatchez, L. 2020. Estimating the contribution of Greenland halibut (*Reinhardtius hippoglossoides*) stocks to nurseries by means of genotyping-by-sequencing: Sex and time matter. Evol. Appl. 13(9): 2155–2167.

Cooke, J.G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. ICES J. Mar. Sci. 56(6): 797–810.

Cooper, D.W., Maslenikov, K.R., and Gunderson, D.R. 2007. Natural mortality rate, annual fecundity, and maturity at length for Greenland halibut (*Reinhardtius hippoglossoides*) from the northeastern Pacific Ocean. Fish. Bull. 105(2): 296–304.

Cox, S., Benson, A., and den Heyer, C.E. 2016. Framework for the assessment of Atlantic halibut stocks on Scotian shelf and Southern Grand Banks. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/001. v + 57 p.

Cox, S., Benson, A., and Doherty, B. 2018. Re-design of the joint industry-DFO Atlantic halibut (*Hippoglossus hippoglossus*) survey off the Scotian shelf and Grand Banks. Can. Sci. Advis. Sec. Res. Doc. 2018/020. v + 50 p.

Cox, S., Holt, K., and Johnson, S. 2019. Evaluation the robustness of management procedures for the sablefish (*Anoplopoma fimbria*) fishery in British Columbia, Canada for 2017-18. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/032. vi + 79 p.

de la Mare, W.K. 1998. Tidier fisheries management requires a new MOP (management-oriented paradigm). Rev. Fish Biol. Fish. 8: 349–356.

DFO. 2006. A Harvest Strategy Compliant with the Precautionary Approach. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2006/023.

Francis, R.C. 2014. Replacing the multinomial in stock assessment models: A first step. Fish. Res. 151: 70–84.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., and Rubin, D.B. 2014. Bayesian data analysis. CRC Press, Boca Raton, U.S.A. 661 p.

Hanselman, D.H., Clark, W.G., Heifetz, J., and Anderl, D.M. 2012. Statistical distribution of age readings of known-age Sablefish (*Anoplopoma fimbria*). Fish. Res. 131: 1–8.

Harris, L.N., Treble, M.A., and Morgan, M.J. 2009. An update of maturity in data for Greenland halibut from trawl surveys of NAFO subarea 0 with emphasis on division 0A. NAFO SCR Doc. 09/025(N5660). 12 p.

Heifetz, J., Anderl, D., Maloney, N., and Rutecki, T.L. 1999. Age validation and analysis of ageing error from marked and recaptured Sablefish, *Anoplopoma fimbria*. Fish. Bull. 97(2): 256–263.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Springer, Dordrecht. 570 p.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B. 2016. TMB: Automatic differentiation and laplace approximation. J. Stat. Softw. 70(5): 1–21.

Monnahan, C.C., and Kristensen, K. 2018. No-u-turn sampling for fast Bayesian inference in ADMB and TMB: Introducing the adnuts and tmbstan R packages. PloS ONE 13(5).

Monnahan, C.C., Thorson, J.T., and Branch, T.A. 2017. Faster estimation of Bayesian models in ecology using hamiltonian monte carlo. Methods Ecol. Evol. 8(3): 339–348.

Morgan, M.J., and Treble, M.A. 2006. A first look at maturity data for Greenland Halibut from Trawl surveys of NAFO Subarea 0. NAFO SCR Doc. 06/05(N5216). 10 p.

Nogueira, A., and Treble, M.A. 2020. Comparison of vessels used and survey timing for the 1CD and 0A-South deep-water surveys and the 1A-F West Greenland shelf surveys. NAFO SCR Doc. 20/015REV3(N7060). 44 p.

Piner, K.R., Lee, H.-H., and Thomas, L.R. 2018. Bias in estimates of growth when selectivity in models includes effects of gear and availability of fish. Fish. Bull. 116(1): 75–80.

Punt, A.E., and Smith, A.D.M. 1999. Harvest strategy evaluation for the eastern stock of gemfish (rexea solandri). ICES J. Mar. Sci. 56(6): 860–875.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rademeyer, R.A., and Butterworth, D.S. 2017. Results for initial candidate management procedure testing for Greenland halibut. NAFO SCR Doc. 17/026REV2(N6678). 41 p.

Regular, P.M. 2020. Update of base case SSM for Greenland halibut in NAFO subarea 2 and divisions 3KLMNO. NAFO SCR Doc. 20/050(N7098). 62 p.

Regular, P.M., Cadigan, N.G., Morgan, M.J., and Healey, B.P. 2017. A simple SAM-style state-space stock assessment model for Greenland halibut in NAFO Subarea 2 and Divisions 3KLMNO. NAFO SCR Doc.17/010(N6659). 35 p.

Sainsbury, K.J., Punt, A.E., and Smith, A.D. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES J. Mar. Sci. 57(3): 731–741.

Schnute, J.T., and Haigh, R. 2007. Compositional analysis of catch curve data, with an application to Sebastes maliger. ICES J. Mar. Sci. 64(2): 218–233.

Smidt, E. 1969. The Greenland halibut *Reinhardtius hippoglossoidis* (walb.), biology and explotation in Greenland waters. Medd. Damn. Fisk-og Havund. N. S. 6: 79–148.

Treble, M.A., and Nogueira, A. 2020. Assessment of the Greenland halibut stock component in NAFO Subarea 0 + 1 (Offshore). NAFO SCR Doc. 20/038(N7086). 31 p.

Varkey, D.A., Regular, P.M., Kunmar, R., Gullage, N., Healey, B., Ings, D.W., Lewis, K., and Dwyer, K. 2020. Review and revamp of the SSM-based Management Strategy Evaluation for Greenland halibut stock in NAFO Subarea 2 and Divisions 3KLMNO. NAFO SCR Doc. 20/047(N7095). 7 p.

Walters, C.J. 1986. Adaptive management of renewable resources. MacMillan Pub. Ltd, Basington, UK. 374 p.

Winker, H., Carvalho, F., Thorson, J.T., Kell, L.T., Parker, D., Kapur, M., Sharma, R., Booth, A.J., and Kerwath, S.E. 2020. JABBA-select: Incorporating life history and fisheries' selectivity into surplus production models. Fish. Res. 222: 1–17.

## APPENDIX A: ADDITIONAL SISCAL-GH SENSITIVITY ANALYSES

During peer review, reviewers requested an analysis of SISCAL-GH sensitivities to the choice of data included for fitting by the stock assessment model, as well as some of the structural assumptions that were made to fit to those data. Specifically, concerns were raised about time-varying parameter assumptions that increased flexibility of the model, and the limits of inferences about stock size and productivity drawn from monitoring data when the statistical power is low, such as with indices, or observations of certain portions of the population are missing due to lack of sampling or gear selectivity, such as larger fish that escape trawl fishing, and the un-surveyed portion of the population that resides in deeper water.

Additional sensitivity runs and simulation tests were defined in response to the reviewers' requests, testing the effects of several model assumptions and data sources on SISCAL-GH estimates. Specifically, the following configuration was suggested for SISCAL-GH:

1. **conM**: constant natural mortality rate for males and females,
2. **noCPUE** removing the BTM_CPUE index for which time-varying catchability was necessary,
3. **domeSel**: dome-shaped selectivity for the GRL_BTM and OTH_BTM trawl fleets in addition to the dome-shaped selectivity already assumed for CAN_BTM.

The original specification and the reviewers' suggested specification for SISCAL were then both subjected to several sensitivity tests:

1. Alternative priors for natural mortality and steepness parameters, i.e., **Mprior**(m,s)$, where $m$ is the mean and $s$ is the standard deviation, and **steepPrior**($\beta_1, \beta_2$), where $\beta_1, \beta_2$ are the parameters for the beta prior on stock-recruit steepness,
2. Lower weightings on the Jeffreys prior for unfished biomass **jeffWtB0_1** and **jeffWtB0_0**,
3. Earlier initial recruitment deviation in 1970 and 1980,
4. later model initialisation in 1988, and
5. **noInshore** testing the exclusion of data from the inshore RV survey that selected for smaller fish.

Additional growth models were also explored, based on two additional sets of age/length observations. The first data set increased the original data from the 2017 0A survey by adding observations of larger fish from the gillnet fishery. The second added 0A survey data from 2014 to the original data from 2017, but excluded the gillnet fishery. The effect of a single-sex model was also emulated by assuming no sexual dimorphism in the growth model for the original 2017 data set, as well as the additional gillnet data.

Detailed results are given in tables and selected figures below, but we summarise the main findings here for readers who do not want to stare at detailed tables.

1. Unfished spawning biomass was not highly sensitive to the suggested changes to the model specification. The new specification had more effect on total biomass, and therefore the ratio of spawning to total biomass, but it was not extreme. Based on further sensitivity analysis, the natural mortality rate is the most significant influence on the total biomass relative to spawning biomass.
2. Unfished spawning biomass was most sensitive to the first year that recruitment process errors were estimated. Earlier recruitment process errors drive unfished spawning biomass higher, but are coupled with a string of negative recruitment deviations in the 1980s that reduce biomass to fit indices, but are unacceptable in a stock assessment context.

3. The 2020 harvest rates range between 1.0 and 1.5 $U_{MSY}$ for both operating model (OM) specifications and across most sensitivities, but jump to above 2.0 for low M estimates (i.e., Mprior(.11,.1)). Overall, it was clear that lower $M$ values produced lower $U_{MSY}$ values and lower recent biomass, and the resulting model estimates of harvest rates indicate overfishing in recent years.
4. Growth was found to affect stock scale and recent harvest rates the most, as the size-at-age of the mostly unobserved older fish in the stock affects the total biomass, which in turn affects exploitable biomass. In some cases there were large changes in optimal harvest $U_{MSY}$ under different growth models, but that was less common.

Overall, the model sensitivities showed that the absolute scale of GH-0+1 biomass is difficult to nail down, but is likely above $B_{MSY}$. Unfished spawning biomass ranged from around 30 kt to around 250 kt at the extremes. The uncertainty in biomass also means that recent harvest rates are very uncertain, making it hard to determine if GH-0+1 is currently subject to overfishing. On the other hand, despite the large range of unfished and current biomass estimates across models, all models in the sensitivity analysis estimated current biomass above the optimal $B_{MSY}$ level, suggesting that while GH-0+1 may or may not be subject to overfishing, GH-0+1 is likely not currently overfished.

*Table A.1. Results of the sensitivity analysis of SISCAL-GH maximum posterior density estimates from the base operating model (OM) and the suggested OM specification initialised in 1968. Models were tested for sensitivity to initial year of recruitment deviations, a Jeffreys prior on unfished biomass, alternative natural mortality and steepness priors, and the inclusion of the inshore SFW survey. Results show MPDEs of unfished biomass ($B_0$), unfished recruitment ($R_0$), time-averaged natural mortality ($M_0$), MSY-based optimal biomass ($B_{MSY}$), optimal harvest rate ($U_{MSY}$), maximum sustainable yield ($MSY$), and mean recruitment deviation.*

| OM | Sensitivity | $B_0$ | $R_0$ | $M_0$ | $h$ | $B_{2020}$ | $D_{2020}$ | $U_{2020}$ | $U_{MSY}$ | $B_{MSY}$ | $MSY$ | meanRecDev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GH-0+1conMnoCPUEdomeSel | basefRecDev1990 | 93.181 | 138.233 | 0.138 | 0.742 | 64.136 | 0.688 | 0.119 | 0.109 | 19.743 | 14.969 | 0.26 |
| GH-0+1conMnoCPUEdomeSel | fRecDev1980 | 227.399 | 342.220 | 0.139 | 0.724 | 104.946 | 0.462 | 0.073 | 0.098 | 50.523 | 36.387 | -0.54 |
| GH-0+1conMnoCPUEdomeSel | jefffWtB00 | 93.181 | 138.233 | 0.138 | 0.742 | 64.136 | 0.688 | 0.119 | 0.109 | 19.743 | 14.969 | 0.26 |
| GH-0+1conMnoCPUEdomeSel | jefffWtB01 | 89.769 | 133.016 | 0.138 | 0.740 | 60.689 | 0.676 | 0.125 | 0.108 | 19.084 | 14.360 | 0.27 |
| GH-0+1conMnoCPUEdomeSel | Mprior(.11,.1) | 106.758 | 26.414 | 0.077 | 0.760 | 50.260 | 0.471 | 0.224 | 0.082 | 24.984 | 7.263 | 0.55 |
| GH-0+1conMnoCPUEdomeSel | Mprior(.14,.1) | 93.397 | 39.817 | 0.094 | 0.755 | 45.416 | 0.486 | 0.202 | 0.086 | 21.191 | 8.381 | 0.46 |
| GH-0+1conMnoCPUEdomeSel | noInshore | 81.748 | 119.628 | 0.138 | 0.739 | 50.029 | 0.612 | 0.152 | 0.108 | 17.407 | 12.899 | 0.25 |
| GH-0+1conMnoCPUEdomeSel | steepPrior(20,10) | 93.488 | 138.711 | 0.138 | 0.674 | 64.285 | 0.688 | 0.119 | 0.092 | 22.752 | 13.221 | 0.26 |
| GH-0+1conMnoCPUEdomeSel | steepPrior(40,20) | 93.508 | 138.742 | 0.138 | 0.670 | 64.296 | 0.688 | 0.119 | 0.091 | 22.923 | 13.125 | 0.26 |
| GH-0+1tvMqCPUE | basefRecDev1990 | 95.210 | 141.587 | 0.125 | 0.778 | 61.884 | 0.650 | 0.117 | 0.116 | 18.491 | 16.420 | 0.21 |
| GH-0+1tvMqCPUE | fRecDev1970 | 157.275 | 240.308 | 0.130 | 0.751 | 79.629 | 0.506 | 0.091 | 0.103 | 32.770 | 26.688 | -0.28 |
| GH-0+1tvMqCPUE | fRecDev1980 | 153.177 | 233.073 | 0.131 | 0.751 | 79.274 | 0.518 | 0.092 | 0.103 | 31.955 | 25.888 | -0.35 |
| GH-0+1tvMqCPUE | jefffWtB00 | 155.385 | 235.420 | 0.081 | 0.794 | 115.864 | 0.746 | 0.066 | 0.117 | 29.024 | 28.124 | 0.33 |
| GH-0+1tvMqCPUE | jefffWtB01 | 139.442 | 209.704 | 0.091 | 0.792 | 103.493 | 0.742 | 0.073 | 0.117 | 26.194 | 24.962 | 0.30 |
| GH-0+1tvMqCPUE | Mprior(.11,.1) | 90.595 | 42.281 | 0.081 | 0.784 | 47.961 | 0.529 | 0.161 | 0.092 | 19.269 | 9.267 | 0.46 |
| GH-0+1tvMqCPUE | Mprior(.14,.1) | 91.732 | 57.474 | 0.092 | 0.779 | 51.700 | 0.564 | 0.147 | 0.096 | 19.224 | 10.620 | 0.38 |
| GH-0+1tvMqCPUE | noInshore | 88.823 | 130.040 | 0.125 | 0.771 | 53.881 | 0.607 | 0.132 | 0.114 | 17.583 | 14.885 | 0.17 |
| GH-0+1tvMqCPUE | steepPrior(20,10) | 93.434 | 138.703 | 0.126 | 0.731 | 61.852 | 0.662 | 0.116 | 0.103 | 20.273 | 14.746 | 0.25 |
| GH-0+1tvMqCPUE | steepPrior(40,20) | 95.259 | 141.655 | 0.125 | 0.704 | 61.900 | 0.650 | 0.117 | 0.096 | 21.877 | 14.315 | 0.21 |

*Table A.2. Results of the sensitivity analysis of SISCAL-GH maximum posterior density estimates from the base operating model (OM) and the suggested OM specification initialised in 1988. Models were tested for sensitivity to initial year of recruitment deviations, a Jeffreys prior on unfished biomass, alternative natural mortality and steepness priors, and the inclusion of the inshore SFW survey. Results show MPDEs of unfished biomass ($B_0$), unfished recruitment ($R_0$), time-averaged natural mortality ($M_0$), MSY-based optimal biomass ($B_{MSY}$), optimal harvest rate ($U_{MSY}$), maximum sustainable yield ($MSY$), and mean recruitment deviation.*

| OM | Sensitivity | $B_0$ | $R_0$ | $M_0$ | $h$ | $B_{2020}$ | $D_{2020}$ | $U_{2020}$ | $U_{MSY}$ | $B_{MSY}$ | $MSY$ | meanRecDev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GH-0+1conMnoCPUEdomeSel1988 | basefRecDev1990 | 87.560 | 131.203 | 0.139 | 0.753 | 59.645 | 0.681 | 0.127 | 0.113 | 18.046 | 14.473 | 0.28 |
| GH-0+1conMnoCPUEdomeSel1988 | jefffWtB00 | 87.560 | 131.203 | 0.139 | 0.753 | 59.645 | 0.681 | 0.127 | 0.113 | 18.046 | 14.473 | 0.28 |
| GH-0+1conMnoCPUEdomeSel1988 | jefffWtB01 | 85.941 | 128.771 | 0.139 | 0.754 | 57.598 | 0.670 | 0.130 | 0.113 | 17.698 | 14.209 | 0.28 |
| GH-0+1conMnoCPUEdomeSel1988 | Mprior(.11,.1) | 106.226 | 27.865 | 0.078 | 0.767 | 54.301 | 0.511 | 0.210 | 0.084 | 24.454 | 7.544 | 0.54 |
| GH-0+1conMnoCPUEdomeSel1988 | Mprior(.14,.1) | 95.802 | 39.701 | 0.093 | 0.763 | 49.574 | 0.517 | 0.196 | 0.089 | 21.445 | 8.573 | 0.46 |
| GH-0+1conMnoCPUEdomeSel1988 | noInshore | 79.093 | 116.129 | 0.138 | 0.751 | 47.439 | 0.600 | 0.159 | 0.111 | 16.403 | 12.806 | 0.24 |
| GH-0+1conMnoCPUEdomeSel1988 | steepPrior(20,10) | 87.916 | 131.761 | 0.139 | 0.695 | 60.053 | 0.683 | 0.126 | 0.097 | 20.551 | 13.025 | 0.28 |
| GH-0+1conMnoCPUEdomeSel1988 | steepPrior(40,20) | 88.012 | 131.912 | 0.139 | 0.681 | 60.166 | 0.684 | 0.126 | 0.094 | 21.139 | 12.695 | 0.28 |
| GH-0+1tvMqCPUE1988 | basefRecDev1990 | 95.748 | 141.545 | 0.204 | 0.754 | 61.451 | 0.642 | 0.118 | 0.110 | 19.740 | 15.734 | 0.02 |
| GH-0+1tvMqCPUE1988 | jefffWtB00 | 112.590 | 166.344 | 0.206 | 0.751 | 81.644 | 0.725 | 0.092 | 0.109 | 23.406 | 18.402 | 0.00 |
| GH-0+1tvMqCPUE1988 | jefffWtB01 | 110.237 | 162.831 | 0.205 | 0.751 | 78.841 | 0.715 | 0.095 | 0.109 | 22.903 | 18.018 | 0.01 |
| GH-0+1tvMqCPUE1988 | Mprior(.11,.1) | 91.140 | 41.937 | 0.160 | 0.776 | 49.406 | 0.542 | 0.158 | 0.089 | 19.785 | 9.189 | 0.36 |
| GH-0+1tvMqCPUE1988 | Mprior(.14,.1) | 91.534 | 53.188 | 0.168 | 0.770 | 50.436 | 0.551 | 0.152 | 0.093 | 19.679 | 10.093 | 0.28 |
| GH-0+1tvMqCPUE1988 | noInshore | 90.545 | 132.107 | 0.211 | 0.751 | 54.293 | 0.600 | 0.132 | 0.109 | 18.825 | 14.615 | 0.00 |
| GH-0+1tvMqCPUE1988 | steepPrior(20,10) | 95.574 | 141.203 | 0.203 | 0.695 | 61.731 | 0.646 | 0.117 | 0.095 | 22.361 | 14.060 | 0.05 |
| GH-0+1tvMqCPUE1988 | steepPrior(40,20) | 95.638 | 141.306 | 0.203 | 0.681 | 61.728 | 0.645 | 0.117 | 0.091 | 22.993 | 13.697 | 0.05 |

*Table A.3. Results of the sensitivity analysis of SISCAL-GH maximum posterior density estimates from the base operating model (OM) and the suggested OM specification initialised in 1968 and usingalternative growth models including 2014 data, gillnet data, and assuming no sexual dimorphism. Models were tested for sensitivity to initial year of recruitment deviations, a Jeffreys prior on unfished biomass, alternative natural mortality and steepness priors, and the inclusion of the inshore SFW survey. Results show MPDEs of unfished biomass ($B_0$), unfished recruitment ($R_0$), time-averaged natural mortality ($M_0$), MSY-based optimal biomass ($B_{MSY}$), optimal harvest rate ($U_{MSY}$), maximum sustainable yield ($MSY$), and mean recruitment deviation.*

| OM | Sensitivity | $B_0$ | $R_0$ | $M_0$ | $h$ | $B_{2020}$ | $D_{2020}$ | $U_{2020}$ | $U_{MSY}$ | $B_{MSY}$ | $MSY$ | meanRecDev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GH-0+1conMnoCPUEdomeSelMprior.11 | identGrowth2017 | 41.036 | 24.601 | 0.093 | 0.763 | 16.151 | 0.394 | 0.319 | 0.094 | 9.851 | 6.839 | 0.51 |
| GH-0+1conMnoCPUEdomeSelMprior.11 | identGrowthGNdata | 29.836 | 26.942 | 0.116 | 0.766 | 8.871 | 0.297 | 0.504 | 0.133 | 6.873 | 7.484 | 0.41 |
| GH-0+1conMnoCPUEdomeSelMprior.11 | sexDim0A201417 | 104.111 | 22.969 | 0.086 | 0.790 | 26.412 | 0.254 | 0.619 | 0.107 | 22.662 | 5.549 | 0.67 |
| GH-0+1conMnoCPUEdomeSelMprior.11 | sexDim2017 | 109.665 | 24.907 | 0.074 | 0.759 | 54.249 | 0.495 | 0.240 | 0.086 | 26.459 | 7.158 | 0.57 |
| GH-0+1conMnoCPUEdomeSelMprior.11 | sexDimGNdata | 61.671 | 24.573 | 0.093 | 0.763 | 16.642 | 0.270 | 0.457 | 0.111 | 13.629 | 6.899 | 0.41 |
| GH-0+1tvMqCPUE | identGrowth2017 | 54.027 | 123.212 | 0.128 | 0.768 | 34.098 | 0.631 | 0.143 | 0.124 | 11.024 | 13.999 | 0.22 |
| GH-0+1tvMqCPUE | identGrowthGNdata | 36.697 | 55.891 | 0.129 | 0.763 | 19.944 | 0.543 | 0.249 | 0.160 | 8.085 | 11.414 | 0.20 |
| GH-0+1tvMqCPUE | sexDim0A201417 | 76.715 | 72.624 | 0.123 | 0.776 | 32.536 | 0.424 | 0.231 | 0.115 | 15.628 | 9.992 | 0.26 |
| GH-0+1tvMqCPUE | sexDim2017 | 95.210 | 141.587 | 0.125 | 0.778 | 61.884 | 0.650 | 0.117 | 0.116 | 18.491 | 16.420 | 0.21 |
| GH-0+1tvMqCPUE | sexDimGNdata | 62.586 | 70.404 | 0.126 | 0.762 | 37.072 | 0.592 | 0.192 | 0.153 | 12.776 | 12.764 | 0.21 |

*Table A.4. Results of the sensitivity analysis of SISCAL-GH maximum posterior density estimates from the base operating model (OM) and the suggested OM specification initialised in 1988 and using alternative growth models including 2014 data, gillnet data, and assuming no sexual dimorphism. Models were tested for sensitivity to initial year of recruitment deviations, a Jeffreys prior on unfished biomass, alternative natural mortality and steepness priors, and the inclusion of the inshore SFW survey. Results show MPDEs of unfished biomass ($B_0$), unfished recruitment ($R_0$), time-averaged natural mortality ($M_0$), MSY-based optimal biomass ($B_{MSY}$), optimal harvest rate ($U_{MSY}$), maximum sustainable yield ($MSY$), and mean recruitment deviation.*

| OM | Sensitivity | $B_0$ | $R_0$ | $M_0$ | $h$ | $B_{2020}$ | $D_{2020}$ | $U_{2020}$ | $U_{MSY}$ | $B_{MSY}$ | $MSY$ | meanRecDev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GH-0+1conMnoCPUEdomeSelMprior.11 | identGrowth2017 | 45.098 | 27.020 | 0.093 | 0.768 | 19.553 | 0.434 | 0.283 | 0.095 | 10.639 | 7.184 | 0.49 |
| GH-0+1conMnoCPUEdomeSelMprior.11 | identGrowthGNdata | 29.661 | 28.608 | 0.119 | 0.769 | 9.884 | 0.333 | 0.462 | 0.137 | 6.758 | 7.708 | 0.40 |
| GH-0+1conMnoCPUEdomeSelMprior.11 | sexDim2017 | 106.226 | 27.865 | 0.078 | 0.767 | 54.301 | 0.511 | 0.210 | 0.084 | 24.454 | 7.544 | 0.54 |
| GH-0+1conMnoCPUEdomeSelMprior.11 | sexDimGNdata | 60.703 | 26.592 | 0.097 | 0.766 | 19.285 | 0.318 | 0.405 | 0.116 | 13.205 | 7.207 | 0.41 |
| GH-0+1tvMqCPUE | identGrowth2017 | 57.917 | 131.542 | 0.204 | 0.753 | 37.334 | 0.645 | 0.133 | 0.119 | 12.272 | 14.579 | 0.05 |
| GH-0+1tvMqCPUE | identGrowthGNdata | 39.436 | 59.994 | 0.212 | 0.756 | 23.162 | 0.587 | 0.220 | 0.154 | 8.820 | 12.108 | 0.12 |
| GH-0+1tvMqCPUE | sexDim2017 | 95.748 | 141.545 | 0.204 | 0.754 | 61.451 | 0.642 | 0.118 | 0.110 | 19.740 | 15.734 | 0.02 |
| GH-0+1tvMqCPUE | sexDimGNdata | 63.170 | 70.913 | 0.203 | 0.752 | 36.420 | 0.577 | 0.194 | 0.149 | 13.197 | 12.657 | 0.09 |

## APPENDIX B: ADDITIONAL SIMULATION TESTS OF SISCAL-GH

Simulation self-tests were re-run with stochastic recruitment simulated randomly off the stock-recruit curve for 1968 - 1988. Three data scenarios were simulated, as in the main body of this paper, where either only indices were simulated (simIdx), only length compositions were simulated (simLengths), or both length and index data series were simulated (simAll). Simulations were completed over 50 Monte-Carlo replicates, due to time constraints, and the simulated SISCAL-GH stock assessment still assumed deterministic recruitment in the 1968-1988 period.

Overall, SISCAL-GH estimates were most sensitive to simulated length composition data. Under the simLengths and simAll scenarios, there was increased variance and bias in unfished biomass and natural mortality parameters, but similar levels of bias and variance for survey catchability and index standard errors.

Biomass time series were also more variable when early recruitment process errors were simulated and tended to be more biased at the beginning and end of the series. Some of the error in model estimates could be attributed to the slow growing nature of Greenland Halibut in the GH-0+1 stock, where changes in recruitment took several years to work into the exploitable and spawning biomass, delaying the bias caused by random recruitment.

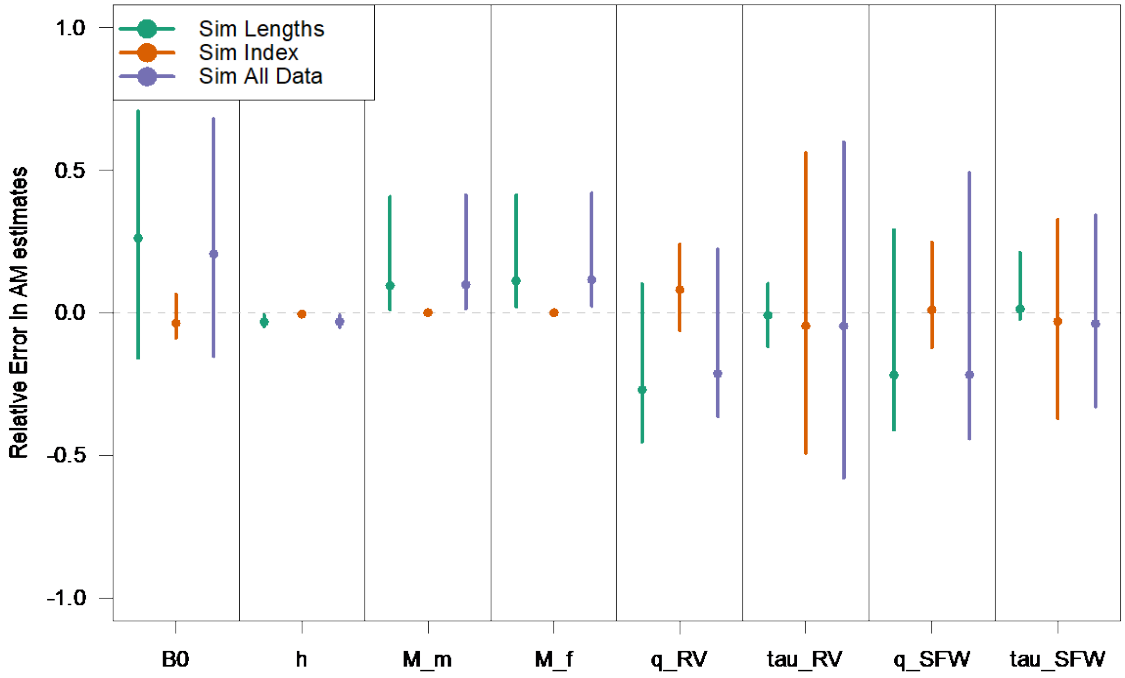


*Figure B.1. Relative errors in leading parameter estimates in the simulation-evaluation self test of SISCAL-GH when recruitments were randomly simulated over the period 1968 - 1988.*

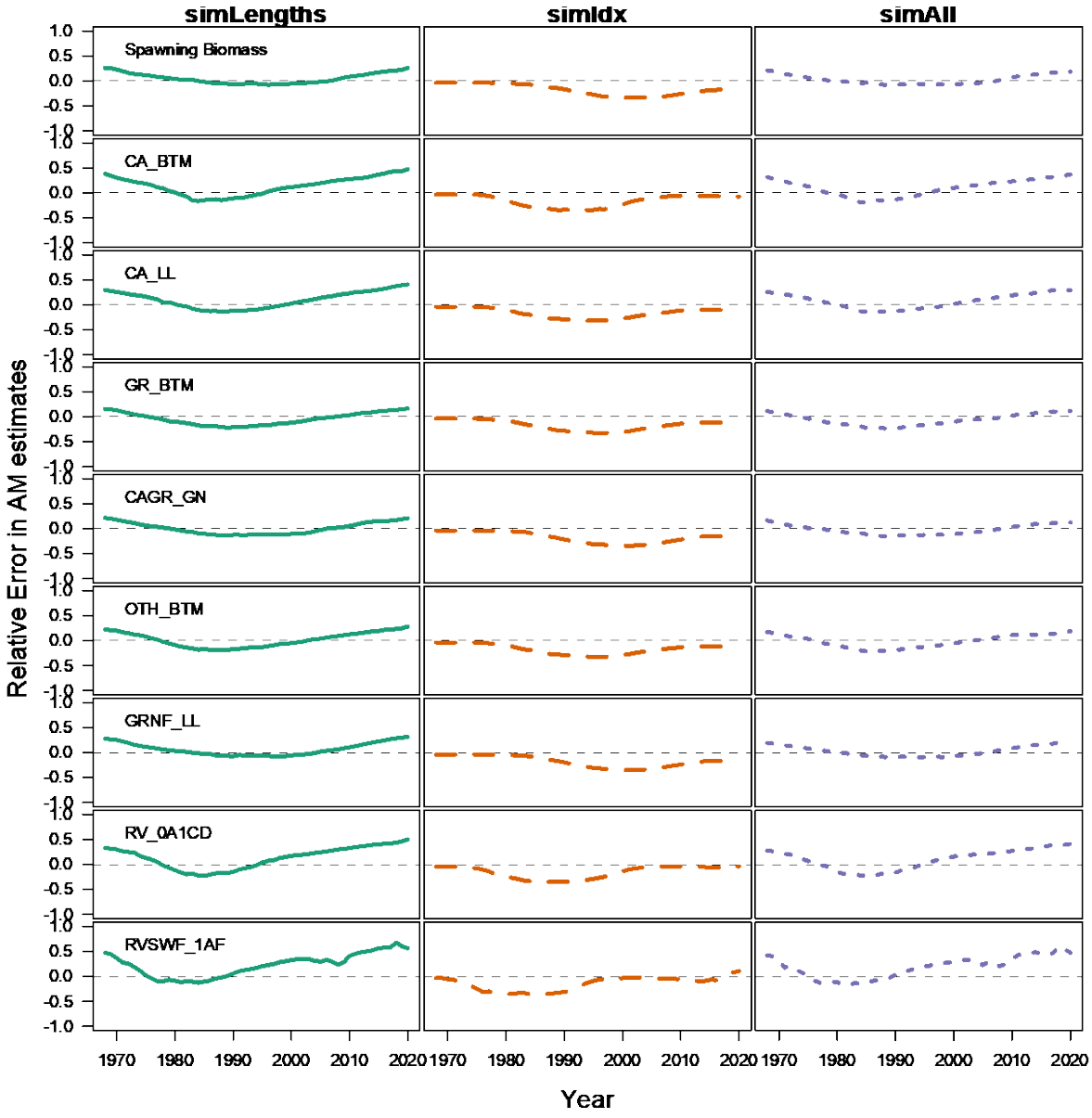


*Figure B.2. Relative errors in biomass time series estimates in the simulation-evaluation self test of SISCAL-GH when recruitments were randomly simulated over the years 1968 - 1988.*

## APPENDIX C: ALTERNATIVE SIMULATED SURVEY CONFIGURATION

An additional simulated projection was presented during peer review that reflects another possibility for the future survey configuration. Under the alternative configuration, the 0A1CD survey series stops after the *Paamiut* was retired in 2017, and the newRV survey begins in 2022 (Figure C.1). The new survey is simulated with a different dome-shaped selectivity that increases sooner than the 0A1CD survey to catch more small fish, representing additional inshore sets similar to the simulation presented in the main body, but has a similar descending limb to the 0A1CD survey selectivity function (Figure C.2). The adjusted newRV index was used by both the adaptive and non-adaptive version of the index-based MP in closed loop simulations to set TACs over 50 simulation replicates, due to time constraints. For both non-adaptive and adaptive MPs, the SISCAL-GH model is fit to the newRV survey data 4 years into the projection to estimate a survey index scalar used to set TACs as in the method presented above. After the first model fit, the non-adaptive MP holds the scalar constant for the remainder of the projection, while the adaptive MP fits SISCAL-GH every 6 years to re-adjust.

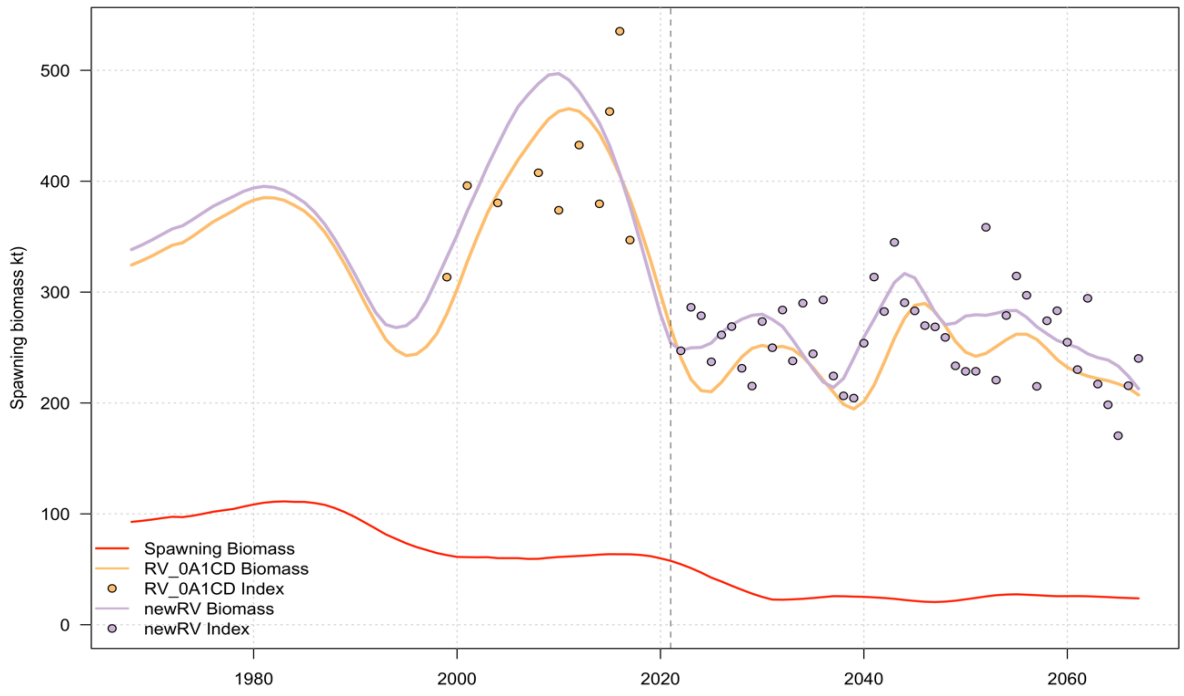


*Figure C.1. GH-0+1 survey biomass (lines) and simulated indices (points) for the 0A1CD survey (yellow) and the newRV survey (purple) under the alternative survey configuration. The vertical dashed line shows the beginning of the simulated projection period.*

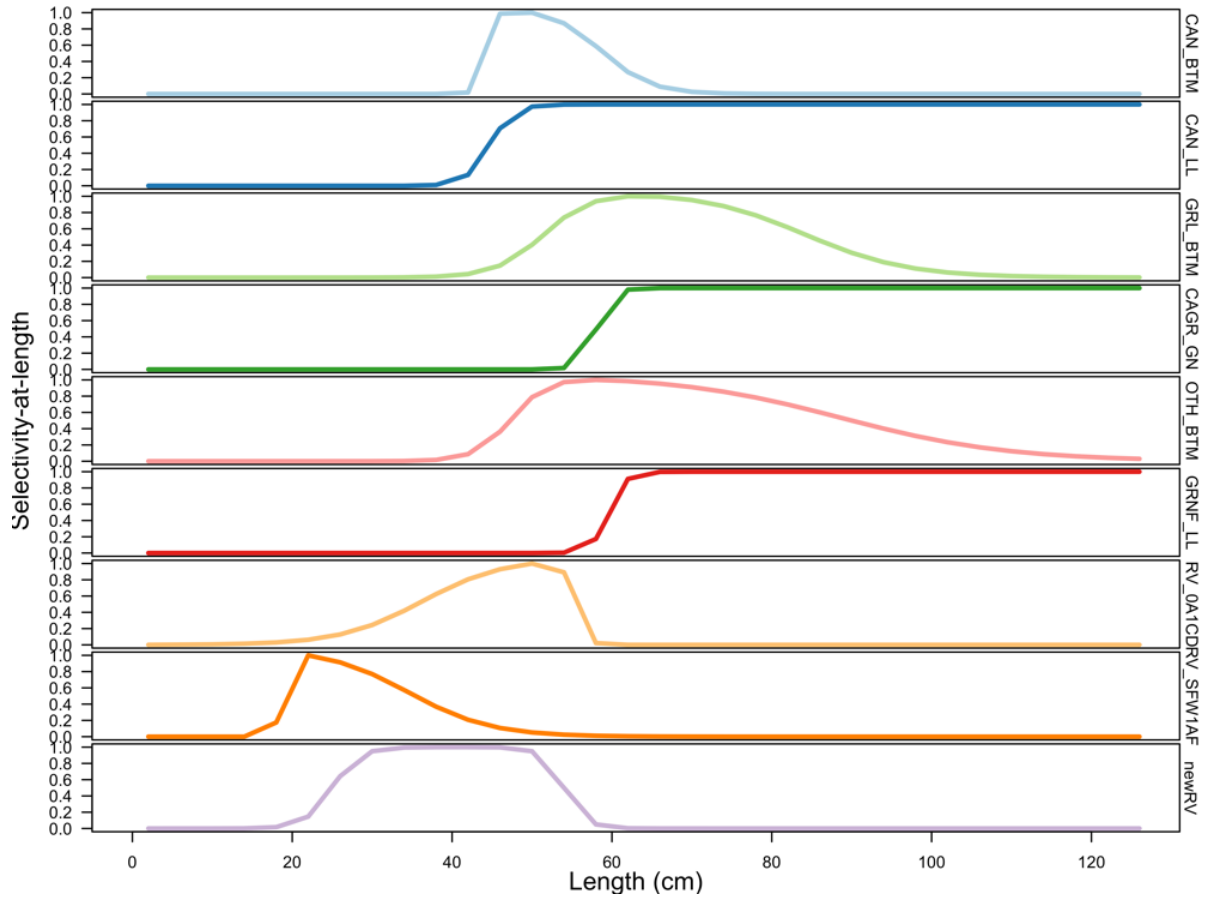


*Figure C.2. GH-0+1 selectivity-at-length for each fleet, highlighting the different newRV survey selectivity (bottom row).*

In the short term (until the adaptive procedure adjusts the scaling factor) both MPs appear to perform similarly, only diverging shortly after the first assessment. In the replicate shown, the adaptive procedure starts out with a close match between survey biomass and exploitable biomass, causing the index based method to follow the survey biomass as a large recruitment enters the fishery (Figure C.3, blue circles). As a result, after the initial spike in harvest rates (caused by the TACs held constant before the first assessment), harvest rates tend to have a saw-tooth pattern between assessments (Figure C.4). They stay closer to the target for a year or two following the assessment, and then begin to drift as new year classes recruit to the exploitable biomass. In the near term, given the closeness of survey and exploitable biomass at in the assessment year, but quick increase of survey biomass right afterwards (Figure C.3), the harvest rates spike until the next assessment is performed and the scalar is adapted (Figure C.4). The adaptation snaps the biomass estimates back towards exploitable biomass (Figure C.3), and the effective harvest rates are reduced (Figure C.4). For the remainder of the projection, the inter-quartile range of harvest rates tends to stay lower, varying between the limit and target harvest rates (Figure C.4), which could be improved to meet NAFO PA policy requirements via tuning of the target harvest rate or other input variables.

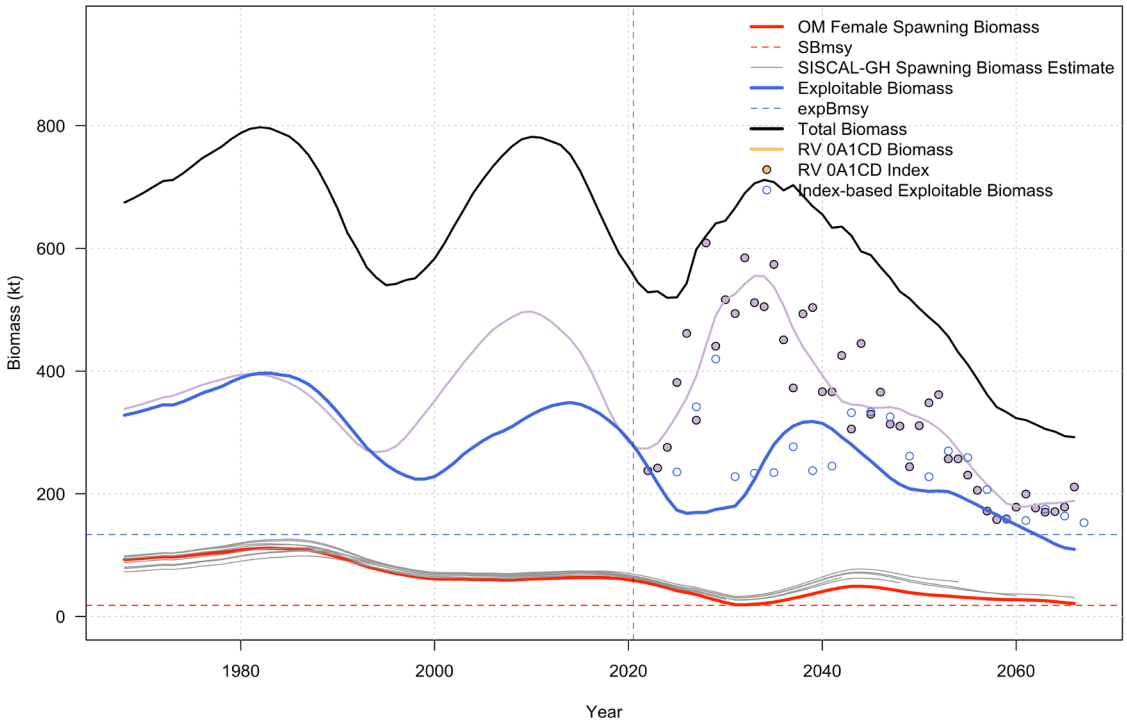


*Figure C.3. A single simulation replicate from MS3-GH under the isPrec scenario and adaptive management procedure (MP) applied to the alternative survey configuration where newRV survey indices are used to set TACS. Model states show true operating model female spawning biomass (red line), exploitable biomass (blue line), total biomass (black line), and research vessel "new RV" survey biomass (purple line), as well as simulated research vessel survey biomass indices (purple filled points). Also shown are MP estimates of female spawning biomass (grey lines) produced by the simulated SISCAL-GH model update and estimates of exploitable biomass from the index-based procedure (blue points).*

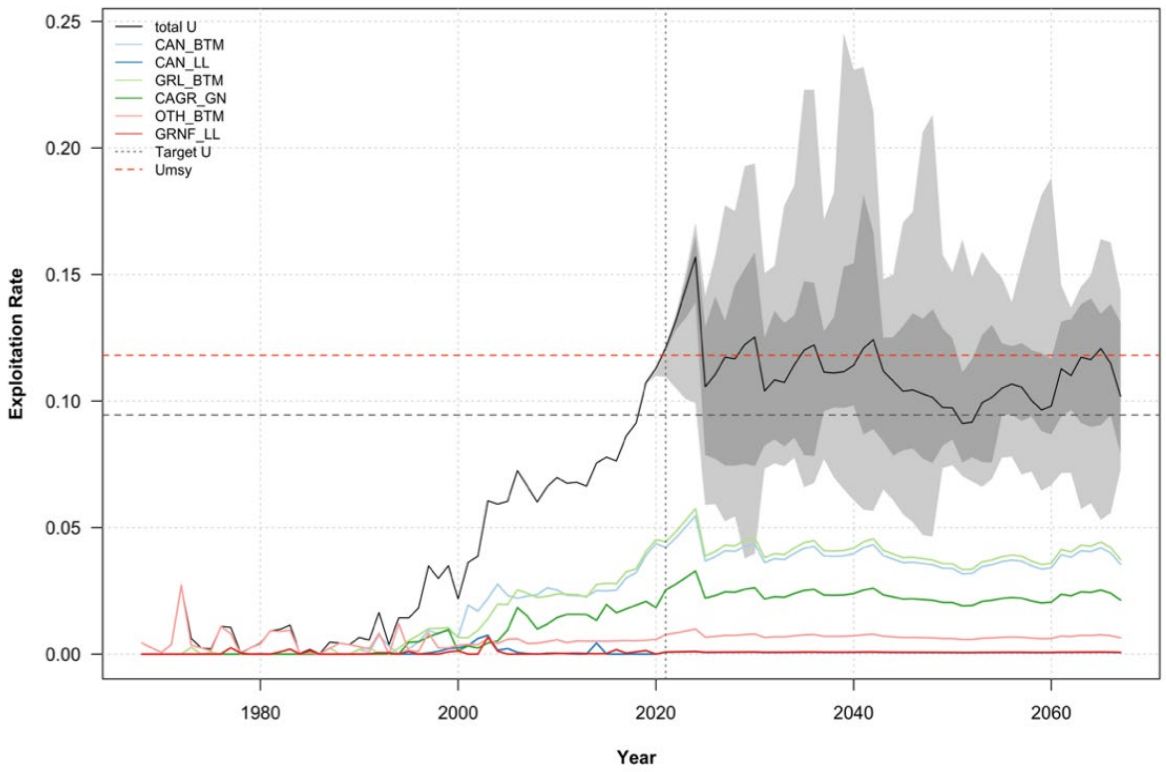


*Figure C.4. Simulated exploitation rates under the adaptive model/index-based management procedure (MP) applied to the newRV survey biomass index, simulated under the isPrec survey precision scenario. Lines show median total (black) and fleet specific (colours) exploitation rates, with the simulation envelope showing the central 60% (dark grey) and central 95% (light grey) of total exploitation rates over the 50 replicates. Target (Target U) and optimal (Umsy) exploitation rates are shown as horizontal dashed lines, and the beginning of the MP is shown as a vertical dashed line in 2021.*

For the non-adaptive procedure, the first 10 years is very similar to the adaptive procedure. The first assessment creates the same adaptative response, leading to similar exploitable and survey biomass estimates (Figure C.5), and therefore similar harvest rates during that time (Figure C.6). After the first 10 years, the non-adapative procedure begins to diverge from the adaptive procedure results, as biomass estimates are not modulated by new 'ground truth' assessments. Therefore, as exploitable biomass declines but survey biomass remains high given a large number of small fish, the TACs remain high, translating into effective harvest rates that climb above the limit with greater than 50% probability (Figure C.6).

The results here highlight the importance of an adaptive procedure. The newRV survey as specified is more sensitive to recruitments, given the higher selectivity for small fish, so the exploitable biomass and survey biomass tend to diverge more often and by larger amounts. Without adaptation, the harvest rates climb with the survey biomass and overfish the stock.

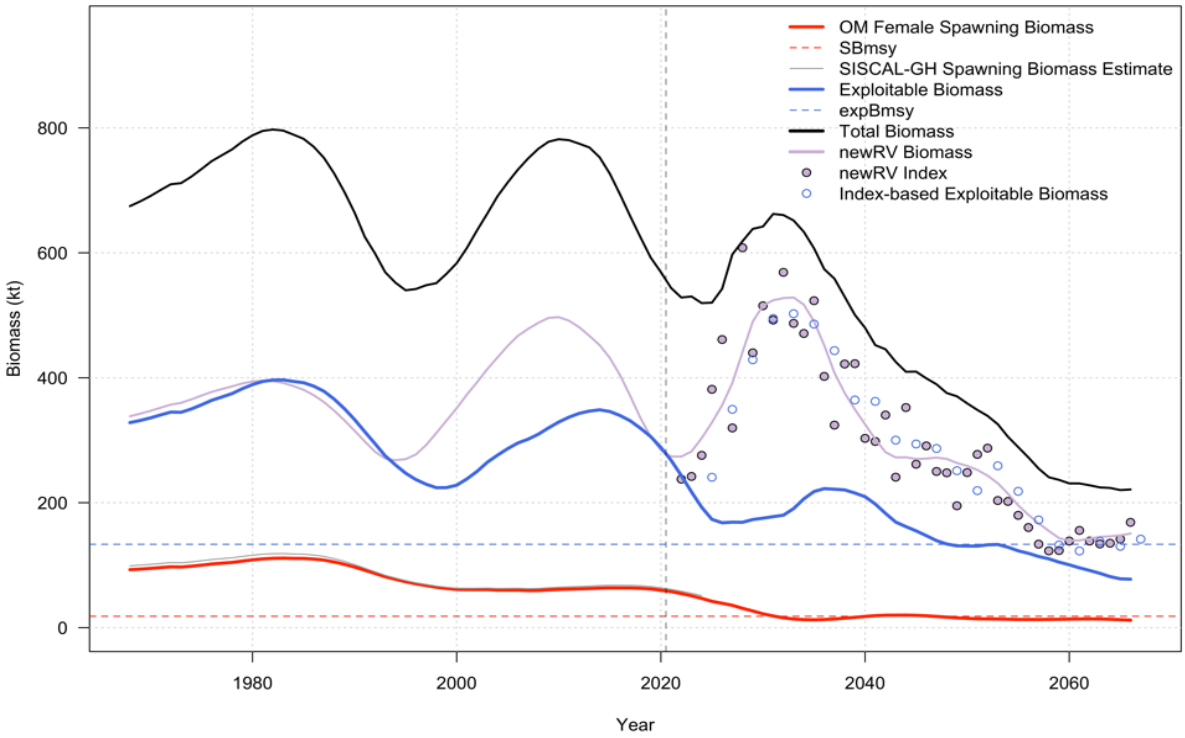


*Figure C.5. A single simulation replicate from MS3-GH under the isPrec scenario and adaptive management procedure (MP) applied to the alternative survey configuration where newRV survey indices are used to set TACS. Model states show true operating model female spawning biomass (red line), exploitable biomass (blue line), total biomass (black line), and research vessel 'newRV' survey biomass (purple line), as well as simulated research vessel 'new RV' survey biomass indices (purple points). Also shown are MP estimates of female spawning biomass (grey lines) produced by the simulated SISCAL-GH model update and estimates of exploitable biomass from the index-based procedure (blue points).*

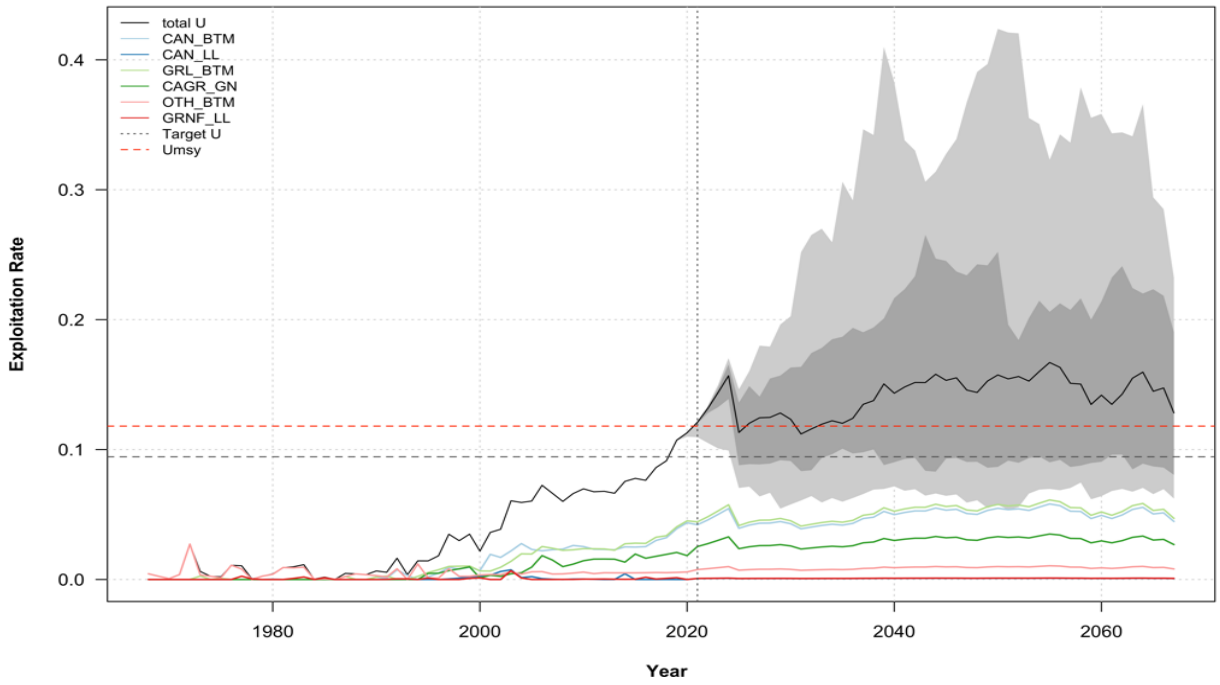


*Figure C.6. Simulated exploitation rates under the adaptive model/index-based management procedure (MP) applied to the newRV survey biomass index, simulated under the isPrec survey precision scenario. Lines show median total (black) and fleet specific (colours) exploitation rates, with the simulation envelope showing the central 60% (dark grey) and central 95% (light grey) of total exploitation rates over the 50 replicates. Target (Target U) and optimal (Umsy) exploitation rates are shown as horizontal dashed lines, and the beginning of the MP is shown as a vertical dashed line in 2021.*

## APPENDIX D: ANNUAL FITS TO LENGTH COMPOSITION DATA

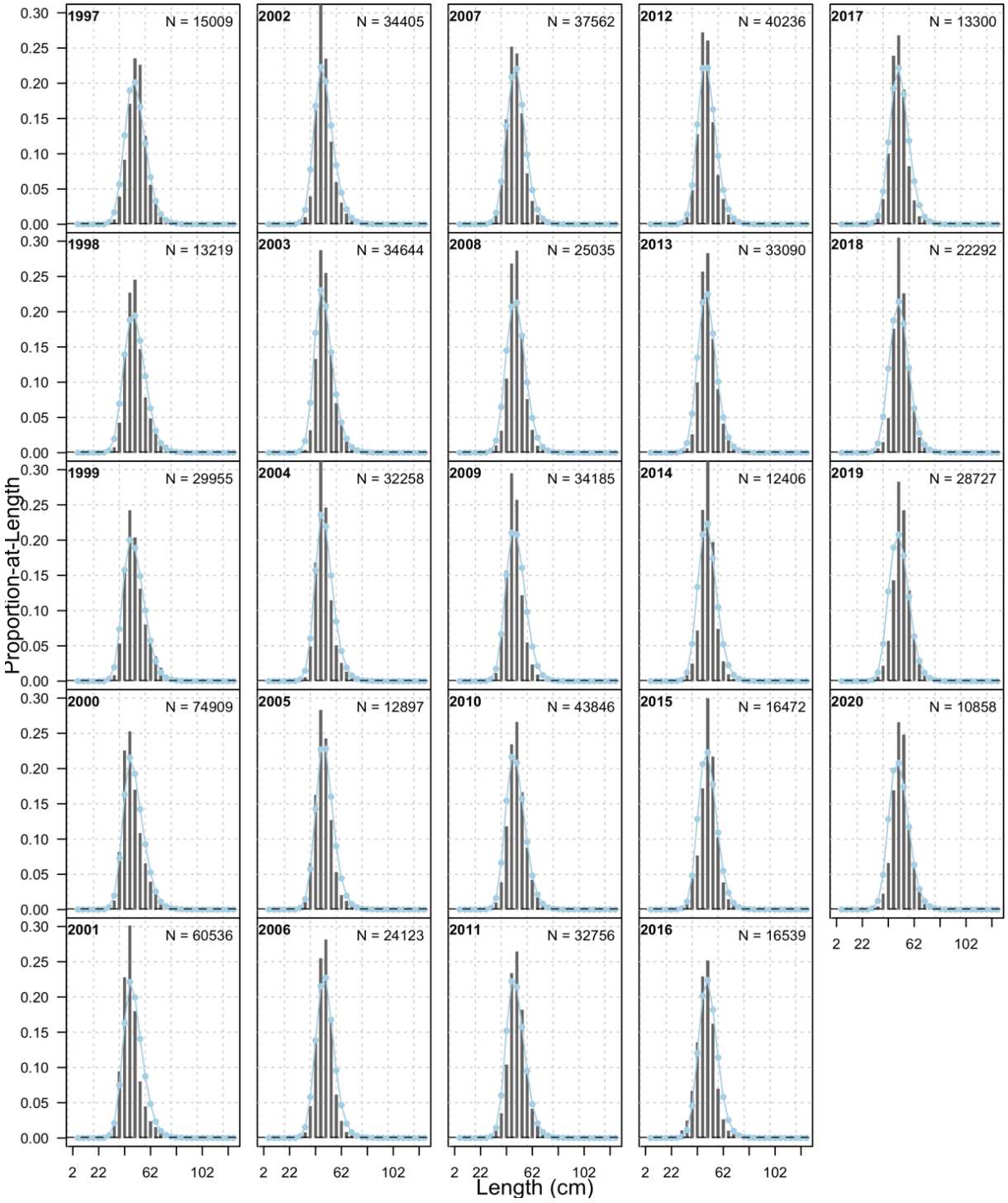


*Figure D.1. Canadian trawl fishery annual estimates (light blue lines and points) of GH-0+1 combined sex length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

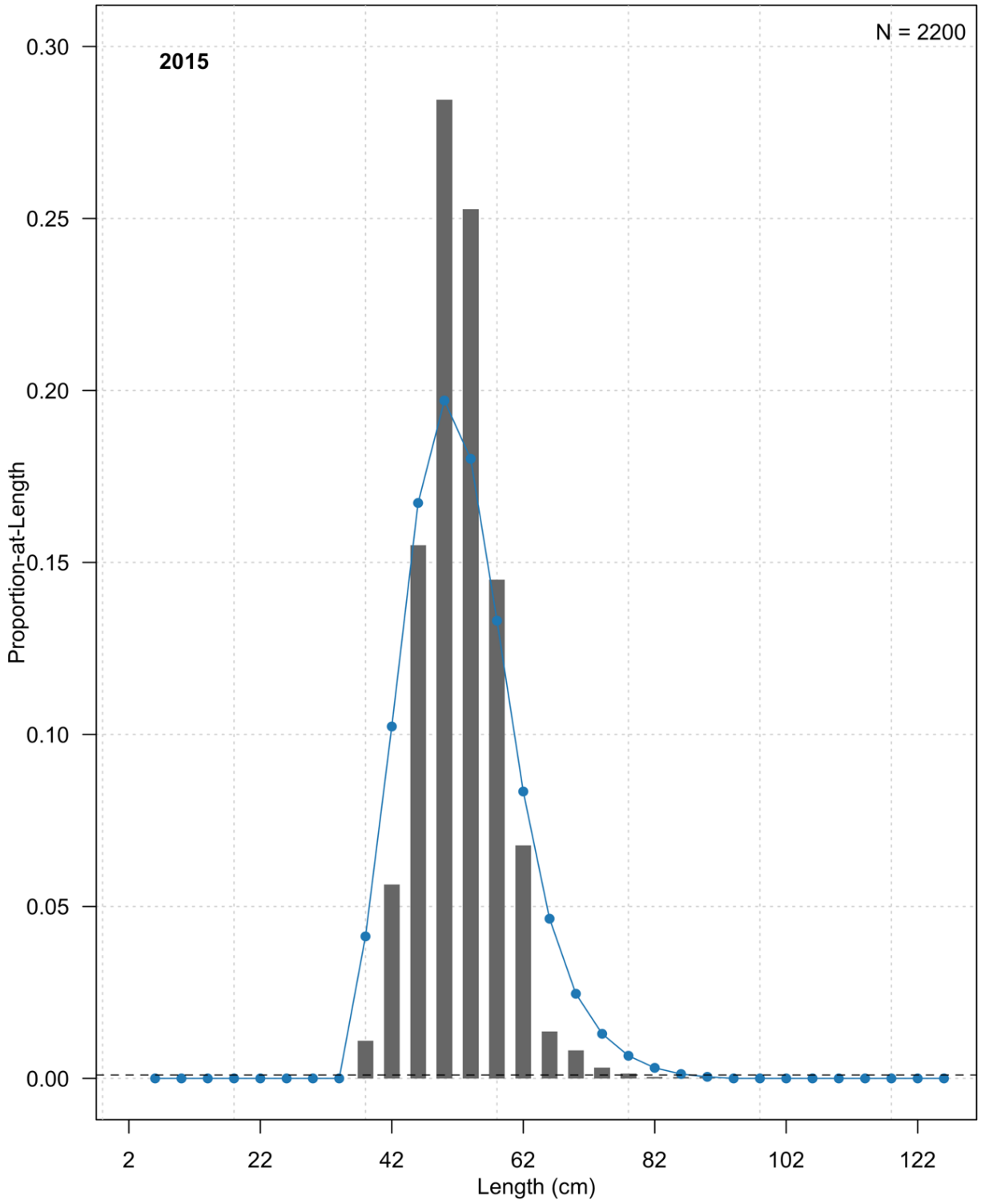


*Figure D.2. Canadian longline fishery annual estimates (lines and points) of GH-0+1 combined sex length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

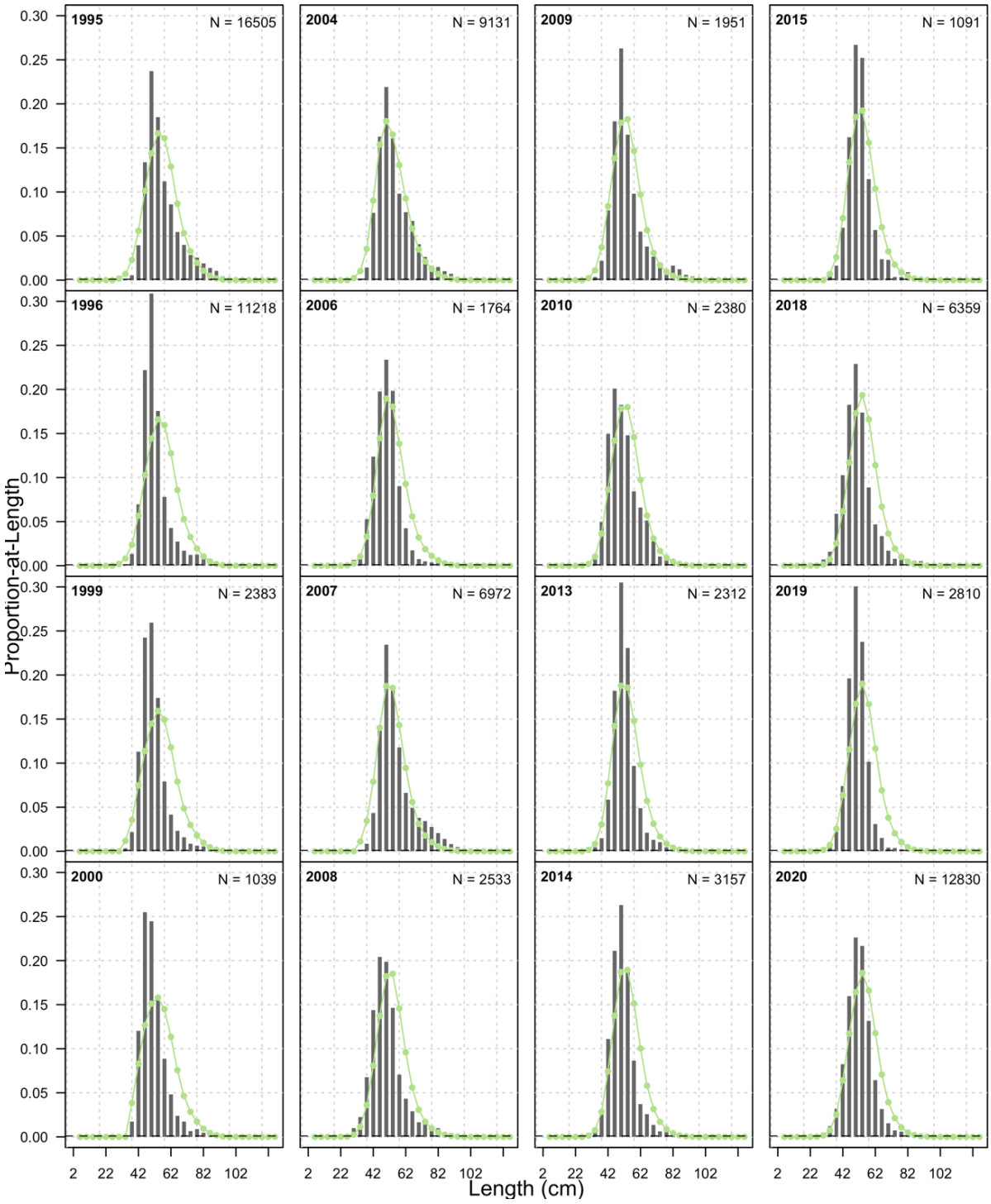


*Figure D.3. Greenland trawl fishery annual estimates (lines and points) of GH-0+1 combined sex length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

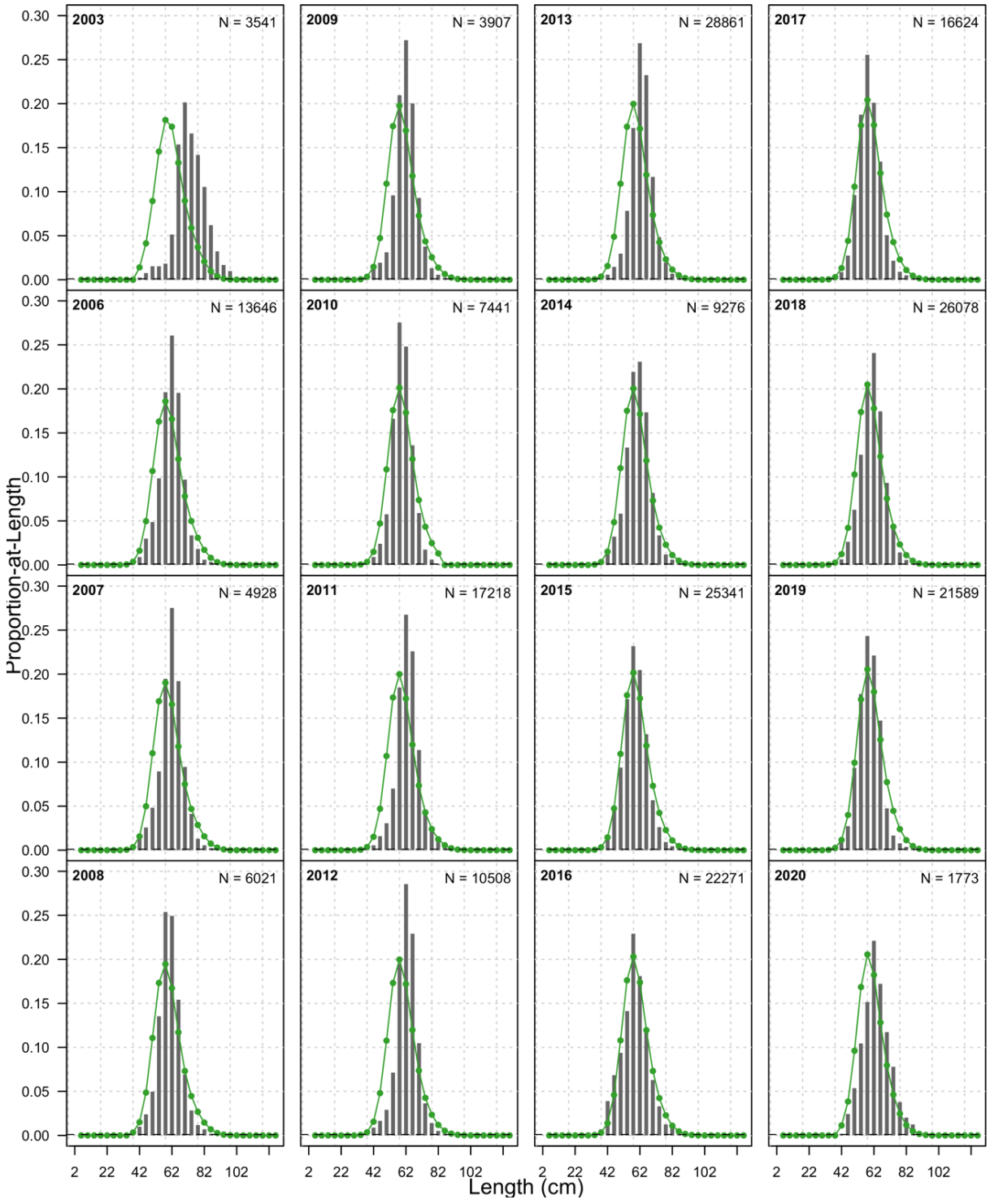


*Figure D.4. Combined Greenland and Canadian gillnet fishery annual estimates (lines and points) of GH-0+1 combined sex length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

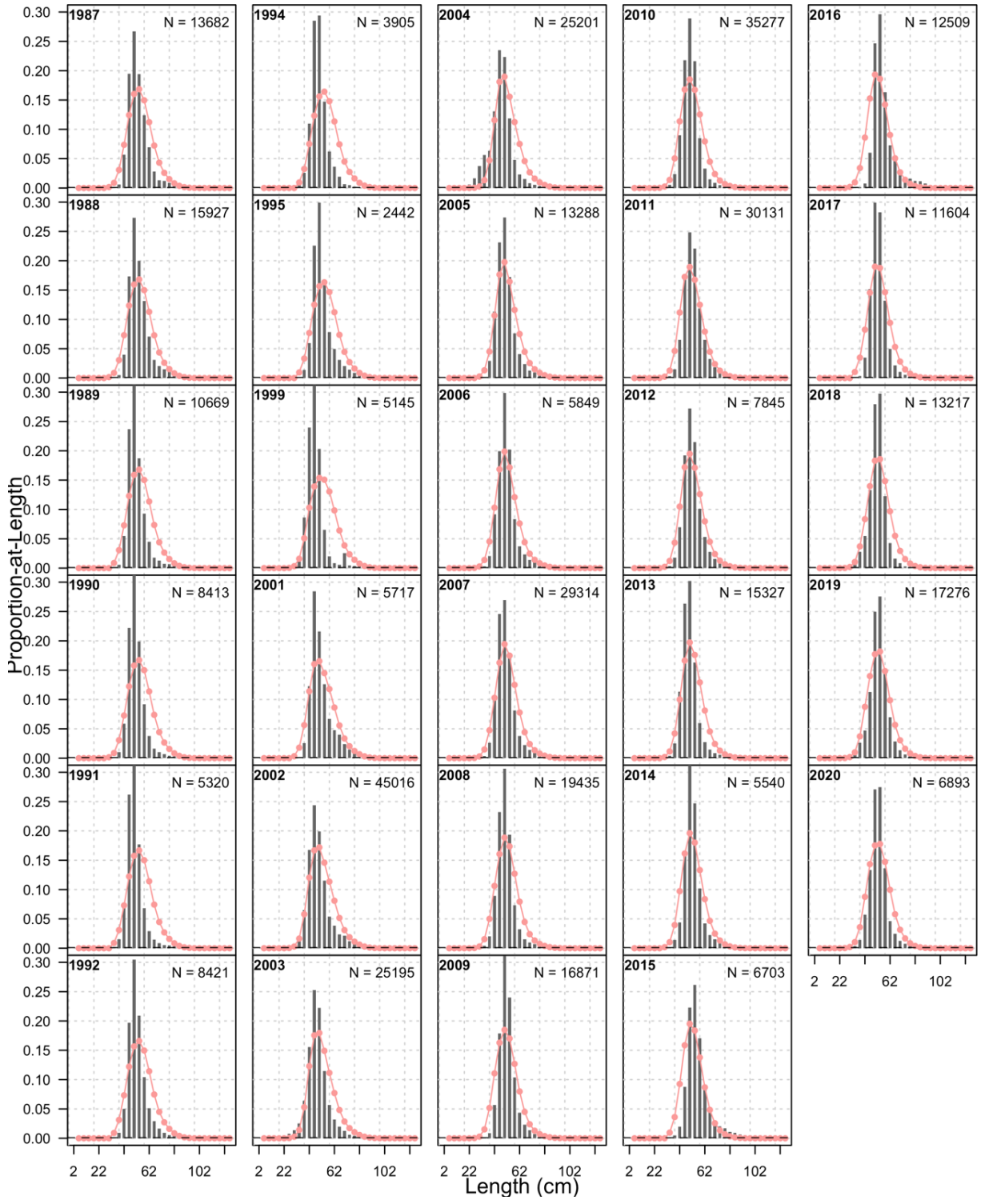


*Figure D.5. Other bottom trawl fishery annual estimates (lines and points) of GH-0+1 combined sex length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

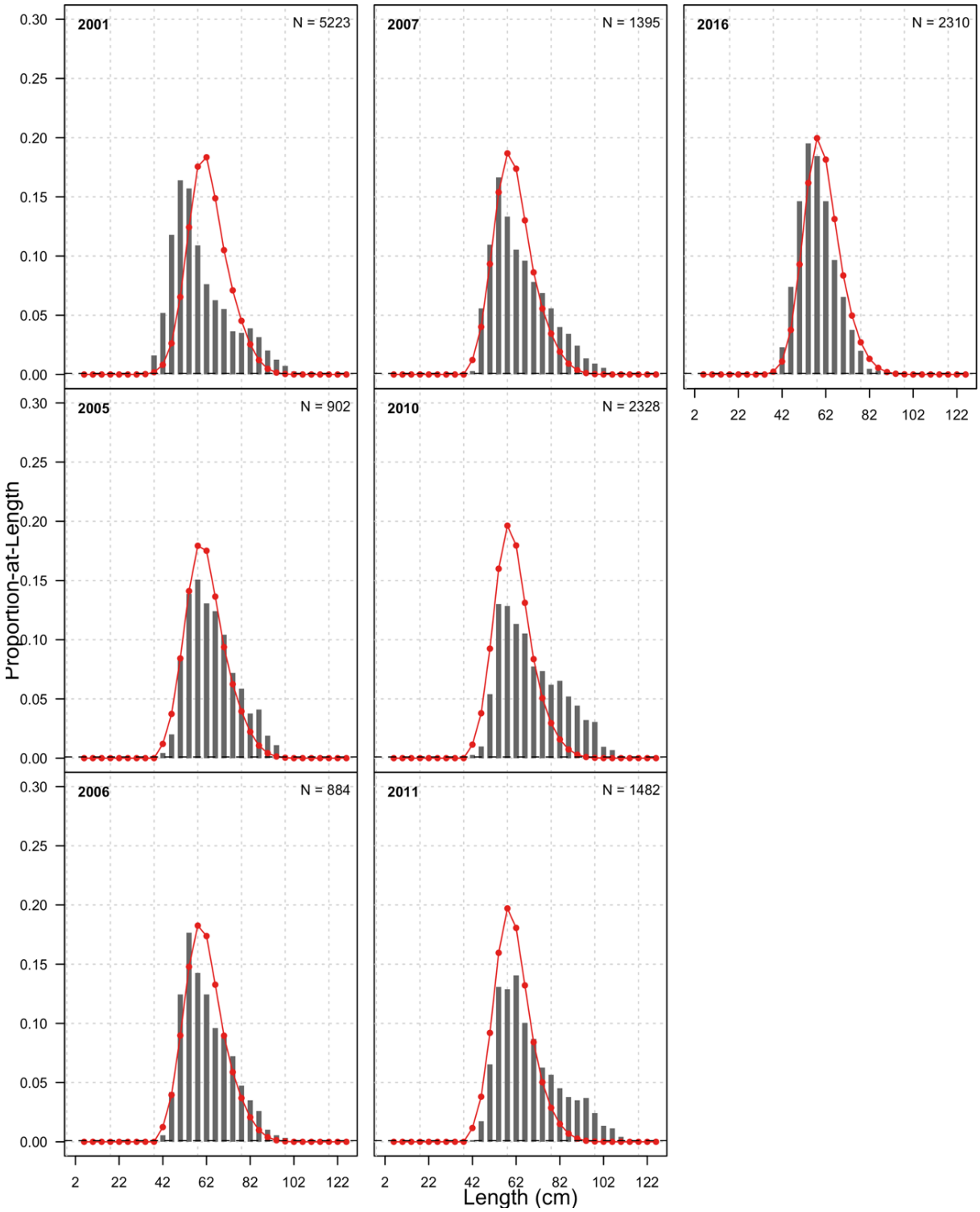


*Figure D.6. GRNF longline fishery annual estimates (lines and points) of GH-0+1 combined sex length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

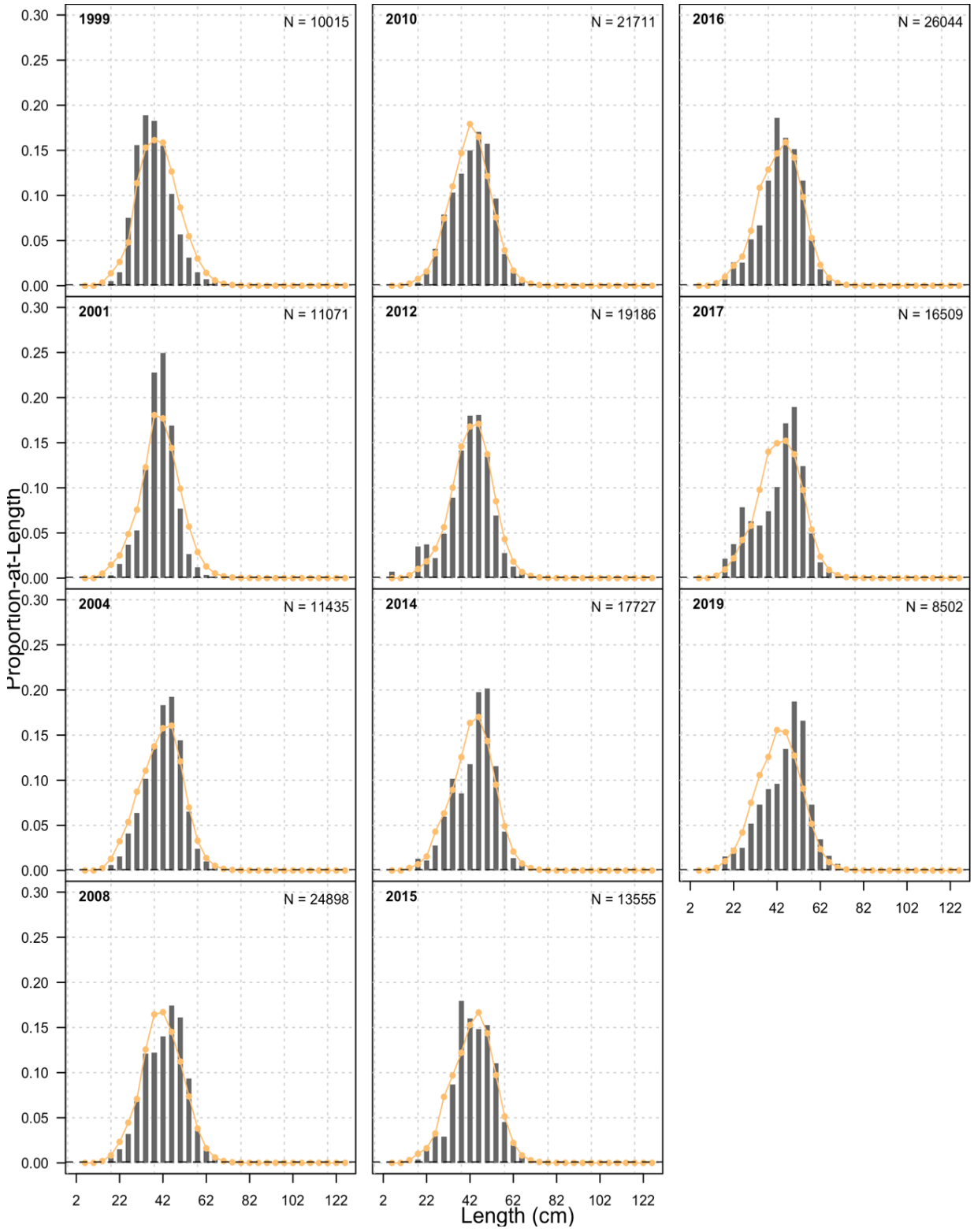


*Figure D.7. Offshore 0A/1CD RV survey annual estimates (lines and points) of GH-0+1 combined sex length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

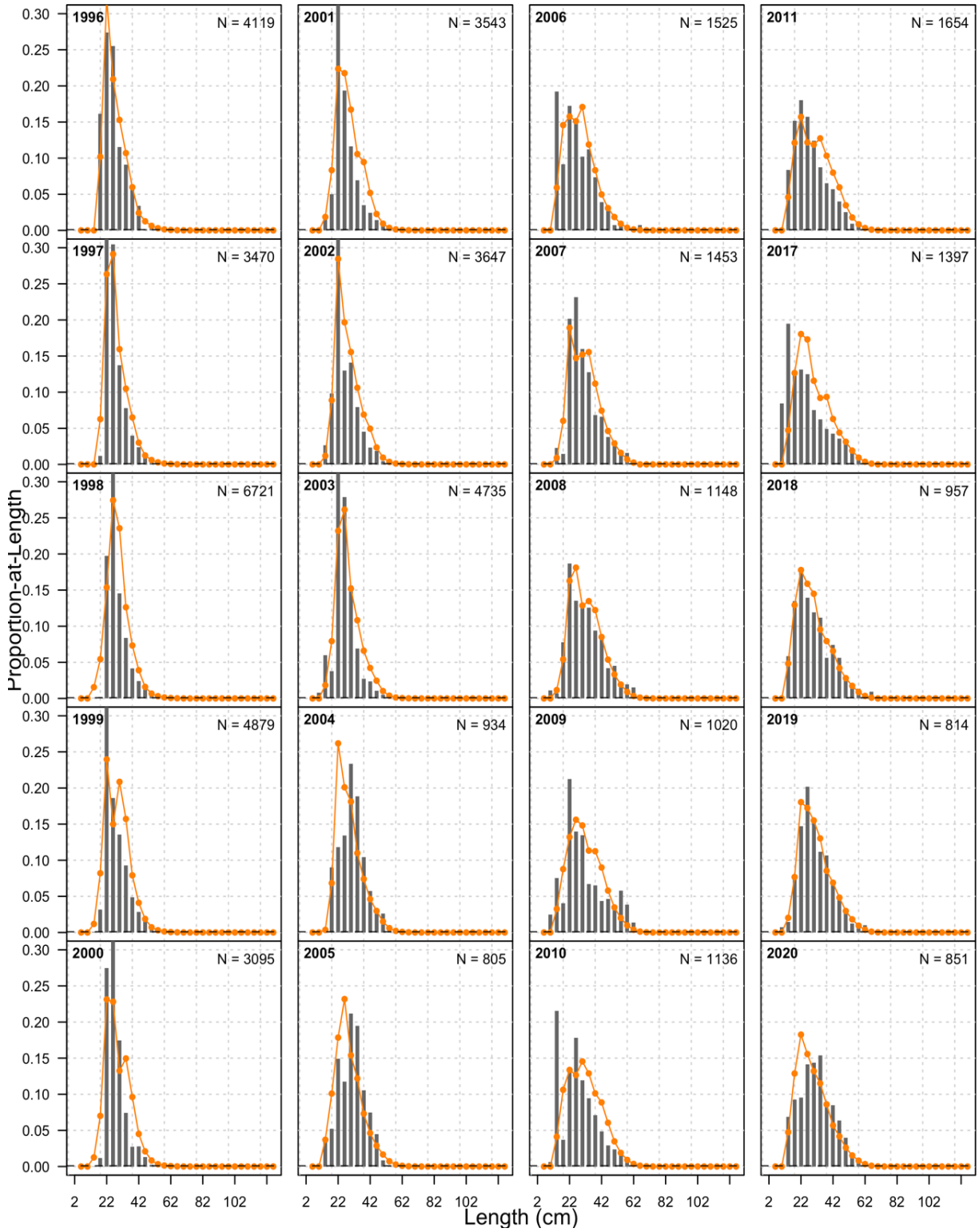


*Figure D.8. Inshore 1AF research vessel survey annual estimates (lines and points) of male GH-0+1 length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

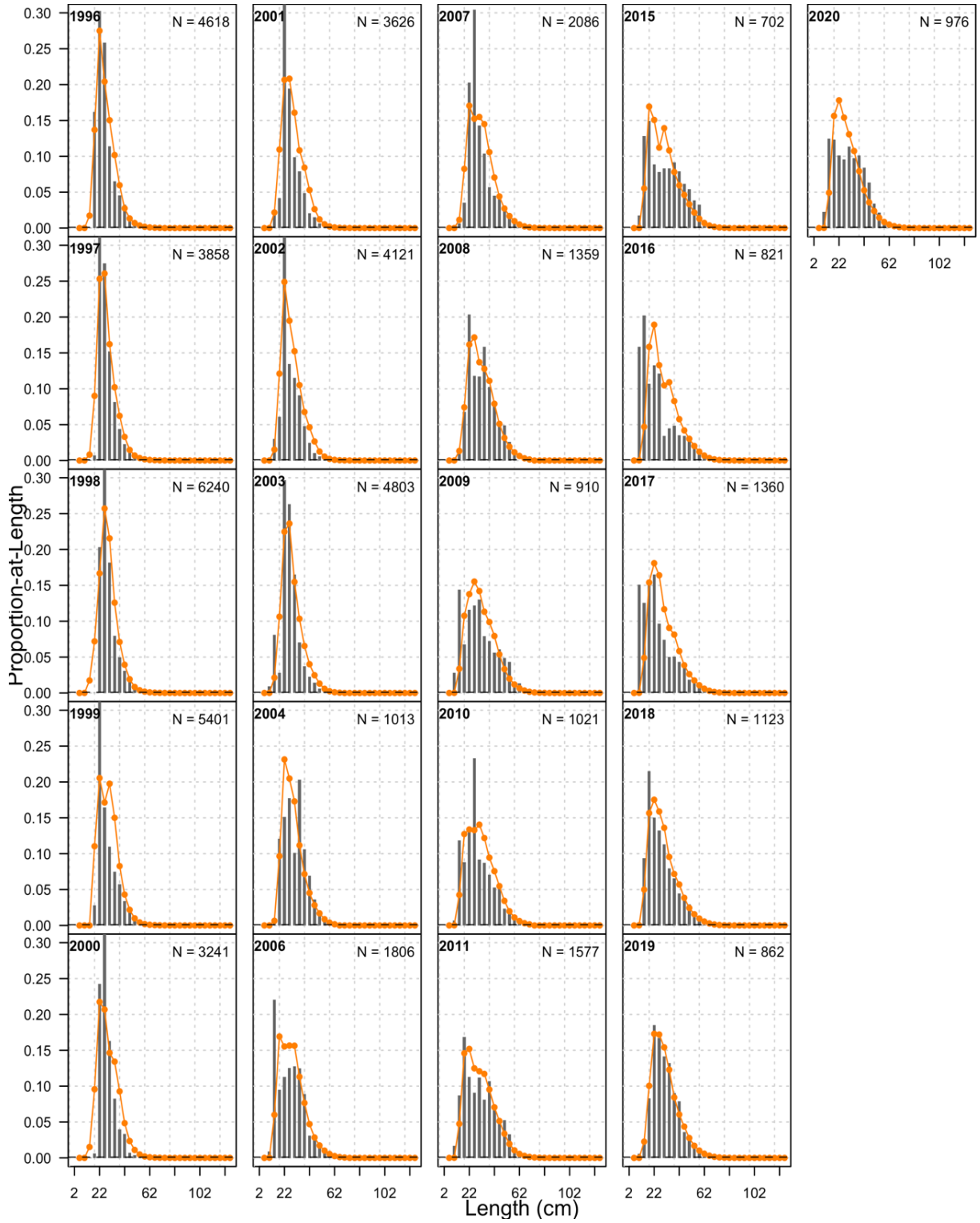


*Figure D.9. Inshore 1AF research vessel survey annual estimates (lines and points) of female GH-0+1 length composition data (bars). Each panel is annotated with the year (bold font, top left) and sample size (N, top right), and the dashed horizontal line shows the minimum proportion required to avoid tail compression.*

## APPENDIX E: ADDITIONAL MANAGEMENT PROCEDURE EVALUATION RESULTS

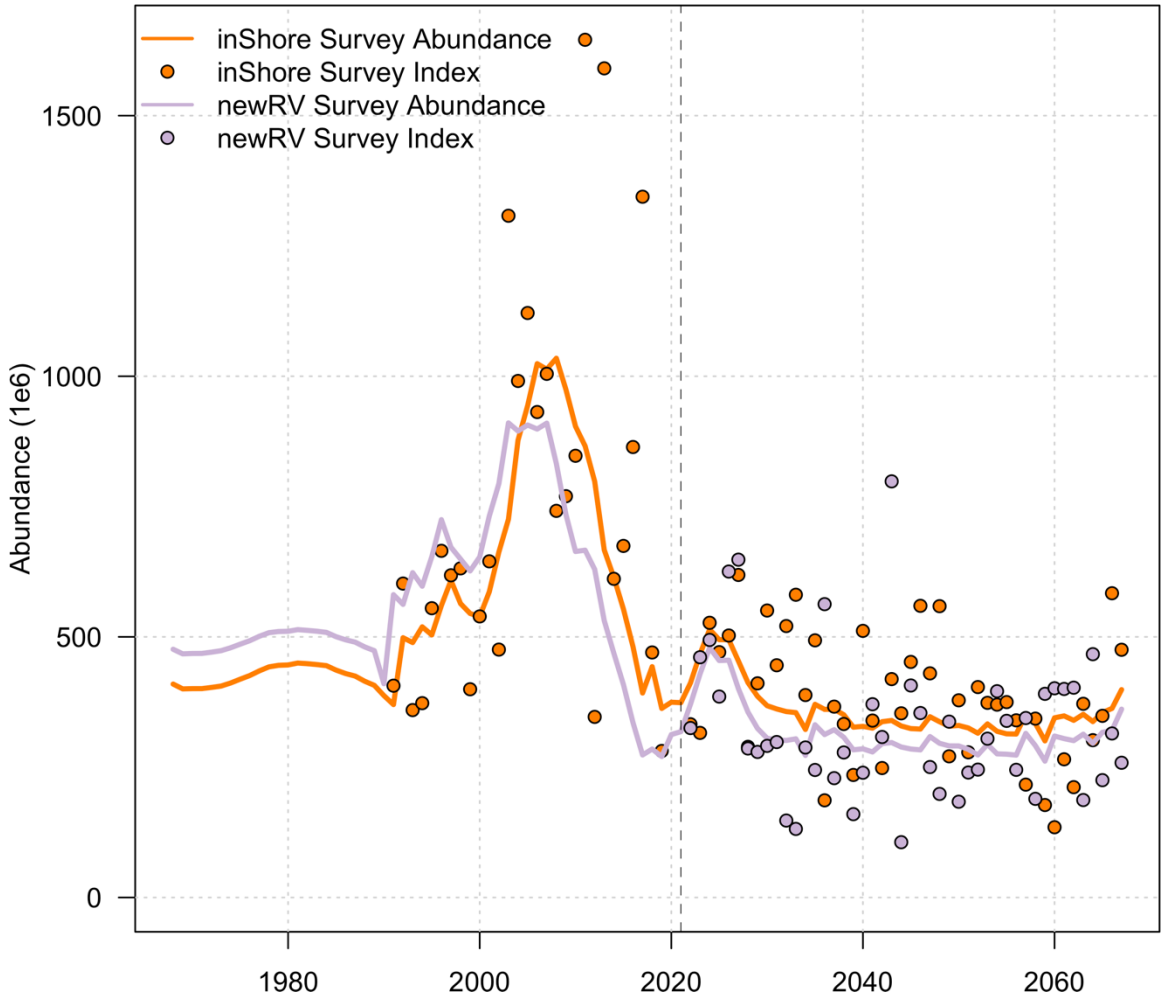


*Figure E.1. Simulated GH-0+1 inshore survey abundances (lines) and stock indices (points) for the existing inshore survey and proposed new inshore survey under the osPrec scenario.*

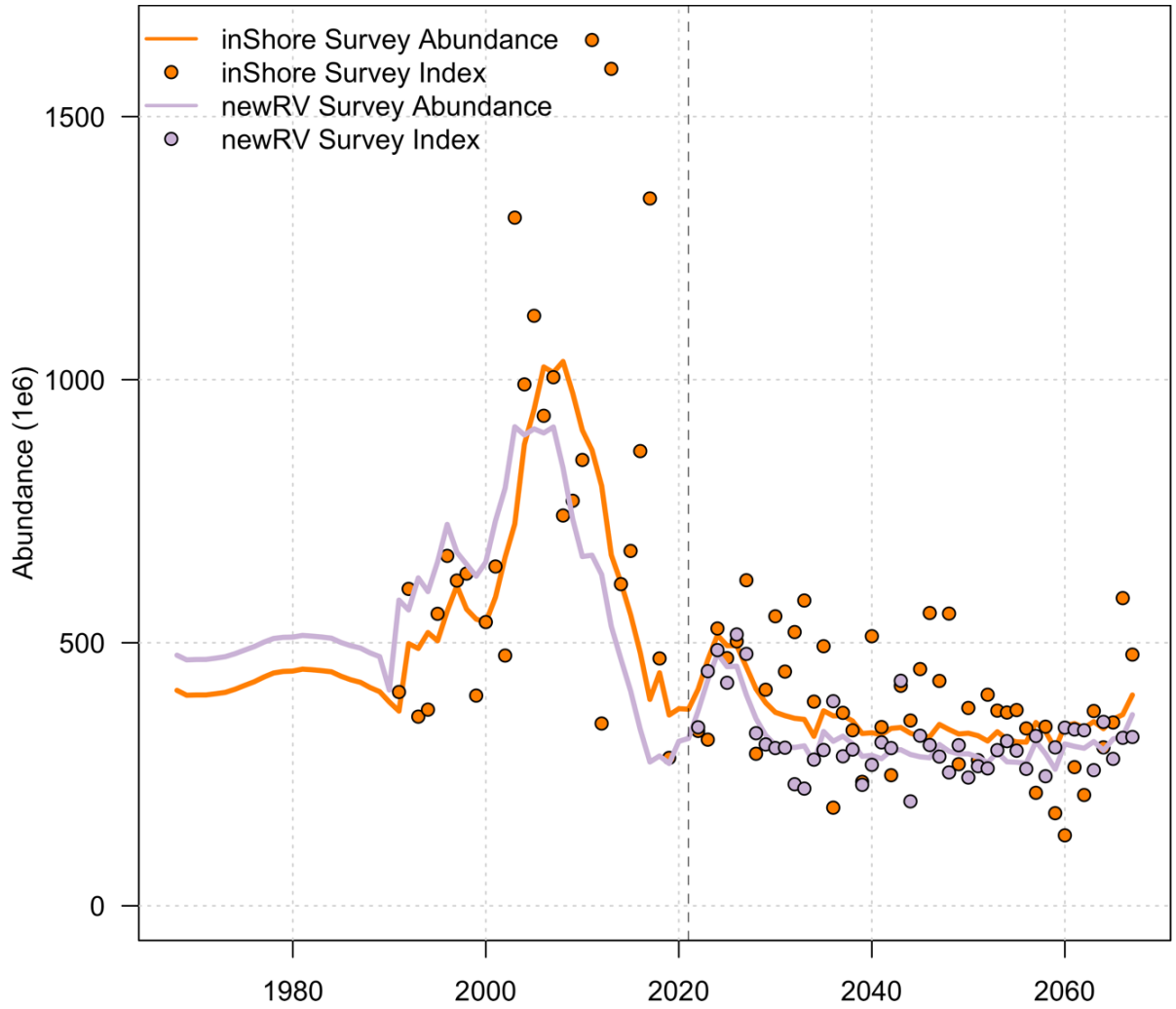


*Figure E.2. Simulated GH-0+1 inshore survey abundances (lines) and stock indices (points) for the existing inshore survey and proposed new inshore survey under the osPrec scenario.*

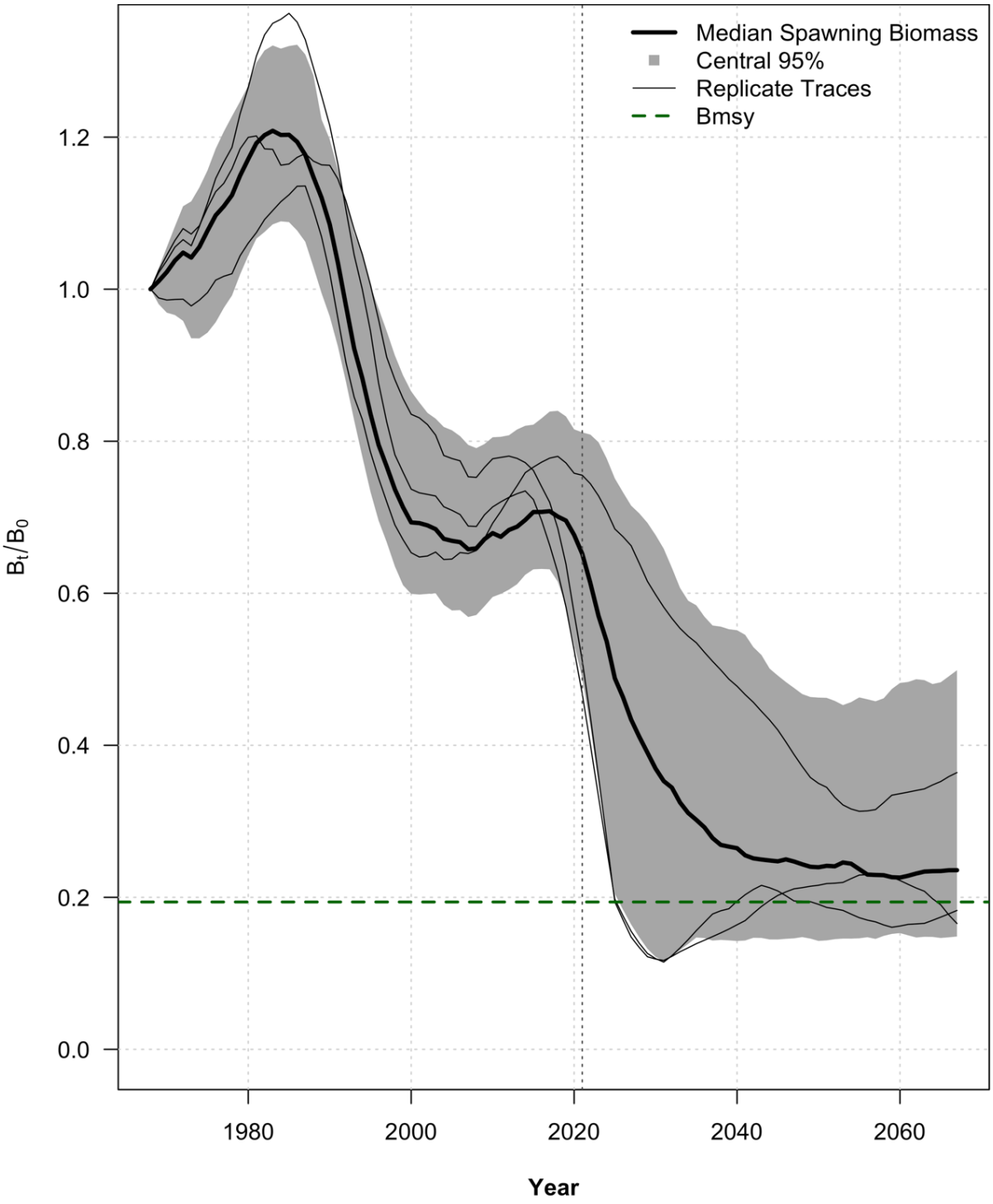


*Figure E.3. Simulated GH-0+1 female spawning biomass depletion under the adaptive model/index-based management procedure (MP) and the osPrec survey precision scenario. Grey simulation envelopes show the central 95% (grey region), median (heavy black line), and 3 random simulation replicates (thin black lines) of spawning biomass depletion, the horizontal green line shows spawning biomass producing $MSY$, and the vertical dashed line indicates the start of the simulated MP in 2021.*

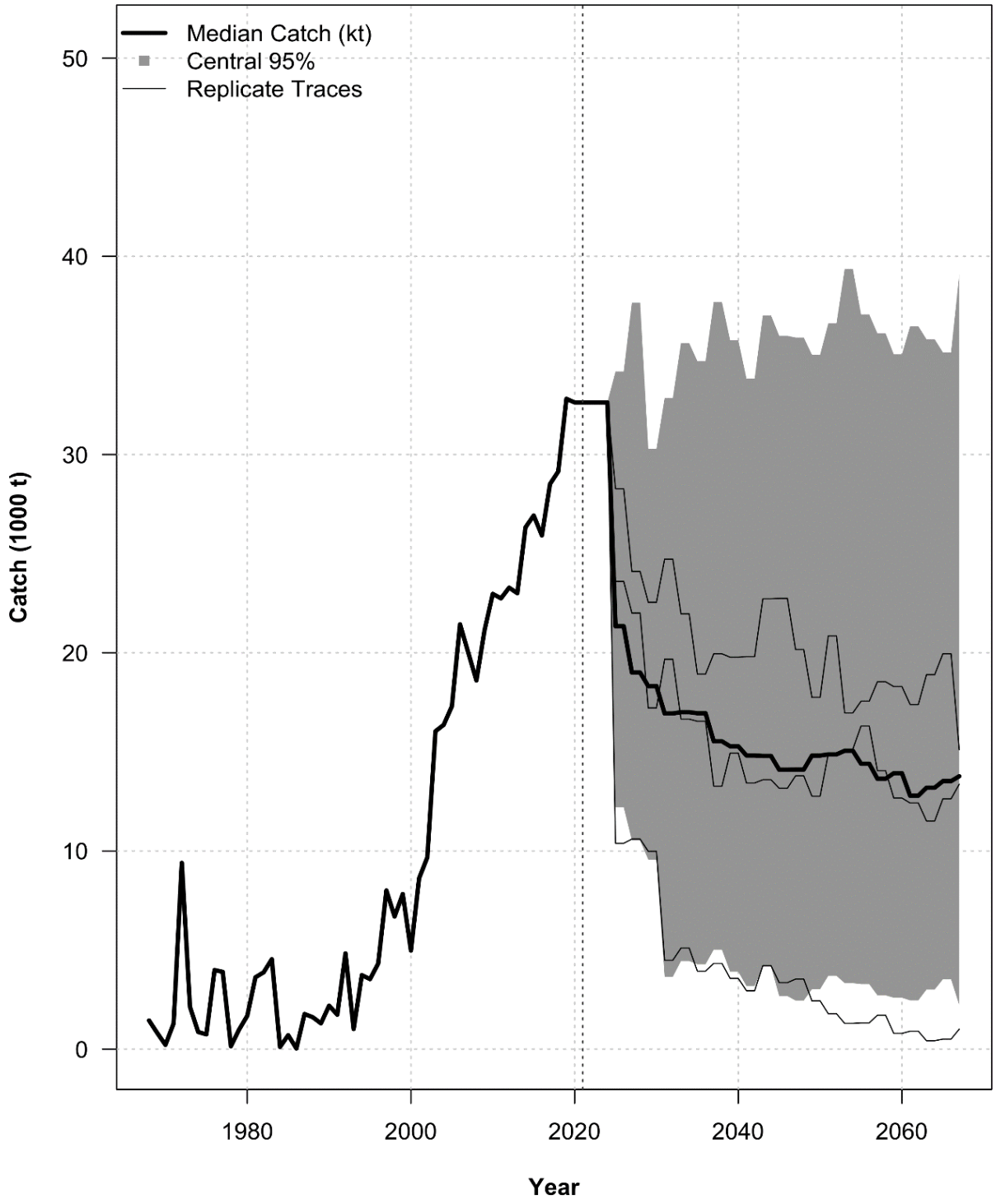


*Figure E.4. Simulated GH-0+1 total catches under the adaptive model/index-based management procedure (MP) and the osPrec survey precision scenario. Lines show median (heavy black line) and three random simulation replicates (thin black lines) with the simulation envelope showing the central 95% (light grey) of total catch over all replicates, and the vertical dashed line shows the beginning of the simulated projection period in 2021.*

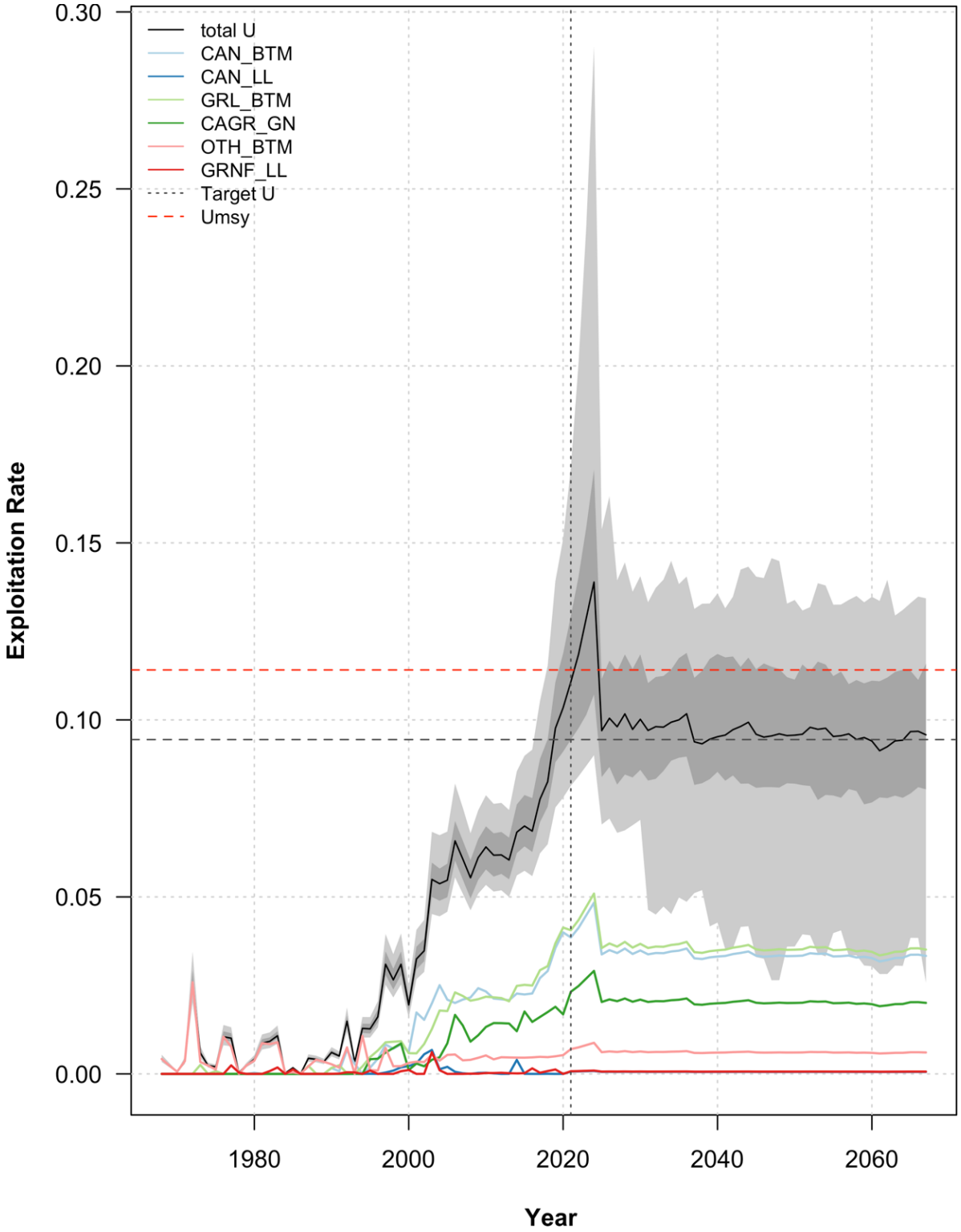


*Figure E.5. Simulated GH-0+1 exploitation rates under the adaptive model/index-based management procedure (MP) and the osPrec survey precision scenario. Lines show median total (black) and fleet specific (colours) exploitation rates, with the simulation envelope showing the central 60% (dark grey) and central 95% (light grey) of total exploitation rates over all replicates. Target and optimal exploitation rates are shown as horizontal dashed lines, and the beginning of the MP is shown as a vertical dashed line in 2021.*

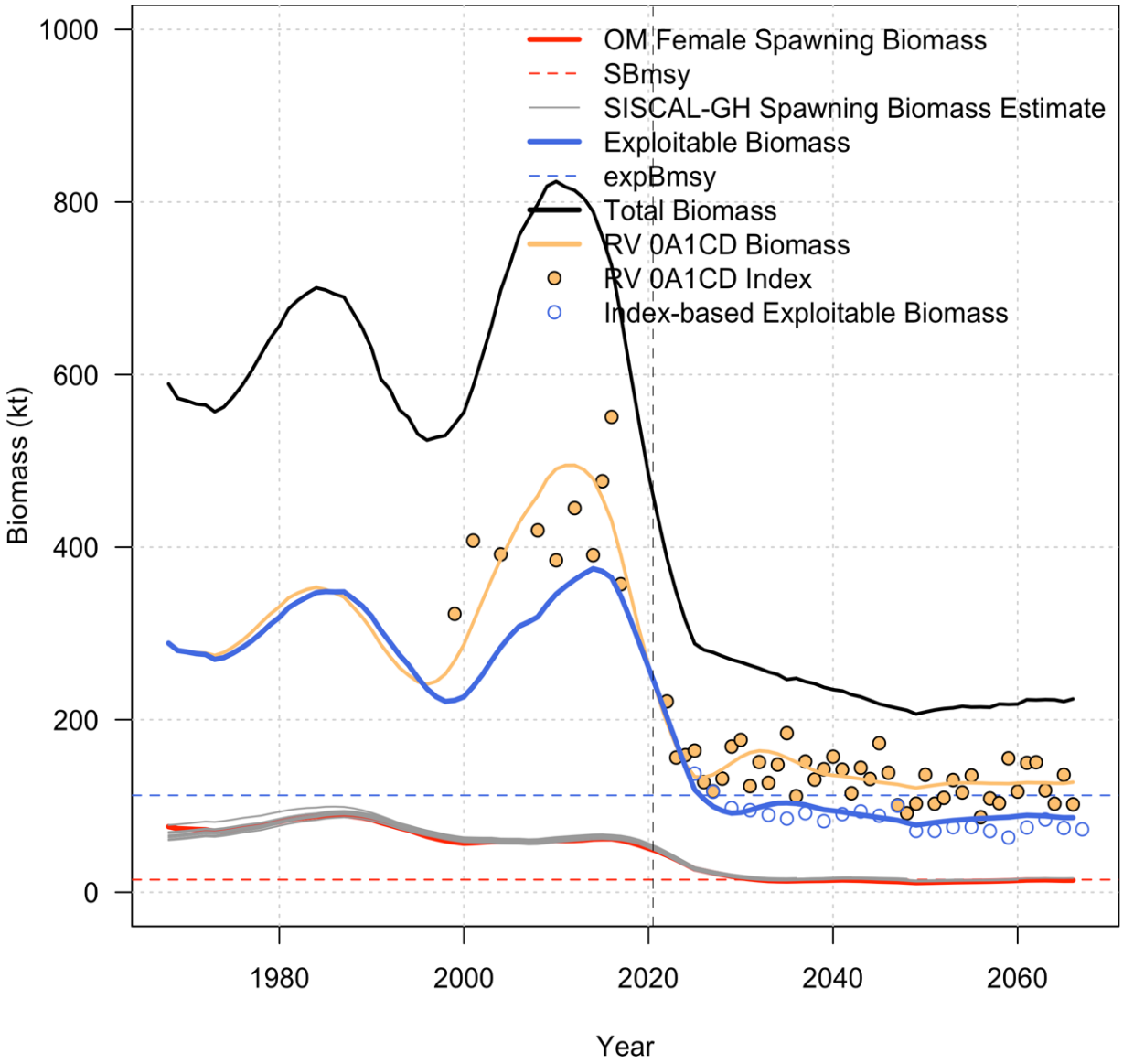


*Figure E.6. A single simulation replicate from MS3-GH under the osPrec scenario and adaptive management procedure (MP) showing the true operating model female spawning biomass (red line), exploitable biomass (blue line), total biomass (black line), and RV survey biomass (orange line), as well as simulated RV survey biomass indices (orange points). Also shown are MP estimates of female spawning biomass (grey lines) produced by the simulated SISCAL-GH model update, and estimates of exploitable biomass from the index-based procedure (blue points).*

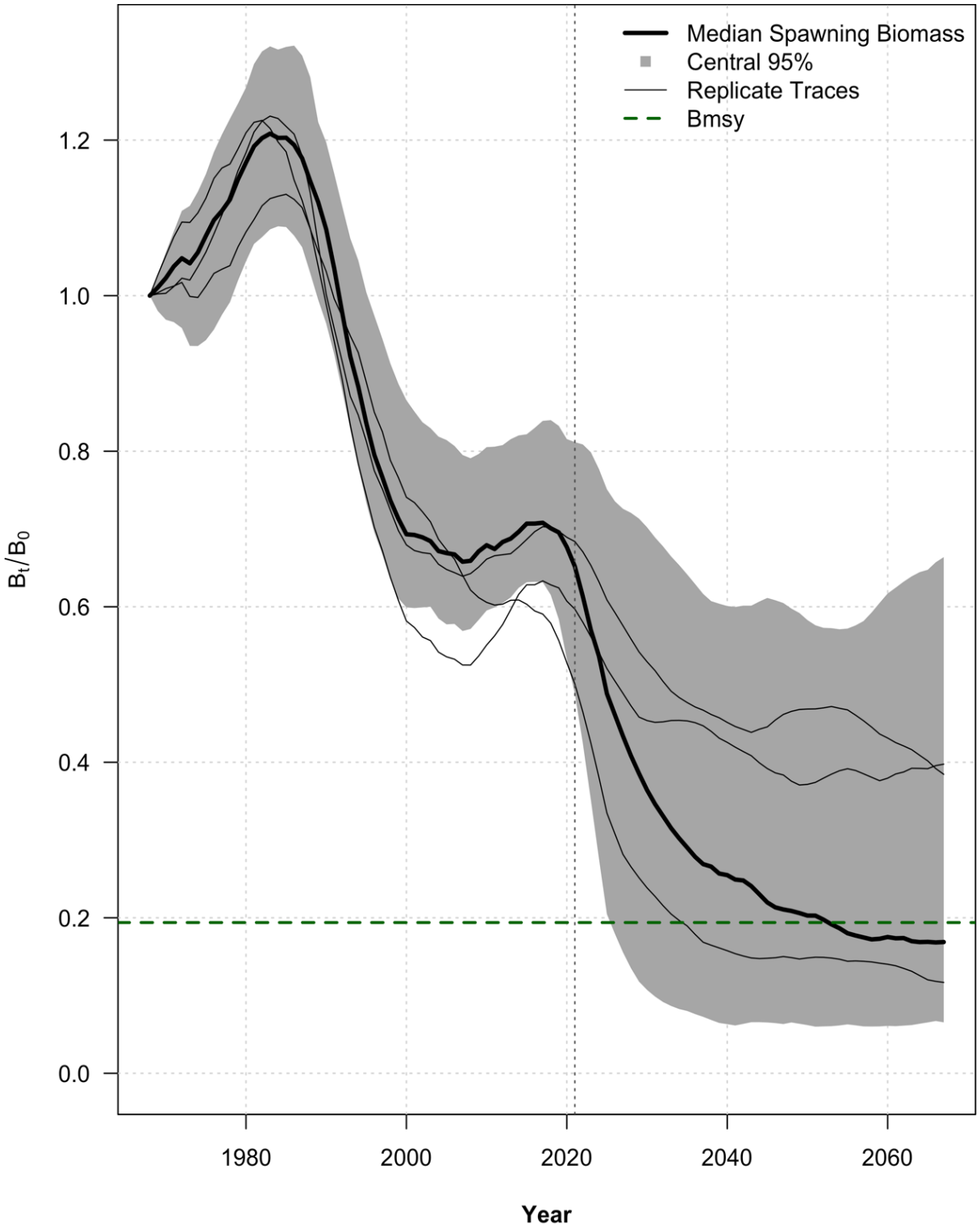


*Figure E.7. Simulated GH-0+1 female spawning biomass depletion under the non-adaptive index-based management procedure (MP). Grey simulation envelopes show the central 95% (grey region), median (heavy black line), and 3 random simulation replicates (thin black lines) of spawning biomass depletion, the horizontal green line shows spawning biomass producing $MSY$, and the vertical dashed line indicates the start of the simulated MP in 2021.*

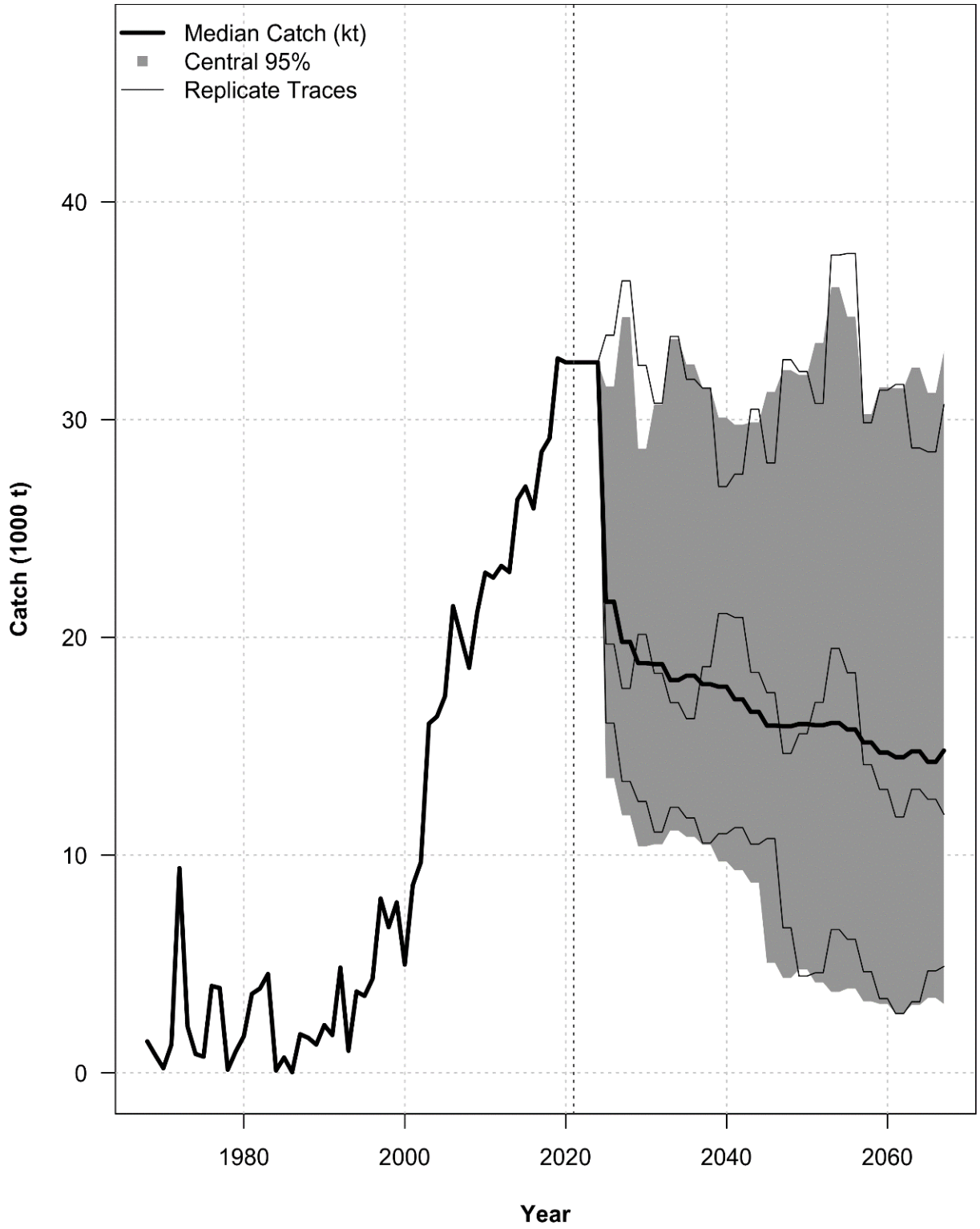


*Figure E.8. Simulated GH-0+1 total catches under the non-adaptive index-based management procedure (MP). Lines show median (heavy black line) and three random simulation replicates (thin black lines) with the simulation envelope showing the central 95% (light grey) of total catch over all replicates, and the vertical dashed line shows the beginning of the simulated projection period in 2021.*

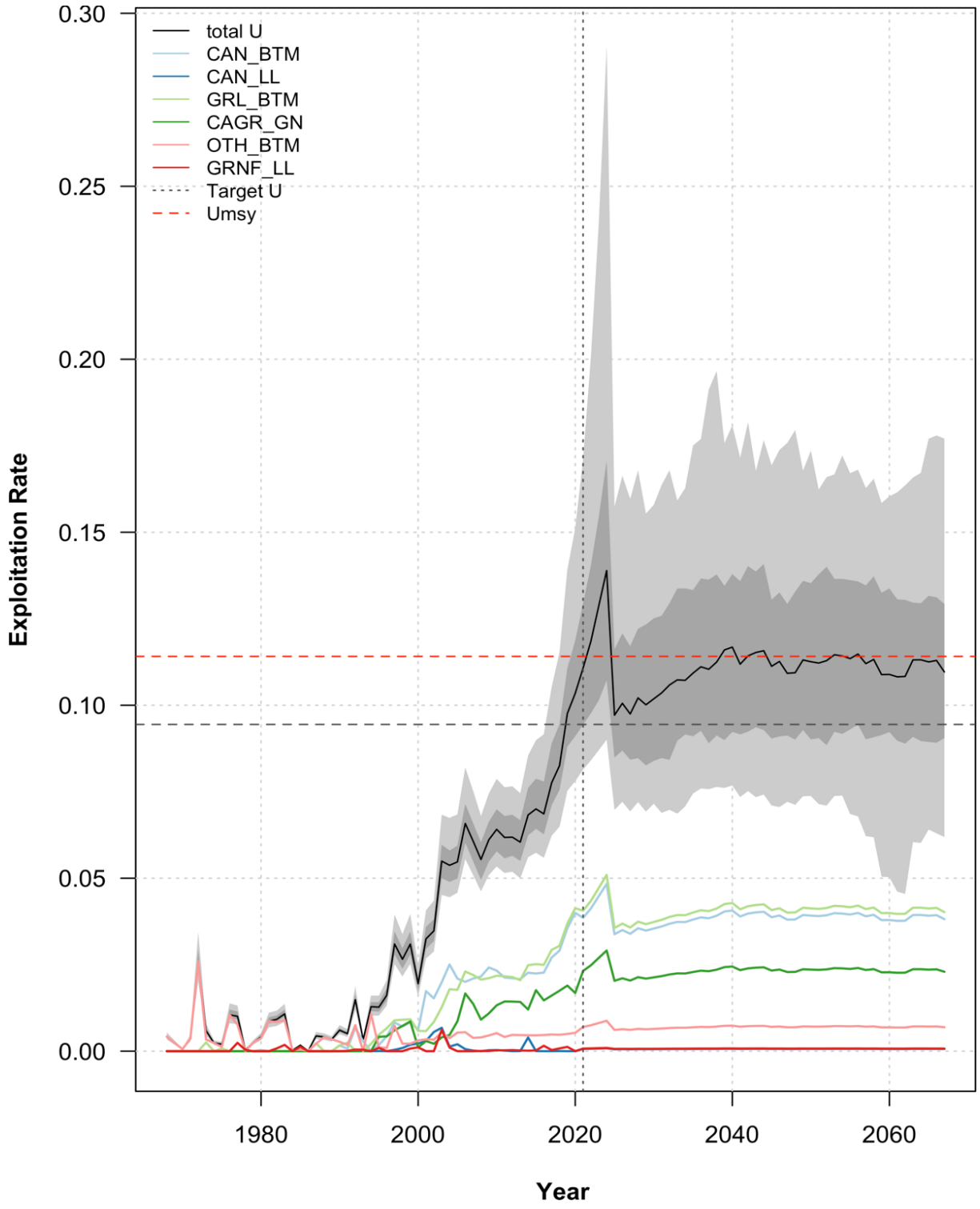


*Figure E.9. Simulated GH-0+1 exploitation rates under the non-adaptive index-based management procedure (MP). Lines show median total (black) and fleet specific (colours) exploitation rates, with the simulation envelope showing the central 60% (dark grey) and central 95% (light grey) of total exploitation rates over all replicates. Target and optimal exploitation rates are shown as horizontal dashed lines, and the beginning of the MP is shown as a vertical dashed line in 2021.*

## APPENDIX F: SIMULATED STOCK ASSESSMENT RELATIVE ERRORS

*Table F.1. Simulated 2024 SISCAL-GH model update median relative errors (MREs) and median absolute relative errors (MAREs) for selected SISCAL-GH parameters.*

| Variable | isPrec | | osPrec | |
|---|---|---|---|---|
| | MRE | MARE | MRE | MARE |
| B0 | -0.02 | 0.06 | -0.03 | 0.07 |
| h | 0.00 | 0.00 | 0.00 | 0.01 |
| M m | 0.00 | 0.07 | 0.00 | 0.08 |
| M f | 0.00 | 0.07 | 0.00 | 0.08 |
| L50A CAN BTM | 0.00 | 0.00 | 0.00 | 0.00 |
| L95A CAN BTM | 0.00 | 0.00 | 0.00 | 0.00 |
| L50D CAN BTM | 0.00 | 0.00 | 0.00 | 0.00 |
| L95D CAN BTM | 0.00 | 0.00 | 0.00 | 0.00 |
| L50A CAN LL | 0.00 | 0.00 | 0.00 | 0.00 |
| L95A CAN LL | 0.00 | 0.00 | 0.00 | 0.00 |
| L50A GRL BTM | 0.00 | 0.00 | 0.00 | 0.00 |
| L95A GRL BTM | 0.00 | 0.01 | 0.01 | 0.01 |
| L50A OTH BTM | 0.00 | 0.00 | 0.00 | 0.00 |
| L95A OTH BTM | 0.00 | 0.00 | 0.00 | 0.01 |
| L50A GRNF LL | 0.00 | 0.00 | 0.00 | 0.00 |
| L95A GRNF LL | 0.00 | 0.00 | 0.00 | 0.00 |
| q RV 0A1CD | -0.04 | 0.05 | -0.05 | 0.05 |
| tau RV 0A1CD | -0.01 | 0.07 | 0.00 | 0.07 |
| L50A RV 0A1CD | 0.00 | 0.01 | 0.00 | 0.01 |
| L95A RV 0A1CD | 0.00 | 0.00 | 0.00 | 0.00 |
| L50D RV 0A1CD | 0.00 | 0.00 | 0.00 | 0.00 |
| L95D RV 0A1CD | 0.00 | 0.00 | 0.00 | 0.00 |
| q RV SFW1AF | -0.05 | 0.07 | -0.05 | 0.08 |
| tau RV SFW1AF | -0.02 | 0.04 | -0.02 | 0.04 |
| L50A RV SFW1AF | 0.00 | 0.00 | 0.00 | 0.00 |
| L95A RV SFW1AF | 0.00 | 0.00 | 0.00 | 0.00 |
| L50D RV SFW1AF | 0.00 | 0.00 | 0.00 | 0.00 |
| L95D RV SFW1AF | 0.00 | 0.00 | 0.00 | 0.00 |
| q newRV | 0.01 | 0.16 | 0.04 | 0.16 |
| tau newRV | -0.71 | 0.71 | -0.56 | 0.56 |
| L50A newRV | 0.04 | 0.10 | 0.01 | 0.09 |
| L95A newRV | 0.02 | 0.06 | 0.00 | 0.05 |
| L50D newRV | 0.00 | 0.02 | -0.01 | 0.02 |
| L95D newRV | 0.00 | 0.04 | -0.02 | 0.04 |

*Table F.2. Simulated 2054 SISCAL-GH model update median relative errors (MREs) and median absolute relative errors (MAREs) for selected SISCAL-GH parameters.*

| Variable | isPrec | | osPrec | |
|---|---|---|---|---|
| | MRE | MARE | MRE | MARE |
| B0 | -0.02 | 0.12 | -0.02 | 0.13 |
| h | 0.02 | 0.03 | 0.03 | 0.04 |
| M m | -0.03 | 0.07 | -0.03 | 0.07 |
| M f | -0.04 | 0.07 | -0.05 | 0.07 |
| L50A CAN BTM | 0.00 | 0.00 | 0.00 | 0.00 |
| L95A CAN BTM | -0.01 | 0.01 | -0.01 | 0.01 |
| L50D CAN BTM | -0.01 | 0.01 | -0.01 | 0.01 |
| L95D CAN BTM | -0.03 | 0.03 | -0.03 | 0.03 |
| L50A CAN LL | -0.01 | 0.01 | -0.01 | 0.01 |
| L95A CAN LL | -0.01 | 0.01 | -0.01 | 0.01 |
| L50A GRL BTM | -0.01 | 0.01 | -0.01 | 0.01 |
| L95A GRL BTM | 0.00 | 0.01 | 0.00 | 0.01 |
| L50A OTH BTM | -0.02 | 0.02 | -0.02 | 0.02 |
| L95A OTH BTM | -0.03 | 0.03 | -0.04 | 0.04 |
| L50A GRNF LL | -0.01 | 0.01 | -0.01 | 0.01 |
| L95A GRNF LL | -0.01 | 0.01 | -0.01 | 0.01 |
| q RV 0A1CD | -0.05 | 0.12 | -0.05 | 0.12 |
| tau RV 0A1CD | -0.02 | 0.06 | -0.02 | 0.06 |
| L50A RV 0A1CD | 0.02 | 0.02 | 0.02 | 0.02 |
| L95A RV 0A1CD | 0.02 | 0.02 | 0.02 | 0.02 |
| L50D RV 0A1CD | -0.01 | 0.01 | -0.01 | 0.01 |
| L95D RV 0A1CD | 0.00 | 0.00 | 0.00 | 0.00 |
| q RV SFW1AF | -0.11 | 0.16 | -0.13 | 0.15 |
| tau RV SFW1AF | -0.02 | 0.05 | -0.02 | 0.05 |
| L50A RV SFW1AF | -0.01 | 0.01 | -0.01 | 0.01 |
| L95A RV SFW1AF | -0.02 | 0.02 | -0.02 | 0.02 |
| L50D RV SFW1AF | 0.00 | 0.01 | 0.01 | 0.01 |
| L95D RV SFW1AF | -0.02 | 0.02 | -0.02 | 0.02 |
| q newRV | -0.02 | 0.12 | -0.04 | 0.13 |
| tau newRV | -0.09 | 0.11 | -0.09 | 0.13 |
| L50A newRV | 0.00 | 0.04 | 0.00 | 0.03 |
| L95A newRV | -0.02 | 0.03 | 0.00 | 0.02 |
| L50D newRV | -0.01 | 0.01 | -0.01 | 0.01 |
| L95D newRV | -0.05 | 0.06 | -0.02 | 0.03 |