Consequences of ignoring climate impacts on New England groundfish stock assessment and management

Mackenzie Mazur[a1]*, Jerelle Jesse[a], Steven X. Cadrin[b], Sam Truesdell[a2], and Lisa Kerr[a]

[a]Gulf of Maine Research Institute, 350 Commercial St., Portland, ME, 04101, USA; lkerr@gmri.org, jjesse@gmri.org

[b]UMass, School for Marine Science & Technology, 836 South Rodney French Boulevard New Bedford, MA 02744, USA; scadrin@umassd.edu

[1]Current address: Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC, V9T 6N7, Canada

[2]Current address: Massachusetts Division of Marine Fisheries Cat Cove Marine Laboratory, 92 Fort Ave, Salem, MA 01970, samuel.truesdell@mass.gov

*Corresponding author, mackenzie.mazur@dfo-mpo.gc.ca

Abstract

The impact of warming on fisheries resources on the Northeast U.S. Shelf is increasingly apparent through shifts in species distribution and productivity changes of economically and culturally important stocks, such as groundfish. Ignoring such impacts can potentially lead to problems with stock assessment performance and effectiveness of fisheries management decisions. Retrospective patterns (i.e., inconsistency of recent estimates after adding another year of data) currently present a large source of uncertainty in the classification of stock status and determination of catch advice for New England groundfish. We evaluated the impact of ignoring historical climate impacts on assessment performance and the resulting management for New England groundfish. We utilized a management strategy evaluation framework to simulate the impacts of climate change on recruitment, natural mortality, and growth for New England groundfish, emulate stock assessment misspecifications, and evaluate the performance of harvest control rules. Three harvest control rules were evaluated: ramped, step in advised fishing mortality (i.e., fishing mortality reduces from 75% F_{MSY} to 70% F_{MSY} when biomass decreases below the biomass threshold) and a constrained ramped harvest control rule (i.e., the target advised catch cannot vary more than 20% from the previous year's catch). Results suggest tradeoffs among control rules, but addressing stock assessment bias resulting from misspecifications may be more important than identifying an optimal harvest control rule for meeting management objectives. Failure to account for changes in stock dynamics from climate change resulted in adverse effects on the performance of New England groundfish assessment and management, but the magnitude of impact varied by harvest control rule. Retrospective patterns caused unintended overfishing because of management actions derived from misperceptions of stock status. Our research shows how management strategy evaluation can be used to test the robustness of harvest control rules to climate change impacts on stock dynamics.

Keywords

Climate change, stock assessment, management strategy evaluation, harvest control rules, groundfish

1. Introduction

1.1. Climate change and stock assessments

Climate change is impacting fish populations throughout the world. Increasing temperatures affect key fish life history parameters, such as growth, recruitment, and natural mortality (M; Drinkwater et al., 2009), and changes in these life history parameters affect fish stock productivity and fisheries (Pörtner and Peck, 2010). These impacts caused by climate change pose challenges to our fisheries management system.

Fisheries stock assessment and the resulting fisheries management do not usually explicitly incorporate the impact of changing environmental conditions on productivity. Skern-Mauritzen et al. (2016) reviewed stock assessments and found that ecosystem processes were only incorporated into 2% of the assessments. Many assumptions in our fishery management process are strongly linked to historical productivity and the concept of stationarity, meaning that stock assessments often do not assume that parameters vary over time.

Ignoring changes in stock dynamics can lead to retrospective patterns, which are inconsistencies of recent estimates after adding another year of data to the stock assessment model (Mohn, 1999). Retrospective patterns are an expression of model instability that in some cases managers attempt to either correct for by adjusting final estimates in accordance with historical retrospective bias (e.g., Legault, 2009; Deroba, 2014) or mitigate by increasing the uncertainty associated with the overfishing limit (e.g., MAFMC, 2020). Retrospective patterns pose challenges for fisheries around the world and have been observed in many types of stock assessment models (Legault, 2009). Many New England groundfish stock assessments , which use virtual population analysis and statistical catch-at-age models, have retrospective patterns (NEFSC, 2022). Retrospective patterns can be a large source of uncertainty and if left unresolved, can lead to unintentional overfishing that undermines efforts to sustainably manage fisheries (Deroba, 2014) or, alternatively, underutilization of the resource that may have social implications. Simulation studies have shown that retrospective patterns can be caused by biased catch, trends in M, or trends in survey catchability (Legault 2009).

1.2. New England groundfish

Over the past several decades, Northeast U.S. Shelf water temperatures have increased significantly, with bottom water temperatures increasing at rates from 0.1 to 0.48°C per decade (Kavanaugh et al. 2017) and sea surface temperatures (SSTs) in the Gulf of Maine (GOM) increasing at a rate of 0.23°C per year since 2004 (Pershing et al., 2015). Groundfish off the coast of New England are expected to have differential responses to climate change due to their exposure and sensitivity to climate impacts (Hare et al., 2016; Klein et al., 2017).

Twenty groundfish stocks are managed under the Northeast multispecies groundfish federal fishery management plan (FMP) by the New England Fishery Management Council. Management of the groundfish fisheries is challenging because of historic overfishing and species-specific climate impacts on the mixed-stock fishery (Brodziak et al., 2008). The current management procedure has not performed well for New England groundfish, including recent failures of stock assessment models, lack of rebuilding or continued overfishing of several overfished stocks, and an economic disaster declaration. Management procedures are the methods for determining catch advice, which include the stock assessment methods, biological reference points (BRPs), and the harvest control rule (HCR). Harvest control rules have an important role in fishery management plans; they define management actions and are based on the status of a stock relative to reference points. The current MP usually consists of 1) a statistical catch-at-age stock assessment model, applied every two years, with many stationary assumptions and retrospective adjustments, 2) F40%, or the fishing mortality (F) expected to maintain 40% of the unfished spawning stock biomass (SSB) per recruit, as the the proxy for maximum sustainable yield and 50% of the long-term equilibrium SSB that corresponds to F40% as the overfished threshold, and 3) a stepped HCR (75% F40% when SSB is above the overfished threshold and 70% F40% when SSB is below the overfished threshold). However, when SSB is estimated to be below the overfished threshold, annual catch limits should ideally be derived from a ramped HCR with a bycatch only catch minimum and F iterated to allow 50% rebuilding probability in ten years (NEFMC, 2009). Several New England groundfish stocks are at or near historic low biomass (e.g., GOM cod, *Gadus morhua*; Georges Bank (GB) cod; GB winter flounder, *Pseudopleuronectes americanus*; GB yellowtail flounder, *Limanda ferruginea*; Southern New England-Mid Atlantic (SNE/MA) yellowtail flounder; witch flounder, *Glyptocephalus cynoglossus*; and GOM-GB windowpane flounder, *Scophthalmus aquosus*), whereas other stocks have increased to record highs (e.g., GB haddock, *Melanogrammus aeglefinus*; GOM haddock; and redfish, *Sebastes fasciatus*; NEFSC 2020, 2022).

The majority of New England groundfish stocks that have analytical assessments (e.g., statistical catch-at-age models) exhibit a retrospective pattern with estimates of SSB revised downward and estimates of F revised upwards with the addition of new data (NEFSC, 2021, 2022). Retrospective patterns represent a large source of uncertainty and pose challenges in the determination of groundfish stock status and catch advice (Brooks & Legault, 2016; Wiedenmann & Jensen, 2018). Mohn's rho (Mohn, 1999) is used to measure the magnitude of retrospective patterns, and a value higher than 0.2 or lower than -0.15 indicates considerable retrospective patterns for long-lived species (Hurtado-Ferro et al., 2015). Mohn's rho for SSB was 0.52 in the 2019 stock assessment of GOM cod when M was assumed constant at 0.2 and 0.69 in the 2019 stock assessment of GB haddock (NEFSC, 2022). Some of the candidate causes of retrospective patterns relevant to Northeast groundfish stocks include misreporting of catch, changes in survey or fishery catchability and/or selectivity, and, more generally, directional ecosystem change such as the impact of ocean warming on population dynamics that may impact those processes. Kerr et al. (2020) found that catch misreporting impacts stock trajectories, assessment, and management performance of GOM cod. Retrospective patterns appeared when catch bias changed overtime.

Given the documented impacts of climate change on New England groundfish population dynamics and significant retrospective patterns in New England groundfish stock assessments, there is a need to evaluate the impacts of stock assessment misspecifications due to climate change. The objectives of this study are: 1) to evaluate the consequences of ignoring historical climate change impacts on New England groundfish stock dynamics and how this affects current stock assessment and management performance, and 2) test robustness of fisheries management alternatives to stock assessment misspecifications. A stock assessment misspecification is when the assessment assumptions are not consistent with true system dynamics, which, if the misspecification varies over time, can lead to retrospective patterns. To address these objectives, we applied management strategy evaluation (MSE), which involves simulating the natural and human aspects of the managed fishery system under different scenarios and evaluating performance for meeting management objectives. We simulated groundfish population dynamics impacted by climate change, stock assessment specifications, and HCRs. The performance of the current New England groundfish management procedure and possible alternatives have not yet been simulation tested. MSE is valuable for identifying HCRs that are robust to natural variation in the system and to uncertainty and error (ICES, 2020). The better uncertainty is represented in an MSE, the more informative such model frameworks are for fisheries management (Punt et al., 2016).

2. Methods

2.1. Model framework overview

In the MSE framework we used to evaluate the impacts of stock assessment misspecifications on groundfish fisheries (https://github.com/lkerr/groundfish-MSE), the operating model (OM) represented the true fish population dynamics and was the basis for evaluating performance relative to the ‘true’ values for the stock and fishery (1; Fig. 1). Through an observation model (2; Fig. 1), simulated trawl survey and catch data were generated with random error to represent the information available for groundfish assessment and management. The simulated survey and catch data informed a statistical catch-at-age stock assessment model (3; Fig. 1) used to estimate stock and fishery metrics. This study emulated current groundfish stock assessment methods and applied the Age Structured Assessment Program (ASAP; Legault & Restrepo, 1998) as the estimation model, which is used for the majority of analytical groundfish stock assessments in the region. Biological reference points (4; Fig.1) were then calculated. The stock assessment output and estimated BRPs were compared to produce estimated stock status. A HCR (5; Fig. 1) then determined a prescribed F based on the estimated stock status. Both the F from the HCR (5) and output from the stock assessment (3) were used in projections (6; Fig. 1) to determine catch advice. Catch advice was generated from projected catch with F determined from the HCR for two years. This catch advice was then applied to simulate harvest in the OM (7; Fig. 1). The advised catch was assumed to be caught in the OM (i.e., there was no bias or uncertainty in catch). Quantities necessary to develop the performance metrics were compiled at each timestep (8; Fig. 1). One hundred iterations were simulated for each scenario. Further description of the MSE components can be found on the Github repository Wiki page (https://github.com/lkerr/groundfish-MSE/wiki).

New England groundfish stock assessments are updated every two years and use data up to the year before the assessment year unless there are unforeseeable circumstances (i.e. a pandemic). This simulated process was designed to be consistent with current New England groundfish management procedure whereby the stock assessment performed in year *t* had a terminal year of *t-1*, and the resulting catch advice was for year *t+1* and *t+2*. Thus, there was a lag in information that informed the catch advice. This simulated fishery resource, management, and harvest feedback loop continued until the end of the management procedure period (2019 - 2040). The MSE approach used in this study was not a full participatory MSE because management objectives were not identified and prioritized by stakeholders. Operating model parameters and stock assessment misspecifications described below were specified in the respective R script in the stockParameters folder on the Github (https://github.com/lkerr/groundfish-MSE/tree/master/modelParameters/stockParameters). Biological reference points and HCRs described below were specified in the csv file for management procedure specifications (https://github.com/lkerr/groundfish-MSE/blob/master/modelParameters/mproc.csv).

2.2. Operating models

We focused OM development on two groundfish stocks: GOM cod and GB haddock to typify a range of conditions experienced by groundfish stocks. The current stock status of GOM cod is overfished and undergoing overfishing, whereas GB haddock is not overfished and not undergoing overfishing (NEFSC, 2021, 2022). GB haddock exemplifies a groundfish stock with a recently increasing stock size. Throughout the manuscript, GOM cod represents an overfished stock and GB haddock represents a not overfished stock, because GOM cod and GB haddock have similar characteristics to other overfished and not overfished New England groundfish stocks, respectively.

The OMs for groundfish stocks in this framework were single species, stochastic, age-structured models designed to emulate population dynamics. The GOM cod and GB haddock historical trajectories were reconstructed by incorporating recruitment and F time series (1982-2018 for cod, 1931-2018 for haddock) from the 2019 stock assessments (NEFSC, 2022) and calculating SSB and catch as emergent properties. Abundance-at-age was calculated using exponential survival. Haddock growth and maturity in the management procedure period were modeled using the average weight-at-age and maturity-at-age from the last five years of the stock assessment. Cod growth and maturity used the weight-at-age and maturity-at-age used in the 2019 projections. The purpose of the historical period was to emulate reality, as it was perceived by groundfish stock assessments. The management procedure period began in 2019. A description of how to modify operating models in the MSE framework can be found at https://github.com/lkerr/groundfish-MSE/blob/master/documentation/Operating_models.pdf.

2.3. Simulating impacts of climate change on groundfish dynamics

We simulated climate and ecosystem change on stock dynamics, specifically changes in recruitment, natural survival, and growth. GB haddock exhibit periodic high recruitment events that are not explained by a theoretical stock-recruitment relationship (SRR), but have been linked to ocean conditions (i.e., autumn bloom; Leaf and Friedland, 2014; Friedland et al., 2015). However, recent abnormally high recruitment events were not explained by ocean conditions. As a result, for GB haddock, recruitment was modeled using empirical cumulative distribution functions based on historical estimated recruitment. Previous studies have documented evidence of the negative impacts of warming water temperatures on GOM cod recruitment (Fogarty et al., 2008, Pershing et al., 2015). For GOM cod, recruitment was modeled using a Beverton-Holt stock recruitment model that included the effect of projected temperature increase on recruitment in the management procedure period. The stock-recruitment model was fit to recruitment and SSB output from the 2019 stock assessment (M=M-ramp; NEFSC, 2022) and annual mean SST anomalies for the GOM. Sea surface temperature anomalies were derived from GOM Optimum Interpolation SST (OISST) data (Huang et al., 2020), which were used to downscale Northeast LME CMIP5 projections. This relationship showed a negative impact of temperature on cod recruitment.

Natural mortality, a highly uncertain parameter for most fish populations, is often assumed to be constant over time, and this assumption is likely not true (Vetter, 1988; Lorenzen, 2022). The survival of GOM cod is negatively impacted by warming waters (Fogarty et al. 2008; Pershing et al. 2015), and including a specified time-varying M in the GOM cod stock assessment model improved model diagnostics and performance (Pershing et al., 2016; NEFSC, 2022). For these reasons, we simulated an increase in M for GOM cod over time in the OM. Natural mortality increased from 0.2 to 0.4 based on the parameterization of increasing M in the GOM cod stock assessment (i.e., M-ramp scenario). In this scenario, M started increasing from 0.2 in 1988 to 0.4 in 2003 where it remained at 0.4 through the rest of the historical period and into the management procedure period (NEFSC, 2022). This OM specification was designed to capture the shift in M to a higher state.

Haddock weight-at-age has decreased over time. This change in growth is explained by density-dependent factors but also temperature (Wang et al., 2021). Warm temperatures at the larval stage for haddock negatively impact adult growth (Brodziak and Link, 2008). For these reasons, we simulated a decrease in weight-at-age over time for haddock during the historical period, based on the parameterization from the GB haddock stock assessment (NEFSC, 2022). During the management procedure period, haddock weight-at-age was constant over time. Values for weight-at-age during the management procedure period were the average of the last five years of the most recent haddock stock assessment (NEFSC, 2022).

2.4. Stock assessment specifications

We simulated scenarios in which the stock assessment accounts and does not account for the historical impacts of a changing climate. For GOM cod, three stock assessment assumptions were simulated; M was assumed constant at 0.2 or 0.3 or correctly assumed to be increasing from 0.2 to 0.4. Constant M assumptions were consistent throughout the historical and management procedure period. The projections had the same assumptions as the final period of the stock assessment. Both true and estimated BRPs were estimated with M at 0.2, even though M increased to 0.4 in the OM, because the stock was at a lower productivity (Legault and Palmer, 2016). Another reason we chose to calculate true and estimated reference points with M=0.2 was to be able to provide insights to other overfished stocks. Additionally, the BRP estimation and projections do not account for the negative impact of temperature on recruitment. BRP estimation and projections assume recruitment to be the mean of the previous 20 years of estimated recruitment.

For haddock, three stock assessment assumptions were simulated; weight-at-age assumed to be high and constant, intermediate and constant, and decreasing as in the OM, based on the parameterization from the GB haddock stock assessment (NEFSC, 2022; Table 1). In the high and constant assumption, weight-at-age was assumed to be the average weight-at-age from the OM from 2000 to 2005. In the intermediate and constant assumption, weight-at-age was assumed to be the weight-at-age from the OM from 2010 to 2015. In the decreasing the same as the OM assumption, the stock assessment correctly assumed time-varying weight-at-age. The projections and BRP estimations had the same assumptions as in the final five years of the stock assessment model to account for the decreased weight-at-age when correctly specified. In both the cod and haddock scenarios, retrospective adjustments were not applied to the estimated values from the stock assessment. A description of how to specify the stock assessment methods in the MSE framework can be found at https://github.com/lkerr/groundfish-MSE/blob/master/documentation/MSEFramework_StockAssessmentMethods.pdf.

2.5. Biological reference points

The F associated with the maximum sustainable yield (F_{MSY}) proxy used in these HCRs was $F_{40\%}$, or the F expected to maintain 40% of the unfished SSB per recruit, which was determined with spawner per recruit (SPR) analysis and is the current overfishing definition for GOM cod, GB haddock and other groundfish. The F_{MSY} proxy $F_{40\%}$ will hereafter be referred to as F_{MSY}. The SSB_{MSY} proxy was the long-term equilibrium SSB that corresponded to F_{MSY} (i.e., $SSBF_{40\%}$) and is the current B_{MSY} proxy for GOM cod, GB haddock and other groundfish in the northeast US. For the estimated and true (OM) SSB_{MSY} proxies, recruitment used in the equilibrium calculation was the mean of the previous 20 years of estimated or true (OM) recruitment values. These recruitment values were dynamic and changed with the addition of years in the simulation. The SSB_{MSY} proxy $SSBF_{40\%}$ will hereafter be referred to as SSB_{MSY}. The SSB threshold used in alternative HCRs was 50% SSB_{MSY}; when SSB is estimated to be below this level the stock is deemed overfished. True reference points were re-calculated every year and are dynamic. A description of how to specify BRPs in the MSE framework can be found at https://github.com/lkerr/groundfish-MSE/blob/master/documentation/MSE_ReferencePoints.pdf.

2.6. Alternative harvest control rules

Three HCRs were evaluated: ramp, F-step, and constrained ramp HCRs. All HCR alternatives included a constraint on catch advice so that it would not be higher than the estimated catch that corresponds to the estimated overfishing limit (OFL) from the stock assessment (i.e., catch at F_{MSY}) to emulate the current in-season quota monitoring system. However, in misspecified scenarios, the true catch could be larger than the catch that corresponds to the true OFL in the OM when there was biased estimation from the stock assessment. All these alternatives also had a minimum catch limit (i.e., the minimum bycatch of the last ten years in the historical period), which would prevent F from declining to zero.

The ramp HCR is intended to promote rebuilding and optimal yield when the stock is not overfished (Fig. 2a). When stock status was greater than 50% SSB_{MSY} (i.e., the 'overfished' threshold), the target F was 75% F_{MSY}. When stock status was perceived to be less than 50% SSB_{MSY}, the target F linearly decreased towards zero as SSB decreased. This HCR is modeled after the current groundfish harvest strategy employed by the NEFMC which allows for fishing at 75% F_{MSY}, but requires a lower fishing mortality if the stock is overfished and fishing at 75% of F_{MSY} does not achieve the mandated rebuilding requirements.

If the SSB decreased below the biomass threshold (50% SSB_{MSY}), the F-step HCR used a target F of 70% F_{MSY}, which has recently been applied to some New England groundfish, such as SNE/MA yellowtail flounder and GB winter flounder, as the $F_{rebuild}$ (Fig. 2b). If the SSB never decreased below the biomass threshold or increased to over SSB_{MSY} (rebuilt) after dropping below the biomass threshold, this HCR used a target F of 75% F_{MSY}. This HCR has been applied after National Standard Guidelines were amended in 2016 (NOAA, 2016). These revisions reduced the need to identify an incidental bycatch ABC for overfished stocks.

In some instances, groundfish catch advice has been characterized by large year-to-year changes that have presented challenges to the fishery. The aim of the constrained ramp HCR was to promote rebuilding, optimal yield, and to provide catch stability if stock biomass were to substantially change from year to year (Fig. 2c). This differed from the ramp HCR in that there was a constraint on variation in target catch from year to year, meaning that the current year's catch limit would not change more than 20% from the previous year's catch limit. The threshold of 20% change in catch was in the middle of the range of change in catch thresholds used in HCRs in other fisheries (Apostolaki and Hillary, 2009; Dankel et al., 2016; IOTC, 2016). Further description of HCRs in the MSE framework can be found at https://github.com/lkerr/groundfish-MSE/blob/master/documentation/MSEFramework_HCRs.pdf.

2.7. Performance Metrics

We evaluated tradeoffs among HCRs by comparing performance metrics. Performance metrics included stock, stock assessment, and management performance metrics. Stock performance metrics included OM catch stability, and SSB, F, catch, and recruitment trajectories. Catch stability was measured as interannual variation in catch (IAV; A'Mar et al. 2009). Stock assessment performance metrics included accuracy (measured as relative error; REE) and Mohn's Rho trajectories for SSB and F and accuracy of estimated reference points (F_{MSY} and SSB_{MSY}). REE for SSB and F was the relative error of the terminal estimated assessment values at each year. REE at each year was calculated as:

$$REE = \frac{SSB_{est} - SSB_{true}}{SSB_{true}} * 100$$

with SSB as an example, where □□□□□□ is estimated terminal SSB from the stock assessment, and □□□□□□□ is true or simulated SSB corresponding to the terminal year of the stock assessment. Mohn's Rho values were calculated with a 7-year peel each year in the management procedure period and plotted over time. Mohn's Rho, which provides measures of retrospective inconsistencies, was calculated as:

$$\hat{p}_T = \frac{\sum_{n=1}^{x} \frac{SSB_{est=T-n,T-n} - SSB_{est=T-n,T}}{SSB_{est=T-n,T}}}{x}$$

with SSB as an example, where □̂□ is Mohn's Rho at year T, x is the desired number of assessments with different terminal years to be used in estimating Mohn's Rho (i.e. the number of "peels"), □□□□□□□1,□2 is estimated SSB from the stock assessment at year y1 and year y2. Management performance metrics included true or OM stock status trajectories, the true frequency of overfishing, and the true frequency of being overfished. Metrics were characterized in the short-term (1-5 years), medium-term (6-10 years), and long-term (11-21 years). Code for more complex stock assessment performance metrics can be found at https://github.com/lkerr/groundfish-MSE/tree/master/functions/performance_metrics.

3. Results

3.1. Performance with no misspecification

When there was no misspecification, the ramp and F-step HCRs were most responsive in adjusting catch advice to changes in SSB, whereas the constrained ramp HCR with limited scope for change in year-to-year catch advice responded more slowly to increasing biomass and consequently conserved a higher SSB regardless of the initial status of the stock (Fig. 4). The constrained ramp HCR tended to produce the lowest catch and F in the short-term. The initial status of the stock led to differences in HCR performance with no stock assessment misspecification. For GOM cod, the ramp and constrained ramp HCRs resulted in less time overfished, while the F-step HCR produced the highest catch and F in the short-term and more time overfished (Figs. 3-5). However, in the medium- and long-term, the ramp HCR led to higher catch and more time overfished. For GB haddock (i.e., not overfished), the constrained ramp HCR produced the lowest catch in the short- and medium- term and resulted in a stock that was always rebuilt (Figs. 3-5). As a result, the constrained ramp HCR produced the highest catch in the long-term. The ramp and F-step HCRs resulted in the highest catch and F in the short-term. In the long-term, the stock fluctuated around SSB_{MSY} under the ramp and F-step HCRs.

3.2. Stock assessment performance with a misspecification

Ignoring climate impacts in the stock assessment led to error and retrospective patterns in the assessment (Figs. 6-8). The M and weight-at-age misspecifications resulted in an overestimation of SSB, underestimation of F, and retrospective patterns. For GB haddock, the large REE values in the beginning of the management procedure period (Fig. 6) may be due to a large recruitment event in 2014. With a misspecification, the stock assessment model had difficulty estimating the increase in biomass that results from that large recruitment event. The stock assessment model assumed a weight-at-age that was higher than the true weight-at-age, which in combination with a large recruitment event, resulted in a large overestimation of SSB and underestimation of F. For both GOM cod and GB haddock, there was no error in F_{MSY} with or without a misspecification. However, SSB_{MSY} was consistently underestimated with a misspecification in the case of GOM cod. For GB haddock, SSB_{MSY} was overestimated in the short-term with a small misspecification, and in the medium- and long-term, the constrained ramp HCR resulted in overestimation of SSB_{MSY}. For GB haddock, estimated SSB_{MSY} was higher in the medium- and long-term under the constrained ramp HCR than under the other HCRs, because estimates of recruitment were higher under the constrained ramp HCR. Mean estimated recruitment from the last 20 years was used in the SSB_{MSY} estimations. Assessment performance interacted and was influenced by the performance of alternative HCRs. Under the constrained ramp HCR with a misspecification, error in SSB tended to be the highest regardless of the initial stock status. For GOM cod, the constrained ramp HCR resulted in the highest absolute values of Mohn's Rho. Also, the ramp and F-step HCRs resulted in the highest absolute values of Mohn's Rho for GB haddock. As the misspecification increased, stock assessment performance changed (Figs. 6-8). For GOM cod, SSB_{MSY} was increasingly underestimated as the magnitude of the M misspecification increased. For GB haddock with an increased misspecification, SSB_{MSY} was underestimated under all HCRs. For GOM cod in the long-term, SSB was underestimated, and error in F was near zero. Mohn's Rho values also increased.

3.3. Overall stock and management performance with a misspecification

Ignoring climate impacts in the stock assessment led to changes in performance of HCRs. An overestimation of SSB and underestimation of SSB_{MSY} resulted in overly optimistic estimated stock status. As a result, the prescribed F from the HCR was higher than it should have been, which resulted in catch advice that was also higher. The catch advice was impacted by the prescribed F as well as the estimates of abundance used in the projections, which were overly optimistic. Through prescribed F and projections, the impact of stock assessment error increased throughout the simulation loop. Consequently, the overly optimistic catch advice resulted in more time overfished and overfishing and a lack of rebuilt status in the OM (Figs. 4 and 5). This corresponded with lower SSB and higher Fs (Fig. 3).

As the misspecification increased, stocks experienced more time overfished and with overfishing (Figs. 5). For GOM cod, SSB decreased and F increased with a greater misspecification (Fig. 3). For GB haddock, with an increase in the misspecification, F increased and SSB decreased even more with SSB hovering around the overfished threshold in the long-term.

3.4. Relative stock and management performance with a misspecification

When a misspecification was introduced, relative performance of HCRs was similar compared to the performance with no misspecification (Figs. 3 and 4). For example, regardless of the initial stock status, the constrained ramp HCR tended to result in the highest SSB, which also occurred with no misspecification. For GOM cod, the F-step HCR resulted in the highest catch in the short-term and highest catch stability in the short- and medium-term with and without a misspecification. The constrained ramp HCR resulted in the lowest F and catch in the medium-term and the highest catch stability in the short-term, which occurred with no misspecification as well. For GB haddock, the constrained ramp HCR resulted in the lowest F and catch and highest catch stability in the short-term, which also occurred with no misspecification. Also, the F-step and ramp HCRs produced the highest catch in the medium-term, and the constrained ramp HCR resulted in the highest catch in the long-term with and without a misspecification.

However, the error introduced by the misspecifications led to a few notable differences in relative performance (Fig. 4). For GOM cod, the constrained ramp HCR resulted in the highest catch in the long-term, which did not occur without a misspecification. Also, the ramp HCR resulted in the highest catch in the medium-term.

The error also caused overfishing and overfished status for the 'true' stock, with the frequencies differing among HCRs (Figs. 4 and 5). The constrained ramp HCR also resulted in the least overfishing. For GB haddock, the constrained ramp HCR resulted in the most time above SSB_{MSY}. The constrained ramp HCR also resulted in the most time overfishing in the long-term but the least time overfishing in the short- and medium-term.

Relative performance of HCRs also changed slightly when the misspecification level increased (Figs. 3-5). For GOM cod, F did not get as high under the constrained ramp HCR with an increased misspecification. For GB haddock with an increased misspecification, the ramp HCR led to more time overfished and overfishing in the short-term, and the F-step HCR led to more time overfished in the long-term. Also, more overfishing occurred in the long-term with an increased misspecification for GB haddock.

4. Discussion

Emulating the impacts of climate on stock assessment and fishery management can be challenging for simulation testing, because the mechanisms and processes of climate impacts are uncertain. However, the conditioning of our OMs and misspecification of estimation models produced similar diagnostic problems as experienced for New England groundfish. For example, the magnitude of retrospective inconsistencies was similar to those from recent stock assessments (NEFSC, 2022). The simulations showed that ignoring climate impacts in stock assessment models led to assessment error, that impacted management performance.

4.2. Management performance with a correctly specified stock assessment

With a correctly specified stock assessment (i.e. perfect knowledge of climate impacts on stock dynamics), there were tradeoffs among management alternatives. The constrained ramp HCR can help meet conservation objectives, because the constraint on year-to-year changes in catch maintains greater SSB. The constrained ramp HCR can also help meet economic objectives, because stability in catch can be beneficial for markets. However, the constrained ramp HCR can also have negative impacts on catch in the short-term. If stock size increases substantially, the constrained ramp HCR produces lower catch in the short-term but the highest catch in the long-term relative to the other HCRs due to slow pace of increase in catch advice. However, if the stock were to collapse, we could expect a different response. In this case, the constrained ramp would result in a slow decrease in catch advice, which may not be conservative. In cases where the stock size was above the overfished threshold, the ramp and F-step HCRs typically performed similarly.

For GOM cod (i.e., the overfished stock), responsive HCRs, such as the ramp HCR and the constrained ramp HCR, were able to increase stock size above the overfished threshold at the fastest rate. Responsive HCRs, in which F changes as a function of biomass, can mitigate negative impacts of climate change (Kritzer et al., 2019). Responsive HCRs are necessary to effectively manage highly variable fisheries (Plagányi et al., 2018). When environmental effects are negligible or beneficial, a fixed F HCR can be effective (Kritzer et al., 2019). However, the environmental effects simulated in the GOM cod OM in this study reduced productivity and were not negligible.

4.3. Consequences of ignoring climate impacts on management performance

Comparing correctly specified scenarios to those with stock assessment misspecifications allowed us to understand how unaccounted changes in population dynamics and resulting stock assessment bias can impact HCR performance. Natural mortality and weight-at-age misspecifications resulted in stock assessment error and retrospective patterns. Stock size was overestimated, and F was underestimated in both cases, which led to overly optimistic stock status estimates. With these misspecifications, simulated Mohn's Rho values were as large as they are for some of the current groundfish assessments (NEFSC, 2022). In the 2019 assessments, GOM cod Mohn's rho was 0.52 for SSB and -0.29 for F under the M=0.2 assumption and 0.29 for SSB and -0.16 for F under the increasing M assumption. Georges Bank haddock Mohn's Rho was 0.69 for SSB and -0.44 for F (NEFSC, 2022). Interestingly, the Mohn's Rho values became larger in absolute value as the M misspecification increased but not as the weight-at-age misspecification increased. The degree of bias in the stock assessment performance and retrospective inconsistencies varied among HCRs and misspecifications and did not always coincide in direction. This is similar to other findings that the direction and magnitude of retrospective patterns are not related to true bias (Hurtado-Ferro et al., 2015). Additionally, Kerr et al. (2020) found that relative errors in SSB were sometimes in the opposite direction of the retrospective bias (i.e. relative error was negative while retrospective bias was positive). The difference in error and retrospective patterns among HCRs suggests that error is not only dependent upon whether the stock assessment is correctly specified but also the trends in stock dynamics. Indeed, a lack of retrospective patterns does not mean that there is no data or model inconsistency (Legault, 2009).

Although HCRs affect stock and management performance, stock assessment misspecifications can alter stock assessment performance, which also impacts stock and management performance. Relative HCR performance was generally similar with and without a misspecification. However, there were a few differences. For example, the constrained ramp HCR met the fishery objective of higher catch even more than the other HCRs in the long-term with a misspecification than without a misspecification for GOM cod. Nevertheless, the constrained ramp HCR consistently conserved SSB regardless of the initial stock status and stock assessment specifications. The F-step HCR also met fishery objectives more than the other HCRs in the short- and medium-term regardless of stock assessment specifications.

Although there were some relative differences in HCR performance, HCRs did not compensate for the impact of a stock assessment misspecification. The demand for fisheries management to implement a precautionary approach has led to the development of HCRs (Kvamsdal et al., 2016), and HCRs are supposed to allow for fisheries management to be effective in the face of uncertainty (Walters and Hilborn, 1976). However, although the precautionary target (e.g., 75%F_{MSY}) may help to mitigate imprecision in stock assessment, these HCRs were not robust to the stock assessment bias created by climate impacts. In these simulations, when the stationary assumptions were violated, fisheries management was negatively affected. Stock assessment misspecifications had a larger impact on stock status in the long-term than HCRs. The impact of stock assessment error influences the catch advice through an overly optimistic F prescribed from the HCR and overly optimistic abundance estimates used in the projections. At the end of the management procedure period with a misspecification, SSB was below the overfished threshold and undergoing overfishing for GOM cod and below the rebuilt threshold or fluctuating around the overfished threshold for GB haddock regardless of the control rule. However, even with no misspecification for GOM cod, the stock was under the overfished threshold in the long-term, and under the F-step HCR, the stock did not increase over the overfished threshold at all during the management procedure period. For GOM cod, the increase in M has decreased productivity for the stock, which no longer has the capacity to rebuild, even when near perfect knowledge of stock dynamics is provided to the stock assessment.

4.4. Future studies

Our results suggest that correctly specifying stock assessments should be a priority, especially for stock assessments that exhibit retrospective patterns. However, it is difficult to identify the source of retrospective patterns, although determining the timing of a misspecification that led to retrospective patterns may be possible (Legault, 2009, Kerr et al. 2022). Findings from research that analyzes the relationship between life history parameters and environmental conditions would be helpful for parameterizing stock assessment models. Population dynamics of the focus stock unit should be understood as well as possible so the stock assessment can be updated with assumptions that are more likely to be correct, resulting in a more accurate stock status determination. Waters in the GOM are continuing to warm (NOAA Fisheries, 2021) with impacts on aspects of groundfish population dynamics; therefore, it is important to continue groundfish research. Previous and ongoing research have focused on groundfish population dynamics. For example, Runge et al. (2010) found that forecasts of environmental conditions for recruitment for GOM cod can be developed with coupled physical-biological models. Buckley et al. (2004) found that Atlantic cod and haddock growth increased with temperature until 7°C and then decreased. However, there are still knowledge gaps in our understanding of the impacts of environmental factors, aside from temperature, on New England groundfish (Klein et al., 2017). As our understanding of groundfish population dynamics improves, stock assessment parameters may need to be updated and other stock assessment models may need to be considered. In a changing climate, time-varying parameters are often needed. Our results also suggest that the HCRs evaluated in this study are not robust to stock assessment model misspecifications, and there is a need for identification of HCRs that are robust to stock assessment model misspecifications that drive retrospective patterns.

This MSE framework can also be used for other groundfish fisheries. The simulation framework used in this study is flexible with the capacity to customize OMs for specific groundfish stocks and evaluate relative performance of all aspects of the management procedure: consideration of alternative monitoring systems (e.g., Kerr et al., 2020), the ability to integrate new stock assessment methods, and alternative HCRs.

4.5. Limitations

Although our simulations represent New England groundfish dynamics and issues with stock, stock assessment, and management performance, there are some limitations that should be noted. Other factors may be causing retrospective patterns that were not explored in this study. Also, BRPs can impact how HCRs perform. In this study, BRPs were calculated to be consistent with the current groundfish stock assessments (NEFSC, 2021; 2022). GOM cod BRPs were calculated with M assumed at 0.2. Even though M increased in the GOM cod OM, both true and estimated BRPs were calculated/estimated with a lower M. However, because M remains at 0.4, the GOM cod stock cannot rebuild even with a correctly specified stock assessment. This fundamental shift in resource productivity due to climate caused the stock to no longer have the capacity to rebuild, even with a precautionary approach (NRC, 2016). Additionally, simulated recruitment was similar among HCRs for GB haddock. Recruitment was drawn from empirical distribution functions, so there was no assumed relationship between recruitment and SSB. By simulating no SRR, the stock would be less impacted by HCRs than with a SRR. However, a traditional stock-recruitment model does not explain GB haddock recruitment, which is impacted by other factors other than SSB. Large GB haddock recruitment events could be explained by autumn phytoplankton blooms the year prior to spawning (Leaf and Friedland, 2014; Friedland et al., 2015). Forecasts of autumn blooms are not available, so a SRR that includes autumn blooms could not be used in the simulations in the management procedure period. However, recruitment was related to SSB and temperature in the GOM cod simulations.

There are also additional HCR forms and adjustments to the features of the HCRs evaluated in this study that could be worthwhile exploring in the future based on the desired outcomes of groundfish fishery management (i.e., management objectives). For example, a constant F HCR was not explored in this analysis. Kerr et al. (2020) did simulate a constant F HCR in the context of catch misreporting scenarios that could be informative for decision making regarding this HCR’s performance. Another limitation of this analysis was that technical interactions were not simulated. For some stocks, the groundfish fishery harvests considerably less than the annual catch limit (ACL) due to technical interactions of the mixed-stock fishery (i.e. choke species issues; Cadrin, 2016). However, if technical interactions were included in the haddock scenarios, HCR performance would be difficult to evaluate if the catch was a small percentage of the ACL. Additionally, the constrained ramp HCR may perform differently if the stock was collapsing. Since the constrained ramp HCR has a 20% limit in change in catch, the catch advice from this HCR could be aggressive if the stock were rapidly and consistently declining. The simulations in this study do not include a stock where the biomass is rapidly and consistently declining in the management procedure period. Also, there is no depensation or strong autocorrelative recruitment simulated, which may result in a collapsed stock.

In addition, the OMs are flexible and can be further tuned to represent additional complexity and variability in groundfish dynamics and operation of groundfish fisheries. For example, declining weight-at-age and density-dependent growth are evident for GB haddock (NEFSC, 2022; Wang et al., 2021), but this was not included during the management procedure period for haddock scenarios. Declining weight-at-age was only included during the historical period. While cod M was constant at 0.4 in the management procedure period, recruitment was negatively impacted by increasing temperatures. However, we can expect similar climate impacts on natural mortality of cod and haddock size in the future. The warming and predation trends in the region are expected to continue. Cod natural mortality and haddock size are positively and negatively related to temperature, respectively (Pershing et al. 2015; Brodziak and Link 2008). Seal predation is also expected to continue, because seal populations are being maintained by the marine mammal protection act. Given these relationships, we could expect cod natural mortality to continue to increase and haddock size to continue to decrease. As a result, the management impacts measured in the current study could likely be a conservative assessment of the long-term management performance. However, since species will occupy habitats that maximize their fitness, cod and haddock spatial distributions are expected to change. Cod natural mortality cannot increase infinitely, and haddock size cannot decrease to zero. Indeed, GOM cod have shifted to greater depths but no significant change in haddock distribution off the Northeast shelf has been observed (Nye et al. 2009). However, haddock size is not only impacted by temperature but also density-dependent mechanisms (Brodziak and Link 2008). Nevertheless, including climate impacts in stock assessments will be increasingly important.

Additionally, unbiased implementation of HCRs was assumed (i.e., no bias in catch observations but some random error), however there is some evidence of an observer effect in catch reporting in the groundfish fishery that could introduce bias (Demarest, 2019; McNamee et al., 2019; Nitschke, 2019). Also, catchability may be changing due to changing species distributions, which may move out of current survey areas. In this study, catchability was assumed constant in the OMs and stock assessment models. However, ASAP has the ability to assume a varying catchability.

4.6. Management implications

Poor stock assessment performance caused by ignoring climate impacts can have negative impacts on management. Tradeoffs among HCRs exist, but addressing stock assessment bias may be the more immediate and important task compared to identifying an optimal HCR in meeting management objectives. If stock assessment models with misspecifications that have a large impact are not able to be corrected or corrected in a timely manner, there is likely to be adverse impacts on management. This study suggests that no commonly applied HCRs, even precautionary HCRs, perform well with large misspecifications. If significant retrospective patterns exist, the stock status determination is likely to be biased. The resulting F from the HCR will then be inappropriate. However, certain HCRs can help mitigate some of the impacts of bias and help to implement a precautionary approach. However, typical HCRs employed for groundfish stock are not sufficiently robust to overcome the level of bias in many of the current stock assessments. Large retrospective patterns can lead to inappropriate fishery management, possibly causing a depleted stock even though stock assessments suggest that the stock is above target levels (Legault, 2009). Legault (2009) recommends that if strong retrospective patterns are present, the stock assessment model as the basis for management advice should be rejected. In this case, an index-based method may perform better. Additionally, future research could identify an HCR that is robust to stock assessment misspecifications. In the short term, improvements in the accuracy of stock assessments and stock status determinations will provide the greatest scope for improvement in New England groundfish management..

Acknowledgements

We thank the Groundfish Plan Development Team, Groundfish Advisory Panel, Recreational Advisory Panel, Groundfish Committee, Scientific and Statistical Committee, MSE experts at the Northeast Fisheries Science Center, and the New England Fisheries Management Council for providing feedback regarding the plan for HCR testing, particularly Jamie Cournane, Robin Frede, and Paul Nitschke. Funding: the New England Fishery Management Council and the NOAA Coastal and Ocean Climate Applications Program. We also thank Andy Pershing, Ashley Weston, Gavin Fay, Jonathan Cummings, Sarah Gaichas, Min-Yang Lee, and Anna Birkenbach for development of the MSE modeling architecture, which was supported by the NOAA COCA program and the NOAA SK program.

References

A'mar, Z. T., Punt, A. E., Dorn, M. W., 2009. The impact of regime shifts on the performance of management strategies for the Gulf of Alaska walleye pollock (Theragra chalcogramma) fishery. Canadian Journal of Fisheries and Aquatic Sciences, 66(12), 2222-2242.

Apostolaki, P., Hillary, R, 2009. Harvest control rules in the context of fishery-independent management of fish stocks. Aquatic Living Resources, 22(2), 217-224. https://doi.org/10.1051/alr/2009022

Brodziak, J., Cadrin, S. X., Legault, C. M., Murawski, S. A., 2008. Goals and strategies for rebuilding New England groundfish stocks. Fisheries Research, 94(3), 355-366. https://doi.org/10.1016/j.fishres.2008.03.008

Brodziak, J., Link, J., 2008. The incredible shrinking Georges Bank haddock (*Melanogrammus aeglefinus*). In 24th Wakefield Symposium Resiliency of Gadid Stocks to Fishing and Climate Change, AK-SG-08-01, pp. 141-160. https://doi.org/10.4027/rgsfcc.2008.08

Brooks, E. N., Legault, C. M., 2016. Retrospective forecasting—evaluating performance of stock projections for New England groundfish stocks. Canadian Journal of Fisheries and Aquatic Sciences, 73(6), 935-950. https://doi.org/10.1139/cjfas-2015-0163

Buckley, L. J., Caldarone, E. M., Lough, R.G., 2004. Optimum temperature and food-limited growth of larval Atlantic cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) on Georges Bank. Fisheries Oceanography, 13(2), 134-140. https://doi.org/10.1046/j.1365-2419.2003.00278.x

Cadrin, S.X., 2016. Management Strategies for Mixed-Species Commercial, Recreational, 01 and Subsistence Fisheries, in: Quinn II, T.J., Armstrong, J.L., Baker, M., Heifetz, J.D., Witherell, D. (Eds.), Assessing and Managing Data-Limited Fish Stocks. Alaska Sea Grant, University of Alaska Fairbanks. http://doi.org/10.4027/amdlfs.2016.01

D'Abramo, L. R., 1979. Dietary fatty acid and temperature effects on the productivity of the cladoceran, *Moina macrocopa*. Biological Bulletin, 157, 234–248. https://doi.org/10.2307/1541051

Dankel, D. J., Vølstad, J. H., Aanes, S., 2016. Communicating uncertainty in quota advice: A case for confidence interval harvest control rules (CI-HCRs) for fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 73(2), 309-317. https://doi.org/10.1139/cjfas-2015-0078

Demarest, C., 2019. Evaluating the Observer Effect for the Northeast U.S. Groundfish Fishery. Groundfish Plan Development Team Document. 27 pp.

Deroba, J.J., 2014. Evaluating the consequences of adjusting fish stock assessment estimates of biomass for retrospective patterns using Mohn's Rho. North American Journal of Fisheries Management, 34, 380-390. https://doi.org/10.1080/02755947.2014.882452

Dichmont, C. M., Deng, A., Punt, A. E., Ellis, N., Venables, W. N., Kompas, T., Ye, Y., Zhou, S., Bishop, J., 2008. Beyond biological performance measures in management strategy evaluation: bringing in economics and the effects of trawling on the benthos. Fisheries Research, 94(3), 238-250. https://doi.org/10.1016/j.fishres.2008.05.007

Drinkwater, K. F., Beaugrand, G., Kaeriyama, M., Kim, S., Ottersen, G., Perry, R. I., Pörtner, H.O., Polovina, J.J., Takasuka, A., 2009. On the processes linking climate to ecosystem changes. Journal of Marine Systems, 79(3-4), 374-388. https://doi.org/10.1016/j.jmarsys.2008.12.014

Indian Ocean Tuna Commission (IOTC), 2016. Resolution 16/02 on harvest control rules for skipjack tuna in the IOTC area of competence. Food and Agriculture Organization of the United Nations. 6 pp.

Fogarty, M., Incze, L., Hayhoe, K., Mountain, D., Manning, J., 2008. Potential climate change impacts on Atlantic cod (*Gadus morhua*) off the northeastern USA. Mitigation and Adaptation Strategies for Global Change, 13(5), 453-466. https://doi.org/10.1007/s11027-007-9131-4

Friedland, K. D., Leaf, R. T., Kristiansen, T., Large, S. I., 2015. Layered effects of parental condition and larval survival on the recruitment of neighboring haddock stocks. Canadian Journal of Fisheries and Aquatic Sciences, 72(11), 1672-1681. https://doi.org/10.1139/cjfas-2015-0084

Hare, J.A., Morrison, W.E., Nelson, M.W., Stachura, M.M., Teeters, E.J., Griffis, R.B., Alexander, M.A., Scott, J.D., Alade, L., Bell, R.J., Chute, A.S., Curti, K.L., Curtis, T.H., Kircheis, D., Kocik, J.F., Lucey, S.M., McCandless, C.T., Milke, L.M., Richardson, D.E., Robillard, E., Walsh, H.J., McManus, C., Marancik, K.E., Griswold, C.A., 2016. A vulnerability assessment of fish and invertebrates to climate change on the Northeast US Continental Shelf. PloS One, 11, e0146756. https://doi.org/10.1371/journal.pone.0146756

Hamazaki, T., Zheng, J., 2012. Norton sound red king crab stock assessment for the fishing year 2012/13. Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. North Pacific Fishery Management Council, Anchorage, AK, USA. 108 pp.

[dataset] Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., Smith, T., Zhang, H.M., 2020. Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1. Journal of Climate*,* 34, 2923-2939. https://doi.org/10.1175/JCLI-D-20-0166.1

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A. E., 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science, 72(1), 99-110. https://doi.org/10.1093/icesjms/fsu198

International Council for Exploration of the Seas (ICES), 2020. Workshop on Catch Forecast from Biased Assessments (WKFORBIAS; outputs from 2019 meeting). ICES Scientific Reports. 2:28. 38 pp. https://doi.org/10.17895/ices.pub.5997

Kavanaugh, M. T., Rheuban, J. E., Luis, K. M. A., Doney, S. C., 2017. Thirty-three years of ocean benthic warming along the U.S. northeast continental shelf and slope: Patterns,

drivers, and ecological consequences. Journal of Geophysical Research, 122, 9399–9414. https://doi.org/10.1002/2017JC012953
Kerr, L.A., Barajas, M., Wiedenmann, J., 2022. Coherence and potential drivers of stock assessment uncertainty in Northeast US groundfish stocks. ICES Journal of Marine Science, fsac140. https://doi.org/10.1093/icesjms/fsac140
Kerr, L.A., Weston, A.E., Mazur, M.D., Cadrin, S.X., 2020. Evaluating the impact of inaccurate catch information on New England groundfish management. Technical report. 62 pp.
Klein, E. S., Smith, S. L., Kritzer, J. P., 2017. Effects of climate change on four New England groundfish species. Reviews in Fish Biology and Fisheries, 27(2), 317-338. https://doi.org/10.1007/s11160-016-9444-z
Kritzer, J. P., Costello, C., Mangin, T., Smith, S. L., 2019. Responsive harvest control rules provide inherent resilience to adverse effects of climate change and scientific uncertainty. ICES Journal of Marine Science, 76(6), 1424-1435. https://doi.org/10.1093/icesjms/fsz038
Kvamsdal, S. F., Eide, A., Ekerhovd, N. A., Enberg, K., Gudmundsdottir, A., Hoel, A. H., Mills, K.E., Mueter, F.J., Ravn-Jonsen, L., Sandal, L.K., Stiansen, J.E., Vestergaard, N., 2016. Harvest control rules in modern fisheries management. Elementa: Science of the Anthropocene, 4: 000114. https://doi.org/10.12952/journal.elementa.000114
Leaf, R. T., Friedland, K. D., 2014. Autumn bloom phenology and magnitude influence haddock recruitment on Georges Bank. ICES Journal of Marine Science, 71(8), 2017-2025. https://doi.org/10.1093/icesjms/fsu076
Legault, C.M.. 2009. Report of the Retrospective Working Group, January 14-16, 2008, Woods Hole, Massachusetts. US Department of Commerce, Northeast Fisheries Science Center Reference Document. 09-01. 30 pp.
Legault, C. M., Palmer, M. C., 2016. In what direction should the fishing mortality target change when natural mortality increases within an assessment? Canadian Journal of Fisheries and Aquatic Sciences, 73(3), 349-357. https://doi.org/10.1139/cjfas-2015-0232
Legault, C. M., Restrepo, V. R., 1998. A Flexible Forward Age-Structured Assessment Program. ICCAT Working Document Scrs/98/58, January 1999. 15 pp.
Lorenzen, K., 2022. Size- and age-dependent natural mortality in fish populations: Biology, models, implications, and a generalized length-inverse mortality paradigm. Fisheries Research, 255, 106454. https://doi.org/10.1016/j.fishres.2022.106454
McNamee, J., Holland, D., Kerr, L., Uchida, H., 2019. Peer Review Report for the Groundfish Plan Development Team Analyses of Groundfish Monitoring. New England Fishery Management Council, Providence, Rhode Island. 24 pp.
Mid Atlantic Fishery Management Council (MAFMC), 2020. Mid-Atlantic Fishery Management Council Scientific and Statistical Committee OFL CV guidance document. MAFMC. 11 pp.
Mohn, R., 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES Journal of Marine Science, 56(4), 473-488. https://doi.org/10.1006/jmsc.1999.0481
New England Fisheries Management Council (NEFMC), 2009. Final Amendment 16 to the Northeast Multispecies Fishery Management Plan including a Environmental Impact

Statement and regulatory flexibility analysis. NEFMC. 931 pp.Northeast Fisheries Science Center (NEFSC), 2020. 2020 Management Track Peer Review Panel Report. NEFSC. 32 pp.

Northeast Fisheries Science Center (NEFSC), 2021. Gulf of Maine Atlantic Cod 2021 Update Assessment Report. NEFSC Ref. Doc. 10 pp.

Northeast Fisheries Science Center (NEFSC), 2022. Operational Assessment of 14 Northeast Groundfish Stocks, Updated Through 2018. NEFSC Ref. Doc. 217 pp.

Nitschke, P., 2019. Comparison of sector vessel landings effort ratios between observed and unobserved trips by gear and broad stock area. Groundfish Plan Development Team Document. 10 pp.

National Oceanographic and Atmospheric Administration (NOAA) Fisheries, 2016. Magnuson-Stevens Act Provisions; National Standard Guidelines. Federal Register, 81(201), 71858-71904.

National Oceanographic and Atmospheric Administration (NOAA) Fisheries, 2021. 2021 State of the Ecosystem New England. NOAA. 43 pp.

National Research Council (NRC), 2016. Evaluating the Effectiveness of Fish Stock Rebuilding Plans in the United States. National Academies Press, Washington, DC. https://doi.org/10.17226/18488

Nye, J.A., Link, J.S., Hare, J.A., Overholtz, W.J. 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Marine Ecology Progress Series, 393, 111-129.

Pershing, A. J., Alexander, M. A., Hernandez, C. M., Kerr, L. A., Le Bris, A., Mills, K. E., Nye, J.A., Record, N.R., Scannell, H.A., Scott, J.D., Sherwood, G.D., Thomas, A. C., 2015. Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. Science, 350(6262), 809–812. https://doi.org/10.1126/science.aac9819

Pershing, A. J., Alexander, M. A., Hernandez, C. M., Kerr, L. A., Le Bris, A., Mills, K. E., Nye, J.A., Record, N.R., Scannell, H.A., Scott, J.D., Sherwood, G.D., Thomas, A. C., 2016. Response to Comments on “Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery”. Science, 352(6284), 423-423. https://doi.org/10.1126/science.aae0463

Plagányi, É., Deng, R. A., Campbell, R. A., Dennis, D., Hutton, T., Haywood, M., Tonks, M., 2018. Evaluating an empirical harvest control rule for the Torres Strait *Panulirus ornatus* tropical rock lobster fishery. Bulletin of Marine Science, 94(3), 1095-1120. https://doi.org/10.5343/bms.2017.1101

Pörtner, H. O., Peck, M. A., 2010. Climate change effects on fishes and fisheries: towards a cause-and-effect understanding. Journal of fish biology, 77(8), 1745-1779. https://doi.org/10.1111/j.1095-8649.2010.02783.x

Punt, A. E., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A., Haddon, M., 2016. Management strategy evaluation: best practices. Fish and Fisheries, 17(2), 303-334. https://doi.org/10.1111/faf.12104

Runge, J. A., Kovach, A. I., Churchill, J. H., Kerr, L. A., Morrison, J. R., Beardsley, R. C., Berlinsky, D.L., Chen, C., Cadrin, S.X., Davis, C.S., Ford, K.H., Grabowski, J.H., Howell, W.H., Ji, R., Jones, R.J., Pershing, A.J., Record, N.R., Thomas, A.C., Sherwood, G.D., Tallack, S.M.L., Townsend, D. W., 2010. Understanding climate impacts on recruitment and spatial dynamics of Atlantic cod in the Gulf of Maine: Integration of observations and modeling. Progress in Oceanography, 87(1-4), 251-263. https://doi.org/10.1016/j.pocean.2010.09.016

Skern-Mauritzen, M., Ottersen, G., Handegard, N. O., Huse, G., Dingsør, G. E., Stenseth, N. C., Kjesbu, O. S., 2016. Ecosystem processes are rarely included in tactical fisheries management. Fish and Fisheries, 17(1), 165-175. https://doi.org/10.1111/faf.12111

Vetter, E. F., 1988. Estimation of natural mortality in fish stocks: a review. Fishery Bulletin, 86(1), 25-43.

Walters, C. J., Hilborn, R. 1976. Adaptive control of fishing systems. Journal of the Fisheries Board of Canada, 33(1), 145-159. https://doi.org/10.1139/f76-017

Wang, Y., Gharouni, A., Friedland, K. D., Melrose, D. C., 2021. Effect of environmental factors and density-dependence on somatic growth of Eastern Georges Bank haddock (*Melanogrammus aeglefinus*). Fisheries Research, 240, 105954. https://doi.org/10.1016/j.fishres.2021.105954

Wiedenmann, J., Jensen, O. P., 2018. Uncertainty in stock assessment estimates for New England groundfish and its impact on achieving target harvest rates. Canadian Journal of Fisheries and Aquatic Sciences, 75(3), 342-356. https://doi.org/10.1139/cjfas-2016-0484

Whiteley, N. M., Robertson, R. F., Meagor, J., El Haj, A. J., Taylor, E. W., 2001. Protein synthesis and specific dynamic action in crustaceans: effects of temperature. Comparative Biochemistry and Physiology Part A, 128, 595–606. https://doi.org/10.1016/s1095-6433(00)00337-8

Figure Captions

Figure 1. The management strategy evaluation framework used in this project.

Figure 2. Alternative harvest control rules (HCRs): a) ramp, b) F-step, and c) constrained ramp harvest control rules. The constrained ramp HCR (c) is the ramp HCR (a) but with a 20% constraint on change in catch from year to year.

Figure 3. True operating model median spawning stock biomass (SSB), fishing mortality (F), and catch with 95% confidence intervals for a) Gulf of Maine cod and b) Georges Bank haddock with no stock assessment model misspecification (none), a small misspecification (intermediate), and a large misspecification (high) from 2019 to 2040.

Color should be used for Figure 3 in print.

Figure 4. Harvest control rule (HCR) performance for a) Gulf of Maine cod and b) Georges Bank haddock with no stock assessment model misspecification (none), a small misspecification (intermediate), and a large misspecification (large) in the short- (1-5 years), and long-term (11-21 years). Some metrics (SSB, catch stability and catch) are

standardized to the maximum value for each metric attained by the different HCRs by dividing the values by the maximum value across HCRs. Frequency not overfished and frequency not overfishing are automatically on a scale from 0 to 1. Metrics are also equally weighted. Spawning stock biomass (SSB) and catch are median SSB and catch for the time period. Black, orange, and blue lines and points represent the ramp, F-step, and constrained ramp harvest control rules, respectively. Sometimes the black and orange lines do not appear due to overlap.

Color should be used for Figure 4 in print.

Figure 5. True stock status trajectories (ratio of fishing mortality to the fishing mortality reference point (F/F_{MSY}) versus ratio of spawning stock biomass to the spawning stock biomass reference point (SSB/SSB_{MSY})) for a) Gulf of Maine cod with no stock assessment model misspecification (none), natural mortality incorrectly assumed constant at 0.3 (intermediate), and natural mortality incorrectly assumed constant at 0.2 (high) and b) Georges Bank haddock with no stock assessment model misspecification (none), a small weight-at-age stock assessment misspecification (intermediate), and a large weight-at-age stock assessment misspecification (high). The dashed line represents the overfished threshold. Black, orange, and blue lines and points represent the ramp, F-step, and constrained ramp harvest control rules, respectively. In the red quadrant, the stock is not rebuilt and undergoing overfishing. In the lower left yellow quadrant, the stock is not rebuilt. In the upper right yellow quadrant, overfishing is occurring. In the green quadrant, the stock is in good status.

Color should be used for Figure 5 in print.

Figure 6. Median percent relative error (REE) in terminal estimated spawning stock biomass (SSB) and fishing mortality (F) with 95% confidence intervals for a) Gulf of Maine cod with no stock assessment model misspecification (none), natural mortality incorrectly assumed constant at 0.3 (intermediate; smaller misspecification), and natural mortality incorrectly assumed constant at 0.2 (high; larger misspecification) and b) Georges Bank haddock with no stock assessment model misspecification (none), a small weight-at-age stock assessment misspecification (intermediate), and a large weight-at-age stock assessment misspecification (high). Black, orange, and blue lines represent the ramp, F-step, and constrained ramp harvest control rules, respectively.

Color should be used for Figure 6 in print.

Figure 7. Median Mohn's Rho values for spawning stock biomass (SSB) and fishing mortality (F) with 95% confidence intervals for a) Gulf of Maine cod with no stock assessment model misspecification (none), natural mortality incorrectly assumed constant and 0.3 (intermediate), and natural mortality incorrectly assumed constant at 0.2 (high) and b) Georges Bank haddock with no stock assessment model misspecification (none), a small weight-at-age misspecification (intermediate), and a large weight-at-age misspecification (high). Black, orange, and blue lines represent the ramp, F-step, and constrained ramp harvest control rules, respectively.

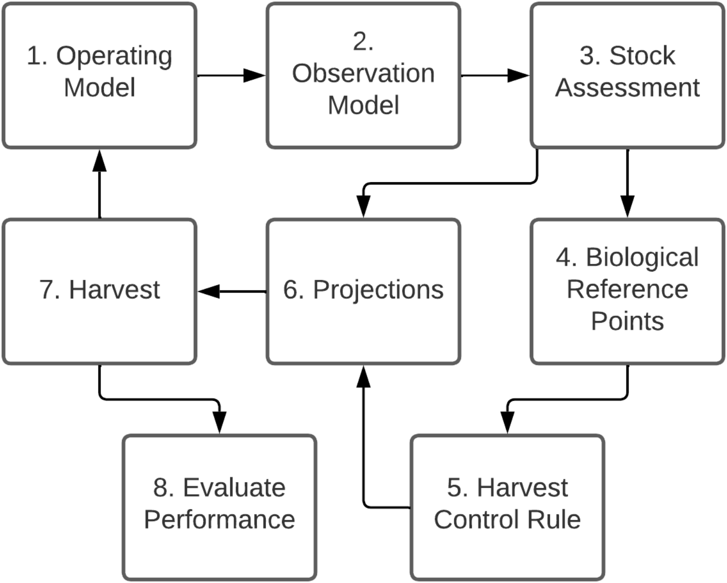
Color should be used for Figure 7 in print.

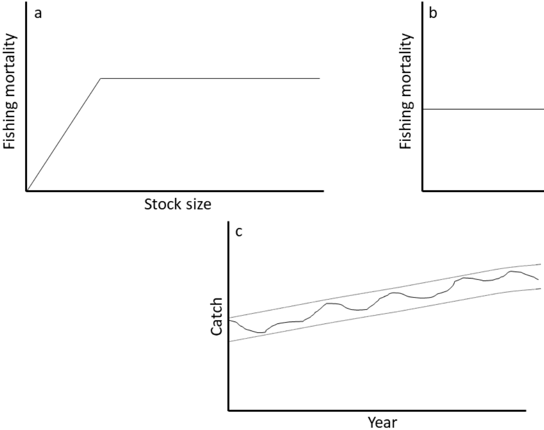
Figure 8. Median ratios of estimated to true stock biomass reference point (SSB_{MSY}) for a) Gulf of Maine cod with no stock assessment model misspecification (none), natural mortality incorrectly assumed constant at 0.3 (intermediate), and natural mortality incorrectly assumed constant at 0.2 (high) and b) Georges Bank haddock with no stock assessment model misspecification (none), a small weight-at-age stock assessment misspecification (intermediate), and a large weight-at-age stock assessment misspecification (high) in the short- (1-5 years), medium- (6-10 years), and long-term (11-21 years).

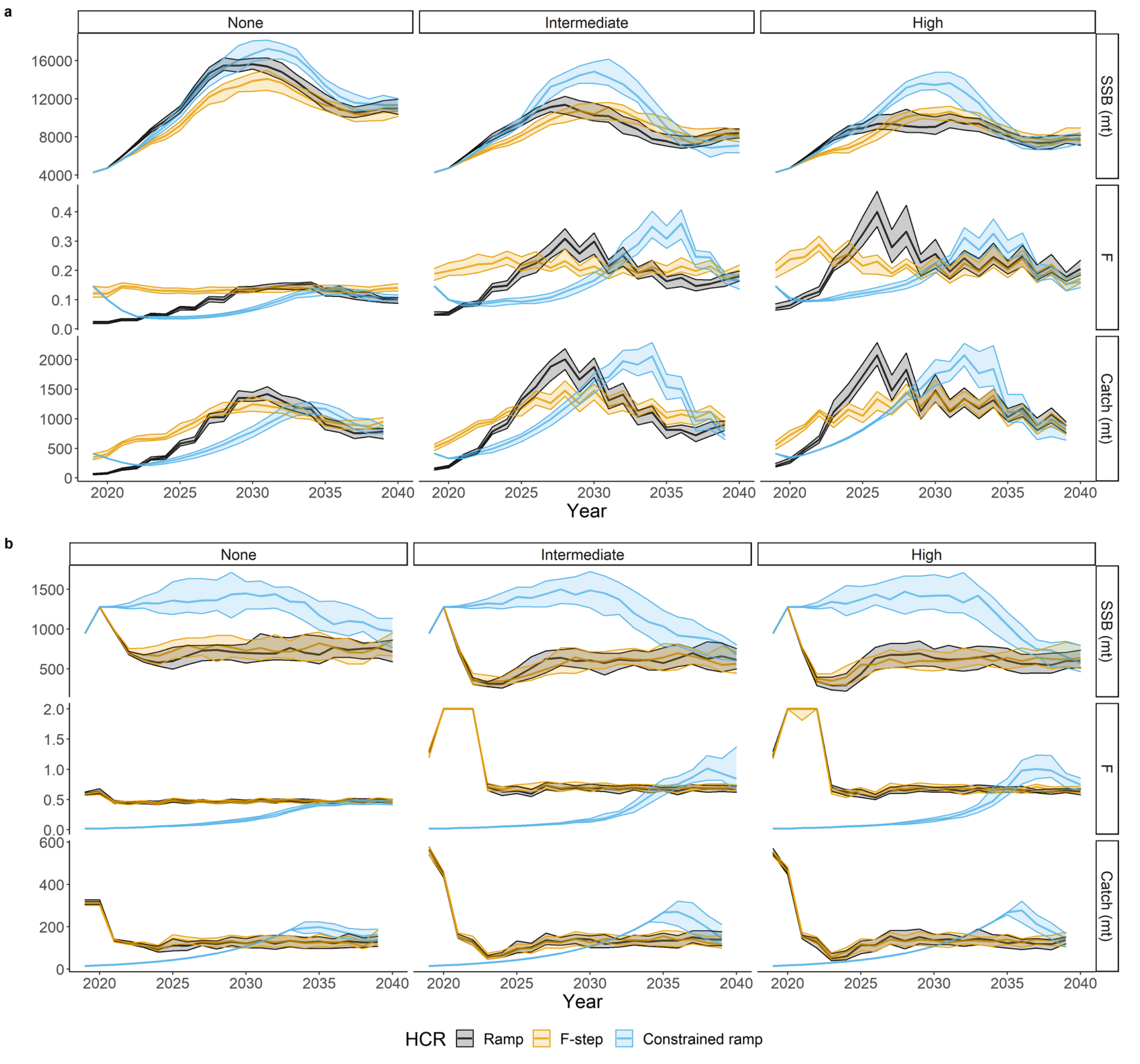
Color should be used for Figure 8 in print.

Data availability statement

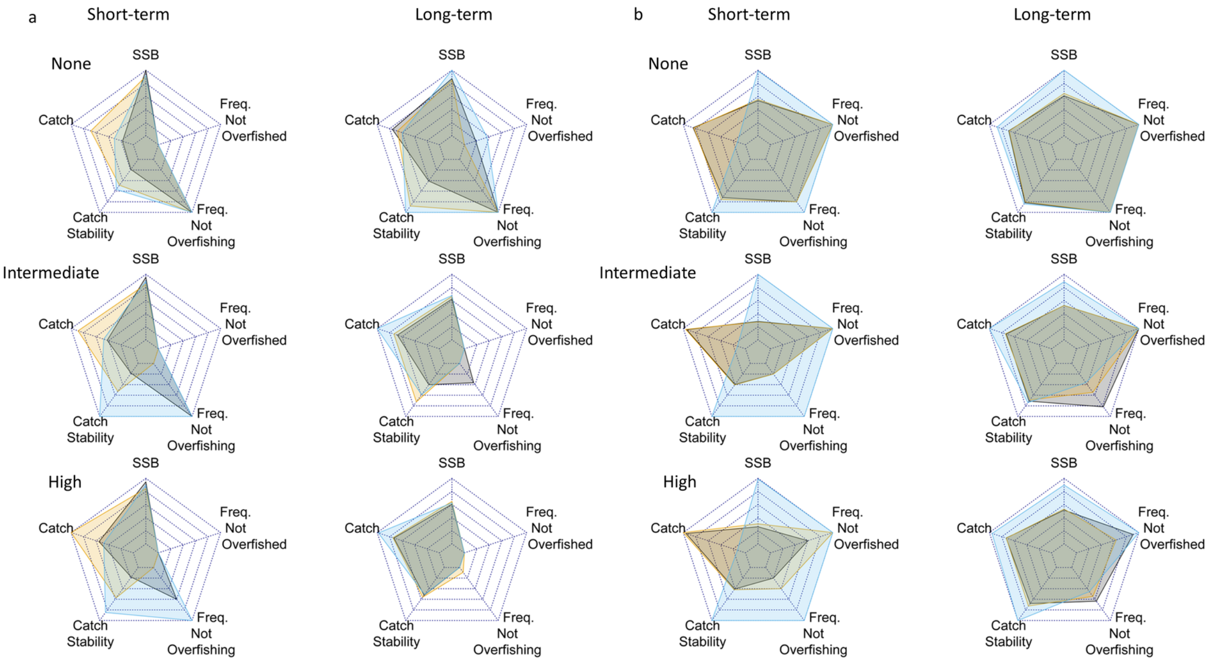
All code and instructions for the management strategy evaluation framework used in this study can be found at the following Github repository: https://github.com/lkerr/groundfish-MSE. Instructions on how to specify the operating models (OMs) and management procedures can be found on the Wiki page: https://github.com/lkerr/groundfish-MSE/wiki. This study was conducted using the 'MDM_Misspecifications' branch. Each scenario uses different specifications of OMs and/or management procedures. Parameters and historical time series from the Gulf of Maine cod and Georges Bank haddock assessments can be found at the Northeast Fisheries Science Center Stock Assessment Support Information: https://apps-nefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php. Gulf of Maine sea surface temperature data can be found at the Optimum Interpolation Sea Surface Temperature (OISST) database: https://www.ncei.noaa.gov/products/optimum-interpolation-sst.


1. Operating Model
2. Observation Model
3. Stock Assessment
4. Biological Reference Points
5. Harvest Control Rule
6. Projections
7. Harvest
8. Evaluate Performance

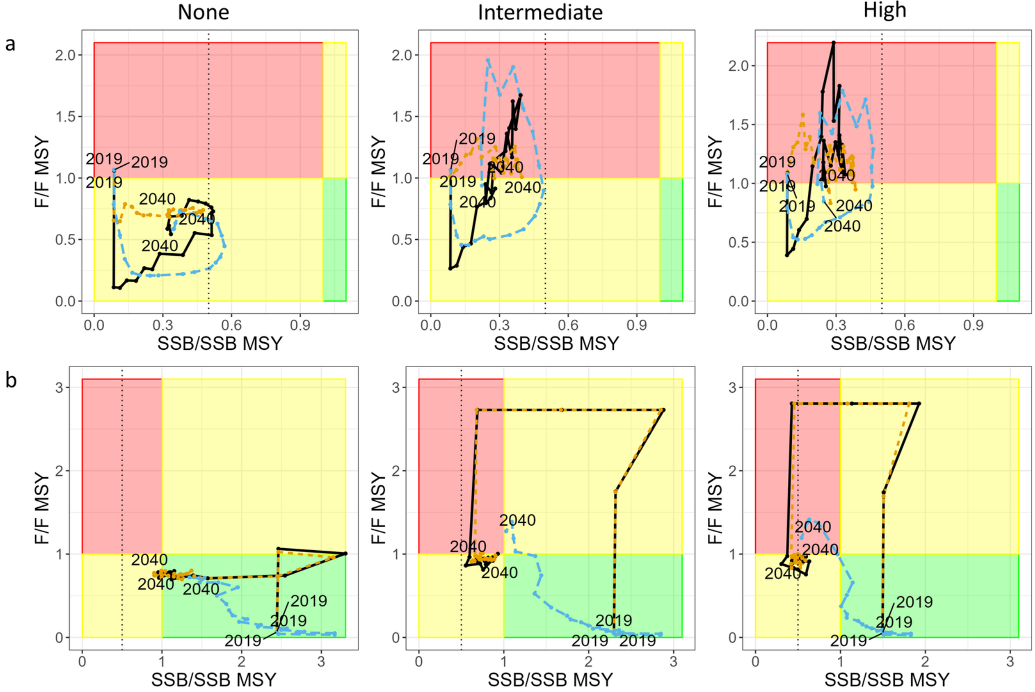

a
Fishing mortality
Stock size
b
Fishing mortality
Stock size
c
Catch
Year



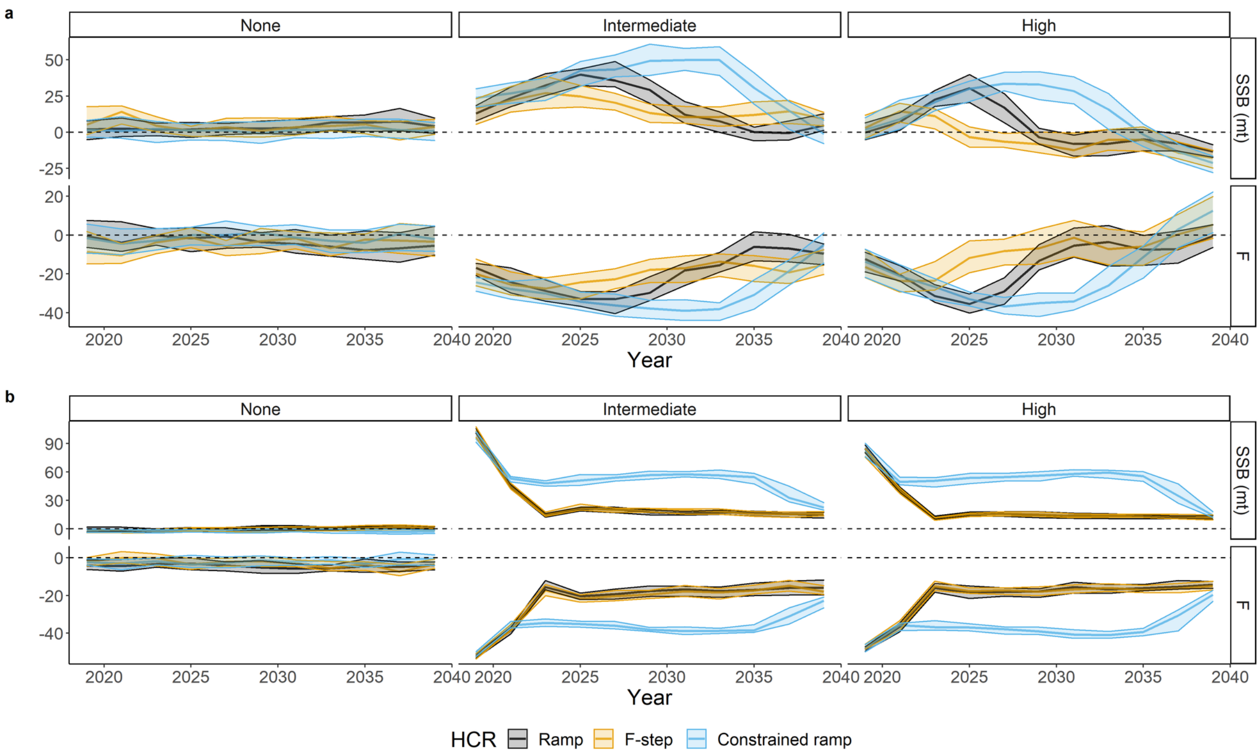
a
None
Intermediate
High
SSB (mt)
F
Catch (mt)
16000
12000
8000
4000
0.4
0.3
0.2
0.1
0.0
2000
1500
1000
500
0
2020
2025
2030
2035
2040
Year
b
None
Intermediate
High
SSB (mt)
F
Catch (mt)
1500
1000
500
2.0
1.5
1.0
0.5
0.0
600
400
200
0
2020
2025
2030
2035
2040
Year
HCR
Ramp
F-step
Constrained ramp

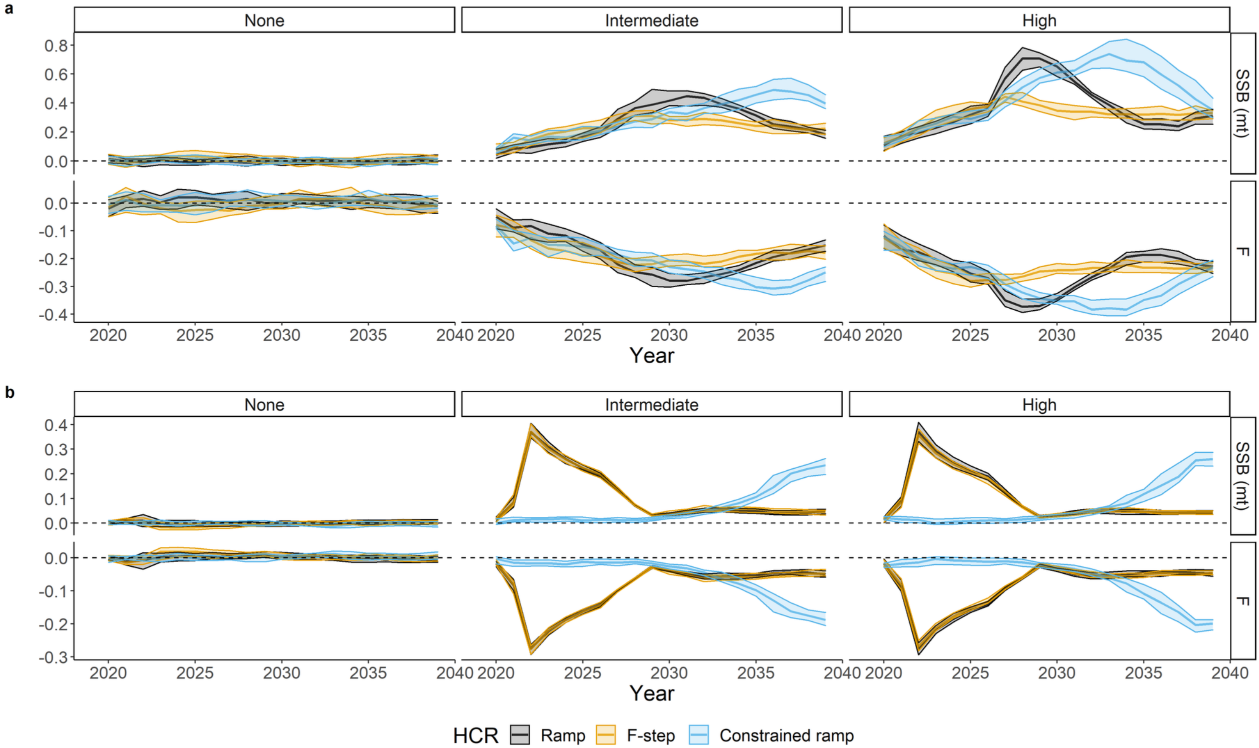


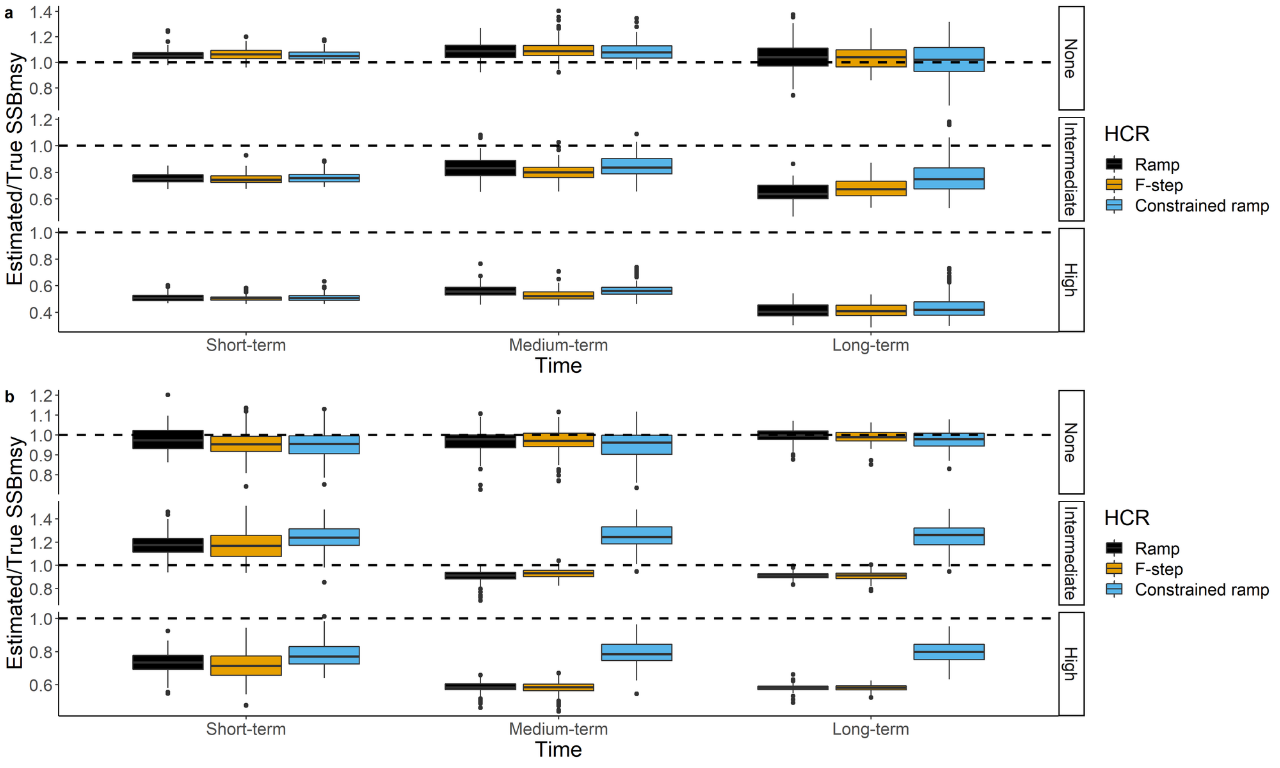
a
Short-term
Long-term
b
Short-term
Long-term
None
Intermediate
High
SSB
Freq. Not Overfished
Freq. Not Overfishing
Catch Stability
Catch



None
Intermediate
High
a
b
F/F MSY
SSB/SSB MSY
2019
2040


a
None
Intermediate
High
SSB (mt)
F
50
25
0
-25
20
0
-20
-40
2020
2025
2030
2035
2040
Year
b
None
Intermediate
High
SSB (mt)
F
90
60
30
0
0
-20
-40
2020
2025
2030
2035
2040
Year
HCR
Ramp
F-step
Constrained ramp


a
None
Intermediate
High
SSB (mt)
F
0.8
0.6
0.4
0.2
0.0
-0.1
-0.2
-0.3
-0.4
2020
2025
2030
2035
2040
Year
b
0.3
HCR
Ramp
F-step
Constrained ramp



a
b
Estimated/True SSBmsy
None
Intermediate
High
HCR
Ramp
F-step
Constrained ramp
Short-term
Medium-term
Long-term
Time

Table 1. Georges Bank haddock weight-at-ages used in the operating model and stock assessments.

	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9+
Weight-at-age in the operating model (2019-2040)	0.00018	0.00043	0.00069	0.00089	0.00109	0.00125	0.00149	0.00158	0.00178
Weight-at-age in stock assessment (high and constant scenario)	0.00033	0.00080	0.00128	0.00166	0.00192	0.00220	0.00252	0.00275	0.00313
Weight-at-age in stock assessment (intermediate and constant scenario)	0.00019	0.00055	0.00090	0.00114	0.00126	0.00137	0.00152	0.00164	0.00188