Executive Summary of Evaluation of Alternative Harvest Control Rules for New England Groundfish

Mackenzie Mazur[1], Steve Cadrin[2], Jerelle Jesse[1], and Lisa Kerr[1]

[1]Gulf of Maine Research Institute, 350 Commercial Street, Portland, ME 04101, lkerr@gmri.org, 207-228-1639
[2]School for Marine Science & Technology, 836 Rodney French Boulevard, New Bedford, MA 02744

June 7 2021 DRAFT

NOTE: This is a working draft and the synthesis of results and conclusions are underway and subject to change as the report is finalized.

Rationale

Management of New England groundfish complex is challenging because of the multispecies nature of the fishery and aspects of groundfish population dynamics that are not completely understood. The majority of groundfish stocks that have analytical assessments exhibit a similar 'retrospective pattern' whereby the addition of new data results in reduced estimates of past stock size and increased estimates of fishing mortality. Retrospective patterns represent a large source of uncertainty and pose challenges in the classification of Northeast groundfish stock status and determination of catch advice.

The determination of the Acceptable Biological Catch (ABC) for each groundfish stock is currently based on the New England Fishery Management Council's harvest control rule (HCR), also known as the ABC control rule. In hindsight it has been recognized that application of the groundfish HCRs did not always prevent overfishing (Brooks & Legault, 2016; Wiedenmann & Jensen, 2018). The accuracy of the stock assessment, retrospective patterns, and the quality of projections have been key contributors to these issues with management performance. In response to the issues raised regarding the current ABC control rule, the NEFMC initiated an evaluation of the performance of the current HCR and possible alternatives through simulation testing.

Goal

The goal of this analysis was to evaluate the performance of alternative HCRs for New England groundfish stocks using management strategy evaluation (MSE). We structured scenarios to address a series of research questions:

a) How do alternative HCRs perform when a stock is overfished?
b) How do alternative HCRs perform when a stock is not overfished?
c) How do alternative HCRs perform when there is a stock assessment misspecification and retrospective patterns?
d) When retrospective patterns exist, do retrospective adjustments result in better performance than no retrospective adjustments?

Approach

MSE, a general framework aimed at simulation testing management strategies, was used to evaluate the performance of alternative harvest control rules (HCRs) for a suite of New England groundfish species. In the MSE framework, 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. Through an observation model, simulated trawl survey data and catch data were generated with plausible random error to represent the information available for groundfish assessment and management. The simulated survey and catch data informed a stock assessment model used to estimate fishery metrics. Biological reference points were calculated with the same assumptions of the stock assessment and stock assessment output. The stock assessment output and estimated BRPs were compared to produce estimated stock status. A HCR then determined fishing mortality (F) based on the estimated stock status. Both the F from the HCR and output from the stock assessment were used in projections to determine catch advice. This catch advice was then applied to simulate harvest in the OM. Catch advice was assumed to be fully caught. Performance of the alternative HCRs were evaluated at each timestep. The stock was assessed every two years unless otherwise noted. We evaluated HCR performance in the context of two groundfish stocks: Gulf of Maine (GOM) cod and Georges Bank (GB) haddock because these stocks typified a range of conditions currently experienced by groundfish stocks. Scenarios with different combinations of stock size, recruitment, and natural mortality assumptions as well as stock assessment model specifications were simulated to evaluate the performance of HCRs when a stock was overfished, not overfished, and when a stock assessment model had a misspecification, which could result in retrospective patterns.

Stock assessment misspecifications included incorrect natural mortality, recruitment, and survey catchability assumptions. For the natural mortality stock assessment misspecification, in the OM, natural mortality was 0.2 at the beginning of the historical period. In 1988, natural mortality started increasing. In 2003, natural mortality was 0.4 and remained at 0.4 through the rest of the historical period and into the projection period. The stock assessment assumed natural mortality was constant at 0.2 in both the historical and projection period. For the recruitment misspecification, the OM assumed recruitment was impacted by SSB and negatively by temperature using a Beverton-Holt stock recruitment relationship with a temperature variable. However, the stock assessment model did not account for the negative impact of temperature. For the survey catchability misspecification, the OM assumed survey catchability was constant until the projection period. At the beginning of the projection period, survey catchability began to decrease with temperature. At the end of the projection period, survey catchability decreased to half of what it originally was. The stock assessment assumed survey catchability was constant at some level.

For each scenario, four different HCRs were evaluated: ramp, P*, F-step and constrained ramp HCRs (Fig. 1). The ramp HCR was designed to emulate the basic structure of the current ABC control rule and promoted rebuilding and optimal yield by decreasing fishing mortality (F) gradually with spawning stock biomass (SSB) if SSB was below the threshold (50% SSBMSY). The P* HCR emulated the P* method, which also ramps down F as SSB decreases below a threshold but avoids overfishing by accounting for uncertainty with a probabilistic approach. The F-step HCR emulated a step in F (between 75% F_{MSY} and 70% F_{MSY}) HCR, which has recently been applied to some New England groundfish. This HCR represented a deviation from the ABC control which defines Frebuild as 70%FMSY. The constrained ramp HCR emulated a ramp HCR that includes a catch variation constraint. With the catch variation constraint, catch advice cannot change more than 20% from the previous year's catch. For the correctly specified GOM cod scenario, we also simulated catch advice with two-year projections (median of catch from each projected year used as catch advice) and with one-year projections (median of catch from the first projected year used as catch advice for both years).

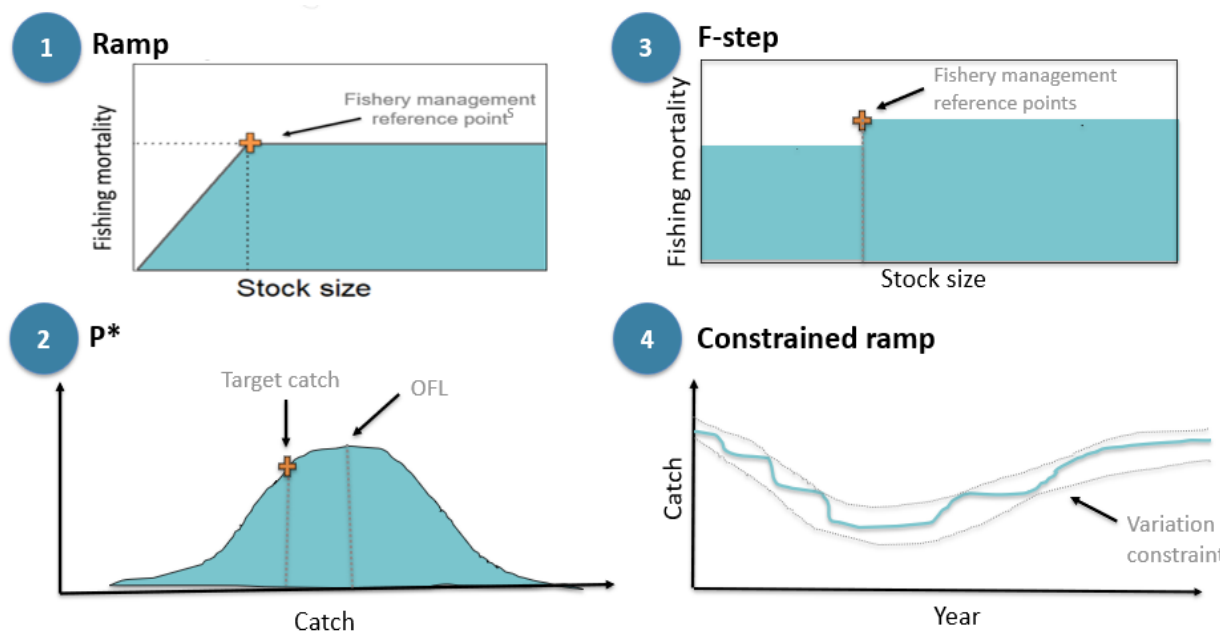

Figure 1. Harvest control rule forms evaluated in this study.

Take Home Results

Overall, the ramp, P*, and F-step HCRs resulted in different catch advice when a stock was overfished, but performed relatively similarly when not overfished. There were trade-offs in the performance of HCRs in the short- (1-5 years), medium- (6-10 years) and long-term (11-21 years) relative to key metrics (e.g. SSB, catch, catch stability, and frequency of overfished and overfishing status). For an overfished stock, the choice of HCRs was most influential in the short and medium-term, as there were more significant differences in HCR performance during this period. In the long-term, the ramp, P*, and F-step HCRs typically performed similarly because stock size increased over the SSB threshold and thus catch advice was similar among HCRs.

With a stock assessment misspecification, population dynamics differed and stock assessment bias was introduced; however, trade-offs among HCRs were similar to those in the correctly specified scenarios. However, misspecifications played a larger role in long-term stock status than the HCRs, because misspecifications can bias reference points and stock estimates. When there was a misspecification, the true rebuilding target may have changed but management was not aware of the change. Thus, the frequency of overfished and overfishing stock status depended more on the type of stock assessment misspecification, rather than the HCR.

HCR performance depended upon OM and stock assessment assumptions. When there were retrospective patterns in a scenario, the patterns were negligible in the beginning of the projection period. Retrospective patterns are a sign that there is a stock assessment misspecification that has greatly impacted our perception of reality and that our understanding of reality may be incorrect. The scenario with the combined natural mortality and recruitment misspecification simulated retrospective patterns that are closest to those in current groundfish assessments and also captures the expected negative impact of temperature on recruitment for cod. Thus, this scenario can provide valuable information for management. The classification of which HCR performs best across a range of conditions will depend on the definition and prioritization of management objectives for the groundfish fishery which was outside the scope of this study.

How do alternative harvest control rules perform when a stock is overfished?

In these scenarios we compared the performance of alternative HCRs for a stock that is overfished (ie. Gulf of Maine cod) with correct stock assessment specifications. For an overfished stock with correct stock assessment specifications, the HCRs performed differently in the short-term as SSB was below the overfished threshold (50% SSBMSY; Fig. 2). In the short-term, the ramp and P* HCRs performed similarly with lower catch compared to F-step and the constrained ramp HCRs. The initial lower catch under the ramp and P* HCRs resulted in slightly faster increases in SSB and decreased the frequency of being overfished, resulting in some gains in medium-term catch. In the long-term, stock trajectories were similar among different HCRs, except under the constrained ramp HCR. Under other HCRs, catch was allowed to increase more than 20% from the previous year's catch. This constraint on catch resulted in higher SSB under the constrained ramp HCR. Interestingly, the constrained ramp HCR did not always result in high catch stability. This was due to the absolute change in allowable catch becoming larger as SSB increased. None of the HCRs allowed for catch to increase to the level of the 1980s and 1990s, because F was not allowed to get as high as it had in the past. Also, all HCRs resulted in rebuilding but with different stock status trajectories to achieve this. In this scenario, median recruitment was fairly constant overtime after SSB increased past the threshold in the cumulative distribution function. There was minimal error in terminal stock assessment and reference point estimates and no evidence of retrospective patterns because there was no stock assessment misspecification.

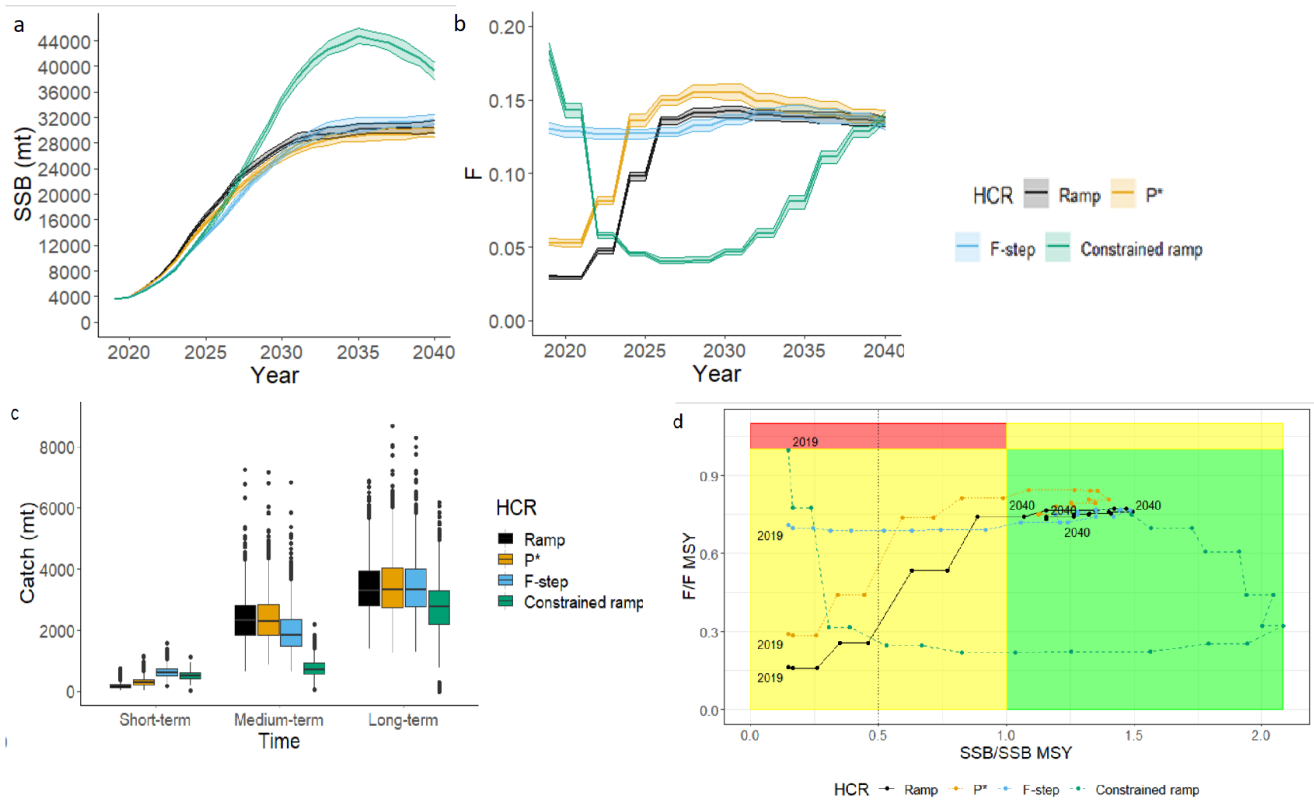

Figure 2. True operating model median spawning stock biomass (SSB; a) and fishing mortality (F; b) with 95% confidence intervals for Gulf of Maine cod with no stock assessment model misspecification (Base Case Overfished Scenario) from 2019 to 2040. c) True median catch for Gulf of Maine cod with no stock assessment model misspecification (Base Case Overfished Scenario) in the short- (1-5 years), medium- (6-10 years), and long-term (11-21 years). d) 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 Gulf of Maine cod with no stock assessment model misspecification (Base Case Overfished Scenario). The dashed line represents the overfished threshold.

How do differences in projections impact performance of alternative harvest control rules?

We compared scenarios with catch advice based on two-year projections and with the year one projections held constant for catch advice. When year one projections were held constant HCRs performed more conservatively than with catch advice informed by two-year projections (Fig. 3). This is because the median of the catch of the first year of the projection was often smaller than that of the second year of the projection. As a result, short-and medium-term catch was lower and SSB rebuilt faster when the year one projection was held constant. Overall, in both of these correctly specified scenarios, all HCRs were able to produce sustainable catch advice.

Year One Projection Held Constant **Two-year Projections**

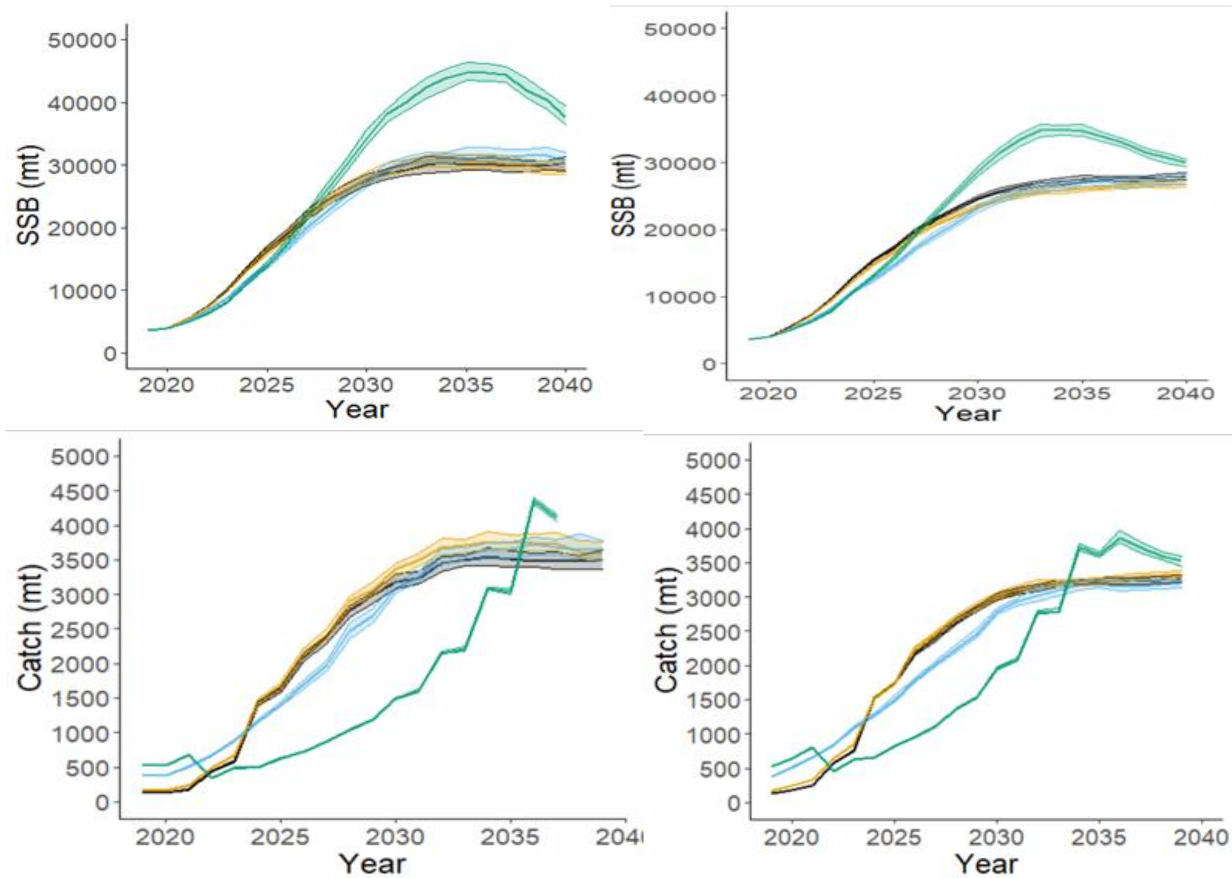

Figure 3. True operating model median spawning stock biomass (SSB) and catch with 95% confidence intervals with year one projection held constant (left) and two-year projections (right) used to determine catch advice for Gulf of Maine cod with no stock assessment model misspecification (Base Case Overfished Scenario) from 2019 to 2040.

How do alternative harvest control rules perform when a stock was not overfished?

These scenarios evaluated the performance of alternative HCRs for a groundfish stock that was not overfished (ie. GB haddock) with a correctly specified stock assessment. Conditioning the simulations on haddock provided a contrast to those conditioned on GOM cod for evaluating HCRs, because the haddock population dynamics were driven by large recruitment events. A large recruitment event occurred near the end of the historical period for all haddock scenarios, and the beginning of the projection period depended on that recruitment event (Fig. 4). Large recruitment events were emulated during the projection period but are not apparent in plots which show median recruitment from all iterations. The ramp, P*, and F-step HCRs performed similarly, because the prescribed F was similar since SSB was above the overfished threshold throughout the projection period. The ramp, P*, and F-step HCRs allowed the fishery to take advantage of the larger recruitment events and achieve higher catch. The HCR with a constraint on variation in catch restricted the ability to take full advantage of large recruitments that resulted in a higher catch in the short-term for the other HCRs. As a result, the constrained ramp HCR conserved SSB which resulted in the highest catch in the long-term. HCRs resulted in a similar stock status in the long-term, although the features of HCRs resulted in different trajectories to get there. The relative HCR performance was different than with the overfished stock. The constrained ramp HCR resulted in highest SSB in short-, medium-, and long-term, whereas for an overfished stock the constrained ramp HCR only resulted in the highest SSB in the long-term. The ramp, P*, and F-step HCRs resulted in the highest catch in the short- and medium-term, whereas for an overfished stock the F-step HCR resulted in the highest catch in the short-term and the ramp and P* HCRs resulted in the highest catch in the medium-term. The constrained ramp HCR resulted in the highest catch in the long-term but it resulted in the lowest catch in the long-term for an overfished stock. There was minimal error in terminal stock assessment and reference point estimates and no evidence of retrospective patterns because there was no stock assessment misspecification.

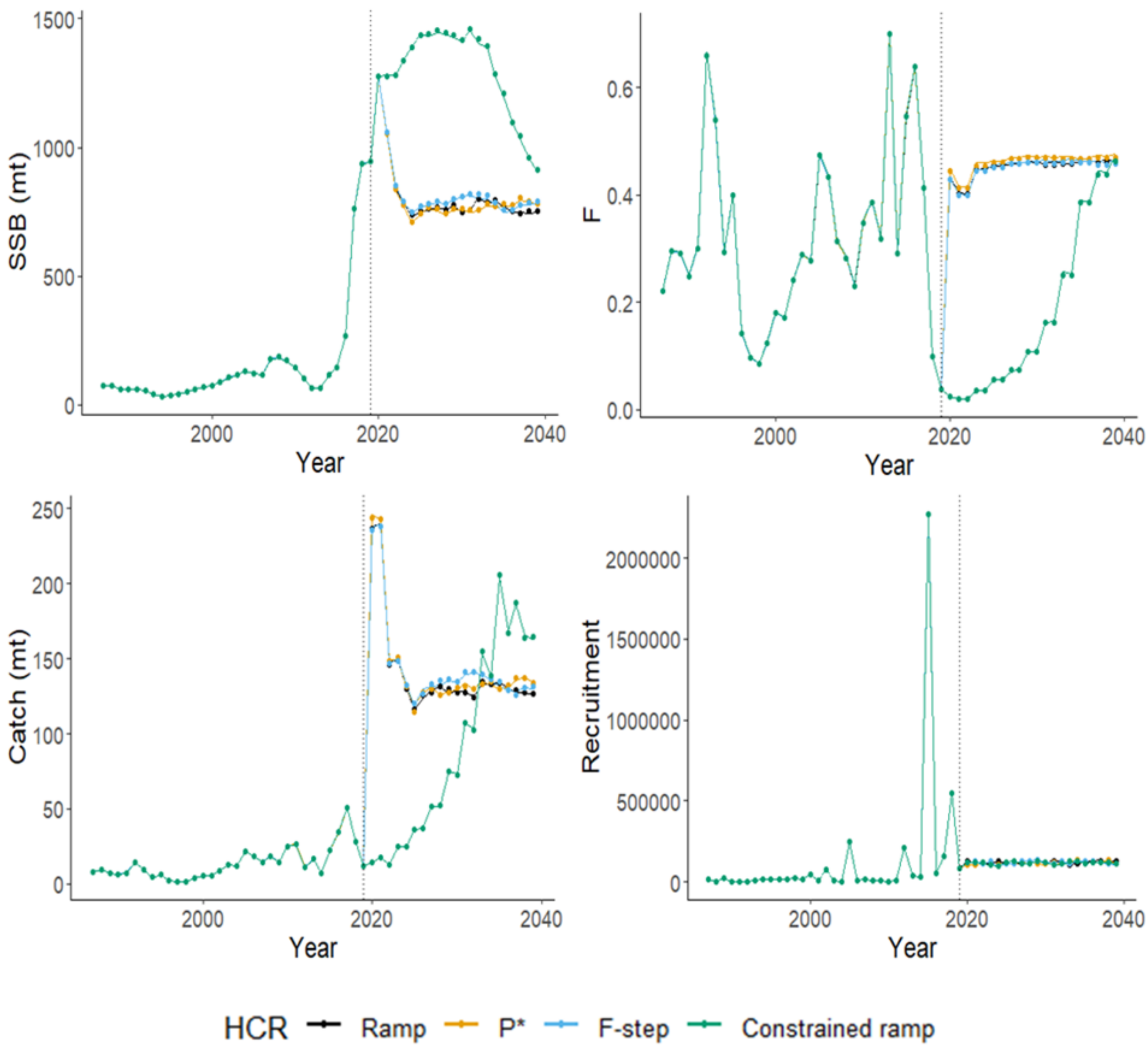

Figure 4. True operating model (closed circles) and estimated stock assessment values from the terminal assessment (lines) of spawning stock biomass (SSB), catch, and recruits for Georges Bank haddock with no stock assessment model misspecification (Base Case Not Overfished Scenario) from 1982 to 2040. The dotted line represents the beginning of the management procedure period (2019).

How do alternative harvest control rules perform when stock assessments are misspecified?

In the following scenarios we compared the performance of HCRs when the stock assessment was correctly specified to scenarios with stock assessment misspecifications.

Natural mortality misspecification

For an overfished stock with a natural mortality stock assessment misspecification, the higher natural mortality contributed to lower catch and SSB. The natural mortality misspecification resulted in assessment bias and retrospective patterns appeared in the long-term. The direction of assessment bias and retrospective inconsistencies tended to coincide. There was overestimation and positive Mohn's rho values for SSB and underestimation and negative Mohn's rho values for F. The natural mortality misspecification biased reference points (Fig. 5). This resulted in more conservative reference points in how they impacted the HCR performance. Estimated SSBMSY was higher, which meant that F would decrease at a higher perceived SSB. Estimated FMSY was lower, which meant that F prescribed from the HCR was lower. The overestimation of SSB and abundance caused a higher F than what was prescribed from the HCR. Estimated stock status indicated overfishing was occuring in the initial years of the projection period and the stock took nearly a decade to rebuild. In this scenario, the perceived stock status (relative to M=0.2 reference points) differed from true stock status (relative to M-ramp reference points). Overall, a misspecification of mortality in the stock assessment led to bias in the stock assessment estimates (ie overestimation of SSB and underestimation of F) and misperception of stock status that doubled the perceived period of stock rebuilding. The mortality misspecification also caused a slight difference in relative HCR performance from the HCR performance under the correctly specified overfished scenario. Long-term catch was slightly higher under the constrained ramp HCR instead of slightly higher under the ramp, P*, and F-step HCRs.

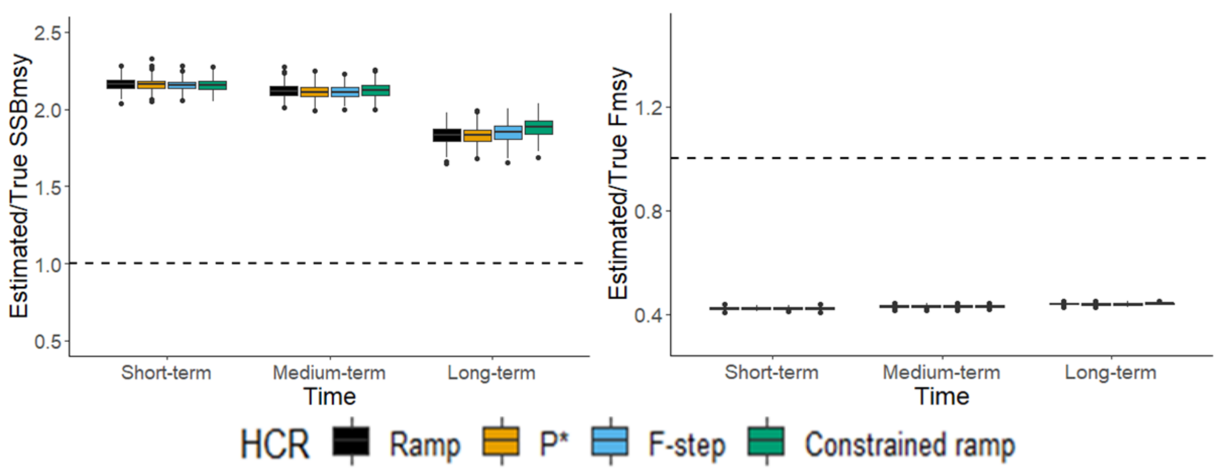

Figure 5. Median ratios of estimated to true stock biomass reference point (SSB_{MSY}) and fishing mortality biomass reference point (F_{MSY}) for Gulf of Maine cod with a natural mortality stock assessment model misspecification (Overfished Mortality Misspecified Scenario) in the short- (1-5 years), medium- (6-10 years), and long-term (11-21 years).

How does performance of HCRs differ with frequency of stock assessment updates?

These scenarios evaluated the impact of annual and alternative year stock assessment updates for the case of an overfished stock with a natural mortality misspecification. The HCRs performed similarly but were more reactive, as catch advice was updated annually (Fig. 6). This caused F to be more similar among HCRs in the long-term.

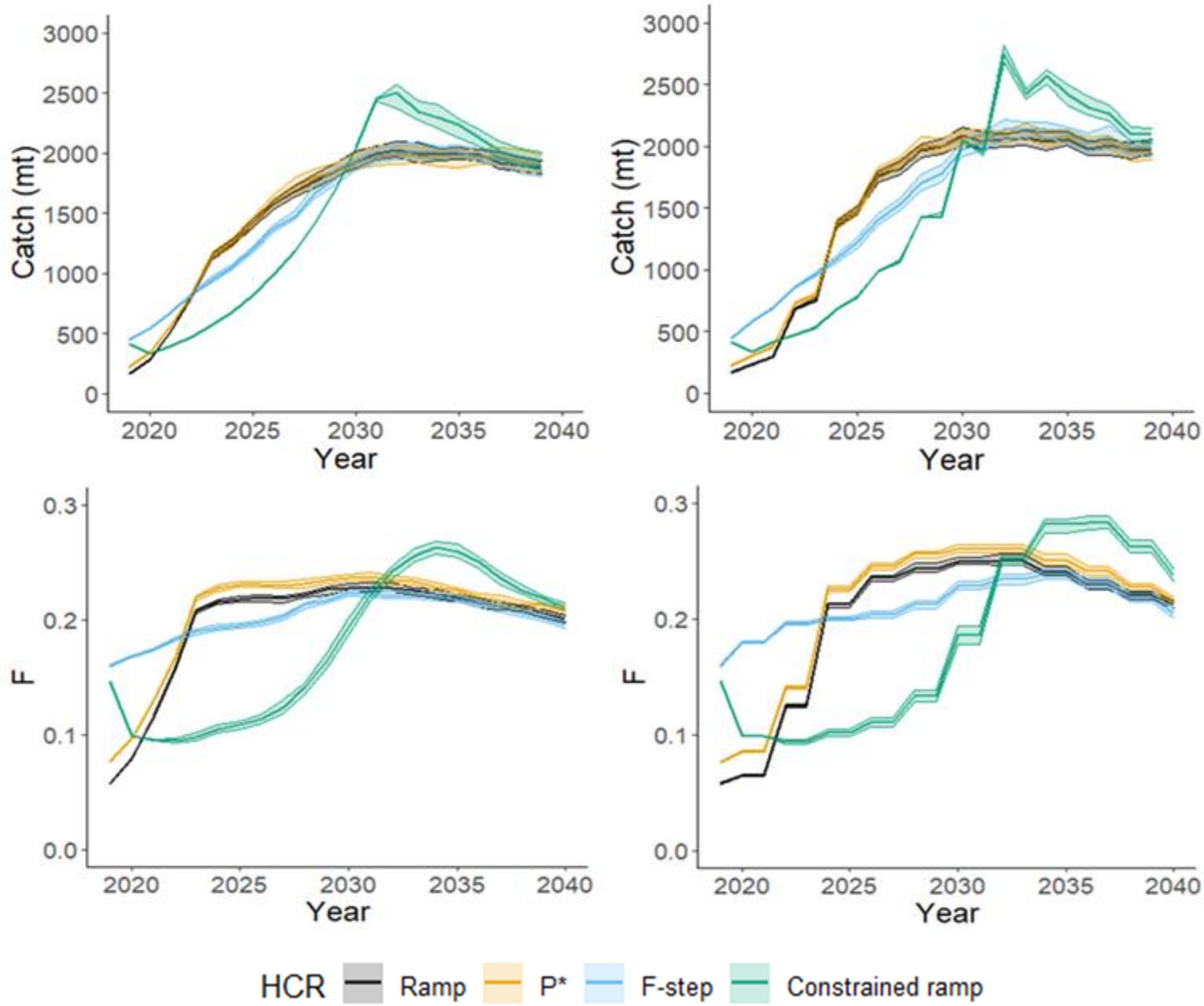

Figure 6. True operating model median catch and fishing mortality (F) with 95% confidence intervals with annual updates (left) and two-year updates (right) for Gulf of Maine cod with a natural mortality stock assessment model misspecification (Overfished Mortality Misspecified Scenario) from 2019 to 2040.

Recruitment Misspecification

For an overfished stock with a recruitment misspecification, the negative impact of temperature on recruitment contributed to lower catch and SSB (Fig. 7). Recruitment declined overtime and this varied with projected temperature. As a result, SSB and catch both declined in the long-term. However, stock assessment error and retrospective patterns for this scenario were negligible with this type of misspecification.

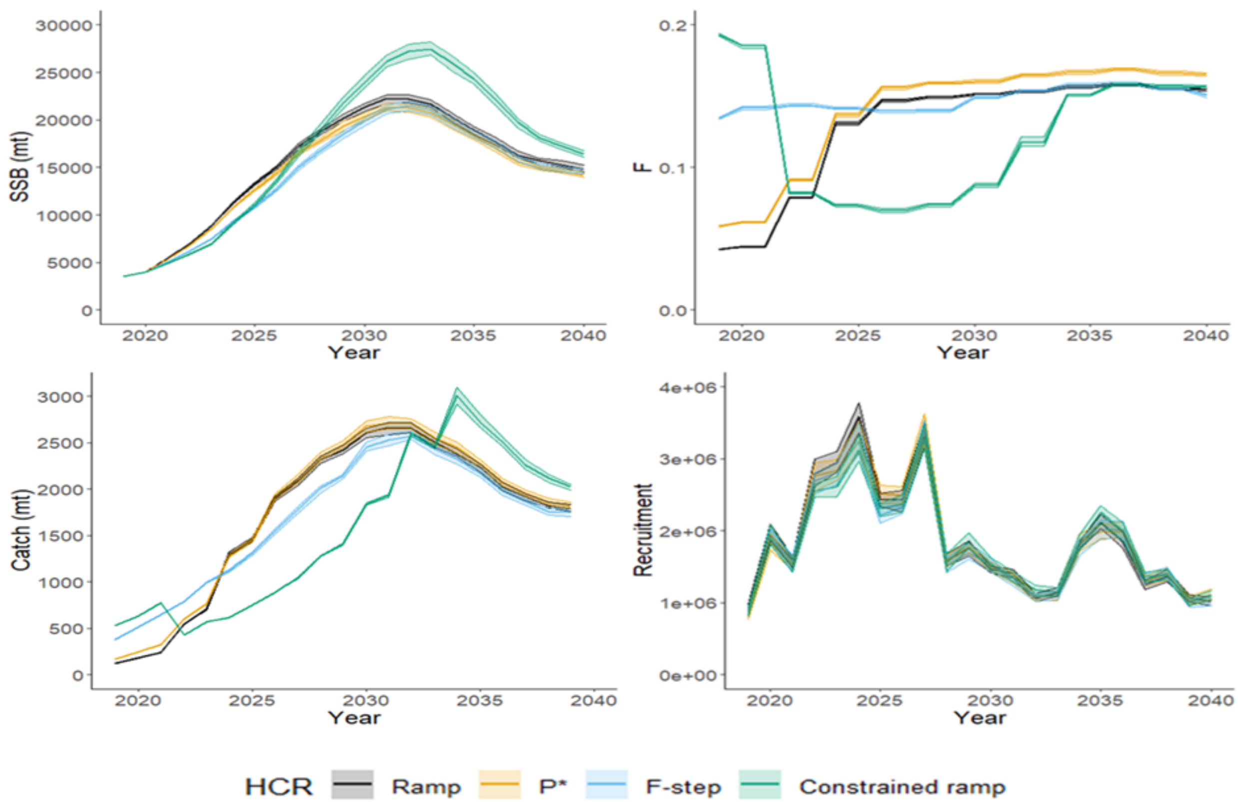

Figure 7. True operating model median spawning stock biomass (SSB), fishing mortality (F), catch (mt), and recruits with 95% confidence intervals for Gulf of Maine cod with a recruitment stock assessment model misspecification (Overfished Recruitment Misspecified Scenario) from 2019 to 2040.

Natural Mortality and Recruitment Misspecification

In previous scenarios, only one parameter was misspecified at a time, but in reality, multiple parameters can be misspecified. These scenarios allowed for testing the performance of alternative HCRs for an overfished stock with both a natural mortality and recruitment misspecification in the stock assessment. This scenario incorporated a negative impact of temperature on recruitment and higher natural mortality which contributed to lower catch and SSB compared to the correctly specified overfished scenario (Fig. 2). The combined natural mortality and recruitment misspecification scenario resulted in the highest levels of stock assessment error and most severe retrospective patterns, which approached the level of retrospective patterns in current groundfish assessments (Fig. 8). In this scenario, the long-term results may be more important to consider, because retrospective patterns have persisted for extended periods for New England groundfish stock assessments, and retrospective patterns were largest in the long-term. Simulating retrospective patterns is an acknowledged challenge in the field and similar efforts have also resulted in no or small retrospective patterns. Stock assessment error and retrospective patterns were sometimes in opposition in this scenario. At first SSB was overestimated but at the end of the projection period it was underestimated. However, Mohn's rho for SSB was always positive. Likewise, at first F was underestimated but at the end of the projection period it was overestimated. However, Mohn's rho for F was always negative.

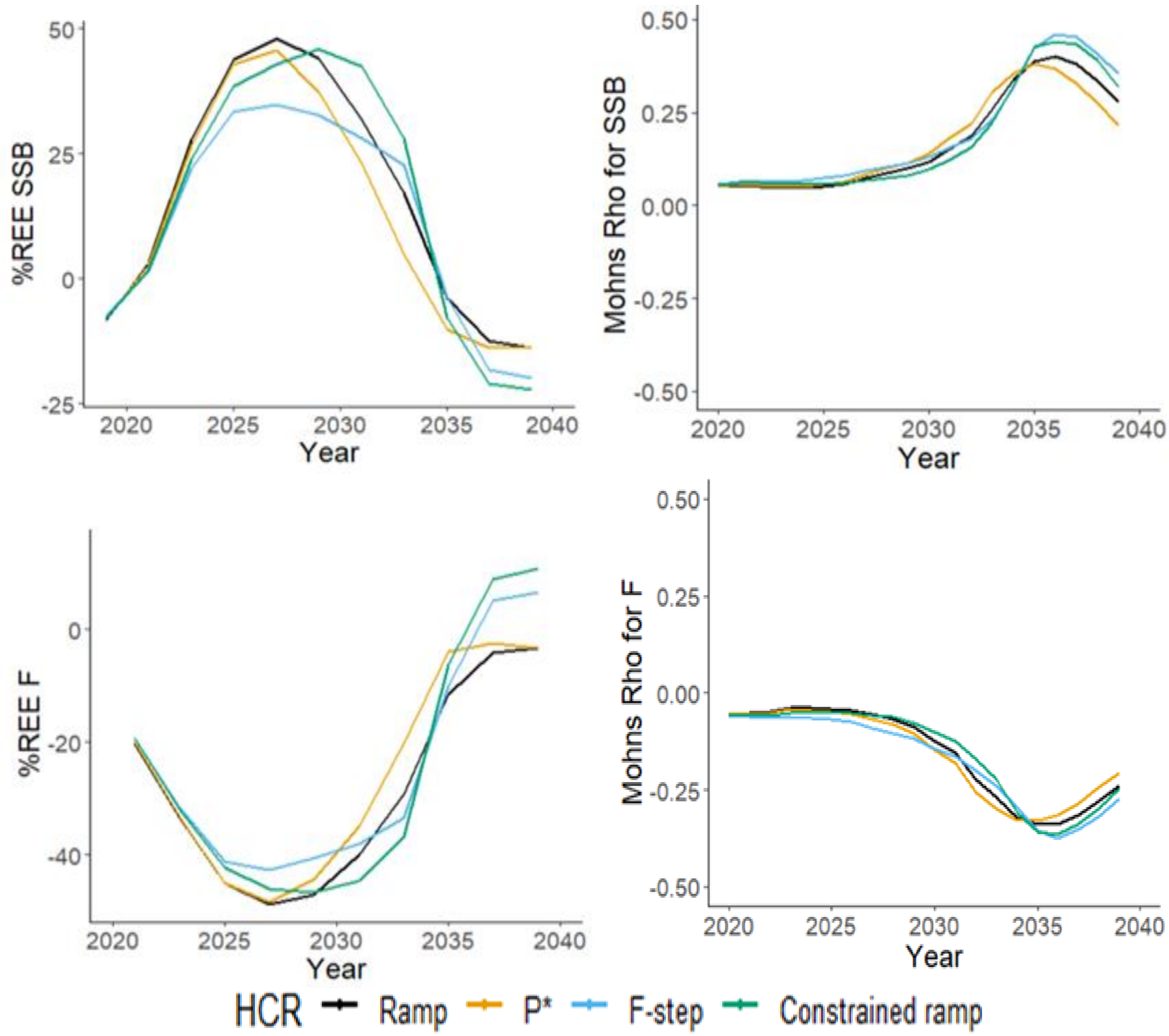

Figure 8. Percent relative error (REE) in terminal estimated spawning stock biomass (SSB) and fishing mortality (F) and Mohn's rho values for SSB and F for Gulf of Maine cod with a natural mortality and recruitment stock assessment model misspecification (Overfished Mortality and Recruitment Misspecified Scenario).

It is important to note that the terminal estimated stock status each year is different from true stock status. At times, the stock was perceived to be overfished and undergoing overfishing and never rebuilt. This estimated stock status is what the catch advice is based on. Although SSB was typically overestimated and F was typically underestimated, the impact of stock assessment bias was lessened to a degree by the estimated reference points being more conservative in how they influenced the HCR. Estimated SSBMSY was higher, meaning that the SSB threshold was higher and estimated FMSY was lower, meaning that prescribed F from the HCR was lower.

The estimated stock status led to more conservative catch advice, but also set a higher bar for overfished and overfishing status. In this scenario, relative HCR performance was generally the same with a stock assessment misspecification, with the exception of true frequencies of overfished and overfishing which differed. SSB was rebuilt in the short- to medium-term, but ultimately the stock was overfished with overfishing occurring in the long-term.

In this scenario, the constrained ramp HCR did not result in dramatically different trajectories because SSB was not increasing fast enough for the catch to increase much more than 20% in the beginning of the projection period. Although the performance of HCRs differed in the beginning of the projection period, near the end the performance was similar except the constrained ramp HCR resulted in a slightly higher F and catch. All HCRs resulted in a lower frequency of an overfished stock due to the conservative estimated reference points and a lower true rebuilt threshold.

Survey Catchability Misspecification

In these scenarios we evaluated the performance of alternative HCRs for a stock that is not overfished (ie. GB haddock) but did have misspecified survey catchability. The survey catchability misspecification led to lower catch in the short-, medium-, and long-term, and higher SSB in the medium- and long-term. In the stock assessment, survey catchability was assumed to be constant, and this caused an underestimation of SSB and overestimation of F (Fig. 9). This misspecification caused the HCRs to be more conservative since the estimated SSB was smaller than the true SSB. SSB and catch levelled off at slightly different magnitudes with the misspecification in comparison to the correctly specified not overfished scenario.

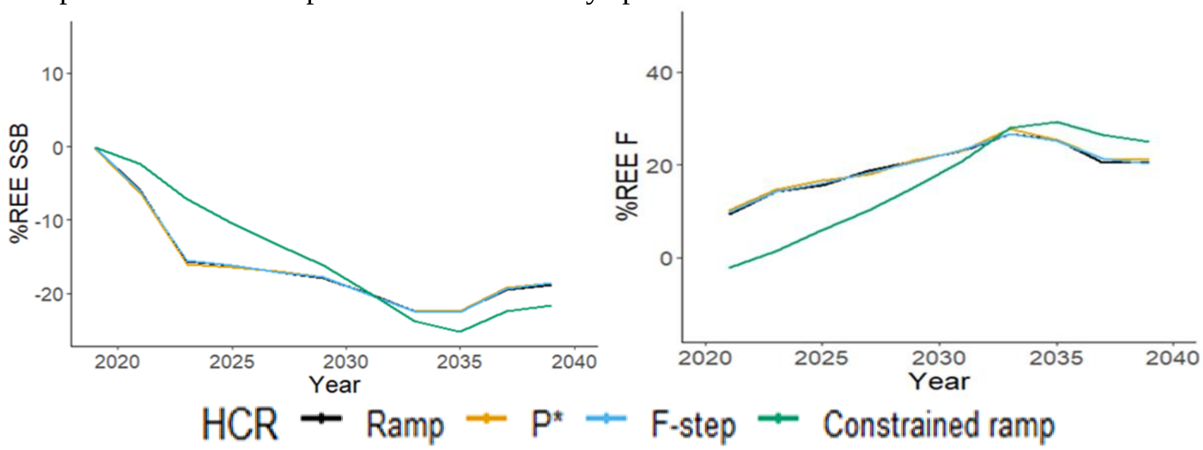

Figure 9. Percent relative error (REE) in terminal estimated spawning stock biomass (SSB) and fishing mortality (F) for Georges Bank haddock with a survey catchability stock assessment model misspecification (Not Overfished Catchability Misspecified Scenario).

When retrospective patterns exist, do retrospective patterns result in better performance than no retrospective patterns?

In these scenarios, we explored the impact of a rho-adjustment on the performance of HCRs. We evaluated this in the context of an overfished stock (GOM cod) with a natural mortality misspecification. However, it is important to note that this result is conditional on the scenario conditions which did not elicit a Mohn's rho value as high as currently observed in many groundfish stocks. Also, in the beginning of the projection period in this study, Mohn's rho values for SSB were less than 0.15, so rho-adjustments did not occur. In the long-term, a rho-adjustment decreased F and catch slightly (Fig. 10). However, retrospective patterns did not appear until the long-term. Additionally, in the long-term, estimated SSB was above the perceived SSB threshold, so rho-adjustments would have little impact on F determined from the HCR but rather the projections instead.

Rho-adjustments **No rho-adjustments**

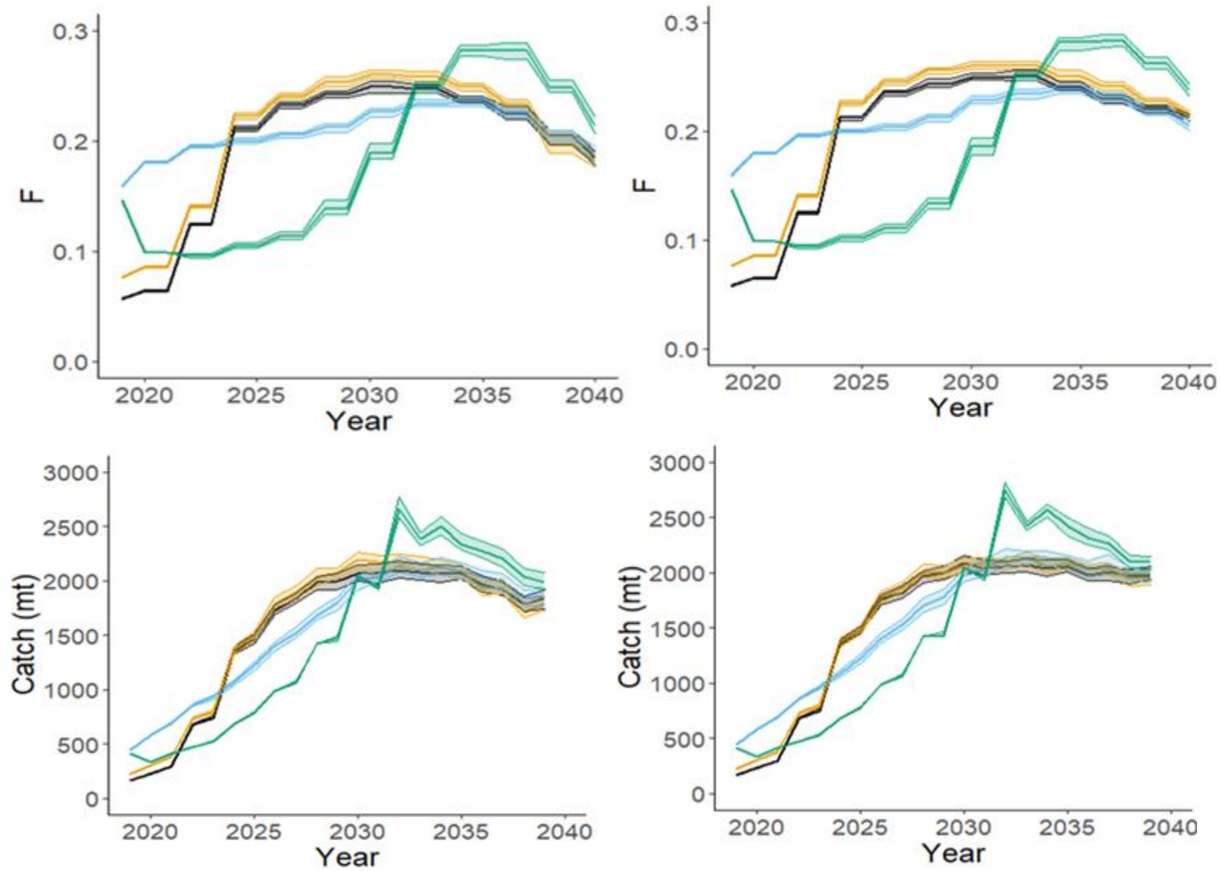

Figure 10. True operating model median fishing mortality (F) and catch (mt) with 95% confidence intervals for Gulf of Maine cod with a natural mortality stock assessment model misspecification (Overfished Mortality Misspecified Scenario) from 2019 to 2040 with rho-adjustments (left) and no rho-adjustments (right).

Caveats

It is important to note some caveats and limitations in this study. The results of this analysis are conditional upon the underlying assumptions of modeled scenarios and the HCRs evaluated. There are additional HCR forms and adjustments to the features of the HCRs evaluated in this study that could be worthwhile exploring in future analyses based on the desired outcomes of groundfish management. One of the limitations of this analysis was that technical interactions were not simulated. The low catch limits on cod have performance implications for several other stocks, such as haddock, that were not included in our evaluations. OMs can also be further tuned to represent additional complexity in groundfish dynamics and operation of groundfish fisheries. However, the reality of groundfish stocks is unknown and recruitment is especially uncertain. For example, GOM cod is currently assessed with two models with different natural mortality assumptions, since true natural mortality is uncertain.

Conclusions

In summary, scenarios with different combinations of stock status, population dynamics, and stock assessment model specifications were simulated to evaluate the performance of alternative HCRs. HCR performance differed between scenarios, metrics, and time periods. When the stock was overfished, all HCRs performed differently in the short-term. HCRs performed differently with a stock assessment misspecification. With a misspecification, the true rebuilding level may be different than the perceived rebuilding level. The frequency of overfished and overfishing depended more on the type of stock assessment misspecification than the HCR. The classification of an optimal HCR will depend on the definition and prioritization of management objectives for the groundfish fishery.