NRC Research Press

ARTICLE

Uncertainty in stock assessment estimates for New England groundfish and its impact on achieving target harvest rates

John Wiedenmann and Olaf P. Jensen

Abstract: Overfishing continues for many stocks in the New England groundfish complex despite efforts to constrain harvest rates and rebuild populations. We evaluated the magnitude and sources of scientific uncertainty in catch targets for groundfish and found that since 2004, annual harvest rates have been 151% above the target while catches have been 29% below the target catch, on average, resulting from overly optimistic catch targets for the majority of stocks. Multiple sources of scientific uncertainty contributed to this overestimation, but the largest contributor was overestimated abundance. By evaluating sequential assessment estimates, we found previous assessments frequently overestimated terminal abundance. Additional uncertainty in catch targets resulted from recent recruitments often being below historical levels. The net effect of overly optimistic catch targets and declining recruitment is that rebuilding has been sluggish and potential yield (and revenue) has been forgone for many stocks. The causes of the overestimation of stock abundance and declines in recruitment remain unknown, but because these patterns were widespread, there may be common mechanisms in the region influencing assessment estimates and stock productivity.

Résumé : La surpêche se poursuit pour de nombreux stocks dans le complexe des poissons de fond de la Nouvelle-Angleterre, malgré des efforts visant à limiter les taux de prélèvement et à reconstituer les populations. Nous avons évalué l'ampleur et les sources de l'incertitude scientifique associée aux objectifs de prises pour les poissons de fond et constaté que, depuis 2004, les taux de prélèvement annuels dépassent les cibles de 151 %, alors que les prises sont de 29 % inférieures aux objectifs de prises, en moyenne, ce qui découle d'objectifs de prises trop optimistes pour la majorité des stocks. Différentes sources d'incertitude scientifique participent à cette surestimation, mais la plus importante est la surestimation de l'abondance. En examinant des estimations découlant d'évaluations séquentielles, nous avons découvert que les évaluations antérieures surestimaient fréquemment l'abondance terminale. Une autre incertitude associée aux objectifs de prises découle du fait que les recrutements récents sont souvent inférieurs aux niveaux historiques. L'effet net d'objectifs de prises trop optimistes et de la baisse du recrutement est que la reconstitution des stocks a été lente et que des rendements (et des revenus) potentiels ne sont pas réalisés pour de nombreux stocks. Les causes de la surestimation de l'abondance de stocks et des baisses du recrutement demeurent inconnues, mais puisque ces phénomènes sont répandus, il pourrait y avoir des mécanismes communs dans la région qui influencent les estimations tirées d'évaluations et la productivité des stocks. [Traduit par la Rédaction]

Introduction

For most fisheries, management aims to keep annual harvest rates for a population at or below some limit that defines overfishing (Mace 2001). While the goal for a particular fishery is a target harvest rate, managers must try to achieve that rate by setting catch limits. When a stock assessment is possible, calculating catch limits is straightforward, but having those catches achieve the target harvest rates is not. Uncertainty in both the science and the management processes is common and can have a large impact on the realized harvest rate for a stock. In some cases, realized harvest rates are far from the target (both above and below), which can have consequences for both the resource and the stakeholders. Identifying the reasons for such discrepancies is essential if we are to effectively manage our fish stocks.

Discrepancies between the target and realized harvest rates can result from uncertainty in both the science and management processes. On the management side, uncertainty typically manifests in the ability of the fishery to achieve the catch limits, also known as implementation error. On the science side, uncertainty manifests in two areas: the current estimates of population size and future dynamics. In other words, how accurate are the estimates, and how predictable are the future stock dynamics? In the assessment model, data issues (e.g., age–length key error or misreported catches) and model assumptions (e.g., fixed natural mortality rate or survey catchability) may interact to increase the error in model estimates (Mohn 1999; Brooks and Deroba 2015). Projections are often used to estimate future catch targets, and the models rely on a number of key assumptions, including future levels of recruitment, natural mortality, masses at age, and fishery selectivity. These inputs are often highly variable and prone to environmental influence, such that the actual values may differ greatly from the assumed levels used in the projection model. Projections that are done over a number of years also assume a target harvest rate, such that error early in the projected time series can propagate over the entire period, exacerbating the discrepancy between the target and realized harvest rates for a given stock.

Scientific uncertainty in current and forecasted population size can have a large impact on a stock. For example, Sinclair et al. (1985) identified a pattern of overestimation of biomass in repeated stock assessments for multiple stocks of Atlantic herring (*Clupea harengus*) in the Northwest Atlantic. Similarly, Walters and

Received 15 November 2016. Accepted 25 April 2017.

J. Wiedenmann. Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ 08901, USA.
O.P. Jensen. Department of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ 08901, USA.
Corresponding author: John Wiedenmann (email: john.wiedenmann@rutgers.edu).

Table 1. The 20 stocks in the New England groundfish complex, whether or not they were undergoing rebuilding in 2003 (NEFMC 2003), and current status (NEFSC 2015).

Full stock name	Scientific name	Abbreviated name	Rebuilding in 2003?	Status in 2014
Georges Bank Atlantic cod	*Gadus morhua*	GB cod	Yes	Overfished
Gulf of Maine Atlantic cod	*Gadus morhua*	GOM cod	Yes	Overfished
Georges Bank haddock	*Melanogrammus aegelfinus*	GB haddock	Yes	Rebuilt
Gulf of Maine haddock	*Melanogrammus aegelfinus*	GOM haddock	Yes	Rebuilt
Georges Bank yellowtail flounder	*Limanda ferruginea*	GB yellowtail flounder	Yes	Unknown, likely overfished*
Cape Cod – Gulf of Maine yellowtail flounder	*Limanda ferruginea*	CC–GOM yellowtail flounder	Yes	Overfished
Southern New England – Mid-Atlantic yellowtail flounder	*Limanda ferruginea*	SNE–MA yellowtail flounder	Yes	Overfished
Georges Bank winter flounder	*Pseudopleuronectes americanus*	GB winter flounder	No	Overfished
Gulf of Maine winter flounder	*Pseudopleuronectes americanus*	GOM winter flounder	No	Unknown
Southern New England – Mid-Atlantic winter flounder	*Pseudopleuronectes americanus*	SNE–MA winter flounder	Yes	Overfished
Witch flounder	*Glyptocephalus cynoglossus*	Witch flounder	No	Overfished
American plaice	*Hippoglossoides platessoides*	Plaice	Yes	Not overfished
Acadian redfish	*Sebastes fasciatus*	Redfish	Yes	Rebuilt
Pollock	*Pollachius virens*	Pollock	No	Rebuilt
White hake	*Urophycis tenuis*	White hake	Yes	Not overfished
Ocean pout	*Zoarces americanus*	Pout	Yes	Overfished
Gulf of Maine – Georges Bank windowpane flounder	*Scopthalmus aquosus*	North windowpane flounder	No	Overfished
Southern New England – Mid-Atlantic windowpane flounder	*Scopthalmus aquosus*	South windowpane flounder	Yes	Rebuilt
Atlantic halibut	*Hippolossus hippoglossus*	Halibut	Yes	Overfished
Atlantic wolffish	*Anarhichas lupus*	Wolffish	NA	Overfished

Note: For status, overfished refers to $S < 0.5S_{\mathrm{MSY}}$, not overfished refers to $0.5S_{\mathrm{MSY}} < S < S_{\mathrm{MSY}}$, and rebuilt refers to $S \geq S_{\mathrm{MSY}}$ (and does not necessarily imply a stock was undergoing rebuilding, like pollock for example).

*Although the status of GB yellowtail is unknown due to uncertainty in the recent assessments, previous assessments that passed the review process indicated very low biomass in recent years (Legault et al. 2013), so it is likely overfished.

Maguire (1996) found repeated overestimation of biomass of the northern cod (*Gadus morhua*) stock in Canada. For both the cod and herring examples, one of the main contributing factors to assessment error was that catch per unit effort (CPUE) in the fishery (used as an index of abundance in the assessment model and assumed proportional to biomass) increased as the stock declined, resulting in inflated biomass estimates at low population sizes. Overestimated abundance can lead to overcapacity in the fleet as well as catch limits being set too high, and both of these contributed to the well-known northern cod stock collapse in the early 1990s (Walters and Maguire 1996).

The groundfish complex in New England currently comprises 20 stocks (Table 1), which have supported large fisheries of great economic and cultural importance in the region. Aggregate landings for groundfish by US vessels fluctuated between 100 000 t to over 200 000 t from 1962 to 1990, generally constituting more than half the total landings of all major fisheries in the northeastern US (comprising both New England and the Mid-Atlantic regions; Sosebee et al. 2006) during this time period. Intense fishing pressure resulted in a number of overfished stocks in need of rebuilding (Rosenberg et al. 2006) as mandated by the Magnuson–Stevens Fishery Conservation and Management Act of 1996 (MSFCMA; NOAA 1996). Early rebuilding efforts were largely ineffective, however, and a lawsuit (Conservation Law Foundation et al. v. Donald Evans et al.) resulted in a federal judge ruling that the Fishery Management Plan (FMP) for groundfish did not comply with the MSFCMA. As a result, the New England Fishery Management Council (NEFMC) drafted Amendment 13 to the groundfish FMP, with the aim of constraining harvest rates and rebuilding overfished stocks, with most of the 14 groundfish stocks undergoing rebuilding estimated to be rebuilt by 2014 (NEFMC 2003). However, after more than a decade of updated stock assessments and management actions, the majority of stocks undergoing rebuilding are not rebuilt (9 of 14, with eight of those still overfished), and three additional stocks are now in need of rebuilding (NEFSC 2015; Table 1 here). Rothschild et al. (2014) noted that aggregate groundfish catches have been below the aggregate catch limit since 2005, suggesting that at least for some groundfish stocks scientific uncertainty is playing a major role in the continued overfishing across groundfish stocks.

Understanding the magnitude and sources of scientific uncertainty is essential for setting sustainable catch limits, and there are different approaches to do so. Brooks and Legault (2016) conducted a "retrospective forecast" to evaluate sources of bias in projections based on a retrospective analysis of the 2008 assessments for New England groundfish and found error in abundance estimates had the largest impact on projection accuracy. Because Brooks and Legault (2016) focused on a single assessment for each stock, their analysis was a measure of uncertainty within a given assessment model. From one assessment to another, a number of changes may occur to both the data inputs (beyond adding new years of data) and the structural assumptions within the assessment model, and such changes can result in large differences in model estimates between assessments. Ralston et al. (2011) measured the uncertainty among repeated assessments by quantifying changes in biomass estimates over time for stocks off the western US, with the aim of informing appropriate buffer sizes in the harvest control rule to comply with the revised MSFCMA. In this paper we expand on these analyses and on that of Rothschild et al. (2014) and investigate how scientific uncertainty in the setting of target catches affected the ability to achieve target harvest rates for New England groundfish stocks from 2004 to 2014 (representing management from Amendment 13 onward). Specifically, we (*i*) compared target and observed catches and harvest rates for all stocks with estimates of harvest rates, (*ii*) evaluated how estimates from repeated stock assessments (a measure of across-model uncertainty) changed over time and how well measures of uncertainty from an assessment (within model uncertainty) captured the magnitude of error in assessment estimates, (*iii*) evaluated the projection models used to set catch targets to identify dominant sources of uncertainty, and (*iv*) quantified where stock biomass would be and the value of current fishery harvests if target harvest rates had been achieved.

Methods

To better understand the roles that scientific and management uncertainty have played in the continued overfishing across groundfish stocks, we first need to know how the observed harvest rates and catches have deviated from the original management targets. We obtained annual target catches and harvest rates (C_{targ} and F_{targ}, respectively) from 2004 onward from amendments and framework adjustments to the New England groundfish FMP (NEFMC 2003, 2006, 2010, 2011, 2012, 2013, 2014). In most cases, F_{targ} was at or below harvest at maximum sustainable yield (F_{MSY}; see Supplemental Information[1] for more details when F_{targ} > was set above F_{MSY}). Stock assessment documents (NEFSC 2002, 2003, 2005, 2008, 2010, 2011, 2012*a*, 2012*b*, 2012*c*, 2013*a*, 2013*b*, 2014, 2015; Legault et al. 2011, 2013; Hendrickson et al. 2015; Palmer 2014) provide estimates of the observed annual catches (C_{obs}), which are fixed inputs in the model, and annual harvest rates (F_{obs}), which are estimated in the model conditioned on C_{obs} and the estimated abundance and selectivity-at-age. Estimates of F_{obs} were provided for all age classes for each stock, but we used the reported value for the fully selected age classes as our annual estimate for F_{obs}. We used the most recent assessment that passed review for a given stock as the source of estimates of "true" harvest rates and catches, but because these values may change in future assessments, we herein refer to them as the updated estimates.

We restricted our analyses to 15 groundfish stocks with estimates of biomass and harvest rates from age-structured or surplus production assessment models. Three stocks were omitted because the current assessment was index-based (ocean pout (*Zoarces americanus*) and two stocks of windowpane flounder (*Scophthalmus aquosus*)), and an additional stock was omitted (Atlantic wolffish, *Anarhichas lupus*) because it was a data-poor stock until recently (NSWG 2009), such that most of the historical catch targets were not based on a target harvest rate. We also omitted Gulf of Maine (GOM) winter flounder (*Pseudopleuronectes americanus*) because assessments since 2008 have relied on survey-based swept-area estimates of biomass. In situations where the most recent age-based assessments were deemed too uncertain to be the basis for management (Georges Bank (GB) cod and GB yellowtail flounder (*Limanda ferruginea*); Legault et al. 2013; NEFSC 2015), the last assessment that passed review was used for the analysis (Legault et al. 2013; NEFSC 2013*a*). It is important to note that many of the uncertainties present in the failed recent assessments were likely present in previous assessments.

With annual estimates of F_{obs}, F_{targ}, C_{obs}, and C_{targ}, we calculated ratios each year: $F_{ratio}(t) = F_{obs}(t)/F_{targ}(t)$, and $C_{ratio}(t) = C_{obs}(t)/C_{targ}(t)$. For some stocks prior to 2010, the target catch calculated using F_{targ} was considered the landings only (i.e., no discards), but the estimates of F_{obs} from the most recent assessment were calculated using the total catch (landings + discards). In such cases we used the observed landings as our C_{obs} and adjusted the estimated harvest rate downward in that year based on the ratio of the landings to the total catch. This adjustment is an approximation and potentially underestimates the F resulting solely from the landings. Discards generally occurred for younger ages, and the F estimate we used was for the fully selected age classes, such that removing the discards may not change the fully selected F estimate by very much. We explored the impact of using the original or adjusted F and found it had a minimal effect on our overall results.

An F_{ratio} above or below 1 indicates the target harvest rate was not achieved in a given year, which could result from either scientific uncertainty in the target catches or management uncertainty. For example, a stock with F_{ratio} = 2 when C_{ratio} = 2 is different from a stock with F_{ratio} = 2 when C_{ratio} = 0.5. In the former case, an overage in the total catch resulted in the higher F. In the latter case, total catch was only half the target, yet F_{obs} was double F_{targ}, indicating that scientific uncertainty played an important role in $F > F_{targ}$. We therefore calculated a relative measure of error in achieving F_{targ} as a quantitative measure of scientific uncertainty:

$$(1)\qquad F_{error} = \frac{F_{ratio}}{C_{ratio}}$$

Values close to 1.0 imply that the discrepancies between the observed and target F are largely the result of the achieved catch (relative to the target) and are examples of low scientific uncertainty, whereas values farther away from 1 indicate that F_{obs}/F_{targ} is disproportionately large or small for a given C_{obs}/C_{targ} and are examples of high scientific uncertainty. This measure assumes proportional changes in F in response to changes in catch, which is not true across the entire range of C_{obs}/C_{targ}, particularly as this ratio approaches 0 or becomes very large ($\gg$1). Furthermore, this measure does not distinguish between the different sources of uncertainty in a given year. Nevertheless, F_{error} is a useful measure for identifying cases where the achieved F was impacted by scientific uncertainty.

We obtained biomass estimates from all available assessments (from 2002 onward) and compared the historical estimates with the most recent (updated) estimates. In some cases stocks required an adjustment to the estimates in the final year of the model (called the terminal year) based on a retrospective pattern. This modification to model output is called a rho (ρ) adjustment, and we used the adjusted values in the terminal year for that assessment. In other cases, multiple models with different estimates were put forward as plausible in an assessment. For GOM cod, recent assessments explored two formulations for natural mortality, with M fixed at 0.2 for all years and M increasing to 0.4 in recent years (called the M-ramp model), and we used both models as our source for updated estimates. For GB yellowtail flounder following the 2005 assessment, multiple models were put forward as plausible (NEFSC 2005), but only output from the "base" model was provided, so we only used the base model estimates here. Uncertainty in terminal estimates is not the only potential source of error in catch advice from a projection model, however, and we conducted a more thorough evaluation of projection estimates for a subset of groundfish stocks (detailed below).

Using annual estimates of biomass and recruitment across multiple stock assessments for a given stock, we calculated the across-model uncertainty in the assessment estimates. Ralston et al. (2011) used a variety of approaches for calculating the across-model uncertainty for stocks in the Pacific, but we focused on the relative error in estimates using the updated estimate as the reference value (B_{up}). For example, the relative error in a biomass estimate from assessment j in year t (REB(j,t)) is calculated with

$$(2)\qquad \mathrm{REB}(j, t) = \frac{B(j, t) - B_{up}(t)}{B_{up}(t)}$$

The analysis of Ralston et al. (2011) only considered uncertainty across benchmark (or "research track") assessments for a given stock, where data inputs may have been modified (e.g., updated catch estimates) and model assumptions changed (e.g., adjusted natural mortality rate or differently shaped selectivity curve) from the previous assessment. We included both benchmark and update assessments (where the existing model is typically rerun with

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2016-0484.

Table 2. Age-structured equations governing the projections used to set catch targets.

Equation		Description
2.1	$N(a,t) = \begin{cases} R(t) & a = a_R \\ N(a-1,t-1)e^{-Z(a-1,t-1)} & a_R < a < a_{\max} \\ N(a-1,t-1)e^{-Z(a-1,t-1)} & a = a_{\max} \\ \quad + N(a,t-1)e^{-Z(a,t-1)} & \end{cases}$	Numerical abundance at age (a) each year (t), determined by the number of recruits (R) entering the population and the number from the previous year discounted by the total mortality rate (Z)
2.2	$R(t) = \dfrac{\alpha S(t-a_R)}{\beta + S(t-a_R)} e^{\theta_R - 0.5\sigma_R^2}$ $R(t) = \mathrm{ECDF}_{\mathrm{all}}$ $R(t) = \begin{cases} \mathrm{ECDF}_{\mathrm{low}} & S < S_{\mathrm{thresh}} \\ \mathrm{ECDF}_{\mathrm{high}} & S \geq S_{\mathrm{thresh}} \end{cases}$	Recruitment, either based on a Beverton–Holt stock–recruit relationship or drawn from a single empirical cumulative distribution function ($\mathrm{ECDF}_{\mathrm{all}}$) based on recruitment estimates over a specified period or from a two-stage ECDF model with low and high values, as well as the distribution used determined by the projected spawning biomass relative to some specified biomass threshold (S_{thresh})
2.3	$S(t) = \sum_a m(a) w_S(a) N(a,t) e^{-\phi Z(a,t)}$	Total spawning biomass (S), calculated using maturity fraction (m) and spawning mass-at-age (w_S) and the fraction of the year when spawning occurs (ϕ)
2.4	$Z(a,t) = M(a,t) + s(a,t)F(t)$	Total mortality (Z), calculated as the sum of the natural mortality rate (M) and the fishing mortality rate (F) scaled by the selectivity at age in the fishery (s)
2.5	$C(a,t) = \dfrac{s(a,t)F(t)}{Z(a,t)} w_C(a) N(a,t)[1 - e^{-Z(a,t)}]$ $C(t) = \sum_a C(a,t)$	Catch (at age and total), calculated using the Baranov catch equation using the mass-at-age in the catch (w_C)

Note: All model parameters for a projection were based on the output from the accepted assessment model at the time.

only new years of data, although in some cases with other small changes) in our analysis because we wanted to include the possibility for large changes in the perception of stock status following an update assessment, and because both benchmark and update assessments carry equal weight in the calculation of target catches.

Changes in biomass estimates across assessments are a measure of across-assessment uncertainty. We explored whether measures of within-assessment uncertainty were useful predictors of whether or not catch advice would be successful (i.e., are measures of uncertainty produced from an assessment of good predictors of how truly uncertain those estimates are and how effective projections will be?). Mohn's ρ (Mohn 1999), a measure of the retrospective uncertainty in terminal estimates, and the coefficient of variation (CV) in the terminal biomass estimate are two metrics often used to quantify within-assessment uncertainty. We obtained estimates of ρ from each assessment for each stock, noting that the number of "peel" years used to calculate it was not consistent across assessments (between 5 and 7 years). Estimates of ρ for a stock were not provided in NEFSC (2002, 2005), but figures showing retrospective estimates were provided. When ρ was not provided, we extracted 5 years of estimates of terminal biomass from the retrospective analysis figure for each stock and calculated ρ using the estimates from the full time series as the reference for each peel year. We repeated this method and calculated ρ using the figures presented in later assessments (NEFSC 2008, 2012*a*) and compared estimates with the reported value to confirm the reliability of this method. The CV of the estimated terminal biomass was also not always available for each stock from each assessment. However, projection output was available for most stocks, so we therefore used the biomass CV in the first year of the projection as a proxy for the CV in the terminal estimate.

Projections

Over- or underestimation of biomass and recruitment are two possible sources of scientific uncertainty, and a more detailed analysis of potential factors influencing the accuracy of target catches is needed. Projection model accuracy is determined by a number of factors, including the assumed initial abundance at age, mass- and selectivity-at-age, future recruitments, and also the assumed catch (or *F*) during the interim years between the terminal year of an assessment and the first year that the target catch is calculated. Identifying the relative impact of each factor requires updating the projections used to set target catches with these inputs. Brooks and Legault (2016) used projections based on retrospective runs of the NEFSC (2008) assessment to evaluate forecast accuracy for New England groundfish, and we modified their approach by (*i*) using the most recent assessment for each stock as our source of updated estimates and (*ii*) using the actual projection models used to set catch targets for each stock. Further discussion of the differences between approaches is provided in the Discussion.

We focused on the projections based on the assessment estimates from NEFSC (2002, 2005, 2008), as these were the basis of target catches from 2004 to 2012. We only included stocks with age-based estimates for each assessment, leaving nine stocks for this portion of our analysis: GOM and GB cod; GB haddock (*Melanogrammus aegelfinus*); Cape Cod – Gulf of Maine (CC–GOM), southern New England – Mid-Atlantic (SNE–MA), and GB yellowtail flounder; SNE–MA winter flounder, witch flounder (*Glyptocephalus cynoglossus*), and American plaice (*Hippoglossoides platessoides*).

We created an age-structured projection model in R (R Core Team 2015) that mimics the AGEPRO model of Brodziak et al. (1998). We created our own projection model because subsequent simulation work required a projection model that can be easily updated as needed in the simulations to test alternative methods for setting target catches. The equations governing the projection dynamics of our model are presented in Table 2, but we provide a summary of the model here. The projection model uses the same input files created to set catch advice for each stock based on the assessments listed above (obtained from NEFMC staff), and we compared our model estimates with the original AGEPRO output provided in NEFMC (2006, 2010) to ensure consistency. The initial abundance at age and all input assumptions are read into the model. The fishing mortality in the first year is either based on the value specified in the input file for that year, or when a catch is specified, *F* is calculated using the Baranov catch equation (eq. 2.5 in Table 2) and the assumed mean catch masses and fishery selectivity. In the second year of the projection, recruitment is determined from the specified recruitment model with the appropriate lag in years (either Beverton–Holt or one- or two-stage empirical cumulative distribution functions; Table 2). For all other age classes, abundance at age in year 2 is determined by the abundance in the previous year discounted by fishing and natural mortality rates (eq. 2.1 in Table 2). Spawning biomass is calculated each year using the estimated abundance at age and the specified mean maturity- and spawning mass-at-age, discounted by a specified fraction of the total mortality that occurs before spawning in a year. Spawning biomass determines recruitment when the Beverton–Holt recruitment model is specified, and it also determines which time series of recruitments should be used when the two-stage empirical model is used. Total catch in a year is calculated using the Baranov catch equation for a given F_{targ}. The stock is projected forward a number of years under the F_{targ}, and this process is repeated 1000 times to account for uncertainty in the initial abundance and future recruitments, producing a distribution of predicted spawning biomass, recruitment, and total catch for each year in the projection model.

We reran each projection for each stock with an updated estimate of key projection inputs and determined whether or not this improved the catch advice from the projection (i.e., resulted in F closer to F_{targ}). We reran the projections with the updated (*i*) initial abundance at age (including the recruited age class in that year), (*ii*) interim catches (between the assessment terminal year and the first year, the target catches are calculated), (*iii*) future recruitments, (*iv*) mass-at-age in the catch, and (*v*) fishery selectivity-at-age. In addition to these updates, Brooks and Legault (2016) updated maturity-at-age, which we omitted because our primary focus was on the target catches from the projections and not the accuracy of spawning biomass estimates. One input was modified for a given projection, and the estimated median catch in each year was compared with the original projected catch, as well as to the updated catch (that would have achieved F_{targ}) to determine the impact that the input had on the error in target catches.

Finally, for each of the nine stocks we projected the population biomass under F_{targ} to determine what the biomass and current catch would be for each stock if F_{targ} had been achieved each year (i.e., no scientific or management uncertainty). Projections were run from 2004 (the first year catch targets were set based on NEFSC 2002) through 2014, using the updated estimates of abundance-at-age in 2004 and nearly identical population dynamics to those detailed in Table 2. The only difference is in how future recruitments were calculated, as we wanted to allow for changes in recruitment that follow changes in spawning biomass resulting from achieving F_{targ}. For each stock we assumed recruitment followed the Beverton–Holt relationship (eq. 2.2 in Table 2) and estimated the parameters α and β using maximum likelihood, assuming lognormal error. With estimates of the annual relative deviations from the best-fitting recruitment curve ($\gamma(t)$), we calculated the projected recruitment each year using the projected spawning biomass, accounting for the lag between spawning and recruitment (determined by the age at recruitment, a_R):

$$(3) \qquad R_{\mathrm{proj}}(t) = \frac{\alpha S_{\mathrm{proj}}(t - a_R)}{\beta + S_{\mathrm{proj}}(t - a_R)}\gamma(t)$$

We used the original F_{targ} for each stock in each year except for GOM cod in 2012, where F_{targ} was 0.88. We set F_{targ} equal to the target from the previous year (0.18), assuming that had F_{targ} been achieved in previous years, the interim measure that set F_{targ} = 0.88 would not have occurred.

Results

Since 2004, observed catches have been at or below the target while harvest rates have exceeded the target for the majority of stocks in most years (Fig. 1). Across stocks the ratio of the mean catch to the target (C_{ratio}) was 0.71 and was above 1.0 for only two stocks, SNE–MA yellowtail flounder and GOM haddock (Table 3). Despite the relative infrequency of catch overages, harvest rates were above the target ($F_{\mathrm{ratio}} > 1.0$) for 10 of 14 stocks, on average, with a mean F_{ratio} of 2.51 across stocks. For each stock we quantified the error in achieving the target F (F_{error}; eq. 1) as a measure of the scientific uncertainty in setting the target catches. For 11 of 14 stocks, F_{error} was above 1, indicating that harvest rates were higher than expected given the observed catch. Only Acadian redfish (*Sebastes fasciatus*), GOM haddock, and pollock (*Pollachius virens*) had a mean $F_{\mathrm{error}} < 1$. Stocks with the highest F_{error} were GB yellowtail flounder, witch flounder, SNE–MA yellowtail flounder, GOM cod, and CC–GOM yellowtail flounder, each with harvest rates more than five times above the expected value given the observed catch (Table 3).

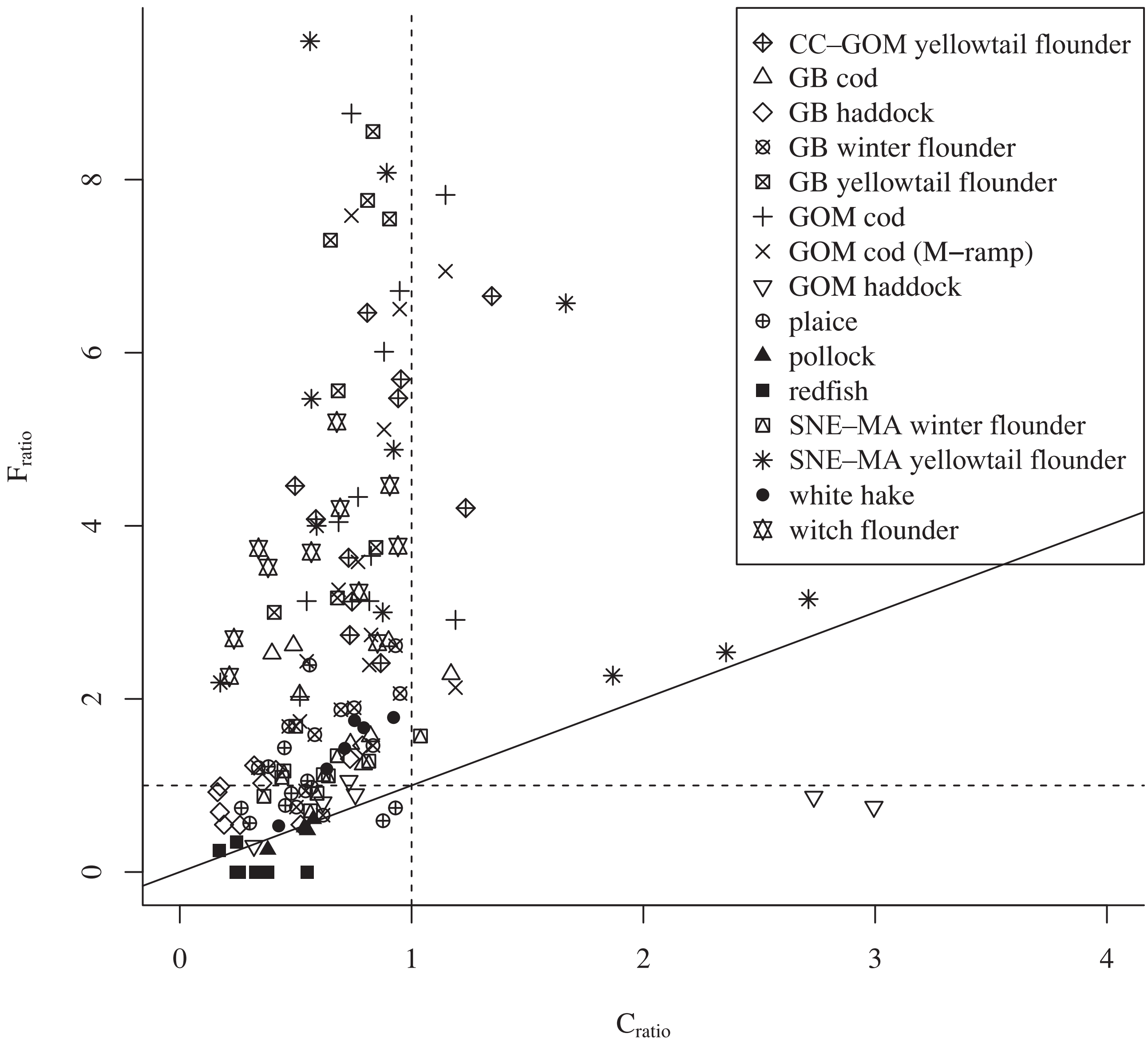
In Fig. 2 we show F_{error} as a function of the lag between the last year of data in the assessment and the year in which the target catch is set based on that assessment (which we call data management lag, or DML). DML ranged from 2 to 5 years, and surprisingly, the median estimate of F_{error} was highest for a DML of 2 years. This result was largely due to GB yellowtail flounder, which had the highest F_{error}, on average, but this stock was also unusual among groundfish stocks in that catch targets were updated annually based on annual assessments (with a DML of 2 years for each). When this stock was excluded, the F_{error} was reduced for a DML of 2 years, although differences in the mean F_{error} were not significant across the range of DML ($p > 0.5$ for all pairwise comparisons of DML using a Tukey's test).

We calculated the relative error in terminal biomass (REB) and recruitment (RER) estimates from 2001 to the terminal year of the second to last assessment (see Fig. 3 for example stocks showing changes in repeated assessments and Figs. S1–S3[1] for all remaining stocks). For most stocks the most recent assessment was completed in 2015, with updated biomass and fishing mortality estimates through 2014 (NEFSC 2015). Estimates of REB and RER in the terminal year of each assessment are shown in Fig. 4 (aggregated across stocks). The general pattern across assessments for most stocks has been to overestimate terminal biomass and recruitment since the assessments of NEFSC (2002). Terminal biomass was estimated to be 64% higher, on average, than updated estimates, with overestimation occurring 71% of the time. Terminal recruitment was overestimated 60% of the time, being 59% higher than the updated estimates, on average. When we only evaluated estimates from the first three comprehensive assessments that were the basis for target catches from 2004 to 2012 (NEFSC 2002, 2005, 2008), the frequency and magnitude of overestimation increased, occurring 78% of the time for biomass estimates (overestimated by 82%, on average) and 69% of the time for recruitment (by 72%, on average).

We compared assessment-specific estimates of REB with two measures of uncertainty from the same assessment to determine if more uncertain assessments (as measured at the time) resulted in greater REB. We used the CV of the biomass in the first year of the projection model as one measure of assessment uncertainty and Mohn's ρ (Mohn 1999), the mean retrospective error in biomass, as the other measure. Linear fits of REB to each measure of assessment uncertainty were not significant ($p = 0.13$ and $R^2 = 0.03$ using ρ as the predictor, and $p = 0.78$ and $R^2 = 0.002$ using the CV as the predictor). The linear fit of REB to ρ was also very sensitive to one point where REB was high for the highest estimated ρ (which occurred for GB yellowtail flounder in the NEFSC (2005) assessment), such that when we removed this point and refit the model, the p value increased to 0.75 (Fig. 5).

In general, overestimates of terminal biomass occurred for assessments with a positive ρ, but there were cases where an assessment had a small positive ρ and the terminal biomass was greatly overestimated, and vice versa. For example, the 2008 assessment for GOM cod had a small retrospective pattern ($\rho = 0.2$), but terminal biomass was overestimated by 293%. In contrast, SNE–MA yellowtail flounder had a large retrospective pattern in biomass

Fig. 1. Annual estimates by stock of the F_{ratio} (calculated as F_{obs}/F_{targ}) as a function of the C_{ratio} (calculated as C_{obs}/C_{targ}). The dashed vertical and horizontal lines at 1 separate the plot space to highlight when catches and F were above or below the target, respectively. The solid black line is the 1:1 line, with values above or below the line indicating the achieved F is disproportionately high or low for a given catch ratio, respectively.



estimates in 2002 (ρ = 1.02), but terminal biomass was overestimated by only 22% (Fig. 5).

Overestimation of terminal abundance can have a large influence on projected target catches, but other factors can also have an impact. In Fig. 6 we show original projection estimates of spawning biomass, recruitment, and target catches for three stocks following NEFSC (2002, 2005, 2008). For GOM cod, both overestimated abundance and below-average forecasted recruitment contributed to the overestimation of catch targets, whereas for SNE–MA winter flounder, catches were overestimated despite underestimated abundance due to below-average forecasted recruitments. For GB haddock, overestimated terminal abundance and declining mass-at-age resulted in large differences between the projected and actual spawning biomasses, particularly following NEFSC (2005). Anomalously high recruitment events have mitigated for these impacts, however, such that in recent years spawning biomass and target catches are similar to the originally projected values (Fig. 6).

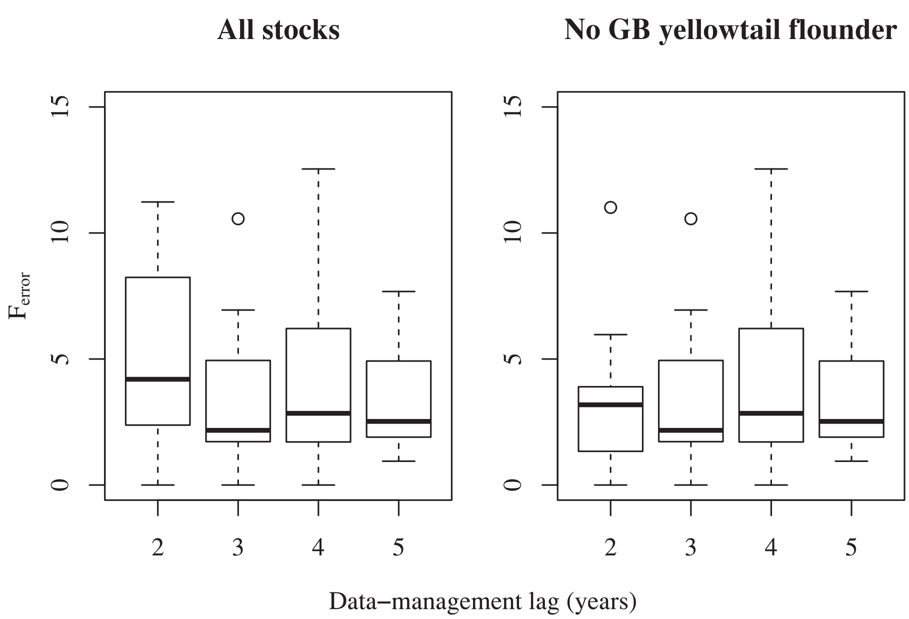
To determine which factor was the largest source of scientific uncertainty across groundfish in the projections used to set target catches, we reran the projections with updated initial abundance at age, forecasted mass, and selectivity-at-age, as well as forecasted recruitments and the actual catches during the DML period. Catch estimates from the projection were averaged across the management periods based on the different assessments used to inform the projections (e.g., 2004–2005 for projection based on NEFSC 2002). A number of factors contributed to the discrepancy between the originally projected and updated catches at F_{targ} across stocks and management periods. In the majority of cases (73%), the initial abundance at age had the largest impact on projection accuracy (Table 4). For some stocks this projection input was the dominant source of error following the NEFSC (2002, 2005, 2008) assessments (GOM cod, GB yellowtail flounder, witch flounder, and plaice), while for other stocks it was the dominant source for two of three projections (GB cod and haddock, GOM cod (*M*-ramp), and SNE–MA and CC–GOM yellowtail flounder). Forecasted re-

Table 3. Mean ratios of the catch and harvest rate (*F*) to the target from 2004 to the most recent year with assessment estimates for each stock.

Stock	No. of SAs	Mean catch (or landings)/target	Mean *F*/target	Error in achieving *F*	Relative error in terminal biomass	Relative error in terminal recruitment
GB cod	4	0.71 (0.4, 1.17)	2.05 (1.24, 2.67)	3.26 (1.57, 6.35)	0.55 (0.14, 0.90)	0.62 (–0.4, 2.05)
GOM cod	7	0.82 (0.52, 1.19)	4.78 (2.03, 8.76)	5.86 (2.45, 11.84)	1.01 (0.0, 2.94)	0.69 (–0.5, 2.77)
GOM cod (*M*-ramp)	7	0.82 (0.52, 1.19)	4.04 (1.74, 7.58)	4.93 (1.79, 10.24)	0.54 (–0.04, 1.92)	–0.14 (–0.77, 0.72)
GB haddock	5	0.37 (0.16, 0.79)	0.95 (0.54, 1.46)	3.14 (1.05, 5.68)	0.36 (–0.11, 0.73)	0.45 (–0.09, 1.63)
GOM haddock	4	1.36 (0.2, 2.99)	0.78 (0.3, 1.05)	0.91 (0.25, 1.44)	–0.34 (–0.41, –0.20)	0.26 (–0.46, 1.32)
GB yellowtail flounder	6	0.70 (0.41, 0.9)	5.37 (1.68, 8.56)	7.49 (3.37, 11.23)	2.13 (1.14, 3.21)	0.90 (–0.68, 1.96)
SNE–MA yellowtail flounder	5	1.20 (0.17, 2.71)	4.70 (2.19, 9.60)	6.47 (1.08, 17.09)	0.93 (0.05, 1.83)	–0.14 (–0.57, 0.83)
CC–GOM yellowtail flounder	5	0.86 (0.5, 1.35)	4.45 (2.41, 6.65)	5.43 (2.78, 8.97)	0.67 (0.14, 1.44)	0.10 (–0.7, 1.08)
Plaice	5	0.53 (0.27, 0.93)	1.04 (0.57, 2.39)	2.18 (0.68, 4.27)	0.59 (–0.23, 1.77)	0.13 (–0.7, 0.92)
Witch flounder	5	0.60 (0.21, 0.94)	3.58 (2.26, 5.2)	7.17 (3.11, 11.54)	1.49 (0.43, 4.08)	2.36 (0.11, 4.59)
GB winter flounder	6	0.66 (0.34, 0.95)	1.52 (0.66, 2.62)	2.37 (1.06, 3.58)	0.53 (0.15, 0.94)	1.62 (0.26, 2.89)
SNE–MA winter flounder	5	0.62 (0.36, 1.04)	1.12 (0.71, 1.57)	1.89 (1.25, 2.6)	–0.17 (–0.47, 0.22)	0.11 (–0.42, 0.41)
Redfish	3	0.34 (0.17, 0.55)	0.31 (0.13, 0.46)	0.92 (0.76, 1.08)	–0.04 (–0.16, 0.08)	1.33 (1.33, 1.33)
White hake	3	0.71 (0.43, 0.92)	1.39 (0.53, 1.79)	1.92 (1.25, 2.32)	0.34 (0.18, 0.50)	0.28 (0.18, 0.38)
Pollock	3	0.51 (0.38, 0.58)	0.47 (0.26, 0.61)	0.90 (0.68, 1.06)	–0.30 (–0.39, –0.22)	–0.22 (–0.78, 0.35)
Mean across stocks		0.71	2.51	3.83	0.64	0.62

Note: Error in achieving *F* is calculated each year for a stock by dividing the mean *F*/target by the mean catch/target (eq. 1), and the mean F_{error} for each stock is calculated using the annual F_{error} estimates. Also shown are the mean relative error in estimates of terminal biomass and recruitment from available assessments through 2008 (NEFSC 2002, 2005, and 2008). For each result, the minimum and maximum values are shown in parentheses. The number of stock assessments (SAs) used to calculate these values is also shown. Estimates from both GOM cod models were combined into a single mean before computing the mean across stocks.

Fig. 2. The relative error in achieving the target *F* (F_{error}; eq. 1) as a function of the number of years between the final year of data in the assessment and the year the catch target was set based on that assessment (called the data management lag, or DML). For example, assessments from NEFSC (2002) had a terminal year of 2001, so catch targets set in 2004 based on this assessment had a DML of 3 years.



cruitment and mass-at-age were the dominant sources of error in four (13%) cases each and were the second most dominant source of error in eight (27%) and 10 (33%) cases, respectively. The assumed selectivity-at-age and assumed interim catches were not the dominant sources of error in any case (Table 4). Although the forecasted recruitment was not the dominant source of projection error in most cases, it did have the largest effect for SNE–MA winter flounder following the NEFSC (2002) and NEFSC (2005) assessments and for SNE–MA and CC–GOM yellowtail flounder following the NEFSC (2002) assessment (Table 4). Error in the forecasted recruitment also contributed to, but was not the dominant source of, the overestimation of target catches in the projection models, indicating below-forecasted recruitment for these stocks. Below-forecasted recruitments during our study period could result from overly optimistic biomass forecasts in cases where recruitment depends on the estimated biomass or from declines in recent recruitment. We used the updated assessment time series of estimated recruitments and calculated the ratio of recent (2004–present) to historical (pre-2004) recruitments. We calculated this ratio for all stocks with recruitment estimates (14) and found that recent recruitments have been below the pre-2004 levels for 10 groundfish stocks (71%), and for six of those stocks (43%), recent mean recruitments were less than half the historical means (Fig. 7). The time period of empirical recruitment estimates used in the projections was not necessarily identical to the time period we used to calculate these ratios. Therefore, our recruitment ratios are not measures of the magnitude of error in forecasted recruitment; rather, they are simply meant to show how recent recruitments compare with historical levels.

Finally, as an illustration of where stock biomass and the values of fishery landings would be if the target harvest rates had been achieved, we projected biomass from 2004 to the most recent year

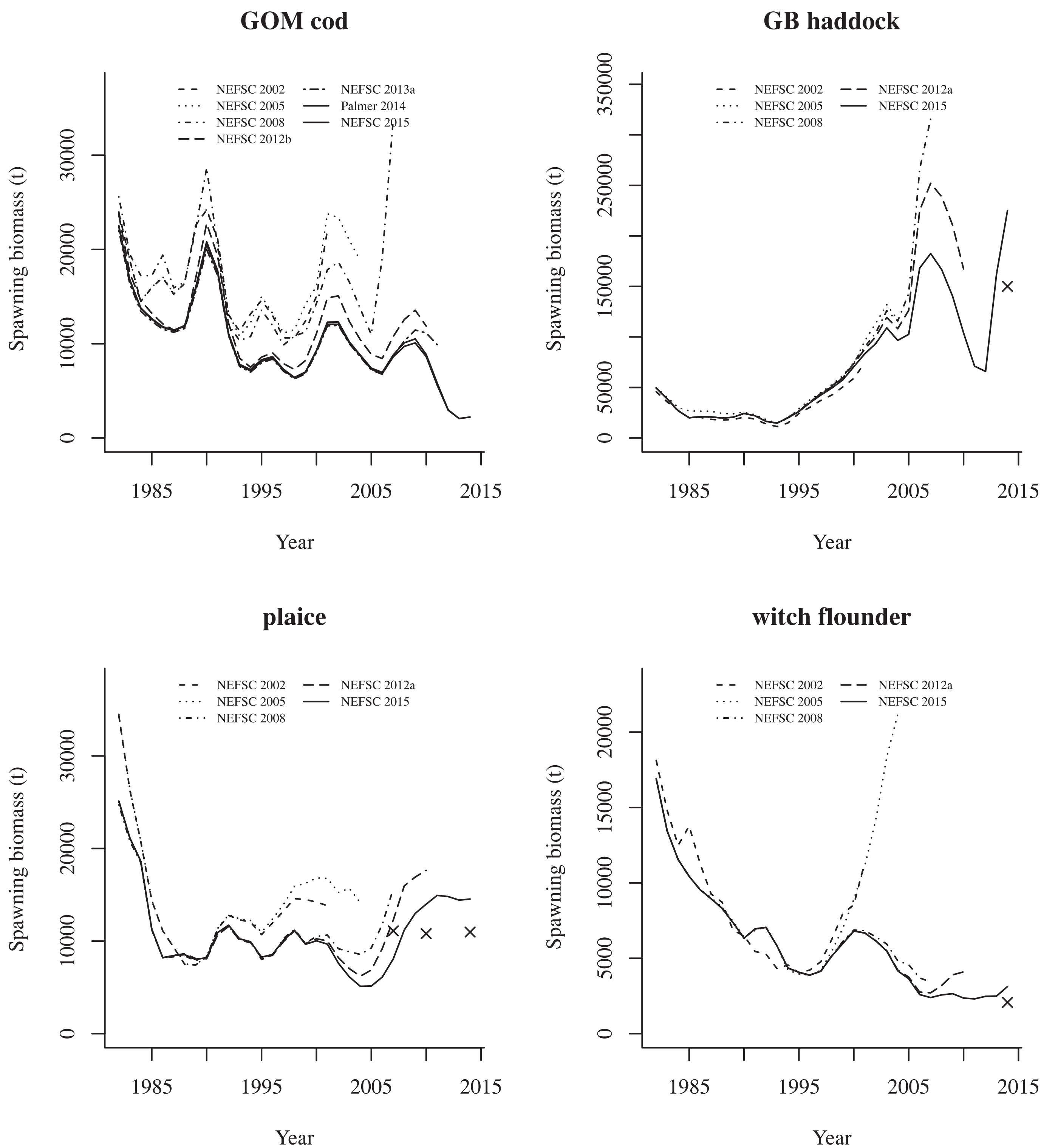


Fig. 3. Estimated spawning biomass across repeated assessments for four groundfish stocks with multiple instances of overestimation of terminal biomass. The solid line in each panel represents the most recent assessment. Biomass estimates shown are the original, unmodified values. When a retrospective adjustment was made to terminal biomass estimate for an assessment, an "×" is shown representing the adjusted biomass estimate.

with estimates for the nine stocks evaluated in the projection update. Under this optimal management scenario, all stocks would have had higher biomass than at present, with four stocks currently undergoing rebuilding being rebuilt (witch flounder, plaice, and SNE–MA yellowtail flounder) and four other stocks showing considerable increases in biomass (GOM cod, GB cod, GB yellowtail flounder, and CC–GOM yellowtail flounder; Fig. 8). For each stock, we multiplied both the observed and projected total catch by the mean ex-vessel price each year (adjusted to 2005 US dollars; Pauly and Zeller 2015) to calculate the difference in total

Fig. 4. Histogram of the relative error in terminal biomass and recruitment aggregated across stocks and across assessments. The dotted and dashed vertical lines represent the median and mean values for the first four assessments (NEFSC 2002, 2005, 2008, 2012*a*), respectively. The lowest possible relative error is −1 based on eq. 2.

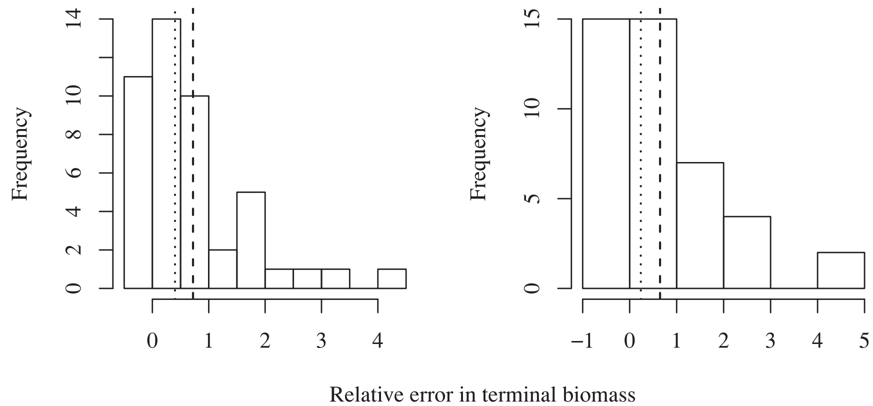
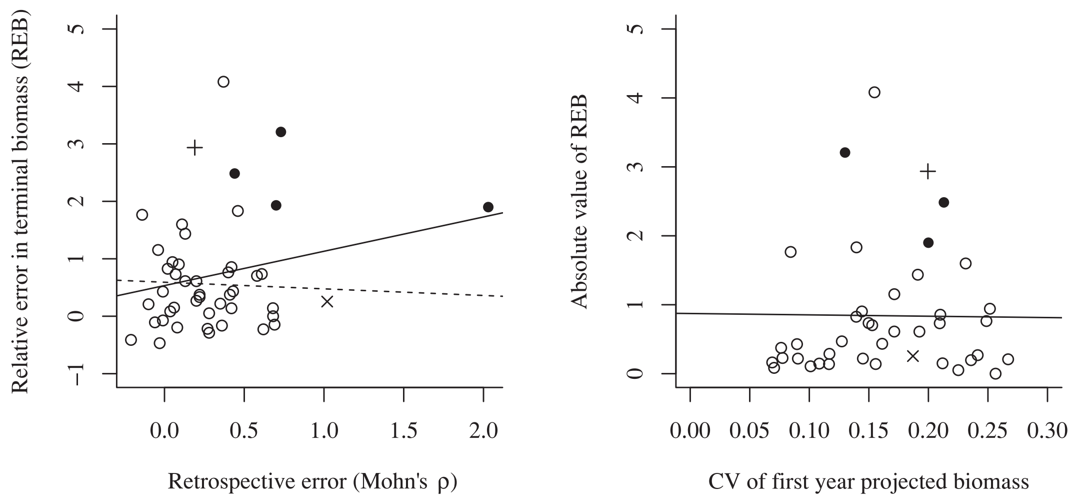


Fig. 5. Measures of uncertainty in terminal estimates from a given assessment (Mohn's ρ and the CV of biomass in the first projected year) were poor indicators of the true uncertainty in terminal biomass, measured as the relative error in the estimate from that assessment. When using the CV of projected biomass as the predictor, we used the absolute value of relative error in estimates because the CV does not indicate direction of the uncertainty, whereas ρ can be positive or negative. The relative error in biomass (REB) was calculated using the original biomass estimates (i.e., not adjusted by the retrospective error) from NEFSC (2002, 2005, 2008, and 2012*a*). The "+" and "×" symbols denote the GOM cod and SNE–MA yellowtail flounder examples referenced in the text, respectively, and the solid circles represent GB yellowtail flounder estimates (one estimate of the CV was not available for this stock, so fewer points are shown in the right panel). The solid black line on the left panel was fit to all data points (REB = 0.53 + 0.6ρ; p = 0.13), while the dashed line was fit removing the point for GB yellowtail where ρ = 2 (REB = 0.83 + 0.9ρ; p = 0.13). The fit using the CV as the predictor was REB = 0.87 − 0.2CV (p = 0.78).

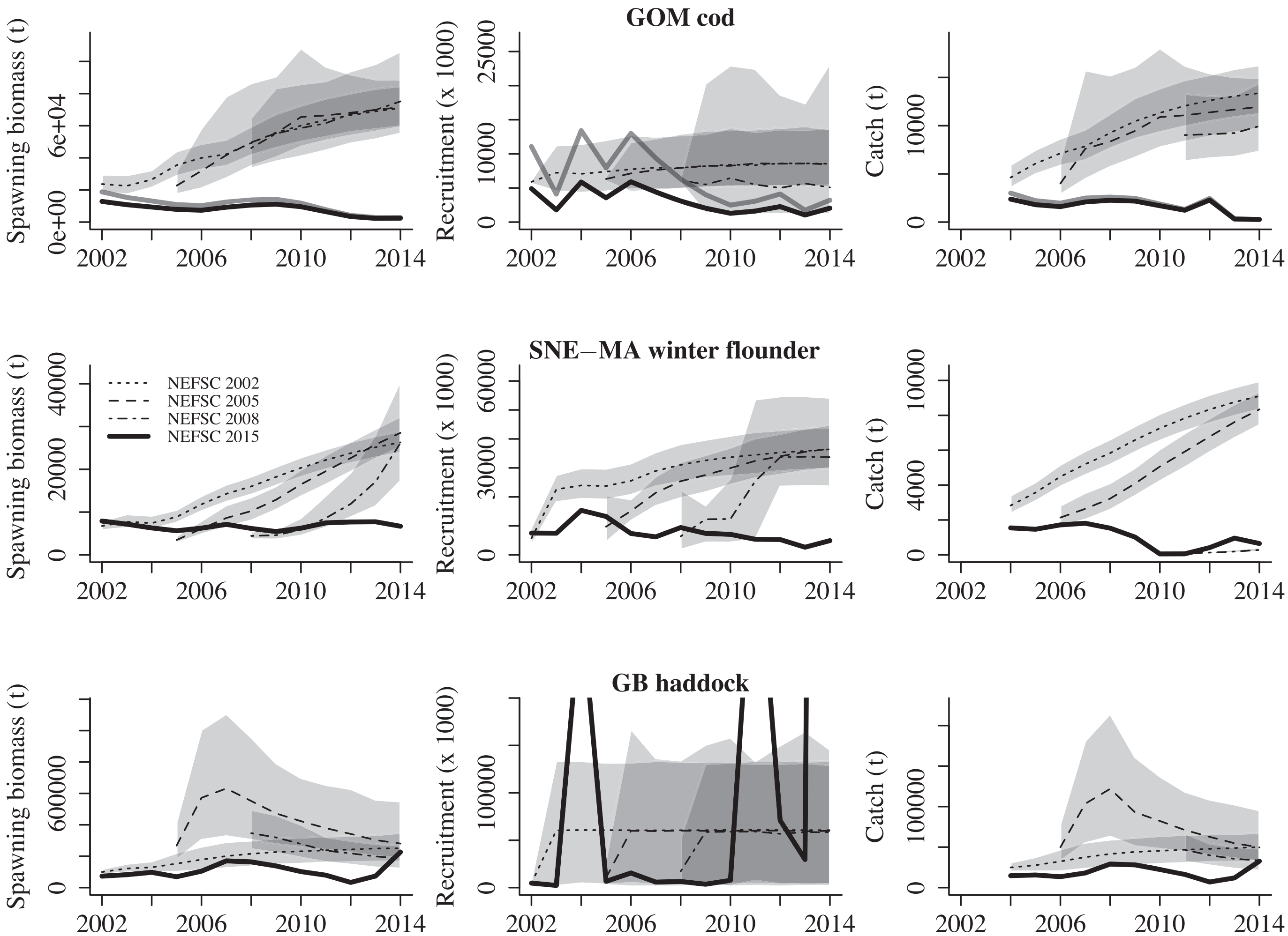


value in landings. Across these nine stocks, total catches by fishing at F_{targ} would have been 45% below the actual aggregate catch in 2004, representing a decrease in ex-vessel value of $54 million ($45 million when GB haddock is excluded). By 2007, catches would have been 4% higher, increasing over the period to 120% higher than the present catch by 2014 across these stocks, representing an increase in revenue in 2014 of $87 million ($50 million excluding GB haddock). Across the entire period, the increased catches would have resulted in an increase in total ex-vessel revenue of $126 million (but only $40 million excluding GB haddock). When using the model for GOM cod that assumes a doubling of *M* in recent years, the impact is reduced total landings (1% to 5% lower) and value (1% to 9.5% lower) since 2007 (Fig. 8). All stocks would have had higher catch than at present, and therefore greater revenue, although the differences were small for SNE–MA winter flounder and yellowtail flounder. This small difference resulted from a small projected increase in biomass for SNE–MA winter flounder due to very poor recruitment (Fig. 8, top panel) and from the fact that the actual current catches are very high (and above C_{targ}) for SNE–MA yellowtail flounder.

Discussion

We evaluated the magnitude and sources of scientific uncertainty in catch advice for stocks in the New England groundfish complex and found that since 2004, annual catches have generally been at or below the target catch in most years for most stocks, yet harvest rates often greatly exceeded the target, resulting from target catches that have been overly optimistic. By evaluating population estimates over time from sequential assessments for each stock, we found a pattern of overestimation of abundance in the terminal year of the assessment. Multiple sources of scientific uncertainty contributed to this overestimation of catch targets, but the largest contributing factor was overestimated terminal abundance in the assessment, which is the starting point in the population projection model used to set future catches. Below-average forecasted recruitments (owing to declines in recent

Fig. 6. Projected spawning biomass (left), recruitment, and total catch for GOM cod (top), SNE–MA winter flounder (middle), and GB haddock (bottom). Dashed lines and shaded regions represent the median and 95% confidence intervals, respectively, for each projection based on output from the 2002, 2005, and 2008 assessments (NEFSC 2002, 2005, 2008). The thick solid black line represents the updated values for biomass and recruitment. The updated target catch was calculated using F_{targ} each year using the updated abundance estimates each year, the fishery selectivity, and catch masses. The first year shown for biomass and recruitment is 2002 because the projections following NEFSC (2002) started that year. Target catches are shown starting in 2004 because this was the first year target catches were calculated under F_{targ} following NEFSC (2002).



recruitment) and mass-at-age also had an effect on the overestimation of catch targets in some cases. The net effect of overly optimistic catch targets and declining recruitment is that rebuilding has been sluggish and potential yield (and revenue) has been forgone for many stocks.

There are numerous potential causes for changes in assessment estimates over time. Example causes include, but are not limited to, revised survey or catch estimates (total and by age class), revised life history information (natural mortality rate, length- and mass-at-age), or changes in model assumptions or structure. Multiple changes to the input data and model are often made between assessments, and such changes are all forms of scientific uncertainty. A full exploration of each change and its impact on the assessment estimates is beyond the scope of this work, but we provide some examples here. For GOM cod between the 2008 and 2012 assessments, updates were made to the recreational and commercial landings and discards at age, the length–mass relationship and mass-at-age estimates, and the survey indices were re-estimated (NEFSC 2012*b*). All of these changes resulted in biomass estimates being between 17% and 37% lower (from 1993 to 2005) in the updated assessment (Fig. 3). However, the largest difference between these assessments was that the 2005 year class did not materialize, estimated at almost 24 million age-1 recruits in NEFSC (2008) but revised downward to 9 million in NEFSC (2012*b*). For pollock, a change in the way the likelihood was calculated in the statistical catch-at-age model between the 2014 and 2015 assessments resulted in a large increase in the magnitude of recruitment estimates over the entire time series, resulting in a roughly 60% increase in biomass estimates over the time period (NEFSC 2015; Fig. S2 here[1]).

Although there are many potential sources of scientific uncertainty in assessment estimates, that the general pattern was to overestimate abundance suggests there may be some common mechanisms in the region influencing assessment estimates. Simulation studies have shown a number of possibilities leading to overestimation of abundance, with increased adult mortality not accounted for in the model being one potential source (Mohn 1999; Deroba and Schueller 2013). Unaccounted adult mortality could result from increased natural mortality, and such a mechanism has been linked to the delayed recovery of many Canadian cod stocks (Shelton et al. 2006). Pershing et al. (2015) noted rapid warming in the Gulf of Maine and suggested it was responsible for increased mortality of maturing GOM cod, although this conclusion has been debated (Palmer et al. 2016; Swain et al. 2016).

Table 4. Projections based on the NEFSC (2002, 2005, and 2008) assessments were rerun with the updated estimates of the initial abundance at age (*N*), catches in the interim period between the terminal year in the assessment and first year of target catches based on that assessment, forecasted recruitment (*R*), forecasted mass-at-age in the catch (w_C), and selectivity-at-age (*s*) from the most recent assessment for each stock.

		Mean catch ratio					
Stock	Assessment year	Original*	Updated *N*	Updated catch	Updated *R*	Updated w_C	Updated *s*
GOM cod	2002	2.04	**1.42**	2.18	1.96	1.82	2.01
	2005	3.43	**1.53**	3.67	3.22	3.12	3.08
	2008	4.23	**1.60**	4.49	3.98	3.74	4.64
GOM cod (*M*-ramp)	2002	1.83	2.35	1.94	1.78	**1.62**	1.76
	2005	3.17	**2.36**	3.44	3.10	2.84	2.87
	2008	4.50	**2.66**	4.76	4.53	3.95	4.93
GB cod	2002	1.41	1.62	1.38	1.23	**1.03**	1.37
	2005	3.34	**1.36**	3.34	3.12	3.02	3.36
	2008	1.53	**1.06**	1.53	1.44	1.45	1.60
GB haddock	2002	1.55	1.59	1.60	1.40	**1.18**	1.26
	2005	3.97	**1.79**	4.01	3.94	2.93	2.75
	2008	1.94	**1.10**	1.94	1.99	1.40	1.99
GB yellowtail flounder	2002	8.41	**2.96**	8.38	7.32	7.90	8.07
	2005	8.28	**1.13**	8.76	6.69	6.09	8.22
	2008	7.79	**1.75**	8.23	6.38	6.45	7.43
CC–GOM yellowtail flounder	2002	2.65	2.02	2.80	**1.99**	2.52	2.14
	2005	3.25	**1.97**	3.37	2.61	3.04	2.76
	2008	3.47	**2.03**	3.44	2.82	3.01	3.41
SNE–MA yellowtail flounder	2002	3.41	3.43	3.58	**1.19**	3.27	3.46
	2005	0.66	**1.18**	0.73	0.64	0.59	0.65
	2008	5.15	**2.05**	5.12	2.60	4.44	4.74
SNE–MA winter flounder	2002	1.45	1.36	1.55	**1.12**	1.51	1.28
	2005	1.37	1.39	1.38	**1.18**	1.28	1.26
	2008	1.06	1.97	1.10	0.96	**1.01**	0.98
Witch flounder	2002	7.63	**1.55**	8.06	7.51	6.99	6.97
	2005	5.26	**1.33**	5.57	5.26	5.15	5.02
	2008	3.29	**1.56**	3.27	3.05	3.26	2.86
Plaice	2002	2.74	**1.45**	3.03	2.72	2.19	2.63
	2005	2.60	**1.22**	2.67	2.58	2.57	2.28
	2008	2.10	**1.33**	2.15	2.07	1.69	2.03

Note: Values in the table are the mean projected catch relative to the mean catch that would have achieved F_{targ} (the projected catch ratio). Values in bold are the updated input that gets the catch ratio closest to 1.0 for each assessment for each stock.

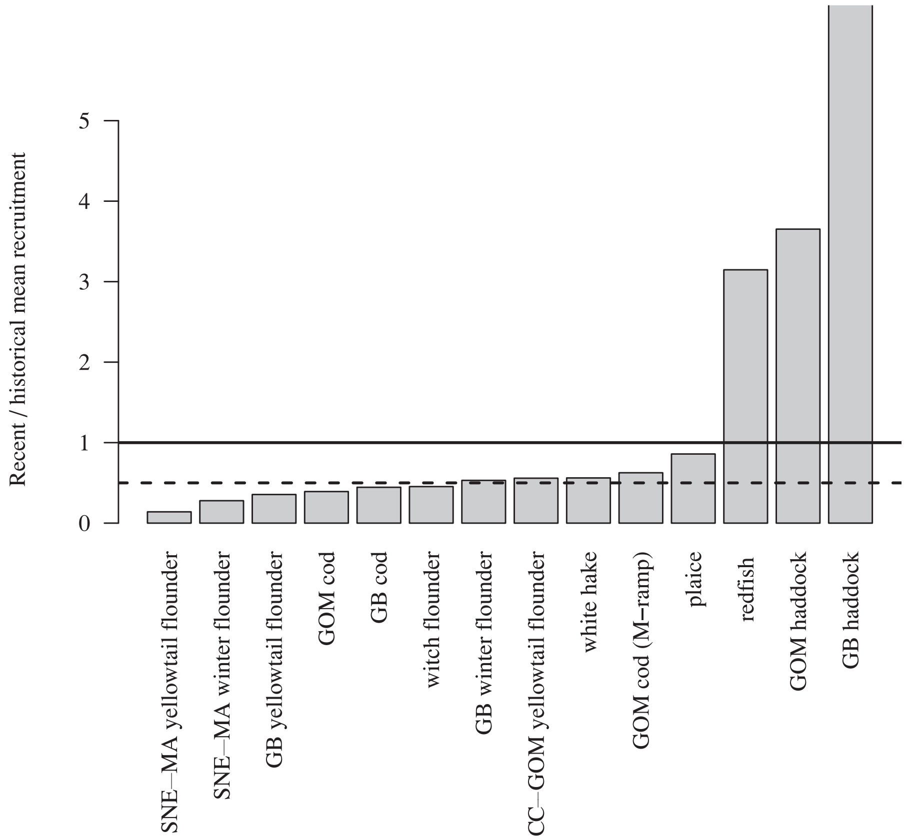
*Original, unmodified projection.

Underreported catches are another potential source of adult mortality, and King and Sutinen (2010) suggested that noncompliance in the groundfish fishery may be considerable, with estimates of illegal harvests between 12% and 24% of the reported total based on surveys of fishermen, enforcement officers, and others involved in the fishery. Underreported catches would also reduce our estimates of F_{error} if the estimates of F_{obs} did not change (i.e., the assessment model estimates higher biomass to account for increased catches). Overestimation of terminal abundance could also result from changes in the distribution of a stock or in the timing of migrations (e.g., Nye et al. 2009; Pinsky et al. 2013) if such changes resulted in more biomass available to the bottom trawl surveys, without changes in total population biomass. A combination of these and other factors may be contributing to the pattern of overestimation we observed, and identifying and addressing these possible sources of scientific uncertainty in the data, assessments models, and projections is essential for the future sustainability of the groundfish fishery.

It is unknown how widespread this pattern of overestimation is beyond groundfish, both within the region and worldwide. Examples of repeated overestimation exist for individual stocks (e.g., Sinclair et al. 1985; Walters and Maguire 1996; Terceiro 2015), but to our knowledge only Ralston et al. (2011) have evaluated assessment estimates across stocks in a region, and they did not identify any patterns of frequent over- or underestimation in assessments for stocks off the western US. Studies that evaluate the performance of assessment estimates and catch targets across management units or broader regions would be beneficial for understanding large-scale patterns possibly controlling assessment accuracy.

We also found reduced recruitment for the majority of groundfish stocks, which often resulted in the recruitment forecasts in the projection models being overestimated. Below-forecasted recruitment did have an impact on the overestimation of projected catch targets, but the contribution was generally small relative to the overestimation of terminal abundance for most stocks. Brooks and Legault (2016) noted forecasted recruitment had a larger effect on the accuracy of projected catch targets after 3 or more years, largely due to when groundfish become available to the fishery. Declining recruitment can have other effects on the sustainability of catch targets. For example, if a target harvest rate is based on a rebuilding projection ($F_{rebuild}$), below-forecasted recruitment would result in overly optimistic rebuilding timelines, and the estimated $F_{rebuild}$ would be biased high. In contrast, increases in recent recruitments can help buffer against other sources of uncertainty. For GB haddock, terminal biomass was repeatedly overestimated, but record year classes (2003, 2010, and likely 2014) have allowed this stock to rebuild (NEFSC 2015). Many studies have looked into the influence of bottom-up environmental variables such as temperature and food availability on recruitment for individual groundfish stocks (e.g., Bell et al. 2014; Friedland et al. 2015; Pershing et al. 2015), and a better understanding of the mechanisms affecting groundfish recruitment is needed. However, attempts to correlate assessment estimates of

Fig. 7. Ratio of the mean recent recruitment (2004–onward) to the mean historical recruitment (pre-2004). Recruitment estimates are based on the most recent assessment for each stock. The solid and dashed lines represent values of 1 and 0.5, respectively. The ratio for GB haddock is >20 due to a very large (and uncertain) estimate of terminal recruitment from the most recent assessment.



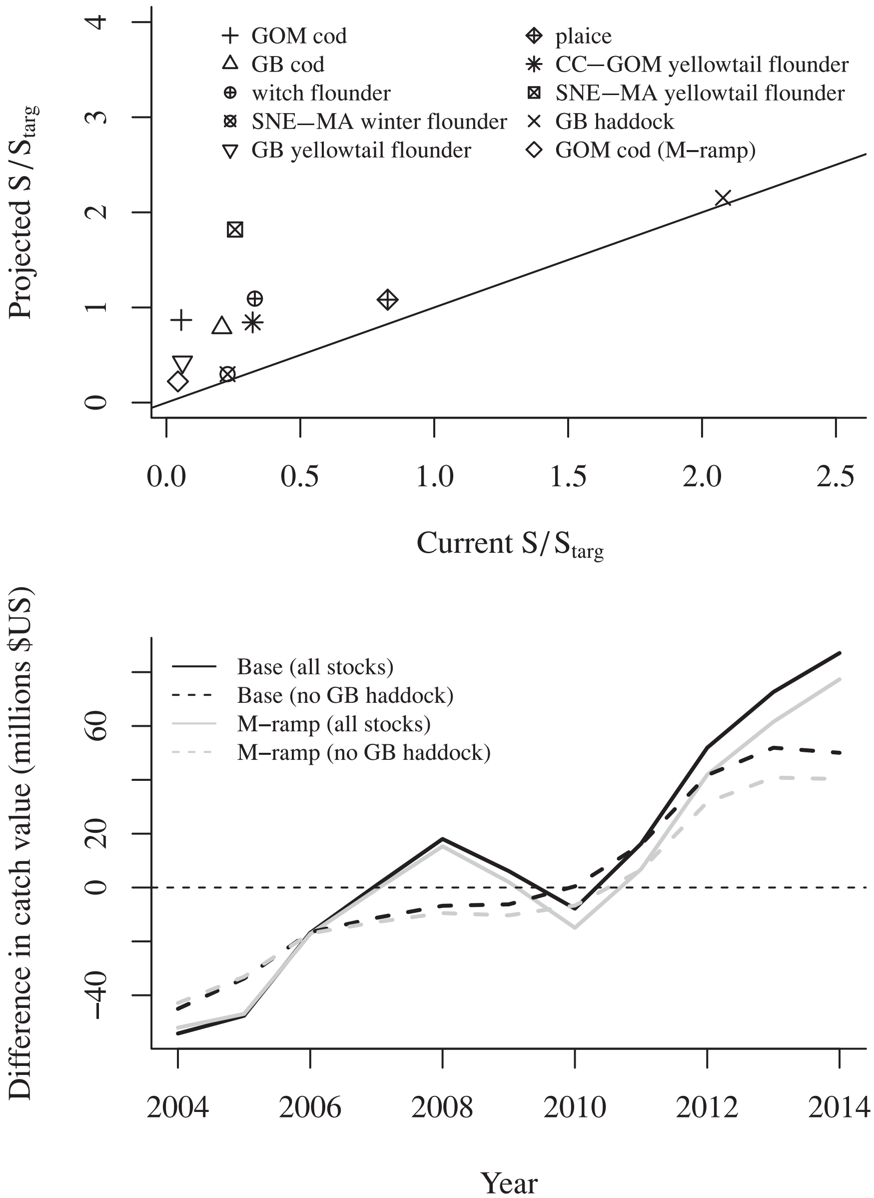
recruitment with environmental variables would benefit from careful consideration of how structural assumptions in the assessment model can influence recruitment estimates, potentially resulting in spurious correlations between recruitment and environmental covariates (Brooks and Deroba 2015). Requiring more frequent assessments could eliminate the need for long-term projections of recruitment (Brooks and Legault 2016), particularly in response to environmental change, but doing so would put an increased burden on the assessment, review, and management systems.

Part of our analysis involved updating projection models to identify the major sources of error, and this approach was similar to the work of Brooks and Legault (2016). There were, however, a few key distinctions between approaches that justified inclusion of this part of our work in the broader analysis. First, Brooks and Legault (2016) used abundance estimates from NEFSC (2008) as their reference for updating the projections, but many of these estimates have changed dramatically with subsequent assessments (see Fig. 3 for examples with GOM cod and plaice). More importantly, Brooks and Legault (2016) did not necessarily evaluate the projections used to set catch targets. Projections in their analysis were based on assessment estimates from a retrospective analysis of the 2008 assessment for each stock, an approach they called retrospective forecasting. For many stocks, there were changes made to the 2008 assessment such that the retrospective estimates differed considerably from the earlier assessment estimates that were the basis for management. In addition, they removed some of the shorter data sets from the assessment models to allow for some of the retrospective runs, and it is unclear what effect this had on the model estimates relative to the original assessment models that were used for management of each stock. Despite these differences, our results were in general agreement with theirs, as they also found that overestimation of terminal abundance had the largest effect on the accuracy of the projections.

A caveat of our work is that we are using updated model estimates to evaluate the performance of past model estimates. The updated models represent the best available science for each stock, but there is uncertainty in the estimates from the updated models, particularly in recent years. Many of the updated assessments had a strong retrospective pattern suggesting overestimation of terminal abundance (NEFSC 2015), such that estimates of abundance in recent years may be revised downward in future assessments. Such a change would likely result in increased estimates of observed harvest rates, which would increase the magnitude of the overage relative to recent years. Future assessments will also likely have major changes in data inputs and model structure and assumptions, and such changes could have a scaling effect on biomass estimates over time (up or down), which occurred for some groundfish stocks in our analysis (see Fig. 3 and Figs. S1–S3[1] for examples). Model changes that alter historical biomass estimates could affect our results, and we recommend that an evaluation of the performance of previous assessment estimates and catch targets be included in the updated assessment (for all stocks, not just New England groundfish) so that managers and scientific advisors can weigh such information in the process of setting future catch limits.

Two stocks included in our analysis (GB yellowtail flounder and GB cod) had histories of age-based assessments, but strong retrospective patterns in recent assessments resulted in these estimates not being used for management. For these two stocks, we relied on the most current assessment that passed review as our

Fig. 8. Projections of biomass and catch under an optimal management scenario where F_{targ} was achieved each year. Top: projected spawning biomass ratio (S/S_{targ}) in the final year of each projection compared with the current biomass ratio. The solid black line represents the 1:1 line, and values above this line indicate higher projected biomass relative to the current state. Bottom: time series of the difference in the ex-vessel value between the optimal management catch and the observed catch. Value was calculated by multiplying the total catch mass by the ex-vessel value per unit mass and was aggregated across the nine stocks shown in the top panel, excluding GB haddock, and for the different population estimates for GOM cod (Base refers to the model with M = 0.2 in all years, and M-ramp refers to the model with M increasing to 0.4 in recent years).



source of updated estimates. The use of older assessment estimates for these stocks would only be problematic for our analyses if the remedies to the current retrospective problem drastically changed historical estimates. For GB cod, the recent assessment was used as the basis for declaring that the stock remained overfished, but the magnitude of the retrospective pattern led to serious concerns about using the model estimate of abundance (adjusted downward based on the estimated ρ) to set catch targets (NEFSC 2015). A range of different age-based and surplus production assessment models were run for GB yellowtail flounder in an attempt to address the retrospective problem as part of a larger modeling exercise for many stocks (i.e., not an official assessment; Deroba et al. 2015). For GB yellowtail flounder, the production models estimated biomass two to six times higher than the age-based models from 1995 onward. If a production model were deemed acceptable for this stock and used as a reference in our analysis, it would indicate past age-based assessments underestimated biomass. However, swept-area biomass estimates for this stock are closer to the estimates from the age-based models (NEFSC 2015), so it is unlikely that a production model would be adopted in the future for this stock.

Our analysis of historical assessment estimates did not reveal a relationship between the within-assessment measures of uncertainty (Mohn's ρ and biomass CV) and the magnitude of relative error in the terminal biomass estimate from that assessment (a measure of across-assessment uncertainty). For ρ, we found that although overestimates of terminal biomass generally occurred for assessments with a positive ρ, large overestimation of terminal biomass occurred for both low and high ρ values. In a simulation study, Hurtado et al. (2015) found no relationship between the magnitude and direction of ρ and the bias in the terminal biomass estimate from the assessment across a range of life histories and different sources of assessment error. In contrast, Miller and Legault (2017) developed an approach for estimating uncertainty in ρ and found through simulations that bias estimates were generally within the confidence bounds for ρ, although this work focused only on misspecification of M for GB yellowtail flounder. Our findings are in general agreement with Hurtado et al. (2015), albeit not based on simulations, and suggest that care is needed when trying to interpret bias in terminal estimates from the magnitude and direction of ρ. However, as Miller and Legault (2017) showed, there may be cases where ρ is informative about bias, and further research into this area is warranted.

In summary, we have identified a pattern of overly optimistic catch targets for New England groundfish since 2004, resulting largely from uncertainty in terminal abundance estimates from age-based assessment models. The question remains, however, as to what can be done to prevent this pattern from continuing? The long-term solution is to identify the major sources of uncertainty and address them in data sets and in the model. In the interim, remedies that can be implemented rapidly are needed to limit overfishing for groundfish in the short term. In future work we will explore the effectiveness of a range of methods that can be applied in the absence of knowledge of the underlying mechanisms causing the overestimation of catch targets. Example methods include alternative control rules, abundance modifications, and modified recruitment forecasts, and we will compare the ability of the target catches estimated from these alternative approaches at reducing the magnitude and frequency of overfishing.

Acknowledgements

We thank Tom Nies and Jamie Cournane from the New England Fishery Management Council for providing all the necessary management and assessment documents and projection files, for helpful discussions throughout the scope of this project, and for comments on earlier versions of this work. We also thank Liz Brooks and Paul Nitschke from the Northeast Fisheries Science Center and Holly Kindsvater, Doug Zemeckis, and Becca Selden from Rutgers University for comments on previous drafts of this work. We thank an anonymous reviewer and Chris Legault for comments on this manuscript and again thank Chris Legault for providing files and answering questions throughout the course of this work. We are grateful to members of the Scientific and Statistical Committee of the New England Fishery Management Council for helpful discussions. This work was funded by the New England Fishery Management Council, and the views expressed herein are not necessarily those of the Council.

References

Bell, R.J., Hare, J.A., Manderson, J.P., and Richardson, D.E. 2014. Externally driven changes in the abundance of summer and winter flounder. ICES J. Mar. Sci. **71**(9): 2416–2428. doi:10.1093/icesjms/fsu069.

Brodziak, J., Rago, P., and Conser, R. 1998. A general approach for making short-term stochastic projections from an age-structured fisheries assessment model. *In* Proceedings of the International Symposium on Fishery Stock Assessment Models for the 21st Century. *Edited by* F. Funk, T. Quinn, II, J. Heifetz, J. Ianelli, J. Powers, J. Schweigert, P. Sullivan, and C.-I. Zhang. Alaska Sea Grant College Program, University of Alaska, Fairbanks.

Brooks, E.N., and Deroba, J.J. 2015. When "data" are not data: the pitfalls of post hoc analyses that use stock assessment model output. Can. J. Fish. Aquat. Sci. **72**(4): 634–641. doi:10.1139/cjfas-2014-0231.

Brooks, E.N., and Legault, C.M. 2016. Retrospective forecasting — evaluating performance of stock projections for New England groundfish stocks. Can. J. Fish. Aquat. Sci. **73**(6): 935–950. doi:10.1139/cjfas-2015-0163.

Deroba, J.J., and Schueller, A. 2013. Performance of stock assessments with misspecified age- and time-varying natural mortality. Fish. Res. **146**: 27–40. doi:10.1016/j.fishres.2013.03.015.

Deroba, J.J., Butterworth, D.S., Methot, R.D., Jr., De Oliveira, J.A.A., Fernandez, C., Nielsen, A., Cadrin, S.X., Dickey-Collas, M., Legault, C.M., Ianelli, J., Valero, J.L., Needle, C.L., O'Malley, J.M., Chang, Y.-J., Thompson, G.G., Canales, C., Swain, D.P., Miller, D.C.M., Hintzen, N.T., Bertignac, M., Ibaibarriaga, L., Silva, A., Murta, A., Kell, L.T., de Moor, C.L., Parma, A.M., Dichmont, C.M., Restrepo, V.R., Ye, Y., Jardim, E., Spencer, P.D., Hanselman, D.H., Blaylock, J., Mood, M., and Hulson, P.-J.F. 2015. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. ICES J. Mar. Sci. **72**(1): 19–30. doi:10.1093/icesjms/fst237.

Friedland, K.D., Leaf, R.T., Kristiansen, T., and Large, S.I. 2015. Layered effects of parental condition and larval survival on the recruitment of neighboring haddock stocks. Can. J. Fish. Aquat. Sci. **2**(11): 1672–1681. doi:10.1139/cjfas-2015-0084.

Hendrickson, L., Nitschke, P., and Linton, B. 2015. 2014 Operational Stock Assessments for Georges Bank winter flounder, Gulf of Maine winter flounder, and pollock [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 15-01. Available from National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/publications/. doi:10.7289/V5930R3C.

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., and Punt, A.E. 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES J. Mar. Sci. **72**(1): 99–110. doi:10.1093/icesjms/fsu198.

King, D.M., and Sutinen, J.G. 2010. Rational noncompliance and the liquidation of Northeast groundfish resources. Mar. Pol. **34**: 7–21. doi:10.1016/j.marpol.2009.04.023.

Legault, C.M., Alade, L., and Stone, H.H. 2011. Stock sssessment of Georges Bank yellowtail flounder for 2011 [online]. TRAC Ref. Doc. 2011/01. Available: http://www.nefsc.noaa.gov/saw/trac/.

Legault, C.M., Alade, L., Gross, W.E., and Stone, H.H. 2013. Stock assessment of Georges Bank yellowtail flounder for 2013 [online]. TRAC Ref. Doc. 2013/01. Available from http://www.nefsc.noaa.gov/saw/trac/.

Mace, P.M. 2001. A new role for MSY in single-species and ecosystem approaches to fisheries stock assessment and management. Fish Fish. **2**: 2–32. doi:10.1046/j.1467-2979.2001.00033.x.

Miller, T.J., and Legault, C.M. 2017. Statistical behavior of retrospective patterns and their effects on estimation of stock and harvest status. Fish. Res. **186**: 109–120. doi:10.1016/j.fishres.2016.08.002.

Mohn, R. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES J. Mar. Sci. **56**: 473–488. doi:10.1006/jmsc.1999.0481.

NOAA. 1996. Magnuson–Stevens Fishery Conservation and Management Act. Technical memorandum NMFS-F/SPO-23, National Oceanic and Atmospheric Administration, Silver Spring, Md.

NEFMC. 2003. Amendment 13 to the Northeast Multispecies Fishery Management Plan [online]. New England Fishery Management Council. Available from http://www.nefmc.org/nemulti/index.html.

NEFMC. 2006. Framework Adjustment 42 to the Northeast Multispecies Fishery Management Plan, Including an Environmental Assessment, Regulatory Impact Review, Initial Regulatory Flexibility Analysis [online]. Newburyport, Mass. New England Fishery Management Council. Available from http://www.nefmc.org/nemulti/index.html.

NEFMC. 2010. Framework Adjustment 44 to the Northeast Multispecies Fishery Management Plan, Including an Environmental Assessment, Regulatory Impact Review, Initial Regulatory Flexibility Analysis [online]. Newburyport, Mass. New England Fishery Management Council. Available from http://www.nefmc.org/nemulti/index.html.

NEFMC. 2011. Framework Adjustment 45 to the Northeast Multispecies Fishery Management Plan, Including an Environmental Assessment, Regulatory Impact Review, Initial Regulatory Flexibility Analysis [online]. Newburyport, Mass. New England Fishery Management Council. Available from http://www.nefmc.org/nemulti/index.html.

NEFMC. 2012. Framework Adjustment 47 to the Northeast Multispecies Fishery Management Plan, Including an Environmental Assessment, Regulatory Impact Review, Initial Regulatory Flexibility Analysis [online]. Newburyport, Mass. New England Fishery Management Council. Available from http://www.nefmc.org/nemulti/index.html.

NEFMC. 2013. Framework Adjustment 50 to the Northeast Multispecies Fishery Management Plan [online]. Newburyport, Mass. New England Fishery Management Council. Available from http://www.nefmc.org/nemulti/index.html.

NEFMC. 2014. Framework Adjustment 51 to the Northeast Multispecies Fishery Management Plan [online]. Newburyport, Mass. New England Fishery Management Council. Available from http://www.nefmc.org/nemulti/index.html.

NEFSC. 2002. Assessment of 20 Northeast groundfish stocks through 2001: a report of the Groundfish Assessment Review Meeting (GARM), Northeast Fisheries Science Center, Woods Hole, Massachusetts, 8–11 October 2002 [online]. Northeast Fish. Sci. Cent. Ref. Doc. 02-16. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/saw/.

NEFSC. 2003. Report of the 36th Northeast Regional Stock Assessment Workshop (36th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fisheries Science Center Reference Document 03-06 [online]. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/saw/.

NEFSC. 2005. Assessment of 19 Northeast groundfish stocks through 2004. 2005 Groundfish Assessment Review Meeting (2005 GARM), Northeast Fisheries Science Center, Woods Hole, Massachusetts, 15–19 August 2005 [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 05-13. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/saw/.

NEFSC. 2008. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, 4–8 August 2008 [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 08-15. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2010. 50th Northeast Regional Stock Assessment Workshop (50th SAW) Assessment Report [oline]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 10-17. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2011. 52nd Northeast Regional Stock Assessment Workshop (52nd SAW) Assessment Report [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 11-17. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2012*a*. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010 [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 12-06. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2012*b*. 53rd Northeast Regional Stock Assessment Workshop (53rd SAW) Assessment Report [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 12-05; 559 pp. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2012*c*. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Report [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 12-18. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2013*a*. 55th Northeast Regional Stock Assessment Workshop (55th SAW) Assessment Report [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 13-11. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2013b. 56th Northeast Regional Stock Assessment Workshop (56th SAW) Assessment Report [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 13-10. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2014. 59th Northeast Regional Stock Assessment Workshop (59th SAW) Assessment Summary Report [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 14-07. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NEFSC. 2015. Stock Assessment Update of 20 Northeast Groundfish Stocks Through 2014 [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 15-24. Available from Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

NSWG. 2009. The Northeast Data Poor Stocks Working Group Report, December 8-12, 2008 Meeting. Part A. Skate species complex, deep sea red crab, Atlantic wolffish, scup, and black sea bass [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 09-02. Available from Northeast Stocks Working Group, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/publications/.

Nye, J.A., Link, J.S., Hare, J.A., and Overholtz, W.J. 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Mar. Ecol. Prog. Ser. **393**: 111–129. doi:10.3354/meps08220.

Palmer, J., Deroba, J., Legault, C.M., and Brooks, E.N. 2016. Comment on "Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery." Science, **352**: 423. doi:10.1126/science.aad9674. PMID:27102473.

Palmer, M.C. 2014. 2014 Assessment update report of the Gulf of Maine Atlantic cod stock [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 14-14. Available from National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/publications/.

Pauly, D., and Zeller, D. (*Editors*). 2015. Sea around us concepts, design and data [online]. Available from www.seaaroundus.org.

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., and Thomas, A.C. 2015. Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. Science, **350**: 809–812. doi:10.1126/science.aac9819. PMID:26516197.

Pinsky, M.L., Worm, B., Fogarty, M.J., Sarmiento, J.L., and Levin, S.A. 2013. Marine taxa track local climate velocities. Science, **341**: 1239–1242. doi:10.1126/science.1239352. PMID:24031017.

R Core Team. 2015. R: a language and environment for statistical computing [online]. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/.

Ralston, S., Punt, A.E., Hamel, O.S., DeVore, J.D., and Conser, R.J. 2011. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fish. Bull. **109**: 217–231.

Rosenberg, A.A., Swasey, J.H., and Bowman, M. 2006. Rebuilding US fisheries: progress and problems. Front. Ecol. Environ. **4**(6): 303–308. doi:10.1890/1540-9295(2006)4[303:RUFPAP]2.0.CO;2.

Rothschild, B.J., Keiley, E.F., and Jiao, Y. 2014. Failure to eliminate overfishing and attain optimal yield in the New England groundfish fishery. ICES J. Mar. Sci. **71**(2): 226–233. doi:10.1093/icesjms/fst118.

Shelton, P.A., Sinclair, A.F., Chouinard, G.A., Mohn, R., and Duplisea, D.E. 2006. Fishing under low productivity conditions is further delaying recovery of Northwest Atlantic cod (*Gadus morhua*). Can. J. Fish. Aquat. Sci. **63**(2): 235–238. doi:10.1139/f05-253.

Sinclair, M., Anthony, V.C., Iles, T.D., and O'Boyle, R.N. 1985. Stock assessment problems in Atlantic herring (*Clupea harengus*) in the Northwest Atlantic. Can. J. Fish. Aquat. Sci. **42**(1): 888–898. doi:10.1139/f85-113.

Sosebee, K., Traver, M., and Mayo, R. 2006. Aggregate resource and landings trends [online]. Available from www.nefsc.noaa.gov/sos/agtt/.

Swain, D.P., Benoît, H.P., Cox, S.P., and Cadigan, N.G. 2016. Comment on "Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery". Science, **352**: 423. doi:10.1126/science.aad9346. PMID:27102474.

Terceiro, M. 2015. Stock assessment update of summer flounder for 2015 [online]. US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 15-13. Available from National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA, or online at http://www.nefsc.noaa.gov/publications/.

Walters, C., and Maguire, J.-J. 1996. Lessons for stock assessment from the northern cod collapse. Rev. Fish Biol. Fish. **6**: 125–137. doi:10.1007/BF00182340.