<div align="center">

Halibut Assessment Report for 2017
Draft
Paul J. Rago
for
New England Fishery Management Council
Last updated: December 1, 2017

</div>

Executive Summary

- Survey, landings and discard estimates were updated for the US stock area.
- Alternative measures of abundance from Maine sources were considered as potential measures of stock trend.
- A review of data poor methods suggests that most have limited utility for Atlantic halibut, however the DCAC model was considered further.
- A ratio method (Rcrit) was developed using randomization methods. Simulation tests suggested that the method had utility as a robust measure of population change and the significance of these changes.
- Application of the Rcrit method to the US and DFO stocks suggest comparable increases of about 9 to 12% per year since the early 2000's
- An "Envelope Method" was applied to estimate relative scale of the population. The Envelope consists of upper and lower bounds of relative abundance that jointly satisfy constraints on abundance based on a range of hypothesized historical fishing mortality and survey catchability estimates.
- Results of the Rcrit and Envelope method were combined to improve the DCAC model but its overall performance is considered unreliable and still governed by strong assumptions.
- A catch forecasting algorithm was developed based on the observed rates of change in one or more indices of relative abundance. The method resembles algorithms commonly used for control of linear systems in engineering applications. The magnitude of catch adjustment depends on the aggregate rate of change in one or more abundance indices in prior years. The method estimates the first and second derivative of population change using loglinear regression. The second derivative is approximated as the difference between successive n-point regressions.
- The method, termed the First and Second Derivative (FSD) model was tested via simulation of a wide variety of initial conditions and trends in stock productivity.
- No simple solution exists but instead the performance should be evaluated with respect to the risk of overfishing, magnitude and variability of projected catch, and the probability of continued increases in stock size.
- FSD model results suggest that the 2018 Atlantic halibut catch should be about 116-118 mt.
- A bootstrap approach was used to compute the uncertainty about the FSD prediction.

- Application of the FSD model to Atlantic halibut and Pacific halibut stocks assessed with advanced statistical catch at age models suggest reasonable agreement between Observed and predicted TACs.

Background and Overview

This is a "Plan B" assessment. It is a consequence of the unsuitability of the existing benchmark approach as a basis for deriving suitable catch limits for resource management. The Plan B assessment was conceived as an interim approach that could bridge the gap between the rejection of an existing methodology and a revised approach. The revised approach will be the product of a benchmark assessment that reviews all the existing information and potential modeling approaches and undergoes extensive levels of peer review. The distinctions between Plan B interim approaches and benchmark assessments are defined by various reports of the Northeast Regional Coordinating Committee (NRCC) (see NRCC 2011) and by precedents accrued over deliberations within the NRCC since 2012.

In general terms, the Plan B process constrains the introduction of new analytical assessment models whose applicability to the stock under consideration has not be subjected to extensive peer review. Practical application of stock assessment models typically requires numerous decisions about the definition of the "stock", inclusion of available data, reliance of parameterization on literature values (e.g., natural mortality rates), plausibility of critical assumptions, and appropriate numerical methods. Such decisions usually benefit from the inclusion of a broad range of expert opinions. In view of the long-term biological and economic consequences of such decision, the benchmark process can lead to greater acceptance of model-based management decisions.

The Plan B process also implies restrictions on the introduction of new time series of indices or changes in estimation methodology. For example, use of time-series that have not been rigorously reviewed for applicability to stock assessments could be problematic, especially if such data are highly influential. Similarly, model parameters that are highly influential, but weakly supported by empirical data (e.g., natural mortality) are typically considered outside the range of Plan B assessments. Changes to the methods for estimating relative abundance or total removals have been allowed but are subject to a case-by-case examination. For example, revision of previous discard estimation methods to a the SBRM method has been allowed. Changes to discard mortality rates (i.e., the post capture survival of release fish) have been incorporated in several assessments (e.g., skates) when strong empirical evidence has been available.

An important management concern arises when the stock is in a rebuilding program. The rejection of the model on which the rebuilding program was based puts rebuilding in a limbo wherein the existing target biomass, rebuilding time-frame, and target fishing mortalities are

obsolete. The management and legal processes are ill-suited to such reversals given the difficulties of putting existing targets in abeyance while new ones are derived.   A rebuilding program without a target is ambiguous at best. For some species, the contrast between underlying biology (especially growth, longevity, geographic range) and existing population structure is sufficient to proceed with a rebuilding program, even when target biomasses and fishing mortality rates are unknown.

The preceding general issues aptly describe the particular issues for Atlantic halibut in US waters.   The previous stock assessment model has been rejected and data to support a true benchmark analytical assessment are insufficient. The rebuilding target is officially defined for 2056 based on the perceived depleted stock abundance and expectations of slow growth and low recruitment. Harvests are restricted to a one fish per trip limit over the federal fishery but a more liberal harvest regulation is allowed in Maine state waters. Estimated discards constitute a large fraction of the fishery removals. While the stock in US waters is considered to be depleted, the immediately adjacent stock in Canada (NAFO areas 3NOPs4VWX5Zc) has been certified by MSC as sustainably harvested. Given the abundant tagging evidence of migrations of fish between the US and Canada, debates about stock definition will likely be a major component of a future benchmark assessment (Shackell et al. 2016, Tryzinski and Bowen 2016, Seitz et al. 2016).

One of the key objectives of this assessment is to employ intuitive and understandable approaches backed by theory and simulation testing. The methodology does not purport to develop biological reference points for a stock that has, by all accounts, declined considerably from a virgin stock size inferred to have existed about 200 years ago (Lear 1998). Given the massive change in fisheries, ecological and environmental conditions since then, it is unlikely that the present environment would support such biomasses in the short term. Or if it could, that a singular focus on catch reductions in halibut would be sufficient to achieve rebuilding to historic levels.   The large historic stocks of halibut were also a function of lightly exploited stocks of many other species and unknown predator-prey and competitive relationships with these species.

The objectives of the approach herein are much more modest. Much of the available data suggests that the stock of halibut residing in US waters is increasing. Encounter rates in nontargeted fisheries and various fishermen reports support such perceptions. Various fishery independent surveys weakly support this hypothesis, however none of the current trawl surveys are efficient at capturing halibut.   Perhaps the most compelling evidence for stock increase is the rapid changes observed in the Canadian stock. Large changes have occurred not only in targeted longline surveys but also in their trawl surveys which have low capture efficiency similar to US bottom trawl surveys

Stock structure and joint management of Atlantic halibut is well beyond the purview of this assessment report. One cannot deny the temptation to note that the current stock boundaries bisect habitats that are thought to be equivalent on either side of national boundaries. Moreover, tagging studies reveal widespread movements (>3000 km, Scott and Crossman 1988) of halibut. More recent data for electronically tagged fish at liberty for up to 210 days reveal maximum travel distances of about 200 km (Seitz et al. 2016). Conventional tagging studies can generate equivocal evidence unless differences in relative fishing effort and report rates of encountered tags are considered in the analyses of perceived migration patterns. Even Data Storage Tags can be problematic if they are not reported by fishermen.

Officially, halibut are in a rebuilding program with a target completion year of 2056. Increases in halibut abundance in US waters are desirable for both legal and economic reasons. Canada's stock has increased rapidly over the past 20 years but their recovery was also preceded by a long period of low abundance and catches. Comparisons shown later in this report reveal comparable patterns in the US stock area, although the baseline begins from a much lower relative abundance.

Under a rebuilding requirement, a desirable harvest control rule is one that does not reduce the rate of increase or decrease the chances for continued abundance increases. A desirable harvest rule should also avoid being overly restrictive. Increases in abundance that arise under contemporary rates of harvest will, under certain restrictive assumptions, continue to occur if the same harvest rate is applied in future years. These restrictive assumptions include constancy of recruitment, natural mortality and growth in future years. Of course, none of these factors are constant, so it is equally important that the aggregate effect of these processes is taken into account when catch limits are set. More critically, catch limits should be responsive to changes over time. Failure to increase catch limits, particularly when catches are driven largely by discards, may lead to accountability measures induced entirely by unavoidable encounters. At the limit, even elimination of all landings may be insufficient to achieve target catch. Failure to decrease catches when indicated can lead rapid increases in fishing mortality, and the loss of biomass accrued during the rebuilding period.

So the overall objective of the methodology described herein is to use readily available empirical data to adjust catches consistent with changes in relative abundance. The methodology is designed to be responsive and in particular, to be sensitive to metrics of changes in underlying productivity. Lessons learned in control of engineering problems suggest that slow responses to signals are one of the most difficult problems to overcome. As examples, low response times often lead to wide temperature swings in HVAC-controlled buildings and production losses in chemical plants. Simulation studies presented in this report support the need for regular updates of stock status information.

__Biology Review__

Among the world's flatfishes Atlantic halibut (*Hippoglossus hippoglossus*) attains the largest size (~2.5 m, but values up to 4.3 m are reported in fishbase.org) and oldest age (>40 y)(Seitz et al. 2016, Armsworthy and Campana 2010)). Maturation for females occurs at about age 9 (Armsworthy and Campana 2010). Despite their well-known history of serial depletion in US waters (Grasso 2008), Trzcinski and Bowen (2016) argue that rapid growth and high fecundity make the population resilient and capable of recovering quickly from a depleted state. Their modeling work suggested that reductions in landings quotas, increases in minimum sizes and reductions in otter trawl fleets were primary factors leading to the rebuilding of Atlantic halibut in Canada. Col and Legault (2009) provide an excellent summary of the history of halibut fishing in the US.

Shackell et al (2016) recently evaluated the distribution of imputed halibut habitat in US and Canadian waters and found occupancy rates four times higher in Canadian waters. Their analyses suggest finer scale stock structure than commonly assumed.   Seitz et al. (2016) reporting on recent electronic tagging results also suggest that the failure of a concomitant increase in US compared to Canada may be due to different stock structures.   Decisions about stock structure are among the most important in stock assessments and this topic will not be considered further in the Plan B assessment.

Past US YPR models (Brodziak 2002) used a natural mortality rate M =0.1 whereas the Canadian assessment model sets M=0.15.

__Available Data__

This report includes updated estimates of catch and discards consistent with estimates provided in past assessments, most recently in Hennen (2015). Estimates for spring and fall NEFSC and ME-NH bottom trawl surveys are also included. Several indices of commercial fishing catch per unit effort are also considered. The focus of this assessment is the interpretation of trends in recent years.

A succinct summary of the primary data considered in this assessment is found in Table1 for the years 2002-2016. Relevant data include recent landings, discards and total catch (FIG 1). Discards by gear type (Table 1.5) revealed that most discards were incidental takes in trawl fisheries until about 2009. Since then the proportion of gill net discards has increased to about 50% of the total.   Estimates of average numbers and weight per tow in the NEFSC spring and fall bottom trawl surveys are summarized in FIG. 2. Catches are near detection limits over much of the survey period and inter-annual fluctuations were very large in the prior to about 2000. Abundance indices that rely on monitoring of commercial fisheries are depicted in FIG.3 for d/k ratios in observed trawl and gill net trips.   These raw estimates are expanded to estimate total discards of halibut using the SBRM approach (Wigley et al. 2008). FIG.3 also includes trends in

ME-NH inshore bottom trawl survey. A standardized CPUE for long line fisheries in Maine was developed by Hansell,et al. (2017) and is shown with error bounds in FIG.4.

Data from an inshore sentinel longline and jig survey originally developed for monitoring cod were received from U. Maine (courtesy of Maddie Rodrigues and Yong Chen). The time series is relatively short (4 years).   It was not possible to resolve key questions about survey methodology for inclusion in this report. The data may be useful as a measure of trend in a future assessment, particularly if state space modeling approaches described by Webster (2017) for Pacific halibut could be developed for Atlantic halibut

Trends in discard rates can be examined for several different measures of effort. In FIG. 5 discard rates in the gill net fishery are depicted as discards per trip, discards per days absent and discard per kept all. All metrics show a striking rise in discard rates (measured on a 6 month interval) beginning about 2002. For observed trawl trips the trends are remarkably similar (FIG. 6).   Data depicted in FIG.5 and 6 were not used in the assessment directly but are shown at a finer temporal resolution to illustrate the consistency of encounters in fishing gears not directing on Atlantic halibut.

Relative abundance data for the Canadian 3NOPs4VWX5Zc stock are not part of this assessment but their results (courtesy of Nell den Heyer DFO, Halifax) are examined for coherence with trends observed in the US.   All of the primary abundance indices in Canada show consistent increases since 2002 (FIG. 7). Trends in US surveys (FIG. 8) have generally increased but not as consistently as those in Canada. Correlations among the US and Canadian abundance indices (FIG. 9) show surprisingly good coherence given the large differences in scale and basis for these observations. Coherence among the Maine-based estimators of relative abundance is a little lower (FIG.10) with some slightly negative correlations for some indices. Similarly,correlations among the NEFSC and Maine-based indices (FIG. 11) is spotty, but this may overly pessimistic owing to the inclusion of all years.

Management Changes
Nies and Cournane (NEFMC, pers. Comm) summarized the major changes in regulations from 2001 to 2017. While many of the effort control measures could have reduced fishing mortality on Atlantic halibut, there have been only two measures directly related to halibut.   In 2009 the one fish per trip regulation was put in place. Amendment 16 later implemented an increase in minimum size from 36 to 41 inches for the 2010 fishing year, beginning on May 1. Such changes would be expected to increase the discard rate, all thing being equal.

Previous Assessment Models

Prior to 2008 Atlantic halibut were assessed using index methods utilizing the NEFSC fall bottom trawl survey. Col and Legault (2009) succinctly summarized the early assessment history of halibut as follows:

> "In previous index-based assessments (NEFSC 2001; Brodziak 2002, Brodziak and Col 2005), Northeast Fisheries Science Center (NEFSC) autumn weight per tow survey indices were expanded to swept-area biomass estimates, and the 5-year average biomass index was compared to B$_{MSY}$ proxy reference points for status determination (FIGure 3). Reference points for Atlantic halibut were originally determined by the New England Fisheries Management Council (Applegate et al. 1998) using Canadian Atlantic halibut length-weight equations (McCracken 1958) and von Bertalanffy growth curves (Nielson and Bowering 1989) to perform yield per recruit (YPR) and biomass per recruit analyses. Natural mortality was assumed to be 0.1, and a Maximum Sustainable Yield (MSY) proxy was chosen to be 300 mt, yielding a B$_{MSY}$ proxy =5400 mt, a ½ B$_{MSY}$ proxy = 2700 mt, and an F$_{MSY}$ proxy (threshold) = F$_{0.1}$ = 0.06. Based on the Groundfish Assessment Review Meeting (GARM) 2005 assessment of Gulf of Maine-Georges Bank Atlantic halibut, the stock was overfished (B$_{2004}$ was 5% of B$_{MSY}$ proxy) and it was unknown whether overfishing was occurring (Brodziak and Col 2005)."

The Replacement Yield Model (RYM) was first applied to US Atlantic halibut in 2008 at the GARM III assessments (NEFSC 2009). The RYM was suggested by Butterworth (refs) at the 2008 GARM III meeting. Col and Legault (2009) implemented the model. The biomass at time t is expressed as

$$B_t = B_{t-1} + R_{t-1} - C_{t-1} \text{ [1]}$$

Where Bt is the biomass at time t, C$_t$ is the catch at time t and R$_t$ is the replacement yield. Replacement yield is based on the logistic growth model and is defined as

$$R_t = rB_t \left(1 - \frac{B_t}{K}\right) \quad \text{[2]}$$

Where K is defined as the estimated population size in 1800. Application of the model required several important assumptions

1. Catches between 1800 and 1893 are unknown but are assumed to increase linearly from zero in 1800 to 798 mt in 1893.
2. The intrinsic rate of increase in population size is assumed to be constant over the entire time series and equal to a life history approximation derived from a YPR analysis of contemporary estimates of growth rates and a natural mortality rate of *M*=0.15. In the Col and Legault (2008) assessment, r was set to twice the value of *F$_{0.1}$*.
3. A penalty function on survey catchabilty with *q*=0.5 was imposed by Col and Legault for the NEFSC fall survey. Col and Legault (2009) also used a penalty function on population size.

4. The combination of an assumed trajectory of catch and a fixed value of *r* are sufficient to allow estimation of $K = B_{1800}$ in the model. The carrying capacity *K* is also assumed to be constant over the assessment period.

When Hennen (2015) updated the assessment in 2015 the model estimates suggested that the stock had completely rebuilt to K, i.e., the population size in 1800 and twice the SSB_{MSY} proxy value (See Fig. 81 in Hennen 2015). Moreover, the population estimates had been well above SSB_{MSY} since the start of the fall survey time series in 1963. Analyses of the log likelihood profile over the *K* parameter revealed extreme sensitivity to *K* (Fig. 16 in Hennen 2015. (Supplemental material).

The combination of implausible estimates of stock status, extreme statistical uncertainty, and dependency on model assumptions, led to the rejection of this assessment approach for Atlantic halibut. While the model incorporates important biological information about growth and natural mortality, model cannot be estimated without imposing constraints on *q* and fixing the intrinsic rate of increase. In Col and Legault, penalty functions were included to help fix $q \sim 0.5$ and to impose bounds on biomass. Sensitivity analyses by Col and Legault revealed that the assumed trajectory for catch between 1800 and 1893 had almost no effect on estimation.

The Review Panel in 2015 concluded that "the updated assessment was not acceptable as a scientific basis for management advice. The updated assessment produced an unstable and unrealistic solution. Estimates of current stock size were highly sensitive to initial conditions and slight changes in assumed parameter values."

Plan B Assessment Process
In the Northeast US rejection of the accepted stock assessment model creates uncertainty about stock status and poses the problem of finding an alternative basis for setting catch limits. These are affectionately known as Plan B assessments (NRCC 2011). The written and implied constraints on Plan B assessments were discussed in the introduction. In the following sections, alternative approaches for providing scientific catch advice are considered. Numerous methods have been proposed for the assessment of data-poor stocks. A number of excellent reviews of both methodology and applications may be found in Berkson et al. (2011), Newman et al. (2015), Carruthers et al. (2014), and especially Edwards (2015). The potential utility of such methods for Atlantic halibut is considered in the following sections.

Life History Methods
Life-history based methods rely on various properties of growth and longevity, and draw upon so-called life history invariants for obtaining suitable target fishing mortality rates. If a population is at equilibrium then length frequency information should be sufficient to obtain a

measure of total mortality. Assuming a rate of natural mortality then allows for derivation of a contemporary fishing mortality rate or target fishing mortality rate.

Length-based Methods

Length-based methods do not provide information on abundance or its trends (Edwards 2015). They are also typically slow to respond to changes fishing mortality because they rely on some degree of constancy in recruitment, fishery selectivity and natural mortality, and adequacy of biological sampling of landings and discards to define a meaningful rate of fishing mortality. Otherwise the derived rates can be biased. Gedamke and Hoenig (2006) developed approaches to address nonequilibrium populations.

Without a measure of scale, one can only interpret current fishing mortality rates with respect of target rates. If $F_t > F_{target}$, then the catch could be reduced by the degree of overage. However, such measures are not useful for setting catch limits unless they are viewed as part of feedback control system. Klaer et al (2012). noted that the feedback control rule had acceptable results for a high productivity demersal stock but that estimates of variability of length at age were essential for proper estimation. The overall sampling frequency for landed and discarded Atlantic halibut has increased in recent years but a full evaluation of the information content of such data is beyond the scope of this project. Introduction of new data requires consideration of potential sources of bias via a working group process.

Productivity-Susceptibility Analysis

The productivity susceptibility analysis method (Patrick et al. 2011) examines multiple attributes of life history, fisheries and habitats to derive a score for productivity P and susceptibility S. The overall vulnerability V of the species to overfishing is a function of P and S.   Results suggest that halibut are only moderately vulnerable to overfishing owing to relatively high productivity scores. Regardless of the underlying PSA score, the widespread absence of halibut in the Gulf of Maine and in deeper waters of Georges Bank suggests that recovery has been slow since the peak periods of fishing in the early 1900's. The PSA method was not considered further for this assessment.

Catch-Based Methods

Catch-based methods are rely primarily on adjustments to recent average catches (Berkson et al. 2011). The basis for the adjustment varies but typically includes a scalar adjustment to recent average catches based on an assumed stock status. For example, Restrepo et al. (1998) employed 3 different scalars, all less than one, depending on whether the stock was below or above the inferred estimate of B_{MSY}. While these methods are widely used in the US in data-poor stocks in the US, Carruthers et al. (2014) concluded that the utility of such methods as control strategies could not be evaluated reliably in a simulation context. Such measures are undoubtedly good starting points for managers until data collection procedures to support more robust measures can

be implemented. However, the degree of data poverty in such stocks is far greater than for Atlantic halibut which has multiple indices of relative abundance and recent biological information. Catch-only based methods were not considered further for this assessment.

Depletion-Corrected Average Catch (DCAC)
Depletion corrected average catch methods were first proposed by MacCall (2009) as a way of interpreting catch histories in terms of an underlying surplus production model. The DCAC model represents an important conceptual advance for fisheries as it applies logical constructs to obtain rough estimates of sustainable yield and more importantly, contemporary catch for data poor stocks. The methodology combines standard principles of surplus production models with various "rules of thumb" from various meta-analyses in fisheries stock assessments. The Depletion-based Stock Reduction Analysis (DB-SRA) is conceptually similar but relies on more detailed biological information. Both DCAC and DB-SRA rely on assumptions about current stock status relative to biological reference points. As this is the usual *output* of an assessment, the need to supply it as an *input* does cause some conceptual problems.
Edwards(2015) provided a succinct summary of the utility of DCAC and DB-SRA as follows:
> "Both DCAC and DB-SRA have been shown to be highly sensitive to the assumed current status of the stock δ, and can easily produce overestimates of the OFL if an optimistic distribution for δ is assumed. This is a major shortcoming, since if depletion of the stock is known already, then it is unlikely to be considered data-poor. Consequently it is difficult to conclude that these methods are an improvement on the scalar methods already in use. Indeed it appears from recent simulation studies that DACS methods produce comparable results (Carruthers et al., 2014)."

Edwards conclusions were tempered somewhat by noting that most data poor assessment models embed such considerations into their definition (e.g., see Restrepo et al 1998 discussion above.)

> "Furthermore, when considering their utility it is worthwhile noting the philosophical stance represented by these catch-only methods. They are centrally based on prior assumptions regarding the state of the fishery (specifically the depletion), which is a departure from previous conceptions of prior information that typically refer directly to parameter values within a particular model specification. Including this type of "soft" information could allow more "sporadic, qualitative or subjective" data to partake in the estimation process (Bentley, 2015), and the methods described by MacCall (2009), Dick & MacCall (2011) and Martell & Froese (2013), represent an important step in that direction."

Rewriting MacCall's (2009) DCAC notation by replacing Y with C leads to

$$C_{sustainable} = \frac{\sum_{t=1}^{n} C_t}{n + \frac{Delta}{0.2\,M}} \quad [3]$$

Where **Delta** is defined as

$$Delta = \frac{B_t - B_{t+n}}{B_{MSY}} \quad [4]$$

MacCall noted that yields are sustainable only if the current biomass is greater than B_{MSY}; otherwise the estimate of catch for the current time step may be approximated as

$$C_t = C_{sustainable} \frac{B_t}{B_{MSY}} \quad [5]$$

While conceptually simple and based on surplus production theory, the DCAC model requires an estimate of the biomass at MSY or equivalently the carrying capacity of the resource. Otherwise the proportion in the denominator of Eq. 3cannot be obtained.

The DCAC approach was first applied to US Atlantic halibut by Col and Legault (2009) as an exploratory exercise for two different cases. First they considered the entire time series of 208 years of catch (i.e., imputed+recorded) used in the replacement yield model (RYM). Using the model biomass estimates as a guide, the derived **Delta**=0.987 estimate of C_t was 35 mt. Using the entire time series of recorded catch (1893-2007), **Delta**=0.098, and the DCAC estimate of C_t is 10 mt. In their application of DCAC, the results of the RYM were used to estimate the key parameter **Delta**.

Theory of DCAC implies that sustainable and current catch can be estimated when the population is increasing as well as decreasing. When the population is declining over time, **Delta**>0 (and vice versa) but the magnitude of **Delta** depends not only the rate of change in abundance indices but also on the relative size of the current population. Col and Legault (2009) were able to use the results of the accepted assessment model to create their estimates. In the absence of such a model the estimation problem can be decomposed into two steps:
1. Estimate the relative rate of change in one or more abundance indices over some period of time. This is described in the section "Ratio Estimation".
2. Obtain an estimate of approximate scale consistent with the catch and relative biomass indices. This is described in the section "Envelope" method.

The methodology for achieving steps 1 and 2 are described in the following sections.

The model requirement that the relative status of the resource must be known in order to estimate relative catch implies that the status must first be inferred from knowledge apart from the model. Several authors have noted the logical difficulty of this approach (Edwards 2015, Carruthers et al 2014) but have also noted that it is it valuable in many fisheries where reasonable guesses of stock status might be made. The approach has been used widely in the US for stocks in which biomass is thought to be well above B_{MSY}. When biomass is well below B_{MSY} the scope for error in the **Delta** parameter is much less. Moreover, the method does not address the management

requirement for rebuilding at very low stock levels.   Both of these conditions are true for halibut, so it is important to define where the current stock is relative to some measure of B_{MSY} and to estimate the relative change that has occurred over the period of extraction.

Eq. 3 applies to instances where the stock has been increasing of the period in which catches have been taken. In this case Delta is less than zero. When *Delta*/(0.2 *M*) equals *n* the sustainable catch is undefined. When *n* <*Delta*/(0.2 *M*) the predicted sustainable catch is negative. Subject to the assumptions underlying the model, the presence of infeasible solutions provides a rough boundary on the current relative state of the stock. The implications of this discontinuity for bounding of abundance estimates will be discussed later.

Ratio Estimation and Randomization Method
A randomization test is developed herein to estimate the magnitude of change in a time series of length t=1,…T. We are interested in the general problem of determining whether the observations at the end of the time series are statistically larger than the observations at the beginning of the time series. For the DCAC issue we are not particularly interested in the trajectory of the change, so a model-based approach is not necessary. Moreover, observation errors tend to be high so that a simple regression model may be misleading.

The first task is to create a test statistic that can be used to compare the population state at the ends of the series. For this exercise I assumed the population state could be estimated as the ratio of the average of the last three observations to the first three observations.   Consider a time series with observations $x_1, x_2, …x_T$. If the times series is simply a random set of observations with no underlying trend, the test statistic should be near the center of the test statistics obtained by randomly shuffling the observations, and computing a new statistic. The collection of all statistics so generated is called the sampling distribution for the test statistic. The approximate significance level of the test statistic from the original time series can be compared to the sampling distribution. If it lies near the tails of the distribution on can assume that the observed value for the original series is improbable due to chance alone. These concepts are formalized in the following equations.

Let $I_{j,t}$ represent the j-th index at time *t* where *j=1,…J* and *t=1,…T*. We compute the endpoint estimates of abundance using an average of multiple years (*n* and *m*) to help reduce the effects of random variation in catchability between years Let *m*= number of years for most recent years and *n* for earlier period. Define test statistic or critical ratio for index I_{jt} as $R_{crit,j}$ as

$$R_{crit,j} = \frac{\sum_{t=T-m+1}^{T} \frac{I_{j.t}}{m}}{\sum_{t=1}^{n} I_{j,t} \frac{I_{j.t}}{n}} \quad [6]$$

If the observations for the various I_j. are not commensurate, then, without loss of generality, the indices can be standardized with respect to their individual means. The composite test statistic for multiple time series can be define as

$$R_{crit,.} = \frac{\sum_{j=1}^{J} \sum_{t=T-m+1}^{T} \frac{s(I_{j,t})}{m}}{\sum_{j=1}^{J} \sum_{t=1}^{n} \frac{s(I_{j,t})}{n}} \qquad [7]$$

Where $s(.)$ refers to a standardization function in which the index is expressed as a ratio to its mean.

$$s(I_{j,t}) = \frac{I_{j,t}}{\sum_{t=1}^{T} \frac{I_{j,t}}{T}} \qquad [8]$$

The sampling distribution of the randomization statistic for **Rcrit** is obtained by shuffling the observed sequence of indices and computing a random realization of the indices. Let $R(.)$ represent the randomization function which shuffles the original indices $I_{j,t}$ with respect to time. Let k represent the index for the kth realization of the random **Rcrit**.

$$R_{crit,k} = \frac{\sum_{j=1}^{J} \sum_{t=T-m+1}^{T} R_k\left(\frac{s(I_{j,t})}{m}\right)}{\sum_{j=1}^{J} \sum_{t=1}^{n} R_k\left(\frac{s(I_{j,t})}{n}\right)} \qquad [9]$$

The sampling distribution of **Rcrit** is obtained by repeatedly applying Eq. 9 over an arbitrarily large number of iterations, $k=1…Nrand$. The approximate significance value of the observed **Rcrit** can be obtained by comparing it to the sampling distribution of realized observations $\{\mathbf{R_{crit,k}}\}$

The probability of obtaining a value greater than $\mathbf{R_{crit,obs}}$ is simply

$$P(R_{crit,k} > Rcrit, obs) = \frac{\sum_{k}^{Nrand} g(R_{crit,k} \geq R_{crit,obs})}{N_{rand}} \qquad [10]$$

where g(.) is an indicator function equal to 1 when the logical argument is true and 0 otherwise. The probability of observing a critical value less the observed value may be obtained by simply reversing the order of the operator in the indicator function $g(.)$

The sampling distribution of the **Rcrit** in Eq. 9 can be enumerated as the product of combinatorials. Total realizations = J* comb(T,n)* comb(T-n,T-m-n) * comb(m,m). For J=6, T=10, n=m=3, the number of potential combinations is 25,200. I approximated the sampling distribution with 2000 iterations.

SIMULATION TESTS for Randomization method
For the purposes of this assessment, the performance of the **Rcrit** statistic is defined as the ability to detect a true rate of change. This is affected by the magnitude a function of the true underlying rate of increase, the underlying observation error for each index and the number of indices. Intuitively one would expect the performance of **Rcrit** to improve with larger true rates

of increase, as the observation error declines and as the number of indices increases. These hypotheses were tested in a series of simulations described below.

Let the true rate of annual increase be defined as λ.
$$I_{true,t+1} = \exp(\lambda)\, I_{true,\,t} \quad [11]$$
The realized observations are assumed to be lognormally distributed random variables with mean defined by the true index value and the SD specifiec by the coefficient of variation CV.
Let

$$I_{realized,t} = LnNormal\big(I_{true,t}, SD\big) \quad [12]$$

Where $SD = \sqrt{\ln(CV^2 + 1)}$

Simulations were conducted for 3 levels of λ,{0.1, 0.05,0.025} 12 levels of CV { 0.1, 0.15, 0.2,…0.65, 0.70} and four different levels of J={1,2,3,5}.   Random times series of 10 observations were computed for 1,000 realizations. For each realization, a $R_{crit,realized}$ was computed. A randomization test with 2,000 iterations was then used to compute the significance level for each random realization   Two million iterations were computed for each of the 3*12*4 combinations of λ, CV and J.

Randomization Simulation Results
Simulation tests suggest relatively little bias in the ratio estimator over a broad range of simulated values except when the true magnitude of increase is small (eg 2.5% per year) and the underlying variability of the observations is low (Table 2). Even then, the bias will decline as the number of indices increases. The probability of successfully detecting a change in population size is given in Table 3.   As expected increases in the true magnitude of change, reductions in the variability of the observations and increases in the number of available indices all act to increase the probability of detecting the true change.

Overall results of the simulation studies are summarized in Tables 4, 5,6 and 7. Each table corresponds to different number of variables used for trend. Within each table lambda ranges from 0.1 to 0.25 and CV ranges from 0.1 to 0.7. The tabulated results are the fraction of test statistics that are significant at the P=0.005, 0.01, 0.05, …, 0.25 probability levels. For example, a value of 0.89 would mean that 89% of the test statistics were less than or equal to the probability level of the columns.   In other words, the entries provide a metric of the ability of the estimator to correctly identify the true ratio.   Color shading is scaled consistently across tables with green shading indicating good performance and red shading indicating poorer performance. As one would expect model performance generally increases with the magnitude of increase (eg. It's easier to find the correct value when the true Rcrit is bigger), as CV gets smaller, and as the number of variables used for detection increases.

.

Application of Rcrit Method to US and Canadian Indices
The Rcrit randomization method was applied to six candidate indices for the US stock and three candidate variables for the DFO 3NOPs4VWX5Zc stock (TABLE 1). Candidate indices for the US stock included the NEFSC fall trawl survey biomass, the d/k ratios for halibut taken in observed trips on gill net and trawl vessels, a modeled index of commercial catch per unit effort, and the fall and spring weight per tow estimates from the ME-NH inshore trawl survey. The inclusion of the d/k ratios for gill nets and trawls should not be interpreted as introduction of a new time series in the model since these are components of the SBRM discard estimate. For the purpose of establishing trend, the selection of the NEFSC trawl survey and d/k ratios should be considered consistent with the time series used in previous assessments. Other indices from Maine are useful for illustrating overall coherence of available information.

No attempt was made to define the "best" set of variables. Instead, the Rcrit method was applied to all possible combinations of indices. For 6 variables, this implies 63 different models based on the sum of combinatorials denoted as (6,6) {ie. 6 items, taken 6 at a time) +(6,5)+ (6,4)+ (6,3)+ (6,2)+ (6,1)=63 possible models. Tables 8, 9, and 10 summarize the results of the complete set of models for 2002-2016, 2005-2016, and 2002-2013, respectively. Nearly all of models configurations were statistically significant. For 2002-2016 95% of the models has significance values less than 5%, for the 2005-2016 period 78% were significant, and for 2002-2013, 95% were significant. The average Rcrit over all models and year ranges went from a low of 2.44 for the 2005-2016 to 3.52 for the 2002-2016 period.

For the Canadian stock three abundance indices are used in their analytical model. The average increase in Rcrit over 6 possible models was 2.92 (P<0.001) (Table 11). The Rcrit for the modeled biomass was 2.763. Overall the comparisons suggest that the US stock has increased at a rate comparable to that observed in Canada. Of course, the scale of these changes is considerably different. Landings in Canada in the last 3 years have averaged ~3400 mt, whereas in the US stock landings have been about 100 mt.

The implied annual rates of increase in relative abundance, given the Rcrit estimates below are on the order of 9 to 15% per year. The similarity in rates of increase between US and Canada stock areas potentially suggests favorable conditions in both areas during the past decade.

	Ratio Definition	Changes in catches			Change in indices		
		Statistics	*Value*		*Statistic*	*Value*	*Model*
US	'02-04:'14-16	Rcrit(Catch)	3.227		Rcrit(Indices)	3.23	(all six indices)
						4.98	(DK_g, DK_t, Survey)
						3.52	average over 63 models
	'05-07:'14-16	Rcrit(Catch)	2.657		Rcrit(Indices)	2.44	average over 63 models
						2.2	(all six indices)
						4.11	DK_g,DK_t, Survey
	'02-04:'11-13	Rcrit(Catch)	2.617		Rcrit(indices)	2.893	(all six indices)
						5.033	(DK_g, DK_t, Survey)
						3.144	average over 63 models
Canada	2002-04: 2014-2016	Rcrit(Catch)	2.259		Rcrit(Indices)	2.703	(two survey , one CPUE
						2.923	average over 6 models
						2.763	Analytical model results

While Rcrit provides a way of quantifying the rate of change in population size, it cannot distinguish the change in scale. For example a population that increase 3 fold during some period could increase from 2% to 6% of the virgin stock size for from 20 to 60%. Application of DCAC requires one to estimate the ***Delta*** parameter in terms of change relative to virgin stock size. To establish scale, the Envelope method was applied, as shown in the next section.

Envelope Method

The envelope method (Miller and Rago 2010) is an approach to establish a range of feasible biomass estimates conditional on an assumed range of feasible catchability estimates q and historic ranges of fishing mortality rates F.   The method combines concepts of swept area biomass and the standard Baranov catch equation. Biomass estimates based on swept area estimation are dependent on an assumed range of catchability estimates q'. Biomass estimates based on observed catches rely on an estimate of fishing mortality rate. For any given time series of catch, one can assume it is the realization of a small population being fished consistently at a high rate, or a large population being fished at a low rate.   Using the standard definition for swept area biomass

$$B_t = \frac{I_t}{q'} \frac{A}{a} \qquad [13]$$

The ratio A/a is the total area A covered by the survey and a is the average area swept per tow. The biomass estimate consistent with observed catch can be obtained from the Baranov catch equation

$$B_{0,t} = \frac{C_t}{\frac{F}{F+M}(1 - e^{-F-M})} \qquad [14]$$

$$B_{f,t} = B_{0,t}e^{-(F+M)f}$$

The second equation in Eq 14. adjusts the biomass estimate to be on the same time scale as the survey estimates in Eq. ss.   If we use the general notation that $B(I,q)$ is the biomass estimate based on the observed I and assumed q, and $B'(C,F,M)$ is the biomass estimate based on the observed C and assumed F and M, then one can generate the following set of biomass estimates:

$$\widehat{B_{1,t}} = B(I_t, q_{Low}) \qquad [15]$$

$$\widehat{B_{2,t}} = B(I_t, q_{High})$$

$$\widehat{B_{3,t}} = B'(C_t, F_{Low}, M)$$

$$\widehat{B_{4,t}} = B'(C_t, F_{High}, M)$$

For many species, prior information on a suitable range of q may obtainable from gear comparison studies. Calibration studies for NEFSC failed to catch sufficient numbers of halibut in either net type to allow estimation of a conversion coefficient (Miller et al. 2008). A plausible range of fishing mortality estimates may be obtained by analogy with other halibut or flatfish fisheries.

The key concept in the envelope method is that the bounds represent extremes in the feasible range of parameter values. The upper and lower bounds of biomass estimates can then be defined as the set of estimates that jointly satisfy both constraints. These values are defined as

$$\hat{B}_{upper,t} = min(\hat{B}_{1,t}, \hat{B}_{3,t}) \qquad [16]$$

$$\hat{B}_{lower,t} = max(\hat{B}_{2,t}, \hat{B}_{4,t})$$

Values of biomass that exceed $B_{upper,t}$ imply catchabilities smaller than q_{low} or fishing mortalities less than F_{low}. Conversely, values of biomass less than $B_{lower,t}$ imply catchabilities greater than q_{high} or fishing mortalities greater than F_{high}. The bounds defined by Eq. 16 describe a set of feasible estimates that are consistent with the assumed ranges of *both* q and F. In theory, a more mechanistic model of stock dynamics should also be within this feasible range. Additional layers of constraints might be applied to the model to further reduce the range of uncertainty. For example, one could reasonably hypothesize that the biomass in US waters in recent years should be less than or equal to the Canadian stock biomass in 3NOPs4VWX5Zc. If a particular constraint is binding, then it can be used to further refine the feasible ranges of q and F for those

years.   A mid-range estimate of central tendency (sensu Tukey 1977) for B can be obtained as the average of $B_{upper,t}$ and $B_{lower,t}$.

Envelope Results
The model was applied to the catch estimates from 1963 to 2016 and NEFSC fall trawl survey estimates of swept area biomass. The assumed range of F is $\{0.02, 0.40\}$ and q is $\{0.02, 1.0\}$. The envelope model can only be applied to catches before 1963. FIG. 12. The Envelope model was also applied to the Kalman smoothed biomass estimates of fall survey biomass (FIG. 13). As might be expected the biomass estimates from the RYM do lie within the boundaries of the Envelope and correspond well with the mid-range estimator of average abundance (FIG.13). Quantitative results of the maximum biomass estimate and 2016 biomasses are summarized in Table12. Depending on the range of years used for estimation of biomass the mid range of the 2016 estimate ranges from 2.4% to 98.2% of the maximum observed value. The derived range is not terribly useful but it does highlight that the perception of the resource varies considerably with the inclusion of more long term data. The wide range of uncertainty in the biomass results is consistent with our limited understanding of the dynamics of halibut in US waters.

Factoring the Rcrit and Envelope Results into DCAC
The results of Rcrit and the Envelope methods can now be factored into the computation of DCAC to obtain estimates of sustainable yield and predictions for 2018 catches. Table 13 combines the observed ratio increases from Rcrit and the Envelope estimates of fraction rebuilt in 2016 to derive a set of possible Delta parameters. In part B of Table 13 the derived Delta are used to estimate sustainable catches for each combination. As noted in the methods above, the DCAC model can become unstable as the denominator in Eq. 3 approaches zero. The model produces infeasible results when the denominator becomes negative. Table13 shows the specific behavior of the model for this application. FIG. 14 general behavior of the estimator for varying values of *Delta* and *Rcrit*.

The overall results of the DCAC approach are not reassuring even when the estimates of Delta are refined by explicit consideration of the recent trends in population indices. The fundamental problem appears to be uncertainty in the absolute biomass estimate. Even if a credible statistical catch at age model can be developed, the uncertainty of any biomass estimates is likely to be very large for the foreseeable future.

Proposed Assessment Approach
The proposed assessment approach is less ambitious in terms of estimating long-term parameters (*r, K*) and instead focuses on short-term changes in stock size and their implications for modifying catch.   As a simplification, it is assumed that the current stock size is well below the historic carrying capacity K such that the expression *(1-B_t/K)* in Eq. 2 is negligible.   This seems reasonable in the context of inferring an initial stock size that existed over 200 years ago and is

estimable only by assuming that the productivity of the stock has been fixed and constant at $2*F_{0.1}$ for a similar period.   The $F_{0.1}$ estimate is based on life history attributes from recent decades and the natural mortality is fixed at M=0.15.

Relaxation of these assumptions leads to a simple linear equation for biomass as a function of a time-varying rate of increase r_t and a time-varying harvest rate h_t.

$$B_{t+1} = B_t + r_t B_t - h_t B_t \qquad [17]$$

Catch is defined as the product of harvest rate and stock size

$$C_t = h_t B_t \qquad [18]$$

Which leads to

$$B_{t+1} = B_t + r_t B_t - C_t \qquad [19]$$

A key assumption in nearly all stock assessment models is that stock size is proportional to one or more indices of abundance I_t as

$$I_t = q B_t \qquad [20]$$

Stock assessment models can be fit to observed data by substituting Eq. 18 into Eq. 17 and by creating a likelihood function for one or more relative abundance indices.   Derivation of q for each index and specifying an appropriate function for r_t can be problematic if r_t is changing and if observation error of I_t is high.

To avoid these problems the equation for biomass dynamics is transformed into a recursive expression for catch over time. The model can be derived as follows. The catch at time t+1 is written as

$$C_{t+1} = h_{t+1} B_{t+1} \quad [21]$$

Dividing Eq. 21 by Eq. 18 gives

$$\frac{C_{t+1}}{C_t} = \frac{h_{t+1} B_{t+1}}{h_t B_t} \quad [22]$$

Rearranging terms a bit provides a prediction of future catch as a function of current catch:

$$C_{t+1} = \frac{h_{t+1}}{h_t} \frac{B_{t+1}}{B_t} C_t \quad [23]$$

Without loss of generality, one can substitute Eq. 20 $I_t=qB_t$ in the Eq. 23 to obtain

$$C_{t+1} = \frac{h_{t+1}}{h_t} \frac{I_{t+1}}{I_t} C_t \qquad [24]$$

The problem with Eq. 24 is that it relies on having an estimate of I_{t+1} in order to estimate C_{t+1}. Furthermore, it also requires h_{t+1} and h_t which are also unknown. By definition the index I_{t+1} is a consequence of the removals at time C_t so it would not generally be available until most or all of the fishery that harvest catch in period t+1 would be complete.

However, from Eq[17] the ratio B_{t+1}/B_t is $1+r_t-h_t$ which is equivalent to I_{t+1}/I_t .

$$\frac{B_{t+1}}{B_t} = 1 + r_t - h_t \qquad [25]$$

Or by substituting Eq. 20 into Eq. 25.

$$\frac{B_{t+1}}{B_t} = \frac{qI_{t+1}}{qI_t} = \frac{I_{t+1}}{I_t} = 1 + r_t - h_t \qquad [26]$$

An approximate estimate of the expression ($1+r_t-h_t$) can be obtained by regression the $log(I_t)$ vs time. This is easily shown by recursively applying Eq. 6 over p time steps to obtain

$$B_{t+1} = (1 + r - h)B_t$$
$$\dots$$
$$B_{t+p} = (1 + r - h)^p B_t \qquad [27]$$

Taking log of both sides results in

$$\ln(B_{t+p}) = p\big(ln(1 + r - h) \ + \ln(B_t) \qquad [28]$$

Therefore, the slope of rate of change in biomass over time is $ln(1+r-h)$. The intercept is simply the log of the initial condition B_t. For an index I_t that is proportional to Biomass B_t as defined in Eq. 28, the slope is independent of the scaling factor q.

$$\ln(I_{t+p}) = p\big(ln(1 + r - h) \ + \ln(I_t) \qquad [29]$$

Using a log-linear regression model for the abundance indices one can approximate as the slope of $ln(I_t)$vs t or the average slope of the composite indices. Recall that we hypothesized that the r and h were functions of time. The regression in Eq. 15 assumes that r and h are constant over the interval $t=t_1$ to t_2. To approximate the change in slope over time one can update the regression equation 15 by computing the slope at each time t for τ time steps (ie. A τ-point regression). For simplicity of notation, let

$$slope_t = \ln(1 + r_t + h_t) \qquad [30]$$

Substituting Eq. 26 and 30 into Eq. 24 gives

$$C_{t+1} \cong \frac{h_{t+1}}{h_t}(1 + \widehat{r_t} - h_t)C_t$$

$$C_{t+1} \cong \frac{h_{t+1}}{h_t}e^{slope_t}C_t \qquad [31]$$

Technically Eq. 31 poses some additional problems since neither h_t nor h_{t+1} are known. The harvest rate h_{t+1} can be written as a function of the biomass and catch at t+1. So one is left with the assumption that the slope at time $t+1$ is approximately equal to the slope at time t.
In most real-world scenarios an index of the biomass at time $t+1$ would not be available at the time when C_{t+1} is being set. In practical terms it means that the slope estimated over the set of points $\{t_i, t_{i+1}, \ldots t_{i+n}\}$ approximates the slope estimated from the set $\{t_{i+1}, t_{i+2}, \ldots t_{i+n+1}\}$. In other words, the slope estimate at time t_{i+n} is used to approximate the estimated slope at t_{i+n+1}.
Under these constraints, the updating function for catch can be written as

$$C_{t+1} \cong e^{slope_t}C_t \qquad [32]$$

Eq. 32 implies that the rate of change in catch should be equal to the rate of change in relative abundance. Note that scale of indices does not affect estimate of slope vs time.
The model can readily be extended to multiple indices by taking the simple average of the rates of change in index values when the **slope**$_j$ is defined by the log-linear regression **$ln(I_{j,t})$** vs **t**.
Hence slopes from multiple indices can be combined without consideration of their underlying scale. One can estimate a common slope via a general linear model in which the various indices are considered factors. Alternatively, the common slope can be estimated as the average of the j=1,…J slope estimates. Without loss of generality Eq. 32 can be written as

$$C_{t+1} \cong e^{average\ \widehat{slope}(I_{j,t})} C_t \qquad [33]$$

The slope parameters of the composite regression incorporate a number of underlying proceses including growth, recruitment, natural mortality and harvesting. Since any or all of these processes can vary over time, it is important that the forecasting equation be responsive to processes as they begin to occur.
This type of control rule has been suggested by Geromont and Butterworth (2015a, 2015b) and several others (Pomaerde et al 2010, Apostaloki and Hillary 2009). Note that the control rule implies that next year's catch can be adjusted based on information about the stock trend in the current year. This basic concept can be extended by applying concepts from control theory as shown in the following section.

APPLICATION OF CONTROL THEORY CONCEPTS TO HARVEST *Control Rules*
Control theory, in the context of this assessment, refers to a general set of principles used to adjust a physical system toward a desired state. A simple example is the use of a thermostat to control temperature in a room. More complicated examples include control of complex chemical

production processes or control of surfaces on airplane wings. In general control theory uses feedback from a monitoring device to adjust some input factor to achieve a desired output (e.g., the thermostat send a signal to the furnace or air conditioner depending on what the temperature is relative to the desired value).

One of the major concepts in control theory is that controls can destabilizing if signals about system state are corrupted by noise or delayed and if the change in input level is too large. Conceptually this could occur if the thermostat sensor drifts or is delayed by some software glitch. Consider the consequences if the daytime output of the furnace is governed by temperatures monitored the previous evening. Destabilization can also occur if the input control is too large relative to the observed deviation of system state. A simple example would be a large furnace in a well-insulated small house. Since most furnaces are simply on/off devices, the likelihood of putting out too much heat is possible, raising the temperature to too high a point which then persists.

So what does this have to do with fisheries management? Consider catch to be the input control and relative abundance to be the output signal. If a target output value is known, then catch can be adjusted to achieve a desired value by monitoring the system state. The system state is simply the slope of the relative abundance index or indices.   If we want to continue to allow the population togrow then catch can be set at some level less than that indicated by the rate of change.   In control theory terms this scalar is referred to as the proportional gain or *Kp.*   A stock in a rebuilding program would be one in which the gain might be less than one, thereby allowing a population to continue to grow.

A stylized schematic of the proposed model is given in FIG 15.   The population dynamics are treated as a black box that outputs one or more abundance indices as a function of changes in the input Catch. If some function of the output variable is less than the previous value, then it is assumed that the previous input signal was too high and the input value of catch for the next time step would be reduced. Conversely, if the index output function increased, there might be some scope for increased catch in the next time step. Of course, devil is always in the details and it is important to consider the responsiveness to the output signal and the magnitude of the inter-annual adjustments. In the following analyses some of these details are explored further for application to Atlantic halibut.

The history of fisheries science is replete with examples of where an underlying process that is assumed constant in a model changes over time. The consequences for management are often overfishing and economic loss. Decreases in growth rate, increases in unobserved mortality, or reduced recruitment will tend to increase the variability in model fit but more importantly, lead to bias in predictions. Analytical models accommodate such changes in varying ways, but many causes can give rise to the same symptoms, such as retrospective patterns. Unfortunately, models

need the most adaptability at the end of the time series, where any emerging trends are difficult to distinguish from noise. As an example, decreases in average size of fish may be due to increased recruitment or changes in fishing areas where smaller fish are more abundant. Model parameters for selectivity are unlikely to reliably estimate this change as it could imply either an increase in the historical recruitment to match the observed catch at age, or it could adjust the fishery selectivity at age.

In the simple model proposed here it is not possible to dissect such changes from the measure of slope. The slope is an aggregate measure of multiple factors. However, it is possible to estimate the rate of change in slope as a measure of acceleration or deceleration of trends. In this context it might be called the second derivative of population change. If we let $\beta(t,n)$=average slope estimated at with terminal year t and based the last n points, then we can compute the second derivative of population change as

$$\beta(t,n) = slope(x_{t-n+1,}x_{t-n}, \ldots .x_{t-1}, x_t)$$
$$\Delta\beta(t,n) = \beta(t,n) - \beta(t-1,n) \qquad [34]$$

The relationship between the slope estimate and its rate of change is important for forecasting future catch. If $\beta(t,n)$ and $\Delta\beta(t,n)$ are positive, then the population would be increasing at an increasing rate. If $\beta(t,n)$ is positive and $\Delta\beta(t,n)$ is negative, then the population is increasing at a decreasing rate. In the former case, one would be more optimistic about continued increase in stock size. The latter case would suggest that population growth may be slowing. There are no hard and fast rules about how to weight the relative importance of these two situations but in the control theory literature, this is called a derivative control, and the weight assigned to this factor is called the derivative gain factor or **Kd**.

With these concepts in mind, the updating function for catch can be improved by considering the proportional change in stock size and the derivative of the rate of change as follows:

$$C_{t+1} = e^{(K_p\beta(t,n)+K_d\Delta\beta(t,n))}C_t \qquad [35]$$

The exponential term in Eq. 35 expresses the rate of change in catch as the weighted sum of the proportional change in abundance (i.e., the first derivative of population size with respect to time) and the derivative of the rate of population change (i.e., the second derivative). This leads to the somewhat hokey name of First and Second Derivative harvest control or FSD control for short. The utility of the model is evaluated over a broad range of simulation scenarios and by application of the model to two managed halibut stocks, the 3NOPs4VWX5Zc Atlantic halibut managed by DFO and Pacific halibut managed by the International Pacific Halibut Commission (IPHC). In both applications, the predicted catch from Eq. 35 is compared to the TAC derived from modern analytical models.

Simulation Experiments for FSD Control Rule

While the control rule (Eq. 35) has some intuitive appeal, its utility is ultimately governed by its ability to control a theoretical population subject to a variety of conditions that are largely unknown or unpredictable in a real world. Relevant factors for a simulation study include

- Observation error for the relative abundance indices *CV={0.005,0.2}*
- Number of abundance indices available *Nvar={2,6}*
- Number of years to consider for estimating average slope. *Ntrend={3,5}*
- Effects of alternative measures of Kp and Kd
- The underlying rate of population increase (r(t)) during the period before and after the control rule is applied.
- The pattern of harvesting (h(t)) prior to the application of the control rule.

In this simulation experiment, the population was harvested without application of the control rule for the first 10 years. During the next 10 years the control was applied. Observation error of the indices was examined by letting the CV range between 0.005 and 0.2. A CV of 0.005 is highly improbable but allows for evaluation of performance in the near absence of observation error. The number of years to use for estimating the slope was varied between 3 and 5 data points. Increasing the number of years decreases the responsiveness of the slope estimator to rates of change. Decreased responsiveness trades off with the increased likelihood of noise-driven estimate of the slope when only 3 points are used. Of course, one expects the shorter interval slope estimators to less reliable as the observation error increases.

Process error in the simulation context was addressed by hypothesizing temporal trends in stock productivity r_t, during the control period. Prior to implementation of the control, all scenarios assumed r=0.2 for the first 10 years. After the control was implemented, *r(t)* in year 11 was assumed to be

- Constant for the next 10 years at 0.2
- Steadily increasing over the next 10 years to 0.3
- Steading decreasing over the next 10 years to 0.1
- Increased as a step function in year 11 to 0.3
- Decreased as a step function in year 11 to 0.2
- Increased steadily for 5 years followed by a steady decrease
- Decreased steadily for 5 years followed by a steady increase.

These scenarios are depicted in FIG. 16.

The history of harvesting prior to implementation of the control rule is important because it defines the set of indices that will be used to develop the average slope estimate. Deterministic simulations suggest that the proximity of the true fishing mortality rate to the true productivity of the stock is critical for the application of the control. Neither of these quantities are estimable, so the control rule should be robust to this uncertainty.

The harvest rate scenarios all assume that the true population is growing during the first 10 years. This mimics the observed pattern for US halibut (See **Application of Rcrit Method to US and Canadian Indices**).   Three scenarios assumed constant harvest rates with h_t=0.15, 0.19 and 0.10, for all *t*.   Other scenarios assume continuously increasing h_t, continuously decreasing h_t, a harvest rate that increases then decreases, and one that decreases then increases (FIG. 17). 49 possible combinations of r_t and h_t were evaluated by pairing each r_t and h_t scenario.

The effect of the **Kp** and **Kd** gain parameters were developed by evaluating performance for each combination of **Kp**={0, 0.25,0.5, 1.0} and **Kd**={0,1,5,10}. Higher values of K_d were used to evaluate the consequences of testing quick responsiveness when the $\Delta\beta$ parameter was changing. There were a total of 16 combinations of **Kp** and **Kd** evaluated for each combination of ht, rt, Ntrend, and Nvar.

Each 20 year simulation was repeated 50 times resulting in 2*2*2*7*7*4*4*50=313,600 applications of the control rule. Summary statistics from each simulation during the control period included the average:
- Number of overfishing events induced by the control rule (i.e., when the predicted catch resulted in overfishing (h_t>r_t)
- Number of "extinctions" when the population is driven to arbitrarily low values
- Catch
- CV of catch
- Net rate of population increase

Simulation results are summarized in Table 14 for two levels of CV ={0.005,0.2} and 16 combinations of K_p and K_d.   Averages are made over the 49 combinations of h_t and r_t and for tow values of Ntrend={3,5}.   Results suggest that low rates of K_p reduce the frequency of overfishing events. As the weighting on the change in slope increases, the frequency of overfishing events tends to increase because of induced oscillations.   The relative precision of the observations appears to have little effect on the probability of overfishing. As expected, increases in K_p result in increases in average catch, but the increases in average catch come at the expense of greater variability in catch.

The frequency of "extinctions" is more complicated to explain as it appears to be driven more by the underlying initial conditions based on r_t and h_t, rather than the control parameters K_p and K_d. When observation error is relatively low, there controls with K_d~5 appear to work well in terms of reducing the frequency of "extinctions". Further simulation work may be necessary to examine the dynamics related to overshooting catches.

The potential interactive effects of the r_t and h_t, can be examined by considering the extinction frequency summed over all values of K_p, K_d, *Ntrend,* and *CV.* The scenarios in the following text table are depicted in Fig. 16 and 17.

		R scenario							
		r=0.2	r_up	r_down	r_step_up	r_step_dn	r_up_dn	r_dn_up	
Descrip-tion	*Harvest Scenario*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	average
h=0.15	1	0.0159	0.0241	0.0178	0.0256	0.0191	0.0225	0.0181	0.020
h=0.19	2	0.0394	0.0375	0.0363	0.0525	0.3281	0.0456	0.0419	0.083
h=0.10	3	0.0053	0.0078	0.0091	0.0078	0.0078	0.0094	0.0053	0.008
h_up	4	0.3722	0.0816	0.5066	0.0691	0.5500	0.1044	0.5113	0.314
h_down	5	0.0016	0.0016	0.0013	0.0028	0.0013	0.0025	0.0019	0.002
h_dn_up	6	0.0469	0.0363	0.3153	0.0316	0.5028	0.0534	0.4019	0.198
h_up_dn	7	0.0119	0.0188	0.0156	0.0231	0.0184	0.0203	0.0141	0.017
	average	0.070	0.030	0.129	0.030	0.204	0.037	0.142	0.092

Overshoots are more frequent when the harvest rates are increasing during the pre control period (Harvest #4) and when harvest rises just prior to the control period (Harvest #6). Declining productivity at the start of the control period is problematic as see in R scenarios #3,5 and 7. Most of the unstable trajectories occur at the intersections of H scenarios 4 and 6 with R scenarios 3, 5 and 7.

One of the primary factors influencing the responsiveness of the system is the number of data points in the regression equation. Slope estimates will always constitute lagged information. The high amount of observation error in available indices leads to concerns that short term changes, deduced from say the ratio of abundance indices in adjacent years is likely to be unreliable. Unfortunately, the most important information in the stock assessment is the changes in the current time period. Hence it is necessary to consider a larger number of time periods when estimating trend. There may be some improved methods for estimating trend that retain sufficient flexibility for signal detection. A Kalman filter that incorporates information about overall observation error may be useful in this regard for future research.

Application of the FSD Model to US Stock Area
Input data for the US stock area are given in Table1. The model was based on a Kp=0.9 and Kd=0.9 for three indices dk_trawl, dk_gillnet, and NEFSC fall bottom trawl biomass. (See Appendix 1 for the input data and relative errors in each index). A five-point slope regression was used compute the slope for all indices FIG 18. The composite average slope has been positive since 2009 but the slope estimates have been decreasing since 2012. The catch multiplier for each year (based on the K_p and K_d factors of 0.75 and 0.50, respectively) suggests that the rate of increase is declining. FIG 19.

Comparison of observed and predicted catches show reasonable coherence (FIG 20) and surprisingly, the forecasted estimates of catch for 2011 to 2014 given in Blaylock and Legault (2012) line up well with observed values.

Model forecasts for 2018 were examined over a range of K_p {0, 0.3,,1.0}and K_d gain factors from {0, 0.25,…, 4.0} in Table 15. The model performance was estimated by computing a total sum of squares differences between the observed and projected values. Over the range of gain factors tested, the 2018 catches range 110.3 to 141.7 mt. Using the region where the SSQ is within 10% of the minimum value, the highest possible catch is 118.1 mt.

Uncertainty Estimation for Catch Forecasts
The uncertainty in the projected catch forecast includes process error related to the potential change in relative productivity of the stock (i.e., r(t)), the inherent lag in information owing to the number of data points in the estimator of the slope, and the gain factors applied to the first and second derivatives. Observation error in the abundance indices themselves also contributes to the uncertainty of the estimate. While simulation experiments address some of the uncertainty for idealized abundance indices (in particular, indices with homogeneous CVs for all indices), simulations do not capture a real-world example with heterogeneous variability over time and among indices. To address the realized uncertainty of the FSD model applied to the US stock area, a parametric bootstrap method was applied to each index time series. Indices were assumed to be lognormally distributed with means equal to the observed value and standard deviations equal to the log of square root of the coefficient of variation squared plus one.

$$Irand_{k,j,t} \sim LogNormal(Iobs_{j,t}, \sqrt{\ln(CV_{j,t}^2 + 1)}, \qquad [36]$$

Note that the **k-**th realization of the **j-**th index at time t is a function of time varying means and variances. The sampling distribution of the projected catches were based on 5000 bootstrap realizations.

Projected catches by year are given in Fig. 21 and the sampling distribution of projected catch in 2018 is shown in Fig. 22. The sampling percentile statistics for this distribution are

1%	5%	10%	25%	50%	75%	90%	95%	99%
98.24	104.98	108.61	114.88	122.65	130.69	138.34	143.16	152.26

The bootstrap mean of projected catch is 123.10 mt with a CV equal to 0.095.

Application of the FSD control rule to DFO Canada and IPHC Pacific halibut

The FSD model was applied to both the DFO Atlantic halibut and the IPHC Pacific halibut stocks. For the DFO application, the K_p and K_d parameters were set to the same values used or

the US application. The forecasted catches tend to be higher than the model based TACs (FIG. 23). While the apparent coherence of the methods is somewhat reassuring, the differences between the two methods are likely to be not inconsequential in any real world application. However further analyses of the differences in model fit may be useful for improving the decision rules for the US stock area.

The application of FSD to the Pacific halibut stock is summarized in FIG. 24. Data used in this application were obtained from Stewart (2017). Given the tight controls on the TACs for Pacific halibut, it is assumed that the realized catch is close to the TAC. Although the FSD model has some large outliers over the entire time series, the model forecasts are relatively close to the realized catches from 2003 onward. Maximum deviations between the observed and predicted catches are less than 8 mt and generally under 4 mt.

Parametric bootstrap results are summarized in Figure 25 and 26. In general the distribution of predicted catches based on the model covered the range of observed total removals. lTabular summaries of the relative error in the survey and commercial WPUE indices from the IPHC were not available. I assumed the CVs for surveys were 0.2 for WPUE were 0.25 based on overall patterns summarized in Stewart (2017).

These results are likely attributable to consistent downward trend in catches and reliable signals from the IPHC surveys. A comparison of these relative trends with modeled biomass is given in Fig. 27

Effects of discard mortality on catch projections
The survival of released halibut is a function of many factors including the type of gear employed and handling of catch on deck. Estimated and assumed discard mortality rates vary widely among fisheries. In Pacific halibut fisheries managed by the International Pacific Halibut Commission (IPHC), discard mortality rates are estimated directly by observers when available. For unobserved trips the IPHC uses a range of discard mortality rates that vary by region (Bering Sea vs Gulf of Alaska), by target species , by tow depth, and by gear type (trawl, pot, and longline). Forty two possible combinations are considered in Table 2 in Dykstra (2017). Averages across gears are as follows: trawls 75.6%, pot 14.3%, longline 9.9%.) The DFO assessments for Atlantic halibut. Davis and Ryer (2003) reported a mortality rate of 100% after 30 days of holding in laboratory setting but the effects of captivity could not be isolated. den Heyer et al. (2015) reported roughly similar results for Atlantic halibut,

> "In general, halibut are thought to be robust to handling relative to other groundfish. Neilson et al. (1989) found that 35% of otter trawl-caught halibut and 77% of longline-caught halibut survived 48 hours in holding tanks. Recent deployments of PSAT tags

suggest that the survival of larger halibut caught by longline gear could be 100% (Armsworthy et al. 2014). Kaimmer and Trumble (1998) found that careful handling of Pacific halibut can increase discard survival and that even those fish with mild or moderate injuries have a higher than expected probability of survival. For example, 69% of Pacific halibut with moderate injuries survived and 43% of halibut with severe injuries survived."

Handling mortality for fish with expensive tags is likely to be low given likely bias in the capture, selection and handling of such fish. Nonetheless, the discard mortality in fixed gear is likely to be lower than in mobile gear. Given the current mixture of gears (primarily longlines) used to prosecute the Canadian fishery, den Heyer et al. (2015) used a discard mortality rate of 23%. DFO (2015) noted that their parameterization of discard mortality in their assessment and management models was based on rather old data and needed to be updated. Similar concerns were expressed by Leaman and Stewart (2017) in their scholarly review of the bases for discard mortality rates for Pacific halibut.

Given that US trawl fisheries are unlikely to target halibut, it seems reasonable to use the average of the discard mortality rates applied to Pacific halibut of 76%. To the best of my knowledge, there are no documented studies of discard mortality rates of halibut in gill nets. It seems reasonable to hypothesize it should be lower than trawl caught mortality but greater than longline estimates. Field studies for spiny dogfish yielded estimates of 30% mortality (Rulifson, East Carolina State Univ., personal communication).

The expected effects of discard mortality rates on total yield can be described simply as

$$C_t = L_t + \sum_{g=1}^{G} \alpha_g D_{g,t} \qquad [37]$$

Where α_g is the discard mortality rate for gear g and Dg,t is the total discard estimate for gear g at time t. The relative effect of α_g on the catch estimate will depend on the magnitude of the discard estimates. The FSD model projects the catch in year t+1 by adjusting the observed catch in year t by the estimated rates of change in the indices in year t. In terms of the observed index data, the effects of changes in the estimated catch do not change the adjustment factor applied. The two quantities are decoupled in this context. Discard mortality will not have any effect on the quota IF the relative magnitudes of the $D_{g,t}$ remains constant. However, if the balance of discards shift to say a less lethal gear, then there may be some room for increased landings, or less penalty for discards.

The relative importance of discarding in gill nets has been increasing relative to trawls (Fig. 28) in recent years. Estimated discards with and without adjustment for discard mortality (Fig. 29)

show some divergence in recent years due the increase in discarding by gill net trip. However, the overall effect on total catch (Fig. 30) shows less divergence.

Use of total rates (discard+kept) in observer data

The proposed methodology was presented to the NEFMC Plan Development Team on November 27, 2017. It was noted that the d/k ratio as a measure of relative abundance may underestimate relative abundance. The team suggested that the total catch of halibut (i.e., landings plus discards) would be a better estimate of relative abundance. Comparisons of the ratio of total halibut caught to the total landings of all species (t/k) with the halibut discard to total landings of all species (d/k) are summarized in Appendix 3 and Figure 3.1. As expected the t/k ratio is consistently greater than d/k but there are no marked changes in recent years (Appendix Fig. 3.2)

1%	5%	10%	25%	50%	75%	90%	95%	99%
98.51	105.14	109.21	115.539	122.80	130.90	138.64	143.37	151.91

The overall mean catch for 2018 using these parameters is 123.43 mt with a CV equal to 0.094. The time series projected catches based on t/k and the sampling distribution of catch in 2018 are shown in Fig. 31 and 32, respectively. FSD model results suggest no significant differences between catch estimates derived using the t/k indices vs the d/k indices. (See also FIG 3.3 to 3.5 in APPENDIX 3).

DISCUSSION

Quantification of the virgin stock size for a halibut fishery that began a quarter century after the Revolutionary War is difficult. Lear (1998) reported that Boston's Atlantic halibut market "began to outstrip the inshore supplies" in the 1820's and by 1836 a fishery was established on Georges Bank. Catches peaked in 1849 and declined rapidly since then. Landings reported in Hennen (2015) show a curious 2 year spike of 4,200 and 4,908 mt in 1895 and 1896 but no other landings have exceeded 943 mt since then. In the first 40 years of recorded landings (1893-1932) landings averaged about 5.5 times greater than during next 60 years (1933-2002).

Two factors, 1) large catches that occurred prior to the collection of synoptic catch recording programs and 2) an apparent dispersal of fishing activity to more fishing grounds even in the earliest years of the fishery, will make it difficult to interpret historic scale in analytic stock assessments. Such estimates will be driven necessarily by strong, but weakly supported assumptions. Whether the derived quantities are useful as an accurate reconstruction of the past is debatable. But the utility of such estimates for contemporary management will be undeniably low. Current stock sizes are likely to be a small fraction of the virgin abundance and rebuilding strategies will likely devolve into debates about the scientific credibility of the targets or dire externalities of restricting other fisheries to achieve halibut rebuilding targets.

These same basic concerns constrain the applicability of data poor methods. The basic methods essentially fall into four categories:

1. Methods that rely some arbitrary scalar adjustment to recent average catches with no rigorous analyses of population consequences.
2. Methods that rely on strong assumptions about current stock status
3. Methods that apply a biologically based harvest rate to a swept area estimate of abundance
4. Methods that adjust current catches based on measures of current trends or trends.

Methods based on category 1 are difficult to justify scientifically even if risk averse reductions are selected. Economic and social considerations, e.g., acceptable inter-annual percentage changes, will necessarily be major considerations. Many poor methods (Category 2) rely on assumptions that are usually the outcomes of complicated assessment models rather than the inputs to data-poor models. Swept area models (Category 3) were not considered for this analysis because catch rates for halibut appear to be very low. Given the low rates of encounter and likely overall low abundance it did not seem prudent to construct a proxy value for halibut capture efficiency. Notably, it was not possible to estimate calibration coefficients for halibut from experimental comparisons (Miller et al. 2010). Instead the calibration coefficient of 2.057 for halibut was taken as an average of estimates for 5 other flatfish species (Blaylock and Legault 2015).

The proposed approach is similar to Management Procedure (MP) approaches or Management Strategy Evaluation (MSE) methods as described by Geromont and Butterworth (2015), Kelly and Codling(2006), and many others. MPA methods have been applied to several ICES stocks, Greenland halibut (NAFO) and advocated by Parma (2002) for Pacific halibut and later by Webster (2017), Hicks and Stewart (2017)for Pacific halibut. Many recent surveys of data-poor methods conclude by supporting MP approaches in one form or another, and often concurrently highlighting poor performance of typical data poor methods (Carrruthers et al. 2014, Wilberg et al. 2011).

A set of papers in Aquatic Living Resources by Apostolaki and Hillary (2009) and Pomarede et al. (2010) provide a nice series of applications on the utility of control theory methods in fisheries assessments. Hillary (2009) illustrates these methods further and provides software appropriate for evaluating a suite of harvest control rules. Pomarede et al appears to be one of the first papers to introduce the PID control theory application. PID stands for Proportional Integral Derivative Controllers (Betts 2011, also http://controlguru.com). Geromont and Butterworth (2015, also 2001) describe a general "slope parameter" that is equivalent to what is typically referred to as a P controller. The FSD model would be referred to as a P-D controller (Betts 2010).

The proposed approach (FSD) to model halibut departs from the RYM in several important ways:

1. Does not assume r and K are constants
2. Focuses on recent changes and implications for catches
3. Does not attempt to estimate long-term reference points
4. Does not utilize M, or YPR concepts to define optimal r or F rates
5. Assumes that stock can be described by linear dynamics, i.e., stock is well below K
6. Applies no estimation of parameters except for aggregate rates of change in indices.

Two parameters are required to apply the FSD model. These are defined as the gain parameters for the proportional and derivative slope components. Their final selection is not currently based on any optimization. Instead, they are based on the likely tradeoffs such parameters imply in terms of average catch, variation in average catch, the likelihood of continued population growth and the risk of overfishing.

Because the FSD model does not compute any of the standard stock status parameters, it is not possible to rigorously define stock status. However, results of the Rcrit analyses do offer some insights into stock status. The review panel for the Operational Assessment in 2015 wrote:

"The GARMIII benchmark assessment and the 2012 update assessment concluded that the stock was overfished but overfishing was not occurring. All information available in the update assessment indicates that stock size has not substantially increased. Therefore, based on the long-term exploitation history and survey trends, the Panel concludes that the stock is still overfished. However, the overfishing status is unknown. Considering the instability of the assessment model, the overfishing threshold was not updated."

Using the results of the Rcrit analyses, it would appear that the stock size has significantly increased since 2005. The overall Rcrit value for the 2005-2016 period suggests an increase of 3 to 5 times (Table 9). In turn these rates suggest annual abundance increases or 9 to 12% per year over the past decade. Randomization tests suggest that all of the increases are statistically significant (P<0.01, Table 9). Catches have increased about 3 fold over this period as well.

Computation of population increases (both relative and absolute) and total catch in the Canadian 3NOPs4VWX5Zc stock reveal increases of 2.25X in catch and 2.92X in relative abundance and 2.73X in modeled absolute abundance (Table 11). Hence the changes in US stock relative abundance have mirrored those observed for the much larger Canadian stock. Such increases in US stock would be unlikely if overfishing were still occurring. Model-based estimates of fishing mortality appear to be decreasing. Taken together, the evidence suggest that recent catches have been sufficiently low to allow the stock in US waters to increase at a rate comparable to that observed in Canada.

Results of the DCAC model based on the combination of Rcrit and Envelope methods were largely inconclusive with respect to the determination of overfished status.   In contrast, the proposed FSD harvest control rule appears to have some desirable properties with respect to detection of underlying trends and with respect to continuation of rebuilding program for halibut. Simulation methods suggest the model can control populations when productivity is changing temporally. Bootstrap analyses of the model forecasts suggest and 80% confidence interval of 109 to 138 mt and median of 123 mt for 2018. Note that this assumes a K_p=0.75 and K_d=0.5

Applications of the model to two other managed halibut stocks suggests potential utility for the US stock of Atlantic halibut and perhaps other stocks in the Northeast. Comparisons between the FSD model and analytical models would be a first step. Another important consideration is the estimation of the slope and change in slope. The 5-point regression might be improved by using a Kalman filter or other state-space model. In theory a MLE based smooth of the index data would be preferable to slopes estimated by an n-point regression.

REFERENCES

Apostolaki, P and R. Hillary 2009. Harvest control rules in the context of fishery-independent management of fish stocks. Aquatic Living Resources 22:217-224.

Arnold, L.M.; Heppell, S.S. (2015). Testing the robustness of data-poor assessment methods to uncertainty in catch and biology: a retrospective approach. *ICES Journal of Marine Science* 72: 243–250.

Armsworthy SL, and Campana SE (2010) Age determination, bomb-radiocarbon validation and growth of Atlantic halibut (*Hippoglossus hippoglossus*) from the Northwest Atlantic. Environ Biol Fishes 89:279–295

Berkson, J.; Barbieri, L.; Cadrin, S.and Cass-Calay, S.; Crone, P.; Dorn, P.; Friess, C.; Kobayashi, D.; Miller, T.; Patrick, W.; Pautzke, S.; Ralston, S.; Trianni, M. (2011). Calculating acceptable biological catch for stocks that have reliable catch data only. NOAA Technical memorandum. National Marine Fisheries Service, National Oceanic and Atmospheric Administration of the US Department of Commerce.

Berkson, J.; Thorson, J.T. (2015). The determination of data-poor catch limits in the United States: is there a better way? *ICES Journal of Marine Science* 72: 237–242.

Betts 2010. Advances in Design and Control. SIAM. See also http://controlguru.com/about/

Brodziak, J. 2002. Atlantic halibut. In NEFSC, *Assessment of 20 Northeast groundfish stocks through 2001*, pp.306-313. US Department of Commerce, Northeast Fisheries Science Center Reference Document 02-16, 509 p. Available at:
http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0216/

Brodziak, J., and Col, L. 2005. Atlantic halibut. In NEFSC, *Assessment of 19 Northeast groundfish stocks through 2004*. 2005 Groundfish Assessment Review Meeting (2005 GARM), Northeast Fisheries Science Center, Woods Hole, MA, 15-19 Aug. 2005.pp.424-431. US Department of Commerce, Northeast Fisheries Science Center Reference Document 05-13, 499 p. Available at: http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0513/

Col, L.A. and C. L. Legault. 2009. The 2008 assessment of Atlantic halibut in the Gulf of Maine-Georges Bank Region. US Department of Commerce, Northeast Fisheries Science Center Reference Document 09-08. 39 p. Available at:
http://nefsc.noaa.gov/publications/crd/crd0908/

Carruthers, T.R.; Punt, A.E.; Walters, C.J.; MacCall, A.; McAllister, M.K.; Dick, E.J.; Cope, J. (2014). Evaluating methods for setting catch limits in data-limited fisheries. *Fisheries Research* 153: 48–68.

Cox, S.P., Benson, A., and den Heyer, C.E. 2016. Framework for the Assessment of Atlantic Halibut Stocks on the Scotian Shelf and Southern Grand Banks. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/001. v + 57 p.

Davis, M.W. 2002. Key principles for understanding fish bycatch discard mortality. Can. J. Fish. Aquat. Sci. 59: 1834-1843.

Davis, M.W., and Olla, B.L. 2001. Stress and delayed mortality induced in Pacific halibut Hippoglossus stenolepis by exposure to hooking, net towing, elevated sea water temperature and air: implications for management of bycatch. N. Am. J. Fish. Manage. 21: 725-732.

Davis, M. W. and C. H. Ryer. 2003. Understanding Fish Bycatch Discard And Escapee Mortality. AFSC Quarterly Report. National Marine Fisheries Service. Seattle WA. 9 pp.

DFO. 2015. Assessment of Greenland Halibut in the Gulf of St. Lawrence (4RST) in 2014. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/056.

DFO. 2015. 2014 Assessment of Atlantic Halibut on the Scotian Shelf and Southern Grand Banks (NAFO Divisions 3NOPs4VWX5Zc). DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/012.

DFO. 2015. Assessment of Greenland Halibut in the Gulf of St. Lawrence (4RST) in 2014. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/056.

DFO. 2015. Stock Assessment of Atlantic Halibut of the Gulf of St. Lawrence (NAFO Divisions 4RST) for 2013 and 2014. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/023.

DFO. 2015. Proceedings of the Regional Framework and Assessment of the Scotian Shelf and Southern Grand Banks Atlantic Halibut (*Hippoglossus hippoglossus*) in NAFO Divisions 3NOPs4VWX5Zc; November 3-6, 2014, and December 8-9, 2014. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2015/040.

DFO. 2016. Update of main indicators of the stock status of Greenland Halibut in the Gulf of St. Lawrence (4RST) in 2015. DFO Can. Sci. Advis. Sec. Sci. Resp. 2016/003.

DFO. 2016. Stock Status Update of Atlantic Halibut on the Scotian Shelf and Southern Grand Banks (NAFO Divs. 3NOPs4VWX5Zc). DFO Can. Sci. Advis. Sec. Sci. Resp. 2016/010.

DFO. 2017. Proceedings of the Regional Framework Discussion on Stock Substructure of Northwest Atlantic Fisheries Organization (NAFO) Divisions 3NOPs4VWX5Zc Atlantic Halibut, *Hippoglossus hippoglossus* (Scotian Shelf and Southern Grand Banks); March 9-11, 2016. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2017/010.

Deroba, J. and J. Bence. 2008. A review of harvest policies: Understanding relative performance of control rules. Fisheries Research 94:210-223.

Dick, E.; MacCall, A.D. (2011). Depletion-Based Stock Reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. *Fisheries Research* 110 (2): 331–341.

Edwards, C.T.T. 2015. **Review of data-poor assessment methods for New Zealand fisheries.** New Zealand Fisheries Assessment Report No. 2015/27. ISSN 1179-5352 (online) ISBN 978-0-477-10596-5 (online)

Gedamke, T.; Hoenig, J.M. (2006). Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. *Transactions of the American Fisheries Society* 135 (2): 476–487.

Geromont, H.F.; Butterworth, D.S. (2015a). Complex assessments or simple management procedures for efficient fisheries management: a comparative study. *ICES Journal of Marine Science* 72: 262–274.

Geromont, H.F.; Butterworth, D.S. (2015b). Generic management procedures for data-poor fisheries: forecasting with few data. *ICES Journal of Marine Science* 71: 251–261.

Grasso, G.M. 2008. What appeared limitless plenty: the rise and fall of the nineteenth-century Atlantic halibut fishery. Environmental History 13: 66-91.

Hansell,A., G. DeCelles, S. Cadrin. 2017. Standardized Catch Per Unit Effort (CPUE) of Atlantic Halibut (*Hippoglossus hippoglossus*) in the Gulf of Maine. Draft Working Paper.
Hillary, R. M. 2009. An introduction to FLR fisheries simulation tools. Aquat. Living Resour. 22:225–232
Hennen 2015. Atlantic halibut operational assessment.

den Heyer, C., Schwarz, C., and Trzcinski, M. K. 2013. Fishing and natural mortality rates of Atlantic Halibut (Hippoglossus hippoglossus) estimated from multi-year tagging and life history. Transactions of the American Fisheries Society, 142: 690–702.

den Heyer, C.E., Hubley, B., Themelis, D., Smith, S.C., Wilson, S., and Wilson, G. 2015. Atlantic Halibut on the Scotian Shelf and Southern Grand Banks: Data Review and Assessment Model Update. DFO Can. Sci. Advis. Sec. Res. Doc. 2015/051. v + 82 p.

Hicks, A. C. and S.Keith 2017. 4.6 Developments in Management Strategy Evaluation/ Management Strategy Advisory Board. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC–2016–RARA–26–R:439-442.

Hilborn, R. 2002. THE DARK SIDE OF REFERENCE POINTS. BULLETIN OF MARINE SCIENCE, 70(2): 403–408, 2002

ICES. 2015. Report of the Inter Benchmark Process on Greenland Halibut in ICES ar-eas I and II (IBPHALI), August 2015, By correspondence. ICES CM 2015\ACOM:54. 41 pp. (have)

Jardim, E., Millar, C.P., Mosqueira, I., Scott, F., Osio, G. C., Ferretti, M., Alzorriz, N., and Orio, A. 2015. What if stock assessment is as simple as a linear model? The a4a initiative. ICES J. Mar. Sci. 72:232-236.

Jensen, A.L. (1996). Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. *Canadian Journal of Fisheries and Aquatic Sciences* 53: 820–822.

Kanwit, J.K. 2007. Tagging results from the 2000-2004 Federal Experimental Fishery for Atlantic halibut (*Hippoglossus hippoglossus*) in the Eastern Gulf of Maine. Journal of Northwest Atlantic Fishery Science 38:37-42.

Kelly, C. J, and E. A. Codling. 2006. "Cheap and dirty" fisheries science and management in the North Atlantic. Fisheries Research 79:233-238.

Kerr, L. A., Hintzen, N. T., Cadrin, S. X., Clausen, L., Worsøe Dickey-Collas, M., Goethel, D. R., Hatfield, E. M.C., Kritzer, J. P., and Nash, R.D.M. 2016. Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. – ICES Journal of Marine Science, 1-15. doi:10.1093/icesjms/fsw188.

Klaer, N.L.; Wayte, S.E.; Fay, G. (2012). An evaluation of the performance of a harvest strategy that uses an average-length-based assessment method. *Fisheries Research* 134-136: 42–51.

Leaman, B. M. and I.J. Stewart. 2017. 2.12 Research basis for estimated Discard Mortality Rates used for Pacific halibut in longline and trawl fisheries. . Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC–2016–RARA–26–R:133-172.

Lear, W. H.. 1998. History of Fisheries in the Northwest Atlantic: The 500-Year Perspective. J. Northw. Atl. Fish. Sci., Vol. 23: 41–73

MacCall, A.D. (2009). Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. *ICES Journal of Marine Science* 66 (10): 2267–2271.

Miller, T, and P. Rago 2012. Empirical Exploration of Feasible Bounds on Butterfish Stock Size and Fishing Mortality Rates, 1975-2011. Report to the Mid-Atlantic Fishery Management Council Scientific and Statistical Committee. Northeast Fisheries Science Center National Marine Fisheries Service. Woods Hole, MA 02543

Miller, T.J., Das, C., Politis, P.J., Miller, A.S., Lucey, S.M., Legault, C.M., Brown, R.W., and Rago, P.J. 2010. Estimation of Albatross IV to Henry B. Bigelow calibration factors. US Department of Commerce, Northeast Fisheries Science Center Reference Document 10- 05; 233 p. Available at: http://www.nefsc.noaa.gov/publications/crd/crd1005/index.html

Neilson, J.D.,Waiwood, K. G., and Smith, S. J. 1989. Survival of Atlantic halibut (Hippoglossus hippoglossus) caught by longline and otter trawl gear. Canadian Journal of Fisheries and Aquatic Sciences, 46:887–897.

New England Fishery Management Council. 1998. Evaluation of existing overfishing definitions and recommendations for new overfishing definitions to comply with the Sustainable Fisheries Act. NEFMC, 50 Water Street, Mill 2 Newburyport, MA 01950. 308 *p.*

Newman, D., J. Berkson, and L. Suatoni. 2015. Current methods for setting catch limits for data-limited fish stocks in the United States. Fisheries Research 164:86-93.

Northeast Fisheries Science Center. 2012. Assessment or Data Updates of 13 Northeast Grounfish Stocks through 2010. US Dept Commerce, Northeast Fish Sci Cent Ref Doc. 12-06; 789 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. CRD12-06

Northeast Fisheries Science Center. 2001. Assessment of 19 Northeast groundfish stocks through 2000. NEFSC Reference Document 01-20, Woods Hole, MA, 02543.

Northeast Fisheries Science Center. 2002. Final Report of the Working Group on Re-Evaluation of Biological Reference Points for New England Groundfish. NEFSC Reference Document 02-04, Woods Hole, MA 02543.

Parma, A. M. 1993. Retrospective analysis of catch at age of Pacific halibut: Implications on assessment of harvesting policies. Pp. 247-289, in Proceedings of International Symposium on Management Strategies for Exploited Fish Populations, Alaska Sea Grant Program. AK-SG-93-02.

Parma, A. M. 2002. In search of robust harvest rules for Pacific halibut in the face of uncertain assessments and decadal changes in productivity. Bulletin of Marine Science 70:423-453.

Patrick, W.; Spencer, P.; Ormseth, O.; Cope, J.; Field, J.; Kobayashi, D.; Gedamke, T.; Cort´es, E.; Bigelow, K.; Overholtz, W.; Link, J.; Lawson, P. (2011). Use of productivity and susceptibility indices to determine stock vulnerability, with example applications to six U.S. fisheries. Technical memorandum. National Marine Fisheries Service, National Oceanic and Atmospheric Administration of the US Deparment of Commerce.

Pomarede, M. , R. Hillary, L. Ibaibarriaga, J. Bogaards and P. Apostolaki. 2010. Evaluating the performance of survey-based operational management procedures. Aquat. Living Resour. 23: 77-94.

Prince, J.; Hordyk, A.; Valencia, S.R.; Loneragan, N.; Sainsbury, K. (2015). Revisiting the concept of Beverton-Holt life-history invariants with the aim of informing data-poor fisheries assessment. *ICES Journal of Marine Science* 72: 194–203.

Prince, J.D.; Dowling, N.A.; Davies, C.R.; Campbell, R.A.; Kolody, D.S. (2011). A simple cost-effective and scale-less empirical approach to harvest strategies. *ICES Journal of Marine Science* 68 (5): 947–960.

Restrepo, V.R.; Thompson, G.G.; Mace, P.M.; Gabriel, W.L.; Low, L.L.; MacCall, A.D.; Methot, R.D.; Powers, J.E.; Taylor, B.L.; Wade, P.R.; Witzig, J.F. (1998). Technical guidance on the use of precautionary approaches to implementing national standard 1 of the Magnuson-Stevens Fishery Conservation and Management act. Technical Memorandum. National Marine Fisheries Service, National Oceanic and Atmospheric Administration of the US Department of Commerce.

Shackell, N.L., Kenneth T. Frank, Janet A. Nye, Cornelia E. den Heyer. 2016. A transboundary dilemma: dichotomous designations of Atlantic halibut status in the Northwest Atlantic. ICES J. Marine Sci.73:1798-1805

Scott, W.B. and M.G. Crossman. 1988. Atlantic fishes of Canada. Canadian Bulletin of Fisheries and Aquatic Sciences No. 219: 731 p.

Sigourney, D.B., Ross, M.R., Brodziak, J., and Burnett, J. 2006. Length at Age, Sexual Maturity and Distribution of Atlantic Halibut, *Hippoglossus hippoglossus* L., off the Northeast USA. Journal of Northwest Atlantic Fishery Science 36:81-90.

Stewart, I. J. 2017. 4.1 Overview of data sources for the Pacific halibut stock assessment and related analyses. . Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC–2016–RARA–26–R:279-364.

Stewart, I.J., and A.C. Hicks. 2017. 4.2 Assessment of the Pacific halibut stock at the end of 2016. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC–2016–RARA–26–R:365-394.

Trzcinski, M. K., and Bowen, W. D. The recovery of Atlantic halibut: a large, long-lived, and exploited marine predator. – ICES Journal of Marine Science, 73: 1104–1114.

Tukey, J. W. 1977. Exploratory Data Analysis. Addison-Wesley. NY

Webster, R. A. 2017. Results of space-time modelling of IPHC fishery-independent setline survey WPUE and NPUE data. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC-2016-RARA-26-R: 241-257.

Weidenmann, J., M. J. Wilberg, and T. J. Miller. 2013. An evaluation of harvest control rules for data-poor fisheries. N. Amer. J. Fisheries Manage. 33:845-860.

Wigley, S.E., Rago, P.J., Sosebee, K.A., and Palka, D.L. 2007. The analytic component to the Standardized Bycatch Reporting Methodology omnibus amendment: sampling design and estimation of precision and accuracy. US Department of Commerce, Northeast Fisheries Science Center Reference Document 07-09, p.156. Available at: http://nefsc.noaa.gov/publications/crd/crd0709/

Wilberg, M.; Miller, T.; Wiedenmann, J. (2011). Evaluation of Acceptable Biological Catch (ABC) control rules for Mid-Atlantic stocks. *Report to the Mid-Atlantic Fishery Management Council. UMCES, CBL 11-029. Technical Report Series No. TS-622-11* .

Wise, J.P., and Jensen, A.C. 1959. Movement of tagged halibut off New England. Trans. Amer. Fish. Soc. 88:357-358.

TABLES

Table1. Summary of model inputs used in the Rcrit, Envelope, DCAC and FSD models for US and Canada (3NOPs4VWX5Zc) Atlantic halibut stocks.

Stock	Year	Discards (mt)	Landings (mt)	Catch (mt)	NEFSC fall survey (kg/tow)	d/k ratio for gill-net trips	d/k ratio for trawl trips	Standardized Longline CPUE	ME_NH inshore trawl surey SPRING	ME_NH inshore trawl surey FALL
US area	2002	20.20	10.01	30.21	0.0041	0.0000000	0.000224	2.431442	0.0521	0.0228
	2003	20.15	16.68	36.83	0.049	0.0003620	0.000125	0.536988	0.2198	0.1736
	2004	15.71	11.22	26.93	0.1119	0.0000950	0.000166	1.22045	0.2864	0.1164
	2005	18.89	16.81	35.70	0.1105	0.0001548	0.000185	1.894313	0.2672	0.2296
	2006	22.45	14.08	36.53	0.0312	0.0002231	0.000224	1.35618	0.9165	0.1528
	2007	17.27	24.61	41.88	0.0774	0.0001075	0.000146	0.894835	0.5177	0.2805
	2008	21.66	28.69	50.35	0.0701	0.0001204	0.000174	1.150346	0.6285	0.7342
	2009	17.85	45.05	62.90	0.0633	0.0000560	0.000227	0.800941	0.9003	0.5314
	2010	34.68	20.20	54.88	0.098	0.0002818	0.00045	0.78386	0.6337	0.5342
	2011	42.34	25.79	68.13	0.0638	0.0005589	0.000652	1.520806	0.6401	1.1621
	2012	52.18	34.80	86.98	0.1241	0.0005674	0.000957	1.61151	0.9459	0.3106
	2013	56.16	34.67	90.83	0.0331	0.0010828	0.001103	1.60406	0.3919	0.3996
	2014	34.33	44.99	79.32	0.1821	0.0009006	0.000595	1.817722	0.4755	0.6448
	2015	46.28	62.00	108.28	0.3011	0.0020334	0.000499	1.573949	0.3535	0.2180
	2016	47.39	68.20	115.59	0.0598	0.0021923	0.000684	1.943505	0.5943	0.1160

Stock	Year			Catch (mt)	RV Summer	RV Spring	LL CPUE	FS Index		
Canada	2002			1,493	0.15	0.0183	111.61	27.67		
	2003			1,600	0.14	0.0258	111.49	28.35		
	2004			1,465	0.25	0.0222	86.21	37.24		
	2005			1,336	0.31	0.0083	104.49	36.12		
	2006			1,395	0.28	0.0099	135.28	42.33		
	2007			1,562	0.47	0.0512	113.48	41.3		
	2008			1,494	0.37	0.0467	136.1	53.84		
	2009			2,144	0.36	0.0474	164.62	67.94		
	2010			1,853	0.69	0.0751	128.1	66.98		
	2011			1,822	0.8	0.0740	127.96	90.25		
	2012			2,220	0.6	0.0911	158.07	84.55		
	2013			2,599	0.5	0.1277	129.54	93.39		
	2014			2,952	0.63		174.4	90.78		
	2015			3,236	0.67		232.1	151.39		
	2016			4,109	0.62		186.73	120.07		

Table 1.5 Discard estimates by gear type 1989-2016 for adjusted for discard mortality rates. Estimates of discard mortality are based on Pacific halibut estimates.

									Adjusted for Discard Mortality Rate			
									0.1	0.6	0.4	
	Discard Estimate (mt)				Fraction by Gear							
Year	Handline	Trawl	Gill net	Total	Handline	Trawl	Gill net		Handline	Trawl	Gill net	Adj Total
1989	0.00	2.88	2.10	4.97	0.000	0.578	0.422		0.00	1.73	0.84	2.56
1990	0.00	12.09	1.46	13.55	0.000	0.892	0.108		0.00	7.25	0.58	7.84
1991	0.00	6.06	0.87	6.93	0.000	0.875	0.125		0.00	3.64	0.35	3.98
1992	0.00	1.92	0.27	2.19	0.000	0.878	0.122		0.00	1.15	0.11	1.26
1993	0.00	0.63	0.44	1.06	0.000	0.590	0.410		0.00	0.38	0.17	0.55
1994	0.00	2.94	0.22	3.16	0.000	0.930	0.070		0.00	1.76	0.09	1.85
1995	0.00	6.30	0.04	6.34	0.000	0.993	0.007		0.00	3.78	0.02	3.80
1996	0.00	0.52	0.14	0.65	0.000	0.791	0.209		0.00	0.31	0.05	0.36
1997	0.00	1.64	0.00	1.64	0.000	1.000	0.000		0.00	0.98	0.00	0.98
1998	0.00	0.00	0.10	0.10	0.000	0.000	1.000		0.00	0.00	0.04	0.04
1999	0.00	68.85	0.25	69.10	0.000	0.996	0.004		0.00	41.31	0.10	41.41
2000	0.00	11.38	0.49	11.87	0.000	0.958	0.042		0.00	6.83	0.20	7.03
2001	0.00	9.29	0.40	9.68	0.000	0.959	0.041		0.00	5.57	0.16	5.73
2002	0.00	20.20	0.00	20.20	0.000	1.000	0.000		0.00	12.12	0.00	12.12
2003	0.00	15.80	4.35	20.15	0.000	0.784	0.216		0.00	9.48	1.74	11.22
2004	0.02	14.81	0.88	15.71	0.001	0.943	0.056		0.00	8.89	0.35	9.24
2005	0.70	16.90	1.29	18.89	0.037	0.895	0.068		0.07	10.14	0.52	10.73
2006	0.00	19.05	3.40	22.45	0.000	0.849	0.151		0.00	11.43	1.36	12.79
2007	0.08	14.65	2.54	17.27	0.004	0.848	0.147		0.01	8.79	1.02	9.82
2008	0.00	18.87	2.79	21.66	0.000	0.871	0.129		0.00	11.32	1.12	12.44
2009	0.00	16.93	0.92	17.85	0.000	0.949	0.051		0.00	10.16	0.37	10.53
2010	2.52	27.55	4.63	34.69	0.073	0.794	0.134		0.25	16.53	1.85	18.63
2011	0.07	33.56	8.71	42.35	0.002	0.793	0.206		0.01	20.14	3.49	23.63
2012	0.00	43.51	8.68	52.19	0.000	0.834	0.166		0.00	26.11	3.47	29.58
2013	0.20	46.27	9.70	56.18	0.004	0.824	0.173		0.02	27.76	3.88	31.66
2014	0.00	23.95	10.39	34.34	0.000	0.697	0.303		0.00	14.37	4.16	18.52
2015	0.00	22.48	23.82	46.30	0.000	0.485	0.515		0.00	13.49	9.53	23.01
2016	0.00	26.00	21.40	47.40	0.000	0.549	0.451		0.00	15.60	8.56	24.16

Table 2.

Table xx. Summary of ratio test simulations for estimation of bias in mean and median of Rcrit as a function of the magnitude of true
rate of change (Rcrit_true), the variation of the observation error (CV) and the number of relative abundance indices (Nvar).
All simulations were based on a time series of length 10, and the ratio of the average of the last 3 to the first 3 observations
for 2000 randomizations of each of 1000 stochastic realizations.

		Relative Bias in Estimated Rcrit vs True Rcrit							
		Nvar=1		Nvar=2		Nvar=3		Nvar=5	
Rcrit_true	CV	Rel Bias (mean)	Rel Bias (median)	Rel Bias (mean)2	Rel Bias (median)	Rel Bias (mean)4	Rel Bias (median)	Rel Bias (mean)6	Rel Bias (median)
2.014	0.1	0.3%	0.2%	0.1%	-0.1%	0.0%	0.1%	0.2%	0.0%
2.014	0.15	0.4%	-0.3%	0.6%	0.1%	0.0%	-0.5%	0.2%	-0.1%
2.014	0.2	0.6%	-1.4%	0.1%	-0.5%	0.4%	0.1%	0.1%	0.1%
2.014	0.25	3.1%	1.2%	1.5%	0.5%	1.2%	0.5%	0.5%	-0.1%
2.014	0.3	2.5%	-0.9%	1.0%	0.4%	1.8%	0.6%	0.7%	-0.1%
2.014	0.35	3.5%	-0.8%	2.5%	0.9%	1.1%	-0.5%	0.7%	0.0%
2.014	0.4	4.9%	-0.9%	3.3%	0.4%	1.8%	0.3%	0.8%	-0.4%
2.014	0.45	10.1%	-0.9%	2.8%	-0.5%	1.9%	0.3%	1.6%	0.5%
2.014	0.5	9.8%	-1.3%	6.1%	0.5%	3.8%	0.2%	1.9%	0.1%
2.014	0.6	-51.9%	-3.0%	6.8%	-1.6%	4.8%	0.2%	2.8%	0.1%
2.014	0.65	18.4%	-1.4%	9.0%	-0.5%	5.1%	0.2%	2.8%	-1.1%
2.014	0.7	7.8%	-5.1%	12.9%	0.7%	3.8%	0.2%	4.1%	0.8%
1.419	0.1	0.1%	-0.1%	0.3%	0.2%	0.2%	0.0%	-0.1%	0.0%
1.419	0.15	1.2%	0.5%	0.8%	0.9%	0.5%	0.5%	0.0%	-0.3%
1.419	0.2	1.5%	1.1%	1.0%	0.1%	0.1%	-0.4%	0.5%	0.2%
1.419	0.25	0.7%	-1.8%	0.5%	-0.2%	0.8%	-0.1%	1.2%	1.1%
1.419	0.3	4.3%	0.1%	2.1%	0.9%	0.6%	-0.4%	0.5%	-0.8%
1.419	0.35	4.5%	0.4%	0.4%	-1.9%	1.0%	-0.1%	0.7%	-0.1%
1.419	0.4	5.9%	1.1%	3.7%	0.7%	1.9%	0.3%	1.0%	-0.3%
1.419	0.45	9.2%	-0.4%	2.0%	-1.2%	2.4%	0.0%	1.3%	-0.1%
1.419	0.5	8.5%	1.8%	5.1%	-0.6%	3.0%	0.5%	2.2%	0.2%
1.419	0.6	24.1%	-0.4%	6.8%	-0.6%	3.2%	-0.5%	2.4%	1.1%
1.419	0.65	17.5%	-0.6%	16.9%	2.1%	4.8%	-1.0%	3.1%	0.5%
1.419	0.7	23.5%	-3.0%	12.1%	3.1%	3.5%	-2.1%	1.6%	-2.1%
1.191	0.1	0.3%	0.5%	0.1%	0.1%	0.2%	-0.1%	-0.1%	-0.2%
1.191	0.15	0.4%	0.0%	0.3%	0.1%	0.2%	0.0%	0.2%	0.0%
1.191	0.2	1.7%	0.2%	0.4%	-0.4%	0.5%	0.1%	0.2%	-0.4%
1.191	0.25	1.5%	-0.4%	1.4%	0.5%	1.3%	1.3%	0.2%	-0.4%
1.191	0.3	2.8%	-0.2%	1.5%	0.7%	0.6%	-0.7%	-0.2%	-0.7%
1.191	0.35	4.6%	1.8%	2.8%	0.9%	2.1%	1.1%	0.6%	-0.9%
1.191	0.4	5.3%	0.1%	2.7%	-0.4%	1.9%	0.5%	1.1%	0.7%
1.191	0.45	8.3%	-0.2%	3.7%	1.7%	2.5%	1.4%	0.8%	-1.1%
1.191	0.5	20.9%	3.8%	3.8%	-1.0%	2.8%	-0.6%	2.1%	0.0%
1.191	0.6	14.8%	1.1%	7.3%	1.7%	3.5%	-0.5%	2.4%	-0.1%
1.191	0.65	26.4%	1.8%	11.9%	0.9%	4.9%	-1.1%	2.3%	-0.9%
1.191	0.7	0.3%	-2.6%	9.7%	-0.6%	8.1%	3.7%	3.1%	-1.8%

Table 3.

Table xx. Summary of ratio test simulations for estimation of the average probability value for simulated Rcrit values as a functio
rate of change (Rcrit_true), the variation of the observation error (CV) and the number of relative abundance indices (N
All simulations were based on a time series of length 10, and the ratio of the average of the last 3 to the first 3 observa
for 2000 randomizations of each of 1000 stochastic realizations.

		Average Probability Value for Rcrit			
Rcrit_true	CV	Nvar=1	Nvar=2	Nvar=3	Nvar=5
2.014	0.1	0.000	0.000	0.000	0.000
2.014	0.15	0.002	0.000	0.000	0.000
2.014	0.2	0.010	0.000	0.000	0.000
2.014	0.25	0.021	0.002	0.000	0.000
2.014	0.3	0.042	0.008	0.001	0.000
2.014	0.35	0.066	0.015	0.003	0.000
2.014	0.4	0.095	0.027	0.005	0.001
2.014	0.45	0.115	0.047	0.012	0.001
2.014	0.5	0.148	0.058	0.020	0.005
2.014	0.6	0.199	0.103	0.040	0.013
2.014	0.65	0.214	0.120	0.052	0.019
2.014	0.7	0.241	0.136	0.070	0.025
1.419	0.1	0.008	0.000	0.000	0.000
1.419	0.15	0.036	0.005	0.001	0.000
1.419	0.2	0.085	0.022	0.006	0.001
1.419	0.25	0.132	0.054	0.020	0.004
1.419	0.3	0.163	0.083	0.046	0.013
1.419	0.35	0.202	0.130	0.076	0.029
1.419	0.4	0.234	0.149	0.098	0.044
1.419	0.45	0.263	0.200	0.123	0.065
1.419	0.5	0.278	0.204	0.143	0.085
1.419	0.6	0.316	0.253	0.192	0.128
1.419	0.65	0.335	0.249	0.205	0.148
1.419	0.7	0.353	0.271	0.229	0.178
1.191	0.1	0.084	0.022	0.005	0.001
1.191	0.15	0.171	0.086	0.044	0.013
1.191	0.2	0.224	0.151	0.094	0.046
1.191	0.25	0.284	0.190	0.145	0.093
1.191	0.3	0.317	0.234	0.198	0.139
1.191	0.35	0.339	0.269	0.218	0.166
1.191	0.4	0.354	0.302	0.250	0.205
1.191	0.45	0.372	0.314	0.270	0.235
1.191	0.5	0.368	0.338	0.304	0.244
1.191	0.6	0.403	0.361	0.330	0.281
1.191	0.65	0.406	0.366	0.342	0.305
1.191	0.7	0.419	0.392	0.328	0.317

Table 4. Summary of fraction of simulations with significance probabilities less than or equal to the column headers (P0.005, P0.01…) for varying levels of lambda, Rcrit and the CV of the simulated observations. Estimates in this table assume **only one** index of abundance is available. The color formatting in this table is consistent for all tables 4,5, 6, and 7.

lambda	Rcrit_true	CV	P0.005	P0.01	P0.025	P0.05	P0.1	P0.15	P0.2	P0.25
0.1	2.014	0.1	0.994	0.998	1	1	1	1	1	1
0.1	2.014	0.15	0.868	0.951	0.989	0.999	1	1	1	1
0.1	2.014	0.2	0.634	0.771	0.914	0.958	0.993	0.994	0.998	0.998
0.1	2.014	0.25	0.426	0.578	0.769	0.891	0.959	0.98	0.99	0.993
0.1	2.014	0.3	0.278	0.41	0.619	0.773	0.883	0.937	0.96	0.973
0.1	2.014	0.35	0.209	0.303	0.475	0.643	0.799	0.869	0.915	0.936
0.1	2.014	0.4	0.142	0.217	0.378	0.527	0.701	0.802	0.864	0.907
0.1	2.014	0.45	0.123	0.181	0.309	0.463	0.642	0.753	0.819	0.863
0.1	2.014	0.5	0.076	0.135	0.248	0.394	0.575	0.679	0.76	0.811
0.1	2.014	0.6	0.06	0.1	0.18	0.288	0.458	0.559	0.645	0.699
0.1	2.014	0.65	0.052	0.086	0.171	0.275	0.44	0.543	0.624	0.685
0.1	2.014	0.7	0.043	0.067	0.142	0.251	0.379	0.487	0.573	0.635
0.05	1.419	0.1	0.644	0.807	0.926	0.969	0.995	0.998	1	1
0.05	1.419	0.15	0.319	0.462	0.641	0.786	0.908	0.949	0.968	0.984
0.05	1.419	0.2	0.177	0.271	0.45	0.6	0.752	0.815	0.87	0.914
0.05	1.419	0.25	0.084	0.132	0.257	0.388	0.6	0.7	0.769	0.817
0.05	1.419	0.3	0.094	0.137	0.231	0.347	0.511	0.624	0.704	0.773
0.05	1.419	0.35	0.051	0.087	0.169	0.283	0.432	0.558	0.643	0.706
0.05	1.419	0.4	0.028	0.061	0.151	0.253	0.374	0.486	0.575	0.646
0.05	1.419	0.45	0.035	0.058	0.114	0.203	0.34	0.432	0.522	0.593
0.05	1.419	0.5	0.021	0.043	0.093	0.16	0.292	0.408	0.504	0.579
0.05	1.419	0.6	0.018	0.031	0.075	0.146	0.253	0.346	0.428	0.502
0.05	1.419	0.65	0.013	0.023	0.062	0.118	0.237	0.322	0.401	0.469
0.05	1.419	0.7	0.016	0.027	0.049	0.112	0.218	0.304	0.368	0.448
0.025	1.191	0.1	0.182	0.268	0.438	0.603	0.75	0.815	0.867	0.9
0.025	1.191	0.15	0.061	0.111	0.227	0.356	0.501	0.609	0.696	0.756
0.025	1.191	0.2	0.037	0.078	0.156	0.26	0.396	0.501	0.589	0.661
0.025	1.191	0.25	0.027	0.047	0.109	0.187	0.318	0.42	0.491	0.556
0.025	1.191	0.3	0.015	0.028	0.075	0.141	0.26	0.352	0.438	0.512
0.025	1.191	0.35	0.02	0.033	0.075	0.138	0.246	0.336	0.412	0.473
0.025	1.191	0.4	0.018	0.029	0.057	0.114	0.199	0.282	0.372	0.435
0.025	1.191	0.45	0.008	0.02	0.047	0.093	0.189	0.27	0.355	0.42
0.025	1.191	0.5	0.011	0.031	0.064	0.109	0.191	0.268	0.352	0.431
0.025	1.191	0.6	0.014	0.02	0.049	0.088	0.172	0.243	0.313	0.382
0.025	1.191	0.65	0.008	0.02	0.045	0.092	0.163	0.229	0.298	0.362
0.025	1.191	0.7	0.005	0.017	0.049	0.085	0.151	0.211	0.287	0.34

Table 5. Summary of fraction of simulations with significance probabilities less than or equal to the column headers (P0.005, P0.01…) for varying levels of lambda, Rcrit and the CV of the simulated observations. Estimates in this table assume **two** indices of abundance are available. The color formatting in this table is consistent for all tables 4,5, 6, and 7.

lambda	Rcrit_true	CV	P0.005	P0.01	P0.025	P0.05	P0.1	P0.15	P0.2	P0.25
0.1	2.014	0.1	1	1	1		1	1	1	1
0.1	2.014	0.15	1	1	1		1	1	1	1
0.1	2.014	0.2	0.989	0.995	1		1	1	1	1
0.1	2.014	0.25	0.906	0.951	0.988		0.999	0.999	0.999	1
0.1	2.014	0.3	0.755	0.839	0.918		0.981	0.992	0.996	0.999
0.1	2.014	0.35	0.604	0.725	0.855		0.972	0.982	0.991	0.994
0.1	2.014	0.4	0.461	0.59	0.761		0.931	0.961	0.97	0.982
0.1	2.014	0.45	0.324	0.454	0.63		0.861	0.911	0.943	0.963
0.1	2.014	0.5	0.247	0.356	0.547		0.828	0.888	0.928	0.954
0.1	2.014	0.6	0.141	0.213	0.373		0.701	0.801	0.844	0.876
0.1	2.014	0.65	0.119	0.206	0.346		0.652	0.733	0.801	0.839
0.1	2.014	0.7	0.102	0.159	0.279		0.613	0.703	0.769	0.815
0.05	1.419	0.1	0.991	0.998	1		1	1	1	1
0.05	1.419	0.15	0.819	0.88	0.948		0.991	1	1	1
0.05	1.419	0.2	0.512	0.639	0.791		0.947	0.972	0.981	0.99
0.05	1.419	0.25	0.302	0.419	0.61		0.842	0.896	0.921	0.938
0.05	1.419	0.3	0.208	0.295	0.466		0.751	0.815	0.874	0.914
0.05	1.419	0.35	0.126	0.179	0.315		0.641	0.723	0.787	0.829
0.05	1.419	0.4	0.11	0.165	0.269		0.562	0.673	0.74	0.798
0.05	1.419	0.45	0.07	0.108	0.21		0.486	0.573	0.636	0.696
0.05	1.419	0.5	0.057	0.101	0.175		0.437	0.544	0.63	0.7
0.05	1.419	0.6	0.039	0.06	0.132		0.36	0.468	0.542	0.617
0.05	1.419	0.65	0.031	0.059	0.118		0.356	0.457	0.544	0.615
0.05	1.419	0.7	0.034	0.062	0.119		0.343	0.434	0.513	0.588
0.025	1.191	0.1	0.513	0.621	0.787		0.945	0.97	0.987	0.993
0.025	1.191	0.15	0.208	0.295	0.476		0.748	0.818	0.867	0.899
0.025	1.191	0.2	0.103	0.162	0.289		0.559	0.658	0.722	0.777
0.025	1.191	0.25	0.053	0.109	0.197		0.479	0.586	0.672	0.728
0.025	1.191	0.3	0.035	0.057	0.15		0.372	0.477	0.576	0.648
0.025	1.191	0.35	0.031	0.054	0.128		0.351	0.447	0.515	0.582
0.025	1.191	0.4	0.039	0.059	0.109		0.264	0.369	0.442	0.514
0.025	1.191	0.45	0.023	0.037	0.079		0.265	0.362	0.445	0.511
0.025	1.191	0.5	0.019	0.037	0.076		0.226	0.317	0.394	0.453
0.025	1.191	0.6	0.015	0.038	0.068		0.212	0.314	0.391	0.452
0.025	1.191	0.65	0.011	0.018	0.051		0.197	0.272	0.344	0.415
0.025	1.191	0.7	0.009	0.018	0.047		0.17	0.259	0.331	0.397

Table 6. Summary of fraction of simulations with significance probabilities less than or equal to the column headers (P0.005, P0.01…) for varying levels of lambda, Rcrit and the CV of the simulated observations. Estimates in this table assume **three** indices of abundance are available. The color formatting in this table is consistent for all tables 4,5, 6, and 7.

lambda	Rcrit_true	CV	P0.005	P0.01	P0.025	P0.05	P0.1	P0.15	P0.2	P0.25
0.1	2.014	0.1	1	1	1	1	1	1	1	1
0.1	2.014	0.15	1	1	1	1	1	1	1	1
0.1	2.014	0.2	1	1	1	1	1	1	1	1
0.1	2.014	0.25	0.996	0.999	1	1	1	1	1	1
0.1	2.014	0.3	0.972	0.984	0.996	0.998	1	1	1	1
0.1	2.014	0.35	0.891	0.927	0.972	0.996	0.999	0.999	0.999	0.999
0.1	2.014	0.4	0.807	0.884	0.945	0.977	0.996	0.998	0.999	1
0.1	2.014	0.45	0.691	0.783	0.884	0.932	0.967	0.985	0.992	0.998
0.1	2.014	0.5	0.596	0.695	0.828	0.901	0.95	0.971	0.983	0.987
0.1	2.014	0.6	0.408	0.522	0.675	0.794	0.882	0.928	0.954	0.969
0.1	2.014	0.65	0.371	0.479	0.642	0.751	0.846	0.892	0.917	0.946
0.1	2.014	0.7	0.292	0.387	0.557	0.674	0.799	0.849	0.887	0.914
0.05	1.419	0.1	1	1	1	1	1	1	1	1
0.05	1.419	0.15	0.971	0.987	0.994	0.999	1	1	1	1
0.05	1.419	0.2	0.771	0.856	0.929	0.972	0.993	0.998	0.999	1
0.05	1.419	0.25	0.566	0.666	0.811	0.897	0.953	0.975	0.984	0.99
0.05	1.419	0.3	0.383	0.487	0.634	0.756	0.862	0.913	0.942	0.964
0.05	1.419	0.35	0.268	0.367	0.527	0.648	0.764	0.835	0.882	0.918
0.05	1.419	0.4	0.221	0.31	0.455	0.572	0.696	0.787	0.843	0.881
0.05	1.419	0.45	0.155	0.233	0.364	0.5	0.637	0.721	0.778	0.83
0.05	1.419	0.5	0.118	0.184	0.321	0.458	0.576	0.673	0.744	0.792
0.05	1.419	0.6	0.071	0.126	0.223	0.332	0.467	0.563	0.648	0.703
0.05	1.419	0.65	0.071	0.121	0.216	0.309	0.435	0.544	0.625	0.699
0.05	1.419	0.7	0.055	0.092	0.157	0.268	0.389	0.489	0.584	0.651
0.025	1.191	0.1	0.796	0.885	0.951	0.98	0.992	0.997	0.998	1
0.025	1.191	0.15	0.354	0.485	0.645	0.784	0.875	0.921	0.948	0.961
0.025	1.191	0.2	0.19	0.279	0.433	0.568	0.718	0.798	0.853	0.881
0.025	1.191	0.25	0.119	0.192	0.327	0.442	0.589	0.684	0.757	0.816
0.025	1.191	0.3	0.083	0.121	0.211	0.319	0.457	0.564	0.641	0.701
0.025	1.191	0.35	0.059	0.098	0.195	0.293	0.451	0.54	0.61	0.665
0.025	1.191	0.4	0.039	0.065	0.139	0.224	0.358	0.465	0.544	0.605
0.025	1.191	0.45	0.036	0.058	0.118	0.202	0.339	0.441	0.524	0.59
0.025	1.191	0.5	0.035	0.063	0.118	0.182	0.288	0.389	0.463	0.517
0.025	1.191	0.6	0.022	0.038	0.076	0.141	0.242	0.328	0.413	0.483
0.025	1.191	0.65	0.031	0.041	0.083	0.142	0.243	0.322	0.391	0.446
0.025	1.191	0.7	0.029	0.045	0.084	0.166	0.266	0.342	0.41	0.482

Table 7. Summary of fraction of simulations with significance probabilities less than or equal to the column headers (P0.005, P0.01…) for varying levels of lambda, Rcrit and the CV of the simulated observations. Estimates in this table assume **five** indices of abundance are available. The color formatting in this table is consistent for all tables 4,5, 6, and 7.

lambda	Rcrit_true	CV	P0.005	P0.01	P0.025	P0.05	P0.1	P0.15	P0.2	P0.25
0.1	2.014	0.1	1	1	1	1	1	1	1	1
0.1	2.014	0.15	1	1	1	1	1	1	1	1
0.1	2.014	0.2	1	1	1	1	1	1	1	1
0.1	2.014	0.25	1	1	1	1	1	1	1	1
0.1	2.014	0.3	1	1	1	1	1	1	1	1
0.1	2.014	0.35	0.996	0.998	1	1	1	1	1	1
0.1	2.014	0.4	0.981	0.991	0.996	0.999	0.999	1	1	1
0.1	2.014	0.45	0.93	0.968	0.986	0.996	1	1	1	1
0.1	2.014	0.5	0.849	0.912	0.956	0.977	0.993	0.998	0.998	0.999
0.1	2.014	0.6	0.709	0.81	0.889	0.934	0.962	0.983	0.991	0.994
0.1	2.014	0.65	0.598	0.697	0.821	0.9	0.956	0.973	0.984	0.989
0.1	2.014	0.7	0.541	0.657	0.797	0.863	0.922	0.958	0.974	0.987
0.05	1.419	0.1	1	1	1	1	1	1	1	1
0.05	1.419	0.15	1	1	1	1	1	1	1	1
0.05	1.419	0.2	0.973	0.992	0.999	0.999	1	1	1	1
0.05	1.419	0.25	0.869	0.921	0.965	0.982	0.994	0.998	0.999	1
0.05	1.419	0.3	0.647	0.747	0.856	0.931	0.976	0.986	0.993	0.997
0.05	1.419	0.35	0.53	0.627	0.761	0.85	0.929	0.953	0.971	0.977
0.05	1.419	0.4	0.382	0.497	0.647	0.765	0.862	0.912	0.943	0.967
0.05	1.419	0.45	0.3	0.413	0.554	0.674	0.819	0.863	0.898	0.927
0.05	1.419	0.5	0.247	0.321	0.473	0.599	0.749	0.818	0.87	0.899
0.05	1.419	0.6	0.151	0.221	0.356	0.496	0.64	0.722	0.775	0.824
0.05	1.419	0.65	0.132	0.199	0.322	0.447	0.591	0.674	0.738	0.782
0.05	1.419	0.7	0.101	0.159	0.262	0.354	0.51	0.599	0.682	0.741
0.025	1.191	0.1	0.964	0.981	0.995	1	1	1	1	1
0.025	1.191	0.15	0.664	0.767	0.869	0.938	0.971	0.988	0.994	0.997
0.025	1.191	0.2	0.381	0.482	0.632	0.745	0.855	0.909	0.936	0.961
0.025	1.191	0.25	0.22	0.308	0.452	0.579	0.719	0.801	0.857	0.893
0.025	1.191	0.3	0.121	0.18	0.314	0.453	0.608	0.699	0.766	0.809
0.025	1.191	0.35	0.103	0.161	0.253	0.363	0.498	0.619	0.693	0.749
0.025	1.191	0.4	0.079	0.14	0.236	0.334	0.481	0.572	0.638	0.692
0.025	1.191	0.45	0.061	0.098	0.18	0.268	0.396	0.498	0.566	0.629
0.025	1.191	0.5	0.055	0.085	0.164	0.263	0.385	0.468	0.551	0.617
0.025	1.191	0.6	0.031	0.056	0.126	0.194	0.323	0.416	0.489	0.554
0.025	1.191	0.65	0.036	0.054	0.119	0.197	0.295	0.379	0.45	0.529
0.025	1.191	0.7	0.035	0.068	0.115	0.181	0.281	0.356	0.423	0.484

Table 8. Summary of Rcrit and Probability values for all possible models based on six candidate indices of relative abundance. Simulations are based US data from 2002-2016. 50,000 replicates were used for each model.

50000 replicates
ratio 2014/2016 to 2002-2004
USA Data (2002-2016)

Model #	Nvars	Combination	Rcrit	Pvalue	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
1	6	1	3.231	0.0000	ME_sprB	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB
2	5	1	3.216	0.0000	ME_sprB	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	
3	5	2	2.436	0.0000	ME_sprB	ME_falB	LLcpueStd	DK_trawl	FallSurvB	
4	5	3	3.196	0.0000	ME_sprB	ME_falB	LLcpueStd	DK_gillnet	FallSurvB	
5	5	4	4.254	0.0000	ME_sprB	ME_falB	DK_trawl	DK_gillnet	FallSurvB	
6	5	5	3.242	0.0000	ME_sprB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB	
7	5	6	3.327	0.0000	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB	
8	4	1	2.184	0.0011	ME_sprB	ME_falB	LLcpueStd	DK_trawl		
9	4	2	3.166	0.0001	ME_sprB	ME_falB	LLcpueStd	DK_gillnet		
10	4	3	2.253	0.0006	ME_sprB	ME_falB	LLcpueStd	FallSurvB		
11	4	4	4.698	0.0000	ME_sprB	ME_falB	DK_trawl	DK_gillnet		
12	4	5	3.140	0.0000	ME_sprB	ME_falB	DK_trawl	FallSurvB		
13	4	6	4.471	0.0000	ME_sprB	ME_falB	DK_gillnet	FallSurvB		
14	4	7	3.228	0.0000	ME_sprB	LLcpueStd	DK_trawl	DK_gillnet		
15	4	8	2.354	0.0003	ME_sprB	LLcpueStd	DK_trawl	FallSurvB		
16	4	9	3.205	0.0000	ME_sprB	LLcpueStd	DK_gillnet	FallSurvB		
17	4	10	4.447	0.0000	ME_sprB	DK_trawl	DK_gillnet	FallSurvB		
18	4	11	3.339	0.0001	ME_falB	LLcpueStd	DK_trawl	DK_gillnet		
19	4	12	2.418	0.0007	ME_falB	LLcpueStd	DK_trawl	FallSurvB		
20	4	13	3.305	0.0001	ME_falB	LLcpueStd	DK_gillnet	FallSurvB		
21	4	14	4.649	0.0000	ME_falB	DK_trawl	DK_gillnet	FallSurvB		
22	4	15	3.352	0.0001	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB		
23	3	1	1.871	0.0122	ME_sprB	ME_falB	LLcpueStd			
24	3	2	3.053	0.0003	ME_sprB	ME_falB	DK_trawl			
25	3	3	5.259	0.0000	ME_sprB	ME_falB	DK_gillnet			
26	3	4	3.040	0.0002	ME_sprB	ME_falB	FallSurvB			
27	3	5	2.033	0.0037	ME_sprB	LLcpueStd	DK_trawl			
28	3	6	3.173	0.0004	ME_sprB	LLcpueStd	DK_gillnet			
29	3	7	2.128	0.0024	ME_sprB	LLcpueStd	FallSurvB			
30	3	8	5.125	0.0000	ME_sprB	DK_trawl	DK_gillnet			
31	3	9	3.142	0.0001	ME_sprB	DK_trawl	FallSurvB			
32	3	10	4.778	0.0000	ME_sprB	DK_gillnet	FallSurvB			
33	3	11	2.106	0.0091	ME_falB	LLcpueStd	DK_trawl			
34	3	12	3.310	0.0009	ME_falB	LLcpueStd	DK_gillnet			
35	3	13	2.196	0.0051	ME_falB	LLcpueStd	FallSurvB			
36	3	14	5.511	0.0000	ME_falB	DK_trawl	DK_gillnet			
37	3	15	3.305	0.0004	ME_falB	DK_trawl	FallSurvB			
38	3	16	5.074	0.0000	ME_falB	DK_gillnet	FallSurvB			
39	3	17	3.374	0.0003	LLcpueStd	DK_trawl	DK_gillnet			
40	3	18	2.319	0.0025	LLcpueStd	DK_trawl	FallSurvB			
41	3	19	3.331	0.0005	LLcpueStd	DK_gillnet	FallSurvB			
42	3	20	4.984	0.0000	DK_trawl	DK_gillnet	FallSurvB			
43	2	1	2.803	0.0042	ME_sprB	ME_falB				
44	2	2	1.611	0.0353	ME_sprB	LLcpueStd				
45	2	3	3.025	0.0014	ME_sprB	DK_trawl				
46	2	4	6.216	0.0000	ME_sprB	DK_gillnet				
47	2	5	3.014	0.0016	ME_sprB	FallSurvB				
48	2	6	1.680	0.0792	ME_falB	LLcpueStd				
49	2	7	3.317	0.0041	ME_falB	DK_trawl				
50	2	8	7.050	0.0003	ME_falB	DK_gillnet				
51	2	9	3.240	0.0045	ME_falB	FallSurvB				
52	2	10	1.901	0.0276	LLcpueStd	DK_trawl				
53	2	11	3.351	0.0046	LLcpueStd	DK_gillnet				
54	2	12	2.033	0.0180	LLcpueStd	FallSurvB				
55	2	13	6.509	0.0003	DK_trawl	DK_gillnet				
56	2	14	3.354	0.0028	DK_trawl	FallSurvB				
57	2	15	5.703	0.0009	DK_gillnet	FallSurvB				
58	1	1	2.550	0.0205	ME_sprB					
59	1	2	3.129	0.0520	ME_falB					
60	1	3	1.274	0.2200	LLcpueStd					
61	1	4	3.447	0.0256	DK_trawl					
62	1	5	11.217	0.0131	DK_gillnet					
63	1	6	3.291	0.0267	FallSurvB					

Average Rcrit value overall models=
3.522825
fraction of models with significance probability <0.05
0.952381

Table 9. Summary of Rcrit and Probability values for all possible models based on six candidate indices of relative abundance. Simulations are based US data from 2005-2016. 50,000 replicates were used for each model.

	Ratio 2014-2016: 2005-2007									
US data 2005-2016 50000 reps										
Model #	Nvars	Combination	Rcrit	Pvalue	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
1	6	1	2.202	0.0002	ME_sprB	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB
2	5	1	2.136	0.0004	ME_sprB	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	
3	5	2	1.644	0.0082	ME_sprB	ME_falB	LLcpueStd	DK_trawl	FallSurvB	
4	5	3	2.090	0.0008	ME_sprB	ME_falB	LLcpueStd	DK_gillnet	FallSurvB	
5	5	4	2.526	0.0001	ME_sprB	ME_falB	DK_trawl	DK_gillnet	FallSurvB	
6	5	5	2.314	0.0002	ME_sprB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB	
7	5	6	2.658	0.0000	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB	
8	4	1	1.428	0.0494	ME_sprB	ME_falB	LLcpueStd	DK_trawl		
9	4	2	1.985	0.0045	ME_sprB	ME_falB	LLcpueStd	DK_gillnet		
10	4	3	1.457	0.0498	ME_sprB	ME_falB	LLcpueStd	FallSurvB		
11	4	4	2.542	0.0006	ME_sprB	ME_falB	DK_trawl	DK_gillnet		
12	4	5	1.782	0.0115	ME_sprB	ME_falB	DK_trawl	FallSurvB		
13	4	6	2.419	0.0013	ME_sprB	ME_falB	DK_gillnet	FallSurvB		
14	4	7	2.267	0.0008	ME_sprB	LLcpueStd	DK_trawl	DK_gillnet		
15	4	8	1.672	0.0098	ME_sprB	LLcpueStd	DK_trawl	FallSurvB		
16	4	9	2.197	0.0012	ME_sprB	LLcpueStd	DK_gillnet	FallSurvB		
17	4	10	2.756	0.0002	ME_sprB	DK_trawl	DK_gillnet	FallSurvB		
18	4	11	2.720	0.0001	ME_falB	LLcpueStd	DK_trawl	DK_gillnet		
19	4	12	1.938	0.0036	ME_falB	LLcpueStd	DK_trawl	FallSurvB		
20	4	13	2.573	0.0004	ME_falB	LLcpueStd	DK_gillnet	FallSurvB		
21	4	14	3.390	0.0001	ME_falB	DK_trawl	DK_gillnet	FallSurvB		
22	4	15	2.913	0.0001	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB		
23	3	1	1.151	0.2737	ME_sprB	ME_falB	LLcpueStd			
24	3	2	1.505	0.0704	ME_sprB	ME_falB	DK_trawl			
25	3	3	2.393	0.0060	ME_sprB	ME_falB	DK_gillnet			
26	3	4	1.534	0.0719	ME_sprB	ME_falB	FallSurvB			
27	3	5	1.418	0.0568	ME_sprB	LLcpueStd	DK_trawl			
28	3	6	2.103	0.0056	ME_sprB	LLcpueStd	DK_gillnet			
29	3	7	1.453	0.0559	ME_sprB	LLcpueStd	FallSurvB			
30	3	8	2.886	0.0008	ME_sprB	DK_trawl	DK_gillnet			
31	3	9	1.857	0.0129	ME_sprB	DK_trawl	FallSurvB			
32	3	10	2.667	0.0019	ME_sprB	DK_gillnet	FallSurvB			
33	3	11	1.727	0.0203	ME_falB	LLcpueStd	DK_trawl			
34	3	12	2.613	0.0017	ME_falB	LLcpueStd	DK_gillnet			
35	3	13	1.721	0.0264	ME_falB	LLcpueStd	FallSurvB			
36	3	14	3.982	0.0001	ME_falB	DK_trawl	DK_gillnet			
37	3	15	2.337	0.0042	ME_falB	DK_trawl	FallSurvB			
38	3	16	3.437	0.0006	ME_falB	DK_gillnet	FallSurvB			
39	3	17	3.110	0.0002	LLcpueStd	DK_trawl	DK_gillnet			
40	3	18	2.049	0.0041	LLcpueStd	DK_trawl	FallSurvB			
41	3	19	2.856	0.0005	LLcpueStd	DK_gillnet	FallSurvB			
42	3	20	4.106	0.0000	DK_trawl	DK_gillnet	FallSurvB			
43	2	1	1.059	0.4341	ME_sprB	ME_falB				
44	2	2	1.067	0.3697	ME_sprB	LLcpueStd				
45	2	3	1.515	0.0945	ME_sprB	DK_trawl				
46	2	4	2.784	0.0097	ME_sprB	DK_gillnet				
47	2	5	1.551	0.1020	ME_sprB	FallSurvB				
48	2	6	1.351	0.1722	ME_falB	LLcpueStd				
49	2	7	2.220	0.0219	ME_falB	DK_trawl				
50	2	8	4.381	0.0015	ME_falB	DK_gillnet				
51	2	9	2.069	0.0399	ME_falB	FallSurvB				
52	2	10	1.820	0.0204	LLcpueStd	DK_trawl				
53	2	11	3.080	0.0023	LLcpueStd	DK_gillnet				
54	2	12	1.793	0.0306	LLcpueStd	FallSurvB				
55	2	13	6.028	0.0002	DK_trawl	DK_gillnet				
56	2	14	2.727	0.0057	DK_trawl	FallSurvB				
57	2	15	4.461	0.0016	DK_gillnet	FallSurvB				
58	1	1	0.837	0.7097	ME_sprB					
59	1	2	1.477	0.2675	ME_falB					
60	1	3	1.287	0.1577	LLcpueStd					
61	1	4	3.205	0.0215	DK_trawl					
62	1	5	10.559	0.0065	DK_gillnet					
63	1	6	2.478	0.0629	FallSurvB					
		[1]	Average Rcrit value overall models=							
		[1]	2.448565							
		[1]	fraction of models with significance probability <0.05							
		[1]	0.777778							

Table 10. Summary of Rcrit and probability values for all possible models based on six candidate indices of relative abundance. Simulations are based US data from 2002-2013. 50,000 replicates were used for each model. The reduced number of years was used to allow comparison with results from Canada.

50000 replicates										
ratio 2011/2013 to 2002-2004										
USA Data (2002-2013)										
Model #	Nvars	Combinati	Rcrit	Pvalue	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
1	6	1	2.893	0.0000	ME_sprB	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB
2	5	1	3.360	0.0000	ME_sprB	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	
3	5	2	2.556	0.0000	ME_sprB	ME_falB	LLcpueStd	DK_trawl	FallSurvB	
4	5	3	2.539	0.0001	ME_sprB	ME_falB	LLcpueStd	DK_gillnet	FallSurvB	
5	5	4	3.686	0.0000	ME_sprB	ME_falB	DK_trawl	DK_gillnet	FallSurvB	
6	5	5	2.629	0.0001	ME_sprB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB	
7	5	6	2.820	0.0001	ME_falB	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB	
8	4	1	3.008	0.0000	ME_sprB	ME_falB	LLcpueStd	DK_trawl		
9	4	2	2.973	0.0001	ME_sprB	ME_falB	LLcpueStd	DK_gillnet		
10	4	3	2.068	0.0011	ME_sprB	ME_falB	LLcpueStd	FallSurvB		
11	4	4	4.865	0.0000	ME_sprB	ME_falB	DK_trawl	DK_gillnet		
12	4	5	3.372	0.0000	ME_sprB	ME_falB	DK_trawl	FallSurvB		
13	4	6	3.320	0.0000	ME_sprB	ME_falB	DK_gillnet	FallSurvB		
14	4	7	3.060	0.0000	ME_sprB	LLcpueStd	DK_trawl	DK_gillnet		
15	4	8	2.207	0.0006	ME_sprB	LLcpueStd	DK_trawl	FallSurvB		
16	4	9	2.196	0.0017	ME_sprB	LLcpueStd	DK_gillnet	FallSurvB		
17	4	10	3.389	0.0000	ME_sprB	DK_trawl	DK_gillnet	FallSurvB		
18	4	11	3.332	0.0001	ME_falB	LLcpueStd	DK_trawl	DK_gillnet		
19	4	12	2.422	0.0007	ME_falB	LLcpueStd	DK_trawl	FallSurvB		
20	4	13	2.406	0.0014	ME_falB	LLcpueStd	DK_gillnet	FallSurvB		
21	4	14	3.710	0.0000	ME_falB	DK_trawl	DK_gillnet	FallSurvB		
22	4	15	2.515	0.0007	LLcpueStd	DK_trawl	DK_gillnet	FallSurvB		
23	3	1	2.411	0.0006	ME_sprB	ME_falB	LLcpueStd			
24	3	2	4.880	0.0000	ME_sprB	ME_falB	DK_trawl			
25	3	3	4.709	0.0000	ME_sprB	ME_falB	DK_gillnet			
26	3	4	2.775	0.0003	ME_sprB	ME_falB	FallSurvB			
27	3	5	2.576	0.0006	ME_sprB	LLcpueStd	DK_trawl			
28	3	6	2.549	0.0017	ME_sprB	LLcpueStd	DK_gillnet			
29	3	7	1.589	0.0278	ME_sprB	LLcpueStd	FallSurvB			
30	3	8	4.633	0.0000	ME_sprB	DK_trawl	DK_gillnet			
31	3	9	2.927	0.0003	ME_sprB	DK_trawl	FallSurvB			
32	3	10	2.883	0.0011	ME_sprB	DK_gillnet	FallSurvB			
33	3	11	2.903	0.0004	ME_falB	LLcpueStd	DK_trawl			
34	3	12	2.864	0.0016	ME_falB	LLcpueStd	DK_gillnet			
35	3	13	1.825	0.0180	ME_falB	LLcpueStd	FallSurvB			
36	3	14	5.243	0.0000	ME_falB	DK_trawl	DK_gillnet			
37	3	15	3.332	0.0002	ME_falB	DK_trawl	FallSurvB			
38	3	16	3.270	0.0012	ME_falB	DK_gillnet	FallSurvB			
39	3	17	2.976	0.0011	LLcpueStd	DK_trawl	DK_gillnet			
40	3	18	2.003	0.0086	LLcpueStd	DK_trawl	FallSurvB			
41	3	19	1.997	0.0170	LLcpueStd	DK_gillnet	FallSurvB			
42	3	20	3.359	0.0007	DK_trawl	DK_gillnet	FallSurvB			
43	2	1	4.607	0.0001	ME_sprB	ME_falB				
44	2	2	1.728	0.0306	ME_sprB	LLcpueStd				
45	2	3	4.505	0.0000	ME_sprB	DK_trawl				
46	2	4	4.307	0.0009	ME_sprB	DK_gillnet				
47	2	5	2.015	0.0176	ME_sprB	FallSurvB				
48	2	6	2.114	0.0194	ME_falB	LLcpueStd				
49	2	7	5.533	0.0000	ME_falB	DK_trawl				
50	2	8	5.234	0.0008	ME_falB	DK_gillnet				
51	2	9	2.521	0.0108	ME_falB	FallSurvB				
52	2	10	2.352	0.0107	LLcpueStd	DK_trawl				
53	2	11	2.328	0.0253	LLcpueStd	DK_gillnet				
54	2	12	1.220	0.2429	LLcpueStd	FallSurvB				
55	2	13	5.033	0.0011	DK_trawl	DK_gillnet				
56	2	14	2.753	0.0068	DK_trawl	FallSurvB				
57	2	15	2.706	0.0173	DK_gillnet	FallSurvB				
58	1	1	3.543	0.0070	ME_sprB					
59	1	2	5.986	0.0031	ME_falB					
60	1	3	1.131	0.3690	LLcpueStd					
61	1	4	5.259	0.0009	DK_trawl					
62	1	5	4.834	0.0411	DK_gillnet					
63	1	6	1.339	0.2620	FallSurvB					
		[1]	Average Rcrit value overall models=							
		[1]	3.144086							
		[1]	fraction of models with significance probability <0.05							
		[1]	0.952381							

Table 11. Summary of Rcrit and Probability values for all possible models based on three indices of relative abundance and the assessment based estimate of absolute abundance in the Canadian 3NOPs4VWX5Zc assessment. Simulations are based Canadian data from 2002-2013. 50,000 replicates were used for each model. The shaded boxes indicate models that include the modeled biomass estimates from the DFO assessment.   Results can be compared with Table 10 for US indices.

Canadian Data								
Model #	Nvars	ombinatio	Rcrit	Pvalue	Var 1	Var 2	Var 3	Var 4
1	4	1	2.719	0.00000	Can.RV.Summer	Can.CRV.Spr	Can.CPUE	Can.SSB.Mod
2	3	1	2.703	0.00000	Can.RV.Summer	Can.CRV.Spr	Can.CPUE	
3	3	2	3.476	0.00000	Can.RV.Summer	Can.CRV.Spr	Can.SSB.Mod	
4	3	3	2.317	0.00000	Can.RV.Summer	Can.CPUE	Can.SSB.Mod	
5	3	4	2.532	0.00000	Can.CRV.Spr	Can.CPUE	Can.SSB.Mod	
6	2	1	3.967	0.00002	Can.RV.Summer	Can.CRV.Spr		
7	2	2	2.101	0.00004	Can.RV.Summer	Can.CPUE		
8	2	3	3.079	0.00000	Can.RV.Summer	Can.SSB.Mod		
9	2	4	2.420	0.00040	Can.CRV.Spr	Can.CPUE		
10	2	5	3.458	0.00004	Can.CRV.Spr	Can.SSB.Mod		
11	2	6	1.948	0.00000	Can.CPUE	Can.SSB.Mod		
12	1	1	3.519	0.00026	Can.RV.Summer			
13	1	2	4.410	0.01296	Can.CRV.Spr			
14	1	3	1.344	0.01606	Can.CPUE			
15	1	4	2.763	0.00000	Can.SSB.Mod			
	[1]	Average Rcrit value overall models=						
	[1]	2.850295						
	[1]	fraction of models with significance probability <0.05						
	[1]	1						
		Rcrit average for models that do NOT include Can.SSB.Mod						
		2.923448						

Table 12. Derived estimates for the K parameter and 2016 biomass using the Envelope method for various ranges of catch data and smoothed vs raw NEFSC fall bottom trawl indices.

Survey Type	Basis for Estimating Max Biomass	Max Catch	Biomass Estimates			Mid Range Biomass 2016	Estimated B(2016)/K		
			max B(C,F)	min B(C,F)	Mid Range B(C,F)		max B(C,F)	min B(C,F)	Mid Range B(C,F)
Kalman	Max catch 1893-2016	4,908	266,850	15,952	141,401	3,363.2	0.013	0.211	0.024
Kalman	Max Catch since 1900	944	51,326	3,068	27,197	3,363.2	0.066	1.096	0.124
Kalman	Constrained range of B(1963-2016)		6,531	1,671	3,425	3,363.2	0.515	2.013	0.982
Raw	Max catch 1893-2016	4,908	266,850	15,952	141,401	3,407.4	0.013	0.214	0.024
Raw	Max Catch since 1900	944	51,326	3,068	27,197	3,407.4	0.066	1.111	0.125
Raw	Constrained range of B(1963-2016)	514	14,680	1,671	7,779	3,407.4	0.232	2.040	0.438

Table 13.

	Table hh. A. Summary of maximum fractional change in population abundance given alternative ranges of proportial stock increase for varying base period year ranges.											
B. Derived Depletion corrected average catches of sustainable harvest alternative levels of rebuilding in 2016 and proportional increase in relative abundance.												
Levels of rebuilding are based on envelope method. Natural mortality is assumed =							0.15				

A

							Maximum Fractional Change (DELTA) in DCAC for varying assumed values of B(t)/K				
	Changes in catches			**Change in indices**			Assume 98.2% rebuilt in 2016	Assume 43.8% rebuilt in 2016	Assume 12.5% rebuilt in 2016	Assume 2.4% rebuilt in 2016	
Ratio Definition	*Statistics*	*Value*		*Statistic*	*Value*	*Model*	*0.982*	*0.438*	*0.125*	*0.024*	*Total Catch*
'02-04:'14-16	Rcrit(Catch)	3.227		Rcrit(Indices)	3.23	(all six indices)	-0.67798	-0.30240	-0.08630	-0.01657	925.3
					4.98	(DK_g, DK_t, Survey)	-0.78481	-0.35005	-0.09990	-0.01918	
					3.52	average over 120models	-0.70302	-0.31357	-0.08949	-0.01718	
'05-07:'14-16	Rcrit(Catch)	2.657		Rcrit(Indices)	2.44	average over 120models	-0.57954	-0.25849	-0.07377	-0.01416	831.4
					4.11	DK_g,DK_t, Survey	-0.74307	-0.33143	-0.09459	-0.01816	
					2.2	(all six indices)	-0.53564	-0.23891	-0.06818	-0.01309	
'02-04:'11-13	Rcrit(Catch)	2.617		Rcrit(indices)	2.893	(all six indices)	-0.64256	-0.28660	-0.08179	-0.01570	622.1
					5.033	(DK_g, DK_t, Survey)	-0.78689	-0.35097	-0.10016	-0.01923	
					3.144	average over 120models	-0.66966	-0.29869	-0.08524	-0.01637	

B

			Derived Delta for Assumed Alpha				Derived Sustainable Average Catch for Assumed Alpha					Y Current --given assumed level of rebuilding			
Nyears	Time Period	Total Catch	0.982	0.438	0.125	0.024	0.982	0.438	0.125	0.024	Obs Ave Catch	0.982	0.438	0.125	0.024
15		925.3	-0.678	-0.302	-0.086	-0.017	-121.8	188.1	76.3	64.0	61.7	-119.6	82.4	9.5	1.5
15	2002-2016	925.3	-0.785	-0.350	-0.100	-0.019	-82.9	277.7	79.3	64.4		-81.4	121.6	9.9	1.5
15		925.3	-0.703	-0.314	-0.089	-0.017	-109.7	203.5	77.0	64.1		-107.7	89.1	9.6	1.5
12		831.4	-0.580	-0.258	-0.074	-0.014	-113.6	245.7	87.1	72.1	69.3	-111.6	107.6	10.9	1.7
12	2005-2016	831.4	-0.743	-0.331	-0.095	-0.018	-65.1	873.0	94.0	73.0		-63.9	382.4	11.7	1.8
12		831.4	-0.536	-0.239	-0.068	-0.013	-142.0	206.0	85.5	71.9		-139.4	90.2	10.7	1.7
12		622.1	-0.643	-0.287	-0.082	-0.016	-66.1	254.3	67.1	54.2	51.8	-64.9	111.4	8.4	1.3
12	2002-2013	622.1	-0.787	-0.351	-0.100	-0.019	-43.7	2067.9	71.8	54.8		-42.9	905.7	9.0	1.3
12		622.1	-0.670	-0.299	-0.085	-0.016	-60.3	304.4	67.9	54.3		-59.2	133.3	8.5	1.3

Table 14.

Table xx Summary of relevant population outputs for varying combinations of Kp and Kd gain parameters by assumed CV level for observation error. The low CV (0.005) assumes almost no observation error in the abundance indices. Effects are averaged over all combinations of r(t) scenarios and harvest scenarios (h(t)) prior to implementation of the control rule. Simulation failures occur when the population size goes to zero because harvest rates are too high.

Average % of overfishing events						**Fraction of simulation failures**					
CV	Kd	Kp=0.25	Kp=0.5	Kp=0.75	Kp=1	CV	Kd	Kp=0.25	Kp=0.5	Kp=0.75	Kp=1
0.005	0	0.171	0.201	0.229	0.282	0.005	0	0.163	0.163	0.153	0.134
	1	0.197	0.228	0.249	0.314		1	0.122	0.112	0.111	0.071
	5	0.244	0.254	0.269	0.280		5	0.000	0.000	0.000	0.000
	10	0.217	0.237	0.255	0.281		10	0.000	0.004	0.010	0.022
0.2	0	0.171	0.201	0.234	0.275	0.2	0	0.166	0.164	0.157	0.151
	1	0.196	0.231	0.271	0.309		1	0.126	0.114	0.097	0.082
	5	0.247	0.253	0.274	0.292		5	0.034	0.035	0.029	0.028
	10	0.226	0.241	0.258	0.279		10	0.147	0.162	0.182	0.194

Average Catch						**Net rate of population change during control period**					
CV	Kd	Kp=0.25	Kp=0.5	Kp=0.75	Kp=1	CV	Kd	Kp=0.25	Kp=0.5	Kp=0.75	Kp=1
0.005	0	228.6	248.9	270.6	296.5	0.005	0	0.078	0.071	0.061	0.047
	1	236.0	254.8	275.3	298.9		1	0.071	0.063	0.056	0.038
	5	256.6	276.3	297.6	320.1		5	0.058	0.054	0.049	0.043
	10	294.4	313.5	332.1	348.2		10	0.054	0.048	0.041	0.033
0.2	0	227.8	248.3	270.6	295.6	0.2	0	0.078	0.071	0.062	0.052
	1	235.0	254.5	276.3	298.8		1	0.072	0.062	0.050	0.038
	5	256.6	277.2	297.1	319.4		5	0.055	0.050	0.042	0.035
	10	292.7	314.0	332.2	346.4		10	0.050	0.044	0.036	0.027

Average CV of Catch				
CV	Kd	Kp=0.25	Kp=0.5	Kp=0.75	Kp=1
0.005	0	0.068	0.135	0.199	0.259
	1	0.097	0.159	0.219	0.283
	5	0.239	0.281	0.322	0.361
	10	0.401	0.430	0.458	0.483
0.2	0	0.070	0.137	0.203	0.266
	1	0.115	0.176	0.239	0.304
	5	0.368	0.410	0.457	0.503
	10	0.675	0.712	0.737	0.769

Table 15. Summary of derived estimates of catch in 2018 based on the FSD model for alternative values of the gain parameters Kp and Kd. The top table shows the effect of the parameters on estimated catch in mt.   The lower table illustrates the effect of the gain parameters on the degree of concordance with historical estimates of observed and predicted catch for the time period 2007 to 2016. Table entries are ratios of the sum of squares difference between observed and predicted to the minimum value.

		Kp								
	122.67	0	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
	0	110.3	114.3	115.7	117.1	118.5	120.0	121.4	122.9	124.4
	0.25	111.2	115.3	116.7	118.1	119.5	120.9	122.4	123.9	125.4
	0.5	112.1	116.2	117.6	119.0	120.5	121.9	123.4	124.9	126.4
	0.75	113.0	117.2	118.6	120.0	121.5	122.9	124.4	125.9	127.4
	1	114.0	118.1	119.6	121.0	122.5	123.9	125.4	127.0	128.5
	1.25	114.9	119.1	120.5	122.0	123.5	125.0	126.5	128.0	129.5
	1.5	115.8	120.1	121.5	123.0	124.5	126.0	127.5	129.0	130.6
	1.75	116.8	121.1	122.5	124.0	125.5	127.0	128.6	130.1	131.7
Kd	2	117.8	122.1	123.5	125.0	126.5	128.1	129.6	131.2	132.8
	2.25	118.7	123.1	124.6	126.1	127.6	129.1	130.7	132.3	133.8
	2.5	119.7	124.1	125.6	127.1	128.6	130.2	131.7	133.3	134.9
	2.75	120.7	125.1	126.6	128.1	129.7	131.2	132.8	134.4	136.1
	3	121.7	126.1	127.6	129.2	130.7	132.3	133.9	135.5	137.2
	3.25	122.7	127.2	128.7	130.2	131.8	133.4	135.0	136.6	138.3
	3.5	123.7	128.2	129.8	131.3	132.9	134.5	136.1	137.8	139.4
	3.75	124.7	129.3	130.8	132.4	134.0	135.6	137.2	138.9	140.6
	4	125.7	130.3	131.9	133.5	135.1	136.7	138.4	140.0	141.7
	min (C(2018))=			111.2	max(C(2018))=		141.7			

	Ratio of (SSQ-Min(SSQ)) to Minimum SSQ									
		Kp								
		0	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
	0	0.12	0.13	0.14	0.16	0.19	0.22	0.25	0.30	0.34
	0.25	0.00	0.04	0.07	0.10	0.14	0.19	0.24	0.30	0.37
	0.5	0.13	0.22	0.26	0.31	0.37	0.44	0.51	0.59	0.69
	0.75	0.58	0.71	0.78	0.85	0.93	1.02	1.12	1.23	1.35
	1	1.40	1.61	1.69	1.79	1.90	2.02	2.15	2.30	2.45
	1.25	2.70	2.98	3.10	3.23	3.38	3.53	3.70	3.88	4.08
	1.5	4.56	4.95	5.11	5.28	5.46	5.66	5.88	6.11	6.36
	1.75	7.13	7.64	7.84	8.06	8.29	8.55	8.82	9.11	9.42
Kd	2	10.54	11.20	11.46	11.73	12.03	12.35	12.69	13.05	13.43
	2.25	14.97	15.81	16.14	16.49	16.86	17.26	17.68	18.13	18.60
	2.5	20.63	21.70	22.11	22.55	23.02	23.51	24.03	24.58	25.16
	2.75	27.78	29.13	29.64	30.19	30.76	31.37	32.01	32.69	33.40
	3	36.71	38.40	39.04	39.71	40.42	41.17	41.96	42.79	43.65
	3.25	47.80	49.90	50.69	51.51	52.39	53.30	54.27	55.27	56.33
	3.5	61.46	64.06	65.03	66.04	67.11	68.23	69.41	70.63	71.92
	3.75	78.20	81.41	82.60	83.84	85.15	86.52	87.94	89.44	90.99
	4	98.65	102.60	104.05	105.57	107.16	108.82	110.55	112.36	114.25

.

Table 16. Estimated discards by gear type and adjustments for discard mortality rates.

| Year | Discard Estimate (mt) | | | | | Fraction by Gear | | | Adjusted for Discard Mortality Rate | | | |
| | | | | | | | | | 0.1 | 0.76 | 0.3 | |
	Handline	Trawl	Gill net	Total		Handline	Trawl	Gill net	Handline	Trawl	Gill net	Adj Total
1989	0.00	2.88	2.10	4.97		0.000	0.578	0.422	0.00	2.19	0.63	2.81
1990	0.00	12.09	1.46	13.55		0.000	0.892	0.108	0.00	9.19	0.44	9.63
1991	0.00	6.06	0.87	6.93		0.000	0.875	0.125	0.00	4.61	0.26	4.87
1992	0.00	1.92	0.27	2.19		0.000	0.878	0.122	0.00	1.46	0.08	1.54
1993	0.00	0.63	0.44	1.06		0.000	0.590	0.410	0.00	0.48	0.13	0.61
1994	0.00	2.94	0.22	3.16		0.000	0.930	0.070	0.00	2.23	0.07	2.30
1995	0.00	6.30	0.04	6.34		0.000	0.993	0.007	0.00	4.79	0.01	4.80
1996	0.00	0.52	0.14	0.65		0.000	0.791	0.209	0.00	0.39	0.04	0.43
1997	0.00	1.64	0.00	1.64		0.000	1.000	0.000	0.00	1.24	0.00	1.24
1998	0.00	0.00	0.10	0.10		0.000	0.000	1.000	0.00	0.00	0.03	0.03
1999	0.00	68.85	0.25	69.10		0.000	0.996	0.004	0.00	52.33	0.07	52.40
2000	0.00	11.38	0.49	11.87		0.000	0.958	0.042	0.00	8.65	0.15	8.80
2001	0.00	9.29	0.40	9.68		0.000	0.959	0.041	0.00	7.06	0.12	7.18
2002	0.00	20.20	0.00	20.20		0.000	1.000	0.000	0.00	15.35	0.00	15.35
2003	0.00	15.80	4.35	20.15		0.000	0.784	0.216	0.00	12.01	1.31	13.32
2004	0.02	14.81	0.88	15.71		0.001	0.943	0.056	0.00	11.26	0.26	11.52
2005	0.70	16.90	1.29	18.89		0.037	0.895	0.068	0.07	12.85	0.39	13.30
2006	0.00	19.05	3.40	22.45		0.000	0.849	0.151	0.00	14.48	1.02	15.50
2007	0.08	14.65	2.54	17.27		0.004	0.848	0.147	0.01	11.13	0.76	11.91
2008	0.00	18.87	2.79	21.66		0.000	0.871	0.129	0.00	14.34	0.84	15.18
2009	0.00	16.93	0.92	17.85		0.000	0.949	0.051	0.00	12.87	0.27	13.15
2010	2.52	27.55	4.63	34.69		0.073	0.794	0.134	0.25	20.93	1.39	22.58
2011	0.07	33.56	8.71	42.35		0.002	0.793	0.206	0.01	25.51	2.61	28.13
2012	0.00	43.51	8.68	52.19		0.000	0.834	0.166	0.00	33.07	2.60	35.67
2013	0.20	46.27	9.70	56.18		0.004	0.824	0.173	0.02	35.17	2.91	38.10
2014	0.00	23.95	10.39	34.34		0.000	0.697	0.303	0.00	18.20	3.12	21.32
2015	0.00	22.48	23.82	46.30		0.000	0.485	0.515	0.00	17.08	7.15	24.23
2016	0.00	26.00	21.40	47.40		0.000	0.549	0.451	0.00	19.76	6.42	26.18

Table 17.Estimated total catch with and without adjustment for discard mortality estimates. Rates are defined in Table 16.

Year	Total Catch (mt)	Adjusted total catch (mt)	% Change
1994	24.9	24.1	-3.4%
1995	16.9	15.3	-9.1%
1996	14.0	13.8	-1.6%
1997	15.6	15.3	-2.5%
1998	8.5	8.4	-0.9%
1999	80.6	63.9	-20.7%
2000	22.9	19.9	-13.4%
2001	20.5	18.0	-12.2%
2002	30.2	25.4	-16.0%
2003	36.8	30.0	-18.6%
2004	26.9	22.7	-15.5%
2005	35.7	30.1	-15.6%
2006	36.5	29.6	-19.0%
2007	41.9	36.5	-12.8%
2008	50.3	43.9	-12.9%
2009	62.9	58.2	-7.5%
2010	54.9	42.8	-22.1%
2011	68.1	53.9	-20.9%
2012	87.0	70.5	-19.0%
2013	90.8	72.8	-19.9%
2014	79.3	66.3	-16.4%
2015	108.3	86.2	-20.4%
2016	115.6	94.4	-18.3%

FIGURES

Year	Disards	Landings	Catch
1994	3.16	21.77	24.93
1995	6.34	10.54	16.88
1996	0.65	13.32	13.97
1997	1.64	14.01	15.65
1998	0.10	8.41	8.51
1999	69.08	11.51	80.59
2000	11.87	11.07	22.94
2001	9.68	10.82	20.50
2002	20.20	10.01	30.21
2003	20.15	16.68	36.83
2004	15.71	11.22	26.93
2005	18.89	16.81	35.70
2006	22.45	14.08	36.53
2007	17.27	24.61	41.88
2008	21.66	28.69	50.35
2009	17.85	45.05	62.90
2010	34.68	20.20	54.88
2011	42.34	25.79	68.13
2012	52.18	34.80	86.98
2013	56.16	34.67	90.83
2014	34.33	44.99	79.32
2015	46.28	62.00	108.28
2016	47.39	68.20	115.59

Figure 1. Landings and discards (mt) for Atlantic halibut in US stock area, 1994-2016, used in this assessment.
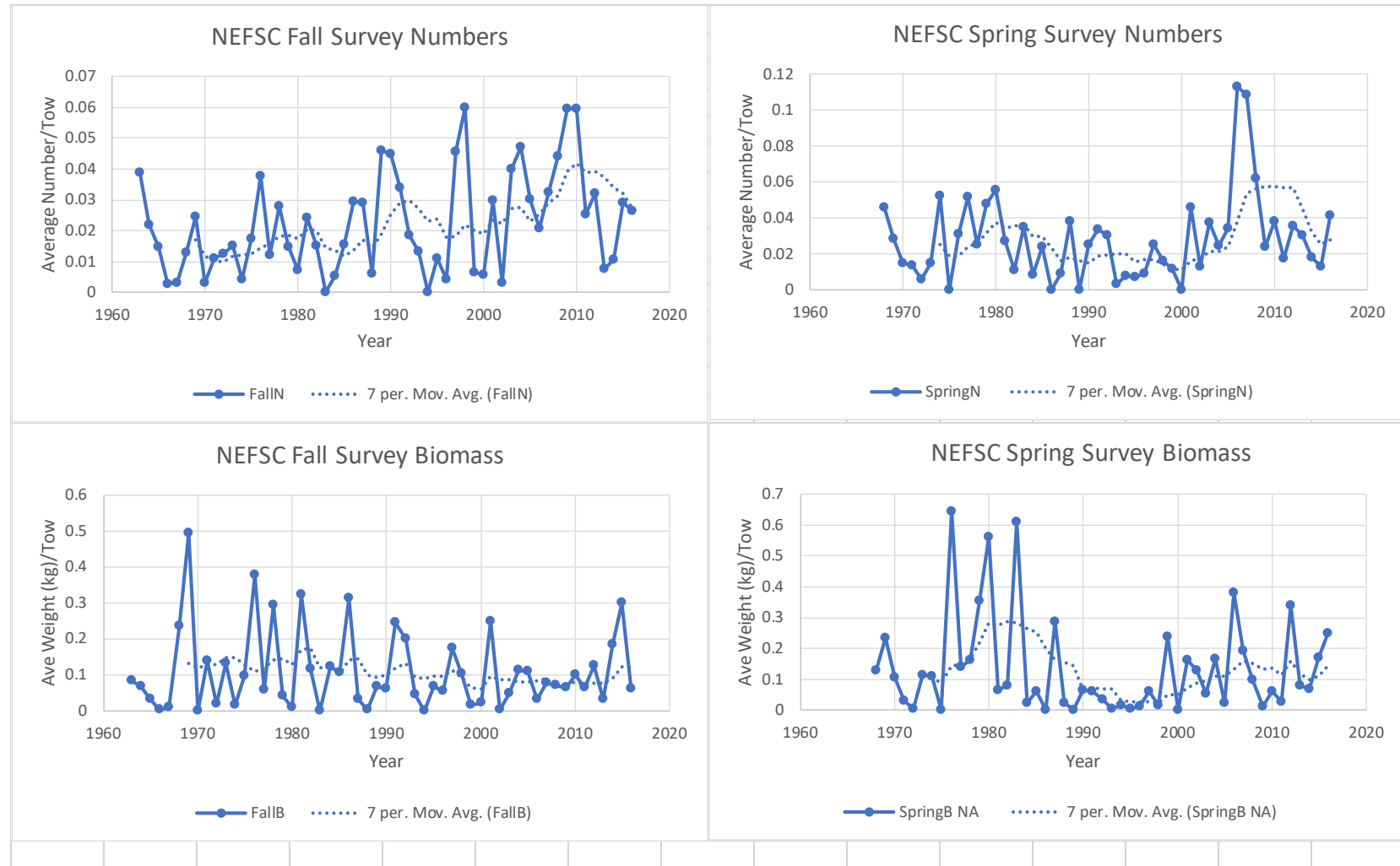
Figure 2. Summary of NEFSC bottom trawl survey indices 1963-2016 for fall and spring surveys expressed in terms of both average numbers per tow and average weight (kg) per tow. A 7 point moving average is used to dampen interannual variability.
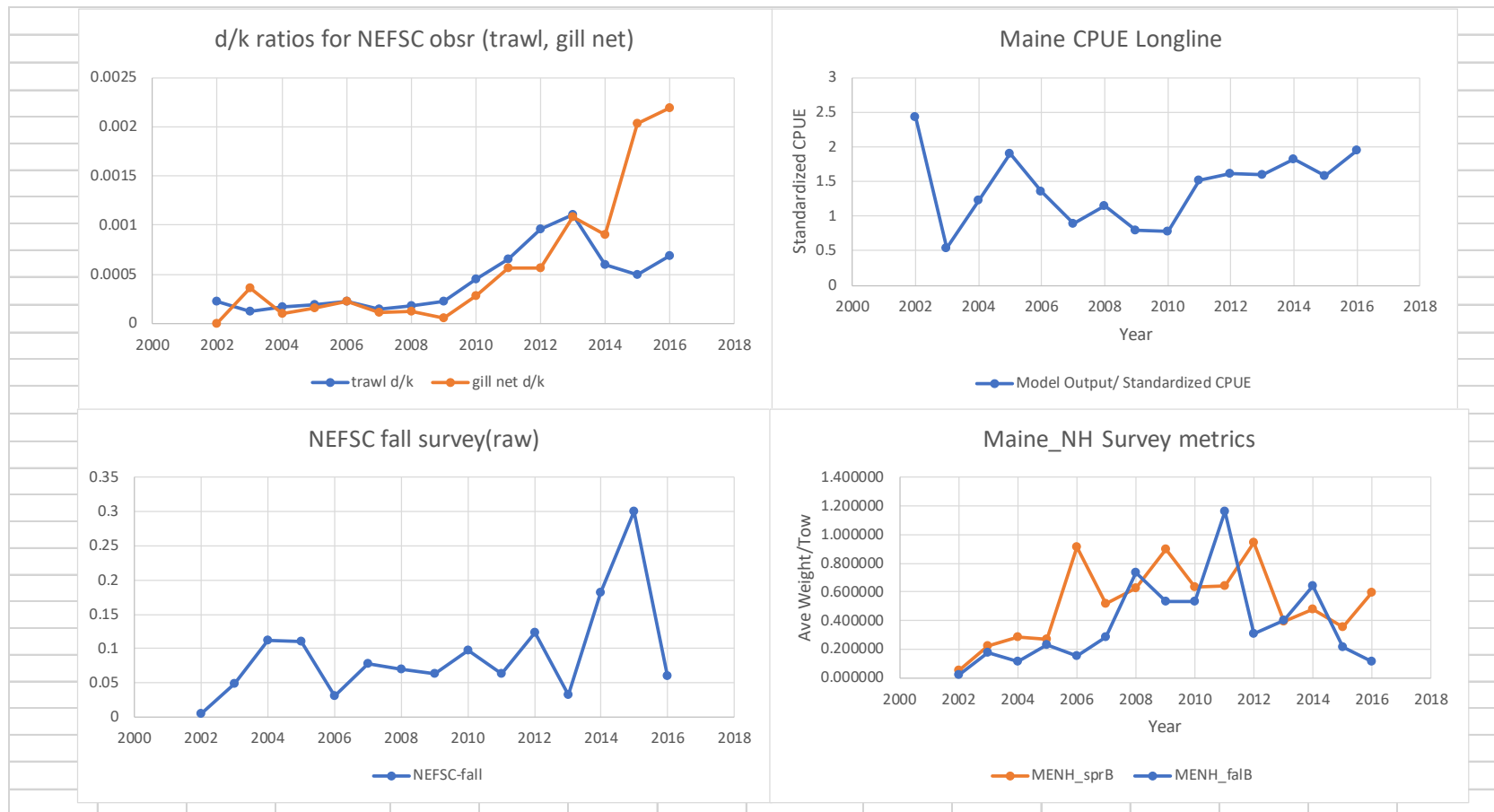
Figure 3.   Candidate relative abundance indices considered in Rcrit and FSD models. The Standardized longline CPUE data were obtained from Hansell et al (2017 ms). Results for the Maine-New Hampshire inshore bottom trawl survey were obtained from Sally Sherman, MEDMR. The d/k ratios were obtained from Susan Wigley, NEFSC. NEFSC bottom trawl estimates were obtained from Daniel Hennen, NEFSC.   See Appendix 4 for variable definitions.
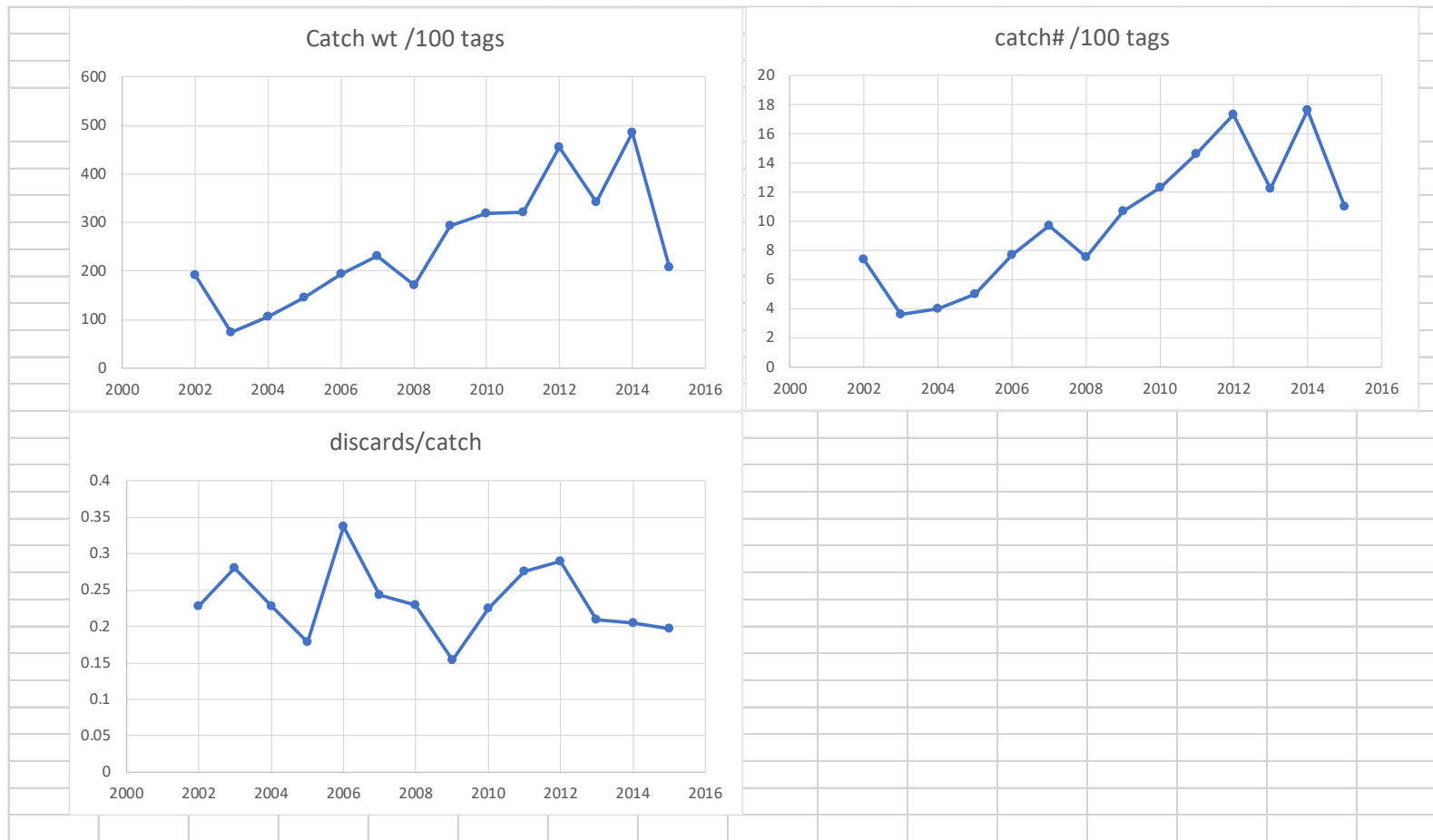
Fig 3.5 Estimates of catch per unit effort based on state data from Maine DMR.   Effort is expressed in terms of number of tags issued, and discard to catch ratios were based on fishermen logbooks. Given uncertainties about the data and the availability of a more refined measure of CPUE from Hansell et al. these data were not considered further in this assessment. See Appendix 4 for variable definitions.
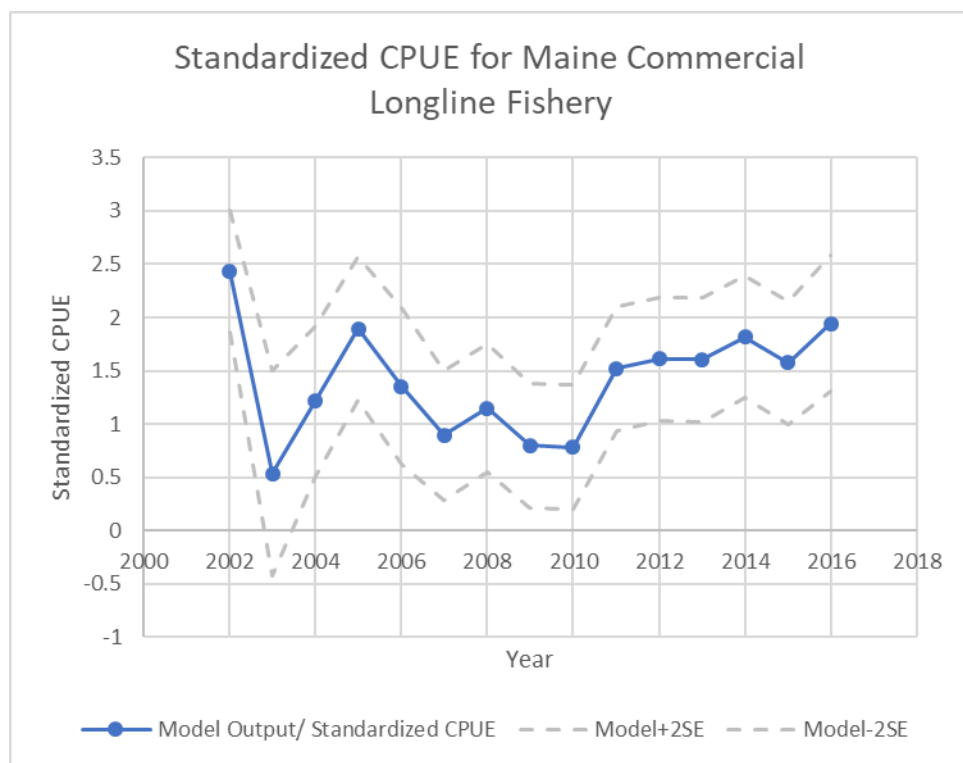
Figure 4. Standardized CPUE from Maine Commercial longline fishery, 2002-2016 provided courtesy of Hansell et al. Estimates and Error bounds are derived from a general linear model analysis.
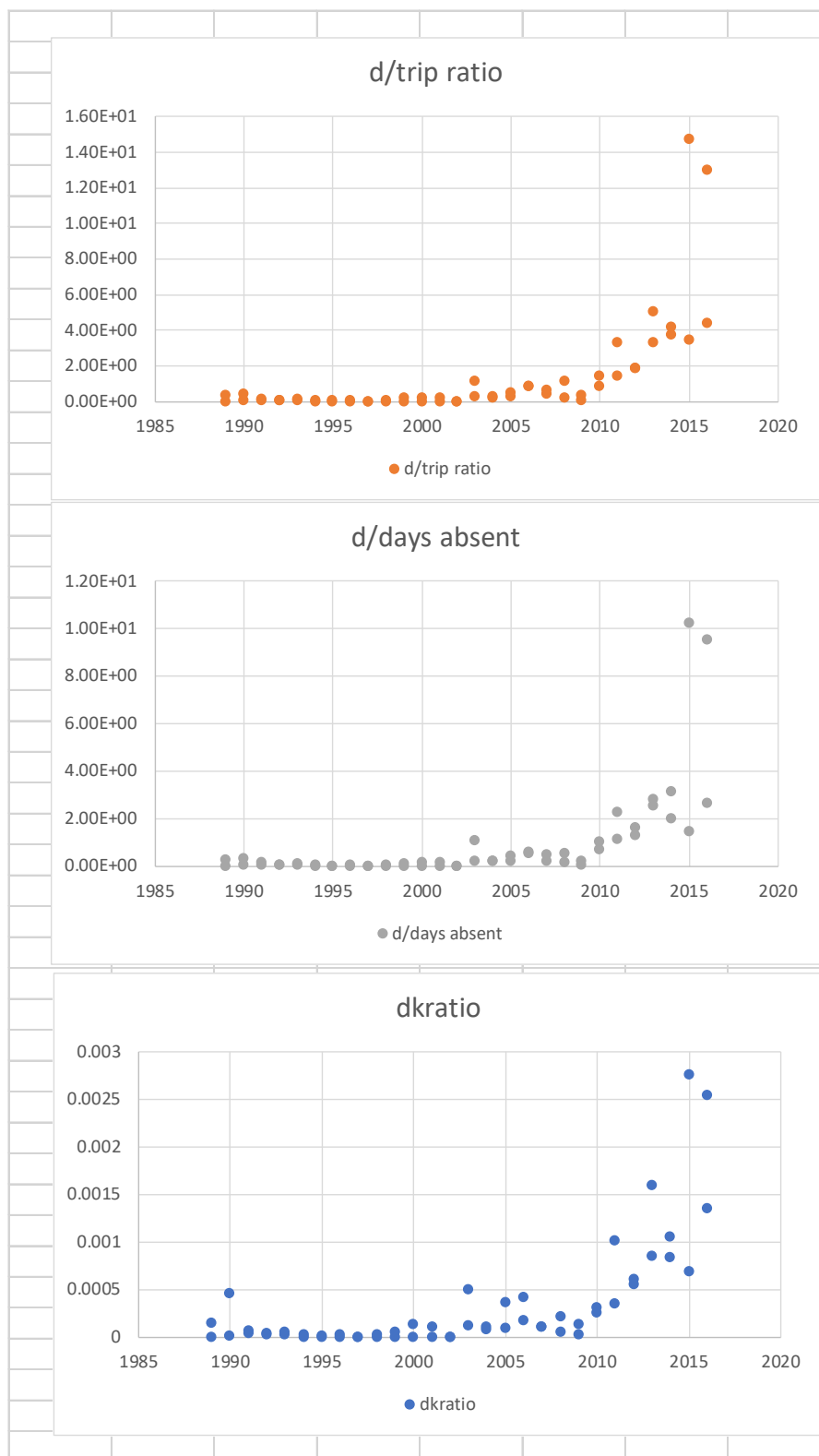
Figure 5. Discard ratios for observed gill net trips from 1989 to 2016 expressed in half year increments. Measures of effort include number of trips, number of days absent and total landed catch. Data courtesy of Susan Wigley, NEFSC
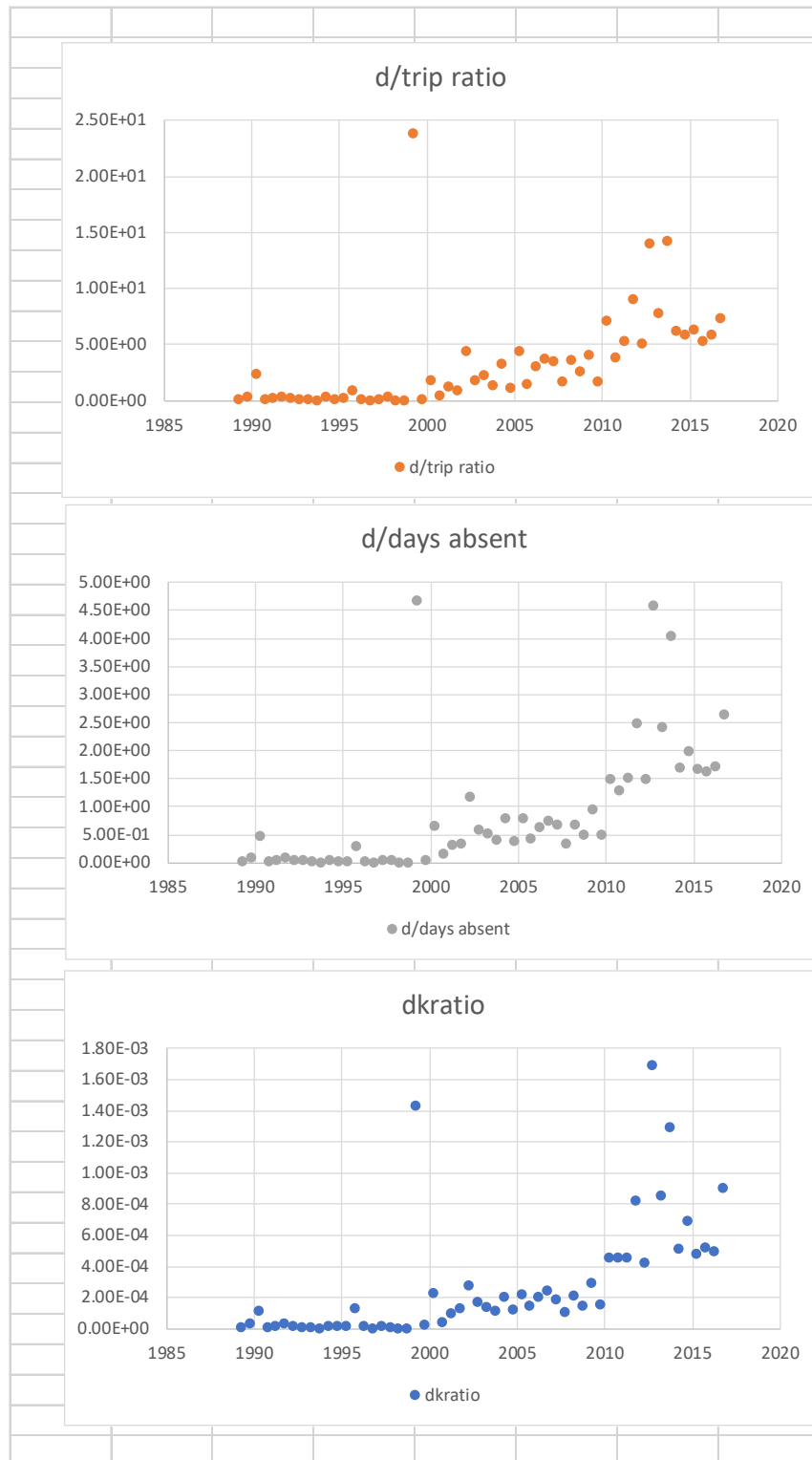
Figure 6. Discard ratios for observed otter trawl trips from 1989 to 2016 expressed in half year increments. Measures of effort include number of trips, number of days absent and total landed catch. Data courtesy of Susan Wigley, NEFSC
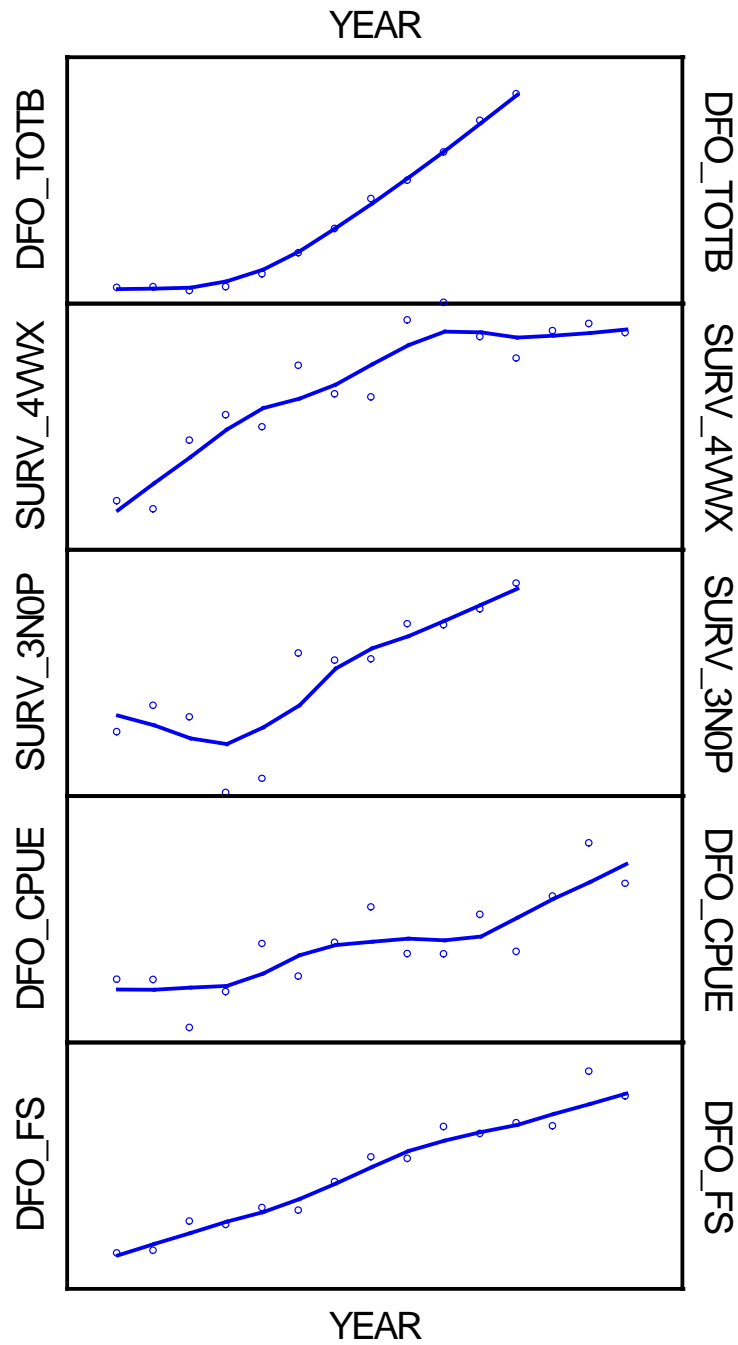
Figure 7. DFO abundance indices 2002-2016. DFO_TOTB is model based estimate of total biomass. Lines represent lowess smooths with tension =0.5. Variable names are DFO_FS, DFO_CPUE, SURV_3NOP, SURV_4VWX. See Appendix 4 for variable definitions.
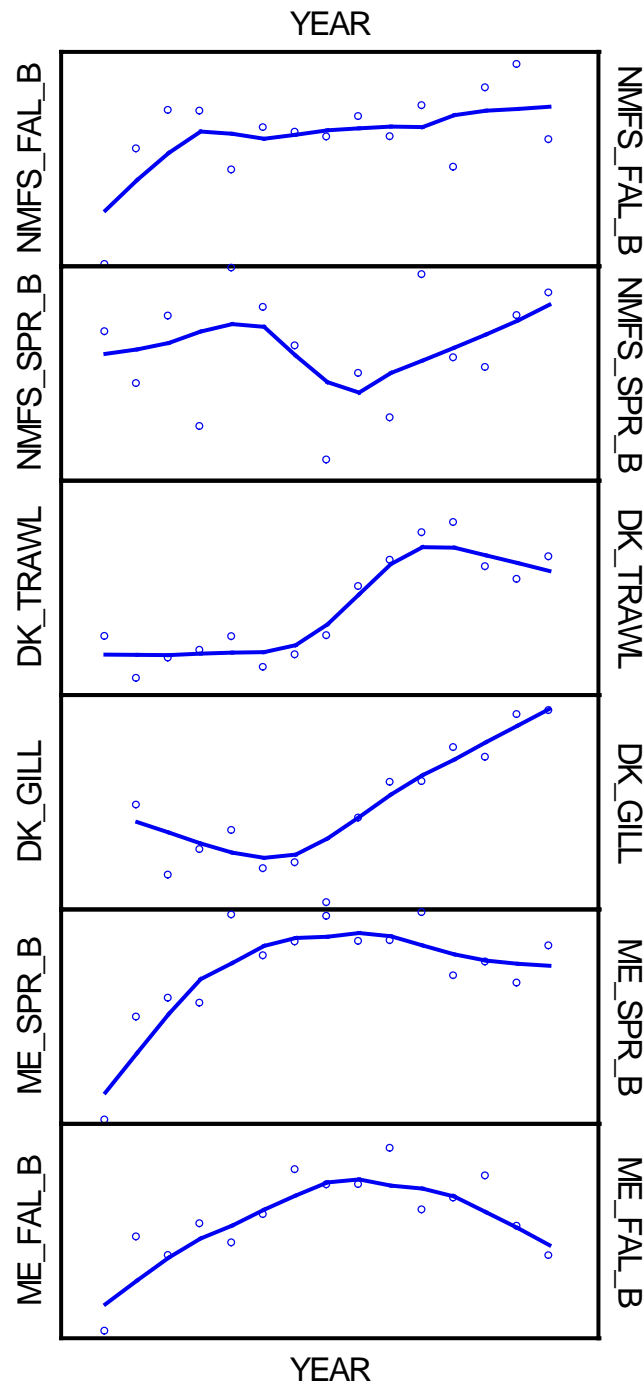
Figure 8. US abundance indices 2002-2016. Lines represent lowess smooths with tension =0.5. Variable names are ME_FAL_B, ME_SPR_B, DK_GILL, DK_TRAWL, NMFS_SPR_B, NMFS_FAL_B. See Appendix 4 for variable definitions.
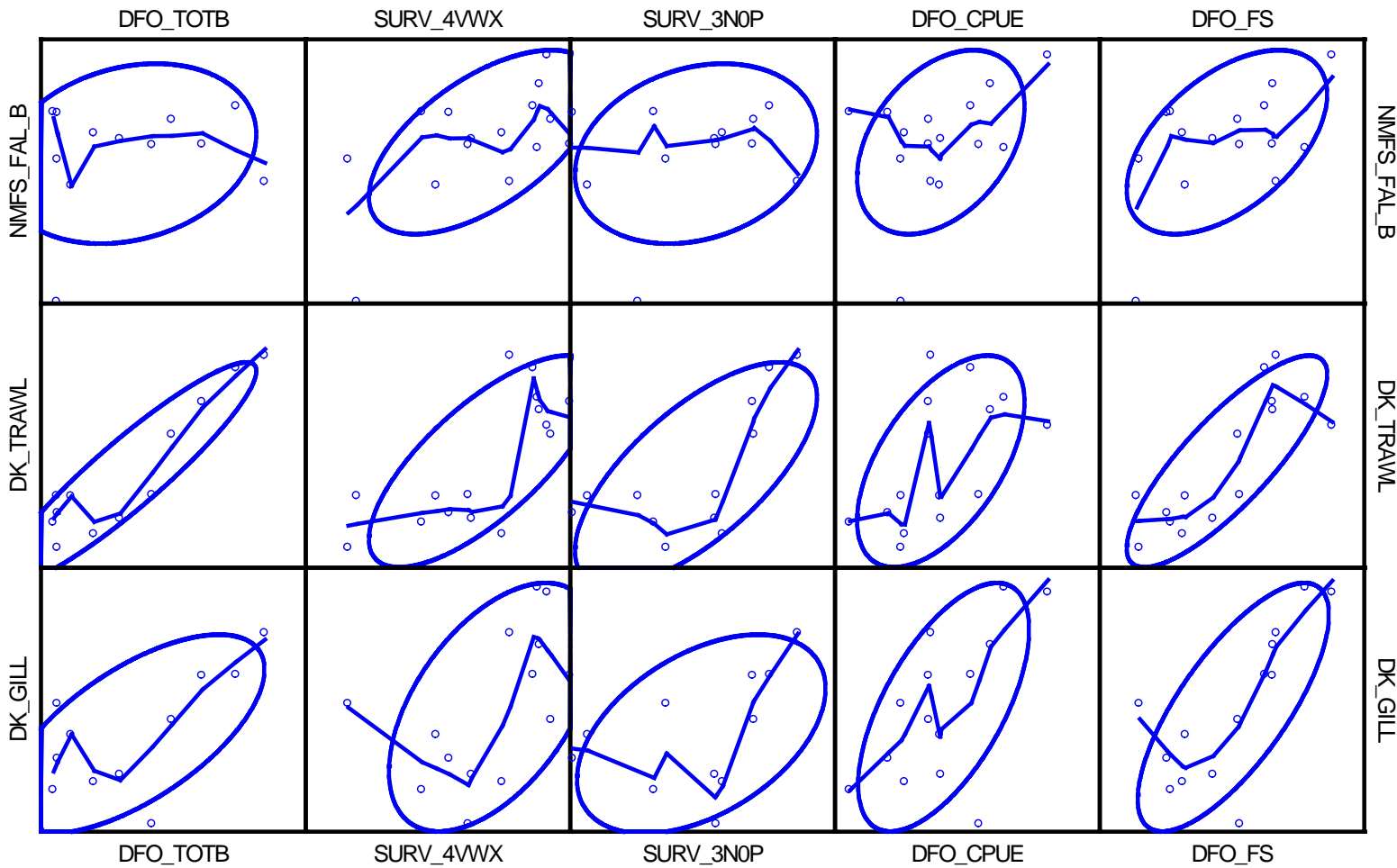
Fig 9. Comparison of DFO and US abundance indices for 2002-2016. Data are smoothed with a lowess mdethod with tension = 0.5. ME_FAL_B, ME_SPR_B, DK_GILL, DK_TRAWL, NMFS_SPR_B, NMFS_FAL_B. DFO_TOTB is model based estimate of total biomass. Variable names are DFO_FS, DFO_CPUE, SURV_3NOP, SURV_4VWX. See Appendix 4 for variable definitions.
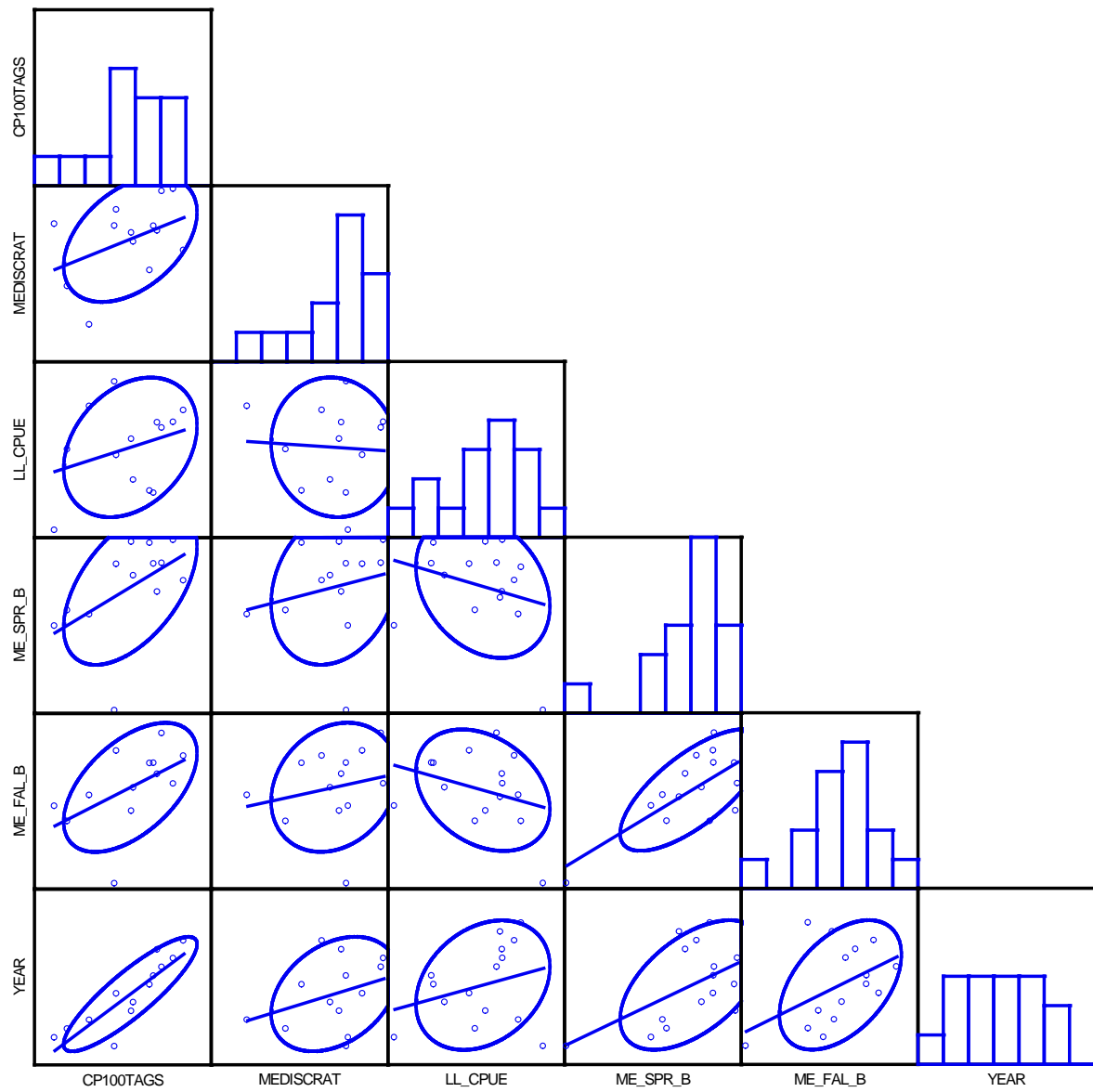
Figure 10. Scatterplot matrix for indices of abundance from Maine surveys and commercial CPUE. Lines represent simple linear regressions. ME_FAL_B, ME_SPR_B, DK_GILL, DK_TRAWL, NMFS_SPR_B, NMFS_FAL_B, LL_CPUE, MEDISCRAT, CP100TAGS. See Appendix 4 for variable definitions.
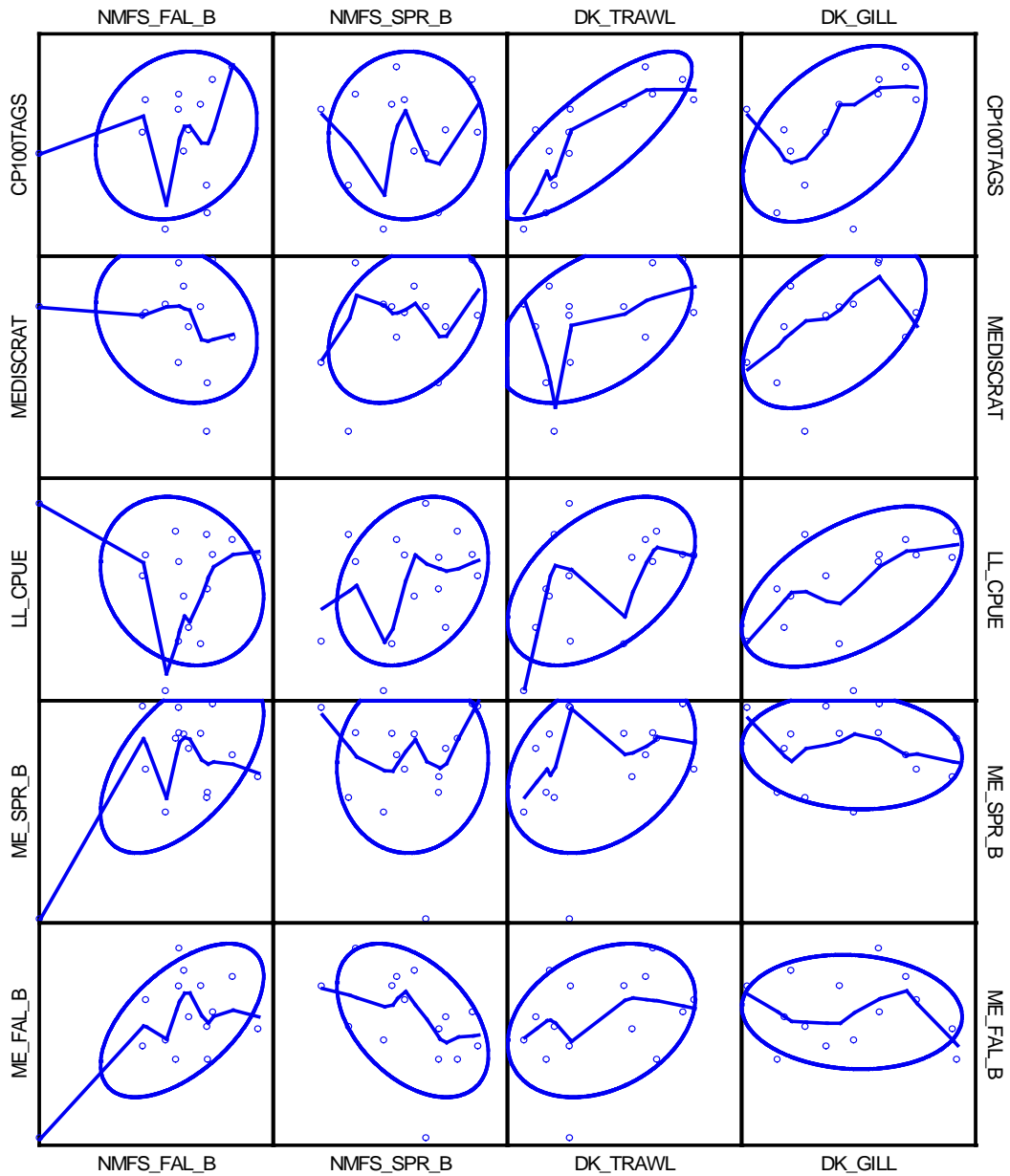
Figure 11. Scatterplot matrix comparison of Maine indices (rows) to NMFS indices. Lines represent lowess smooths with tension =0.5. ME_FAL_B, ME_SPR_B, DK_GILL, DK_TRAWL, NMFS_SPR_B, NMFS_FAL_B, LL_CPUE, MEDISCRAT, CP100TAGS. See Appendix 4 for variable definitions.
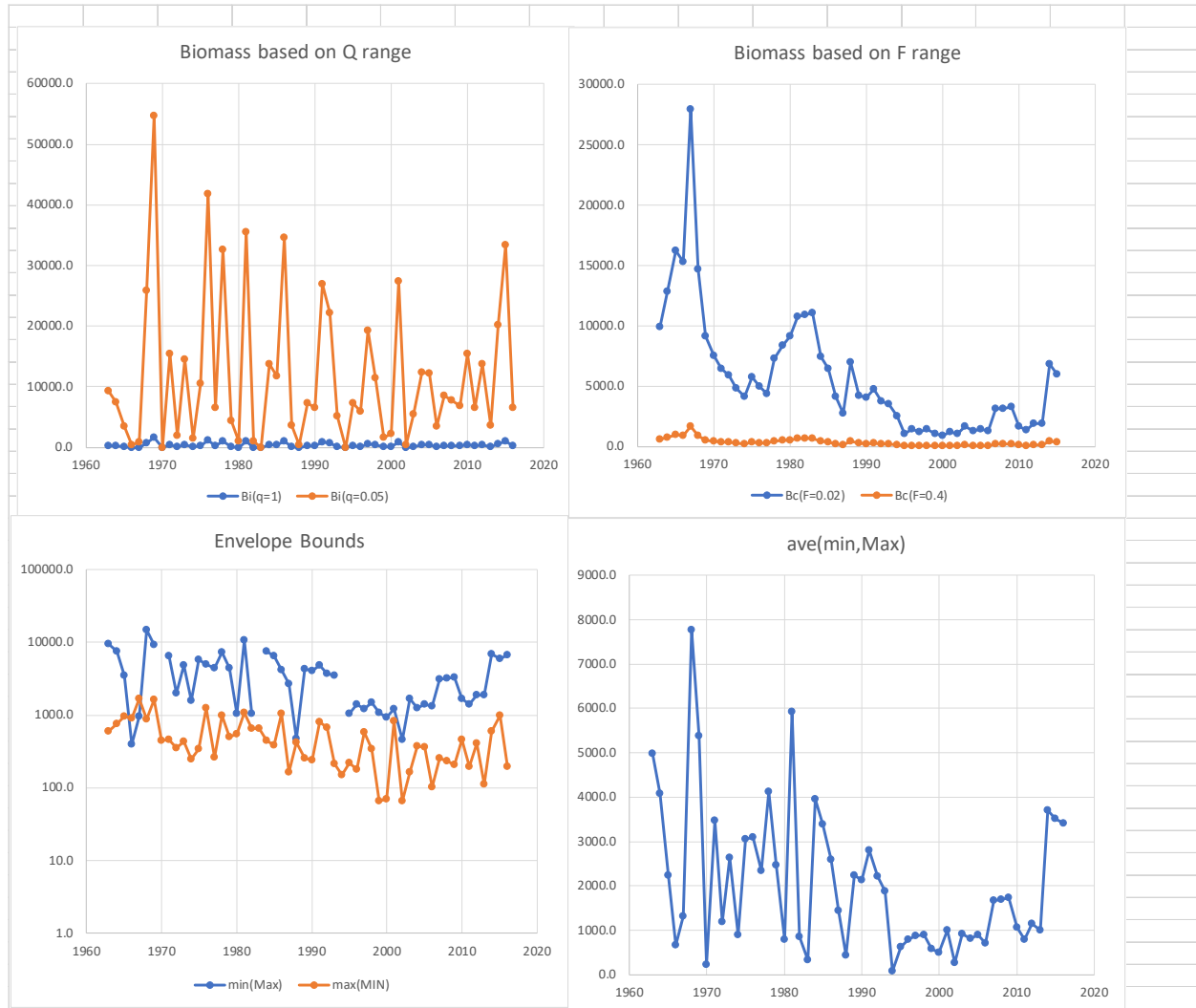
Figure 12. Result of Envelope model analyses of landings and NMFS fall bottom trawl survey estimates. Envelope bounds(lower left plot) represent constrained limits on a log scale. The lower right graph shows the average of the min and max envelope values. Estimates are based on raw survey data.
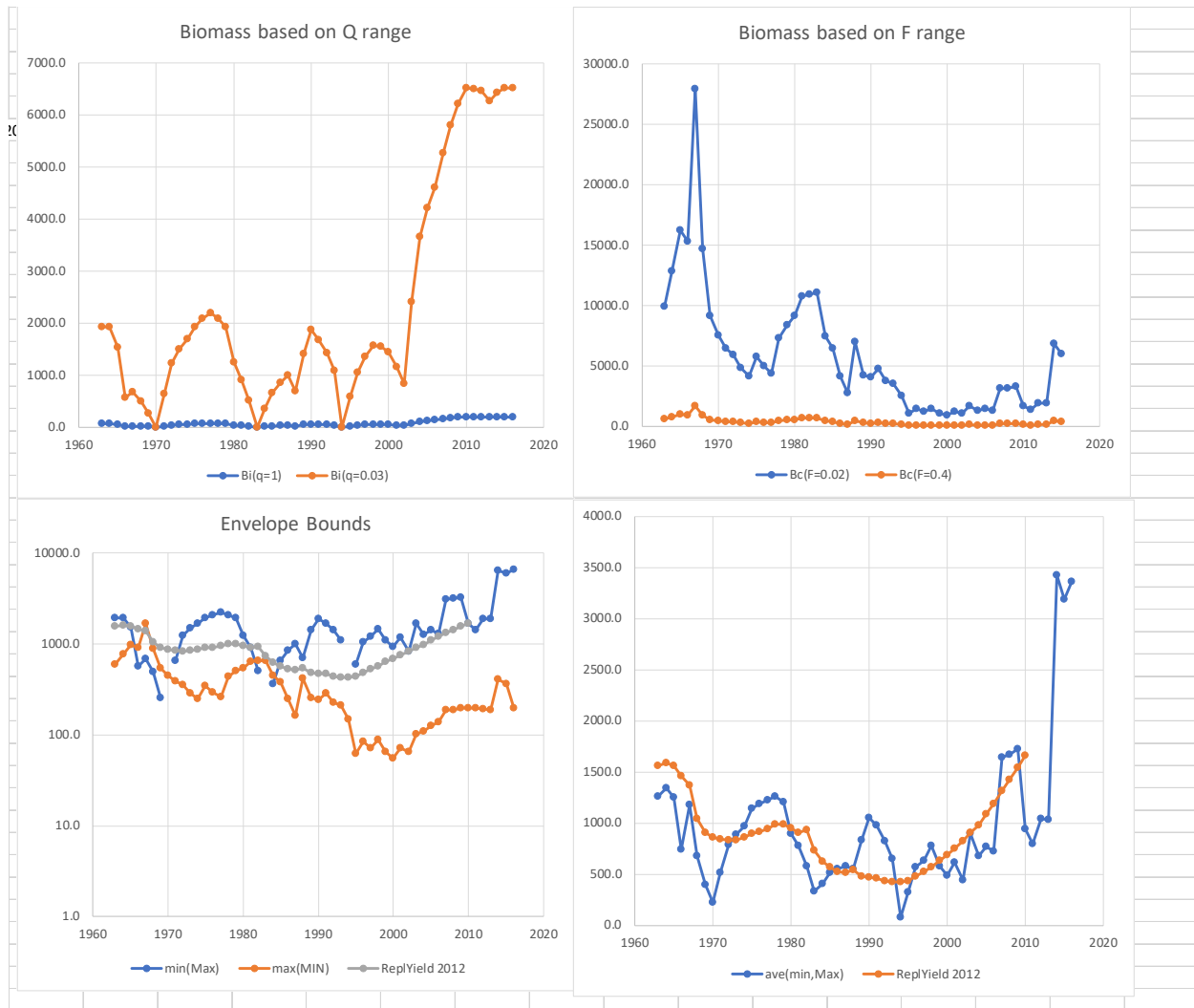
Figure 13. Result of Envelope model analyses of landings and NMFS fall bottom trawl survey estimates. Envelope bounds(lower left plot) represent constrained limits on a log scale.   The lower right graph shows the average of the min and max envelope values. Estimates are based on Kalman smoothed survey data. Biomass estimates from Replacement Yield model (Col and Legault 2012) are provide for reference.
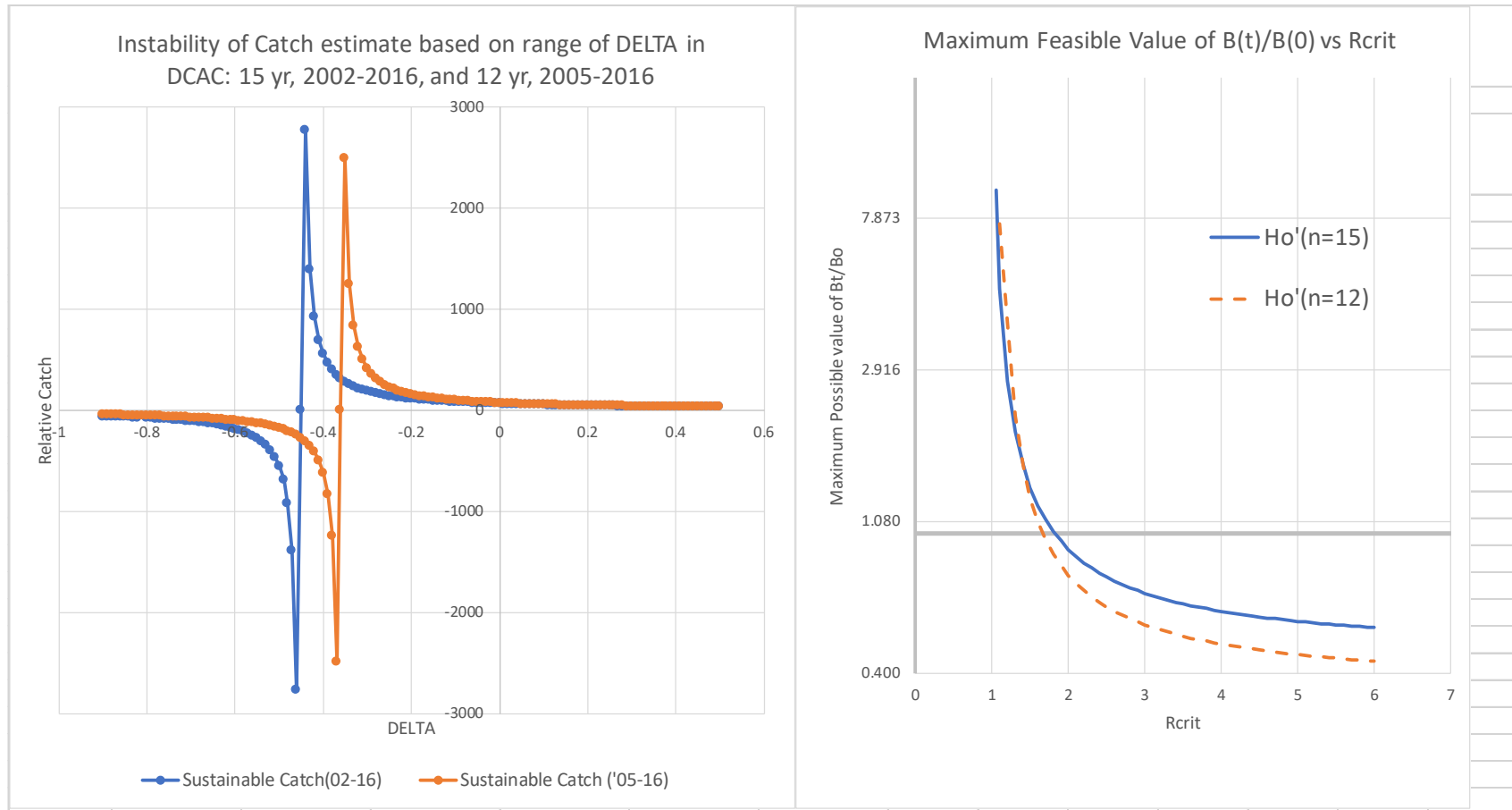
Figure 14. Singularity issues in DCAC for instances where the denominator becomes negative (left plot). The right plot shows the maximum possible value of Bt/B0 for alternative values of Rcrit given 12 or 15 years of catch data.

$$C_t$$

$$\downarrow$$

Population
Black Box

$$\downarrow$$

$$I_{t+\Delta t}$$

$$\downarrow$$

$$F(I_{t+\Delta t}) < F(I_t)?$$

NO   YES

$$C_{t+\Delta T} > C_t \qquad C_{t+\Delta T} < C_t$$

Population
Black Box

$$\downarrow$$

$$I_{t+2\Delta t}$$
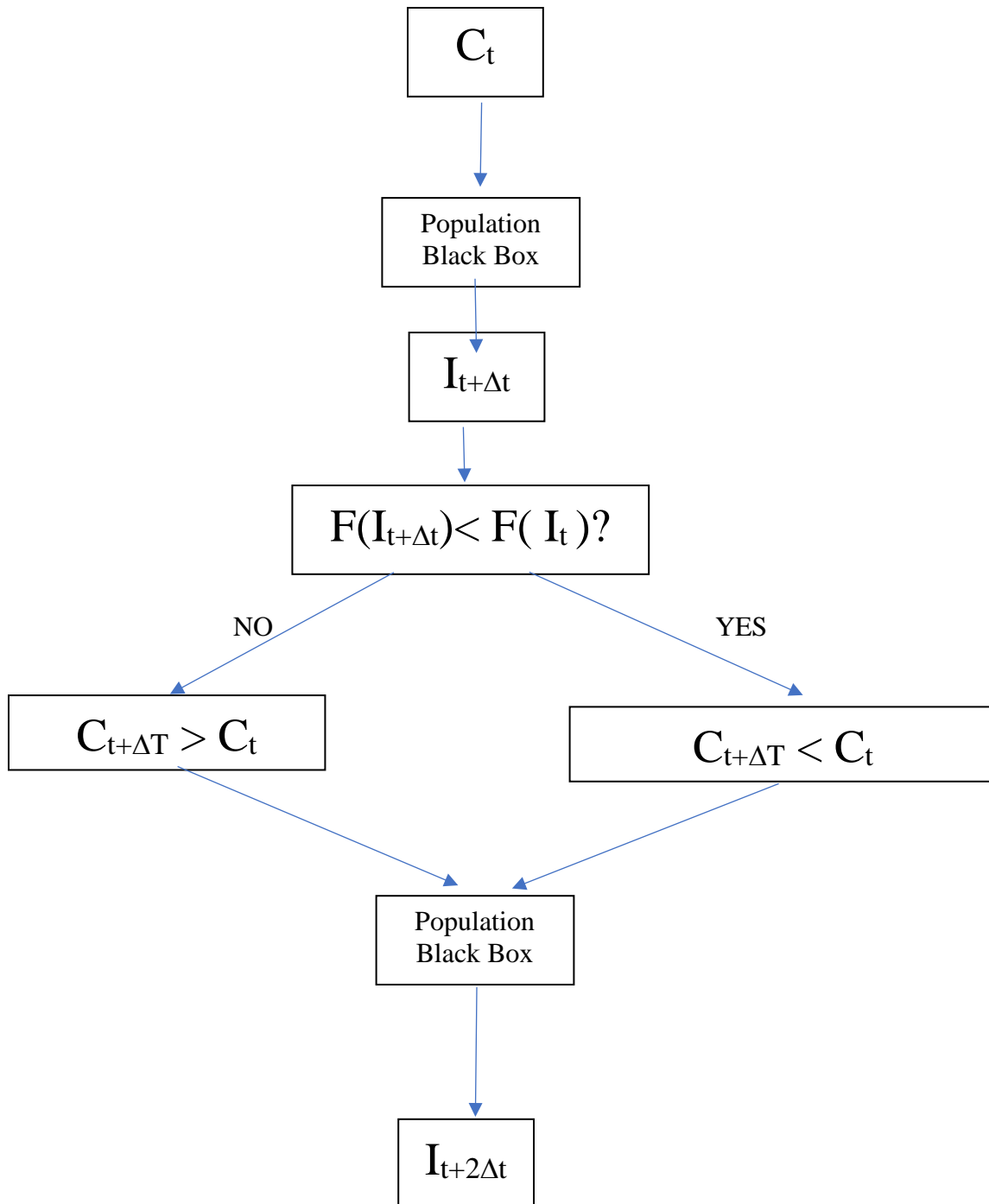
Figure 15. Schematic depiction of the feedback loop used in the FSD model.
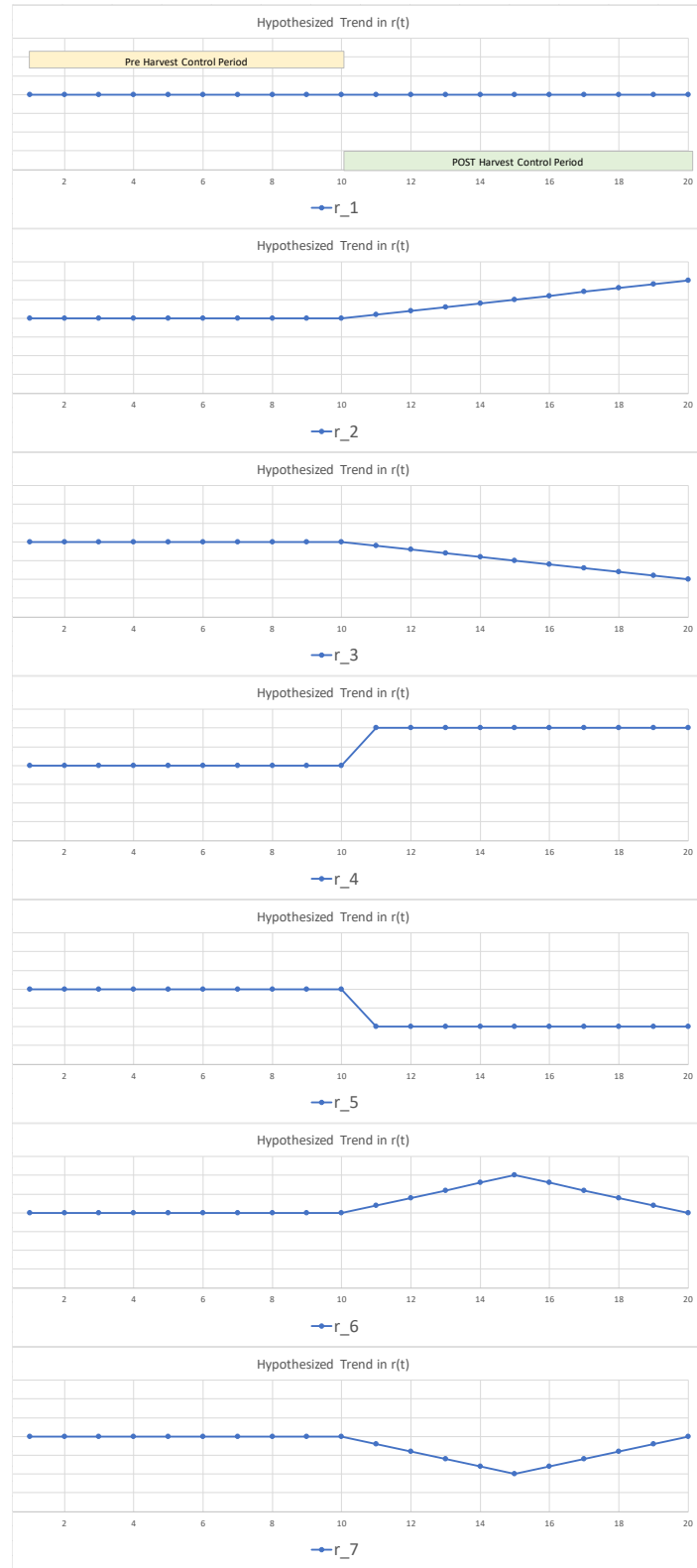
Figure 16. Scenarios for the intrinsic rate of increase r(t) used in the simulation analysis of the FSD model. Note that r(t) is constant during the first ten years, prior to implementation of the FSD control rule.
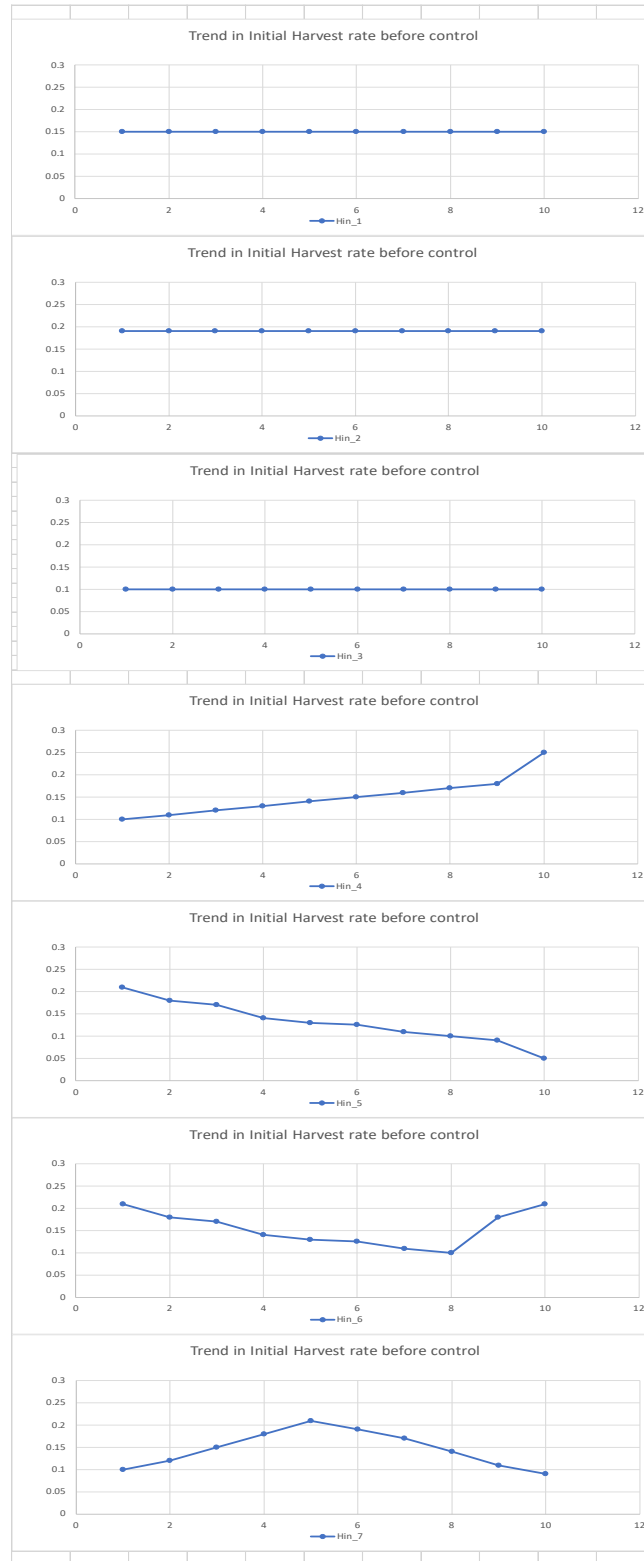
Figure 17. Scenarios for the initial harvest rate h(t) used in the simulation analysis of the FSD model. Note that the harvest rate is used only for the first 10 years. After that, harvest is controlled by the FSD harvest control rule.
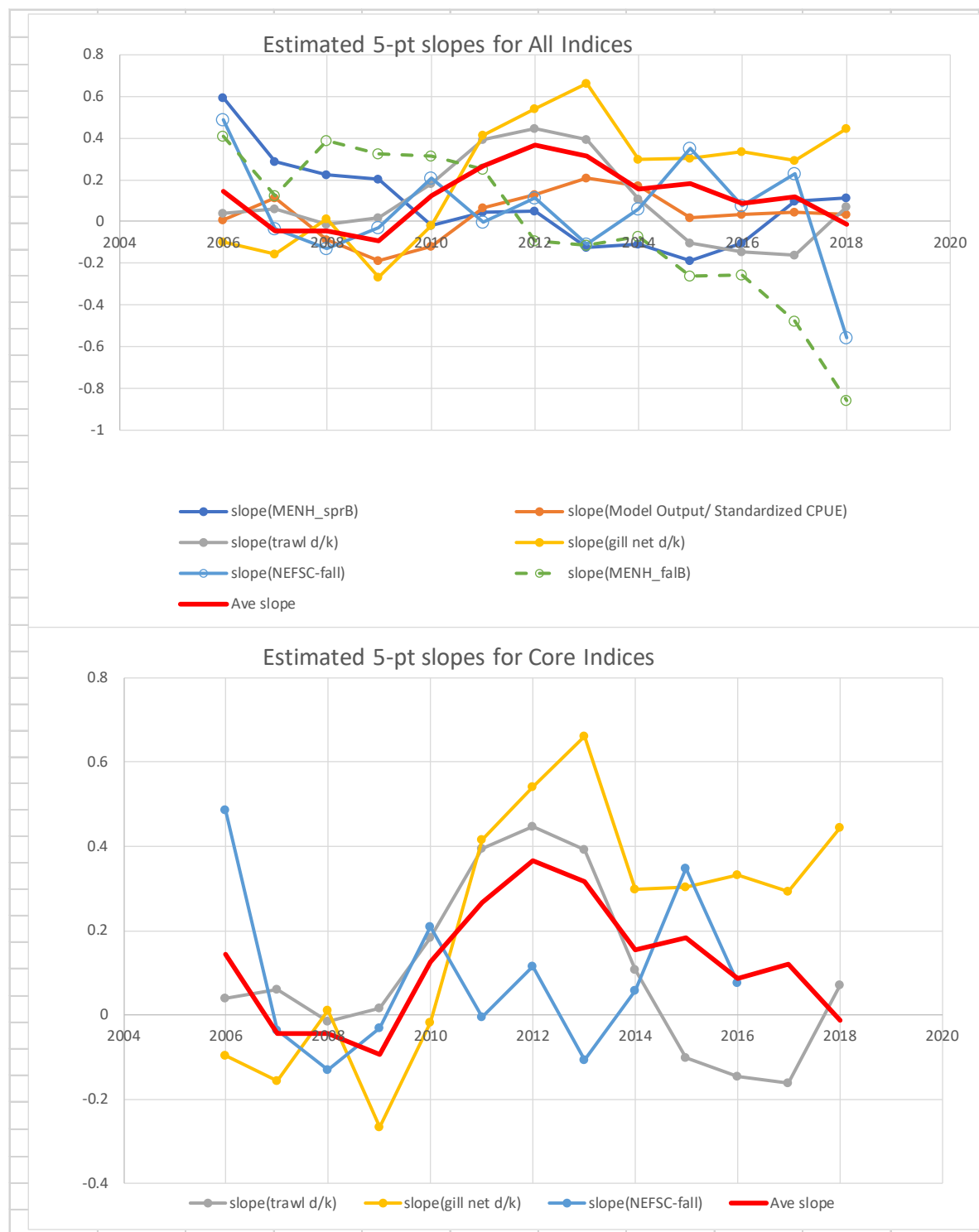
Figure 18. Trends in abundance indices based on 5 point regressions. Top plot shows full model with 5 indices. Bottom plot shows reduced model using only existing indices.
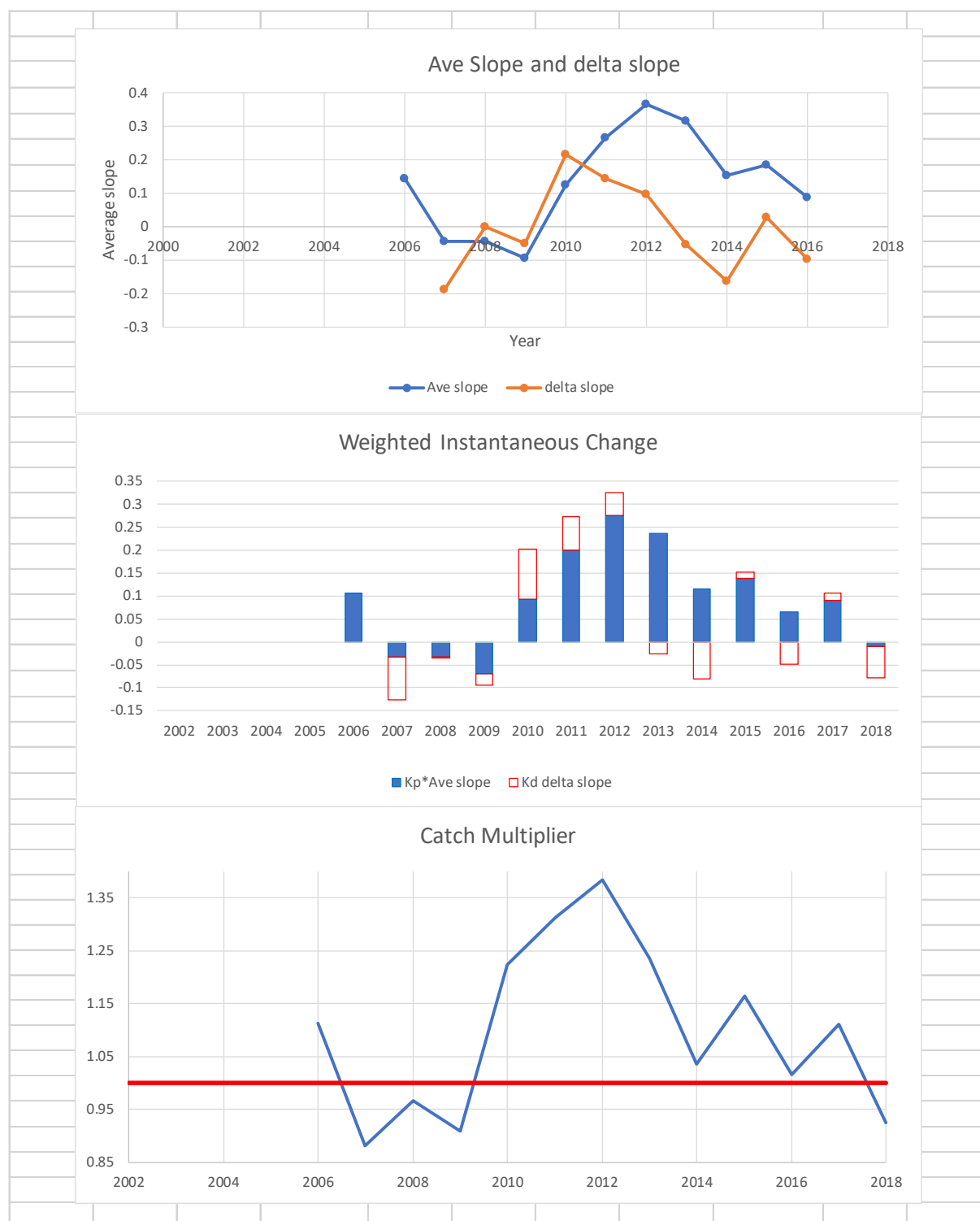
Figure 19. Summary of FSD model results for US Atlantic halibut based on d/k trawl, d/k gill net and NEFSC fall survey abundance indices (See Fig. 18 bottom). Instantaneous rates of change represent Kp and Kd weighted values of the first and second derivative (top figure). The bottom figure show the Catch multiplier used to forecast catch. . The gain parameters for proportional and derivative were set at Kp=0.75 and Kd=0.50, respectively.
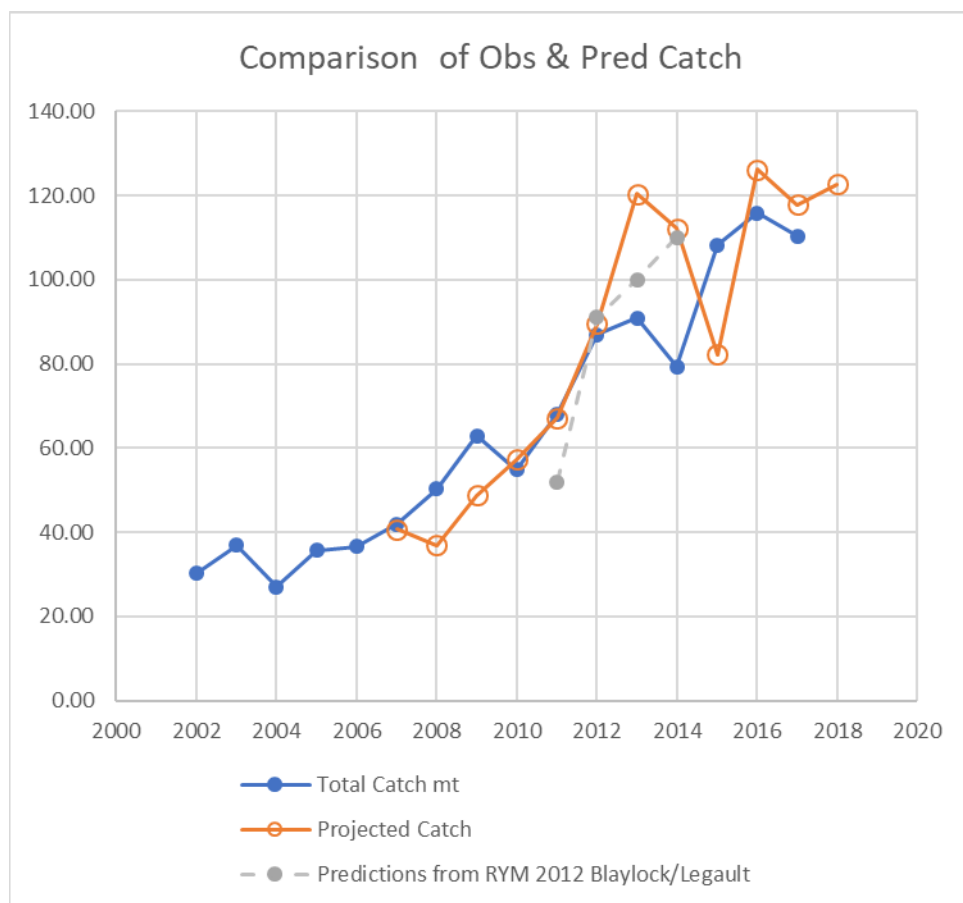
Figure 20. Comparison of observed vs predicted catches based on the FSD model applied to US stock area. Forecasts from RYM application (Blaylock and Legault, 2012) are included for comparison. The gain parameters for proportional and derivative were set at Kp=0.75 and Kd=0.50, respectively.
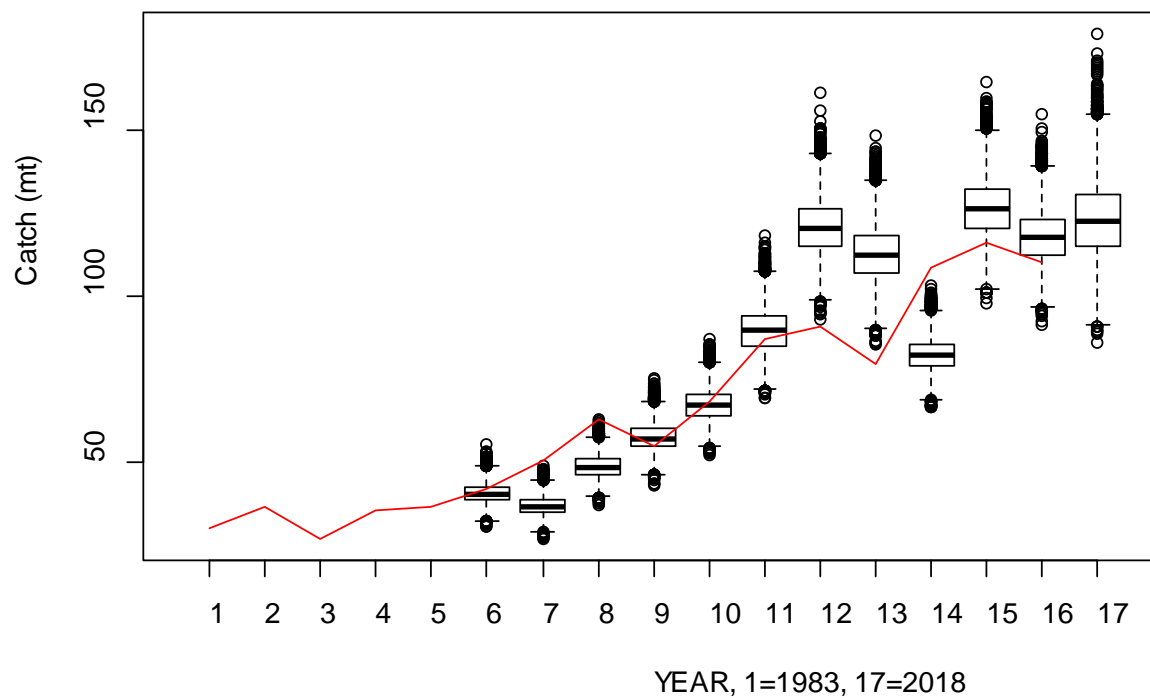
Figure 21. Comparison of observed and predicted distribution of catches for FSD model applied to the US stock area. Uncertainty estimates are based on a parametric bootstrap method described in the text. The gain parameters for proportional and derivative were set at Kp=0.75 and Kd=0.50, respectively.

Histogram of C.hat.boot[Tmp1,]
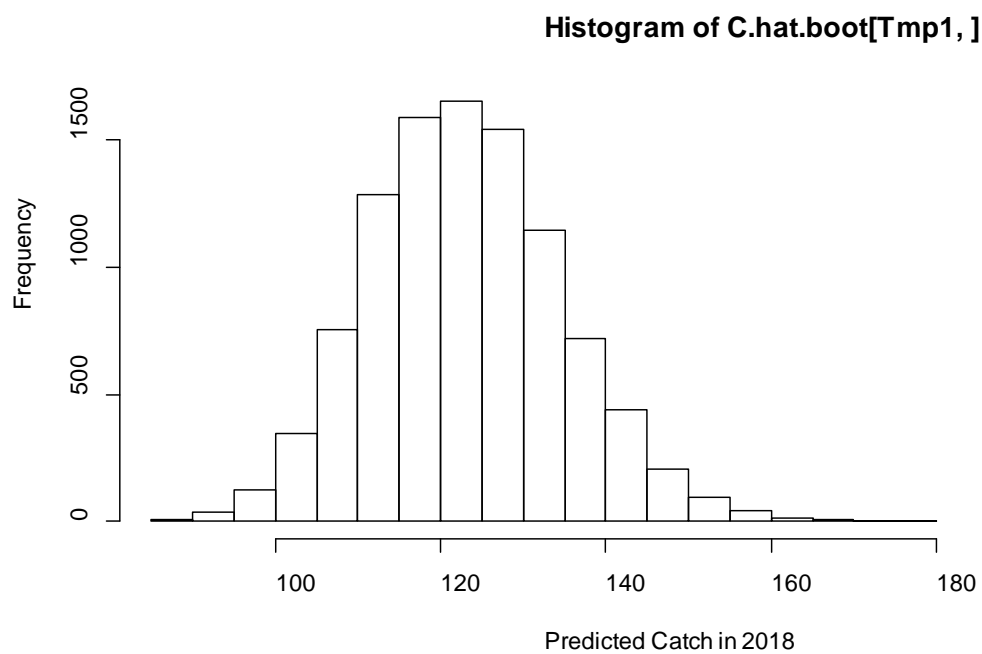
Figure 22. Sampling distribution of predicted catch for 2018 based on parametric bootstrap method with 5000 replications.
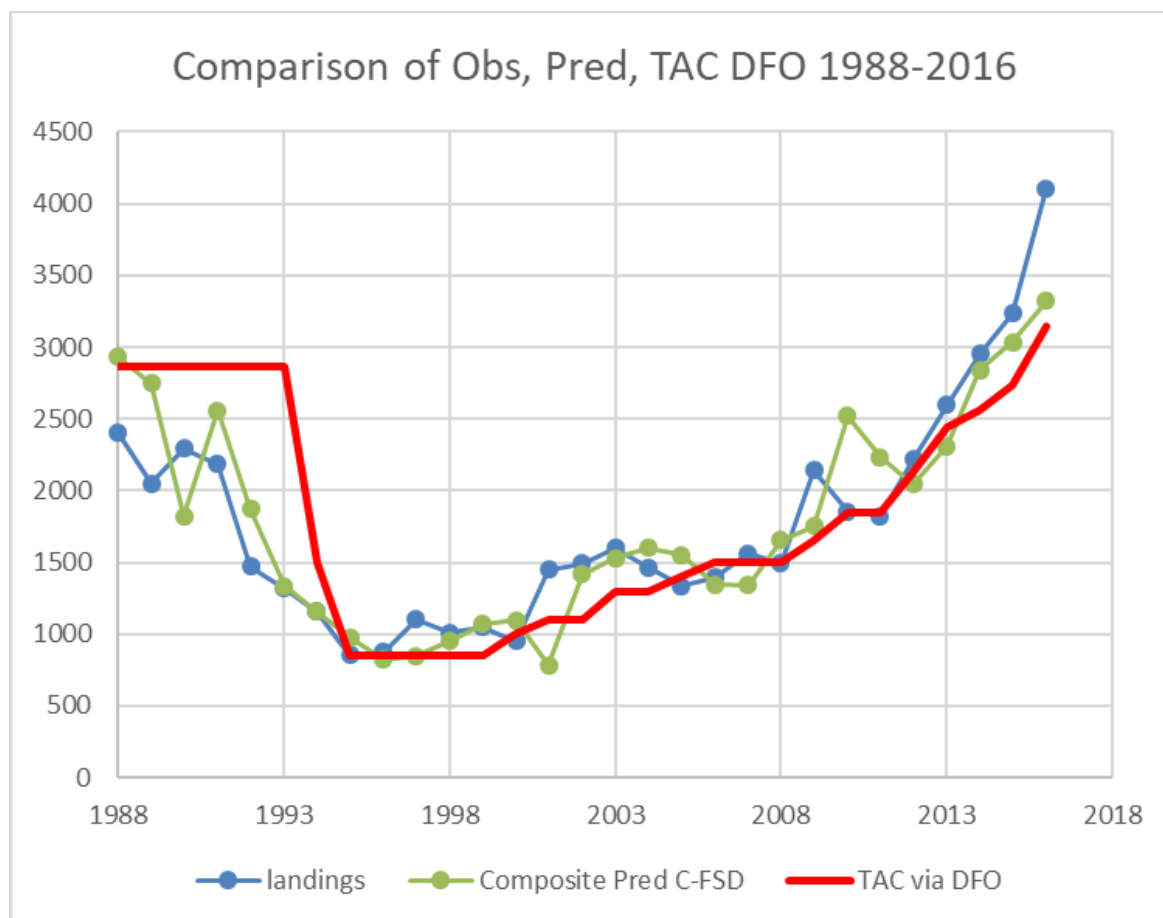
Figure 23. Comparison of FSD prediction with observed landings and TAC for DFO 3NOPs4VWX5Zc stock of Atlantic halibut.
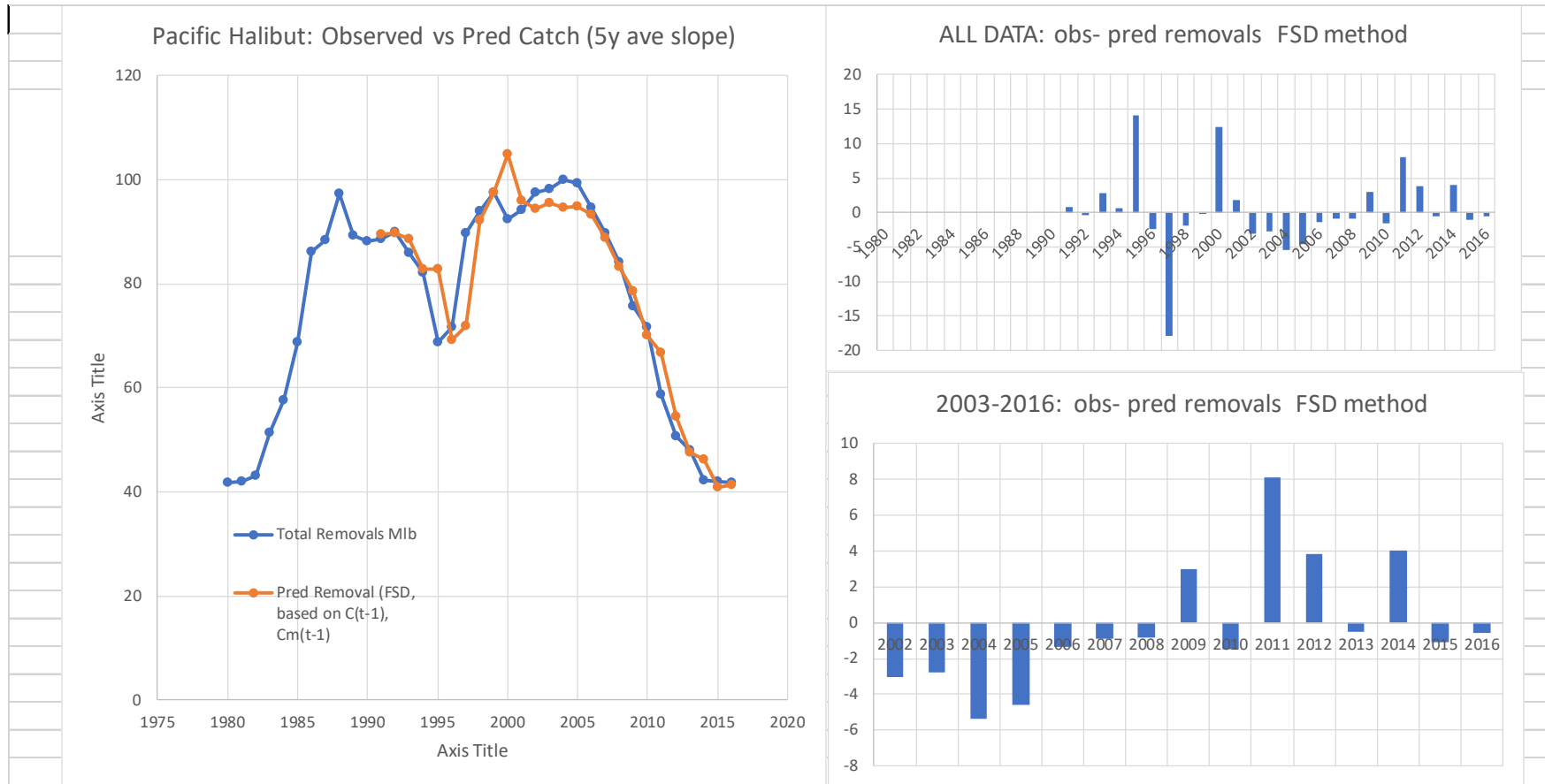
Figure 24. Example application of FSD model to observed catches of IPHC Pacific halibut. Residuals are shown on right hand plots for the full and reduced time series.
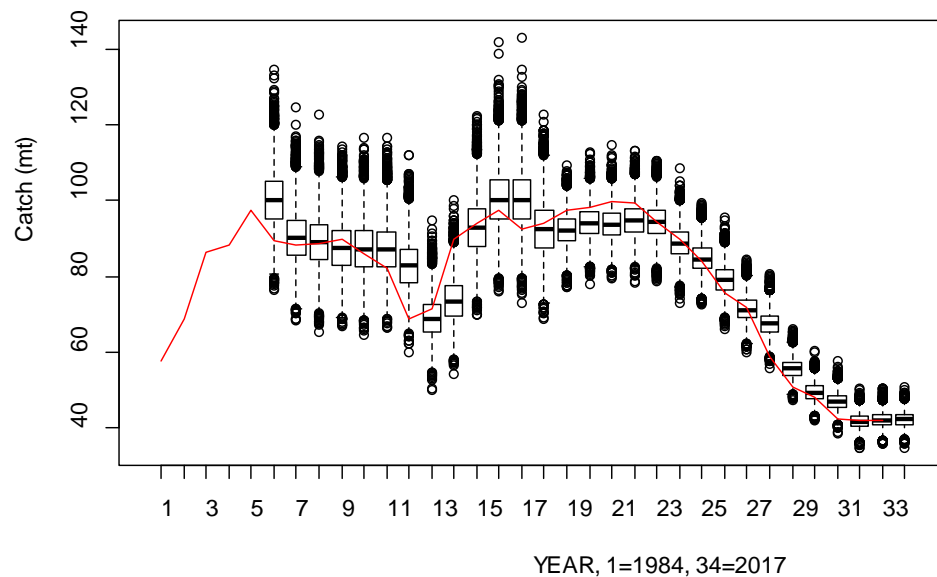
Figure 25. Comparison of observed and predicted catches based on the FSD model with Kp=0.778 and Kd=0.1. Research survey indices were assumed to have a CV=0.2 and Commercial CPUE indices were assumed to have a CV=0.25.
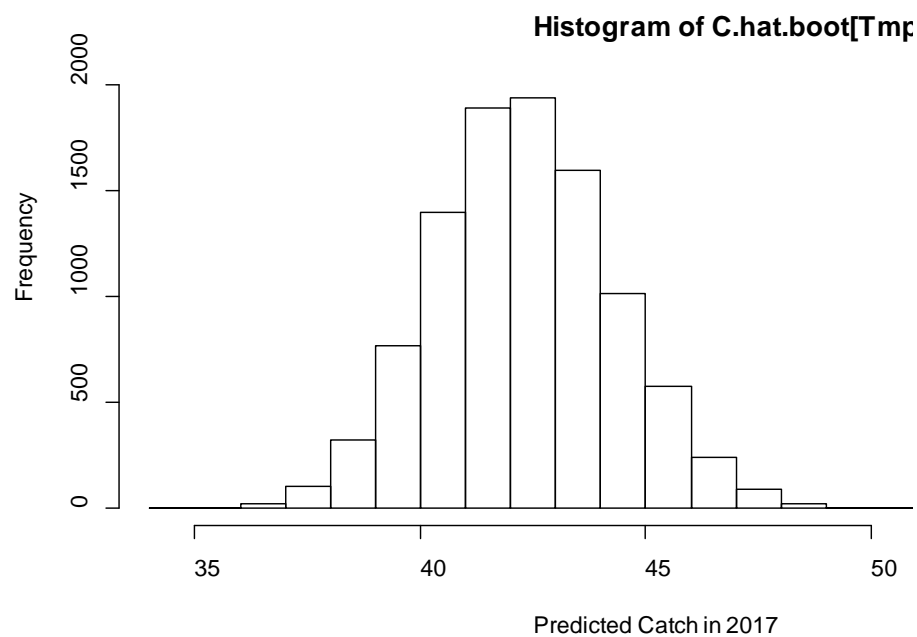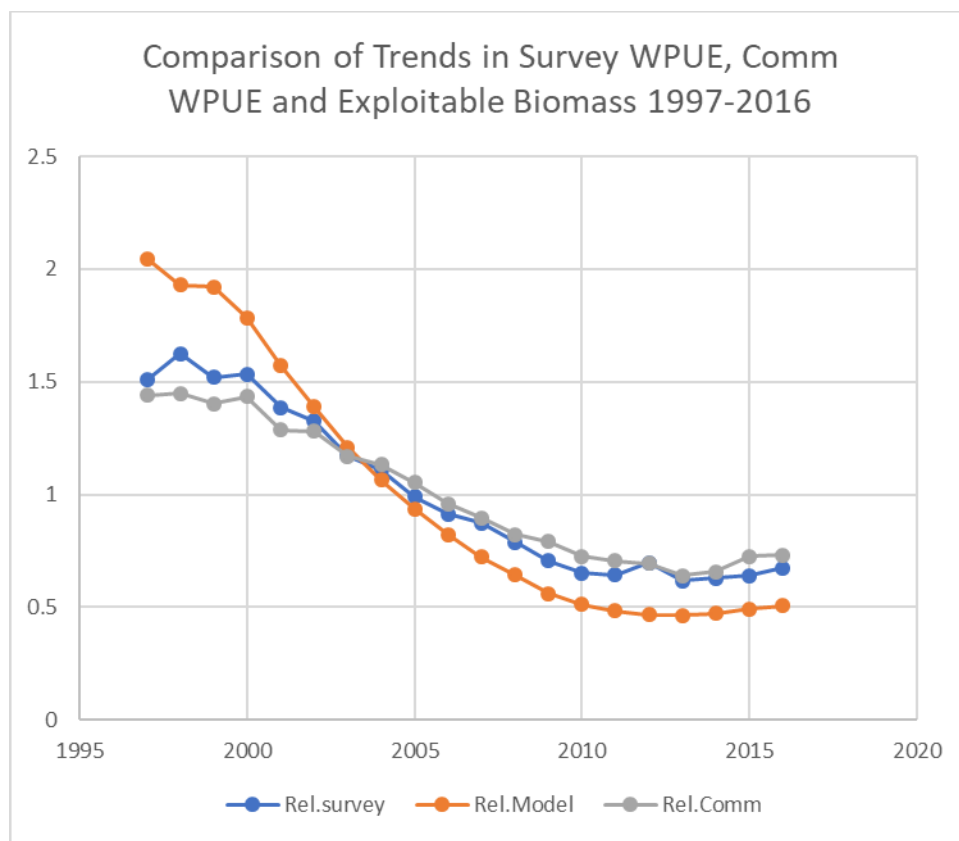

Figure 26. Predicted sampling distribution of estimated total removals in 2017 for Pacific halibut. The FSD model used Kp=0.778 and Kd=0.1. Research survey indices were assumed to have a CV=0.2 and Commercial CPUE indices were assumed to have a CV=0.25

Figure 27. Comparison of relative trends in biomass for IPHC Pacific halibut surveys and modeled biomass.   Quantities are expressed as ratio of observed value to its mean for the period 1997-2017. Data were obtained from Stewart(2017) and Stewart and Hicks (2017).
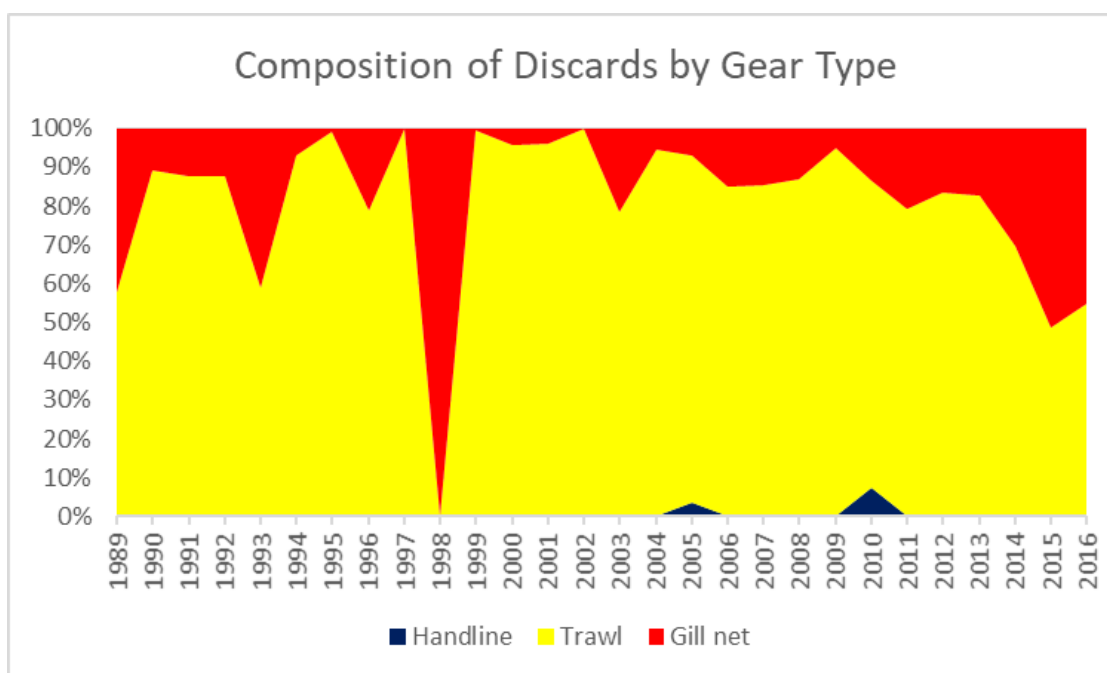
Figure 28. Estimated proportions of discards by gear type, 1989-2016. Estimates are based on an assumed 100% mortality.
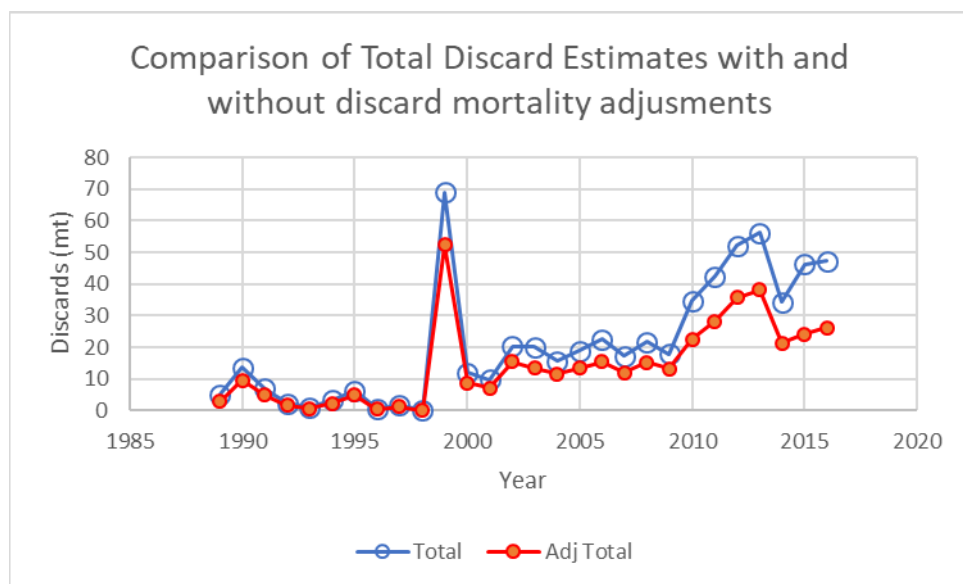

Figure 29. Comparison of total discard estimates based on assumed rates of discard mortality: Trawls 76%, Gill nets 30% and hook gear 10%.
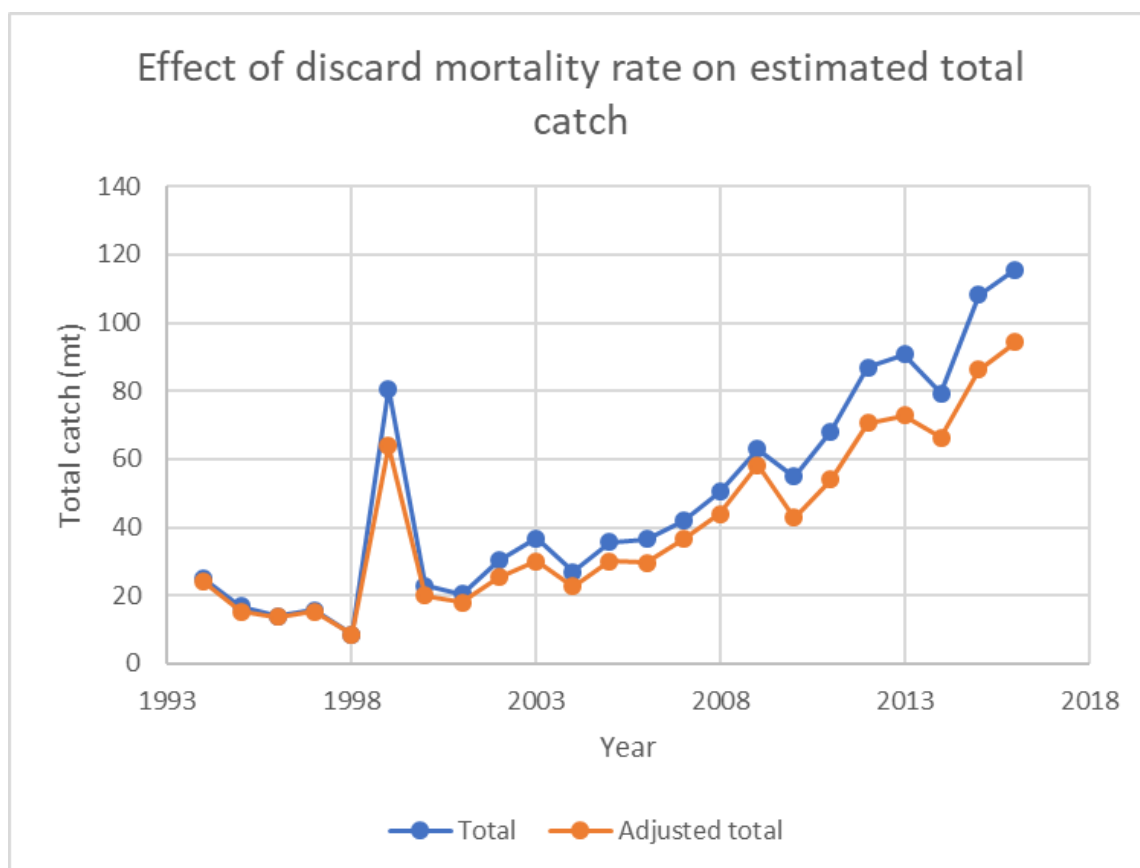
Figure 30. Comparison of total catch estimate based on alternative assumptions about discard mortality. Total assumes 100% mortality, Adjusted total assumes gear specific discard mortality rates of 76% for trawls, 30% for gill nets and 10% for hook gear.
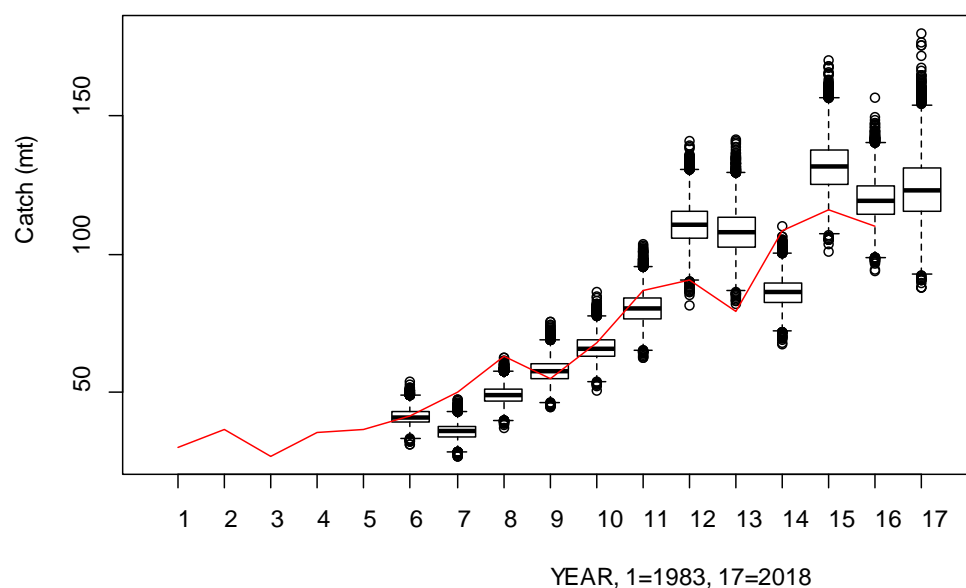
Figure 31. Comparison of observed and predicted distribution of catches for FSD model applied to the US stock **area using total catch to total kept all indices as measures of relative abundance**. Uncertainty estimates are based on a parametric bootstrap method described in the text. The gain parameters for proportional and derivative were set at Kp=0.75 and Kd=0.50, respectively.
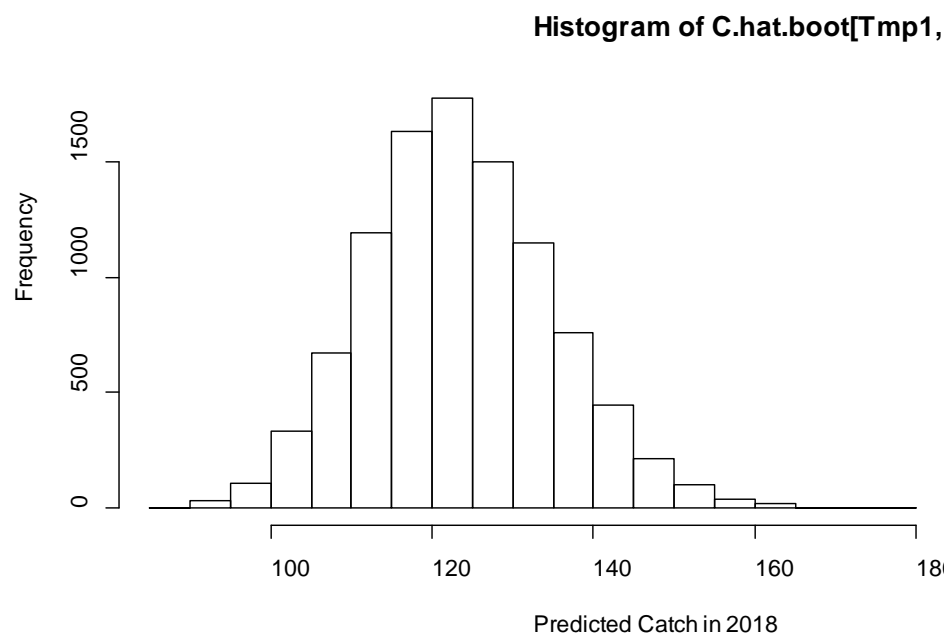

Figure 32. Distribution of forecasted catch for 2018 based on t/k ratio for the US stock area Atlantic halibut. Kp=0.75, Kd=0.5. The model uses t/k ratio for gill nets and trawl trips as measure of relative abundance.

APPENDIX 1. Summary of Input Data considered for use in FSD model for US stock area.

Table 1.1 Summary of average weight per to in the NEFSC bottom trawl survey for fall and spring seasons, 1968-2016.

NEFSC bottom trawl biomass indices (kg/tow)					NEFSC bottom trawl biomass indices (kg/tow)				
Year	Spring B	Spring B_CV	Fall B	Fall B_CV	Year	Spring B	Spring B_CV	Fall B	Fall B_CV
1963	NA	0	0.0848	0.6024	1990	0.0638	1	0.0596	0.4172
1964	0	0	0.0669	0.5215	1991	0.0618	0.9497	0.2434	0.829
1965	0	0	0.0316	0.5376	1992	0.0368	0.6815	0.2007	0.738
1966	0	0	0.0036	1	1993	0.0058	0.9999	0.0462	0.4607
1967	0	0	0.0086	1	1994	0.0172	0.9997	0	NA
1968	0.1294	0.5851	0.2335	1	1995	0.0051	0.7132	0.0661	1.0001
1969	0.2363	0.5754	0.4943	0.9433	1996	0.0126	0.7075	0.0532	1
1970	0.1054	0.7436	0	NA	1997	0.0626	0.6585	0.1735	0.6996
1971	0.0329	0.7144	0.1393	1	1998	0.0173	0.6558	0.103	0.5894
1972	0.0055	1	0.0182	0.8064	1999	0.2394	0.9686	0.0147	0.6154
1973	0.1129	0.8448	0.1314	0.913	2000	0	NA	0.0209	1
1974	0.1116	0.5555	0.0141	1	2001	0.1626	0.8797	0.2474	0.8066
1975	0	.	0.0951	0.9542	2002	0.128	0.6385	0.0041	1
1976	0.6439	0.9105	0.3775	0.6905	2003	0.0525	0.9486	0.049	0.5685
1977	0.1418	0.4812	0.0588	0.699	2004	0.1676	0.9827	0.1119	0.2902
1978	0.1628	0.7433	0.2943	0.797	2005	0.0251	0.6704	0.1105	0.6199
1979	0.3565	0.4123	0.04	0.5102	2006	0.383	0.46	0.0312	0.6105
1980	0.5625	0.6764	0.0095	0.7326	2007	0.1946	0.6034	0.0774	0.6127
1981	0.0659	0.7243	0.3214	0.6741	2008	0.1005	0.5723	0.0701	0.4966
1982	0.0817	0.7678	0.115	0.862	2009	0.0141	0.4794	0.0633	0.3948
1983	0.6108	0.5743	0	NA	2010	0.0625	0.3045	0.098	0.3592
1984	0.0224	0.8456	0.1237	1	2011	0.0291	0.667	0.0638	0.4975
1985	0.063	0.8692	0.1064	1	2012	0.3418	0.8601	0.1241	0.5256
1986	0	NA	0.3129	0.7392	2013	0.0819	0.5129	0.0331	0.7426
1987	0.2873	1	0.0328	0.6816	2014	0.0693	0.3737	0.1821	0.5923
1988	0.0231	1	0.0043	0.9993	2015	0.169	0.5296	0.3011	0.6603
1989	0	NA	0.0665	0.6767	2016	0.2499	0.355	0.0598	0.3691

Table 1.2. Summary of average weight per tow in the Maine-New Hampshire Inshore Survey for fall and spring, 2001 to 2016. Estimates courtesy of Sally Sherman (MEDMR).

Maine-New Hampshire Inshore Survey				
Year	Spring (kg/tow)	CV_spring	Fall (kg/tow)	CV_fall
2001	0.49	2.50	0.31	1.32
2002	0.05	1.38	0.08	2.49
2003	0.22	1.16	0.02	0.84
2004	0.29	0.94	0.17	0.83
2005	0.27	0.86	0.12	0.65
2006	0.92	0.90	0.23	0.37
2007	0.52	1.00	0.15	0.88
2008	0.63	1.00	0.28	0.62
2009	0.90	0.75	0.73	0.64
2010	0.63	0.56	0.53	0.92
2011	0.64	0.56	0.53	0.49
2012	0.95	0.55	1.16	0.51
2013	0.39	0.73	0.31	0.45
2014	0.48	1.04	0.40	0.86
2015	0.35	1.62	0.64	1.05
2016	0.59	0.77	0.22	1.49

Table 1.3 Summary of CPUE analyses for Maine logbook data. Standardization model results are courtesy of working paper by Hansell, DeCelles and Cadrin (2017).

Year	Raw CPUE	Model Output/ Standardized CPUE	2 SE for standardized CPUE	Model+2SE	Model-2SE
2002	2.3938	2.4314	0.5691	3.0006	1.8623
2003	0.5306	0.5370	0.9593	1.4963	-0.4223
2004	1.1996	1.2205	0.7039	1.9243	0.5166
2005	1.8928	1.8943	0.6751	2.5694	1.2192
2006	1.3472	1.3562	0.7358	2.0920	0.6204
2007	0.8881	0.8948	0.6098	1.5046	0.2850
2008	1.1387	1.1503	0.6023	1.7527	0.5480
2009	0.7890	0.8009	0.5808	1.3817	0.2201
2010	0.7673	0.7839	0.5832	1.3671	0.2007
2011	1.5123	1.5208	0.5848	2.1056	0.9360
2012	1.5831	1.6115	0.5775	2.1890	1.0340
2013	1.5768	1.6041	0.5787	2.1827	1.0254
2014	1.7945	1.8177	0.5748	2.3926	1.2429
2015	1.5739	1.5739	0.5756	2.1495	0.9984
2016	1.9209	1.9435	0.6355	2.5790	1.3080

Table 1.4. Estimated average **discard to kept all ratios** for observed gill net and   trawl fishing trips, 1989-2016, originating from ports in New England (Rhode Island and north). Trips departing from ports in the Mid-Atlantic ports had negligible encounters with halibut over this period.

	Discard/kept_all indices								
	gear=100, region=NE					Gear=50, region=NE			
YEAR	gill net_d/k	CV_gill net	mean-SE	mean+SE	YEAR	trawl_d/k	CV_trawl	mean-SE	mean+SE
1989	7.58E-05	0.648	2.66E-05	1.25E-04	1989	0.00002	0.66587	0.00001	0.00003
1990	6.50E-05	0.415	3.80E-05	9.20E-05	1990	0.00007	0.89169	0.00001	0.00014
1991	5.09E-05	0.344	3.34E-05	6.84E-05	1991	0.00002	0.40011	0.00001	0.00003
1992	2.46E-05	0.357	1.58E-05	3.34E-05	1992	0.00001	0.51843	0.00001	0.00002
1993	4.18E-05	0.480	2.17E-05	6.19E-05	1993	0.00000	0.99321	0.00000	0.00001
1994	1.47E-05	0.743	3.78E-06	2.56E-05	1994	0.00001	0.61458	0.00001	0.00002
1995	1.92E-06	1.021	-3.98E-08	3.89E-06	1995	0.00005	0.84180	0.00001	0.00008
1996	5.11E-06	1.064	-3.27E-07	1.05E-05	1996	0.00001	0.67974	0.00000	0.00001
1997	0.00E+00		0.00E+00	0.00E+00	1997	0.00001	0.67339	0.00000	0.00001
1998	7.19E-06	1.013	-9.09E-08	1.45E-05	1998	0.00000		0.00000	0.00000
1999	1.85E-05	0.808	3.55E-06	3.34E-05	1999	0.00101	0.97915	0.00002	0.00200
2000	4.85E-05	0.912	4.26E-06	9.28E-05	2000	0.00011	0.31436	0.00008	0.00014
2001	3.82E-05	1.009	-3.49E-07	7.67E-05	2001	0.00011	0.27023	0.00008	0.00014
2002	0.00E+00		0.00E+00	0.00E+00	2002	0.00022	0.40807	0.00013	0.00032
2003	3.62E-04	0.498	1.82E-04	5.42E-04	2003	0.00012	0.16616	0.00010	0.00015
2004	9.50E-05	0.221	7.41E-05	1.16E-04	2004	0.00017	0.23997	0.00013	0.00021
2005	1.55E-04	0.338	1.02E-04	2.07E-04	2005	0.00019	0.10528	0.00017	0.00020
2006	2.23E-04	0.430	1.27E-04	3.19E-04	2006	0.00022	0.16604	0.00019	0.00026
2007	1.08E-04	0.322	7.29E-05	1.42E-04	2007	0.00015	0.12751	0.00013	0.00016
2008	1.20E-04	0.380	7.47E-05	1.66E-04	2008	0.00017	0.11345	0.00015	0.00019
2009	5.60E-05	0.364	3.56E-05	7.64E-05	2009	0.00023	0.12131	0.00020	0.00025
2010	2.82E-04	0.153	2.39E-04	3.25E-04	2010	0.00045	0.10346	0.00040	0.00050
2011	5.59E-04	0.241	4.24E-04	6.94E-04	2011	0.00065	0.06147	0.00061	0.00069
2012	5.67E-04	0.144	4.86E-04	6.49E-04	2012	0.00096	0.08508	0.00088	0.00104
2013	1.08E-03	0.159	9.10E-04	1.26E-03	2013	0.00110	0.08859	0.00101	0.00120
2014	9.01E-04	0.113	7.99E-04	1.00E-03	2014	0.00059	0.08587	0.00054	0.00065
2015	2.03E-03	0.176	1.67E-03	2.39E-03	2015	0.00050	0.11531	0.00044	0.00056
2016	2.19E-03	0.322	1.49E-03	2.90E-03	2016	0.00068	0.23433	0.00052	0.00084

Table 1.5. Estimated **average total halibut catch (landings +discard) to kept all** ratios for observed gill net and   trawl fishing trips, 1989-2016, originating from ports in New England (Rhode Island and north). Trips departing from ports in the Mid-Atlantic ports had negligible encounters with halibut over this period

total catch/kept_all indices								
gear=100, region=NE					Gear=50, region=NE			
YEAR	gill net_t/k	CV_gill net	mean-SE	mean+SE	YEAR	trawl_t/k	CV_trawl	mean-SE	mean+SE
1989	0.0001	0.4948	0.0001	0.0002	1989	0.0002	0.6536	0.0001	0.0003
1990	0.0003	0.3469	0.0002	0.0004	1990	0.0003	0.5359	0.0001	0.0004
1991	0.0003	0.2268	0.0002	0.0003	1991	0.0004	0.3055	0.0003	0.0006
1992	0.0001	0.2631	0.0001	0.0002	1992	0.0006	0.2623	0.0004	0.0007
1993	0.0002	0.3209	0.0001	0.0003	1993	0.0004	0.3417	0.0003	0.0005
1994	0.0001	0.5553	0.0000	0.0001	1994	0.0008	0.2602	0.0006	0.0010
1995	0.0000	1.0207	0.0000	0.0000	1995	0.0008	0.4758	0.0004	0.0012
1996	0.0000	0.8262	0.0000	0.0001	1996	0.0002	0.3312	0.0001	0.0002
1997	0.0000		0.0000	0.0000	1997	0.0001	0.3379	0.0001	0.0002
1998	0.0000	1.0126	0.0000	0.0000	1998	0.0001	0.9919	0.0000	0.0002
1999	0.0000	0.8078	0.0000	0.0000	1999	0.0010	0.9146	0.0001	0.0020
2000	0.0001	0.6977	0.0000	0.0001	2000	0.0002	0.2693	0.0002	0.0003
2001	0.0000	1.0091	0.0000	0.0001	2001	0.0002	0.3038	0.0002	0.0003
2002	0.0000		0.0000	0.0000	2002	0.0003	0.3523	0.0002	0.0004
2003	0.0005	0.4293	0.0003	0.0008	2003	0.0004	0.1332	0.0003	0.0004
2004	0.0002	0.1916	0.0002	0.0002	2004	0.0003	0.1612	0.0003	0.0004
2005	0.0005	0.3640	0.0003	0.0006	2005	0.0004	0.0911	0.0004	0.0004
2006	0.0003	0.3815	0.0002	0.0005	2006	0.0004	0.1285	0.0003	0.0004
2007	0.0002	0.2691	0.0002	0.0003	2007	0.0003	0.1130	0.0003	0.0004
2008	0.0002	0.3584	0.0002	0.0003	2008	0.0004	0.0951	0.0004	0.0005
2009	0.0001	0.4044	0.0001	0.0002	2009	0.0005	0.1029	0.0005	0.0006
2010	0.0004	0.1485	0.0003	0.0005	2010	0.0006	0.0923	0.0006	0.0007
2011	0.0007	0.2017	0.0006	0.0009	2011	0.0009	0.0583	0.0008	0.0009
2012	0.0007	0.1277	0.0006	0.0008	2012	0.0012	0.0767	0.0011	0.0013
2013	0.0015	0.1479	0.0013	0.0017	2013	0.0015	0.0864	0.0014	0.0016
2014	0.0013	0.1040	0.0011	0.0014	2014	0.0009	0.0753	0.0009	0.0010
2015	0.0026	0.1553	0.0022	0.0030	2015	0.0010	0.0917	0.0009	0.0011
2016	0.0028	0.2751	0.0020	0.0036	2016	0.0011	0.1722	0.0009	0.0012

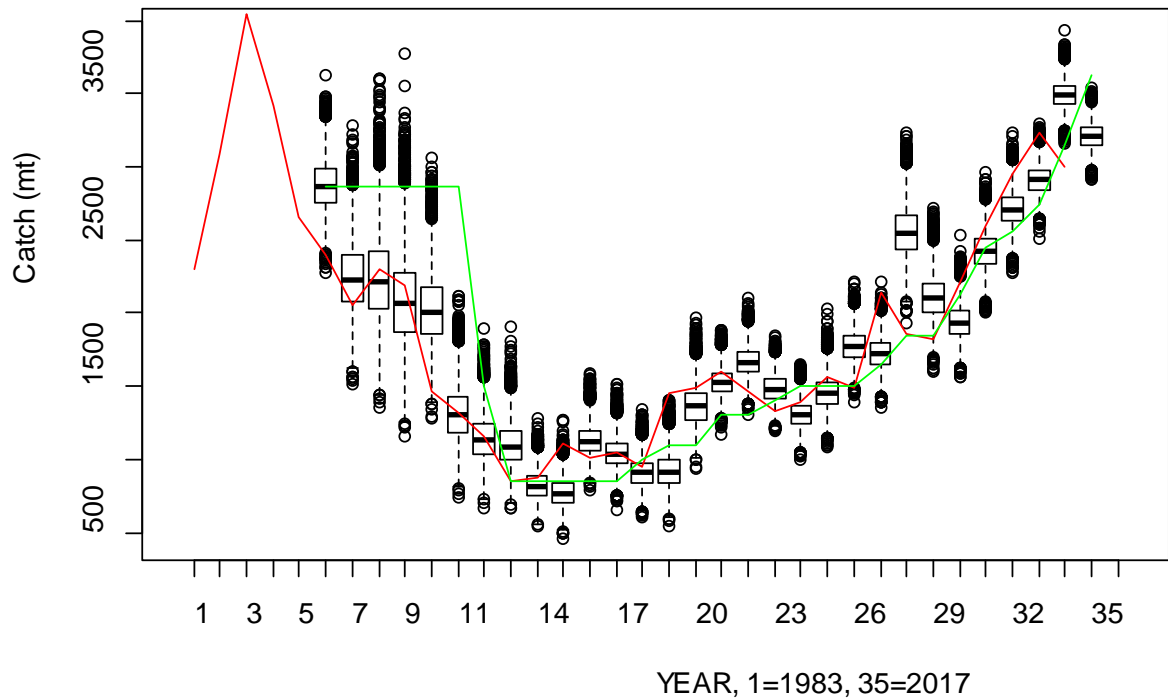**APPENDIX 2**.   Bootstrap analyses of DFO 3NOPs4WX5Zc

Figure. 2.1 Comparison of observed and predicted distribution of catches for FSD model applied to the DFO 3NOPs4VWX5Zc   Uncertainty estimates are based on a parametric bootstrap method described in the text. The gain parameters for proportional and derivative were set at Kp=0.75 and Kd=0.50, respectively. Red line is observed catch. Green line is Canadian TAC.

DFO survey and catch and TAC

1%	5%	10%	25%	50%	75%	90%	95%	99%
3016.2	3070.7	3101.9	3152.2	3211.0	3270.1	3321.4	3355.3	3412.6

Mean=3211.0, CV=0.026

Histogram of C.hat.boot[Tm,]
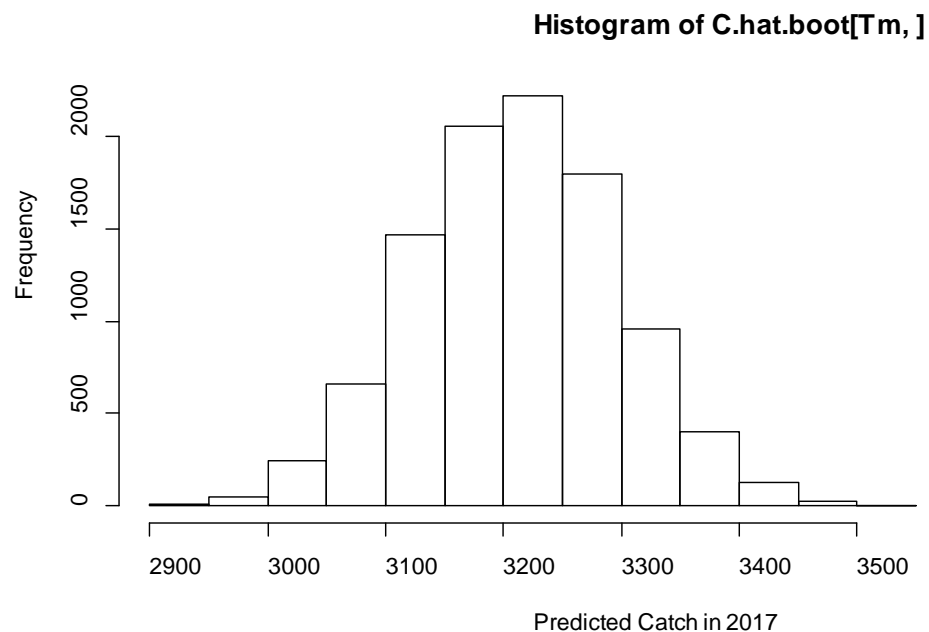
Figure 2.2 Distribution of forecasted catch for 2017. for FSD model applied to the DFO 3NOPs4VWX5Zc   Uncertainty estimates are based on a parametric bootstrap method described in the text. The gain parameters for proportional and derivative were set at Kp=0.75 and Kd=0.50, respectively

APPENDIX 3. Comparison of d/k to t/k ratios and implications for assessment.

The following set of graphs and figures represent results of using total halibut catch by weight per weight of total all species combined on observed trips. The t/k ratio was used as a measure of relative abundance instead of the d/k ratio. In general terms the t/k ratio mirrored the d/k ration for both gill nets and trawls (Fig. 3.1 and 3.2)
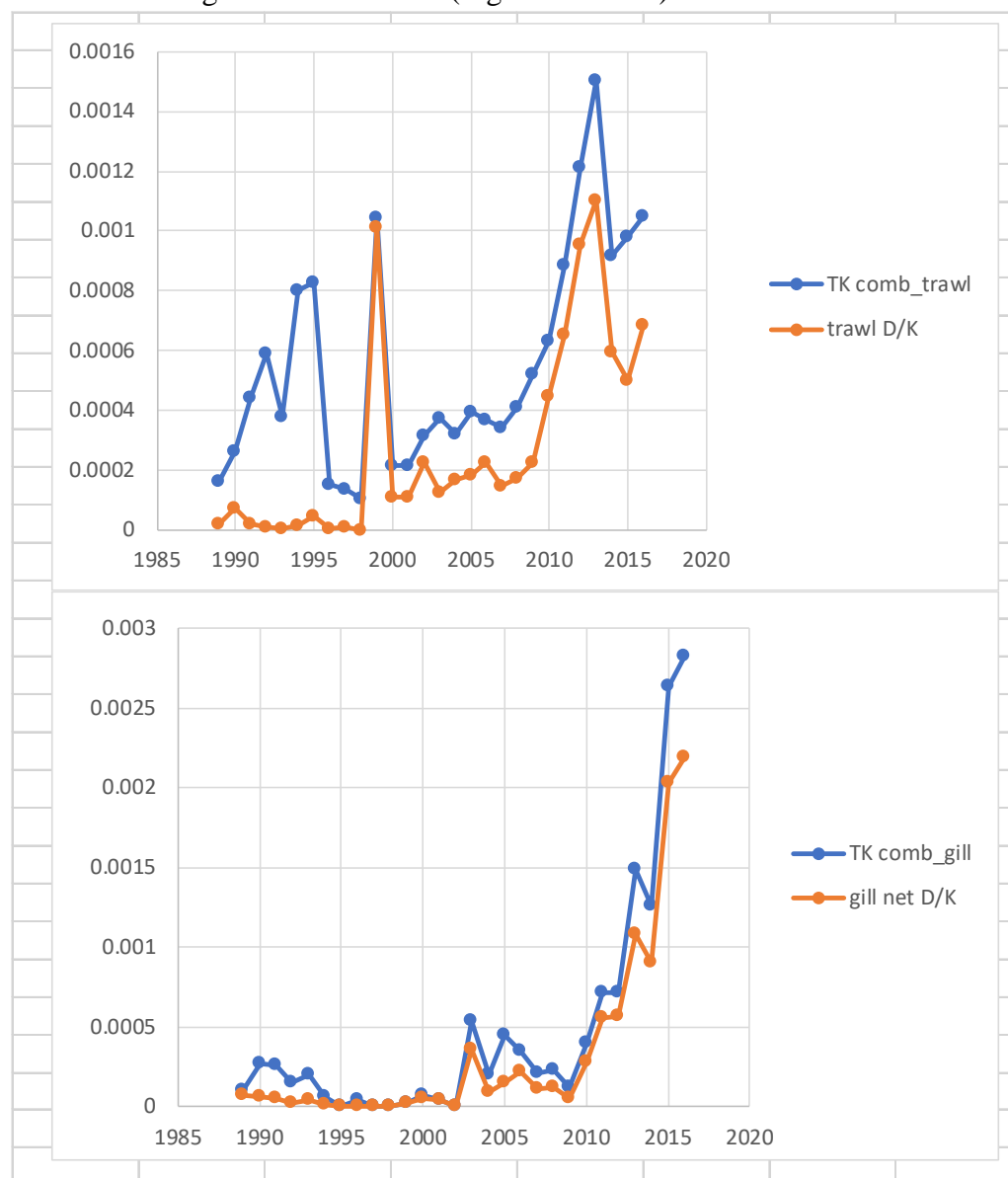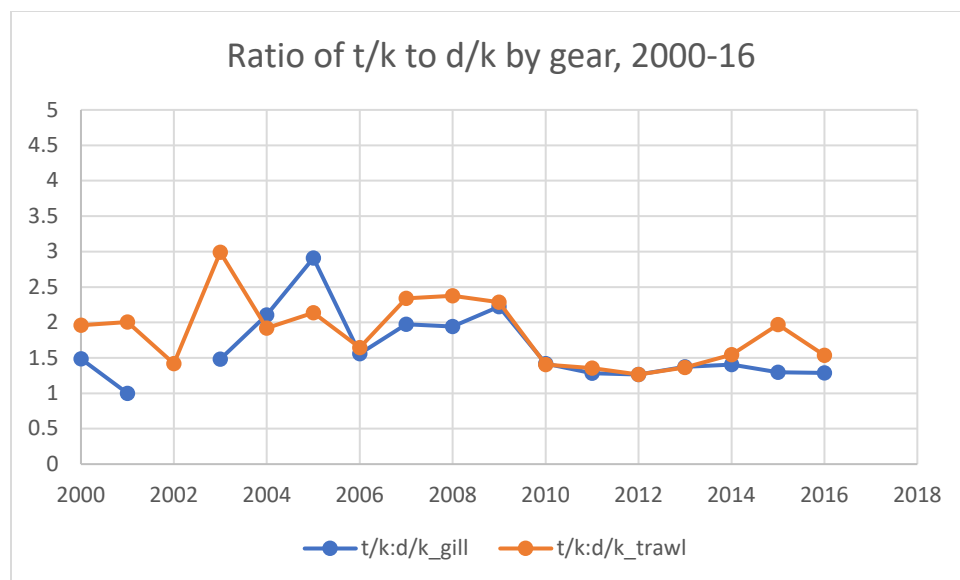

Fig. 3.1 Ratio of discard rates to total catch based on observer data

Fig. 3.2 Ratio of t/k to d/k for observed trips on gill net and trawl vessels.
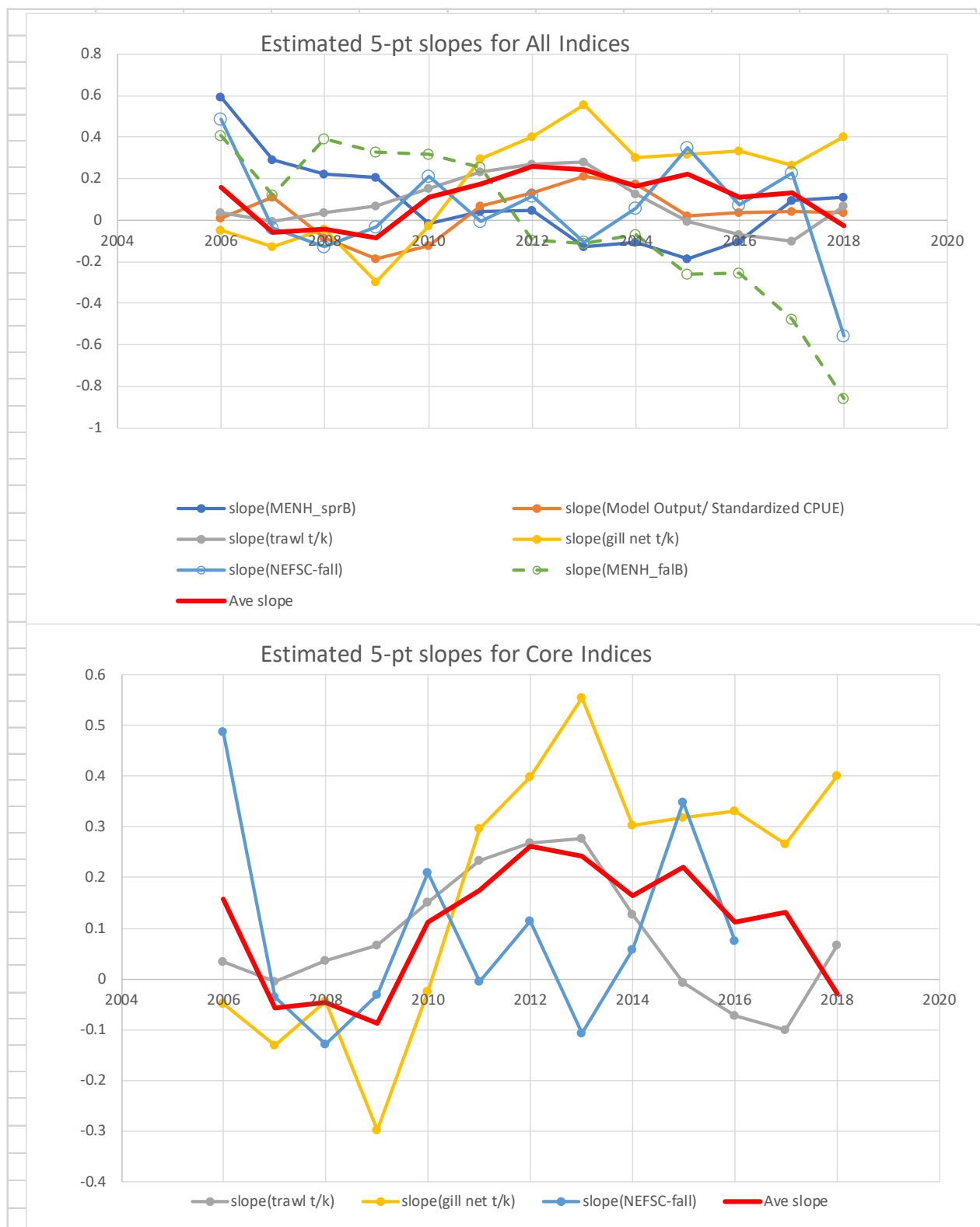
Figure 3.3 Trends in abundance indices based on 5 point regressions. Top plot shows full model with 5 indices. Bottom plot shows reduced model using only existing indices. **Model uses the t/k ratio for gill nets and trawls rather than the d/k ratio as measures of relative abundance**
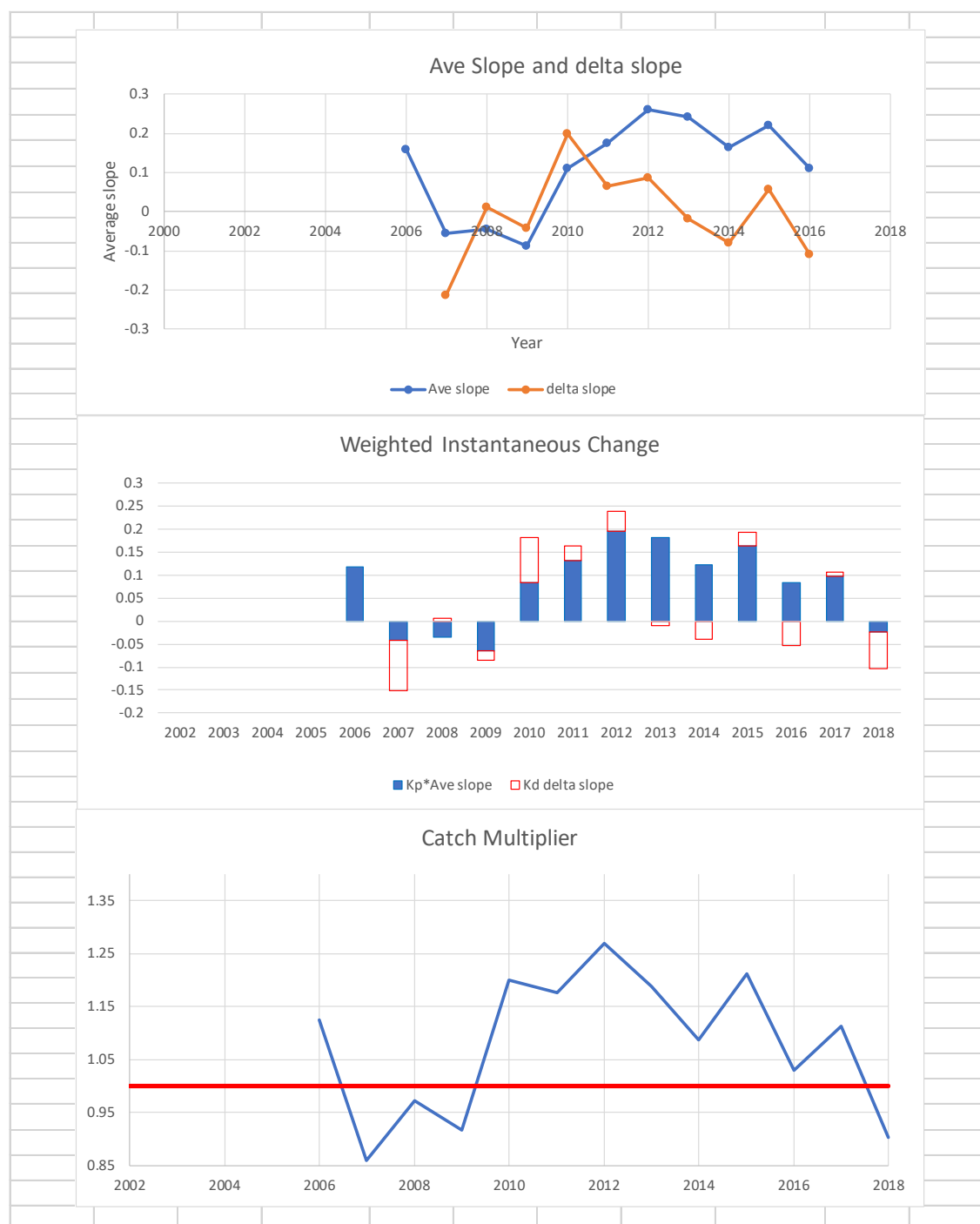
Figure 3.4. Summary of FSD model results for US Atlantic halibut based on d/k trawl, d/k gill net and NEFSC fall survey abundance indices (See Fig. 18 bottom). **Model uses the t/k ratio for gill nets and trawls rather than the d/k ratio as measures of relative abundance** Instantaneous rates of change represent Kp and Kd weighted values of the first and second derivative (top figure). The bottom figure show the Catch multiplier used to forecast catch. . The gain parameters for proportional and derivative were set at Kp=0.75 and Kd=0.50, respectively.
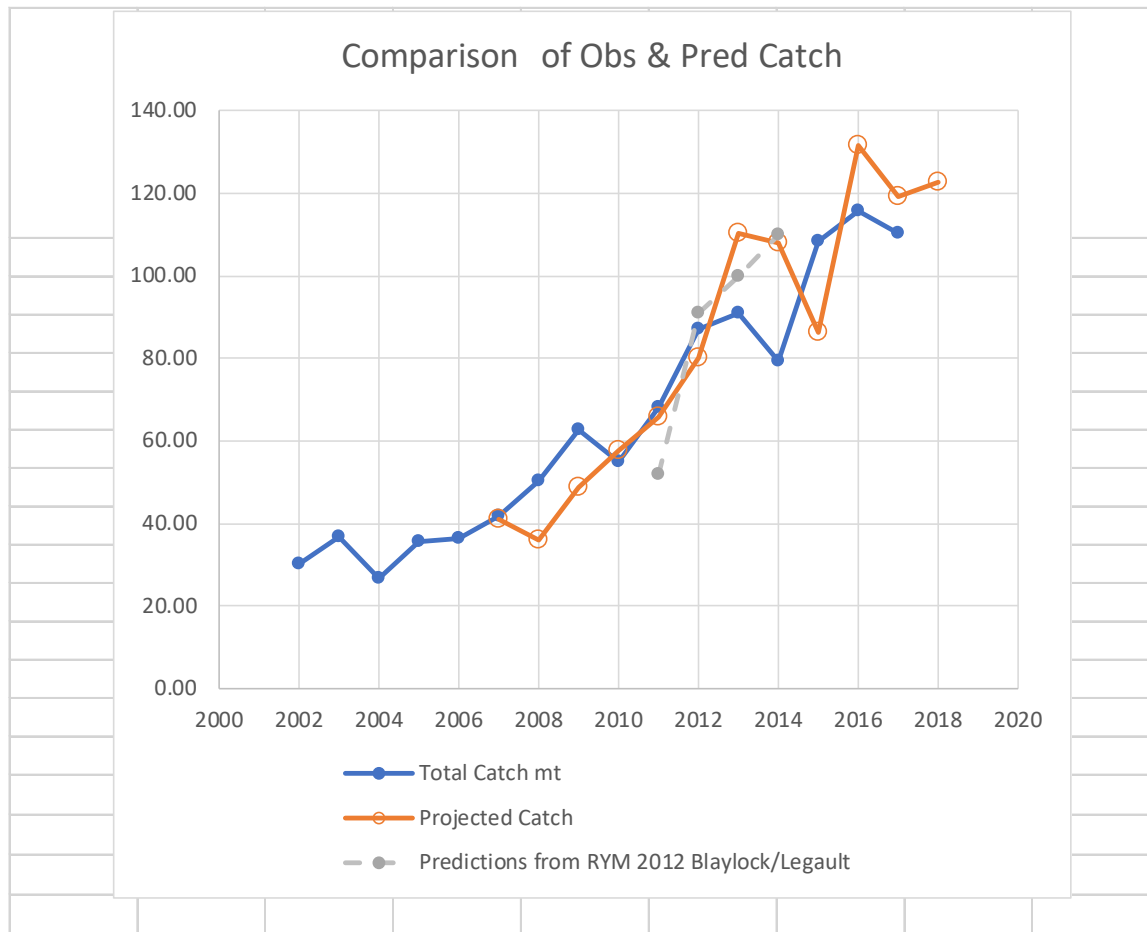
Figure 3.5. Comparison of observed vs predicted catches based on the FSD model applied to US stock area. **Model uses the t/k ratio for gill nets and trawls rather than the d/k ratio as measures of relative abundance**. . Forecasts from RYM application (Blaylock and Legault, 2012) are included for comparison. The gain parameters for proportional and derivative were set at Kp=0.75 and Kd=0.50, respectively.

APPENDIX 4. Summary of variable acronyms used in report.

Variable Name	Definition
ME_FAL_B	Average weight per tow in the Maine-New Hampshire bottom trawl survey conducted in Fall, 2001-2016.
ME_SPR_B	Average weight per tow in the Maine-New Hampshire bottom trawl survey conducted in Spring 2001-2016.
DK_GILL	Average ratio of weight of halibut discarded to total weight of all species kept (ie landed) for observed trips on gill net vessels departing from ports in New England Region (Rhode Island and north), 1989-2016.
DK_TRAWL	Average ratio of weight of halibut discarded to total weight of all species kept (ie landed) for observed trips on trawl vessels departing from ports in New England Region (Rhode Island and north), 1989-2016.
TK_GILL	Average ratio of total weight of halibut kept plus discarded to total weight of all species kept (ie landed) for observed trips on gill net vessels departing from ports in New England Region (Rhode Island and north), 1989-2016.
TK_TRAWL	Average ratio of total weight of halibut kept plus discarded to total weight of all species kept (ie landed) for observed trips on trawl vessels departing from ports in New England Region (Rhode Island and north), 1989-2016.
NMFS_SPR_B	Average weight per tow in the Northeast Fisheries Science Center bottom trawl survey conducted in Spring, 1968-2016.
NMFS_FAL_B	Average weight per tow in the Northeast Fisheries Science Center bottom trawl survey conducted in Fall, 1963-2016.
LL_CPUE	Model adjusted estimates of longline fishing catch per unit effort for commercial fishing vessels in Maine, 2002-2016. Based on work of Hansell et al. 2017
MEDISCRAT	Halibut discard rate reported by Maine longline harvesters
CP100TAGS	Total catch of halibut per 100 tags issued by Maine DMR
DFO_FS	Average weight per set for halibut captured in a scientific long line survey conducted jointly by DFO and commercial fishermen, 1998-2016. Fixed stations only. Based on results of GLM.
DFO_CPUE	Average weight per set for halibut reported by commercial fishermen. CI index is just mean and se of all ci sets (600*catchkg/(#hooks)/(DURATION), last reported in data update for 2014 asssessment, 1998-2016
SURV_3NOP	Average number per tow for halibut captured in DFO bottom trawl survey in NAFO area 3NOPs in spring, 1971-2013.
SURV_4VWX	Average number per tow for halibut captured in DFO bottom trawl survey in NAFO area 4VWX in summer,1970-2016
DFO_TOTB	Total biomass estimate derived from analytical model 1970-2013.