Fisheries and Oceans Canada Pêches et Océans Canada

NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT	***TRAC***
Comité d'évaluation des ressources transfrontalières	**Transboundary Resource Assessment Committee**
Document de référence 2016/04	**Reference Document 2016/04**
Ne pas citer sans autorisation des auteurs	Not to be cited without permission of the authors

Developing an Empirical Approach for Providing Catch Advice for Eastern Georges Bank Cod (TOR 4)

Elizabeth N. Brooks[1], Irene Andrushchenko[2], Yanjun Wang[2], Loretta O'Brien[1]

[1]NOAA/NMFS Northeast Fisheries Science Center
166 Water Street
Woods Hole, Massachusetts 02543
USA

[2]Fisheries and Oceans Canada
531 Brandy Cove Road
St. Andrews, New Brunswick E5B 3L9
Canada

Canadä

(page intentionally left blank)

ABSTRACT

An empirical method was developed for providing quota advice for Eastern Georges Bank cod. This method adjusts recent quotas by recent population abundance trends. The average of three surveys (DFO spring, NMFS spring, and NMFS fall) is fit by a *loess* smoother, and the slope in 3 year intervals is calculated (on log-scale). The slope is used to adjust recent quotas. Uncertainty is characterized quantitatively by a bootstrap analysis on the fit of the *loess* smoother, and qualitatively with a table of secondary indicators. The estimated slope from the most recent 3 survey years (2014-2016 for DFO and NMFS spring; 2013-2015 for NMFS fall) was applied to the average quota for years 2013-2015 (650 mt) to provide a range of quota advice for 2017.

Introduction

A benchmark assessment for cod in the Eastern Georges Bank management unit was held in 2013. Two model formulations were presented, one based on virtual population analysis (VPA) and one based on a statistical catch at age model (ASAP). The primary difference between the models was the assumption about natural mortality (M)—the ASAP model assumes that M= 0.2 at all ages and for all years, while the VPA assumes that M changes from 0.2 to 0.8 for ages 6+ beginning in 1994. There was no consensus on the most appropriate model. However, it was agreed that catch advice would be based on the VPA analysis including a consequence analysis comparing the risks of implementing either the VPA or ASAP catch advice in each of the models (Claytor and O'Brien 2013, Wang and O'Brien 2013).

The lack of consensus on an assessment model, the corresponding complexity of adequately characterizing risk, and the growing disparity between catch advice for the eastern portion of Georges Bank (EGB) versus the whole of Georges Bank, has been a source of consternation for both scientists and managers. This came to a head in 2015 when there was a possibility that catch recommendations for cod in the eastern Georges Bank management unit could exceed those recommended for the management unit on the whole of Georges Bank. In the end, that scenario did not occur in 2015. However, the threat of this potential imbalance in catch advice and the divergent advice for EGB cod led the TMGC to express concern about the difficulty of reconciling the advice from the VPA "M 0.8" model and the sensitivity analysis from the ASAP "M 0.2" model (TMGC Guidance Document 2015/01). In addition, at the Canada-US Steering Committee meeting held in Dartmouth, NS, in September 2015, the Steering Committee Co-chairs encouraged TRAC to explore other analyses for Cod that would assist further informing catch advice given the very different results yielded by the VPA and ASAP model approaches.

This reference document describes an empirical method to providing catch advice, which shares many similarities with the method used to provide catch advice for the Georges Bank management unit during the 2015 update assessment (O'Brien 2015). The main idea of the proposed approach is that quota advice is derived by adjusting recent quotas by recent trends in the resource. While this approach requires no explicit assumption about M, recent quotas have been negotiated based on the VPA results, thus the empirical model will be implicitly linked to recent VPA output for the first few years of application.

The methods were developed collaboratively between the TRAC scientists. Decisions about the details of the method are summarized, and where alternatives were considered, the justification for those decisions are provided. The empirical method does not have a means of performing projections or characterizing risk in the same manner as the VPA or the ASAP models. Risk will be characterized both quantitatively (by bootstrapping the smoother) as well as qualitatively by considering secondary indicators, which are tabulated below (similar to the risk table for Georges Bank yellowtail flounder (see Legault *et al.* 2015). The reader is also referred to Brooks *et al*.(2016) where biological and fishery indicators are presented.

Methods

1. Identify how surveys will be used

Series included

Three indices of abundance are included in this analysis: the DFO spring, the NMFS spring and NMFS fall surveys.

Series length

The three indices began at different points in time, however, 1987 was chosen as the starting point for this analysis because it is the first year where all 3 surveys had complete coverage. This is important, so that at any point in the time series, the trend reflects a consistent measure of the population.

Series units (Biomass/numbers)

The three surveys can be examined in either units of biomass or in numbers of fish. Because the quantity being adjusted is in biomass units, it is appropriate to examine survey trends in the same units.

Ages included in surveys

Different age range combinations were considered for the empirical approach (ages 1-8, 3+, 3-8, 1+, etc.). Total survey biomass (1+) was chosen. It was noted that fishery catch on eastern GB is dominated by ages 3-5 in the most recent years and the proportions at age of cod captured by the fishery are within the bounds of those sampled by the three surveys (Figures 1a-b). Also, analytical CV estimates are not available for age-range specific biomass indices, which would reduce the ability to characterize annual uncertainty in the index.

Normalizing and CV weighting

The three survey indices were normalized prior to combination by scaling each series to its mean over the most recent 30 years (1987-2016 for spring surveys, 1986-2015 for fall survey). By putting all surveys on the same relative scale, it removes the effect of averaging indices with different magnitudes caused by different catchabilities. However, averaging the three surveys ignores differences in selectivity and catchability, and effectively considers trends from all three surveys to be equally representative of changes in the population. It should be noted that there is a lack of older fish in the fall survey catch.

The combined survey biomass index consists of the NMFS ($S_{(Y)}$) and DFO ($D_{(Y)}$) spring surveys in a given year and the NMFS fall survey ($F_{(Y-1)}$) from the previous year [1]. Given the timing of these three surveys on eastern Georges Bank (~October for NMFS fall, ~February for DFO, and ~April for NMFS spring), the average survey is expected to correspond to January 1 population biomass. Thus, average relative biomass index in year Y reflects the population immediately after fishery catch for year Y-1 has been removed from the population.

$$\overline{I_Y} = f[F_{(Y-1)}, S_{(Y)}, D_{(Y)}] \qquad [1]$$

Because the precision for any index value in any given year varies, a straight arithmetic average was not taken. Instead, a weighted average was calculated with inverse coefficient of variation (CV) weights [2] so that, in a given year, more precise points were given more weight than less precise points.

$$\bar{I}_{(Y)} = \frac{\left({F_{(Y-1)}}/{CV_{F_{(Y-1)}}} + {S_{(Y)}}/{CV_{S_{(Y)}}} + {D_{(Y)}}/{CV_{D_{(Y)}}} \right)}{\left({1}/{CV_{F_{(Y-1)}}} + {1}/{CV_{S_{(Y)}}} + {1}/{CV_{D_{(Y)}}} \right)} \qquad [2]$$

An alternative to using inverse CV weights is inverse variance weights. However, unlike CVs, variance is not scale-invariant and the average resulting from using inverse variance weights would be pulled towards the lower survey values. TRAC scientists believed that using inverse CV weights was the most appropriate choice for averaging the three surveys.

The CV associated with this average index reflects the uncertainty associated with the annual variations and mean from each survey. For the survey biomass *(B*) from survey *i*, the variance of inverse of *n* years of survey mean was calculated as:

$$var(1/\bar{B_i}) = \frac{1}{n^2}\sum_{y=1}^{nyear} var(B_{i,y}) \,/\bar{B_i}^4 \qquad [3]$$

The variance of normalized survey indices (*I*) for survey *i* in year *y* was calculated as (Goodman, 1960):

$$var\left(I_{i,y}\right) = B_{i,y}^2\, var(1/\bar{B_i}) + \frac{1}{\bar{B_i}^2} var\left(B_{i,y}\right) + var(1/\bar{B_i}) var\left(B_{i,y}\right) \qquad [4]$$

Fitting a loess smoother

A *loess* smoother (Cleveland, 1979) will be fit to the average index of abundance. The resulting smooth through the survey data will be used to calculate a slope (on log scale) for the most recent three year interval. Estimating this slope from the smooth, rather than the average index itself, is based on the assumption that the smooth reflects the main features of population trend without being pulled sharply by interannual variability. Slopes are calculated on log scale so that the estimated trend is expressed as a proportion – i.e., a slope of 1 indicates no change in population trend, a slope less than 1 implies a decreasing population trend, while a slope greater than 1 implies an increasing trend. Calculating slopes on log scale also avoids negative scalars being applied to adjust quotas. A method with these same fundamental features was presented in Garamont and Butterworth (2015), and reviewed by Carruthers et al. (2016) in a summary of the performance of simple management procedures.

Loess diagnostics

Applying a *loess* smoother requires specifying the degree of polynomial, the “family” used for applying the local smooth, the number of iterations for calculating the smoothed value at each point, the span (fraction of available data used in calculating the fit at each point), and the option to specify additional weights when evaluating local fitting.

The degree of polynomial, d, is an integer value with $d \geq 0$. A polynomial with d = 0 assumes a locally constant function, d = 1 assumes a locally linear function, d = 2 assumes locally quadratic, and so on. Selecting the degree of polynomial should coincide with the amount of flexibility required to capture the main trends in the data being smoothed, and the main criterion for evaluating the appropriateness of the degree is by visually inspecting the result. Cleveland (1979) only discusses polynomials of degree 0 (locally constant) and degree 1 (locally linear), and concludes that d=1 should be sufficient for most cases because higher degrees introduce computational difficulties. This computational limitation no longer exists due to advances in computers and software. More recently, Jacoby (2000) suggested that only $d = 1$ or $d = 2$ be evaluated as degrees higher than 2 were not likely to improve local fits of the smoother and were also not parsimonious, and recommended a general guideline that if the data indicated curvilinearity then the degree should be 2. After examining the data (the average index calculated with inverse CV weights), the TRAC agreed that the appropriate degree was 2.

There are two options available for fitting the local smooth (option “family” in the *loess* package in R; R Core Team, 2016): family=Gaussian or family=symmetric. Specifying family=Gaussian will result in a local least squares fit, and assumes normality in the residual errors. As a result of the normality assumption, there are a variety of diagnostic plots that can be examined to assess fit, for example, qq plots to check the normality assumption, and residuals plotted against year to assess homoscedasticity (Cleveland and Devlin, 1988). If family=symmetric is specified then the fitting is a robust local regression procedure with multiple iterations, where residuals from the first iteration (which is a local least squares fit) are used as additional weight factors in a second iteration, and so on. There is no assumption about normality of residual errors in this case, so there are no formal diagnostics to evaluate fit, just visual inspection. The TRAC agreed that there was value in looking at both fits, as the Gaussian option offers straightforward diagnostics, while the symmetric option offers potentially more robust fits by downweighting outliers.

Choosing the number of iterations is only relevant if the family is specified to be symmetric (i.e. use robust regression). Cleveland (1979) recommended 2 iterations based on experimentation with real and artificial data. The TRAC agreed that 2 iterations (beyond the first initial least squares fit) would be specified for the robust fitting routine.

The span refers to the proportion of data considered in the local fitting procedure at each point. For example, if the span=0.1, then 10% of the data would be used, while span=0.75 would use 75% of the data. A large span will produce a smoother curve while a smaller span will closely follow the ups and downs in the data being fitted. The choice reflects typical bias/variance trade-offs, and to some extent the choice can be based on fits by eye. As noted above in the discussion regarding choice of family (Gaussian or symmetric), if errors are assumed to be normally distributed (family=Gaussian), then the choice of span can be based on the value that produces the best diagnostics (qq plot and plot of residuals versus year). The TRAC approach to selecting the appropriate span was to use diagnostics to select the best fit when the Gaussian option was used, and to compare a range of spans for the symmetric option.

Default weights in the *loess* algorithm are tricubic weights, which assigns higher weights to points closer to each observation within the span about that point. The group considered whether additional weighting factors, based on the inverse CV at each point in the average index, should be specified. The documentation for the package *loess* did not describe how the additional optional weights were implemented, and the literature suggested that the default weights were recommended (Cleveland,1979; Jacoby, 2000). The TRAC compared results with and without specifying the additional weight factors and found no real difference in the fits; therefore, the TRAC agreed to proceed by using the default tricubic weights.

Alternative analyses considered but not pursued

Use of age-range specific biomass indices was explored but not pursued, as lack of analytical CV estimates for age specific indices made their combination problematic.

Initially, concerns were raised about combining surveys with known differences in age contributions (e.g. fall survey has historically caught younger fish), and using this combination to advise fishery catch. Further examination of proportional biomass at age contributions of the surveys and fishery indicated that the biomass at age composition between the four data sources has not differed drastically in recent years (Figure 1a-b).

2. Identify and explain an appropriate starting point (catch amount or quota) for applying the empirical approach. To the extent possible, characterize uncertainties and sensitivities.

Slope calculation

In calculating the slope from the fitted *loess* curve, the TRAC discussed the appropriate number of years that should be used. This decision reflects a balance between having a sufficient number of points to estimate the slope with some precision, and restricting those points to 'recent years' so that the population that produced that trend still comprises the population to be harvested in the recommended catch. The TRAC considered three, four, or five years for calculating the slope. Considering that the years being averaged would impact catch advice 2 years later, and the fact that catch is mainly composed of fish aged 3-5, the TRAC agreed that three years was the appropriate number of years to use in estimating the slope.

Starting point for applying empirical approach

There were two possible quantities that could be adjusted by the estimated slope: recent catch or recent quota. The TRAC debated the merits of each. It could be argued that catch is an appropriate choice, because the population size, and changes between years, are directly affected by catch rather than quota. However, choosing catch as the quantity to be adjusted implies that population abundance is the limiting factor when catches<quota in a given year. The alternative option is to adjust recent quotas, which is consistent with the approach used for the Georges Bank Yellowtail stock (Legault et al. 2015). The argument for adjusting quotas is that it removes the dependence on fishing effort, which can be driven by factors other than population abundance. Although it was concluded that adjusting quota is a more appropriate approach, it was noted that since recent catches have been 77-98% of quota, adjusting one or the other was not expected to produce very different quota advice.

Quantitative risk characterization

The uncertainties about the estimated slope were derived using the bootstrap method (Efron, 1979). In this approach, the bootstrap replicates are obtained by resampling with replacement from the *loess* fitted survey biomass residuals and adding these to the *loess* model predicted values for the survey biomass index; 1000 bootstrap replicates were sampled.

Due to non-linearity associated with *loess* and log-transformed regression, estimation bias is expected. The bias of the slope estimates was calculated as the difference between the mean of all bootstrapped samples and the point estimate from the fit to the original data. The bias-corrected percentile method (Efron, 1982) was used to determine the confidence interval about the estimated slope. This method improves on the percentile method by adjusting for differences between the median of the bootstrap samples and the estimate from the original data sample.

An alternative characterization of risk was explored, which derived a confidence interval about the estimated slope using the standard error of that estimate. However, this approach was deemed unacceptable; with only 3 points from which to estimate the slope, the SE was very large and thus the 25th and 75th percentiles for the slopes were extremely broad, at times producing conflicting advice (i.e., 25th percentile suggested a dramatic reduction in quota, while the 75th percentile suggested an increase in quota).

Qualitative risk characterization (Table of secondary indicators)

The quantitative characterization of risk reflects uncertainty in the fit of the *loess* smooth. The TRAC felt it was important to consider qualitative indicators that reflect the condition of the stock and recent performance of the models. The TRAC agreed to produce a table of reasons to adjust catch advice, similar to what has been provided for Georges Bank yellowtail flounder, and similar to what is being presented in the haddock interim report.

3. Recommend thresholds for annual increases/decreases in catch advice.

The TRAC considered the introduction of a threshold, or cap, on annual increases or decreases in the catch advice. The purpose of a threshold was to maintain some stability in catches and avoid large fluctuations that could be due to a year effect, especially given that only 3 years are being used to estimate the slope. The TRAC agreed that specifying a cap was appropriate, and propose 20% as the maximum amount by which catch could increase or decrease between years. This is the same value specified in the control rule for a Management Strategy Evaluation of western component Pollock (DFO, 2011).

Results and Discussion

The three normalized indices are reported in Table 1 and are plotted in Figure 2. The NMFS fall index decreased in 2015, and has mostly been fluctuating below the series average since 2005. The DFO index in 2016 remained comparable to its 2015 value, while the NMFS spring and fall both showed substantial increases from 2015 to 2016. The combined index of abundance increased in 2016 from its series low in 2015 (Table 1, Figure 3).

Two *loess* fits were made to the combined index, one with local least squares (the Gaussian family assuming normal errors) and one with robust least squares with two additional iterations to downweight outliers. For the Gaussian fit, a span of 0.3 had the best diagnostics (Figure 4), and the same span seemed appropriate for the robust fit as well. The fits of both *loess* smooths is shown in Figure 5. The fits are very similar, and only the years immediately around 1990 have non-negligible differences. The difference is due to the robust iterations downweighting the very high point in 1990. This high point is driven by the DFO survey in that year, where several very large tows were observed in survey stratum 5Z2. Confidence intervals from 1000 bootstrap replicates for the *loess* fit are shown in Figure 6.

For all 1000 bootstrapped fits of the *loess*, slopes were estimated on log scale in 3 year blocks. The slopes were then transformed to arithmetic scale, and bootstrap bias correction was applied. The resulting 5th-25th-median-75th-95th percentiles for the estimated slope are shown in Figure 7. The values of the median slope, and other percentiles, are reported in Table 2. Similar to the result of fitting the two *loess* smooths, the estimated slopes are extremely similar except for a small window around 1990, where the two fits differ in whether to treat that point as an outlier.

The recent three year average of quotas is 650 mt (600 mt in 2013; 700 mt in 2014; 650 mt in 2015). The median of bootstrap bias-adjusted slope estimates for survey years 2013-2015 (NMFS fall) or 2014-2016 (DFO and NMFS spring) is 1.03 for the least squares fit and 1.06 for the robust regression fit. The [25th, 75th] percentiles are [0.94, 1.14] and [0.99, 1.17] for the local least squares and robust least squares, respectively. The differences in distributions are related to the differences in residuals between the two smoother applications. Applying the median slope to the recent average quota (650 mt) produces a 2017 quota of 669.5 or 689 mt for the least squares and robust least square fits, respectively. Uncertainty due to the fit of the *loess* curve, and its impact on catch advice, can be characterized by multiplying the 25th and 75th percentiles of bootstrapped slope estimates to the recent average quota. This produces quota advice of 611 or 643.5 (25th percentile for local least squares or robust least squares) and 741 or 760.5 (75th percentile for local least squares or robust least squares) (Table 3).

Qualitative indicators of risk, which are related to stock condition, productivity, and recent model performance, are summarized in Table 4. These indicators suggest a low risk strategy is appropriate.

Conclusions/Summary/Recommendations

At the 2016 TRAC meeting, the robust least squares loess smooth was used for catch advice. The TRAC recommends that low risk quotas are appropriate for the cod resource. Productivity, which includes growth and recruitment, is low, and the stock has shown no signs of rebuilding.

Literature Cited

Brooks, L., Andrushchenko, I., Wang, Y. and L. O’Brien 2016. Biological and Fishery Indicators and Models Projection Performance for Eastern Georges Bank Cod (TOR 1 and 2). *TRAC Ref. Doc.* 2016/03.

Carruthers, T.R., L.T. Kell, D.S. Butterworth, M.N. Maunder. H.F. Geromont, C. Walters, M.K. McAllister, R. Hillary, P. Levontin, T. Kitakado, C.R. Davies. 2016. Performance review of simple management procedures. ICES J. Mar. Sci. 73(2): 464-482.

Claytor, R., and L. O'Brien. 2013. Transboundary Resources Assessment Committee Eastern Georges Bank cod benchmark assessment. TRAC Proceedings 2013/01, 29p.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. J.A.S.A. 74(368): 829-836.

Cleveland, W.S and Devlin, S.J. 1988. Locally weighted regression: An approach to regression analysis by local fitting. J.A.S.A. 83(403):596-610.

DFO. 2011. Western Component (4Xopqrs5) Pollock Management Strategy Evaluation. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2011/054.

Efron, B. 1979. Bootstrap methods: another look at the jackknife. *Ann. Statist*. 7: 1-26.

Efron, B. 1982. The Jackknife, the Bootstrap, and Other Resampling Plans. Philadephia: Society for Industrial and Applied Mathematics. 38: 92p

Legault, C.M., Alade, L., Busawon, D., and Stone, H. (2015). Stock assessment of Georges Bank Yellowtail Flounder for 2015. TRAC Reference Document 2015/01.

Garamont, H.F. and D.S. Butterworth. 2015. Complex assessments or simple management procedures for efficient fisheries management: a comparative study. ICES J. Mar. Sci. 72(1):262-274.

Goodman A. L. 1960. On the exact variance of products. J. of the American statistical association. Vol. 55, 1960: 708-713.

Jacoby, W.G. 2000. Loess: a nonparametric, graphical tool for depicting relationships between variables. Electoral Studies 19:577-613.

O'Brien, L. (2015). Georges Bank Atlantic cod in Northeast Fisheries Science Center. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 15-24. 251 p.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Transboundary Management Guidance Committee. 2015. Guidance Document. 2015/01.
Wang, Y., and L. O'Brien. 2013. 2013 Benchmark Assessment of Eastern Georges Bank Atlantic Cod. TRAC Ref. Doc. 2013/01, 62p.

Table 1. Normalized ('Norm') swept area biomass indices and their associated coefficients of variation (CV), and the combined index and its CV.

	NMFS Fall		NMFS Spring		DFO		Combined	
Year	Norm	CV	Norm	CV	Norm	CV	Norm	CV
1987	0.87	0.51	1.14	0.27	0.57	0.36	0.89	0.21
1988	1.05	0.50	1.39	0.35	1.26	0.32	1.26	0.21
1989	2.71	0.48	1.81	0.32	0.95	0.21	1.58	0.21
1990	2.06	0.45	1.62	0.32	3.68	0.28	2.56	0.20
1991	2.05	0.73	1.41	0.19	1.63	0.25	1.57	0.18
1992	0.17	0.56	0.91	0.22	0.95	0.25	0.80	0.16
1993	0.69	0.49	0.99	0.33	1.08	0.23	0.96	0.18
1994	0.24	0.61	0.19	0.33	0.85	0.51	0.40	0.34
1995	0.69	0.85	1.51	0.50	0.53	0.35	0.88	0.33
1996	0.49	0.65	0.99	0.43	2.09	0.31	1.38	0.24
1997	0.72	0.47	0.61	0.24	0.72	0.27	0.67	0.18
1998	0.75	0.99	1.84	0.48	0.33	0.24	0.81	0.34
1999	1.19	0.33	0.79	0.26	0.60	0.36	0.86	0.19
2000	0.42	0.69	1.10	0.28	2.10	0.46	1.27	0.27
2001	0.36	0.39	0.56	0.46	1.17	0.45	0.68	0.28
2002	0.46	0.61	0.66	0.28	1.32	0.43	0.82	0.25
2003	4.61	0.69	0.96	0.55	0.40	0.20	1.25	0.45
2004	0.24	0.48	3.15	0.63	0.37	0.30	0.96	0.48
2005	3.34	0.80	0.66	0.25	1.70	0.60	1.40	0.40
2006	0.58	0.48	1.10	0.28	0.82	0.30	0.87	0.19
2007	0.87	0.81	0.91	0.30	0.73	0.28	0.82	0.21
2008	0.17	0.42	0.80	0.28	0.89	0.29	0.67	0.19
2009	0.32	0.35	0.65	0.33	1.51	0.68	0.69	0.33
2010	0.48	0.43	0.54	0.24	1.71	0.66	0.75	0.31
2011	0.29	0.44	0.26	0.37	0.55	0.27	0.39	0.20
2012	0.92	0.73	0.73	0.28	0.16	0.22	0.49	0.26
2013	0.24	0.43	1.44	0.63	0.72	0.50	0.73	0.38
2014	1.03	0.68	0.49	0.33	0.16	0.29	0.44	0.32
2015	0.55	0.54	0.26	0.24	0.23	0.39	0.32	0.24
2016	1.43	0.44	0.54	0.24	0.24	0.25	0.62	0.24

Table 2a. Estimated slope from the local least squares fit (bc=bias corrected; bs=bootstrap) for 5-25-50-75-95 percentiles, for both the local least square (a) and robust least square (b) *loess* fits.

Year	Slope	Mean_bs	bias_adj	bc_0.5	bc_0.05	bc_0.95	bc_0.25	bc_0.75
1991	0.97	0.96	0.98	0.98	0.92	1.03	0.96	1.00
1992	0.77	0.80	0.74	0.74	0.67	0.85	0.70	0.79
1993	0.62	0.70	0.54	0.56	0.49	0.73	0.50	0.62
1994	0.81	0.79	0.84	0.83	0.70	0.93	0.78	0.89
1995	1.17	1.06	1.28	1.36	1.07	1.48	1.19	1.46
1996	1.17	1.17	1.18	1.25	0.98	1.55	1.13	1.38
1997	1.02	1.06	0.99	0.96	0.90	1.14	0.93	1.04
1998	0.93	0.99	0.87	0.91	0.87	1.07	0.90	0.94
1999	0.99	1.00	0.99	1.01	0.83	1.22	0.92	1.07
2000	1.06	1.01	1.11	1.08	0.94	1.21	1.05	1.17
2001	1.00	1.01	0.99	0.99	0.92	1.08	0.95	1.04
2002	0.98	1.02	0.93	0.93	0.87	1.07	0.90	0.99
2003	1.06	1.08	1.04	1.05	0.90	1.27	0.98	1.13
2004	1.15	1.10	1.20	1.21	1.06	1.29	1.15	1.25
2005	1.05	1.04	1.07	1.07	0.97	1.17	1.03	1.12
2006	0.91	0.94	0.89	0.91	0.84	1.00	0.87	0.95
2007	0.84	0.87	0.82	0.81	0.78	0.88	0.79	0.85
2008	0.84	0.86	0.83	0.83	0.78	0.93	0.80	0.88
2009	0.92	0.89	0.96	0.95	0.88	0.98	0.92	0.97
2010	0.92	0.91	0.92	0.92	0.82	1.01	0.90	0.98
2011	0.89	0.92	0.85	0.86	0.82	0.96	0.84	0.89
2012	0.95	0.93	0.97	0.97	0.83	1.07	0.90	1.04
2013	0.96	0.95	0.97	0.97	0.83	1.12	0.91	1.02
2014	0.95	0.98	0.92	0.92	0.85	1.04	0.88	0.97
2015	1.00	1.01	0.99	0.99	0.88	1.12	0.93	1.04
2016	1.03	1.03	1.04	1.03	0.83	1.22	0.94	1.14

Table 2b. Estimated slope from the robust least squares fit (bc=bias corrected; bs=bootstrap) for 5-25-50-75-95 percentiles, for both the local least square (a) and robust least square (b) *loess* fits.

Year	Slope	Mean_bs	bias_adj	bc_0.5	bc_0.05	bc_0.95	bc_0.25	bc_0.75
1991	0.96	0.95	0.98	0.97	0.87	1.01	0.91	0.99
1992	0.83	0.84	0.82	0.81	0.71	0.92	0.76	0.88
1993	0.69	0.79	0.60	0.62	0.55	0.86	0.58	0.73
1994	0.83	0.85	0.81	0.77	0.59	0.95	0.70	0.85
1995	1.13	1.00	1.25	1.32	0.96	1.55	1.17	1.44
1996	1.13	1.07	1.18	1.23	0.92	1.46	1.07	1.39
1997	1.01	1.06	0.97	0.99	0.90	1.21	0.93	1.09
1998	0.96	1.04	0.87	0.92	0.89	1.07	0.91	0.96
1999	1.03	1.04	1.01	1.01	0.84	1.24	0.94	1.12
2000	1.05	1.02	1.09	1.07	0.91	1.18	0.99	1.13
2001	0.99	1.01	0.98	0.97	0.92	1.06	0.94	1.01
2002	0.99	1.03	0.95	0.94	0.90	1.08	0.91	1.00
2003	1.08	1.08	1.07	1.08	0.91	1.30	1.00	1.17
2004	1.15	1.09	1.21	1.21	1.05	1.29	1.15	1.26
2005	1.05	1.03	1.06	1.06	0.96	1.15	1.02	1.11
2006	0.91	0.94	0.88	0.88	0.84	0.97	0.85	0.91
2007	0.84	0.88	0.81	0.80	0.78	0.91	0.79	0.85
2008	0.85	0.87	0.83	0.83	0.79	0.93	0.81	0.88
2009	0.93	0.89	0.96	0.95	0.88	0.97	0.92	0.96
2010	0.92	0.91	0.93	0.92	0.82	1.01	0.90	0.98
2011	0.89	0.92	0.86	0.87	0.82	0.97	0.84	0.90
2012	0.94	0.93	0.96	0.96	0.82	1.06	0.88	1.03
2013	0.95	0.95	0.96	0.95	0.82	1.10	0.89	1.01
2014	0.95	0.97	0.93	0.92	0.85	1.06	0.88	0.97
2015	1.00	1.01	1.00	1.01	0.89	1.15	0.96	1.07
2016	1.05	1.04	1.06	1.06	0.87	1.24	0.99	1.17

Table 3. Quota advice (mt) resulting from application of the empirical method.

Year	25%	50%	75%
2017 (local least square)	**611**	**669.5**	**741**
2017 (robust least square)	**643.5**	**689**	**760.5**

Table 4. Table of secondary indicators to characterize risk to the cod resource.

Reasons to choose low risk	Reasons to choose neutral risk
High total mortality	Relative exploitation rate at the lowest level since 1978
Poor recruitment for the last 20 years	2013 year class is somewhat stronger than recent year classes
Mean weight is below average	Recent improvement in condition (Fulton’s K)
Low spawning stock biomass	
Lack of rebuilding	
Assessment models overestimating biomass, projections have been optimistic*	
**There is a lack of consensus on this entry.*	

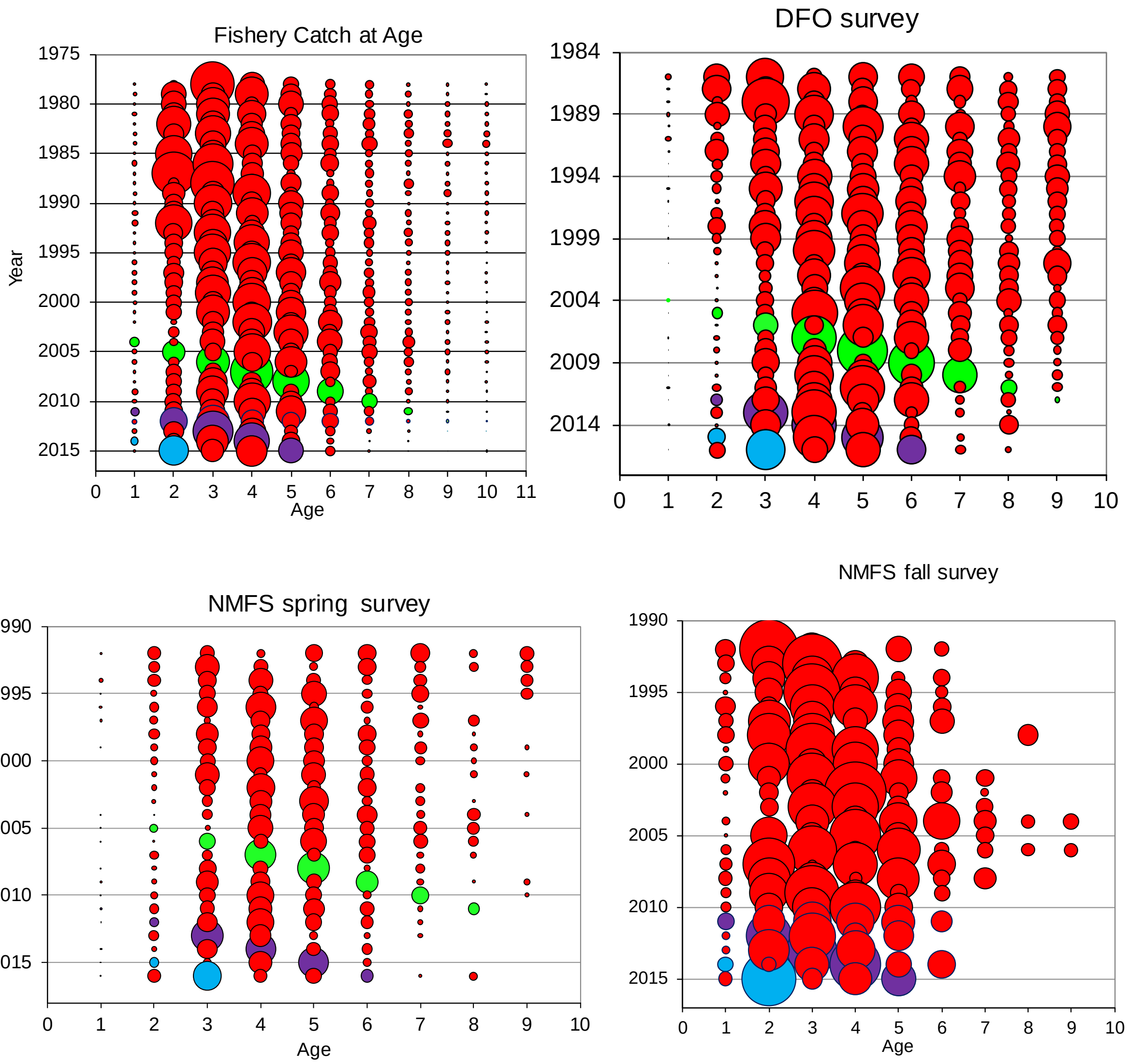

Figure 1a. Proportion of biomass at age from eastern Georges Bank for commercial fishery (top left), DFO spring survey (top right), NMFS spring survey (bottom left) and NMFS fall survey (bottom right). The bubble area is proportional to the magnitude.

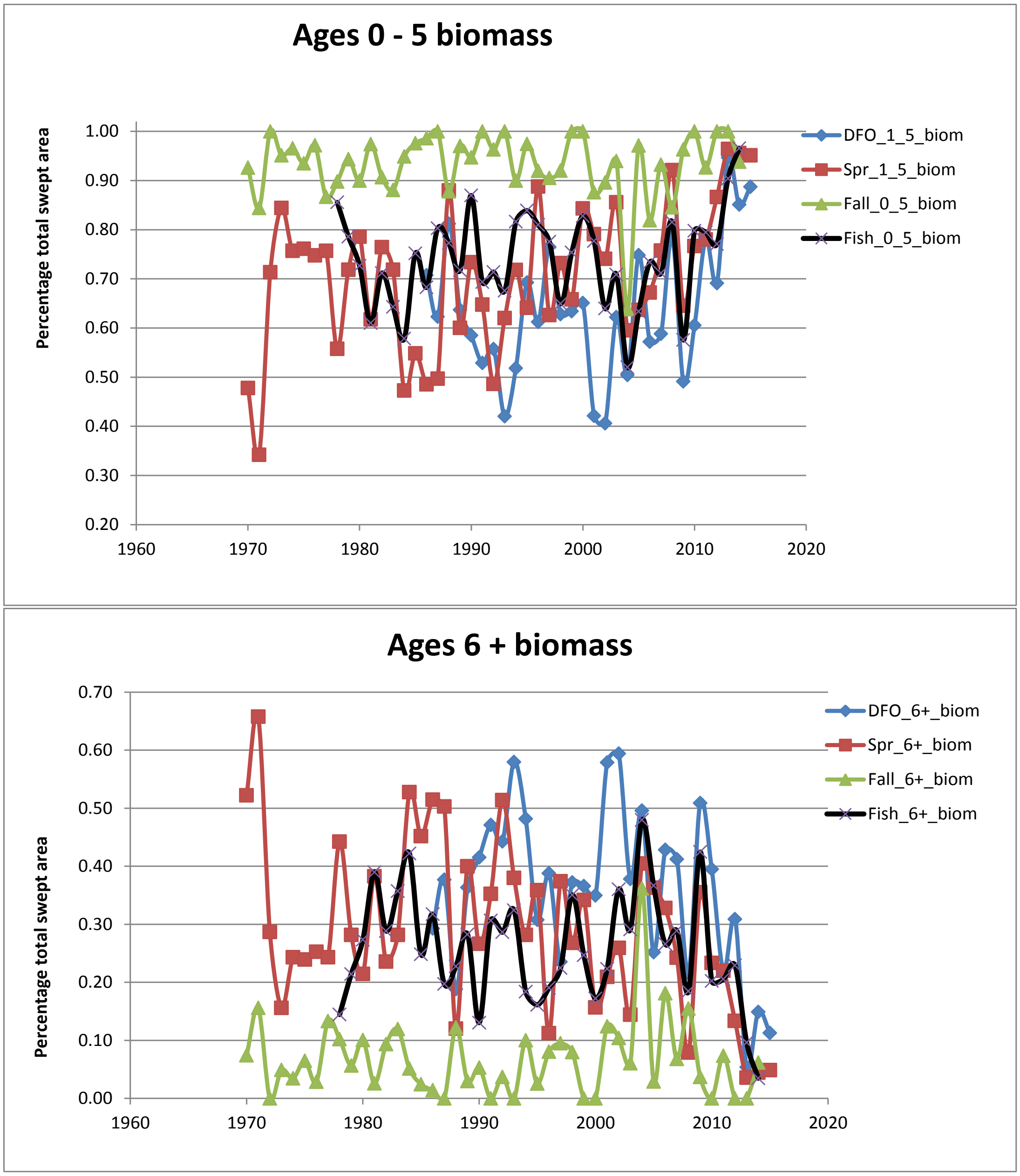

Figure 1b. Temporal change in percent biomass contribution by 0-5 ages (top) and 6+ ages for eastern Georges Bank Cod by from DFO spring Survey (blue), NMFS spring Survey (Red), NMFS fall Survey (Green) and fishery catch (black)

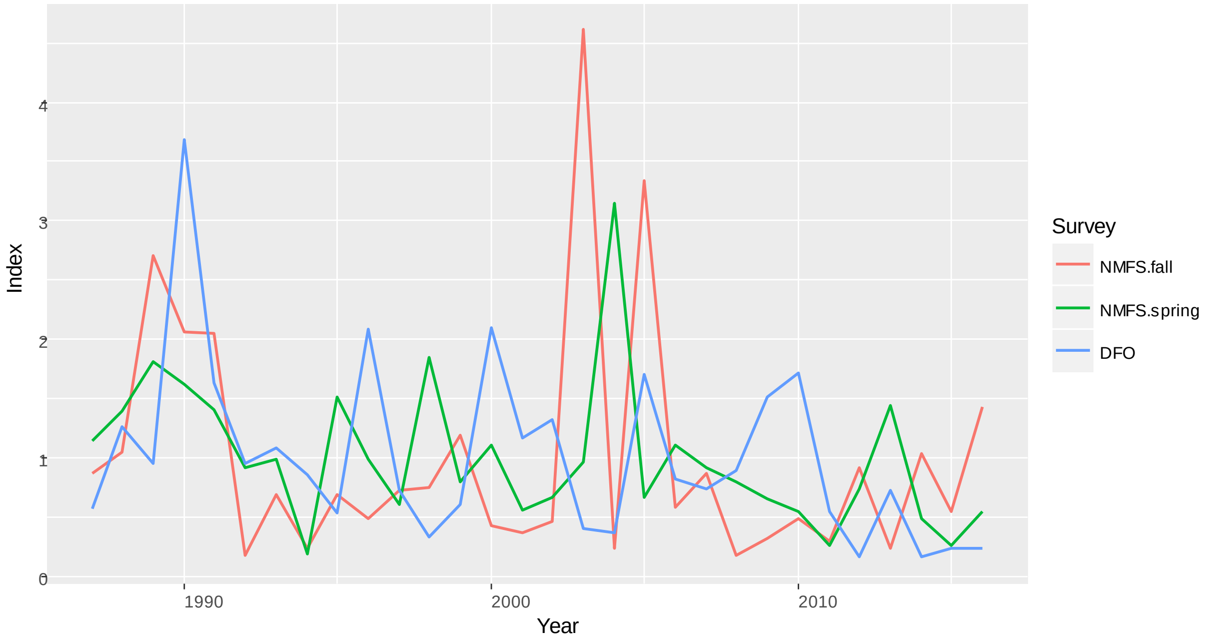

Figure 2. Plot of the normalized NMFS fall, DFO, and NMFS spring indices from 1987 (1986 fall) through 2016 (2015 fall). All three indices were divided by their mean and are plotted on the same scale.

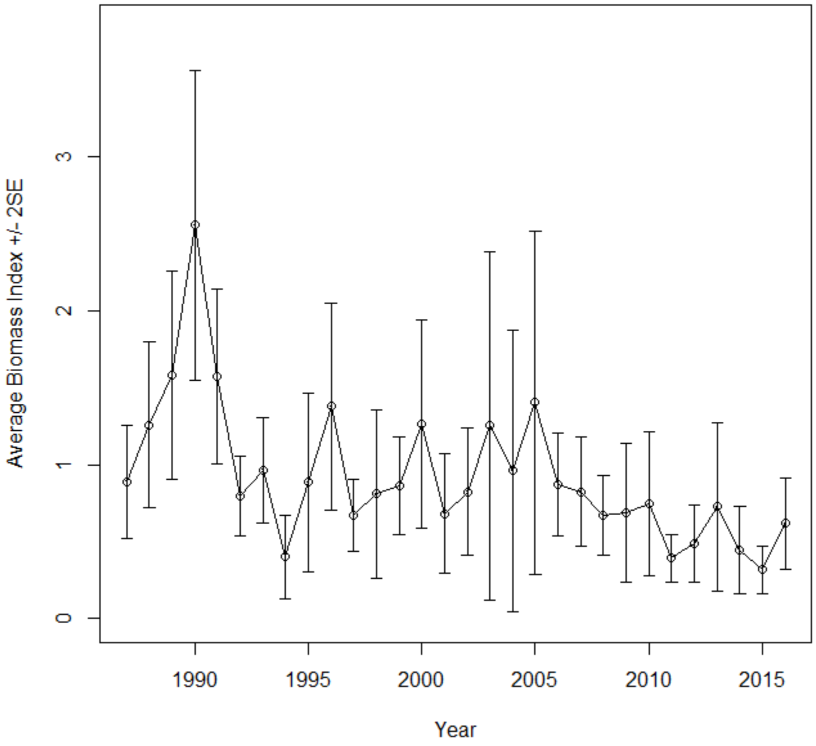

Figure 3. Plot of the Combined Index from CV weighted average of 3 surveys.

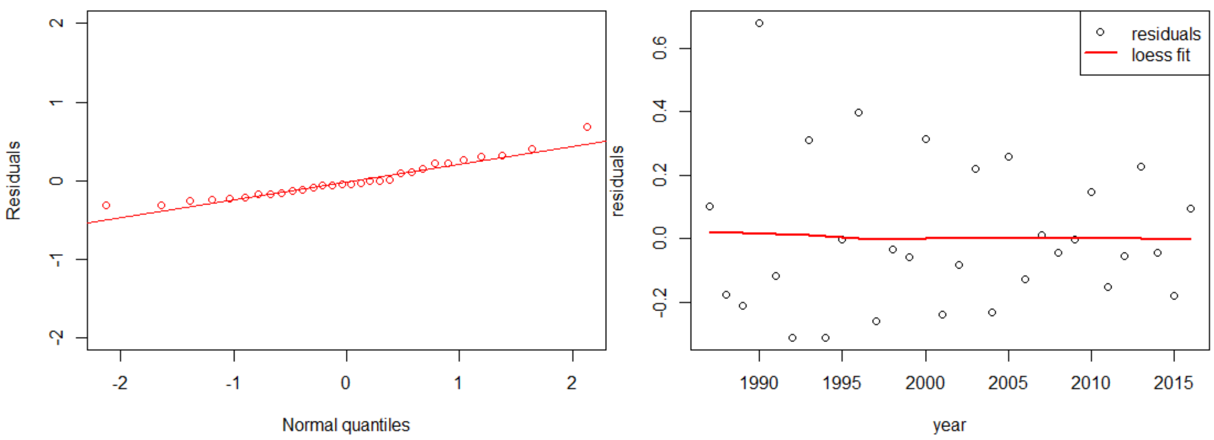

Figure 4. Diagnostics for assessing the fit of the *loess* smooth based on local least squares (family=Gaussian).

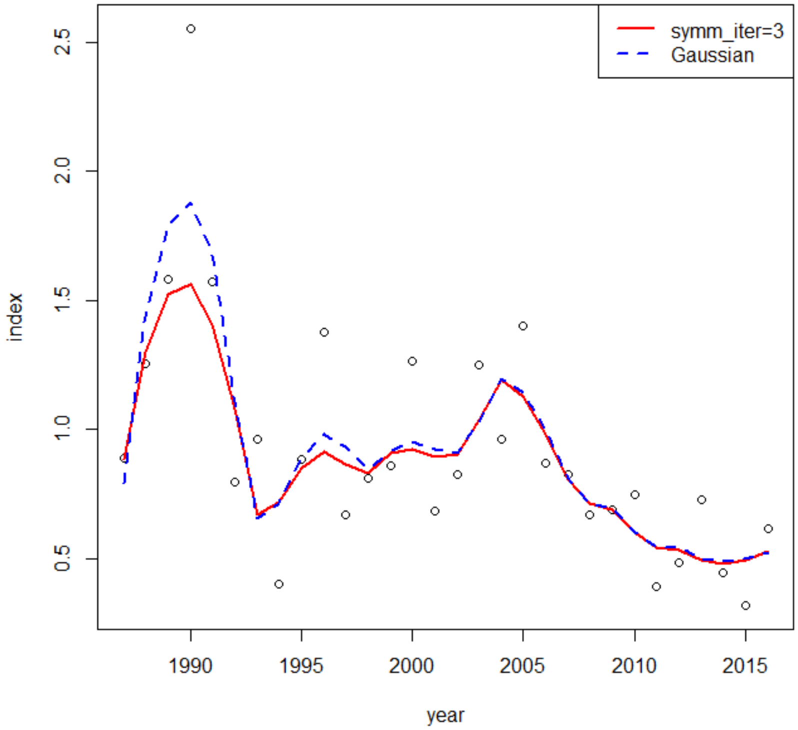

Figure 5. Comparison of fit of local least squares versus robust least squares.

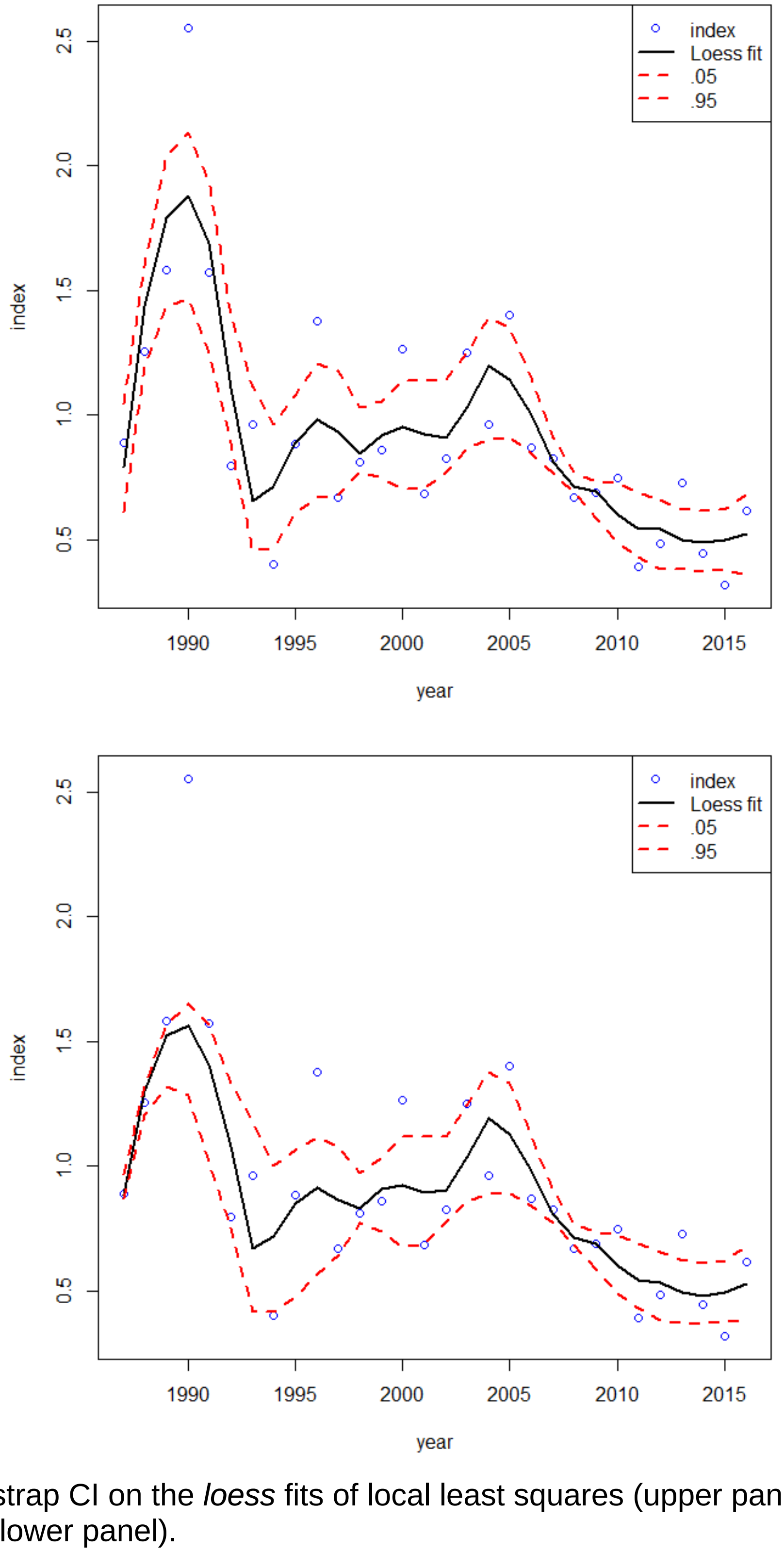

Figure 6. Bootstrap CI on the *loess* fits of local least squares (upper panel) versus robust least squares (lower panel).

Figure 7. Bootstrap CI for the estimated slope of local least squares (upper panel) versus robust least squares (lower panel).