

Further Remarks on Gulf of Maine-Georges Bank Witch Flounder Assessment Results

Doug S. Butterworth and Rebecca A. Rademeyer

November 2016

Summary

An SCAA assessment variant is argued to be preferable to the ASAP final model advocated by the SAW WG as the basis for witch flounder status assessment and management advice. The key differences introduced in this assessment variant are the possibility of domed selectivity, the inclusion of the LPUE times series, and an objectively-based down-weighting of the contribution to the likelihood of the CAA data. The main reasons that this variant is preferred (with comparisons to an SCAA assessment specified to mimic the ASAP final model) are a much reduced retrospective pattern (virtually eliminated for spawning biomass), notably better (though certainly still not complete) agreement with the results from the Sweep study in regard to estimates of abundance in absolute terms, and appreciably improved fits to the NEFSC survey indices of abundance.

Introduction

There are two major difficulties with the “ASAP final model” advocated by the SAW WG as the “best” assessment of witch flounder and hence the basis to determine status evaluation and advice:

- a) a strong retrospective pattern; and
- b) an underestimation by a factor $q \sim 4$ of biomass in absolute terms as indicated by the “Sweep study” results.

During the SAW meeting certain results became available only over the last two days. For the alternative SCAA methodology, this was somewhat inevitable because exploration of variants could be conducted effectively only after the meeting had agreed on baseline specifications to provide an appropriate basis from which to work and allow for inter-method comparisons. One such SCAA variant developed late during the WG meeting (see sens7 in Table 5 together with Figures 15-18 and the associated text towards the end of Appendix B3 of the WG report) largely addresses the first and in part the second of these two key concerns. This contributed in large part to the minority view recorded on pg 63 of the WG Report against adopting the ASAP final run as the best basis for provision of scientific advice on witch flounder management.

Since, however, new SAW WG procedures preclude the motivation for minority views being included in SAW Reports, this document summarises those motivations for the information of SARC reviewers.

Those reviewers are also requested to note some differences of recollection regarding certain components of the SAW Witch flounder report as to whether these had been accepted as the agreed view of all WG members, or constituted place-holders to which to return later in the meeting (something which was overlooked in the rush as the meeting concluded).

Data and Methodology

The SCAA data and methods are as detailed in Appendix B3 of the main report, except that for RCvar1 and RCvar2 an extra parameter M_{plus} was estimated to allow for the possibility of M increasing at large ages:

$$M_a = \begin{cases} 0.15 & a \leq 7 \\ 0.15 + M_{plus}(a - 7) & 7 < a \leq 11 + \end{cases} \quad (1)$$

Projections

These are deterministic and made the following assumptions:

- A catch of 460mt in 2016 and fishing at $F_{40\%}$ in 2017-2019;
- Future recruitment taken to be mean recruitment over the period 1982 to 2015;
- Future commercial selectivity taken to be the same as the 2005-2015 selectivity,
- Future mean weights-at-age taken to be the average of the 2011-2015 mean weights; and
- Maturity taken to be the same as the average of the 2012-2016 maturity vectors.

Calculation of $F_{40\%}$

The equilibrium catch for a fully selected fishing proportion F is calculated as:

$$C(F) = \sum_a w_a^{mid} S_a F N_a(F) e^{-(M_a/2)} \quad (2)$$

where numbers-at-age a are given by:

$$N_a(F) = \begin{cases} R_1(F) & \text{for } a = 1 \\ N_{a-1}(F) e^{-M_{a-1}} (1 - S_{a-1} F) & \text{for } 1 < a < m \\ \frac{N_{m-1}(F) e^{-M_{m-1}} (1 - S_{m-1} F)}{(1 - e^{-M_m} (1 - S_m F))} & \text{for } a = m \end{cases} \quad (3)$$

with

$$R_1(F) = \frac{\alpha B^{sp}(F)}{\beta + B^{sp}(F)} \quad (4)$$

$F_{40\%}$ is then found by searching over F so that the associated spawning biomass is equal to 40% of the biomass at $F=0$:

$$B_{F_{40\%}}^{sp} = \sum_a f_a w_a^{str} N_a(F_{40\%}) e^{-M_a \delta} = 0.4 B_{F=0}^{sp} \quad (5)$$

The selectivity, mean weight-at-age and maturity are taken to be the same as assumed for the projections.

RY analyses

These have followed the methodology set out in Appendix B2 exactly, except for an adjustment to permit RY to be estimated to be less than zero.

Results

SCAA

Table 1 lists estimates of primary parameters and management-related quantities for the RC and two variants:

- RC: this is sens7 of Appendix B3 of the main report: dome selectivities for both the fishery and the surveys, fits to LPUE (40% trips series) and SCAA final BC and downweighting the CAA data ($W_{CAA}=0.36$ for commercial, 0.54 for spring survey and 0.34 for autumn survey);
- RCvar1: as for the RC but with flat survey selectivities and allows for an increase in M from age 7;
- RCvar2: as for the RC but with flat survey and commercial selectivities and allows for an increase in M from age 7.

Figures 1-3 give plots of fits and diagnostics for the three runs, while Figure 4 plots the retrospective analysis for the RC.

The 2016-2019 projected catches and associated spawning biomass are given in Table 2 for the RC and the two variants.

RY analyses

Figure 5 plots the estimated RY (together with standard error estimates) from analysis starting from 1990 to starting from 2006 for the cases with $q=1$ and with and without the LPUE series.

Discussion

All “best assessment” selections involve some trade-off to balance the good and poor properties of the results for different assessment variants. In our view, the results for SCAA sens7, renamed Reference Case (RC) herein, reflect a much improved trade-off choice to those for the ASAP final model.

It is first important to be clear as to the differences between these two runs. Sens7 is a variant of the SCAA Final Base Case (see Table 5 of Appendix B3), which was developed to mimic the specifications of that ASAP final model as closely as the differences between the ASAP and SCAA frameworks allowed. Sens7 involves three variations to those specifications. These, together with their motivations, are as follows.

- 1) *Domed (estimated) selectivities for both the commercial fishery and the NEFSC surveys rather than ones which are asymptotically flat*

These domes are not (in isolation) strongly preferred in likelihood terms (see Table 4 of Appendix B3). However they do assist reduce the estimate of q in RC compared to the final SCAA Base Case, and independent arguments offered by the SAW WG for preferring flat selectivity are not strong: catch curve features are subject to multiple confounding effects, and though length distributions by bottom type were argued not to be supportive of domed selectivity, the WG was informed that no power analysis had been conducted to check the ability of that approach to detect such differences. Furthermore, lesser numbers of older fish in catches may be reflective of higher natural mortality instead of (or in combination with) domed selectivity, which is the reason for reporting results for variants that consider this possibility in this document.

- 2) *Inclusion of the LPUE index of abundance in the fit*

An important result which became evident on the final day of the SAW WG meeting, when comparing ASAP results with those from SCAA for which the additional variance associated with survey indices of abundance is estimated to be additive (see equation B.18 of Appendix B3) rather than the multiplicative approach implemented for ASAP, was that the multiplicative factors applied in the ASAP approach were much too small, thus overweighting these data.

Adjusting for this made little difference to the ASAP results, which did not include fitting to LPUE data, but the importance of this becomes clear when residual variances for fits to both sets of series are compared (see Table A.1 of Appendix B2, which shows that these variances are some three times bigger for the NEFSC surveys compared to the LPUE data, so that in this respect the latter are much more informative than the former as regards abundance trends. This reinforces impressions we have obtained in assessments of South African resources where fits to both CPUE and research survey indices of abundance point to the much lesser variance of the former.

In these circumstances we do not consider it justifiable to omit the LPUE series from the data to which the selected assessment model is fit. Clearly there is the question of possible bias that arises when the LPUE series is introduced, but this has to be traded off against the considerable gains regarding variance/precision that including the LPUE series brings.

In these circumstances we disagree with the majority SAW WG decision to exclude the LPUE series from the ASAP final run. Simply to argue “unknown bias concerns” as a reason is inadequate in the face of the variance advantages brought by including these data. The burden to be met in such a situation, if such LPUE data are to be excluded from the “best” assessment, is a quantitative demonstration that issues such as management regulation changes could bring about effects of sufficient magnitude to indeed give rise to substantial bias.

3) *Down-weighting of the CAA data in the likelihood*

CAA data are certainly correlated, and taking this into account using an objective procedure to estimate the associated down-weighting factors (see Appendix B3) would seem preferable to failing to do so in circumstances where there is some conflict amongst the different data types. In such a situation, appropriate choices for relative weighting become the more important.

Why prefer the RC to the ASAP final model?

There are a number of strong reasons for this.

- The retrospective pattern (see Figure 4) is considerably reduced, especially for spawning biomass, compared to that for the SCAA Final Base Case (see Figure 11 of Appendix B3) which mimics the ASAP final model. Mohn’s rho falls in magnitude from 1.99 to -0.17 for spawning biomass, and from -2.18 to -0.85 for apical fishing mortality.
- The survey q values drop from more than 4 to near to 3 – not ideal, but certainly moving closer towards the “Sweep study” results. Furthermore, if this q is forced to be 2 or less, the overall $-\ln L$ value for the fit deteriorates by only some 1.5 units (see Table 5 of Appendix B3), suggesting that no great inconsistency with the data fitted arises for yet smaller values of q .
- Compared to the SCAA equivalent to the ASAP final model, the fit to the NEFSC survey abundance estimates improve by 15 $-\ln L$ points (see Table 5 of Appendix B3), which is substantial. Although, naturally, this is achieved at the expense of the fits to the CAA data, these still appear broadly reasonable (see Figure 3).

Some other remarks

The estimated catch for 2017 under $F_{40\%}$ is 2 720 mt (Table 2), compared to the 762 mt for the ASAP final model. Clearly the selection amongst assessment variants has important management implications.

Allowing for natural mortality M to increase with age above age 7 leads to yet smaller q estimates (down to about 2 – see Table 1), but this comes with the 1982 spawning biomass being estimated to be above K , which is primarily a reflection of the estimation of a large 11+ abundance in this (assessment-commencing) year. However, given that the RY model suggests a drop in resource productivity over time (see Figure 5), it might be that there has also been a fall in the value of K over time in contrast to the time-invariance assumed for the assessment model.

Notwithstanding the above, if assessment estimates of q notably greater than 1 remain considered unacceptable given the Sweep study results, the RY model (which can force $q=1$) becomes a candidate for the basis for the provision of management advice. For this reason, Figure 5 has been provided to extend the results provided for this model in Appendix B2. Clearly results for RY, which pertain to a potential catch limit, are dependent on the length of the time series of catches and abundance indices considered in this case. Lesser lengths yield lower estimates but also higher variances. For the reasons given above we would advocate inclusion of the LPUE series in these analyses. Considering the bias-variance trade-off involved in choosing the year to consider in commencing such analyses, we would advocate a commencement year somewhere in the 1995-2000 period.

Table 1: Negative log-likelihood components and estimates of abundance and related quantities for witch flounder for the RC and two variants. The RC corresponds to sens7 of Appendix B3 of the main report. Biomass units are thousand mt.

	RC	RCvar1	RCvar2
	Both dome, incl. LPUE, down. CAA	As RC, but flat survey sel. and increasing M from age 7	As RC, but flat survey and comm. sel. and increasing M from age 7
-lnL: overall	-656.2	-660.3	-647.0
-lnL: survey	41.0	44.3	47.0
-lnL: LPUE	9.1	13.3	20.7
-lnL: comCAA	-229.5	-236.5	-229.0
-lnL: survCAA	-457.6	-460.6	-466.4
-lnL: RecRes	27.2	25.5	26.5
-lnL: Catch	-46.4	-46.4	-45.8
K^{sp}	37.13	24.61	21.59
B^{sp}_{1982}	31.85	36.90	30.45
B^{sp}_{2015}	14.78	14.61	11.78
B^{sp}_{2015}/K^{sp}	0.40	0.59	0.55
$B^{sp}_{2015}/B^{sp}_{1982}$	0.46	0.40	0.39
F_{2015}	0.07	0.06	0.05
ϕ	0.13	0.00	0.00
M_7	0.15	0.15	0.15
M_{11+}	0.15	0.38	0.41
1982 N-at-age ($\times 10^6$)			
$N_{1,1982}$	16.54	20.85	19.19
$N_{2,1982}$	21.32	25.93	23.67
$N_{3,1982}$	21.35	25.57	23.58
$N_{4,1982}$	16.18	19.58	17.58
$N_{5,1982}$	10.08	13.52	11.76
$N_{6,1982}$	8.12	11.63	10.12
$N_{7,1982}$	6.50	10.01	8.71
$N_{8,1982}$	4.84	8.62	7.50
$N_{9,1982}$	3.77	7.00	6.05
$N_{10,1982}$	3.02	5.37	4.58
$N_{11+,1982}$	14.51	12.31	9.73
q :			
NEFSC spring	3.36	2.01	2.58
NEFSC autumn	3.18	1.90	2.48

Table 2: Projected catches and spawning biomass under $F_{40\%}$ for the RC and two variants. Hessian-based 95% CIs are given in parentheses. The catch in 2016 is taken to be 460 mt. Both catch and biomass units are thousand mt.

	RC		RCvar1		RCvar2	
	Both dome, incl. LPUE, down. CAA		As RC, but flat survey sel. and increasing M from age 7		As RC, but flat survey and comm. sel. and increasing M from age 7	
F_{2016}	0.05		0.04		0.04	
$F_{40\%}$	0.26		0.36		0.32	
C_{2016}	0.46	(0.35; 0.61)	0.46	(0.33; 0.63)	0.46	(0.29; 0.67)
C_{2017}	2.72	(1.96; 3.49)	4.41	(3.06; 5.76)	4.31	(2.78; 5.85)
C_{2018}	2.85	(1.99; 3.71)	4.32	(2.93; 5.70)	4.29	(2.85; 5.73)
C_{2019}	3.05	(2.03; 4.08)	4.34	(2.81; 5.88)	4.30	(2.81; 5.78)
B^{sp}_{2016}	16.76	(11.01; 22.51)	16.29	(11.49; 21.09)	13.55	(8.80; 18.30)
B^{sp}_{2017}	19.02	(13.06; 24.97)	18.26	(13.05; 23.47)	15.61	(10.71; 20.50)
B^{sp}_{2018}	20.08	(13.93; 26.24)	17.94	(12.61; 23.27)	15.46	(10.69; 20.22)
B^{sp}_{2019}	20.81	(14.31; 27.30)	17.37	(11.76; 22.99)	15.08	(10.06; 20.11)

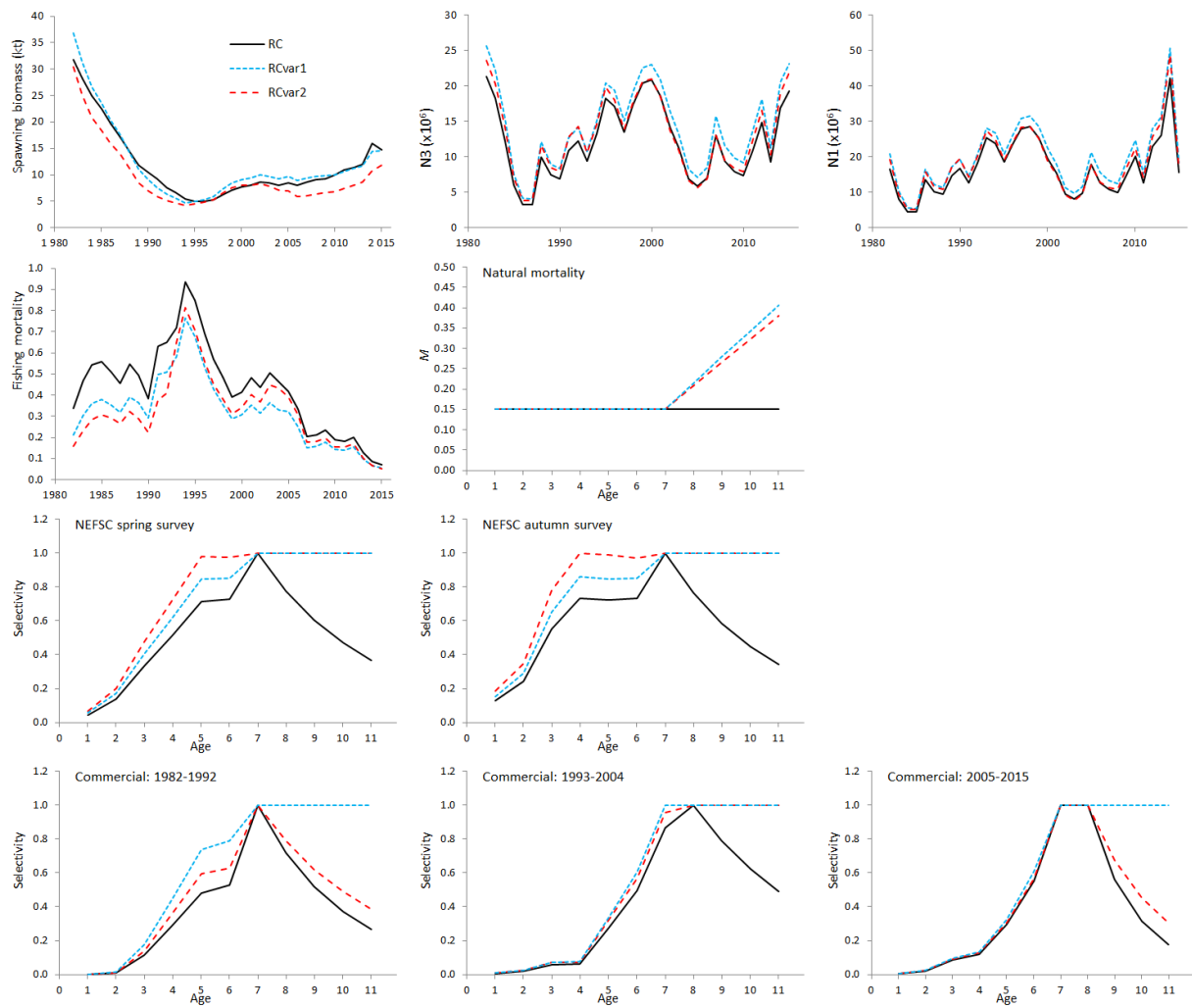


Figure 1: Results for the RC and two variants. Fishing mortality is apical (for age 7).

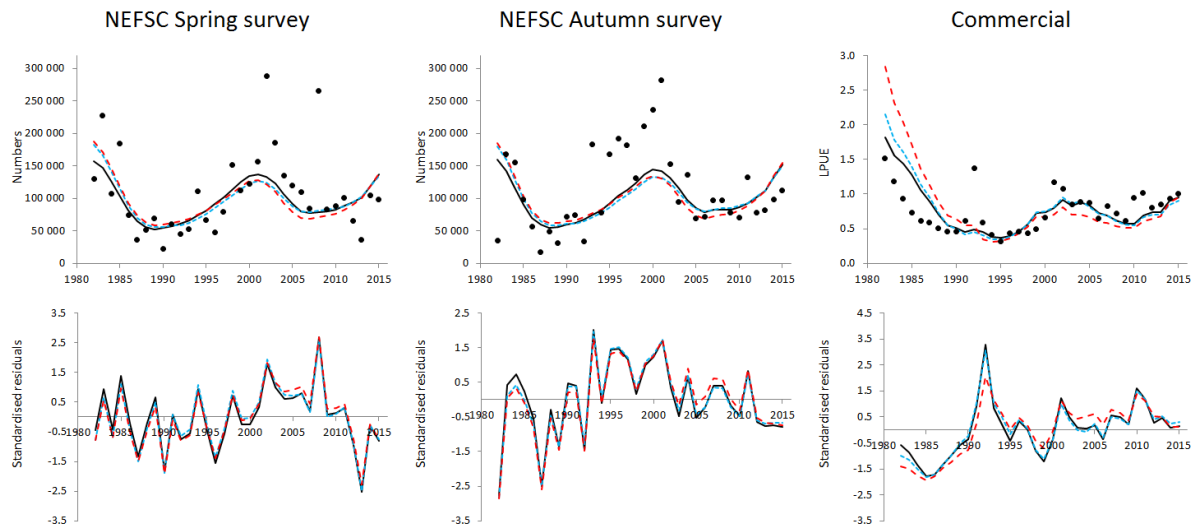


Figure 2: Fits to the survey and LPUE series for the RC and two variants.

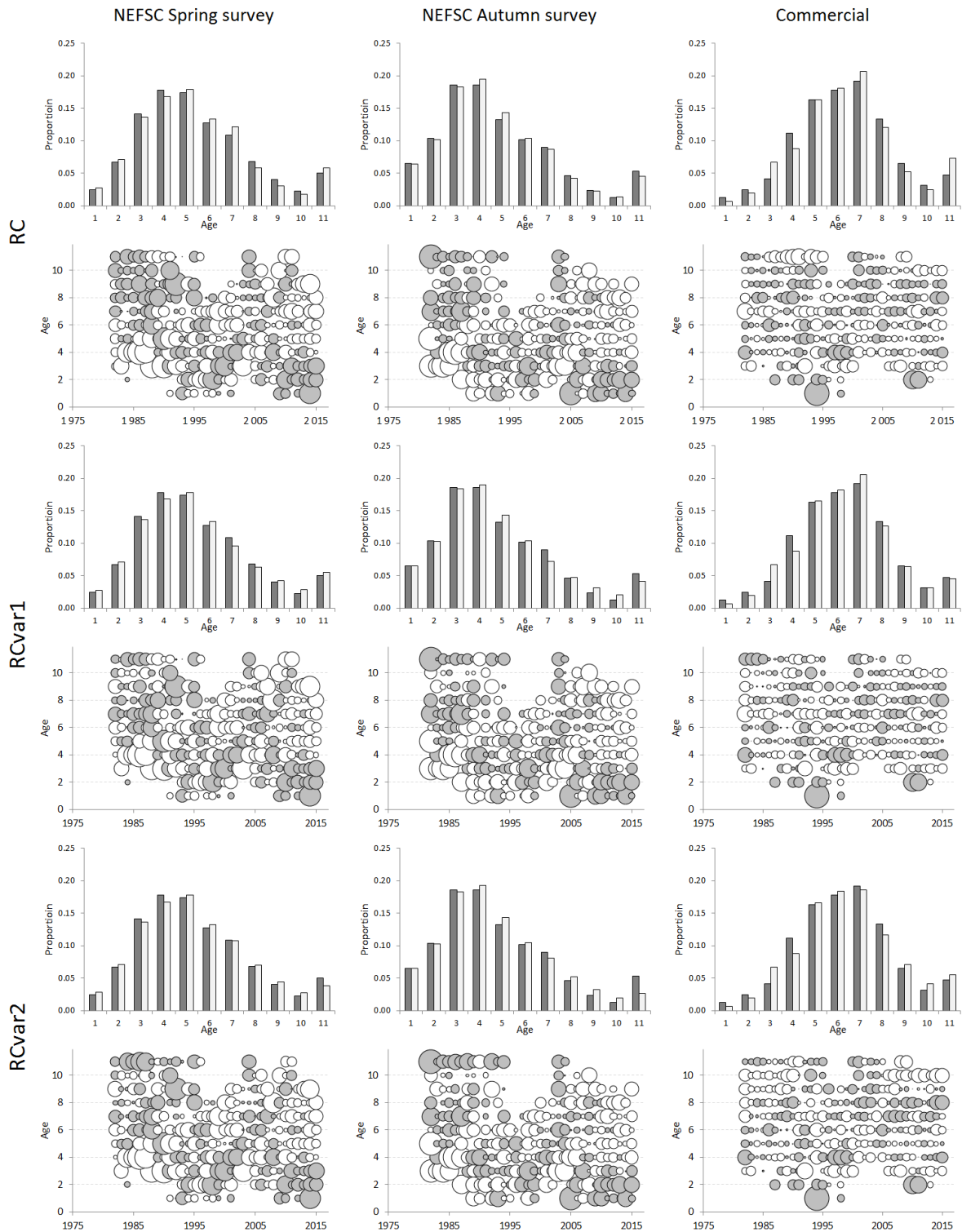


Figure 3: Fits to the survey and commercial catch-at-age data (averaged over all the years for which data are available and bubble plots of residuals, with grey bubbles representing positive residuals) for the RC and two variants.

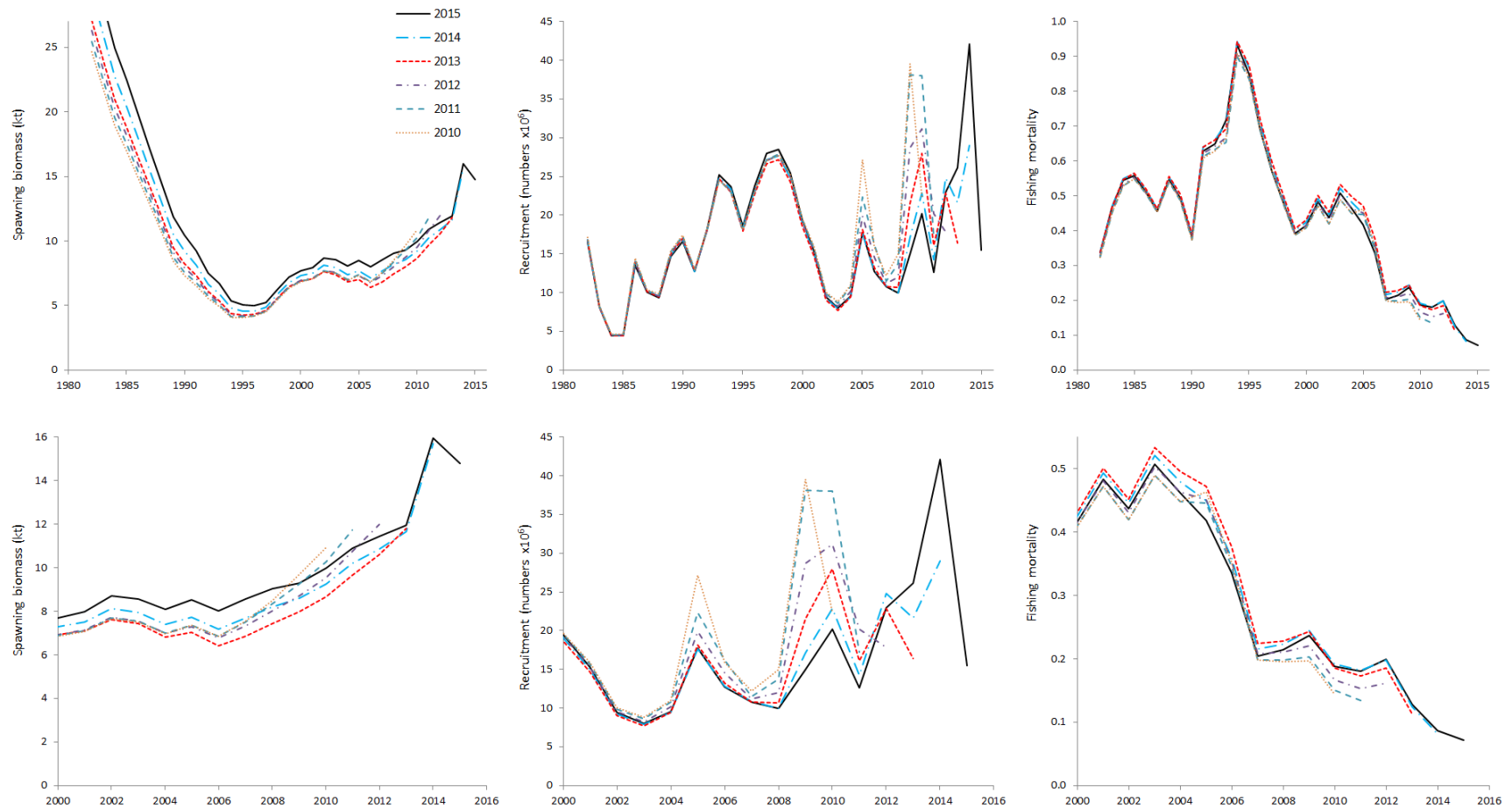


Figure 4: Retrospective plots of spawning biomass, recruitment and (apical – age 7) fishing mortality for the RC (domed fishery and survey selectivities, including LPUE and downweighting the CAA data). The bottom row replicates the top row but with different scales.

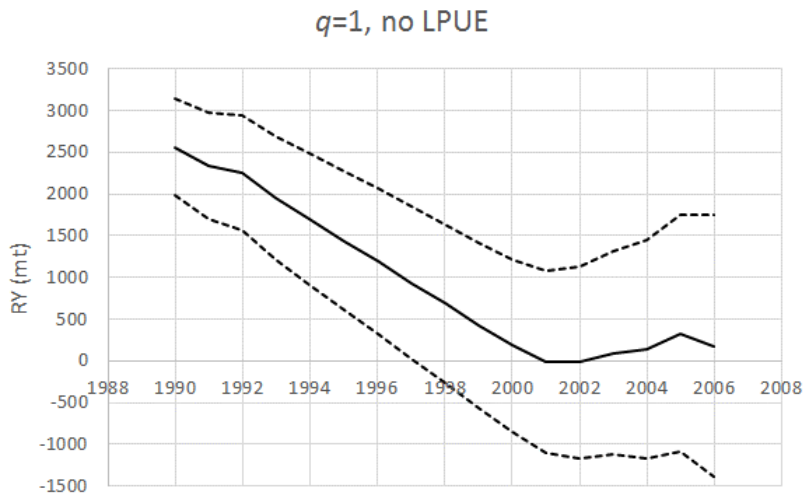
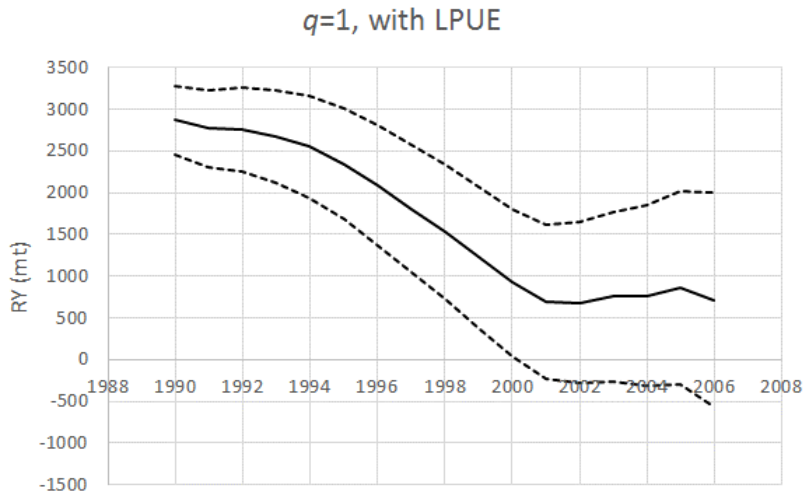


Figure 5: RY estimates for starting years from 1990 to 2006. The dashed lines are with one (Hessian based) standard error added and subtracted.