

On the Precision of Absolute Abundance as Estimated by VPA

Doug S. Butterworth and Rebecca A. Rademeyer

MARAM (Marine Resource Assessment and Management Group)
Department of Mathematics and Applied Mathematics
University of Cape Town, Rondebosch 7701, South Africa

The higher degree of imprecision in the scale of the estimates of halibut biomass in absolute terms shown for the baseline SCAA assessment in Butterworth and Rademeyer (2009) may seem surprising. How can this be reconciled with XSA results (Healey and Mahé, 2008) for which much greater precision is reported, and which manifest the backwards-convergence property of VPA?

We suspect that the seemingly high precision of the XSA results might be an artefact of the manner in which XSA, like many VPA approaches, forces a deterministically exact relationship between the fishing mortalities on the two oldest age classes considered in such analyses, e.g. that $F_{y,m-1}$ equals $F_{y,m}$, or some constant multiple thereof if selectivity is not asymptotically flat. However, since in VPA the selectivity pattern amongst younger ages shows variability over years, why not for the two oldest ages as well?

The two Annexes to this paper reproduce contributions by the authors to review meetings last year which considered the assessment of the US Gulf of Maine cod resource. Annex I develops a VPA approach which happens to admit such variability in this asymptotic relationship amongst the fishing mortalities-at-age. Annex II shows a particular application of this approach.

The rather interesting feature of the results shown in Annex II is that results can change appreciably depending on the extent of variability admitted in this asymptotic relationship. As this extent is increased, the estimated resource trajectory does not always change smoothly, but suddenly jumps to another solution corresponding to appreciably higher biomass in absolute terms (see Fig. A2.1 in Annex II). This is in line with our suspicions that aspects of the algorithms usually applied in

implementing VPA may lead to a false impression of the precision associated with the overall scale of the biomass estimated.

[Note: For the example in Annex II, asymptotically flat selectivity (on average) was assumed, and natural mortality on the plus group treated as an estimable parameter. This is very similar to fixing natural mortality, but allowing a constant ratio between the fishing mortalities on the two oldest age-groups in the analysis to be estimated.]

References

Butterworth DS and Rademeyer RA. 2009. Extensions to SCAA applications reported in: “Further applications of statistical catch-at-age assessment methodology to the 2J3K-O Greenland halibut assessment. Document presented to this meeting.

Healey BP and Mahé J-C. 2008. An assessment of Greenland halibut (*Reinhardtius hippoglossoides*) in NAFO Subarea 2 and Divisions 3KLMNO. NAFO SRC Doc. 08/48, Ser. No N5550.

ANNEX I

Appendix A - The ADAPT-VPA Model

Note that the specifications set out in the first part of section A1 are not in their most general form (see Anon., 2003), but rather as implemented for the Mayo and Col (2006) application to Gulf of Maine cod.

A.1. Population Dynamics

The resource dynamics are modelled by the following set of equations:

$$N_{y,a} = N_{y+1,a+1}e^{M_a} + C_{y,a}e^{M_a/2} \quad \text{for } 1 \leq a \leq m-2 \quad (\text{A1})$$

$$Z_{y,a} = \ln\left(\frac{N_{y,a}}{N_{y+1,a+1}}\right) \quad (\text{A2})$$

$$F_{y,a} = Z_{y,a} - M_a \quad (\text{A3})$$

where

$N_{y,a}$ is the number of fish of age a at the start of year y (which refers to a calendar year),

M_a denotes the instantaneous rate of natural mortality for fish of age a ,

$C_{y,a}$ is the number of fish of age a caught in year y ,

m is the maximum age considered (taken to be a plus-group),

$Z_{y,a}$ is the instantaneous rate of mortality during year y from all causes (total mortality) on fish of age a , and

$F_{y,a}$ is the instantaneous rate of fishing mortality on fish of age a .

The numbers of the oldest true age ($m-1$) and the plus-group (m) are computed as follows:

$$N_{y,a} = \frac{Z_{y,a}C_{y,a}}{F_{y,a}(1 - e^{-Z_{y,a}})} \quad \text{for } a = m-1 \text{ and } a = m \quad (\text{A4})$$

Fishing mortality on the oldest true age is defined as:

$$F_{y,m-1} = PR_{m-1}F_y^{full} \quad (\text{A5})$$

where

$$F_y^{full} = \ln \left[\frac{\sum_{a \in R} N_{y,a} e^{-M_{y,a}}}{\sum_{a \in R} N_{y+1,a+1}} \right]$$

is the fully-recruited fishing mortality in year y , R denoting the set of fully-recruited age classes, excluding the oldest true age $m-1$, and

PR_{m-1} is the partial recruitment for fish of age $m-1$, which is input. (Note the partial recruitment PR_a is essentially the selectivity S_a of the ASPM approach of Appendix B.)

Fishing mortality on the plus-group is defined as:

$$F_{y,m} = \alpha F_{y,m-1} \tag{A6}$$

where

α is the plus-group ratio, which is input.

In the RC-VPA, $PR_{m-1} = 1$ and $\alpha = 1$; further the set of fully recruited age-classes in equation (A5) is taken to be $R = \{4,5\}$ where $m=7$ and $F_{y,m-1}$ is set equal to F_y^{full} .

Alternative approach (Alt-VPA)

There is a problem with the overall approach above used to compute plus-group abundances. Essentially that approach consists of fitting a model to the data up to age $m-1$ to estimate a numbers-at-age matrix $N_{y,a}$ for ages 2 to $m-1$, and then applying equation (A4) for each year in conjunction with equation (A6) to provide the plus-group abundance for that year. The difficulty with this is that plus-group abundance is governed by the equation:

$$N_{y+1,m} = \left(N_{y,m} e^{-M_{m/2}} - C_{y,m} \right) e^{-M_{m/2}} + \left(N_{y,m-1} e^{-M_{m-1/2}} - C_{y,m-1} \right) e^{-M_{m-1/2}} \tag{A7}$$

and results obtained from the combined application of equations (A4) to (A6) will not necessarily satisfy equation (A7), because of the specification of potentially contradictory conditions. In other words, the overspecification of the approach above leads to incorrect estimates of plus-group abundance.

In circumstances of asymptotically flat selectivity (partial recruitment) at higher ages, together with heavy fishing mortality so that few fish survive to reach the plus-group, any errors to which these inconsistencies give rise are likely to be small. However, this is not necessarily the case in circumstances of lesser fishing mortality and particularly selectivity that declines with age at larger ages.

This problem can be rectified by replacing equation (A4) by equation (A7) together with the equations following:

$$Z_{y,m-1} = \ln \left(\frac{N_{y,m-1}}{N_{y,m-1} e^{-M_{m-1}} - C_{y,m-1} e^{-M_{m-1/2}}} \right) \tag{A8}$$

and

$$F_{y,m-1} = Z_{y,m-1} - M_{y,m-1} \tag{A9}$$

and for $a=m$:

$$Z_{y,m} = \ln \left(\frac{N_{y,m}}{N_{y,m} e^{-M_m} - C_{y,m} e^{-M_m/2}} \right) \quad (\text{A10})$$

and

$$F_{y,m} = Z_{y,m} - M_{y,m} \quad (\text{A11})$$

All VPA assessments of numbers-at-age $N_{y,a}$ are computed taking $m=7$.

A.2. The Objective Function

The model is fit to survey abundance and CPUE indices. Contributions by each of these to the objective function (maximised in the fit) are computed as follows.

Calculations assume that the observed abundance indices are log-normally distributed about their expected values:

$$I_{y,a}^i = \hat{I}_{y,a}^i \exp(\varepsilon_{y,a}^i) \quad \text{or} \quad \varepsilon_{y,a}^i = \ln(I_{y,a}^i) - \ln(\hat{I}_{y,a}^i) \quad (\text{A12})$$

where

$I_{y,a}^i$ is the observed abundance index for year y , age a and series i ,

$\hat{I}_{y,a}^i$ is the corresponding model estimate, where

$\hat{I}_{y,a}^i = q^i N_{y,a}$ for begin-year indices or

$\hat{I}_{y,a}^i = q^i N_{y,a} \frac{1 - e^{-Z_{y,a}}}{Z_{y,a}}$ for mid-year indices, and

\hat{q}^i is the constant of proportionality (catchability) for abundance series i .

The objective function of Mayo and Col (2006) is then given by:

$$SS = \sum_{i,y,a} \left[\ln(I_{y,a}^i) - \ln(\hat{I}_{y,a}^i) \right]^2 \quad (\text{A13})$$

The function is minimised by treating the abundances for ages 2 to $m-1$ in year $T+1$ as estimable parameters, where T is the final year. These then define $F_{T,a}$ for $a=1$ to $m-2$, $F_{T,m-1}$ is obtained from equation (A5), and $F_{T,m}$ from equation (A6). Given $F_{T,m-1}$, $N_{T,m-1}$ follows from equation (A4), and then for each year in sequence backwards $N_{T-1,m-2}$ and $F_{T-1,m-2}$ are calculated, with $F_{T-1,m-1}$ and $F_{T-1,m}$ following from equations (A5) and (A6) as in the preceding sentence.

Alternative approach (Alt-VPA)

With this approach, the $N_{y,m}$ are estimated directly for each year to year T and a penalty is added to the objective function so that equation (A7) is satisfied:

$$P_1 = \sum_y \left[\ln(N_{y,m}) - \ln(\hat{N}_{y,m}) \right]^2 / 2\sigma_{plus}^2 \quad (\text{A14})$$

where

$$\hat{N}_{y,m} = N_{y-1,m-1} e^{-M_{m-1}} - C_{y-1,m-1} e^{-M_{m-1}/2} + N_{y-1,m} e^{-M_m} - C_{y-1,m} e^{-M_m/2}$$

and σ_{plus} is set sufficiently small to ensure the equality required.

A further penalty is added so that equation (A6) is satisfied:

$$P_2 = \sum_y \left[\ln(F_{y,m}) - \ln(\hat{F}_{y,m}) \right]^2 / 2\sigma_F^2 \quad (\text{A15})$$

where

$$\hat{F}_{y,m} = \alpha F_{y,m-1} \cdot \quad (\text{A16})$$

and σ_F is set small in the same way as σ_{plus} .

While the process for solving for $N_{y,m}$ and $F_{y,m}$ could be taken sequentially back in time in one year steps as for the previous approach, this becomes more complicated here as each time the solution to two simultaneous non-linear equations is required; thus the minimisation process immediately above is easier to implement.

In implementation, however, it was found that setting σ_F very small (i.e. forcing the equality of equation A16) could lead to unstable estimation behaviour. This arises because of the very small numbers of plus group fish estimated to be caught in some years (see Table CD.5). More robust behaviour was achieved by allowing some variability about the relationship of equation A16 by not setting σ_F too small; results presented in this paper set $\sigma_F=0.35$. Thus the relationship of equation A16 is achieved in an ‘‘average’’ sense, rather than exactly each year.

A.3. Calculation of MSY

If the years with catch-at-age data considered in the VPA are $y = 1$ to T , then the computations above provide a matrix of numbers-at-age estimates, $\{N_{y,a} : y = 1, \dots, T; a = 1, \dots, m\}$. These in turn provide a series of spawning stock-recruitment pairs $\{(B_y^{sp}, N_{y+1,1}) : y = 1, \dots, T-1\}$ where

$$B_y^{sp} = \sum_a f_{y,a} w_{y,a}^{strt} N_{y,a} e^{-M_{a+1} F_{y,a} / 6} \quad (\text{A17})$$

where the formulation of this equation is to allow for cod spawning two months after the start of the year, and

$w_{y,a}^{strt}$ is the mass of fish of age a during spawning, and

$f_{y,a}$ is the proportion of fish of age a that are mature.

A stock recruit function with estimable parameters $p : R_y = f\left(p, B_y^{sp}\right)$ is then fit to these estimates by minimising:

$$SS = \sum_{y=1}^{T-1} \left[\ln(N_{y+1,1}) - \ln \left(f \left(p, B_y^{sp} \right) \right) \right]^2 \quad (\text{A18})$$

to obtain estimates of the parameters p .

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

$$C(F^*) = \sum_a w_a^{mid} PR_a F^* N_a(F^*) e^{-(M_a/2)} \quad (\text{A19})$$

where numbers-at-age a are given by:

$$N_a(F^*) = \begin{cases} R_1(F^*) & \text{for } a = 1 \\ N_{a-1}(F^*) e^{-M_{a-1}} (1 - PR_{a-1} F^*) & \text{for } 1 < a < m \\ \frac{N_{m-1}(F^*) e^{-M_{m-1}} (1 - PR_{m-1} F^*)}{(1 - e^{-M_m} (1 - PR_m F^*))} & \text{for } a = m \end{cases} \quad (\text{A20})$$

where

$$R_1(F^*) = \frac{\alpha B^{sp}(F^*)}{\beta + B^{sp}(F^*)} \quad (\text{A21})$$

for a Beverton-Holt stock-recruitment relationship

or

$$R_1(F^*) = \alpha B^{sp}(F^*) e^{-\beta [B^{sp}(F^*)]^\gamma} \quad (\text{A22})$$

for a modified Ricker stock-recruitment relationship, where the Ricker form results when fixing $\gamma=1$.

The maximum of $C(F^*)$ is then found by searching over F^* to give F_{MSY}^* , with the associated spawning biomass and yield given by

$$B_{MSY}^{sp} = \sum_a f_a w_a^{strt} N_a(F_{MSY}^*) e^{-M_a/6} (1 - PR_a F_{MSY}^*/6) \quad (\text{A23})$$

$$MSY = \sum_a w_a^{mid} PR_a F_{MSY}^* N_a(F_{MSY}^*) e^{-(M_a/2)} \quad (\text{A24})$$

Typically the inputs to MSY calculation are $K^{sp} = B^{sp}(F^* = 0)$ and stock-recruitment steepness h where $h = R_1(B^{sp} = 0.2K^{sp}) / R_1(B^{sp} = K^{sp})$. The values of α and β for equations A21 or A22 are then obtained by solving two simultaneous equations: the first is obtained by substituting equation A20 with $F^* = 0$ into equation A17 to provide an expression for K^{sp} in terms of $R_1(B^{sp} = K^{sp})$, and the second follows from the fact that under equilibrium at K^{sp} , recruitment $R_1(F^* = 0)$ must exactly balance the number of fish dying from natural mortality over the year.

In application (for both VPA and ASPM), the maturity- (f_a) and begin-year weight-at-age (w_a^{strt}) vectors are taken as those for the last year available. The mid-year weight-at-age vector (w_a^{mid}) is

taken as the average over the period with data available (1982-2004) and the partial recruitment (PR_a , equivalently selectivity S_a – see Appendix B) is computed as:

$$PR_a = \frac{\sum_{y=1992}^{2004} F_{y,a} / 13}{\max\left(\sum_{y=1992}^{2004} F_{y,a} / 13\right)} \quad (\text{A25})$$

The relationship between the fishing proportion F^* and fishing mortality F is given by:

$$F = -\ln(1 - F^*) \quad (\text{A26})$$

In these calculations, the plus-group is taken as 11+.

ANNEX II

ADDENDUM 2

Further SCAA/ASPM Assessments of Gulf of Maine Cod Including Data for 2007 and Exploring the Impact of Age-Dependence in Natural Mortality

Doug S. Butterworth and Rebecca A. Rademeyer

MARAM (Marine Resource Assessment and Management Group)
Department of Mathematics and Applied Mathematics
University of Cape Town, Rondebosch 7701, South Africa

INTRODUCTION

The purpose of this Addendum is to provide some further background to one of the arguments offered in the main text for preferring SCAA/ASPM over VPA for the assessment of the Gulf of Maine cod. Specifically this relates to the third bullet of the text section concerned, which argues that when the plus-group is treated in a mathematically consistent way within VPA, results can fail to show robustness to the manner in which asymptotically flat commercial selectivity assumption is imposed on the computations.

In Butterworth and Rademeyer (2008a, Appendix 1, reproduced here as Annex I), an “Alt-VPA” approach was developed to correct the mathematical inconsistency in the manner in which plus-group computations had been carried out for previous ADAPT-VPA applications to this stock. For the asymptotically flat commercial selectivity assumption, this involves associating an effective standard deviation σ_F with the distribution of differences between $F_{y-1,m}$ and $F_{y,m}$, where m is the age of the plus-group. In a penalty added to the sum of squares (SS) function minimised in fitting the Alt-VPA to the abundance index data to promote nearer equality of $F_{y-1,m}$ and $F_{y,m}$, the weight associated with their differences is then taken to be $1/(\sigma_F)^2$, i.e. the lower the σ_F value input, the greater the weight given to satisfying the asymptotically flat selectivity assumption exactly for every year.

In the process of considering alternative mechanisms to account for the relative paucity of older cod in the commercial and surveys catches-at-age, the Alt-VPA method was applied to the updated data for the Gulf of Maine cod stock for the default $M=0.2$ assumption, but treating M for the plus-group ($m=8+$) as a further estimable parameter.

RESULTS AND DISCUSSION

Applications of Alt-VPA for this scenario (an estimable M for the plus-group) were considered across a wide range of values of σ_F . Interestingly the results effectively bifurcated once σ_F fell much below 0.35. Fig. A2.1 shows results for two illustrative values for σ_F (and they differ little from the one or the other of these solutions for other choices for σ_F). For the lower values of σ_F , lower estimated biomasses are indicated, with M for the plus-group at a constraint boundary of 0.01; but for the higher values of σ_F , estimated biomasses are appreciably larger, and the plus-group M is estimated to be large (0.86 for the case shown). The fit to the abundance index data is better for the higher of the σ_F values considered ($SS = 251.2$ compared to 266.0, i.e. the abundance indices “prefer” some variability rather than exactness in the asymptotically flat commercial selectivity assumption), though the difference between these values drops from 14.9 to only 1.5 when the penalty associated with deviations from flatness is taken into account.

Table A2.1 compares the F matrices for the two values of σ_F considered. When more weight is given to achieving exact flatness in commercial selectivity between ages $m-1$ and m each year, it is noticeable that at times the associated F reaches very high values (higher perhaps than might be thought realistic).

Nevertheless, however one might interpret these F matrices, the salient point is that VPA can yield very different results depending on the manner in which the asymptotically flat selectivity assumption is imposed. In actuality there would be variation about such a relationship, not exactness, so the solution corresponding to $\sigma_F \rightarrow 0$ is not necessarily to be preferred. Results depend (in the case examined, at least) on the value input for σ_F , but there is no immediately clear basis for the choice for this value.

It is this indeterminacy that gave rise to the comment about a lack of robustness of VPA in the main text.

$\sigma_F=0.01$

	1	2	3	4	5	6	7	8+
1982	0.010	0.221	0.670	0.668	0.664	0.532	0.719	0.719
1983	0.002	0.262	0.587	0.850	0.824	1.196	0.813	0.813
1984	0.003	0.145	0.573	0.793	1.042	0.639	0.670	0.670
1985	0.007	0.147	0.736	1.139	0.924	1.009	0.533	0.533
1986	0.001	0.039	0.696	0.832	0.999	0.811	1.312	1.312
1987	0.008	0.125	0.417	1.111	0.892	1.178	1.021	1.021
1988	0.000	0.060	0.492	0.764	1.233	0.289	0.410	0.410
1989	0.001	0.029	0.380	0.985	0.947	1.030	0.418	0.418
1990	0.000	0.084	0.335	0.935	0.981	1.197	1.022	1.022
1991	0.000	0.157	0.838	0.866	1.258	0.962	1.379	1.379
1992	0.000	0.068	0.326	1.010	1.383	1.288	1.671	1.671
1993	0.000	0.028	0.721	0.987	0.817	1.559	0.556	0.556
1994	0.000	0.008	0.508	1.481	2.026	1.546	1.763	1.763
1995	0.005	0.128	0.245	1.042	1.575	1.187	2.162	2.162
1996	0.000	0.034	0.476	0.823	1.280	1.356	0.567	0.567
1997	0.000	0.034	0.263	0.771	1.159	0.890	0.657	0.657
1998	0.000	0.029	0.269	0.582	0.683	0.915	1.033	1.033
1999	0.000	0.007	0.242	0.589	0.583	0.575	1.415	1.415
2000	0.000	0.037	0.226	0.559	0.397	0.432	0.306	0.306
2001	0.000	0.051	0.312	0.508	0.540	0.418	0.509	0.509
2002	0.000	0.004	0.119	0.407	0.525	0.636	0.636	0.636
2003	0.000	0.005	0.190	0.305	0.757	0.830	0.800	0.800
2004	0.000	0.001	0.131	0.770	0.516	0.655	0.712	0.712
2005	0.000	0.001	0.142	0.490	0.776	0.764	0.720	0.720
2006	0.000	0.000	0.027	0.498	0.557	0.669	0.719	0.719
2007	0.000	0.000	0.039	0.190	0.500	0.721	0.366	0.366

$\sigma_F=0.35$

Year	1	2	3	4	5	6	7	8+
1982	0.010	0.208	0.634	0.536	0.470	0.223	0.208	0.228
1983	0.001	0.248	0.538	0.757	0.543	0.587	0.220	0.288
1984	0.003	0.136	0.530	0.669	0.792	0.309	0.185	0.267
1985	0.007	0.145	0.666	0.951	0.644	0.541	0.183	0.203
1986	0.001	0.038	0.679	0.675	0.639	0.401	0.358	0.497
1987	0.008	0.122	0.405	1.046	0.578	0.448	0.291	0.316
1988	0.000	0.058	0.477	0.725	1.028	0.146	0.088	0.121
1989	0.001	0.029	0.364	0.920	0.840	0.643	0.175	0.161
1990	0.000	0.082	0.329	0.861	0.825	0.871	0.387	0.515
1991	0.000	0.147	0.796	0.838	0.998	0.645	0.598	0.710
1992	0.000	0.065	0.299	0.890	1.251	0.679	0.607	0.589
1993	0.000	0.026	0.676	0.836	0.615	1.077	0.172	0.149
1994	0.000	0.007	0.461	1.224	1.124	0.752	0.558	0.410
1995	0.004	0.103	0.220	0.852	0.823	0.248	0.313	0.237
1996	0.000	0.028	0.361	0.686	0.753	0.284	0.055	0.034
1997	0.000	0.027	0.209	0.483	0.755	0.295	0.057	0.055
1998	0.000	0.023	0.212	0.417	0.307	0.370	0.167	0.075
1999	0.000	0.006	0.188	0.414	0.342	0.175	0.256	0.111
2000	0.000	0.030	0.182	0.390	0.236	0.200	0.067	0.035
2001	0.000	0.040	0.243	0.376	0.309	0.206	0.178	0.115
2002	0.000	0.003	0.091	0.289	0.330	0.272	0.232	0.233
2003	0.000	0.004	0.164	0.222	0.439	0.385	0.215	0.343
2004	0.000	0.001	0.099	0.613	0.330	0.266	0.204	0.238
2005	0.000	0.001	0.113	0.342	0.500	0.371	0.185	0.251
2006	0.000	0.000	0.023	0.366	0.322	0.315	0.224	0.211
2007	0.000	0.000	0.035	0.159	0.314	0.304	0.127	0.140

Table A2.1: Fishing mortality at age for the Gulf of Maine cod VPA assessments for which M for the plus-group is treated as an estimable parameter for $\sigma_F = 0.01$ and $\sigma_F = 0.35$.

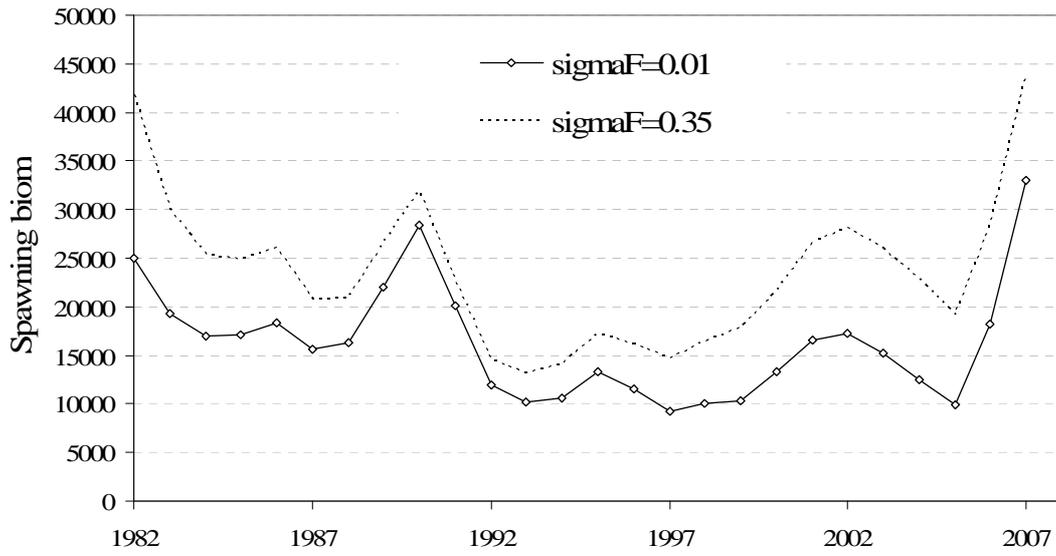


Fig. A2.1: Spawning biomass trajectories for the Gulf of Maine cod VPA assessment for which M for the plus-group is treated as an estimable parameter for $\sigma_F = 0.01$ and $\sigma_F = 0.35$.