

SPECIFICATIONS FOR THE ADAPT VPA CODE, SEPTEMBER 1996 (REVISED)

Geromont, H.F., D.S. Butterworth¹

1 Introduction

The report of the Bluefin Tuna Methodology Session, which took place at the ICCAT Headquarters in Madrid from 16-19 April 1996, prioritised extensions to made to software for use at the September 1996 Bluefin Tuna Assessment meeting in Genoa. Brief descriptions of the ADAPT code of André Punt which was utilised at the 1994 ICCAT SCRS meeting and the subsequent extensions are provided below. For readers' convenience, this paper has subsequently been revised to include indications of the selections made for the base case ADAPT assessments for the western and eastern Atlantic bluefin tuna at the Genoa meeting — these indications are given in italics. A copy of this code is held by the ICCAT Secretariat.

2 Model

The fundamental age-structured equations describing the dynamics are:

$$N_{y+1,a+1} = N_{y,a} e^{-(F_{y,a} + M_{y,a})} = N_{y,a} e^{-Z_{y,a}} \quad (1)$$

and:

$$C_{y,a} = \frac{N_{y,a} F_{y,a} (1 - e^{-Z_{y,a}})}{Z_{y,a}} \quad (2)$$

or, substituting equation 1 into equation 2:

$$C_{y,a} = \frac{N_{y+1,a+1} F_{y,a} (e^{Z_{y,a}} - 1)}{Z_{y,a}} \quad (3)$$

where

$N_{y,a}$ is the number of fish of age a at the start of year y , estimated by the model,

$F_{y,a}$ is the instantaneous fishing mortality rate on fish of age a during year y , estimated by the model,

$M_{y,a}$ is the instantaneous natural mortality rate on fish of age a during year y , which is assumed time-invariant (i.e. independent of y) and is input, and

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

Given initial guesses for the terminal population numbers ($N_{n+1,a}$ — see sections 3.1 and 3.2 below), the VPA consists of estimating the remaining numbers-at-age ($N_{y,a}$) and fishing mortalities-at-age ($F_{y,a}$) by proceeding backwards up each cohort by solving equations 2 and 3 successively. The plus-group population numbers are calculated as follows:

$$N_{y+1,m} = N_{y,m-1} e^{-Z_{y,m-1}} + N_{y,m} e^{-Z_{y,m}} \quad (4)$$

where

m is the oldest age, taken to be a plus-group,

$N_{y,m-1}$ and $N_{y,m}$ are given by equation 2, and

$F_{y,m}$ is defined by equation 5 below.

3 Parameters

3.1 Estimating terminal population numbers

The population numbers ($N_{n+1,a}$) in the year following the last year ($y = n$) for which catch-at-age data are available are estimated in the model fitting procedure.

The 1996 base case ($n = 1995$) involved estimating population numbers at the start of 1996 for ages $a=3, 5, 7$ and 9 for the western Atlantic, and ages $a=3, 7$ and 9 for the eastern Atlantic.

3.2 Selectivity of terminal year ($y = n$)

Given values for the estimable parameters specified above ($N_{n+1,a}$), fishing mortalities ($F_{n,a-1}$) can then be computed. For those ages a for which population numbers are not estimated, fishing mortalities cannot be calculated directly. For such ages, the F s need to be linked to those fishing mortalities that are directly estimated via a vector specifying which ages are grouped, as well as the relative selectivities within each group.

Note: The 1996 base case assessment assumes for the western Atlantic that ages 1-3, 4-5, 6-7 and 8-9 are grouped and that:

$$F_{1995,1} = 0.318F_{1995,2} = 0.318F_{1995,3}; F_{1995,4} = F_{1995,5}; F_{1995,6} = F_{1995,7}; \text{ and } F_{1995,8} = F_{1995,9},$$

and for the eastern Atlantic that the groupings are 1-3, 4-7 and 8-9 and that:

$$F_{1995,1} = 0.6F_{1995,2} = 0.6F_{1995,3}; F_{1995,4} = F_{1995,5} = F_{1995,6} = F_{1995,7}; \text{ and } F_{1995,8} = F_{1995,9}.$$

3.3 Estimating the F ratio for the plus group

Fishing mortalities F for the oldest age m (taken to be a plus-group) and any year y are given by:

$$F_{y,m} = r_y F_{y,m-1} \quad (5)$$

where r_y is the F ratio which can either be input or now also estimated for any year y , or blocks of years.

Note: The 1996 base case assessments correspond to $m = 10$ and to the F ratio being pre-specified ($r_y = 1.0$) for the period 1970-1973, a single value estimated for the period 1974-1981, and pre-specified (western Atlantic: $r_y = 1.14$, eastern Atlantic: $r_y = 2.02$) for the most recent period (1982-1995). The values assumed for the most recent period are as adopted for the 1994 assessments.

¹ Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7700, South Africa.

3.4 Natural mortality rates

The natural mortality rate M_a input can either be constant or age-dependent.

Note: M_a is assumed age-independent (=0.14) in the 1996 base case assessments.

4 Data

Aggregated, as well as fleet disaggregated, catches-at-age are input for a period (1970 to 1995 for the 1996 base case assessments), in addition to mid-year masses-at-age for the same period.

Values of indices of abundance used to fit the model are input. They can either be indices in terms of population numbers or biomass, and relate to the beginning or middle of the year as well as to any pre-specified range of ages. The standard errors (i.e. square roots of variances) associated with each value (σ_y^i) may also be input. When computing the expected value of each index (\hat{I}_y^i) from the model estimates of biomass-at-age or numbers-at-age, "selectivities" are specified to give a relative weight to each age within the range specified. There are three possible choices:

- "uniform": equal weighting by age
- "partial catches": first partial fishing mortalities are calculated for the associated fleet (f):

$$F_{y,a}^f = F_{y,a} \frac{C_{y,a}^f}{C_{y,a}} \quad (6)$$

and then the weights for a year are given by $F_{y,a}^f$ divided by the largest $F_{y,a}^f$ for that fleet and year within the age range specified.

- "pre-specified": weights for each year and the age range under consideration are input on the basis of consideration of the catch-at-age matrix ($C_{y,a}^f$) for the fleet concerned. *Note: For the 1996 assessments, this applies to the Canadian tended line index for the western Atlantic only, and is intended to make allowance for this index applying to fish aged 13+ rather than 10+. (A similar adjustment was made to the Canadian Hell Hole fishery index which was included in a sensitivity test.)*

5 Model fitting and variance estimation

5.1 Fitting criteria

5.1.1 Index variances for weighting individual index values

Note: 1996 base case option, which assumed log-normally distributed errors.

In this case the model parameters are estimated by minimising the function (negative of the log-likelihood, less constants):

$$-\ln L = \sum_i \left[\sum_y (\epsilon_y^i)^2 / 2(\sigma_y^i)^2 + \ln \sigma_y^i \right] \quad (7)$$

where σ_y^i is the standard error of the value for abundance series i in year y , which is input, and

$$\epsilon_y^i = I_y^i - \hat{q}^i \hat{I}_y^i \quad \text{for normally distributed errors} \quad (8)$$

or:

$$\epsilon_y^i = \ln(I_y^i) - \ln(\hat{q}^i \hat{I}_y^i) \quad \text{for log-normally distributed errors} \quad (9)$$

where

I_y^i is the abundance index for year y and series i ,

\hat{I}_y^i is the corresponding model estimate, and

\hat{q}^i is the catchability coefficient for abundance series i , estimated by its maximum likelihood value:

$$\hat{q}^i = \frac{\sum_y 1/(\sigma_y^i)^2 I_y^i \hat{I}_y^i}{\sum_y 1/(\sigma_y^i)^2 \hat{I}_y^i \hat{I}_y^i} \quad \text{for normally distributed errors} \quad (10)$$

or:

$$\ln \hat{q}^i = \frac{\sum_y 1/(\sigma_y^i)^2 (\ln I_y^i - \ln \hat{I}_y^i)}{\sum_y 1/(\sigma_y^i)^2} \quad \text{for log-normally distributed errors} \quad (11)$$

5.1.2 Input weighting

In this case the model parameters are estimated by minimising a function of the form:

$$-\ln L = \sum_i w_i \left[\sum_y (\epsilon_y^i)^2 \right] \quad (12)$$

where

ϵ_y^i is defined by equations 8 and 9,

w_i is the weight given to each abundance index ($w_i = 1$ for "equal weighting"), and

\hat{q}^i is the catchability coefficient for abundance index i , estimated by its maximum likelihood value:

$$\hat{q}^i = \frac{\sum_y I_y^i \hat{I}_y^i}{\sum_y \hat{I}_y^i \hat{I}_y^i} \quad \text{for normally distributed errors} \quad (13)$$

or:

$$\ln \hat{q}^i = \frac{\sum_y (\ln I_y^i - \ln \hat{I}_y^i)}{n^i} \quad \text{for log-normally distributed errors} \quad (14)$$

where n^i is the number of data points in abundance series i .

Note that "equal weighting" means that each point within each series receives equal weight, so that a series of twice the length (double the number of points) receives twice the weight. Further, "input weighting" corresponds to the special case of index variance weighting where σ_y^i is independent of y .

5.1.3 Maximum likelihood (equivalent to "iterative re-weighting")

In this case index variances and catchabilities are estimated in the fitting procedure by their maximum likelihood values. Unlike section 5.1.1 above, this assumes that $\sigma_y^i = \sigma^i$, i.e. all points in the same series have the same variance. The objective function minimised is then given by (negative of the log-likelihood, less constants):

$$-\ln L = \sum_i n^i \ln \sigma^i + \frac{n^i}{2} \quad (15)$$

where

n^i is the number of data points in abundance series i ,

σ^i is the standard deviation for abundance series i , estimated by its maximum likelihood value:

$$\hat{\sigma}^i = \sqrt{\frac{\sum_y (\epsilon_y^i)^2}{n^i}} \quad (16)$$

where

ϵ_y^i is defined by equation 8 or 9, and

q^i is defined by equation 13 or 14.

5.2 Variances

5.2.1 Bootstrapping of the index data

Note: 1996 base case option, which assumed log-normally distributed errors.

Multiple pseudo-data sets are generated from each abundance index by adding error using the input standard errors, estimated by (say) a GLM, of each index:

$$I_y^{i,U} = I_y^i + \epsilon_y^{i,U} \quad \text{for normally distributed errors} \quad (17)$$

or:

$$I_y^{i,U} = I_y^i \exp \epsilon_y^{i,U} \quad \text{for log-normally distributed errors} \quad (18)$$

where

$$\epsilon_y^{i,U} \sim N(0, (\sigma_y^i)^2),$$

$I_y^{i,U}$ is the abundance index for year y and series i in bootstrap data set U ,

I_y^i is the actual abundance index for year y and series i , and

σ_y^i is the standard error for year y for abundance series i estimated by (say) a GLM, which is input.

5.2.2 Bootstrapping of the index residuals

The corresponding formulas for the parametric bootstrap which generates pseudo data from model-fitted values are:

$$I_y^{i,U} = \hat{q}^i \hat{I}_y^i + \epsilon_y^{i,U} \quad \text{for normally distributed errors} \quad (19)$$

or:

$$I_y^{i,U} = \hat{q}^i \hat{I}_y^i \exp \epsilon_y^{i,U} \quad \text{for log-normally distributed errors} \quad (20)$$

where

$$\epsilon_y^{i,U} \sim N(0, (\hat{\sigma}^i)^2),$$

\hat{I}_y^i is the pertinent model estimate, and

\hat{q}^i and $\hat{\sigma}^i$ are also as estimated in the fitting procedure, as defined by equations 13 or 14 and 16.

5.2.3 Bootstrapping of the F ratios

For any year, or block of years, for which the F ratio (r) is estimated in the fitting process, this is simply re-estimated given the corresponding pseudo data for each replicate. For years for which a fixed value is input for the fit itself, to allow for the F ratios in any year y to fluctuate about the constant assumed for that year, bootstrap replicates are generated as follows:

$$F_{y,m}^U = r_y \hat{F}_{y,m-1} \exp \eta_y^U \quad (21)$$

where

$$\eta_y^U \sim N(0, \sigma_y^2),$$

$$\sigma_y = 0.4 \text{ for } y < 1975,$$

$$\sigma_y = 0.25 \text{ for } y \geq 1975,$$

$\hat{F}_{y,m-1}$ is the model estimated fishing mortality for year y and age $m-1$ where m is the oldest age taken to be a plus-group, and

r_y is the F ratio input by the user.

6 Additional options

Note: None of these options were applied for the 1996 base case assessments.

6.1 Penalty function basis to include stock-recruit relationship

The number of recruits is assumed to be related to the spawner stock size by a Beverton-Holt stock-recruit relationship with auto-correlated stochastic deviations:

$$N_{y,a_{min}}^R = \frac{\alpha B_{y-a_{min}}^{sp}}{\beta + B_{y-a_{min}}^{sp}} e^{\chi_y} \quad (22)$$

where

$$\chi_y = \rho \chi_{y-1} + \sqrt{1 - \rho^2} \epsilon_y,$$

$$\epsilon_y \sim N(0, \sigma_\epsilon^2),$$

ρ is the auto-correlation coefficient, which is input,

a_{min} is the smallest age of animals present in the catch (input as 0 or 1),

α and β are the stock-recruit relationship parameters estimated in the fitting procedure, and

B_y^{sp} is the spawner stock size at the start of year y :

$$B_y^{sp} = \sum_{a=a_{min}}^m f_a w_a N_{y,a} \quad (23)$$

where

w_a is the begin-year mass of a fish of age a , and

f_a is the proportion of fish of age a that are mature.

The contribution of the penalty function (if included) to the quantity minimised (the negative of the log-likelihood function) is therefore:

$$n_R \ln \sigma_R + 1/(2\sigma_R^2) \sum_{y=y_1}^{y_2} \left[\frac{\chi_y - \rho \chi_{y-1}}{\sqrt{1-\rho^2}} \right]^2 \quad (24)$$

where

$$\chi_y = \ln N_{y,a_{min}} - \ln N_{y,a_{min}}^R$$

y_1 and y_2 are the first and last years considered in the penalty function, which are input ($y_1 > 1970$ and $y_2 \leq y_n$),

$$n_R = y_2 - y_1 + 1, \text{ and}$$

σ_R is the associated standard deviation, which is input.

Note that in the interests of simplicity, expression 24 omits a term in χ_{y_1} , which generally is of little quantitative consequence to values estimated.

6.2 Selectivity of terminal year ($y = n$) from average last p years

Let S^* denote the relative selectivity for age $a' - i$ and year y , such that

$$S_{y,a'-i}^* = F_{y,a'-i} / F_{y,a'} \quad (25)$$

where

$$a'' < a' - i < a', \text{ and}$$

a'' and a' are successive ages for which the $F_{n,a}$ are treated as estimable parameters in the most recent year n ,

then $F_{n,a'-i}$ is given by

$$F_{n,a'-i} = F_{n,a'} \frac{1}{p} \sum_{y=n-p}^{n-1} S_{y,a'-i}^* = F_{n,a'} \frac{1}{p} \sum_{y=n-p}^{n-1} \frac{F_{y,a'-i}}{F_{y,a'}} \quad (26)$$

where p is the number of years over which the relative selectivity is averaged for each age class.

6.3 Constraints to reduce retrospective patterns

A penalty function to reduce trends in recent F values may be included within the minimisation as follows: for each age a , a linear regression is fitted to $\log F$ as a function of year y for the last q years:

$$\ln F_{y,a} = m_a y + c_a \quad \text{for } n - q + 1 \leq y \leq n \quad (27)$$

where

n is the last year considered, and

m_a is the slope of the q most recent F values for a particular age a (from standard linear regression formula).

The contribution of the associated penalty function to the quantity minimised is then:

$$w^T \sum_{a=1}^m w_a m_a^2 \quad (28)$$

where

m is the maximum age considered,

w^T is the overall weight, and

w_a is a relative weight vector to allow for possible differential weighting (even omission) of each age a .

6.4 Retrospective adjustments

Population numbers N for any age $a > a_{min}$ at the start of the most recent year n can be adjusted to reduce retrospective patterns as follows:

$$N_{n,a}^{adj} = \hat{N}_{n,a} \delta_a \quad (29)$$

where

$N_{n,a}^{adj}$ is the adjusted number of fish of age a at the start of year n ,

$\hat{N}_{n,a}$ is the number of fish of age a at the start of year n estimated by applying ADAPT, and

δ_a is the retrospective adjustment factor for age a ,

given by:

$$\delta_a = \left\{ \prod_{y=n-p}^{n-1} \hat{N}_{y,a}^{y+1} / \hat{N}_{y,a}^y \right\}^{\frac{1}{p}} \quad (30)$$

where

p is the number of years over which the population numbers are averaged for each age $a > a_{min}$ (a_{min} being the smallest age of animals present in the catch), and

$\hat{N}_{y,a}^y$ is the number of fish of age a at the start of year y estimated by applying ADAPT with the most recent year set equal to y' .

Note: No retrospective adjustment is applied to $\hat{N}_{n,a_{min}}$, as it is set equal to the geometric mean of past recruitments over the period 1981 to $n - 3$.

If a bootstrap procedure is implemented to estimate precision in this case, when the parametric bootstrap is applied (see equation 19), the model estimates \hat{F}_y^i are based on the ADAPT estimates $\hat{N}_{y,a}$, NOT on the values of $N_{y,a}$ that follow from the adjustment equation 29.