



# Modeling terminal-year fishing mortality rates in western Atlantic bluefin tuna virtual population analyses

John F. WALTER<sup>a</sup> and Clay E. PORCH

U.S. Department of Commerce, NOAA National Marine Fisheries Service, Southeast Fisheries Science Center, Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, FL33149, USA

Received 1st February 2012; Accepted 6 November 2012

**Abstract** – Virtual population analysis requires information on the fishing mortality rate (or abundance) for one age group from each cohort. In many cases available data are insufficient to estimate these rates for every age group and structural assumptions must be imposed to reduce the number of effective parameters. Past assessments of western Atlantic bluefin tuna (*Thunnus thynnus*) have reduced the number of parameters by assuming pre-specified values for the ratios of the fishing mortality rates on adjacent age groups. More recent bluefin tuna assessments have estimated terminal fishing mortality rates ( $F$ ) on each age with a constraint that restricts change from one year to the next. We explore the implications of these methods of estimating terminal-year  $F$  through retrospective analyses of the 2006 bluefin tuna assessment and through stochastic simulations. The use of pre-specified ratios for  $F$  created strong retrospective biases and may have led to overly optimistic projections. Constraining annual changes in the terminal-year  $F$  appeared to mute retrospective patterns and resulted in abundance projections less prone to spurious initial leaps. Simulation results indicate that the constraint improves estimation, particularly with moderate to low interannual changes in selectivity.

**Keywords:** Stock assessment / Virtual population analysis / Fishing mortality rate / Retrospective patterns / Age-based population model / *Thunnus thynnus*

## 1 Introduction

Atlantic bluefin tuna are among the most highly sought after and valued of marine fish and their stock status is of great international concern. They are managed through the International Convention for the Conservation of Atlantic Tuna (ICCAT) under the presumption that there are distinct eastern and western populations that spawn in the Mediterranean Sea and Gulf of Mexico, respectively (Fromentin and Powers 2005). Recent stock assessments suggest that the abundance of both stocks may be below the Convention objectives.

The status of Atlantic bluefin tuna has been assessed primarily with virtual population analysis (VPA), a class of age-structured stock assessment tools that assumes the catch history of any given year class is known with negligible error (Butterworth and Rademeyer 2008). While a number of alternatives allow for errors in catch, VPAs continue to enjoy widespread use, owing to the relatively minimal assumptions and data required (Gulland 1983; Megrey 1989). Typically, VPAs employ a backwards recursion that calculates the annual abundance of each year class from its abundance in a

subsequent year by adding in catch and natural mortality. Doing so requires information about the abundance or fishing mortality rates ( $F$ ) on all ages in the last year of the model and also on the oldest age in each year. Of particular interest is the  $F$  on the incomplete cohorts observed during the terminal year (Megrey 1989), which form the basis of the future fishery and spawning stock biomass (SSB). In this paper we address only the terminal-year  $F$  and not the terminal-age  $F$ .

Various approaches to estimating terminal-year and terminal-age  $F$  have been proposed, some by employing ad hoc structural assumptions to reduce the number of parameters (Smith and Gavaris 1993; Patterson and Kirkwood 1995) and others by incorporating ancillary data in the form of tagging experiments (Parks 1976), abundance indices (Doubleday 1981; Parrack 1986; Chen et al. 2008) or effort (Pope and Shepherd 1985). The latter group of methods, generally described as “tuning”, work best when ancillary information is available for each age group or each year for which terminal- $F$  are needed (Haddon 2001). Otherwise, it might still be necessary to reduce the effective number of parameters. ICCAT assessments of western Atlantic bluefin tuna conducted between 1994 and 2008 were, for example, tuned via a statistical fit to several indices of abundance but also reduced the number

<sup>a</sup> Corresponding author: john.f.walter@noaa.gov

of estimated terminal-year  $F$  parameters by linking adjacent age groups through the use of a pre-specified vector (Anon. 1997):

$$F_1 = 0.318 \times F_2; \quad F_3 = F_2; \quad F_5 = F_4; \quad F_7 = F_6; \quad F_9 = F_8 \quad (1)$$

where  $F_i$  is the fishing mortality rate for age  $i$ , and only  $F_2$ ,  $F_4$ ,  $F_6$  and  $F_8$  are estimated. This vector in (1) was originally derived from a separable VPA (Pope and Shepard 1982) conducted for an earlier ICCAT assessment and reflects historical selectivity patterns. The oldest age group of the VPA is a plus group (ages 10 and older) and its corresponding fishing mortality rate in the terminal and in all prior years was determined by a ratio of  $F_{10}$  to  $F_9$ . Previous assessments have determined this ratio according to a schedule which fixed it at 1 for 4 years, then a single value was estimated for the next 8 years and then another value was estimated for the remainder of the time series (Anon. 1998). While this ratio and its specification was of concern for previous VPAs, the most recent ICCAT stock assessment for bluefin tuna (Anon. 2011) expanded the modeled ages to 16+ allowing estimation of changes in selectivity that might have been important between ages 9 and 10 but are less consequential between ages 15 and 16. Hence, we do not address the  $F_{10}/F_9$  ratio here.

The assignment of assumed ratios to the fishing mortality rates on adjacent ages in the terminal year can introduce important biases, which may manifest themselves in the form of strong retrospective patterns (a situation where the parameter estimates for a given year exhibit a systematic bias as successive years of data are removed from the model, Mohn 1999). Such retrospective biases may explain the observation that projections of the future abundance of bluefin tuna appear to have been consistently overly optimistic (Gavaris et al. 2008). An alternative approach is to estimate the terminal-year  $F$  for each age, either as free parameters (recognizing that they may not be well-determined, particularly without age-specific abundance indices) or, as we propose in this paper, subject to constraints that restrict the amount of change in the relative fishing mortality rate (selectivity) at age vector in the last several years of the model.

The objectives of this paper are 1) to examine previous assessments for empirical evidence of whether the fixed-ratio method of estimating terminal-year  $F$  may have resulted in misspecified ratios when viewed in the hindsight of later assessments; 2) to perform retrospective analyses using the 2006 assessment data to examine the effects of allowing relative  $F$  to vary freely or with a restriction; and 3) to perform stochastic simulations to evaluate the efficacy of a penalty restricting change in relative fishing mortality rate and to determine appropriate values for the penalty. While bluefin tuna is the focus of this paper, comparisons of the methods of modeling terminal- $F$  are applicable to many VPA analyses.

## 2 Materials and methods

### 2.1 Virtual population analysis (VPA) model

We employ the assessment program VPA-2BOX version 3.01 (Porch 2003) based upon the ADAPT framework

(Parrack 1986; Gavaris 1988) and available from the ICCAT website ([www.iccat.es](http://www.iccat.es)). The virtual population analysis (VPA) configuration consists of a single catch at age matrix and 12–14 separate catch per unit effort (CPUE) indices which correspond to different fleets or surveys (Anon. 2011). None of the indices represent a single age; rather they reflect an amalgam of ages which is specified in the model by the survey or fleet-specific catch at age. Furthermore, all of the indices are measured with considerable uncertainty such that they do not give a clear interpretation of age-specific relative abundance and hence provide further motivation for the simplifying assumptions described in equation 1 above. Parameter estimates (terminal-age and terminal-year fishing mortalities, index catchability and variances and others) are estimated by minimizing the penalized negative log-likelihood which is the sum of the likelihood terms quantifying the fits to the indices and any constraints (Porch 2003).

### 2.2 Proposed constraints on changes in relative fishing mortality

In the absence of informative age-specific abundance indices, terminal-year  $F$  estimates tend to be poorly determined, resulting in imprecise VPA results. One solution is to take advantage of the common observation that the evolution of fishing practices tends to occur gradually over a period of years. Such a rationale forms the basis for the concepts of regularization and shrinkage in extended survivors analysis (Shepherd 1999), where the amount of annual change in  $F$ -at-age is restricted by regressing current estimates of fishing mortalities on a moving average (Lassen and Medley 2001).

The proposed constraint is similar in concept, but is added to the likelihood function as a penalty term and involves only the relative fishing mortality rates (rather than the absolute values).

Specifically, the change in the relative  $F$  vector is modeled as a correlated random walk:

$$\text{penalty} = \sum_{y=Y-n+2}^Y \sum_{a=\alpha}^A \left\{ 0.5 \left( \frac{\ln v_{a,y} - \ln v_{a,y-1}}{\sigma_v} \right)^2 + \ln \sigma_v \right\} \\ v_{a,y} = F_{a,y} / \max_a F_{a,y} \quad (2)$$

where  $v_{a,y}$  and  $F_{a,y}$  are the relative and absolute fishing mortality rate, respectively, on age  $a$  in the year  $y$  ( $Y$  is the last year in the time series,  $A$  is the oldest age and  $\alpha$  is the youngest age, assumed to be 1 for bluefin tuna). The variable  $n$  represents the number of years to be included in the penalty and  $\sigma_v$  is the log-scale standard error (which controls the severity of the penalty). Fishing mortality parameters for each age group are estimated independently without imposing a selection curve.

The penalty in (2) allows the absolute level of fishing mortality to vary freely from one year to the next, while restricting only the amount of change in the age-specific pattern of  $F$  or, equivalently, selectivity. This is consistent with a situation of a relatively static fleet capacity in terms of vessels, trawl, mesh or hook sizes, but with variable absolute levels of effort in terms of days or nets fished (due to changing catch quotas, for example). Porch (2003) initially recommended a value of  $\sigma_v$

between 0.2 and 1.0, as values of  $\sigma_v$  much smaller than the standard errors of the indices of abundance cause the penalty to dominate the likelihood. Porch (2003) further recommended that the value of  $\sigma_v$  should be large ( $>1.0$ ) if there is evidence to suggest that the selectivity (age or size-based relative  $F$ ) has changed substantially during the last  $n$  years of the time series, for instance due to a major change in regulations. We evaluate appropriate penalty values in *stochastic simulations*, below.

The penalty, below, is added to the negative log-likelihood function to be minimized:

$$L(\vec{I}) = \sum_i \sum_y \left\{ 0.5 \left( \frac{\ln I_{iy} / \hat{I}_{iy}}{\sigma_{iy}} \right)^2 + \ln \sigma_{iy} \right\} \quad (3)$$

*+vulnerability penalty*

where  $I_{iy}$  is the observed and  $\hat{I}_{iy}$  the expected index values for series  $i$  in year  $y$  and  $\sigma_{iy}$  is the index standard deviation using a lognormal error structure.

### 2.3 Retrospective analyses of empirical data

We used the same VPA-2BOX data input and control files as used for the 2006 bluefin tuna assessment (Anon. 2007) to perform retrospective analyses going back five years using three methods of estimating terminal-year fishing mortalities:

- 1) fixed-ratio- terminal-year  $F$  for ages 2, 4, 6, and 8 estimated as free parameters and the remaining terminal-year  $F$  determined from the schedule in equation (1);
- 2) terminal-year  $F$  for all ages (1–9) estimated as free parameters (no penalty);
- 3) terminal-year  $F$  for all ages (1–9) estimated subject to the penalty in equation (2).

In this case we have linked the relative  $F$  for ages 1–9 over 3 years with a standard deviation ( $\sigma_v$ ) of 0.5 but also explore the implications of a range of standard deviation values between 0.1 and 0.9 through the stochastic projections described in the next section. A link over 3 years was chosen as the fish do not enter the catch at age or the indices until about age 3. In the simulations, described below we change the link to two years because we impose a change in relative  $F$  at age in the last year, but we also explore the sensitivity of the results to 2, 3, 4 and 5 years of linkage.

We employed two statistics to describe a) the degree of retrospective bias ( $\rho$ ) and b) the degree of retrospective error ( $|\rho|$ ). The statistic  $\rho$ , is similar to that proposed by Mohn (1999) except we calculate it for all years going back to the number of years retrospectively removed ( $k$ ) rather than just the terminal year. The statistic is calculated as the sum of the differences between the retrospective estimates for a given year and the estimates obtained from the entire time series, divided by the full time series estimates:

$$\rho = \sum_{y_2-k}^{y_2} [X_{(y_1:y_2-k),y_2-k} - X_{(y_1:y_2),y_2-k}] / X_{(y_1:y_2),y_2-k} \quad (4)$$

Where  $X$  represents a quantity of interest,  $X_{(y_1:y_2-k),y_2-k}$  represents the value of  $X$  for year  $y_2 - k$  based on running the

assessment with 1, 2, ...  $k$  years removed. In this analysis we use estimates of numbers at age and apical fishing mortality ( $F$  on the fully selected age) for  $X$ . The statistic represents the sum of the differences between retrospective estimates and the full time series estimates. As this statistic is signed, it measures consistent retrospective bias above or below the values estimated for the full time series.

A second statistic,  $|\rho|$ , measures the absolute deviation between retrospective values and the values estimated for the full time series and is the same equation as (3) but with the absolute value of the quantity in the numerator. Both statistics are useful measures of performance as  $\rho$  measures the degree of retrospective bias (consistent under or overestimation of terminal year values with successive removals of a year of data) and measures absolute variability. High values for either statistic indicate poor VPA performance.

### 2.4 Stochastic simulations

Simulations were performed to test the performance of the three methods (fixed ratio, penalty and no penalty) and to explore the strength of the penalty that should be applied under various degrees of change in the terminal year relative fishing mortality. Simulations were conducted with datasets conditioned on the 2006 bluefin tuna VPA. The catch at age matrix and indices of abundance were altered to reflect the relative  $F$  changes described below. A set of 6 relative abundance indices (for ages 1, 3, 5, 7, 9 and 10) were created by dividing the simulated abundance at the given age for each year by the mean for that age over all years. For simplicity, we created relative abundance indices for these specific ages, rather than using indices that span several ages as in the VPA.

Stochastic simulations were conducted by randomly altering the terminal year relative  $F$  vector and calculating the corresponding expectations for the catch at age and indices of abundance in the last year. To explore the strength of the penalty that should be applied versus the degree of change in terminal year  $F$ , we replaced the  $F$  at age by a random value drawn from a uniform distribution with lower bound equal to the maximum of either zero or the original  $F$  minus a predetermined value and an upper bound equal to the original  $F$  plus the predetermined value. A vector of predetermined values (0.01, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8) produced vectors that diverged from the original vector very little (for 0.01) to very much (for 0.8). Simulations used 100 replicates for each value. To categorize the divergence we used the angular deviation (AD) which provides a quantitative measure of the degree of divergence between two vectors (*vec1* and *vec2*):

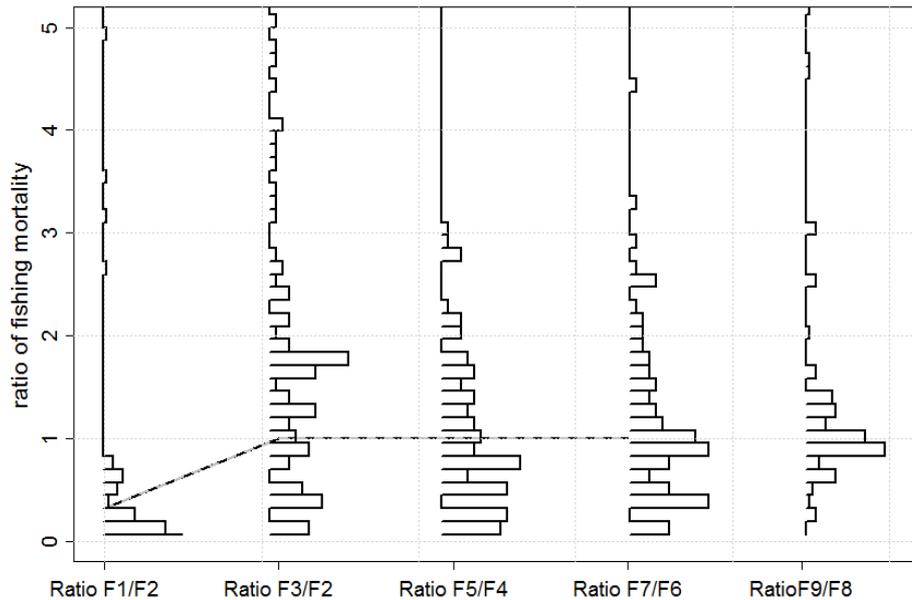
$$angular\ deviation = a \cos \left( \frac{a}{b \times c} \right) \times \left( \frac{180}{\pi} \right) \quad (5)$$

where:

$$a = \sqrt{\sum vec1 \times vec2}$$

$$b = \sqrt{\sum vec1 \times vec1}$$

$$c = \sqrt{\sum vec2 \times vec2}.$$



**Fig. 1.** Ratio of fishing mortality rates estimated in three previous western bluefin tuna assessments. The dashed line is the fixed ratio used for the terminal years, and histograms display the distribution of values around the assumed ratios.

AD values can range between 0 and 180 degrees but the selectivity vectors that we evaluate are all positive, they can vary only between 0 and 90 degrees. Parallel vectors have AD of 0 whereas orthogonal vectors have an AD of 90 degrees. To test the value of the penalty restricting the amount of change in relative  $F$  over time we explored penalty ( $\sigma_v$ ) values ranging from 0.1 (strong) to 0.9 (weak) and then no penalty.

To add error to the CPUE we created new indices  $I_{ky}$  from the simulated CPUE data by drawing at random with replacement from a lognormal distribution as follows:

$$I_{ky} = I_{ky} e^{\pi}. \quad (6)$$

Where  $\pi \sim N(0, \hat{\sigma})$ ,  $k$  denotes a particular series and  $y$  denotes year and  $\hat{\sigma} = 0.61$ , as estimated in the 2008 VPA. To compare the performance of the three estimation scenarios we used the root mean square error (RMSE) between the predicted and the known numbers at age over all ages and years.

### 3 Results

#### 3.1 Analysis of previous VPA models

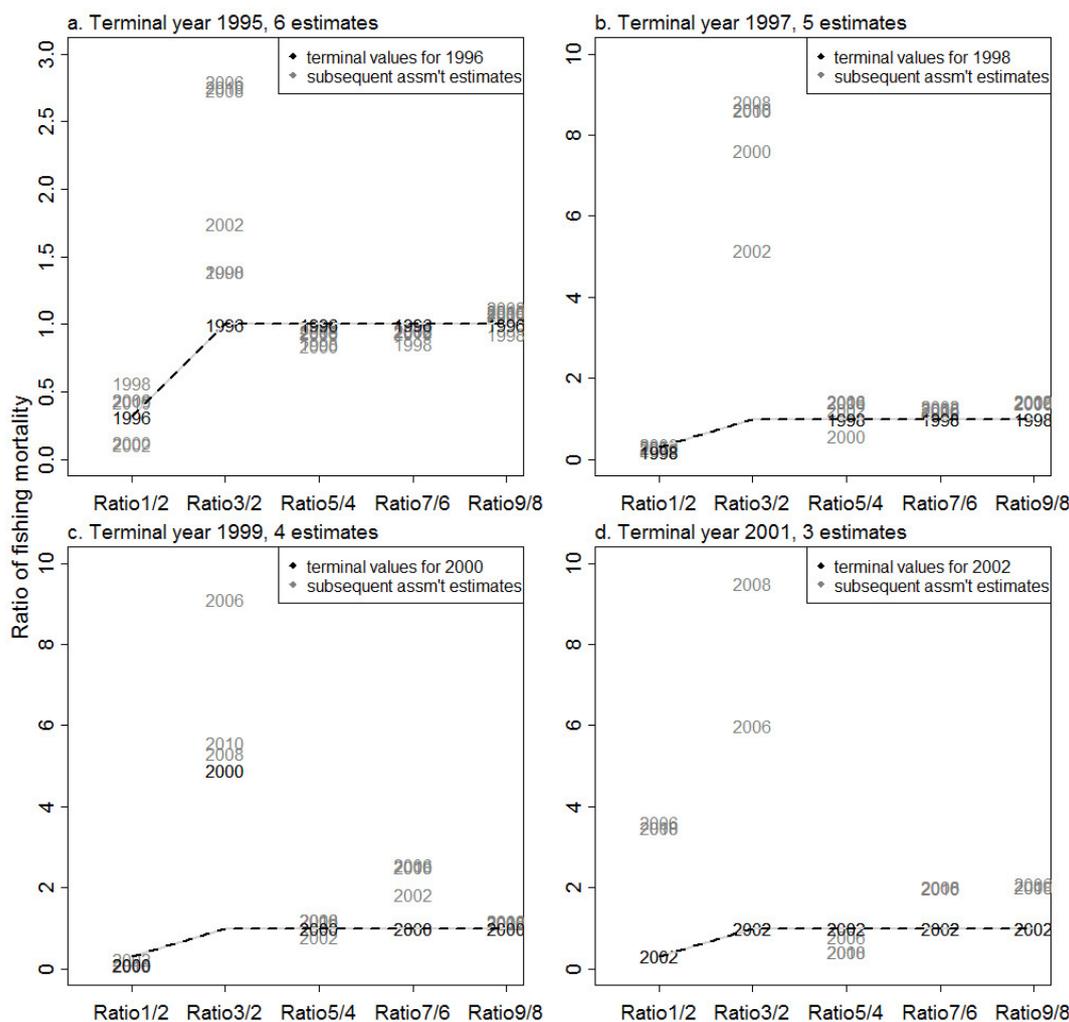
Age-specific fishing mortality estimates were obtained from five previous bluefin tuna VPA assessments (Anon. 2001, 2003, 2007, 2009 and 2011) and compared against the fixed  $F$  ratio linkage assumptions. Note that the publication year lags the assessment by one year and for the 2010 assessment we use only the “continuity case run” which was configured exactly the same as the previous models. The actual base model had 16 ages and so was not directly comparable to previous models. Figure 1 compares the assumed terminal age- $F$ -ratios of equation (1) with those estimated for all years by the 2000, 2002, 2006, and 2008 assessments and the 2010 continuity case run. There was a systematic underestimation of  $F_1/F_2$ ,

i.e.,  $F_1/F_2$  was generally estimated to be less than 0.318. In contrast,  $F_3/F_2$  often appeared to be higher than the assumed value of 1, though the other estimated  $F$ -ratios appeared to be more symmetrically distributed around 1.

Another way to examine misspecification of the terminal age  $F$  ratios is to take advantage of the fact that estimates of  $F$  at age in VPA tend to converge and become increasingly precise, though not necessarily more accurate as one goes back in time (Pope 1972). This implies that the estimates of  $F$  at age for a given year (say 1998) should be more precise in recent versus older assessments (2006). Since the same assumption about the terminal-year  $F$  (Eq. (1)) was made for the last several assessments, an idea of the degree to which that assumption was met can be obtained by evaluating the  $F$ -ratios for four previous assessments with the same year’s estimate in subsequent assessments (Fig. 2). The ratio of  $F_3/F_2$  was subsequently estimated to be much higher than the assumed value of 1 for the 1997, 2000 and 2002 assessments. The misspecification of the ratio when it is really higher than 1 would lead to overestimates of age 2 fish. For most assessment years, the subsequently estimated ratios of  $F_5/4$ ,  $F_7/6$  and  $F_9/8$  were also estimated to be higher than 1, leading to overestimates of age 4 and 6 fish (Fig. 2).

#### 3.2 Retrospective analysis

Analysis of the retrospective patterns indicated that both the fixed-ratio and unconstrained methods tended to overestimate the most recent numbers at ages 4–7 (Fig. 3) when compared with the full data series (solid black line). For ages 8–10 (only 8 shown for brevity) all three methods appeared to retrospectively underestimate numbers at age, a likely byproduct of estimating a constant ratio of fishing mortality between ages 9 and age 10+ for several previous years. For 6 of the 11 population metrics (10 abundances at age and apical fishing



**Fig. 2.** Ratios of fishing mortalities for years a) 1995, b) 1997, c) 1999 and d) 2001 as estimated by subsequent stock assessments. The solid line is the median for subsequent assessments.

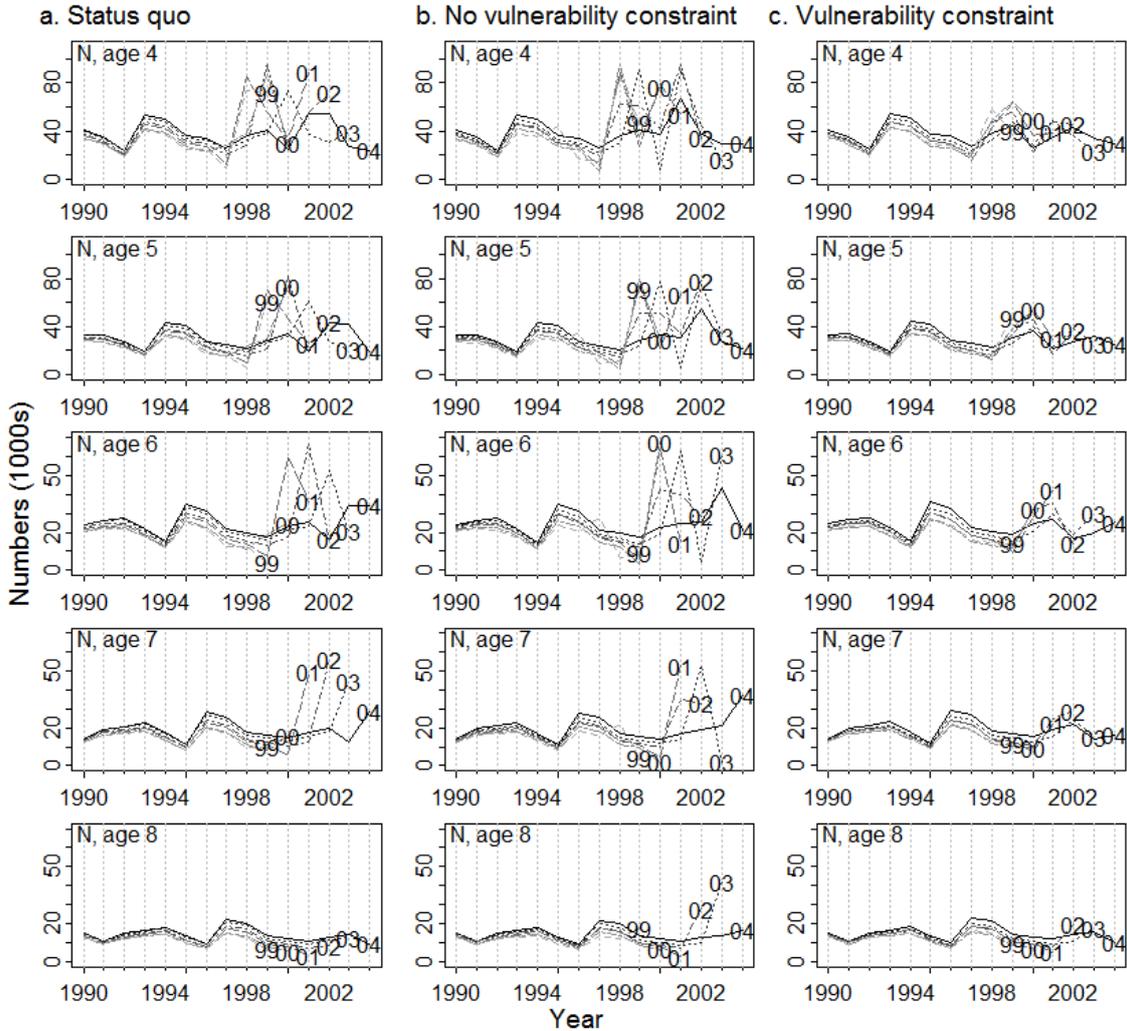
mortality rate) the penalty resulted in the lower  $\rho$  values indicating that it produced less biased retrospective patterns (Table 1). A similar pattern was observed for the absolute value of the Mohn statistic  $|\rho|$ , in which case the penalty improved retrospective error in 7 of the 11 estimates and, particularly in ages 4, 7, and 8, represented a substantial improvement over the other two methods (Table 1). For ages 1 and 2 where the penalty performed the worst of the three methods, assessments of bluefin tuna usually replace estimates of age 1 and 2 abundance in the last and second to last year with predictions from the stock recruitment curve, on the premise that these estimates are highly unstable, regardless of the estimation method. No method completely eliminated the systematic overestimation of ages 5–7 and systematic underestimation of ages 8–10; all three methods still produced rather high values of the Mohn statistic for these ages.

The potential impact on management advice of adopting the relative  $F$  constraint was examined by applying it to the 2006 bluefin tuna model (Anon. 2007). The results indicated higher estimates of spawning stock biomass (SSB) in the terminal year (2004) than for the 2006 assessment using the fixed

$F$  ratio vector, but less optimistic projections under the 2100 t quota. The fixed ratio method estimated larger numbers of 5–8 year old fish in the last few years than did the penalty method, which likely produced the conspicuous bump in SSB and SSB/SSB<sub>MSY</sub> early in the projections not seen with the penalty (Fig. 4).

### 3.3 Stochastic simulations to test the performance and value of the penalty

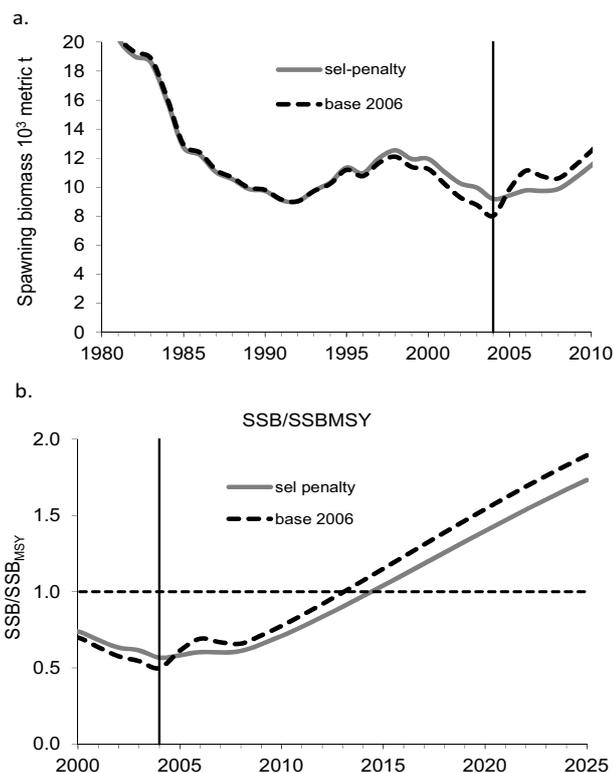
Stochastic simulations with error in the CPUE indices and a single year change in relative  $F$  demonstrate the value of applying the penalty, even across most levels of AD in relative  $F$  (Fig. 5). When there is very little change in relative  $F$  (AD between 0–5) a strong penalty improves estimation substantially (Fig. 5). However as AD increases above 10, the RMSE diminishes with penalties greater than 0.5. Averaging across all levels of relative  $F$ , a penalty of between 0.5–0.6 provides the lowest RMSE (Fig. 5).



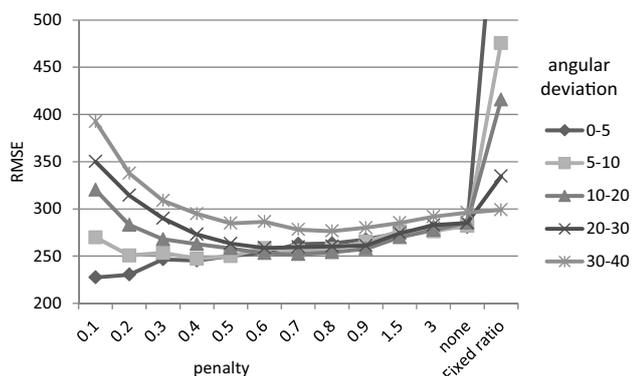
**Fig. 3.** Retrospective analysis of number at age for the three methods of estimating terminal year fishing mortality: a) fixed ratio b) estimate all terminal-year  $F$  with no constraint c) estimate with penalty. The numbers on the end of the lines represent the estimate for the last year estimated by each model run.

**Table 1.** Value of the Mohn statistic as a measure of retrospective bias. The SQ method is the status quo fixed  $F$ -ratio method (Eq. (1)).

Parameter	Mohn statistic			Absolute value Mohn statistic		
	status quo	no constraint	relative $F$ penalty	status quo	no constraint	relative $F$ penalty
$F$	78.2	1406.5	70.4	78.2	1406.5	70.4
Age 1	-73.4	-99	152.5	73.4	99	232.4
Age 2	-64.9	4.4	96.7	68	117.1	153
Age 3	9	15.7	11.6	35.3	75.4	37.8
Age 4	46.5	9.8	12	46.5	38.4	28.5
Age 5	42.1	64.7	33.6	62.1	69.6	41.1
Age 6	-12.9	31.7	19	32.9	67.2	35.8
Age 7	115.5	5.3	-1.9	131	104.5	25.4
Age 8	-44.4	50.1	-17	44.4	100.1	37.6
Age 9	-38.3	-41.2	-37.6	38.3	41.2	37.6
Age 10	-40.3	-43.6	-39.2	40.3	43.6	39.2
shading represents relative performance	1st	2nd			3rd	



**Fig. 4.** Projections of a) spawning stock biomass and b)  $SSB/SSB_{MSY}$  with a constant 2100 mt quota for the 2006 assessment with the fixed  $F$  ratio method and the penalty. Data beyond the solid vertical lines represent projections.



**Fig. 5.** Root mean square error when the terminal year relative fishing mortality varies over wide degrees of angular deviation with penalties ranging from 0.1 (very strong) to 0 (no penalty). Note that for the 0–5 angular deviation, the fixed ratio case goes so far above the y-axis that it is not shown to preserve legibility of the other lines.

## 4 Discussion

The problem of how to estimate terminal year fishing mortalities is a critical issue for VPAs that can have substantial consequences for assessments of current stock status as well as projections of future abundance under proposed management measures. For example, Gavaris et al. (2008) noted that projected SSB trajectories from the various ICCAT assessments

of bluefin tuna all featured a “bump” in the SSB that appeared early in the projection (Fig. 4) and may have led to overly optimistic appraisals of future abundance. Similar patterns of a bump in SSB in the early years of the projections can be seen in previous assessments (Anon. 1999, Fig. 25, 26; Anon. 2001, Fig. 13, 14; Anon. 2003, Fig. 40–42; Anon. 2007, Fig. 39). This bump may have led to overly optimistic appraisals of future abundance.

Our analyses indicated that this spike in projected SSB may have been largely an artifact of incorrectly specifying the ratios of the fishing mortality rates for linked ages. In particular, the ratios of the  $F$  on ages 5, 7, and 9 relative to the next younger age groups tended to be specified too low, thereby leading to overestimates of the abundance the younger age groups. This overestimation of ages 4–7 can be seen in the retrospective plots (Fig. 3). As a reviewer pointed out, the effect of incorrectly specifying these ratios might be mitigated to the extent that overestimation of  $F$  on one age group is offset by under estimation of  $F$  on an adjacent age group. However, the extent of this offset is difficult to predict and will depend on the information content of the indices of abundance relative to the affected age groups. At least for bluefin tuna, the proposed method of restricting the change in relative  $F$  leads to reduced retrospective biases and may provide for more realistic projections.

The stochastic simulations provide a broader context for the utility of the penalty as well as guidance for selecting the appropriate penalty strength as a function of the degree of inter-annual change in the relative fishing mortality rate vector. Moderate penalty ( $\sigma$  between 0.4 and 0.8) substantially improved estimation performance (Fig. 5). Stronger penalties ( $\sigma < 0.4$ ) tended to bias estimation except when the change in the AD of relative  $F$  was very low (between 0 and 5 degrees, see Fig. 5). Overall for most levels of AD that we have explored, the potential loss in performance obtained by using a moderate penalty is not as severe as when either no penalty or a strong penalty is used. There appears to be little to lose by use of moderate penalties regardless of how relative  $F$  actually changes from year to year unless it really has not changed. In this case, stronger penalties on the order of 0.2 may be justified if it is known a priori that the fishery has changed very little in the most recent few years.

The choice of penalty value for any given application may be aided by calculating the angular deviation (AD) from previous analyses and comparing those with results in Figure 5. For example, the ADs calculated from the relative fishing mortality at age matrix estimated by the 2006 bluefin tuna assessment ranged from 0.53 to 34 with a mean of 17.4, suggesting that optimal performance for bluefin tuna would be obtained with a penalty of about 0.6. For many other species similar annual changes might be observed indicating values between 0.5–0.6 are likely to be appropriate compromises between free estimation and imposing too great of a penalty on change.

In summary, all methods that constrain the degree of change in the relative  $F$  or vulnerability schedule over time represent a tradeoff between model fit versus reliability of the estimates (Lassen and Medley 2001; Porch 2003). If the penalty on the amount of change is too great, the model will not allow sufficient variation in relative fishing mortality rates

over time, even as catch and effort varies, potentially resulting in poor model solutions. This can be problematic when fishing mortality at age changes rapidly such as with the imposition of a size limit or a rapid change in targeting strategy. In contrast, if no penalty is applied, the results may be “unbiased” in a statistical sense, but so imprecise as to be rendered meaningless in practical application. In such cases actual changes in recruitment and year class strength may be masked and attributed simply to changing relative fishing mortality.

In the case of the 2008 bluefin tuna assessment, the values used for the fixed ratios were legacies of a separable VPA conducted during the 1994 assessment (Anon. 1997). It is clear that the relative  $F$  pattern has changed since then owing to changes in fishery targeting and management measures designed to shift fishing mortality to older ages, rendering the legacy specifications obsolete. One could argue for recalculating the ratios based on a separable VPA using the more recent data, but this still presupposes that the selectivity in the most recent year deviated little from the average pattern exhibited during the period to which the separable VPA was applied. Given the above results, the more flexible modeling approach afforded by use of the penalty on changes in relative  $F$  appears to be the better option. Fortunately, stock status with the penalty was similar to the fixed-ratio during the 2008 assessment, suggesting that the management advice would not have been substantially different (Fig. 4b). Both indicate the spawning biomass in 2004 was well below the level that would produce the maximum sustainable yield. However, this may not have been true for earlier assessments, such as in 2002, where unusually strong year classes were erroneously estimated for 1997 and 1999. Given the above results, the most recent ICCAT bluefin tuna assessment (Anon. 2011) adopted the method of estimating all terminal year relative fishing mortalities using the constraint on interannual changes in relative  $F$  rather than the fixed  $F$  ratio schedule used in previous assessments.

*Acknowledgements.* The authors wish to thank Shannon Calay, Jerry Scott and Alex Chester who reviewed the manuscript and provided constructive comments and Michael Prager who suggested the use of angular deviation and Chris Legault who recommended the Mohn statistic.

## References

- Anonymous, 1997, Report of the 1996 SCRS Bluefin Tuna Stock Assessment Session. Col. Vol. Sci. Pap. ICCAT 46(1), 1–186.
- Anonymous, 1999, Report of the 1998 ICCAT SCRS Bluefin Tuna Stock Assessment Session. Col. Vol. Sci. Pap. ICCAT 49(2), 1–191.
- Anonymous, 2001, Report of the 2000 ICCAT SCRS West Atlantic Bluefin Tuna Stock Assessment Session. Col. Vol. Sci. Pap. ICCAT 52 (3), 831–958.
- Anonymous, 2003, Report of the 2002 Atlantic Bluefin Tuna Stock Assessment Session. Col. Vol. Sci. Pap. ICCAT 55(3), 710–937.
- Anonymous, 2007, Report of the 2006 Atlantic Bluefin Tuna Stock Assessment Session. Col. Vol. Sci. Pap. ICCAT 60(3), 652–880.
- Anonymous, 2009, Report of the 2008 Atlantic Bluefin Tuna Stock Assessment Session. Col. Vol. Sci. Pap. ICCAT 64(1), 1–352.
- Anonymous, 2011, Report of the 2010 Atlantic Bluefin Tuna Stock Assessment Session. Col. Vol. Sci. Pap. ICCAT 66(2), 505–714.
- Butterworth D.S., Rademeyer R.A. 2008, Statistical catch-at-age analysis vs. ADAPT-VPA: the case of Gulf of Maine cod. ICES J. Mar. Sci. 65, 1717–1732.
- Chen Y., Jiao Y., Sun C.L., Chen X., 2008, Calibrating virtual population analysis for fisheries stock assessment. Aquat. Living Resour. 21, 89–97.
- Doubleday W.G., 1981, A method for estimating the abundance of survivors of an exploited fish population using commercial catch-at-age and research vessel abundance indices. Can. Spec. Publ. Fish. Aquat. Sci. 58, 164–178.
- Fromentin J.M., Powers J., 2005, Atlantic bluefin tuna: population dynamics, ecology, fisheries and management. Fish Fish. 6, 281–306.
- Haddon M., 2001, Modelling and quantitative methods in fisheries. New York, NY. Chapman & Hall.
- Gavaris S., Hazin F., Neilson J.N., Pallares P., Porch C., Restrepo V.R., Scott G., Shelton P., Wang Y. (eds.), 2008, Proc. The joint Canada-ICCAT 2008 workshop on the precautionary approach for western bluefin tuna, Halifax, Nova Scotia, 17–20 March, 2008.
- Gavaris S., 1988, An adaptive framework for the estimation of population size. CAFSAC. Res. Doc. 88/29.
- Gulland J.A., 1983, Fish stock assessment: a manual of basic methods. New York, NY Wiley.
- Lassen H, Medley P., 2001, Virtual Population Analysis – A Practical Manual for Stock Assessment. FAO Fish. Tech. Pap. 400.
- Megrey B.A., 1989, Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Am. Fish. Soc. Symp. 6, 8–48.
- Mohn R., 1999, The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56, 473–488.
- Parks W.W., 1976, Cohort analysis, equilibrium yield per recruit analysis and predicted effects of minimum size limit regulation in the Atlantic bluefin tuna fisheries system. ICCAT Coll. Vol. Sci. Pap. 5, 313–331.
- Parrack M.L., 1986, A method of analyzing catches and abundance indices from a fishery. ICCAT Coll. Vol. Sci. Pap. 24, 209–221.
- Patterson K.R., Kirkwood G.P., 1995, Comparative performance of ADAPT and Lauec-Shepherd methods for estimating fish population parameters and in stock management. ICES J. Mar. Sci. 52, 183–196.
- Pope J.G., Shepherd J.G., 1982, A simple method for the interpretation of catch at age data. ICES J. Mar. Sci. 40, 176–184.
- Pope J.G., Shepherd J.G., 1985, A comparison of the performance of various methods for tuning VPAs using effort data. ICES J. Mar. Sci. 42, 129–151.
- Pope J.G., 1972, An investigation of the accuracy of virtual population analysis using cohort analysis. ICNAF Res. Bull. 9, 65–74.
- Porch C.E., 2003, VPA-2BOX Version 3.01 Users Guide. NOAA Fisheries Southeast Fisheries Science Center Sustainable Fisheries Division Contribution SFD/2003-0004.
- Shepherd J.G., 1999, Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. ICES J. Mar. Sci. 56, 584–591.
- Smith S.J., Gavaris S., 1993, Evaluation of the accuracy of projected catch estimates from sequential population analysis and trawl survey abundance estimates. In: Smith S.J., Hunt J.J., Rivard D. (Eds.) Risk evaluation and biological reference points for fisheries management, Can. Spec. Publ. Fish. Aquat. Sci. 120, 163–172.