SCIENTIFIC EXAMINATION OF WESTERN ATLANTIC BLUEFIN TUNA STOCK-RECRUIT RELATIONSHIPS

Andrew Rosenberg,[1] Andrew Cooper, Mark Maunder, Murdoch McAllister,
Richard Methot, Shana Miller, Clay Porch, Joseph Powers, Terrance Quinn, Victor Restrepo,
Gerald Scott, Juan Carlos Seijo, Gunnar Stefansson, John Walter

SUMMARY

A workshop was convened in Washington, DC on 19-21 June 2012 to review the available information on the relationship between spawning stock and recruitment for western Atlantic bluefin tuna (Thunnus thynnus), a critical part of the scientific advice for management of the stock. The workshop participants noted that considerable attention has been paid to the form of the relationship between stock and recruitment, but that insufficient attention had been paid to the consequences of different hypotheses about that relationship. We suggest a simple decision table approach, using well-developed methods, for informing on the consequences of different policy choices given the uncertainty in the productive capacity and future trajectory of the stock. This will enable the scientific advice to move beyond a debate about the form of the stock and recruitment relationship and provide a more informative set of information for management.

Based on the decision table analysis, considering alternative hypotheses about stock and recruitment using the available data, the results indicate that for rebuilding of spawning stock and yield in the long-term, fishing mortality rates should be kept low. Higher values are only advantageous if the primary goal is short-term yield at the expense of long-term rebuilding of the biomass. Higher values pertain to tradeoffs with yield in short and long term. An intermediate goal would be to reduce F enough to rebuild SSB to near 30,000 mt. At this level of SSB, information on the potential for the stock to increase recruitment and rebuild to a higher level should be apparent. Finally, the workshop participants identified a set of issues in the stock assessment that need further investigation and development to continuously improve the scientific advice provided through the assessment process.

KEYWORDS

Bluefin tuna, Recruitment, Stock assessment

1. Workshop goals

This workshop was convened to review the available data and information on the relationship between spawning stock and recruitment for western Atlantic bluefin tuna and, if possible, come to consensus among the participants concerning this relationship. The members of the working group were experts in fishery population dynamics and the modeling of fishery data, some with extensive experience with this stock in particular, and some who have not previously considered the stock. The goal was to take a fresh look at the available information and focus solely on the relationship between stock and recruitment because of its fundamental role in providing scientific advice for policy-makers on the potential productivity of the stock.

In recent years, two forms of the stock recruitment relationship have been considered by the International Commission for the Conservation of Atlantic Tuna's (ICCAT) Standing Committee on Research and Statistics (SCRS): a conventional Beverton-Holt relationship and a two-line model. These two suggested forms for the relationship have been considered to be equally likely in advising managers, in other words, that there is no information that would suggest that one is more representative of the underlying relationship than the other. The working group critically examined this hypothesis.

[1] Union of Concerned Scientists, 2 Brattle Square, Cambridge, MA 02138, USA, arosenberg@ucsusa.org

Fundamentally, the real question is whether recruitment is expected to increase if spawning stock abundance is allowed to increase through fishery management action to rebuild the stock. This question pertains even if the precise form of the stock and recruitment relationship remains uncertain. However, of the two models considered, the two line model essentially presumes that mean recruitment will not increase above recent levels. The Beverton-Holt relationship would indicate that recruitment will increase if higher stock sizes are obtained.

2. Review of western Atlantic bluefin tuna stock-recruitment data and modeling at ICCAT

2.1 Western Atlantic bluefin tuna stock-recruitment data

The available data indicate that restoration of a higher level of spawning biomass could produce a higher level of recruitment that can both sustain that level of biomass and support a higher level of catch than is currently attainable. In general, data used in contemporary assessments do not include the pre-1970 period of available catch, effort, and catch-at-size information for estimating stock abundance. This is because, although the catches were large and catch per effort was high in that period, the information on catch-at-size is relatively weak for pre-1970. In addition, the methodology used to monitor stock response to the current rebuilding plan requires additional assumptions that are difficult to justify in order to incorporate those data, adding uncertainty to the overall evaluation. Alternative forms of stock abundance estimation, which better accommodate the sampling characteristics from the early part of the fishery, tend to support the view that both spawning stock and recruitment levels during the 1950s and 1960s were similar or greater than the early 1970s.

The current assessment of western bluefin tuna indicates that recruitments of approximately 300,000 age 1 fish were common prior to 1975, and these recruitments were produced by an average spawning biomass near 50,000 mt in the period 1970-1975 (Figure 1). When assessment of western bluefin tuna is extended back to 1960 (Porch *et al.* 2001), the results demonstrate that similarly high recruitment and spawning biomass levels existed prior to 1970 (Figure 2). In contrast, for the years 1978-2005, recruitment and spawning biomass have averaged 91,000 and 16,400 mt, respectively. The transition from high biomass and recruitment to low biomass and recruitment occurred rapidly after the mid-1970s. It appears that during this period low recruitment began to occur and juveniles were fished heavily; this moved the spawning biomass down to a level that has not produced recruitments as high as those that were common prior to 1975.

Estimates of annual spawning stock biomass and recruitment produced by the 2010 assessment (Figure 1) constituted the key information evaluated by the group. As spawning biomass is estimated to have been rapidly depleted during the 1970s, there are relatively few data points with which to establish a relationship between spawning biomass and subsequent recruitment of one year-old juveniles. The western bluefin stock assessment does not have any true indices of recruitment and only ‘sees’ recruits through the catch-at-size converted to catch-at-age once they have entered the fisheries and through fishery-dependent indices. In the early 1970s, the fishery selectivity focused on smaller fish and hence the lag between recruitment and capture by the fishery was relatively short, particularly when the purse seine fisheries were capturing smaller fish. After these early years, selectivity appears to have shifted to larger fish so the lag between recruitment and appearance into the fishery and detection by the VPA is even greater. This is evidenced by retrospective analysis of the 2010 VPA, which only really detected the apparently large 2003 year class within the past 2 years of the terminal year of the VPA (2009). The absence of strong indicators of recruitment coupled with the slicing method of aging the catch at size further blurs the detection of recruitment, making these data imperfect.

Two other sources of spawning stock and recruit data were also presented to the group. Stock and recruitment estimates were output from the MAST statistical catch-at-age model (Figure 3, Taylor et al. 2011). These estimates represent fits from an assumed Beverton-Holt stock-recruit function and incorporate potential stock mixing between East and West.

An additional dataset was available for years 1960 to 1997, using a statistical catch-at-age model that did not require complete size data for all years (Figure 2, Porch et al. 2001). The model also did not impose an underlying stock-recruitment function. The results suggested that the high levels of recruitment estimated for the 1970s were not anomalous, but typical of the era from 1960-1975 when the spawning stock biomass was also much higher. The contrast gained by adding data from the 1960s (Figure 2) lends further support to the existence of a relationship between spawning biomass and subsequent recruitment. It is important to note that the catch-at-age data used by Porch et al. (2001) were derived using a growth curve that has since been revised substantially (Restrepo et al. 2010). Use of the new growth curve would likely change the magnitude of the spawning biomass and recruitment estimates but probably not their relative trends through time. Two other facets of the Porch et al. (2001) analysis that should be examined further are the Japanese longline CPUE series for the Gulf of Mexico (which shows perhaps unrealistically large fluctuations during the early 1960s) and the assumption that the bluefin tuna caught by the Japanese longline fishery off Brazil during the 1960s were of western origin.

Other important sources of uncertainty in the western bluefin data include the relative degree of maturation of females with age, growth dynamics for the stock over time, and the degree to which perception of western spawning population abundance and subsequent recruitment estimates are influenced by the amount and direction of fish movement between management zones and, accordingly, the relatively greater uncertainty associated with stock abundance estimates for eastern Atlantic and Mediterranean bluefin.

2.2 Modeling of the stock-recruitment relationship by ICCAT

Assessment modeling of the stock-recruitment relationship (SRR) has occurred over the period from the early 1980s through 2010. Before 1984, individual scientists presented competing assessments which often showed different relationships depending on the model and data used. Since then, Working Groups began assessing the stock and making short-term projections using "recent" recruitment levels to forecast if the stock would increase or decrease under the extant total allowable catch (TAC). In 1993, maximum sustainable yield (MSY)-related benchmarks were computed for the first time from the VPA results combined with an SRR. This approach was not used again until 1998, when the western rebuilding plan was adopted. During the intervening years and several times into the future, SSB_{1975} was often used as a proxy for SSB_{MSY}, assuming constant recruitment.

The assessments made between 1998 and 2004 used the two-line (Figure 1, red line) and/or Beverton-Holt SRRs (Figure 1, blue line). The terms "low" and "high" recruitment scenario were introduced in 2000 for the two-line and Beverton-Holt, respectively, and mention of a possible ecosystem explanation for a regime shift was made in 2002 (ICCAT 2004). The two-line model assumes that recruitment increases linearly with SSB from zero with no spawners to a maximum value (R_{MAX}) when SSB reaches a certain threshold. The SSB threshold or hinge point was set to an average over a period of years where SSB was low. Most recently at the 2010 assessment, this hinge point was calculated as the average SSB during 1990-1995, the period with the lowest estimated SSB, and R_{MAX} was calculated as the geometric mean recruitment during 1976-2006, a period over which recruitment was relatively constant. VPA recruitment estimates for the most recent three years have generally been deemed unreliable and were not used in the modeling.

The 2006 assessment made only short-term projections and estimated MSY benchmarks based on recent recruitment. The two-line and Beverton-Holt SRRs were used again in the 2008 and 2010 assessments.

We summarize below the various supporting evidence that has been provided for the two-line and Beverton-Holt SRRs.

The Beverton-Holt function is a commonly applied stock-recruit function in fisheries stock assessment modeling of marine fish stocks (Hilborn and Walters 1992). It is typically applied as the default stock-recruit function when there is no strong evidence to suggest an alternative form of a stock-recruit function. This model predicts that density-dependent processes increase the larval/juvenile survival rates as spawning stock abundance becomes low. Conversely, at high population sizes, density-dependent processes reduce survival, resulting in an asymptotic level of maximum average recruitment.

The Beverton-Holt model has been applied under the presumption of long-term stationarity in the environmental conditions affecting recruitment. It presumes that under long-term unfished and lightly fished conditions, recruitment could be expected, on average, to be larger than it would be when the stock is depleted. The formulation of the model that has been applied, allows for recruitment deviates to follow a lag 1 autoregressive pattern in which there is a positive correlation in deviates from the deterministically predicted annual recruitment values. This allows the model to track environmental or multi-species ecological conditions that impact recruitment over series of years, and thus accommodates what might be termed to be mild shifts in such conditions that impact recruitment.

The two-line recruitment model fitted to data after 1980 was advanced as a parsimonious statistical model to describe the stock-recruit data and predict future recruitment in the evaluations of the potential consequences of alternative catch quotas. The main rationale provided for this model was that the best estimate of future recruitment was the average of recent values for recruitment, that it was unlikely that the historic high recruitment values and high spawning stock size values will reoccur and that if SSB drops below the lower end of observed values, that recruitment could be expected to decline proportionally with SSB.

Skepticism about whether recruitment could be rebuilt to the levels seen prior to the mid-1970s, and therefore support for the two-line SRR, has arisen because:

1. It has been noted that the VPA constructed recruitment fell abruptly and somewhat more rapidly than the apparent large declines in SSB in the middle to late 1970s. A large drop in recruitment preceding a large drop in SSB could potentially be an indication that extrinsic conditions impacting recruitment have deteriorated and could potentially be part of the cause of the decline in SSB.

2. The apparent sharp decline in recruitment in the 1970s also causes the VPA constructed recruitment to consistently fall short of the long-term average recruitment predicted by the "best-fit" Beverton-Holt model for the spawning stock sizes in that period (Figure 1). When fitted to the data, the Beverton-Holt function shows a series of positive recruitment deviates 1970-1974. The deviates turn quite strongly negative from 1975 to 1982. Following the early 1980s, the VPA predicted recruitment has remained consistently low. This over-prediction of the Beverton-Holt model for series of years since 1974 and the consistently low recruitment since 1975 has been conjectured by some to reflect a so-called "regime shift" such that average conditions that impact the survival of juveniles have deteriorated and persisted and will persist in the future.

3. Some have suggested that components of the western (and eastern) population(s) appear to have been lost for good after large harvests of large Atlantic bluefin tuna in the 1950s and 1960s and the lack of recurrence of aggregations of large fish in these areas since then. Some very large catches of mature fish were taken off of Brazil in the 1960s with total annual catches from this area amounting to well over 10,000 t, and large catches of large fish were also taken off of Norway in the 1950s and 1960s (Tangen 2009).

In turn, other concerns have been raised over the two-line, regime shift recruitment model.

1. The ICCAT stock-recruit data show a one way trip (i.e., both recruits and spawning stock size have shown large systematic decreases since the early 1970s). All of the low recruitment values have occurred at the lowest stock sizes. However, the very large cohort in 2003, which was about two thirds the average recruitment predicted by the 2010 VPA for the early 1970s, suggests that potentially very large cohorts can be produced with relatively small stock sizes and that, should stock size be rebuilt, similarly large cohorts could still be possible.

2. The hypothesis that key components of the western spawning population have been lost for good and that therefore recruitment will not return to the higher levels that occurred when these components were present stands in contrast to the competing hypotheses of reversible range contraction and long-term fluctuations in productivity (Petitgas 1998; Ravier and Fromentin 2001; Huse et al. 2002; Melvin and Stephenson 2007; Frisk et al. 2009; Fromentin and Kell 2009; Tamdrari et al. 2011).

3. The regime shift hypothesis appears to be a post hoc proposal to explain the data without independent evidence of the regime shifting.

For many fish populations, there does not appear to be a stock recruit model that fits the data well. If the purpose is to calculate an acceptable biological catch level and to make short-term projections, then a spawner-recruit model is not immediately needed. Instead one could use a proxy for the fishing mortality rate that produces MSY. For instance, a proxy fishing mortality rate (such as $F_{40\%}$, the F that reduces spawning biomass per recruit to 40% of that at F=0) could be used. One would then calculate the corresponding catch and project the population forward in time using the average or median of past recruitments thought likely to occur in the near future. This approach still entails making an assumption regarding the nature of future recruitment, does not provide targets for long-term rebuilding plans and provides little protection against continued decline for severely depleted stocks. Such an approach is used for crab and groundfish fisheries managed by the North Pacific Fishery Management Council, among others and has been explored for ICCAT species (ICCAT 2010b).

3. Statistical considerations for the examination of recruitment relationships

3.1 A traditional statistical approach

Traditional statistical methods can be used with essential biological assumptions to select appropriate models to describe the stock and recruitment relationship. This exercise was designed to provide a fresh look at the stock and recruitment data and was performed by one of the workshop participants who does not have a history with Atlantic bluefin tuna stock assessment.

Any biological model to be used for projections that entertains some potential for extinction needs to go through the origin. A minimal model is therefore R=aS (i.e., a straight line through the origin). This model is a submodel of all models considered below, and they can therefore be formally evaluated by comparing with this as the null model. The next simplest model includes density dependence. This is needed to avoid having a stock which can only explode or go extinct, but is not essential for management advice. Density dependence can be implemented through a two-line relationship and a Beverton-Holt (BH) relationship, among many others.

Let R denote recruitment and S the spawning stock biomass generating those recruits. Both are measured with error, but the errors in S will be ignored in the equations. The models will predict R conditionally on the observed S and assume errors in Ln(R) to be independent and identically distributed measurements from a normal distribution. Statistically the parameters are simply regression parameters, but the interpretation depends on the specific context. Note that there are several fundamental differences in going from a null R=aS to e.g. R=aS/(1+S/K) vs going from a null R=b to R=aS/(1+S/K):

- The former comparison starts with models which are essential for management, the latter does not.
- For an exploited stock the former implies a search from a null of proportionality to find where density dependence occurs whereas the second implies a search from a null of constancy to find where a collapse occurs.
- The former encapsulates a possible collapse from the outset whereas the latter implies infinite resistance to fishing.
- Finally, the latter is not nested so statistical comparisons are not trivial.

From a statistical viewpoint there is no real difference between starting with one null model or the other since they each have one parameter to estimate.

This approach does not solve all problems. For example the BH curve may fit better but behave unpredictably outside the range of the data, as may a Ricker curve.

Models are fitted using (nonlinear) least squares based on log-transformed recruitment. An alternative approach is to inversely transform recruitment and SSB, which transforms the BH relationship to a straight line, but this is only used to verify results.

The two-line model is fitted by estimating a value, B* of biomass, such that the S-R relationship is linear to the origin on the left but flat to the right of B*.

The regime shift assumption, as considered in SCRS analyses to date, can be tested by explicitly including a regime shift into the model. In the two-line model this is done by estimating two different values of B*, one for each time period (before and after 1980). For the BH model two values of K are estimated.

This implies a total of 5 models: The null, R=aS plus the BH and two-line, each of the latter with and without a regime shift.

3.2 Statistical analysis results

Stock and recruitment data were obtained from the 2010 ICCAT western Atlantic bluefin tuna stock assessment. Fitting different BH curves for the two time periods (i.e., before and after 1980) is not feasible as the optimal solutions send K either to infinity or to negative values, depending on which time period or estimation method is used. Results are therefore only considered for the remaining 4 combinations of models. Figure 4 gives all the fitted curves. Note that several of these completely overlap.

The resulting summary statistics are seen in the following text table, giving values for the linear model, the Beverton-Holt model, the two-line model and the two-line model with a regime shift:

Model	**# params**	**Residual_df**	**Residual_SS**	**F**	**P**
Straight line	1	38	7.6612		
Beverton-Holt	2	37	7.4458	1.07	0.31
2-line	2	37	7.6612	0	1.00
2-line (regime shift)	3	36	7.6486	0.06	0.94

The best fitting model of the lot is the simple BH model fitted to the entire time series. However, this is not significantly better than the simple straight line (P=0.31), and the reason is obvious from Figure 4, where it is seen that the BH is very close to a straight line for most of the data points. Since they are so close, the choice to include density dependence is merely one of convenience in modeling projections. It is unlikely to have any effects at all on the outcomes of short-term projections but results in different equilibrium interpretations. Note a straight line form of the SRR is unrealistic as it implies the population can increase without limit, but over the range of the available data may be sensible.

Given that the BH is not significantly better than the linear model, none of the other models will be either, since they include the same number of parameters or more and they give a (slightly) worse fit when compared to the BH. It follows that a formal (statistical) modeling approach (basically using Occam‘s razor) would simply select the linear model since none of the others are significantly better, i.e., further complications are not needed. However, it must be stressed that the entire above exercise uses simply the VPA-estimated spawning stock and recruitment and so any estimated stock-recruitment relationships are dependent upon the adequacy of the VPA and the range of the available data. Some participants were concerned that the inability of either of the stock-recruitment models to fit the data well may be due to inadequacies in the VPA model or data inputs and that more sophisticated statistical catch at age models may be better able to capture the stock-recruitment relationship.

The two-line and the BH cannot be compared using traditional statistical methods. However, it is seen from Figure 4 and the text table that the two-line alone provides no additional information to the straight line. The reason for this is that the best estimate of B* lands outside the range of the data. Thus the two-line model reduces automatically to the straight line.

Finally, consider the regime shift. This can, in principle, be added to the two-line or the BH models but results in non-convergence or implausible parameter estimates for the BH and is therefore rejected as an addition to the BH. The improvement to the two-line model is very marginal (see summary table above). It follows that there is little evidence in the data supporting the assumption of a regime shift. It should be noted that these methods assume that each estimated recruitment is independent and identically distributed and that there is no error in the spawning biomass estimate. When considering the fitted curves, the common slope of the double two-line curves (Figure 4) coincides almost exactly with the null slope, and the B* values for each time period are basically placed outside the range of the data for each time period.

A corresponding analysis based on stock-recruitment data for year classes 1961-1997 (based on Porch et al. 2001) is given in Figure 5. As in the earlier analysis, there is precious little difference between the different curves, which almost exactly coincide within the ranges of data to which they are fitted. The sole exception is the 3-parameter two-line model, which now has 4 data points to the right of the first breakpoint. The corresponding summary statistics are:

Model	**# params**	**Residual_df**	**Residual_SS**	**F**	**P**
Straight line	1	36	11.136		
Beverton-Holt	2	35	11.127	0.027	0.87
2-line	2	35	11.136	0	1.00
2-line (regime shift)	3	34	10.336	1.316	0.28

In particular, an F-test comparing the 3-parameter two-line model to a 1-parameter straight line gives a P-value of 0.27.

3.3 Conclusions based on the frequentist statistical analysis

The use of any particular stock-recruitment relationship is a model extrapolation with little foundation in the stock and recruitment estimates from the VPA. A straight line through the origin would suffice to explain the data and provide short-term projections, but, of the models examined, a BH curve would be the statistically best approach to incorporate density dependence. Because nearly all the recent stock and recruitment estimates are clustered around low spawning stock and low recruitment (except for 2003), there is little contrast in these recent data from which a statistical analysis might determine which model is more likely to better predict the level of recruitment that would come from larger stock levels.

In principle, one could try several variations on the above analysis (e.g., estimate a common value of B* but different slopes for the two-line models, or implement/estimate the purported regime shift as a pivot point.) These modifications are unlikely to improve the fit of the data. The above analysis includes the entire time series as independent observations, but one might also try to drop some of the most recent years and/or include an autoregressive lag(1) model for the residuals.

3.4 Using Bayesian methods to better understand the stock-recruitment relationship of western Atlantic bluefin tuna

An updated analysis of McAllister *et al.* (2001), which used a Bayesian computational approach to obtain empirically based weightings for conflicting stock assessment results, was presented. This approach offers a methodologically rigorous basis for computing credibilities or weights (i.e., Bayesian posterior probabilities) for different hypotheses. Importance sampling was presented as a methodology to compute such probabilities. The approach was illustrated using stock-recruit data for western Atlantic bluefin tuna. Alternative hypotheses were considered on (a) regime shifts affecting future recruitment and (b) the mathematical form of the stock-recruit function. The updated analysis uses stock-recruit data obtained from ICCAT's 2010 stock assessment of Atlantic bluefin tuna. Bayes factors were presented for the alternative hypotheses, which reflect the ratio of Bayes' probability of the data for a particular hypothesis to that for some reference hypothesis. Improved formulations of prior probabilities for estimated parameters were identified that removed the influence of the priors on the values obtained for Bayes factors.

When applied to stock-recruit data from 1970 to 2005 and presuming a regime shift in 1977, the Bayes factors favored neither hypothesis. When the 2006 recruitment data point was added, according to the methodology of the 2010 assessment, Bayes factors favored the Beverton-Holt model by factor of 4.8 (*i.e.*, the probability of the data was 4.8 times higher for the Beverton-Holt model compared to the two-line regime shift model). The Bayes factors increased to 4.9 and 5.4, respectively, when the 2007 and 2008 recruitment data points were added as sensitivity runs. When Bayes factors were computed for the 1960-1998 stock-recruit data set provided by Porch *et al.* (2001), Bayes factors continued to support the Beverton-Holt model by a factor of 2.7. The sensitivity of Bayes factors to numerous different statistical assumptions was evaluated, including settings for autocorrelation in stock-recruit deviates, different priors, differences in variance in recruitment deviates between regimes and the stock-recruit data to analyze. The detailed results of the sensitivity analyses are given in McAllister (2012).

4. Decision theory

4.1 Background on decision theory approaches and criteria

Within the western bluefin tuna fishery context described in previous sections, an application of decision theory for systematic choice under uncertainty considering different risk attitude criteria, with and without mathematical probabilities, is suggested. Under this approach, decision-makers in the bluefin tuna fishery are expected to select one management strategy, *d*, out of a set of *D* alternative strategies. When selecting a strategy, the fishery manager should be aware of the corresponding consequences. These consequences are likely to be a function of the cause-effect relationships specified in the fishery model, the estimated bio-ecologic parameters, and the possible states of nature.

In decision analysis, it is important to be able to estimate a loss of opportunities function, $L(d,\theta)$, also called a regret matrix (Parmigiani and Inoue 2009), which reflects the resulting losses of having selected strategy d when the state of nature occurring is θ.

If prior or posterior probabilities are available to build decision tables for fishery managers, the expected values (EV) and their corresponding variance (VAR) can be estimated for the selected fishery performance variable, FPV (*e.g.*, spawning stock biomass, yield), but this can be done only if the distributions for the variables of interest overlap under the different hypotheses. There are, however, different degrees of risk aversion and instances where probability values are either unavailable or controversial. Should there be skepticism about subjectively or empirically derived probabilities for different hypotheses, an associated uncertainty question arises: How should uncertainties inherent to data collection and recruitment modeling of western bluefin tuna be addressed without the use of probabilities of occurrence of possible states of nature? Decision theory provides alternative criteria for such instances, including those with increasing degrees of caution in decision-making (Shotton and Francis 1997, Seijo *et al.* 1998). Applying these concepts to the western Atlantic bluefin fishery, criteria with differing degrees of caution and criteria with and without mathematical probabilities were also presented and discussed during the workshop.

4.1.1 Bayesian criterion

The Bayesian criterion is a procedure that uses prior or posterior probabilities to aid the selection of a management strategy (i.e., bluefin tuna TAC). It requires that the fishery manager select the decision that minimizes the expected loss of opportunities. Decisions without experimentation and data analysis use prior distributions estimated out of experiences that are translated subjectively into numerical probabilities. Fishery decisions that are based on experimentation and data analysis can use posterior probabilities. Posterior probabilities are the conditional probability of state of nature θ, given the experimental data. The criterion prescribes the selection of the fishery management option with the lowest expected value of loss of opportunities. For this, we need to build a loss of opportunities matrix or regret matrix as mentioned above.

4.1.2 Decision criteria without mathematical probabilities: Minimax and Maximin

In the absence of sufficient observations to assign probabilities to possible states of nature in bluefin tuna fisheries, there are two decision criteria reflecting different degrees of precaution concerning selection of fishery management strategies: Minimax and Maximin (Anderson and Seijo 2010; Seijo et al. 2004). The Minimax criterion uses the regret or loss of opportunity matrix to calculate the maximum loss of opportunities of each management strategy and selects the one that provides the minimum of the maximum losses. This criterion proceeds as if nature would select the probability distribution, defined for all possible states of nature, which is least favourable for the decision-maker of the bluefin tuna fishery.

Maximin uses the performance variable decision table (payoff table) that estimates the resulting values for a set of combinations of alternative decisions and states of nature. The criterion calculates a vector of the minimum values for the performance variable (e.g., bluefin tuna spawning stock biomass or yield) resulting from each alternative management decision. Then, the fishery manager proceeds to select the maximum of the minimum of those values. This is the most cautious of the decision theory criteria.

4.2 Understanding the potential consequences of a wrong choice

The management advice last provided by the SCRS to ICCAT (2010c) included "Kobe 2 Strategy Matrices" (K2SM) as requested by the Commission (Figure 6), presented separately for the two recruitment scenarios, as well as combined. The K2SMs present useful information regarding the implications that management decisions can have in terms of rebuilding the stock if a given recruitment scenario is true.

The workshop participants felt that managers would also benefit from being presented with additional information related to the potential consequences of making the wrong choice. For example, what are the repercussions if the TACs for the rebuilding plan were chosen on the basis of the low recruitment scenario when, in fact, the high recruitment scenario applies? And vice-versa.

To illustrate the potential consequences of basing a decision on the wrong recruitment assumption, the workshop participants used a simple approach of tabulating consequences in terms of short-term (average 2012-2018) and long-term (average 2048-2050) yield and short-term (2018) and long-term (2050) spawning biomass. SSB and yield values were obtained as medians across the 500 simulation replicates (re-estimates of the VPA obtained by bootstrap resampling of the data inputs) projected forward in time to 2050. This approach made use of the same projections used by SCRS to construct the K2SMs but summarized the information differently, in a way that more clearly presents tradeoffs depending on choices. Additional metrics that could be analyzed in such a framework include variability in yield, the number of years in which yield falls below (or rises above) a threshold, or even the time in which it takes to discern which state of nature has the most support in the data. In addition, the workshop participants calculated various risk-averse criteria commonly used in operations research for decision-making under uncertainty (see Appendix 1).

The following two questions were investigated using projection methodologies employed in the 2010 stock assessment (ICCAT 2010c):

4.2.1 Question 1. If the stock is managed following a strategy X based on assuming one of the recruitment scenarios, what happens if it is the "wrong" recruitment scenario?*

The strategies examined, X*, included:

- F_{MSY}
- $F_{0.1}$
- $F=0.5^* F_{MSY}$

These translate into five fishing mortality strategies along a continuum from $0.5^*F_{MSY\ B\text{-}H}= 0.03$ to $F_{MSY,\ 2L} = 0.16$ (Table 1).

For illustration purposes, we will describe in detail the results of a single decision table and then summarize the overall results. Evaluating the short-term tradeoffs in SSB under the five fishing mortality strategies (Table 1) indicates that the lowest F strategy ($0.5^*F_{MSY\ B\text{-}H}$) is both the maximin and minimax. The maximin represents the maximum value of the minimum outcome, in other words, the maximum SSB, that would be obtained even if the recruitment assumption was incorrect. In this situation, that is obtained as the median of the 500 bootstrap estimates of SSB in 2018 (26,585 mt) projected under $0.5^*F_{MSY\ B\text{-}H}$ for either recruitment hypothesis. The minimax represents the minimum of the maximum loss of opportunity, where the maximum opportunity is the maximum SSB that could be obtained under all considered scenarios. For each F scenario the expected SSB is subtracted from this maximum opportunity to obtain the associated loss and the minimum of the maximum loss is the minimax.

For both short and long-term SSB (Tables 1 and 2) the lowest fishing mortality rate ($0.5*F_{MSY\ B\text{-}H}= 0.03$) is both the minimax and maximin. Yield, however, demonstrates the contrast between short and long-term benefits (Tables 3 and 4). Short-term minimax and maximin are obtained by fishing at the highest F (d5: TAC at $F_{MSY,\ 2L}$; F= 0.16) but the long-term minimum gain in yield is maximized by d3: TAC at $0.5F_{MSY\ 2L}$; F= 0.08 strategy, and maximum loss in yield is minimized by d4: TAC at $F_{0.1}$; F= 0.11 strategy (Table 5). This illustrates the tradeoff between high early yields and much greater future yields that are obtained by an intermediate F strategy that allows biomass to build.

Note that neither of these strategies necessarily achieves ICCAT convention objectives of SSB_{MSY} or F_{MSY} as fishing above F_{MSY} for a given state of nature would represent overfishing and may not allow SSB to build to the SSB_{MSY}. Decision tables developed by the SCRS may need to consider alternative metrics such as SSB performance relative to SSB_{MSY}, rather than simply SSB since the management objective is SSB_{MSY}.

4.2.2 Question 2. How long could it take to rebuild the stock under various strategies X to a level of biomass B* where the differences in expected recruitment between the two scenarios would be substantially different?*

The goal of this exercise was to determine how long it might take to observe the substantially higher level of recruitment that is anticipated under the BH hypothesis and what levels of fishing mortality would allow these levels to be observed, if indeed the BH hypothesis is correct. Observing these much higher levels of recruitment would lend substantial credibility to the BH model over the two-line model. The strategies X* were the same as in (1) above, and encompass the strategy implied by ICCAT Recommendation [10-03]. The workshop participants suggested that reasonable values for B* would be between 30,000 and 40,000 tons. Gavaris *et al.* (2009) indicated such a biomass level could inform on the likelihood of gains in average recruitment with increased spawning biomass. Because of the longevity of bluefin tuna, the workshop participants chose to carry out simulations for 50 years.

Under the two-line hypothesis, spawning biomass and yield increase under fishing mortality scenarios lower than the two-line F_{MSY} as the stock is already at the two-line B_{MSY} (ICCAT 2010c). However, recruitment does not increase substantially over current levels and remains at approximately 85,000 age 1 individuals (Figure 7a-d). Further, spawning stock biomass does not increase above 30,000 t during the 2030 time window except at very low F or at quotas less than 1500 mt. In contrast, under the Beverton-Holt spawner-recruitment curve (Figure 8a-d), average recruitment increases substantially above the maximum two-line recruitment level fairly early in the projection time period. For reference, the Beverton-Holt recruitment at a SSB of 30,000 is shown on Figure 7d. Under most levels of fishing mortality equal to or less than F0.1 (0.11), expected recruitment diverges substantially from that expected under the two-line model. At SSB levels of 30,000 t, achieved at F levels below half of the two-line F_{MSY} (0.08) by 2030, expected recruitments are much higher than those expected under the two-line hypothesis. These results indicate that modest reductions in fishing mortality rate to increase spawner biomass to an interim milestone of 30,000 mt should be sufficient to demonstrate the higher average levels of recruitment that would then be a foundation for higher long-term sustainable catches. It should be noted that both projections are products of the estimated stock status trajectory of the VPA fit to the historical data (Figure 9). It should also be noted that the projected recruitment is the median expected recruitment level unadjusted for the lognormal bias correction and that any realized annual recruitment is likely to diverge substantially from expected values.

4.3 Insights Provided by Decision Theory Approach

The decision tables and figures presented above illustrate the expected short-term and long-term performance of a range of harvest rates. Each of these selected rates has a logical basis, but they also can be considered as points along a continuum of possible fishing mortality rates. Lower fishing mortality rates have the greatest positive benefit for long-term biomass and yield, and the greatest negative effect on short-term yield. Conversely, fishing mortality rates near the top end of the possible range have little short-term or long-term effect because these rates are close to status quo. Examination of results from a continuum of possible fishing mortality rates would allow for identification of an intermediate rate that trades off some short-term loss of yield for some long-term gain in biomass and yield. It also would be feasible to calculate the likely consequences of a step-down, say over 5 years, from the current harvest policy down to the F_{MSY} B-H policy.

5. Stock assessment model issues and considerations

Finally, the workshop participants identified a set of issues in the stock assessment that need further investigation and development to continuously improve the scientific advice provided through the assessment process (see Appendix 2). The following issues were discussed:

- Ageing
- Disappearance of the 1973 cohort
- Age-specific M
- Reliability of the early spawning biomass estimates
- Density-dependent catchability
- Other sensitivities and diagnostics
- Development of a statistical stock assessment model
- Stock-recruit relationships in other *Thunnus* species
- Comprehensive analysis of the stock-recruitment relationship and reference points

While many of these issues have been previously raised, our perspective will hopefully provide new impetus for addressing these important questions. Within the next few years, a move to a statistical assessment model is warranted. The model would combine both the western and the eastern populations into the same model, with mixing parameters to separate out eastern fish in the west and vice versa. The model would handle data uncertainty in a more realistic manner and should improve estimates of recruitment and spawning stock biomass over that in the currently use virtual population analysis (VPA) model.

Acknowledgements

The workshop participants kindly thank the National Geographic Society (NGS) and the Pew Environment Group for sponsoring this meeting. In particular the participants would like to recognize the inspirational venue of the NGS boardroom, mission control for many grand, historic explorations. While our explorations occurred largely on paper and with little personal danger, we hope that they contribute in some small way to improved scientific management of bluefin tuna.

References

Anderson, L.G. and J.C. Seijo. 2010. Bioeconomics of Fisheries Management. Wiley-Blackwell, New Jersey. 305 p.

Diaz, G.A. 2011. A revision of western Atlantic bluefin tuna age of maturity derived from size samples collected by the Japanese longline fleet in the Gulf of Mexico (1975-1980). Collect. Vol. Sci. Pap. ICCAT 66(3): 1216-1226.

Everson, J.P., Laslett, G.M., and Polacheck, T. 2004. An intergrated model for growth incorporating tag–recapture, length–frequency, and direct aging data. Can. J. Fish. Aquat. Sci. 61: 292–306.

Fournier, D.A., Sibert, J.R., Majkowski, J., and Hampton, J. (1990). MULTIFAN: a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). Canadian Journal of Fisheries and Aquatic Sciences 47: 301-317.

Frisk, M.G., Duplisea, D.E. and Trenkel, V.M. 2009. Exploring the abundance–occupancy relationships for the Georges Bank finfish and shellfish community from 1963 to 2006. Ecological Applications 21(1): 227–240.

Fromentin, J.M. and Kell, L.T. 2007. Consequences of variations in carrying capacity or migration for the perception of Atlantic bluefin tuna (*Thunnus thynnus*) population dynamics. Can. J. Fish. Aquat. Sci. 64: 827-836.

Gavaris, S., F. Hazin, J.N. Neilson, P. Pallares, C. Porch, V.R. Restrepo, G. Scott, P. Shelton, and Y.Wang . 2009. Proceedings of the joint Canada-ICCAT 2008 workshop on the precautionary approach for western bluefin tuna (Halifax, Nova Scotia, March 17-20, 2008). Col. Vol. Sci. Pap. ICCAT 64(2): 353-379.

Hilborn, R. and Walters, C. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York. 570 p.

Huse, G., Railback, S., and Ferno, A. 2002. Modelling changes in migration pattern of herring: collective behaviour and numerical domination. Journal of Fish Biology 60:571–582.

ICCAT. 2004. Report for biennial period, 2002-03. Part II- Vol. 2.

ICCAT. 2010a. Report of the 2010 ICCAT bluefin data preparatory meeting. BFT Data Prep. Meeting, Madrid, Spain, June 14-19, 2010. 65p.

ICCAT. 2010b. Report of the 2010 ICCAT working group on stock assessment methods. Madrid, April 21-23, 2010. 33p.

ICCAT. 2010c. Report of the 2010 Atlantic Bluefin Tuna Stock Assessment Session (Madrid, Spain - September 6 to 12, 2010. ICCAT Col. Vol. Sci. Pub. 66(2): 505-714.

ISSF. 2011. Report of the 2011 ISSF Stock Assessment Workshop. Rome, Italy, March 14-17, 2011. 18p.

McAllister, M., E.A. Babcock, and E. K. Pikitch. 2001. Using Bayesian methods to improve stock assessment and management of stock rebuilding when there is uncertainty in processes affecting future recruitment. Col. Vol. Sci. Pap. ICCAT 52(3): 1078-1093.

McAllister, M. 2012. Using Bayesian methods to evaluate the credibility of stock-recruitment relationships for western Atlantic bluefin tuna. In press.

Melvin, G.D. and Stephenson, R.L. 2007. The dynamics of a recovering fish stock: Georges Bank herring. ICES Journal of Marine Science 64:69–82.

Methot, R.D. 2000. Technical description of the stock synthesis assessment program. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-43, 46 p.

Parmigiani, G. and L. Inoue. 2009. Decision Theory: Principles and Approaches. Wiley Series of Probability and Statistics, John Wiley & Sons Ltd. West Sussex, UK. 372p.

Petitgas, P. 1998. Biomass-dependent dynamics of fish spatial distributions characterized by geostatistical aggregation curves. ICES Journal of Marine Science 55: 443–453.

Porch, C.E., S.C. Turner, and G.P. Scott. 2001. Updated catch-at-age analyses of West Atlantic bluefin tuna 1960-1997. Col. Vol. Sci. Pap. ICCAT 52(3): 1052-1066.

Ravier, C., and Fromentin, J. M. 2001. Long-term fluctuations in the eastern Atlantic and Mediterranean bluefin tuna population. ICES Journal of Marine Science 58: 1299–1317.

Restrepo, V. R., G. A. Diaz, J. F. Walter, J. Neilson, S.E. Campana, D. Secor, and R. L. Wingate. 2010. Updated estimate of the growth curve of Western Atlantic bluefin tuna. Aquatic Living Resources. 23: 335–342.

Richards, F.J. 1959. A flexible growth function for empirical use. J. Exp. Bot. 10(2):290-301.

Seijo, J.C.; E. Pérez and J.F. Caddy. 2004. A simple approach for dealing with dynamics and uncertainty in fisheries with heterogeneous resource and effort distribution. Marine and Freshwater Research 55:249-256.

Seijo, J.C.; O. Defeo & S. Salas. 1998. Fisheries bioeconomics: theory, modelling and management. FAO Fisheries Technical Paper. 368:108p.

Shotton, R.; R.I.C.C. Francis. 1997. "Risk" in fisheries management: a review. Can. J. Fish. Aquat. Sci. 54: 1699-1715.

Tamdrari, H., Martin Castonguay, M. Brêthes, J-C. Galbraith, P.S. and Duplisea, D.E. 2011. The dispersal pattern and behaviour of Atlantic cod (*Gadus morhua*) in the northern Gulf of St. Lawrence: results from tagging experiments. Can. J. Fish. Aquat. Sci. 69: 112–121.

Tangen, M. The Norwegian fishery for Atlantic bluefin tuna. ICCAT Col. Vol. Sci. Pap. Vol. 63:79-93.

Taylor NG, McAllister MK, Lawson GL, Carruthers T, Block BA. 2011. Atlantic bluefin tuna: a novel multistock spatial model for assessing population biomass. PLoS ONE 6(12): e27693. doi:10.1371/journal.pone.0027693.

		Alternative States of Nature		
F	Alternative Decisions	1. (BH)	2. (2L)	Min
***0.03**	**d1: TAC at 0.5*Fmsy b-h**	**26585**	**27725**	**26585**
0.06	d2: TAC at Fmsy, b-h	22535	23680	22535
0.08	d3: TAC at 0.5*Fmsy 2L	20720	20430	20430
0.11	d4: TAC at F0.1	18870	18130	18130
0.16	d5: TAC at Fmsy, 2L	13760	14295	13760

Performance indicator: SSB in 2018 (mt)

*Maximin Criteria: Select the maximum of the minimum performance vector

		Alternative States of Nature		
F	Alternative Decisions	1. (BH)	2. (2L)	Max
+0.03	**d1: TAC at 0.5*Fmsy b-h**	**0**	**0**	**0**
0.06	d2: TAC at Fmsy, b-h	4050	4045	4050
0.08	d3: TAC at 0.5*Fmsy 2L	5865	7295	7295
0.11	d4: TAC at F0.1	7715	9595	9595
0.16	d5: TAC at Fmsy, 2L	12825	13430	13430

Performance indicator: SSB in 2018 (mt)

+ Minimax Criteria: Select the decision that generates the minimum of the maximum losses vector

Table 1. Short-term decision tables based upon spawning stock biomass in 2018. The recommended decision is shown in bold.

F	Alternative Decisions	Alternative States of Nature 1. (BH)	2. (2L)	Min
***0.03**	**d1: TAC at 0.5*Fmsy b-h**	**106150**	**41790**	**41790**
0.06	d2: TAC at Fmsy, b-h	63405	30465	30465
0.08	d3: TAC at 0.5*Fmsy 2L	47285	25940	25940
0.11	d4: TAC at F0.1	29570	20500	20500
0.16	d5: TAC at Fmsy, 2L	11285	13150	11285

Performance indicator: SSB in 2050 (mt)

*Maximin Criteria: Select the maximum of the minimum performance vector

F	Alternative Decisions	Alternative States of Nature 1. (BH)	2. (2L)	Max
+0.03	**d1: TAC at 0.5*Fmsy b-h**	**0**	**0**	**0**
0.06	d2: TAC at Fmsy, b-h	42745	11325	42745
0.08	d3: TAC at 0.5*Fmsy 2L	58865	15850	58865
0.11	d4: TAC at F0.1	76580	21290	76580
0.16	d5: TAC at Fmsy, 2L	94865	28640	94865

Performance indicator: SSB in 2050 (mt)

+ Minimax Criteria: Select the decision that generates the minimum of the maximum losses vector

Table 2. Long-term decision tables based upon spawning stock biomass in 2050. The recommended decision is shown in bold.

		Alternative States of Nature		
	Alternative	1. (BH)	2. (2L)	Min
F	Decisions			
0.03	d1: TAC at 0.5*Fmsy b-h	980	972	972
0.06	d2: TAC at Fmsy, b-h	1713	1691	1691
0.08	d3: TAC at 0.5*Fmsy 2L	2011	2167	2011
0.11	d4: TAC at F0.1	2386	2400	2386
***0.16**	**d5: TAC at Fmsy, 2L**	**2813**	**2900**	**2813**

Performance indicator: Average yield 2012-2018 (mt)

*Maximin Criteria: Select the maximum of the minimum performance vector

		Alternative States of Nature		
	Alternative	1. (BH)	2. (2L)	Max
F	Decisions			Loss of Opp.
0.03	d1: TAC at 0.5*Fmsy b-h	1833	1928	1928
0.06	d2: TAC at Fmsy, b-h	1101	1209	1209
0.08	d3: TAC at 0.5*Fmsy 2L	802	733	802
0.11	d4: TAC at F0.1	427	500	500
+0.16	**d5: TAC at Fmsy, 2L**	**0**	**0**	**0**

Performance indicator: Average yield 2012-2018 (mt)

+ Minimax Criteria: Select the decision that generates the minimum of the maximum losses vector

Table 3. Short-term decision tables based upon average yield 2012-2018 (mt). The recommended decision is shown in bold.

Long-term

F	Alternative Decisions	Alternative States of Nature: 1. (BH)	2. (2L)	Min
0.03	d1: TAC at 0.5*Fmsy b-h	3526	1384	1384
0.06	d2: TAC at Fmsy, b-h	4430	2090	2090
0.08	d3: TAC at 0.5*Fmsy 2L	4304	2327	2327
0.11	**d4: TAC at F0.1**	**3758**	**2559**	**2559**
0.16	d5: TAC at Fmsy, 2L	2242	2683	2242

Performance indicator: avg yield in 2048-2050 Yield (mt)

*Maximin Criteria: Select the maximum of the minimum performance vector

F	Alternative Decisions	Alternative States of Nature: 1. (BH)	2. (2L)	Max Loss of Opp.
0.03	d1: TAC at 0.5*Fmsy b-h	903	1299	1299
0.06	d2: TAC at Fmsy, b-h	0	593	593
0.08	**d3: TAC at 0.5*Fmsy 2L**	**126**	**356**	**356**
0.11	d4: TAC at F0.1	671	124	671
0.16	d5: TAC at Fmsy, 2L	2188	0	2188

Performance indicator: avg. yield in 2048-2050 Yield (mt)

+ Minimax Criteria: Select the decision that generates the minimum of the maximum losses vector

Table 4. Long-term decision table based upon average yield in 2048-2050. The recommended decision is shown in bold.

Criterion	Minimax	Maximin
short term SSB	d1: TAC at Fmsy, B-H; F= 0.03	d1: TAC at Fmsy, B-H; F= 0.03
long term SSB	d1: TAC at Fmsy, B-H; F= 0.03	d1: TAC at Fmsy, B-H; F= 0.03
short term yield	d5: TAC at Fmsy, 2L; F= 0.16	d5: TAC at Fmsy, 2L; F= 0.16
long term yield	d3: TAC at 0.5Fmsy 2L; F= 0.08	d4: TAC at F0.1; F= 0.11

Table 5. Summary of decision criteria results.

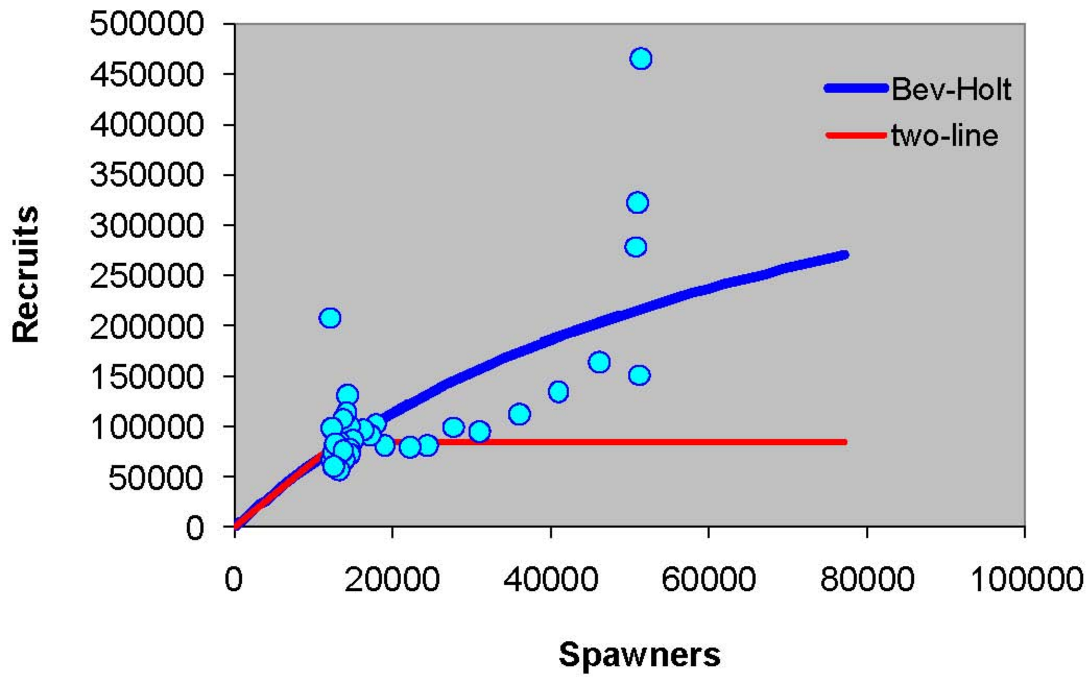

Figure 1. The spawner-recruit relationships fit to the 2010 VPA base model. Points represent the estimates from the VPA.[2] (ICCAT 2010a)

[2] SSB and recruits (age 1) represent VPA outputs. The two-line model was not fit to the data but is a product of an assumed hinge point and an assumed maximum recruitment.

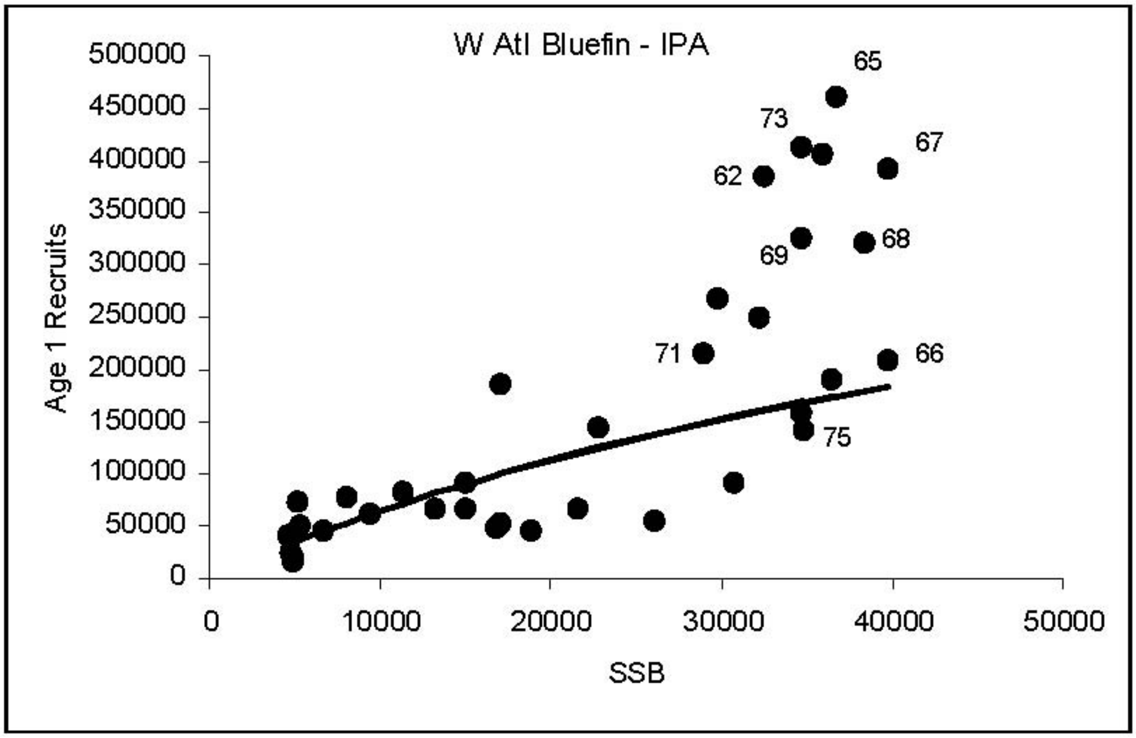

Figure 2. Estimated relationship of spawners to recruits from 1960 to 1997 with fitted Beverton and Holt curve using the age-structured model CATCHEM.[3] (Porch et al. 2001)

[3] Model assumes that bluefin tuna caught by the Japanese longline fishery off Brazil during the 1960s were of western origin and uses an older growth curve that has since been substantially revised.

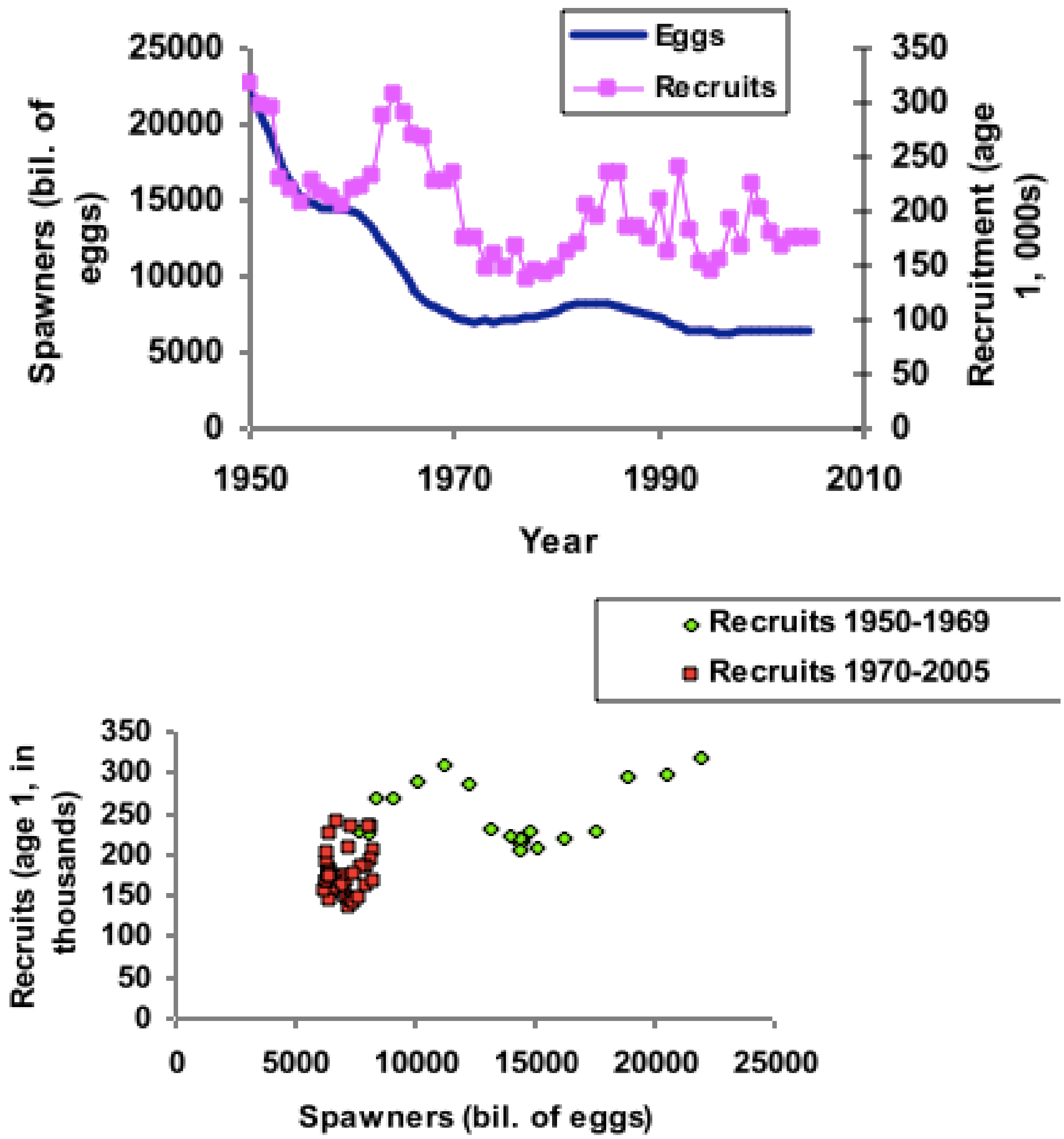

Figure 3. Trends in spawners, eggs and recruits for the western stock calculated using the MAST model.[4] (Taylor *et al.* unpublished data)

---

[4] The MAST catch-age-model was fitted to ICCAT stock assessment data for both stocks, electronic tag track data, otolith microchemistry data, and conventional tag data. The model removes the effects of migrations of fish from the eastern spawning stock into the western Atlantic Ocean from the construction of recruitment and spawner abundance for the western stock. The model is fitted presuming a Beverton-Holt model and allows for annual deviates from model predicted recruitment. The median age at maturity for the western stock was 12 years, and the maximum age of fish in the model was 30 years.

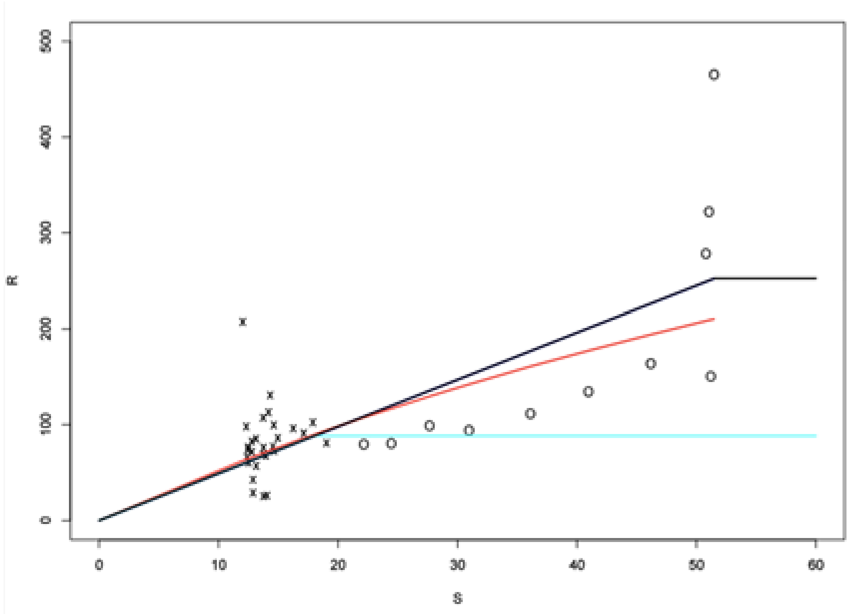

Figure 4. Stock recruitment models fit to the 2010 VPA results for the years 1971-2005.

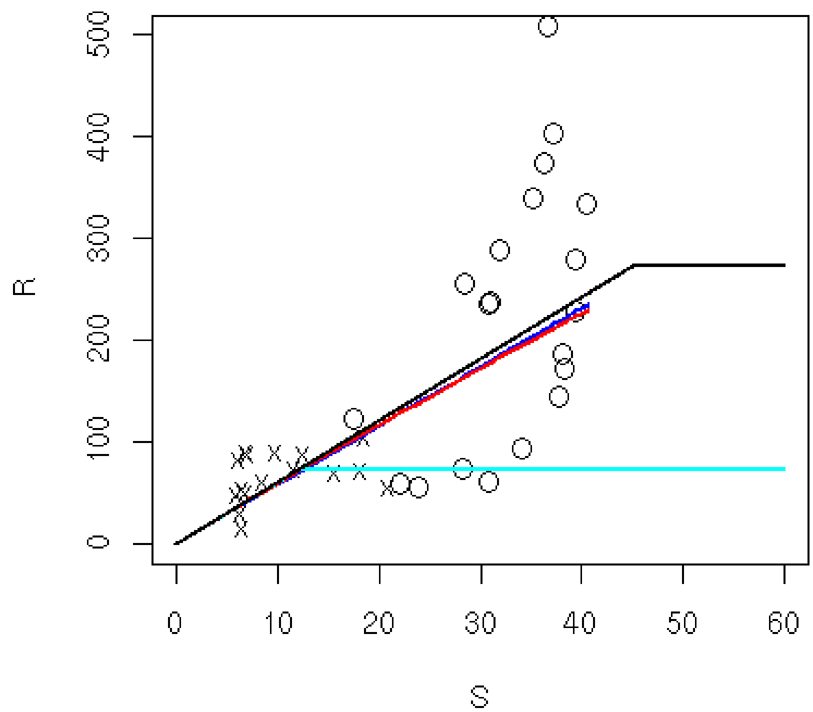

Figure 5. Stock recruitment models fit to the Porch *et al.* (2001) assessment for the years 1961-1997.

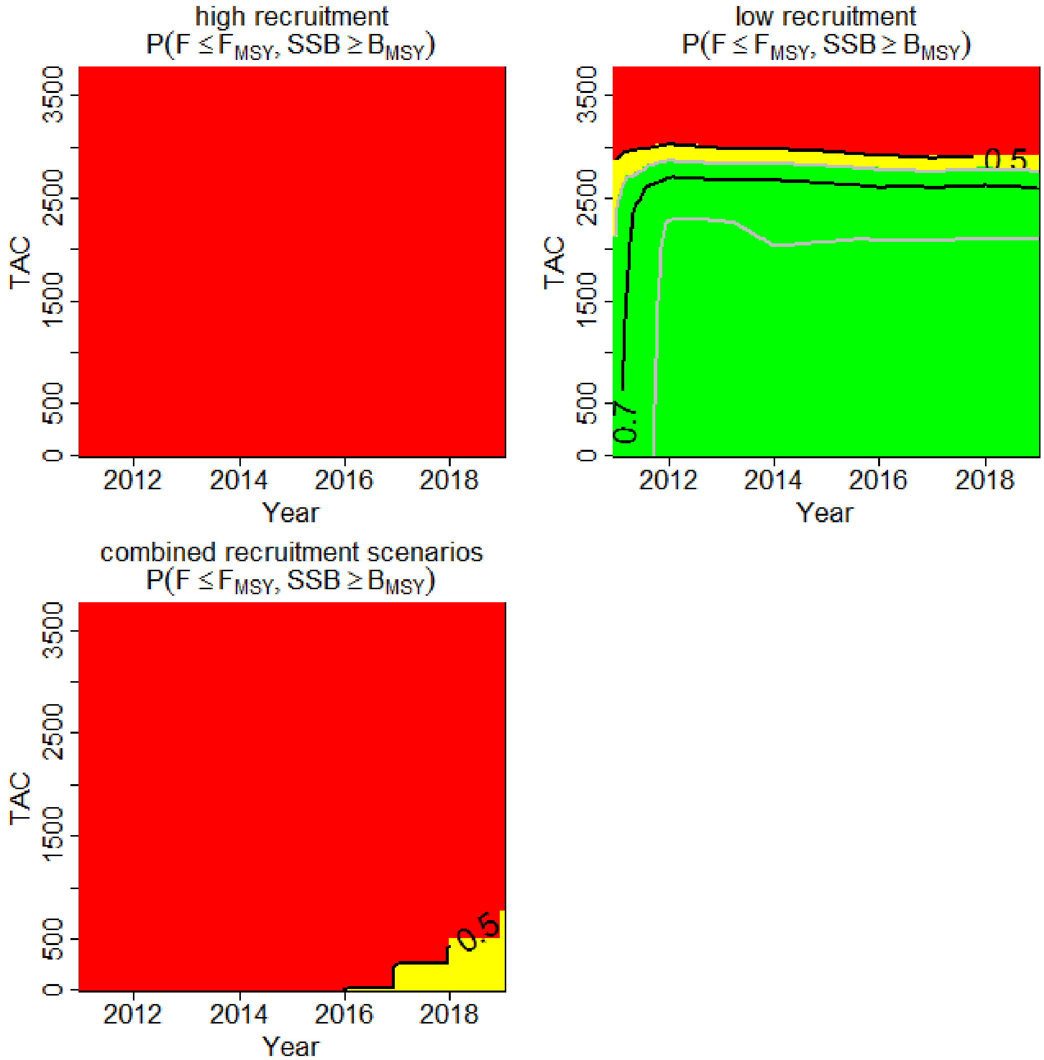

Figure 6. K2SMs provided by SCRS in 2010 to support the Commission's decision for TAC setting for WBFT, consistent with its rebuilding plan. The graphics depict projected odds of success in rebuilding provided TAC's indicated are adhered to, based on quantified uncertainty in the WBFT 2010 stock assessment. Rebuilding objectives, including time-frame, were first defined in the Commission's Rec [98-07]. In these graphics, red coloration implies probability of achieving goal, given the future recruitment hypothesis indicated and quantified uncertainties is lower than 50%. Yellow coloration implies probability of between 50% and 60% and green >60%. Certain probability isopleths at 50% and above are also depicted. SCRS indicated that success in achieving rebuilding plan objectives depends mostly on belief about future recruitment potential.

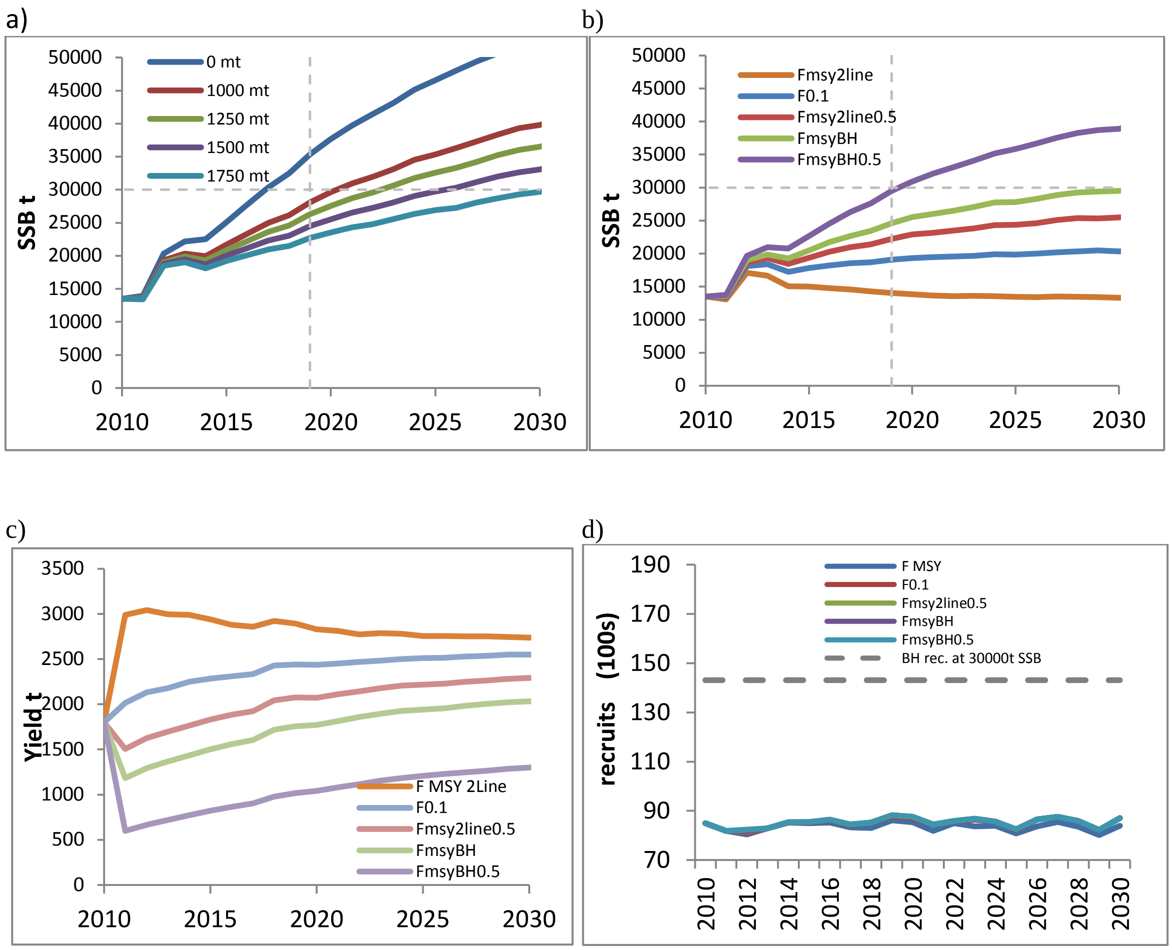

Figure 7. a. Projected SSB at fixed TACs, b. Projected SSB under fixed F, c. Projected yields and d. recruitment at fixed F trajectories, some of which would allow growth in SSB to an informative level of 30,000t (horizontal dashed line in a and b) under the assumption that recruitment will not increase with increasing biomass (**two-line hypothesis**). Vertical dashed line in (a) and (b) represent the 2019 terminal point of the current ICCAT rebuilding plan. The horizontal line in (d) is the recruitment estimated by the Beverton-Holt model at a biomass level of 30,000 t.

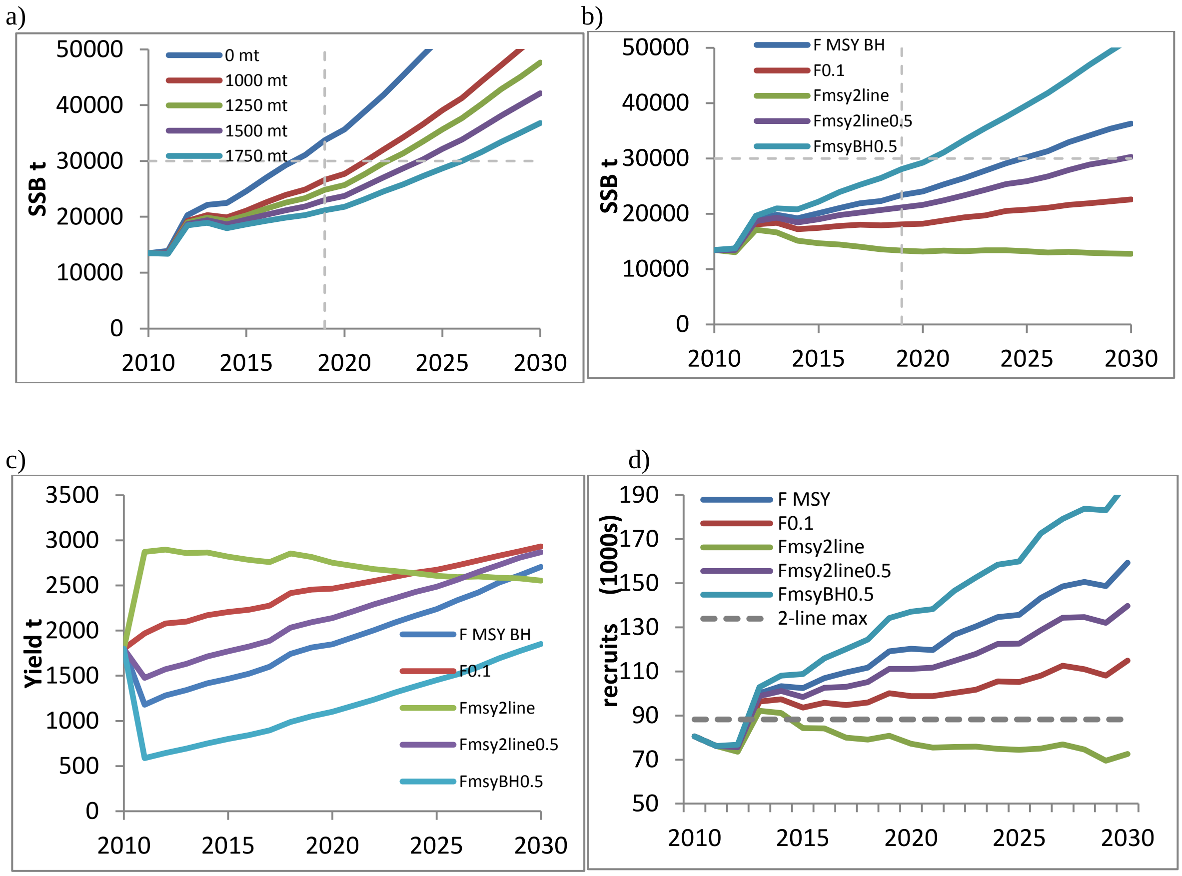

Figure 8. a. Projected SSB at fixed TACs, b. SSB under fixed F, c. yields and d. recruitment at fixed F trajectories, some of which would allow growth in SSB to an informative level of 30,000 t (horizontal dashed line in a and b) under the assumption that recruitment will increase with increasing biomass (**Beverton-Holt model**). Vertical dashed line in (a) and (b) represents the 2019 terminal point of the current ICCAT rebuilding plan. The horizontal line in (d) is the maximum recruitment estimated for the two-line model.

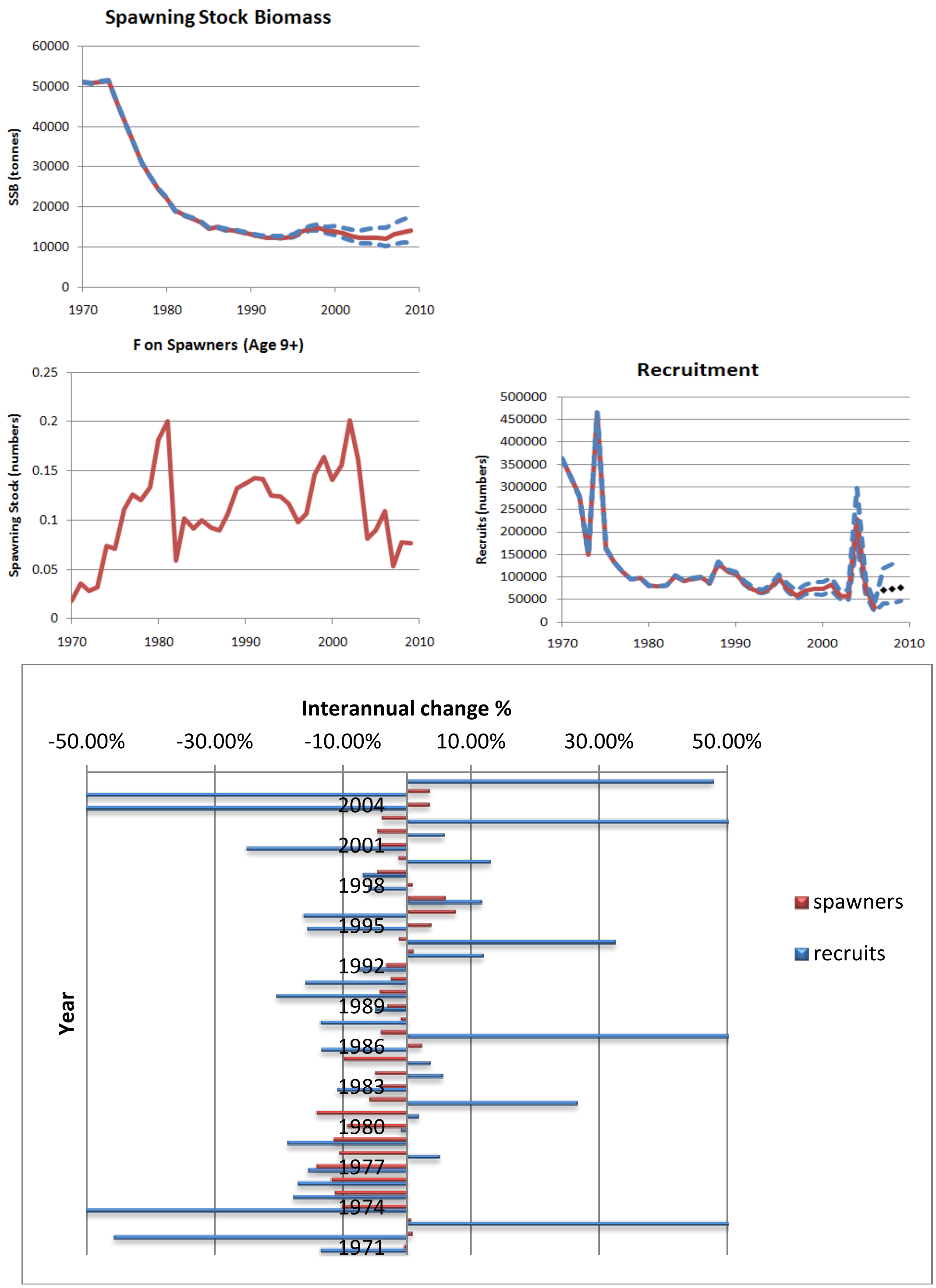

Figure 9. Upper plate: Time-trend of median estimates of spawning stock biomass (t) from 1970 to 2010 with approximate 80% confidence bounds (dashes) based on bootstrapping. Center left plate: Time-trend of median estimates of apical F on spawning age fish (ages 9 and older) from 1970 to 2010. Center right plate: Time-trend of median estimates of recruits (age 1) from 1970 to 2007 with approximate 80% confidence bounds (dashes) based on bootstrapping (note that age 1 abundance estimates from 2008-2010 (dots) are considered unreliable and substituted based on an assumed SRR for the purposes of forward projections. Lower plate: Time-trend of interannual rate of change (%) in median estimates of spawning biomass and recruits for the time-period 1971-2007.

Appendix 1. Decision-making under uncertainty

Under a risk analysis framework applied to the question at hand (Section 4.2 Understanding the potential consequences of a wrong choice), the calculated consequences of a given decision will generally be a function of (i) the cause-effect relationships represented in the recruitment model built to explore management questions, (ii) the parameter set used, and (iii) the possible states of nature concerning the alternative recruitment model/parameters used.

The MINIMAX and MAXIMIN criteria (Seijo *et al.* 1998) were used. These are risk-averse criteria that work as follows:

MAXIMIN: This criterion is used to select a decision that maximizes a minimum gain. For example, if a high SSB is desirable: For each decision d, calculate the minimum SSB that would result from applying that decision across the possible states of nature. Then, select the decision d that results in the largest of these minima.

MINIMAX: This criterion is used to select a decision that minimizes the maximum possible loss of opportunities for a worse-case scenario. For example, if large losses in yield are to be avoided: For each decision d, calculate the maximum Yield loss that would result from applying that decision across the possible states of nature. Then, select the decision d that results in the smallest of these losses.

A formal decision-theoretic framework makes use of the probabilities that the various states of nature apply (iii, above) to weight the outcomes by the corresponding probabilities. The workshop participants discussed at length the calculation of the probability that the high or low recruitment scenario is true. Methods such as those of McAllister et al. (2001) can be used to calculate these from Bayes weights. But the results obtained from these methods can be influenced by the choice of many of the statistical settings for the analysis, *e.g.*, the final cutoff year of the data and the form of the priors for model parameters; a careful inspection of these is needed by the analysts to ensure that the most scientifically defensible and objective choices are made. The workshop participants generally supported the approach but recognized that the application was still undergoing some refinements.

In this application, the workshop participants decided to illustrate the use of decision-making under uncertainty with and without making use of these probabilities. To illustrate the methodology of including Bayes factors, a set of decision tables in which the Beverton-Holt and two-line relationships are considered to be equally plausible (50:50), as was assumed in the 2010 assessment, was developed. All the Bayes factor tables were identical to the minimax criteria, which resulted in the same decision as the maximin, except in the case of long-term yield. The long-term yield table with Bayesian criteria equal to 0.5 is therefore included below. (Appendix Table 1). Note that if the Bayes weights developed by McAllister (2012), which had indicated a 4.8:1 higher likelihood for the Beverton-Holt model, were used, then the Bayesian minimax becomes the Beverton-Holt F_{MSY}.

		Alternative States of Nature		
	Alt. decision	1. (BH)	2.(2L)	min
F	Weight	0.5	0.5	
0.03	d1: TAC at 0.5Fmsy b-h	903	1299	1101
0.06	d2: TAC at Fmsy, b-h	0	593	296
#0.08	**d3: TAC at 0.5Fmsy 2L**	**126**	**356**	**241**
0.11	d4: TAC at F0.1	671	124	397
0.16	d5: TAC at Fmsy, 2L	2188	0	1094

Performance indicator: Average yield in 2048-2050 (mt)

Bayesian minimax criteria: Select the decision that generates the minimum expected value of loss of opportunities

Appendix Table 1. Long-term decision table based upon average yield in 2048-2050. The recommended decision is shown in bold. The Bayesian criteria use equal (0.5) weighting of hypotheses.

Appendix 2. Stock assessment model issues and considerations

Neither the two-line model nor the Beverton-Holt model fit the stock-recruitment data well. This suggests that either the processes determining recruitment do not conform very well to either model or that the stock assessment model used to estimate the spawning biomass and recruitment may be misspecified, causing bias in the estimates of either the spawning biomass, recruitment, or both. Several issues identified with the current stock assessment should be addressed, and this may improve the scientific advice about the form of the stock-recruitment model. Most of these issues have been previously identified and, in some cases, investigated through several different analyses. However, further research and/or data collection is needed to resolve these issues. In particular, a statistical stock assessment model should be developed to provide a more detailed framework to investigate the model assumptions and data consistency.

Ageing

The current assessment model requires a time series of total catch at age. The method used to estimate the catch-at-age matrix is based on cohort slicing using observed modes in the length composition data for young ages and the growth curve for old ages. The catch-at-age matrix shows substantial smearing of fish into adjacent ages, particularly starting around age 5. Estimates of recruitment require reliable estimates of catch at age, so improving the estimation of the catch-at-age matrix should also improve the estimates of recruitment. For example, methods that model the distribution of length-at-age (e.g., MULTIFAN) and use information on the growth curve might be more appropriate.

Due to difficulty in determining cohorts from modes in the length composition data, a growth model is used to slice the length composition data into age data. Biases in the growth model result in biases in the catch-at-age matrix. A recent change in the growth model changed the results of the VPA stock assessment. Analyses should be carried out to determine the sensitivity of the model to different assumptions about growth (e.g., sensitivity to the asymptotic length). The growth curve is estimated from both age-length data and tag-increment data. There are now statistically rigorous methods to simultaneously use both age-length data and tag-increment data (e.g., Everson et al. 2004), and these methods should be applied to the western Atlantic bluefin data. The length composition data also provides information about the growth curve and using a statistical stock assessment model that includes length-composition data, age-length data, and tag-increment data should be considered to improve the growth estimates in addition to its other advantages. A more flexible growth curve such as the Richards (1959) growth curve and time varying growth should also be considered.

Disappearance of the 1973 cohort

A large cohort was born in 1973 and appears as a very high number of age 1 fish in 1974. This cohort is a dominant feature of the catch-at-age data in the early to mid 1970s, but it disappeared quickly. By age 6 this cohort was no longer dominant. The model estimates very high fishing mortalities for this cohort, but low exploitation rates for neighboring cohorts. This implies that the fishery is able to target cohorts. This is most likely to occur if there is spatial separation of age classes. An alternative hypothesis is that these cohorts came from the eastern Atlantic and moved back to the eastern Atlantic when they matured at around age 5. A logical test of this hypothesis is to see if the cohort appears in the catch-at-age data in the Gulf of Mexico in the mid 1980s or in the eastern Atlantic catch-at-size. This would be conveniently done in a statistical stock assessment model comparing observed and expected data given the model assumptions (e.g., constant selectivity). Intuitively, if constant selectivity is used, the model would predict a high proportion of this cohort compared to neighboring cohorts. However, since the fishing mortality rates were estimated to be high for this cohort, using cohort specific selectivity as implied by the VPA would cause this cohort to be absent from the Gulf of Mexico. Another test of the hypothesis would be to look at the growth rates of this cohort compared to other cohorts in the western Atlantic and to cohorts in the eastern Atlantic.

Age-specific M

The catch-at-age matrix shows a substantial decrease in the catch of a cohort a few years after it enters the fishery. This could be due to the selectivity of the gear, but it also could be due to age-specific natural mortality. The current assessment assumes the same value of natural mortality for all ages. The eastern Atlantic assessment (and other bluefin tuna assessments) assume age-specific natural mortality. Age-specific natural mortality should be considered for the western Atlantic bluefin assessment.

Reliability of the early spawning biomass estimates

The estimates of spawning biomass in the early 1970s may be unreliable because they are composed of cohorts that have not been observed at all ages in the catch-at-age data and are therefore dependent to some extent on the terminal age fishing mortality assumptions. Bootstrap analysis may underestimate uncertainty in the early spawning biomass because they do not take the model uncertainty into consideration and the early spawning biomass estimates may be particularly sensitive to some model assumptions (e.g. asymptotic length). Similarly, recent recruitments that have not been observed in all ages in the catch-at-age data may also be uncertain.

The recruitment and spawning biomass estimates from the model starting in 1970 should be compared with estimates from a model starting earlier with and without Brazilian catch. This may provide insight into the reliability of the estimates and determine if the stock-recruitment conclusions are robust to the catch off Brazil, which may be from the eastern Atlantic spawning stock.

Density-dependent catchability

The indexes of abundance used in the western Atlantic Bluefin tuna assessment are principally based on catch rates by the fishery. Fishery catch rates are a measure of the rate of encounter between fish and fishermen in the areas where fishing occurs. The assessment model interprets these data as being proportional to the total abundance of the stock throughout its range. The ecology of bluefin tuna and the searching behavior of fishermen can cause a non-proportional relationship between catch rates and total abundance. For the bluefin tuna assessment, this means that it is possible that the current assessment has underestimated the degree to which the abundance of bluefin tuna has declined.

If bluefin tuna tend to expand/contract the range over which they occur as total abundance increases/ decreases, then the density of fish where the fishery is occurring may remain relatively constant. This will cause the catch rate observed in the fishery to be an insensitive indicator of the change in total abundance. Changes in the distribution of fishing effort can also cause a stability of catch rates because the spatial distribution of bluefin tuna is patchy. Effective fishermen are successful predators capable of searching out these patches of higher bluefin abundance. Since it is extremely difficult to account for the amount of their effort that goes into this searching, the measured catch rates principally reflect the density of fish after they have found the good patches. Assessment scientists do not receive catch rates averaged over the range of the stock, as they would if a randomized fishery-independent survey could be conducted. Instead they get catch rates averaged over the places where fishing occurs.

The degree to which catch rates have a non-proportional link to total abundance cannot be directly measured. However, assessment models can use an asserted degree of non-proportionality and, in some cases, they can internally estimate the degree of non-proportionality. Allowing the assessment model to estimate the degree of non-proportionality for all indexes will result in greater uncertainty in the overall assessment result. This occurs because some of the information from the catch rate trends is being used to estimate these non-proportionality factors rather than to provide information about the trend in the stock itself. Although this result is less precise, it should no longer be biased by an incorrect assumption of proportionality. A hybrid approach in which some well-standardized indexes are treated as proportional to abundance and other indexes have estimated non-proportionality factors is a common practice in assessment modeling.

Other sensitivities and diagnostics

Several diagnostics might provide insight into the reliability of the recruitment estimates. The proportion of each recruitment attributed to the catch of each fishery should be calculated to determine how much of the recruitment comes from poorly sampled fisheries. The proportion of each recruitment attributed to catch at each age should be calculated to determine how much recruitment comes from poorly represented ages.

Several sensitivity analyses should be carried out to investigate the robustness of the stock-recruitment analysis. For example, the model should be run with different maturity ogive assumptions.

Development of a statistical stock assessment model

In the next few years, a move to a statistical assessment model (e.g., MAST, Taylor *et al.* 2011; Stock Synthesis, Methot 2000, Methot and Wetzel in press; Multifan-CL, Fournier *et al.* 1990) is warranted. The model would combine both the western and the eastern populations into the same model, with mixing parameters to separate out eastern fish in the west and vice versa. It is likely that the model would have local stock-recruitment relationships, not a global one. There may be a need for more spatial areas to deal with local fisheries and stock composition of the fisheries.

The model would use data disaggregated by region, season, and fleet, which would lead to more detailed understanding of population dynamics, including recruitment and spawning stock biomass. The model would be fitted to otolith microchemistry, genetics, tag, and archival tag data, along with the standard stock assessment data (age and length compositions, surveys, CPUEs).

The model would handle data uncertainty in a more realistic manner and could increase accuracy and precision in R and SSB over that in the VPA model, due to errors in age data caused by the inaccurate method of cohort slicing and potential errors in the oldest ages due to incorrect terminal fishing mortalities. The catch-at-age would be constructed by the model using internally estimated growth and variation of length-at-age while taking the population dynamics into consideration rather than substituting data for missing catch-at-age. Another advantage of a statistical assessment model over VPA is that better diagnostics are available to examine the fits of the model to the datasets and the statistical distributions of the data. A data point from an archival tag (aggregated in an appropriate form) can be treated as coming from a single new release.

One complication that would have to be dealt with in the statistical model is the specification of gear selectivity. (This is not an issue in VPA because age-selectivity automatically varies over time.) In examining the catch-age data, some participants in the workshop noticed that some fisheries may have the capability to target strong year-classes. For example, the strong 1973 year-class had fishing mortalities for ages 1-4 that were about 0.6, whereas other cohorts had fishing mortalities near 0.1-0.2. Potential solutions could be to have different selectivity parameters for each strong year-class, density dependence selectivity, or to model selectivity as a random walk. Participants also noticed that cohorts adjacent to strong cohorts appeared to have larger catches as they get older, suggesting ageing error might cause smearing of the age composition data. The statistical model uses the length data directly and should address this issue. Workshop participants support efforts to collect data that are necessary to construct a statistical assessment model, such as tagging information, otolith chemistry, and ageing techniques.

Stock-recruit relationships in other Thunnus species

The workshop participants reviewed tuna stock-recruitment relationships used worldwide. A key parameter of most stock-recruitment models is steepness, h, or the ratio of recruitment expected when the spawning stock size is 20% of the virgin stock to the recruitment expected at virgin stock size. Steepness in the Beverton-Holt model ranges from 0.2-1 with higher values of steepness corresponding to a more productive and more resilient stock. A recent report (ISSF 2011) conducted a meta-analysis of steepness values used for tropical, temperate and bluefin tunas throughout the world’s oceans. Western Atlantic and Southern BFT h values were between 0.5 and 0.6, among the lowest of all the tuna stocks examined. In contrast, Eastern Atlantic BFT had h values above 0.9. Similarly, the annual variations in recruitment were analyzed with temperate tunas and WBFT and EBFT, in particular, exhibiting higher levels of recruitment variability than tropical tunas.

While there is some consistency for tropical tuna stocks, the inconsistency in steepness values between Eastern and Western bluefin and in recruitment variability between southern bluefin tuna and EFBT and WBFT make these meta-analytic results of limited use in western Atlantic bluefin. There were several reasons for this and for the ongoing debate about Atlantic bluefin both eastern and western. First, it has long been recognized that individual bluefin move across the ocean. However, the current working hypothesis is that there is mixing on feeding grounds but that the preponderance of spawning fish return to their natal grounds. Therefore, the fisheries exploiting western Atlantic bluefin are likely to be exploiting eastern bluefin as well, but the degree to which the mixing impacts the western Atlantic breeding population assessment is not accounted for in ICCAT's assessment. Additionally, there are large differences in the age of maturity between Western and Eastern breeding stocks (9+ Western, 4+ Eastern), and one study suggests that age at 50% maturity in the Gulf of Mexico may be as high as 15 (Diaz 2011).

Comprehensive analysis of the stock-recruitment relationship and reference points

In general, reliable estimates of the steepness parameter of the Beverton-Holt stock-recruitment relationship are unlikely, while common reference points are highly dependent on steepness. Due to the flat yield curve when steepness is high, underestimating steepness has lower risk of lost equilibrium yield than overestimating steepness. The implied survival of the Beverton-Holt and Ricker stock-recruitment models declines more rapidly with abundance at low stock sizes, which is counterintuitive, and there is a new more general stock-recruitment model that allows survival to decline more rapidly as the population reaches the carrying capacity. Changes in the spatial distribution of the spawning stock as abundance changes might be the most important determinant of the stock-recruitment relationship.

Appendix 3. List of Workshop Participants

Cooper, Andrew
School of Resource and Environmental Management, Simon Fraser University
TASC 1 Building, Room 8405, 8888 University, Burnaby, British Columbia, CANADA V5A 1S6
778-782-3954, andrew_cooper@sfu.ca

McAllister, Murdoch
University of British Columbia
Fisheries Centre, 2202 Main Mall, Vancouver, BC CanadaV6T 1Z4
604-822-3693, m.mcallister@fisheries.ubc.ca

Maunder, Mark
Inter-American Tropical Tuna Commission
8604 La Jolla Shores Drive, La Jolla, CA, 92037-1508, USA
(858) 546-7027, mmaunder@iattc.org

Methot, Richard
NOAA Fisheries – Office of Science and Technology
2725 Montlake Blvd E., Seattle, WA 98112
206-860-3365, richard.methot@noaa.gov

Porch, Clay
NOAA Southeast Fisheries Science Center
75 Virginia Beach Drive, Miami, FL, USA 33149
305-361-4232, Clay.Porch@noaa.gov

Powers, Joseph
Louisiana State University, Department of Oceanography and Coastal Sciences
2147 Energy, Coast and Environment Building, Baton Rouge, LA 70803
225-578-7659, jepowers@lsu.edu

Quinn, Terrance
University of Alaska, Fisheries Division
17101 Point Lena Loop Rd, Juneau, AK 99801
907-796-5457, Terry.Quinn@alaska.edu

Rosenberg, Andrew – Chairman
Conservation International
2011 Crystal Drive, Suite 500, Arlington, VA 22202
703-341-2436, arosenberg@conservation.org

Restrepo, Victor
International Seafood Sustainability Foundation
805 15th Street NW, Suite 650, Washington, DC 20005
703-226-8101, vrestrepo@iss-foundation.org

Scott, Gerald
NOAA Southeast Fisheries Science Center
75 Virginia Beach Drive, Miami, FL, USA 33149
305-361-4596, gerry.scott@noaa.gov

Seijo, Juan Carlos
Marista University of Merida, Periferico Norte Tablaje 13941
Carretera Merida-Progreso, Merida, Yucatan, Mexico 97300
52-999-9429727, jseijo@marista.edu.mx

Stefansson, Gunnar
University of Iceland
Science Institute, Taeknigardur
Dunhaga 5, 107 Reykjavik, Iceland
+354-695-8264, gstefans@gmail.com

Walter, John F. – Rapporteur
National Oceanic and Atmospheric Administration
75 Virginia Beach Drive, Miami, FL, USA 33149
305-365-4114, john.f.walter@noaa.gov