

**PROGRESS ON CATCH-AT-AGE ANALYSES OF ATLANTIC BLUEFIN TUNA
USING A TWO-AREA MIXING MODEL THAT ASSUMES FISH CAN REMEMBER
WHERE THEY CAME FROM**

Porch, C.E.

*National Marine Fisheries Service, Southeast Fisheries Center,
75 Virginia Beach Drive, Miami, Florida 33149, U.S.A.*

SUMMARY

An integrated analysis approach, which allows for the possibility of errors in the catch at age matrix and accommodates mixing, is being developed to estimate the abundance and mortality of the eastern and western populations of Atlantic bluefin tuna. The mixing model employed is more general than those previously considered by the SCRS. It includes the possibility that fish traveling outside their native area may be more or less likely to return home (provided they survive) than they were to leave. Thus, instead of two mixing parameters, there are four.

RESUME

Une méthode d'analyse intégrée permettant d'éventuelles erreurs dans la matrice de captures à un âge donné et tenant compte du mélange, est actuellement élaborée pour évaluer la mortalité et l'abondance des stocks de thon rouge de l'Atlantique Est et Ouest. Le modèle de mélange utilisé est plus général que ceux qui étaient envisagés auparavant par le SCRS. Il tient compte de la possibilité que les poissons qui se déplacent hors de leur zone de naissance sont plus ou moins susceptibles d'y retourner (dans la mesure où ils survivent) qu'ils ne le sont de la quitter. Par conséquent, il existe quatre paramètres de mélange au lieu de deux.

RESUMEN

Se está desarrollando en la actualidad un enfoque de análisis integrado, que permite la posibilidad de errores en la matriz de captura por clases de edad y admite la mezcla, para estimar la abundancia y mortalidad de las poblaciones de atún rojo en el Atlántico este y oeste. El modelo de mezcla empleado es más general que los que el SCRS tuvo en cuenta anteriormente. Incluye la posibilidad de que los peces que salen fuera de su zona de nacimiento puedan tener más o menos las mismas posibilidades de regresar a su origen (suponiendo que sobrevivan) que las que tuvieron para salir. En consecuencia, en vez de dos parámetros de mezcla, hay cuatro.

INTRODUCTION

SCRS assessments of North Atlantic bluefin tuna populations have relied heavily on tuned virtual population analyses (VPA), and to a lesser extent, age-structured production models. The VPA approach is convenient because it is well-known, easy to apply, and employs only a few parameters. Consequently, the SCRS's attention to other recent advances in catch-at-age analysis has been limited to applications of Deriso et al.'s (1985) CAGEAN program by Collie (1988) and Methot's (1989, 1990) stock synthesis model by Porch et al. (1994a). The SCRS concluded that the abundance patterns estimated by the stock synthesis model were similar to those calculated with the tuned VPA's, but did not produce any recommendations regarding the viability of the stock synthesis model in future SCRS deliberations (ICCAT 1993).

The stock synthesis and CAGEAN models are two of several approaches that explicitly consider errors in the catch data. For convenience I will refer to these as integrated population analyses (IPA). Both the tuned VPA and IPA approaches employ auxiliary data (such as CPUE indices of abundance) to estimate the parameters. The VPA's assume the catch-at-age data is perfect and compute the population statistics sequentially from estimates of the population structure in the terminal year. The IPA's allow that the catch data is imperfect and estimate the population statistics simultaneously over all years. In theory the IPA approach results in a more determinate overall solution than VPA (Methot 1989), but at the expense of an increased number of parameters and a more difficult search.

Atlantic bluefin catch statistics are imprecise, due largely to inadequate reporting by parties not contracting with ICCAT (ICCAT, 1993), suggesting that an integrated approach would be helpful. The matter is complicated, however, by the large number of selectivity parameters required by IPA methods. Most authors have reduced this number by assuming the age and year effects on fishing mortality are independent (separable) and by using simple functions to model the pattern of selectivity by age.

In recent years the SCRS has become increasingly concerned with the importance of mixing between the western and eastern bluefin stocks. Butterworth and Punt (1994) introduced a two-area VPA that allows mixing as a fixed fraction of the numbers in each area at the end of the year. Similar models with some additional refinements were examined in NRC (1994), Punt and Butterworth (1994), Porch et al. (1994b), and Porch (1994). They found that even fairly low mixing rates on the order of a few percent could allow for significant departures from the single-stock assessments.

At the 1994 SCRS meeting on east Atlantic bluefin in Malaga, scientists expressed two concerns regarding the models applied in the papers above. First, the models assume that bluefin tuna spawn opportunistically, whereas the scientists felt that bluefin were faithful to their natal spawning grounds. Second, the scientists questioned the justification for treating the eastern Atlantic and Mediterranean Sea as a single homogenous unit (ICCAT 1994). It was proposed that the two matters be considered separately at first, and together if both were demonstrated to have a

significant effect on the outcomes of the assessments. In either case, however, it was recognized that a forward projection approach with the recruitments treated as estimable parameters would be much easier to implement than the backward projection approach of conventional VPA's.

The purpose of this paper is to present an alternative movement model, intended to address the SCRS's concern that bluefin tuna appear to exhibit spawning site fidelity, within the framework of an integrated analysis estimation procedure. It is hoped that this preliminary paper will spark further discussion and suggestions for maximizing the utility of the software to the SCRS.

THE MIXING MODEL

The model developed here considers two populations that share a common range but are faithful to different spawning grounds. The model divides the range into two areas that each include within their borders one of the two spawning grounds. For convenience, the two areas (and spawning grounds) will heretofore be referred to as 'east' and 'west'. Members of a population now occupying the area in which they were born will be referred to as residents. Members of a population currently occupying the area other than the one in which they were born will be referred to as visitors.

Individuals originating from either population are assumed to be equally likely to survive a fixed period of time in a given area, but the probability of surviving may differ between the two areas. This is tantamount to assuming that the residents and visitors survive at the same rate and that fishing effort is randomly distributed within each area (or that the fish instantaneously redistribute themselves). While obvious abstractions of reality, these assumptions are implicit to all models that do not consider space as a continuous variable (including conventional VPA's).

Movement between the areas is affected by allowing a fixed fraction E_α of the resident population in area α to emigrate to the other area at the beginning of each year. At the end of the year a fixed fraction I_α of the visitors from area α are allowed to return to the resident population. Thus, the model keeps track of four groups and uses four movement parameters. The mixing model of Butterworth and Punt (1994) and subsequent SCRS-related publications is essentially a special case of the present model where the emigration rate from one area is identical to the immigration rate to the other.

The equation describing the dynamics of the eastern visitors to the west is

$$V_{w,y+1,a+1} = (V_{w,y,a} + E_e R_{e,y,a}) e^{-(M_w + F_{w,y,a})} (1 - I_e), \quad (1)$$

where $V_{w,y,a}$ is the number of age a individuals originating in the east but located in the west at the beginning year of y , F is the fishing mortality rate, and M is the natural

mortality rate. Likewise, the equation describing the dynamics of the western residents is

$$R_{w,y+1,a+1} = R_{w,y,a} (1-E_w) e^{-M_w - F_{w,y,a}} + I_w (V_{e,y,a} + E_w R_{w,y,a}) e^{-M_e - F_{e,y,a}} \quad (2)$$

Similar equations, with the e and w subscripts interchanged, apply to the residents of and visitors to the east.

The total number of fish in the west (or east) areas is simply the sum of the residents and visitors. The corresponding catch equation is then

$$C_{w,y,a} = \frac{F_{w,y,a}}{Z_{w,y,a}} (V_{w,y,a} + E_e R_{e,y,a} + (1-E_w) R_{w,y,a}) (1 - e^{-Z_{w,y,a}}) \quad (3)$$

where Z is the sum of the natural and fishing mortality rates.

THE ESTIMATION SCHEME

The basic tenet behind IPA analyses is that the data are error prone and that these errors need to be accounted for explicitly. The problem is to find estimates of the recruitments and mortality rates that minimize the disparity between the observed and predicted values of the data subject to the constraints imposed by equations (1)-(3). Many measures of the disparity between the observed and predicted values can be used, depending on the assumptions one is willing to make about the behavior of the data. The most popular approach to date has been to assume that the data follow some particular probability distribution and use maximum likelihood techniques.

The IPA procedure currently under development (IPA-2BOX) follows the underlying structure presented in Deriso et al. (1985), but with several important modifications. The selectivities associated with specific indices of abundance, for example, can be estimated outside the search using the method of partial catches described in Powers and Restrepo (1992). Also, an option is provided to incorporate tagging data into the likelihood function.

Parameters

The parameters used in program IPA-2BOX include the natural mortality rate, von Bertalanffy growth coefficients, average age of the plus group, age-specific movement rates, tag shedding rate, tag reporting rate, initial mortality of tagged fish, age and year-specific selectivities (scaled to 1.0), year-specific maximum fishing mortality rates (the rate when the selectivity equals 1.0), year-specific recruitments, and the numbers at age in the first year of the time series. Any of these parameters

can be fixed to specific values at the discretion of the user. Thus, for A number of age groups and Y number of years there are a total of $5A + Y(1+A) + 9$ parameters for each area, which is prohibitively large given the data available for most fisheries (including Atlantic bluefin tuna). The selectivities account for $Y(A-1)$ of these since A-1 selectivities must be estimated each year (one can be held fixed and the others estimated relative to the fixed value). The solution of most investigators has been to assume the selectivities are constant from year to year-- i.e., the age and year effects on F are separable-- and so reduce the number of selectivities to be estimated to A-1 and the total number of parameters to $6A + 2Y + 8$.

A common criticism especially germane to Atlantic bluefin is that the separability assumption is unreasonable over long time series. I compromise by allowing the selectivity to vary among B blocks of years, reducing the number to be estimated to $B(A-1)$. With regard to bluefin tuna, I propose to define these blocks in accordance with the blocks used by the SCRS in recent assessments. Future versions of the program will allow distinct fleets within the fishery in a manner analogous to the stock synthesis model (Methot, 1990; Porch et al., 1994a). This will further relax the separability assumption and allow the selectivity vectors estimated for each fleet to be applied to corresponding CPUE indices of abundance.

Another option available to reduce the number of parameters in the search is to assume that the total catch is known exactly, or at least that it is measured much better than its age-composition. In this case it is reasonable to adjust the maximum fishing mortality rates so that the estimated total catch matches the observed total in each year.

Some of the parameters, such as those associated with growth, can often be estimated independently. This should be done whenever possible to minimize the number of parameters in the search.

Structure of the objective function

The objective function to be minimized includes log-likelihood terms for age-composition of the catch, indices of abundance, frequency of tag recaptures, and weight at age. The data are generally assumed to be lognormally distributed, however an option allowing for normally-distributed abundance data is provided to accommodate the possibility of negative indices (as sometimes happens when the data are standardized using ANOVA techniques). There is also an option to allow the tagging data to be Poisson distributed (see Hilborn, 1990).

The likelihood expressions based on normal and lognormal error assumptions require estimates of the variances. I follow the standard approach of supplying estimates of the variances as inputs while providing an option that allows one the freedom to vary the emphasis on each likelihood component (Fournier and Archibald, 1982; Deriso et al., 1985; Methot, 1990). In this regard it is important to realize that if the data were perfect (or at least unbiased and exceedingly abundant) the magnitude of the variance estimates would be immaterial. However, given that the data are imperfect, different likelihood components will favor different sets of

parameter estimates. If these differences are large one would obviously want to emphasize the data sets in which one has the most confidence.

The objective function uses the log-likelihood expressions (with the constant terms omitted) rather than the likelihoods to improve the computational efficiency of the search algorithm. This does not affect the solution because the parameters that maximize the logarithm of a function are identical to those that maximize the function itself. Nevertheless, the complete likelihood expressions are output to facilitate comparisons of the fits to each data set when different error assumptions apply.

The age-composition likelihood component compares the observed proportion of catch at age, $p_{ya} = C_{ya}/C_y$, with the model predictions assuming a lognormal error structure. The program automatically computes the proportions from the data, so the input matrix may be either the full catch at age matrix or the actual sampling data. The input variances, however, must be on the proportions. Alternatively, one may input the CV's and the program will automatically convert them to the appropriate variances. Future versions of IPA-2BOX will include an option for choosing a multinomial error structure, which is theoretically more appropriate for age-composition data and more convenient because the variances can be determined directly from the sample size.

The index likelihood treats series of annual measures of relative abundance. The input data for each series are automatically scaled relative to the series mean. (As above, the input variances must be scaled appropriately or CV's used.) The predicted value for the i 'th index I_i is given by

$$\hat{I}_{iya} = q_i \sum_a S_{iya} N_{ya} D_{iya}$$

The abundances N are determined strictly by parameters estimated in the search, however the index-specific catchabilities q_i and selectivities S_{iya} are not. (D is a conversion factor that accounts for the time of year the survey takes place and whether or not the index is of biomass or numbers.)

It is generally considered unnecessary to estimate the catchabilities within the search since they can be computed directly from the estimated N 's and observed values of the index (see Powers and Restrepo, 1992, for the appropriate formula). Inasmuch as the program does not yet assign likelihood components to index-specific age-composition data, it is unreasonable to attempt to estimate the index-specific selectivities within the search. They can, however, be estimated outside the search using the formula

$$S_{iya} = \frac{C_{iya}}{C_{ya}} F_{ya} \quad (4)$$

and then rescaling the S_{iya} to make the maximum over age (a) equal to one. This is the approach employed by the ADAPT VPA (e.g., Powers and Restrepo, 1992)-- less

the division by the summation over all ages, which is unnecessary when the selectivities are rescaled. Equation (4) makes the implicit assumption that the catches are known exactly, which is inconsistent with the IPA procedure. Accordingly, program IPA-2BOX replaces the observed catches with the estimated values in equation (4).

The average weight at age can either be input or estimated via the growth equation. In case of the latter, a caveat of some concern is the estimation of the average weight of the plus group. Program IPA-2BOX computes this from the growth curve using the estimated average age of the plus group (a parameter in the search). This approach is accurate provided growth is relatively slow and nearly linear for ages greater than the plus age-- an assumption easily tested by inspecting the growth curve. A likelihood component is included which compares the predicted weight of the plus group to the observed values assuming a lognormal error structure. Preliminary simulations have demonstrated that the average age of the plus group is poorly estimated if this likelihood component is not included.

The likelihood component for the tagging data examines the observed number of recaptures in a given area, which necessitates estimating the initial tagging mortality, tag shedding rate, and the probability that a recaptured tag will be reported in each area (see, e.g., Hilborn, 1990; and Hampton, 1991). Options for lognormal or Poisson errors are currently available. This approach should produce estimates of the movement rates that are superior to those that would be estimated from the tagging data alone because the fishing mortality rates are better determined with catch data and because the abundance data necessarily reflects the effects of movement.

Searching the objective function for parameter estimates

At this point program IPA-2BOX has been applied mostly to artificial data generated without error from a simple one-area model. Two search algorithms were tested: the Nelder-Mead simplex search of Press et al. (1988) and an adaptive nonlinear least-squares search (Dennis et al., 1981). The latter failed to converge when the number of parameters became large, whereas the simplex search converged readily for any number of parameters. The simplex algorithm has the additional advantage of being able to handle larger discontinuities in the objective function than most other search algorithms. This is particularly important when discrete movement models (such as those discussed here) are used because the partial derivatives are not continuous.

The main draw back to the simplex algorithm is that it converges very slowly. Preliminary runs on data for west Atlantic bluefin tuna took over an hour to converge on a 50 Mz 486 PC microcomputer.

The parameters read by the search are scaled to an order of magnitude near one. In this way all of the parameters will be adjusted in a similar manner. When the objective function is called the parameters are rescaled to their absolute values.

REFERENCES

- Anonymous. 1994.
Report of the National Research Council review of Atlantic bluefin tuna. National Academy Press. Washington, D.C. 148 pp.
- Butterworth, D. S., and A. E. Punt. 1994.
The robustness of estimates of stock status for the western north Atlantic bluefin tuna population to violations of the assumptions underlying the associated assessment models. ICCAT Collective Volume of Scientific Papers 42(1): 192-
- Collie, J. S. 1988.
Evaluation of virtual population analysis tuning procedures as applied to Atlantic bluefin tuna. ICCAT Collective Volume of Scientific Papers 28: 203-220.
- Dennis, J. E., Gay, D. M., and R. E. Welsh. 1981.
ALGORITHM 573 NL2SOL-- An adaptive nonlinear least-squares algorithm. ACM Transactions on Mathematical Software 7(3):369-383.
- Deriso, R. B., Quinn, T. J., and P. R. Neal. 1985.
Catch-age analysis with auxiliary information. Can. J. Fish. Aq. Sci. 42:815-824.
- Fournier, D., and C. P. Archibald. 1982.
A general theory for analyzing catch at age data. Can. J. Aq. Sci. 39:1195-1207.
- Hampton, J. 1991.
Estimation of southern bluefin tuna *Thunnus maccoyi* natural mortality and movement rates from tagging experiments. Fish. Bull., U.S. 89:591-610.
- Hilborn, R.
1990. Determination of fish movements patterns from tag recoveries using maximum likelihood estimators. Can. J. Fish. Aquat. Sci. 47:636-643.
- ICCAT. 1993.
Report for the biennial period 1992-1993. Part I (1992). 375 pp.
- ICCAT. 1994.
Report for the biennial period 1992-1993. Part II (1993). 395 pp.
- Methot, R. D. 1989.
Synthetic estimates of historical abundance and mortality for northern anchovy. American Fisheries Society Symposium 6:66-82.
- Methot, R. D. 1990.
Synthesis model: an adaptive framework for analysis of diverse stock assessment data. In Loh-Lee Low (ed.). Proceedings of the Symposium on Application of Stock Assessment Techniques to Gadids. International North Pacific Fisheries Commission Bulletin 50:259-277.
- Porch, C. E. 1994.
A two-area VPA with discrete mixing: Can we discriminate between mixing rates given the present condition of the data? ICCAT Working Document SCRS/94/75.
- Porch, C. E., Turner, S. C., and R. D. Methot. 1994a.
Estimates of the abundance and mortality of west Atlantic bluefin tuna using the stock synthesis model. ICCAT Col. Vol. Sci. Doc. 42(1): 229-239.
- Porch, C. E., Restrepo, V. R., Turner, S. C., and G. P. Scott. 1994b.
Virtual population analyses of Atlantic bluefin tuna incorporating movement and tagging data. ICCAT Working Document SCRS/94/73
- Powers, J.E., and V. R. Restrepo. 1992.
Additional options for age-sequenced analysis. ICCAT Col. Vol. Sci. Pap. 39(3):540-553.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling.
1986. Numerical Recipes. Cambridge University Press. Cambridge, MA. 818 pp.
- Punt, A. E., and D. S. Butterworth. 1994.
Use of tagging data within a VPA formalism to estimate migration rates of bluefin tuna across the north Atlantic. ICCAT Working Document SCRS/94/72 (revised).