SCRS/1987/063 Rev.

Col.Vol.Sci.Pap. ICCAT, 28 : 203-220 (1988)

EVALUATION OF VIRTUAL POPULATION ANALYSIS TUNING PROCEDURES AS APPLIED TO ATLANTIC BLUEFIN TUNA

J. S. Collie
Natural Resources Management Program, Simon Fraser University
Burnaby, British Columbia, Canada

SUMMARY

The purpose of this report is (1) to evaluate the method of tuning virtual population analysis (VPA) developed by Parrack (1986), (2) to evaluate the applicability of this method to estimating the abundance of bluefin tuna as performed at the 1985 and 1986 ICCAT meetings and (3) to recommend modifications to the bluefin assessment methods. In theory, the VPA tuning procedure used in the program CAL is a valid method of estimating population abundance, and CAL produced unbiased results when run with correct data. With data sets likely to contain large errors, such as bluefin tuna, it is not possible to estimate the natural mortality rate with any reliability. The most serious source of error is probably in measuring CPUE and consequently the biggest problem in applying CAL is in selecting appropriate abundance indices. I recommend that the accept-or-reject criteria currently used by ICCAT be replaced by a weighted sum of squares. The program for catch-at-age analysis, CAGEAN, should be considered as an alternative to CAL. Although I did not analyze the Atlantic swordfish data, the same recommendations apply to the estimation of swordfish abundance.

RESUME

Le but du présent rapport est: (1) d'évaluer la méthode d'ajustement de l'analyse des populations virtuelles (VPA) élaborée par Parrack (1986), (2) de juger dans quelle mesure cette méthode peut être appliquée, comme ceci a été fait aux réunions de 1985 et 1986 de l'ICCAT, pour estimer l'abondance du thon rouge, et (3) de recommander des modifications des méthodes d'évaluation de cette espèce. La méthode d'ajustement de la VPA utilisée dans le programme CAL est théoriquement valide pour estimer l'abondance de la population, et le CAL a donné des résultats non biaisés lorsqu'il était utilisé avec des données correctes. Dans le cas de jeux de données qui contiennent probablement des erreurs importantes, comme ceux qui concernent le thon rouge, il n'est pas possible d'estimer de façon fiable un taux de mortalité naturelle. La plus grave source d'erreur surgit probablement lorsque l'on mesure la CPUE, et le problème le plus important que suscite l'application du CAL concerne le choix d'indices adéquats de l'abondance. Je recommande que les critères d'acceptation ou de rejet actuellement employés par l'ICCAT soient remplacés par une somme pondérée des carrés. Le programme d'analyse de la prise à un âge donné, le CAGEAN, doit être considéré comme une alternative au CAL. Bien que je n'aie pas analysé les données sur l'espadon de l'Atlantique, les mêmes recommandations s'appliquent à l'estimation de l'abondance de cette espèce.

RESUMEN

Los objetivos de este informe son: (1) evaluar el método de ajuste del análisis de población virtual (VPA) desarrollado por Parrack (1986), (2) evaluar la aplicabilidad de este método para estimar la abundancia del atún rojo, como se hizo en el curso de las reuniones ICCAT en 1985 y 1986, y (3) recomendar las oportunas modificaciones en los métodos de evaluación del atún rojo. En teoría, el procedimiento de ajuste del VPA que se aplicó en el programa CAL es un método válido de estimación de la abundancia del stock, y CAL dió resultados no sesgados cuando se realizó con datos correctos. Con conjuntos de datos susceptibles de contener errores importantes, como es el caso del atún rojo, no es posible estimar con fiabilidad alguna la tasa de mortalidad natural. Los mayores errores se producen probablemente al medir la CPUE y por lo tanto, el principal problema al aplicar CAL es la selección de los índices de abundancia apropiados. Se recomienda en el documento que los criterios de aceptación o rechazo que actualmente se aplican en ICCAT sean remplazados por una suma ponderada de cuadrados. El programa de análisis de la captura por edad (CAGEAN), debe considerarse como alternativa al CAL. Si bien en el documento no se analizan los datos del pez espada atlántico, las mismas recomendaciones pueden aplicarse a la estimación de la abundancia de pez espada.

Theoretical evaluation of VPA tuning

The basic approach to calibrating the VPA is to minimize the sum of squared deviations (SS) between observed and predicted abundance:

$$SS = \sum_i [\tilde{X}_i - q(\sum_j N_{ij})^b]^2 \qquad (1)$$

where $\tilde{X}_i$ is the observed relative abundance (e.g. CPUE) in year i, N_{ij} is the number of fish of age j in year i predicted by VPA, q is a catchabilaity coefficient and b is a parameter that determines the shape of the relationship between X and N. This sum of squares is minimized in a two-step grid search. Firstly, the series of N_{ij} values is calculated by VPA for a given combination of natural mortality (M) and fully-recruited fishing mortality in the final year of the data (FF). The vector of partial recruitments at age, P_j, is assumed known. Secondly, Eq. 1 is minimized with respect to q (and optionally b).

This two-step procedure masks some potential problems with the method. One problem is that in minimizing Eq. 1 it is implicitly assumed that all the error is in $\tilde{X}$, and that N is measured without error. In fact, there are errors associated with estimating N from C and errors in the catch-at-age data as well. This means that the estimates of q and N must be interpretted with caution. The assertion that the procedure "develops least-squares estimates of stock sizes and fishing mortalities" is not strictly true because only q is estimated by least squares while N_{ij} is calulated by VPA. Another way to consider this problem is that, if the N_{ij} values were calculated by some procedure other than VPA, SS in Eq. 1 might be different and possibly lower. These objections may seem semantic but they are meant to suggest that, instead of using a two-step procedure, it may be preferable to expand N in Eq. 1, as a function of C, and to estimate all the model parameters in one step by nonlinear regression.

In theory it is possible to estimate natural mortality (M) from catch-at-age and relative abundance data, but in practice M is not estimable with any degree of confidence from most real data sets, with a few exceptions such as the Opeongo lake trout data. The reason is that estimates of M are highly correlated with estimates of q, such that if q is over-estimated M is under-estimated.

The same applies to the dispersion parameter b. There are well-known examples of underdispersion (b<1, e.g. schooling fishes) and overdispersion (b>1, e.g. sessile, territorial species) but b can not generally be estimated from the data because it is correlated with the estimates of q, M, etc. Parrack's (1986) assertion that overdispersion (b>1) is a very desirable trait because CPUE is very sensitive to stock size changes is only true at high stock sizes; at low stock sizes CPUE is insensitive to changes in abundance. The most desirable situation is thus a direct relationship between CPUE and abundance (b=1) which obviates the need to estimate b.

Multiple abundance indices

The advantage of Eq. 1 is that it can flexibly include multiple abundance indices (k),

$$SS = \sum_k \lambda_k \sum_i [\tilde{X}_{ki} - q_k(\sum_j N_{ij})^b]^2 , \qquad (2)$$

which may span different years and different ages. I included an extra term λ_k to indicate the relative weight assigned to abundance index k. Parrack's approach is to fit each CPUE set separately and then assign λ_k a value of 0 or 1 according to a predetermined set of criteria. The criteria for accepting an index adopted by the 1985 Bluefin working group (BFWG, Anon 1986) were:

(1) The probability of significant positive correlation between the index and VPA stock size estimates was greater than 0.8.

(2) The residuals between observed and predicted abundance do not exhibit a trend over time. For example, a linear trend could occur because the regression of CPUE on N is forced through the origin.

(3) The minimum sum of squares occurred for a terminal fishing mortality (FF) within the biologically realistic range 0.0001 to 6.0.

The 1986 BFWG (Anon 1987) changed the first criterion such that the correlation coefficient itself should be greater than 0.8 instead of the probability of a significant correlation. The danger of this change is that the correlation coefficient does not account for different numbers of observations in each abundance index. A higher correlation coefficient should be obtained with fewer observations; in using this criterion one would tend to accept indices of shorter duration and reject indicies of longer duration. The 1986 BFWG also narrowed the range of acceptable FF values to 0.00001<FF<1.5.

The problem with these criteria is that they do not necessarily accept correct abundance indices and reject incorrect ones. Choosing incorrect M and FF values to initialize VPA results in spurious abundance trends during years containing incomplete cohorts (Pope 1972, Agger et al. 1973, Ulltang 1977, Sims 1984). An incomplete cohort is one that has not reached the maximum age by the final year of the analysis. In the case of bluefin tuna with 31 age classes and 16 years of data, all years contain incomplete cohorts and it is therefore possible to reproduce any linear trend in abundance (correct or incorrect) by choosing the appropriate M and FF values. On the other hand, it is possible to reject a correct abundance index if that index is measured with a large amount of error.

It is important that the various abundance indices be scaled to the same order of magnitude so that they receive approximately equal weighting in Eq. 2. This point is not mentioned by Parrack (1986) but was recognized by the BFWG which scaled the CPUE indices such that the largest values equals one.

Program notes

The major advantage of the FORTRAN program CAL is that it permits multiple relative abundance indices to be included in the sum of squares, indices which need not span the same number of years or ages. Another advantage is that the program is entirely self contained and does not require any proprietary software except a FORTRAN compiler. The grid search over M and FF values is made more efficient by means of a course and then fine precision search. CAL is written as a structured set of subroutines which makes it easy to understand and to modify. The program is very flexible in the types of relative abundance data it accepts (numbers or weights), and there are options for the

number of parameters to be estimated (F, M, q and b) and the type of output desired. The full output, including parameter estimates and plots, makes it easy to tell how the program has worked.

Test runs of CAL on known data

Parrack (1986) used Opeongo lake trout data to demonstrate that CAL estimated values of M and FF which were very similar to the values estimated by Paloheimo's (1980) regression method. This showed that CAL does produce consistent estimates but it did not test the accuracy of the estimates, nor is it an indication of how CAL would perform on other data sets. The Opeongo lake trout data are in fact among the best quality fisheries data sets in existence because of representative sampling of the catch, direct aging of the fish, and good records of effort kept by the outfitter. The estimates of M, FF and/or q are very sensitive to which ages of lake trout are included in the analysis; this is true of CAL as of Paloheimo's method. For example, if ages 8-12 are used in Paloheimo's method the correlation between M and q is -0.94. If ages 7-11 are used instead, the estimate of q is negative! Unfortunately, the fact that M can be estimated in this example is due more more to good data and good luck than to good methods of parameter estimation.

In my first test run of CAL I used the same data for west Atlantic bluefin tuna as were used in the final run at the 1986 ICCAT meeting (catch data in Table 7 in Anon 1987, five abundance indices, partial recruitment vector M1, and M=0.1). These are referred to as the observed data in subsequent trials. The first test was simply to ensure that my version of the program gave the same results (Table 19 and 20 in Anon 1987) which it did. Next, I used the estimated number-at-age data, summed from ages 1-5, 6-9, 10-30 (bottom rows of Table 19) as correct relative abundance data without error, instead of the observed CPUE indices. The estimated parameters were very close to the true values (Table 2, rows 1 and 2), showing that CAL gives the correct answer when given correct data.

I did not try to estimate the b parameter because, as mentioned above, it is not estimable from most data sets and no attempt appears to have been made to do so by the BFWG. A priori it is difficult to predict whether b should be less than or greater than one because underdispersion caused by fishermen exploiting schools of tuna may be partly counteracted if new areas or subpopulations were discovered as the fishery developed.

I tested the estimation of M using the same correct data and observed data as those described above except that M was allowed to vary. The correct data used above were derived assuming M=0.1 and FF=0.46. The same data, refit by CAL resulted in parameter estimates that were only slightly over-estimated (M=0.11, FF=0.48), showing that CAL can accurately estimate M from correct data. However, even with correct data the 95% confidence region (isopleth 3 in Fig. 1) is quite large, especially in the direction of low M and high FF. If the observed data are fit by CAL and M is allowed to vary, the result is no longer a point estimate but an unbounded valley of equally probable M and FF values (Fig. 2). The high negative correlation between the estimates of M and FF as illustrated by the diagonal isopleths indicates that one cannot expect to estimate both M and FF from the observed data; at best, one can estimate feasible combinations of the two.

Information content of catch-at-age data

The reason that the parameters estimated by CAL are so uncertain is that the catch-at-age data contain relatively little information concerning the impact of fishing versus natural mortality on the bluefin tuna population. The history of the bluefin fishery appears to be a "one-way trip": high catches until 1982 were sustained only by increasing fishing effort (Table 3 and Fig. 7). Abundance has presumably decreased but it is uncertain whether the population is large and unproductive or small and productive; only by allowing the population to recover from exploitation can its productivity be assessed.

The age structure of the catch data does not provide much additional information because, although the data contain 31 age groups, most of the catch occurs at age seven or less (Fig. 3). A few of the cohorts appear stronger than average (e.g. 1971, 1973, 1982) but these peaks cannot be traced beyond age five because of aging errors and high cumulative fishing mortality. Classifying the catch by age is obviously necessary for VPA but beyond age 10 there is virtually no information in catch-at-age that is useful for estimating FF and M.

Tests of CAL on data with known errors

Potential sources of error in the data used to run CAL include sampling of the catch, aging errors, the partial recruitment vector and error in measuring CPUE. Of these, aging errors and CPUE measurement errors are probably the most serious. Catch sampling is obviously a problem in a fishery in which data from several fleets must be combined into a single catch-at-age matrix. Aging errors arise because bluefin tuna age is not measured directly but estimated from an age-length relationship. In this analysis I did not consider the effect of sampling errors on parameter estimation because sampling errors are probably less important than aging errors. Similarly, I ignored the effect of choosing the wrong partial recruitment vector because it can not be estimated by CAL and must be derived independently (e.g. by SVPA or CAGEAN).

Aging errors

The BFWG (Anon 1986) noted that if a growth curve is used to estimate catch-at-age, year class strengths are smoothed for fish older than age four and there is little age discrimination of fish longer than 200 cm (approx. age 10). To investigate the effect of aging errors on parameter estimation I assumed that estimated age was normally distributed about the true age. The standard deviation of aging errors increased with age such that there was no error until age five and the maximum degree of error was reached at age 10. This resulted in a matrix of true versus estimated age (Table 1) in which the severity of aging errors is an age-specific function of the standard deviation, smax.

I ran CAL using the "correct CPUE" data derived in previous trials and catch data with errors (A) derived by multiplying the catch-error matrix (E) by the observed catch-at-age matrix (C)

for various levels of smax (A = EC). Interestingly, there was very little bias in the estimates of FF and q even at high values of smax (Table 2). If the proportion of fish at age were uniform there would be a tendency to overestimate catch at ages 4 to 6 and to underestimate ages 7 to 11 (final column of Table 1). In the case of bluefin tuna, the catch is dominated by fish less than seven years old; only a small proportion of the abundant 5-and-6 year-old fish need to be misaged as older fish to compensate for the larger proportion of the less abundant 7-to-11-year-old fish which are misaged as younger fish. The age distribution of catch (Fig. 3) has been mostly unimodal and fairly stable over time and therefore, the smoothing of year classes due to aging errors does not appear to bias the parameter estimates.

The conclusions of the aging error trials are contingent on the assumption that estimated age is symmetrically distributed around the true age. If there is a systematic bias in estimating age, the estimated FF would also be biased. Errors in sampling the age distribution of catch could be important if the true age compositions were markedly different than those plotted in Fig. 3. For example, if strong year classes were actually present but smoothed away by aging errors, the precision in estimating FF might be much less than it would be without aging errors.

Errors in measuring CPUE

To investigate the effect of measurement errors on parameter estimation, I ran CAL using the observed catch-at-age data and the "correct" abundance indices derived for ages 6-9 and 10-30. Abundance indices ($\hat{X}_{ki}$) were simulated by multiplying the correct indices (X_{ki}) by random lognormal measurement errors (w_{ki}):

$$\tilde{X}_{ki} = X_{ki}\exp(w_{ki} - \sigma_w^2/2) \ . \qquad (3)$$

The constant $\sigma_w^2/2$ is subtracted from each w_{ki} to ensure that the expectation of the measurement error is 1. The standard deviation σ_w was assigned a value of 0.4 which means that half the $\tilde{X}_{ki}$ values are within plus or minus one third of X_{ki}. Judging by the scatter in the observed CPUE indices, this is not an unrealistic level of error.

Having modified CAL to perform multiple trials, I ran the program on 100 sets of randomly generated relative abundance data. The output of each trial included the average error of prediction, the correlation between input and estimated abundance indices (ρ) and the estimate of FF. For each relative abundance index the program recorded the estimated catchability coefficient (q_k), the partial sum of squares (SS_k), the partial correlation coefficient (ρ_k) and the probability that ρ_k was greater than zero.

The estimated FF values are only slightly biased upward (median = 0.52) but the distribution has a long tail of high FF values (Fig. 4). Six of the 100 estimates were at the upper limit of 3.0 set in the grid search. The extreme FF estimates are a consequence of assuming multiplicative lognormal measurement errors; extreme errors in some abundance sets make it appear that the stock has been severely depleted. The corresponding catchability coefficients are unbiased but the distributions are also skewed (Fig. 5). The extremely high q values correspond to the trials with extremely high FF estimates.

The partial correlation coefficients are all quite large (Fig. 6) and all are greater than zero at the 0.01 significance level. The lines at ρ=0.8 divide the data into quadrants with roughly equal numbers of observations. If ρ>0.8 were used as a criterion for accepting or rejecting abundance indices, in roughly one quarter of the trials both indices would be accepted, in one quarter both indices would be rejected and in half the trials one or other of the indices would be rejected. The trials resulting in unrealistically high FF values were not the ones with low ρ values (Fig. 4). Of the six trials with FF>3.0, three had both ρ_k>0.8 and three had one ρ_k>0.8.

The results of these Monte Carlo trials have several important implications for choosing appropriate abundance indices to use in CAL. One should not be surprised that different abundance indices or different combinations of indices result in very different estimates of FF, because even the correct abundance indices with realistic measurement errors result in a wide distribution of FF values (Fig. 4). The wide range of FF values corresponds to a wide range of q_k values (Fig. 5) and a wide range of abundance estimates (not shown). The relative abundance indices that result in the best estimates of FF and q_k do not necessarily have the highest partial correlation coefficients.

As explained in the theory section, any linear trend in abundance can be matched in VPA by choosing the appropriate FF and M values and one is just as likely to obtain a high correlation coefficient for an index with a spurious trend in abundance as one is for a correct index. Because all the partial correlation coefficients were significantly greater than zero, there are no grounds for rejecting any of the simulated abundance indices on this criterion. The consistently high probabilities of a positive correltaion (P>0.99) imply that all of the simulated indices resulted in better fits than any of the observed indices. One can arbitrarily reject the abundance indices resulting in extreme FF values but there is no justification for doing so, other than they don't fit one's preconceived idea of what FF should be.

Recommendations

The natural mortality rate (M) can not be reliably estimated from the bluefin tuna data because the joint confidence region for M and FF is unbounded (Fig. 2). The best that can be done is to try to estimate M independently and to estimate FF for a realistic range of M values. Aging errors appear to have a neutral effect on parameter estimation. Classifying the catch by age beyond age 10 adds virtually no useful information (Fig. 3) but errors in aging these fish may not bias the parameter estimates.

The most serious source of error in the bluefin tuna data is probably error in measuring CPUE and consequently the biggest problem in applying CAL is in selecting appropriate indices. I think that the criteria adopted by ICCAT to accept or reject indices are inappropriate for reasons previously mentioned. The result of using these criteria has been to reject the earlier indices (West 1, West 2, Medium 5B) and to accept the more recent indices (U.S. observer data and larval indices). The risk is that the recent declines observed in the Canadian hand-line data and larval indices may be over-emphasized relative to earlier CPUE data. Unless the earlier CPUE data are also included, estimates of abundance in the early 1970s are constrained only by later catches.

It has been argued that, because the VPA estimates of abundance become insensitive to the input FF value as one goes back in time, it is not necessary to use the earlier CPUE data in tuning the VPA. While it is true that VPA converges in the earlier years, it does not necessarily converge on the correct values. Therefore, if one wishes to estimate abundance in the initial years of the analysis (as well as FF), relative abundance indices for the early years should be included in addition to CPUE data from the most recent years.

I propose that instead of assigning the abundance indices weights of 0 or 1, all the indices be included and assigned relative weights from 0 to 1 (λ_k in Eq. 2). Any prior information that exists about the reliability of the various indices could be used to assign relative weights. Another possibility is to weight the indices of different fleets by the proportion of total catch caught by each fleet. I think the best approach would be to run CAL on each index individually and weight each index by its residual error variance (REV) calculated as the sum of squares (SS_k) divided by the degrees of freedom,

$$REV_k = SS_k/(n-2) \quad , \qquad (4)$$

where n is the number of years in the index and two parameters are estimated (q and FF). The relative weight assigned to each index would be the reciprocal of the residual error variance ($1/REV_k$) as is standard in weighted regression analysis (Steel and Torrie 1960).

This weighting procedure would assign relatively larger weights to abundance indices which produced a higher correlation coefficient when fitted individually. Dividing the partial sum of squares by the degrees of freedom corrects the tendency to obtain a higher correlation coefficient from indices with fewer observations. The weighting procedure would allow all the relative abundance indices to be included and would ensure that CAL would try equally hard to fit each index. CAL could easily be modified to weight the abundance indices; alternatively the weights could be multiplied directly to the data and the program left unchanged.

Alternative methods of parameter estimation

The most appropriate alternative to CAL for estimating the abundance of bluefin tuna is the program CAGEAN developed by Deriso et al. (1985). In CAGEAN it is assumed that fishing mortality is a separable product of age-specific selectivity ($\bar{P}_j$) and full-recruitment fishing mortality (F_i). The current procedure used by ICCAT to assess bluefin tuna involves estimating P_j in the terminal year by SVPA, estimating abundance by VPA and calibrating the VPA to abundance indices. CAGEAN performs these three steps in a single nonlinear estimation and derives confidence intervals for the parameter estimates by a bootstrapping technique.

CAGEAN is flexible about the amount of auxilliary information it uses. Effort data are required but missing years are allowed. A stock-recruitment relationship can optionally be included; it may do little to improve the parameter estimates but is useful for predicting future stock abundance. CAGEAN also allows multiple gear types, multiple age selectivity groups, and multiple catchability coefficients, options that are relevant to bluefin tuna. However, the inclusion of multiple groups requires more parameters to be estimated and there is a risk of over-parameterizing the problem.

CAGEAN is slightly less flexible than CAL in the types of auxiliary data it accepts. Because CAGEAN minimizes the differences between estimated and observed effort, the actual effort data are required. CAL minimizes the difference between observed and estimated relative abundance, and any index (numbers or weights) can be used. To use survey data in CAGEAN one must create "psuedo-effort" data by dividing commercial catch by standardized survey abundance.

Another disadvantage of CAGEAN is that it is user unfriendly. To run the program a number of voluminous data files must first be created. The output of the program includes a list of the logarithms of the parameter values but no tables of estimated numbers-at-age, etc. Nor are there any plots of observed and estimated effort, recruitment or age selectivities. Thus it is quite difficult to tell how well the model fits the data without doing a lot of extra work.

Trial run of CAGEAN on bluefin tuna data

To run CAGEAN using bluefin tuna data one would ideally need to partition the catch-at-age matrix by gear types and one would also need effort data for each gear type. Lacking these data, I ran CAGEAN on the aggregate catch-at-age matrix just to illustrate whether the program is appropriate for bluefin tuna. The distribution version of CAGEAN needed to be redimensioned to accommodate a 16 by 30 element catch-at-age matrix and corresponding effort data. I pooled the data for ages 29 and 30 so that the total number of observations (catch-at-age plus effort) would be less than the maximum of 500.

I derived psuedo-effort data (E_i) from the bluefin tuna data by dividing the commercial catch (C_i) by the standardized CPUE data:

$$E_i = C_i q_k / X_{ki} \quad , \qquad (5)$$

where C_i is summed over the ages appropriate to index k, and q_k is the catchability coefficient estimated by CAL. (This q_k is not the same parameter as the constant of proportionality between fishing mortality and effort, as estimated by CAGEAN and called q' in Fig. 7) This is obviously a circuitous and indefensible way of calculating a standardized index of effort but it suffices for illustrative purposes (Fig. 7). The effort data display an increasing trend in the 1970s and a sharp decrease in 1982 due to catch restrictions. In years with two effort values I took the geometric mean value to use in CAGEAN.

Without effort data CAGEAN predicts that fishing has had little effect on bluefin tuna and that abundance has been

monotonically increasing. The inclusion of effort data forces CAGEAN to realize that to maintain constant or decreasing catches in face of increasing effort, abundance must have decreased. The linear relationship between fishing effort (E_i) and fishing mortality (F_i) as estimated by CAGEAN is quite good except in the final years, 1983 to 1985 (triangles and squares in Fig. 7). The sharp decrease in effort between 1981 and 1982 provides good contrast to help define the model parameters. The CAL estimates of fishing mortality agree remarkably well with those of CAGEAN (plus signs and squares in Fig. 7) except in the initial years, 1970 to 1972, for which there are no effort data to constrain the estimates.

In fitting CAGEAN I assumed that partial recruitment to the fishery is constant for ages 16 to 30, similar to the assumption made in running CAL. Age selectivities could be estimated for all ages, but it is better to estimate as few parameters as possible. The age selectivities estimated by CAGEAN (Fig. 8) are maximal for ages 2 and 3, minimal for ages 7 to 12, and increase for older fish, which is consistent with the proportions of fish caught at age (Fig. 3).

The CAGEAN age selectivities apply to all years; the partial recruitment vector used to run CAL (Fig. 8) is different because it applies only to the terminal year (1985). In 1985, 5 and 6 year old fish constituted a much larger proportion of the catch than in previous years, presumably in response to the minimum size restrictions imposed in 1983. To check whether the age selectivities have changed in recent years, one could examine the residuals in catch-at-age, estimated by CAGEAN from 1983 to 1985. If these residuals indicate a shift in age selectivity, CAGEAN could be rerun with a separate selectivity group for the most recent years. (At least three years should be included in this group to produce good parameter estimates.) The departure in the relationship between fishing effort and fishing mortality (Fig. 7) from 1983 to 1985 could be due to changed age selectivity.

The recruitments estimated by CAGEAN agree in magnitude and trend with those estimated by CAL (Fig. 9). The CAL estimates indicate above average recruitments in 1971, 1973, and 1982 as expected from the proportional catch-at-age (Fig. 3). The CAGEAN estimates have the same trends but are less variable than the CAL estimates because the former are more tightly constrained by the separable fishing mortality assumption used in CAGEAN.

I tried including a stock-recruitment relationship in the CAGEAN fit by assuming the age of maturity is 10 years and after that fecundity is proportional to weight. This resulted in a fit similar to that without effort data: abundance appeared to increase irrespectively of fishing mortality. It is not surprising that the stock-recruitment realtionship does little to stabilize the parameter estimates because the data set is only slightly longer than the age of maturity. The stock-recruitment relationship might be better defined if the catch data from 1960 to 1970 were included in the analysis. It is somewhat surprising that the stock-recruitment relationship caused the CAGEAN fit to degenerate; perhaps the assumption of fecundity proportional to weight is grossly incorrect.

In conclusion, CAGEAN estimated recruitments and fishing mortalities that agree remarkably well with those of CAL. Of course the two fits are not completely independent because I used the CAL catchability coefficients to standardize the effort data, but this could be rectified by using the actual effort data for each gear type. Given the relatively low catches at ages 16 to 30, the error in aging these fish and the assumption of equal selectivity of these ages, it would make sense to pool ages 16 to 30 in future CAGEAN trials. The advantage of CAGEAN, as mentioned above, is that all the parameters are estimated in a single minimization. One must still select which groups of parameters to estimate, and the danger remains, with CAGEAN as with other methods, that with enough fiddling of age selectivity groups one can usually obtain some preconceived result. This caution aside, I strongly recommend that CAGEAN be considered as a useful alternative to CAL for estimating the abundance of Atlantic bluefin tuna.

References

Agger, P., I. Boetius and M. Lassen. 1973. Error in the virtual population analysis: the effect of uncertainties in the natural mortality coefficient. J. Cons. Int. Explor. Mer 35:93.

Anon 1986. Report of the meeting of the bluefin working group, Miami FL, USA, September 1985. Int. Comm. for the Conserv. of Atlantic Tunas, Col. Vol. Sci. Pap. XXIV:254p.

Anon 1987. Report of the meeting of the bluefin working group, 1986. Int. Comm. for the Conserv. of Atlantic Tunas, Col. Vol. Sci. Pap. In press.

Deriso, R.B., T.J. Quinn II and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Paloheimo, J.E. 1980. Estimation of mortality rates in fish populations. Trans. Am. Fish. Soc. 109:378-386.

Parrack, M.L. 1986. A method of analyzing catches and abundance indices from a fishery. Int. Comm. for the Conserv. of Atlantic Tunas, Col. Vol. Sci. Pap. XXIV:209-221.

Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. Res. Bull. Int. Comm. Northw. Atl. Fish. 9:65-74.

Sims, S.E. 1984. An analysis of the effect of errors in the natural mortality rate on stock-size estimates using virtual population analysis (cohort analysis). J. Cons. Int. Explor. Mer 41:149-153.

Steel, R.G.D. and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill, New York. 481p.

Ulltang, O. 1977. Sources of errors in and limitations of virtual population analysis (cohort analysis). J. Cons. Int. Explor. Mer 37:249-260.

Table 1. Aging error matrix (E) for smax=1.0. This example has only 20 ages so that it will fit on the page; the actual trials had 30 ages.

Estimated age (j)	True age (a) 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Σ_a
1	1.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.00
2	0.0	1.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.00
3	0.0	0.0	1.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.00
4	0.0	0.0	0.0	1.00	0.02	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.02
5	0.0	0.0	0.0	0.0	0.96	0.14	0.02	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.12
6	0.0	0.0	0.0	0.0	0.02	0.72	0.23	0.06	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.04
7	0.0	0.0	0.0	0.0	0.0	0.14	0.50	0.24	0.07	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.96
8	0.0	0.0	0.0	0.0	0.0	0.0	0.23	0.40	0.24	0.07	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.95
9	0.0	0.0	0.0	0.0	0.0	0.0	0.02	0.24	0.36	0.24	0.07	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.95
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.06	0.24	0.35	0.24	0.07	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.97
11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.07	0.24	0.35	0.24	0.07	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.98
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.35	0.24	0.07	0.01	0.0	0.0	0.0	0.0	0.0	1.00
13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.35	0.24	0.07	0.01	0.0	0.0	0.0	0.0	1.00
14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.35	0.24	0.07	0.01	0.0	0.0	0.0	1.00
15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.35	0.24	0.07	0.01	0.0	0.0	1.00
16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.35	0.24	0.07	0.01	0.0	1.00
17	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.35	0.24	0.07	0.01	1.00
18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.36	0.24	0.07	1.00
19	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.44	0.24	1.00
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.07	0.24	0.68	1.00
Σ_j	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	20.0

Table 2. Effect of aging errors on parameter estimation by CAL. M was set at 0.1 in these trials.

Maximum standard deviation of aging errors	Parameter			
	Full F	q_1	q_2	q_3
True value	0.46	1.00	1.00	1.00
0.0	0.45	0.98	1.00	1.00
1.0	0.46	0.99	1.00	1.00
2.0	0.48	0.99	1.02	1.00
3.0	0.50	1.00	1.03	0.99
5.0	0.52	1.02	1.07	0.94

Table 3. Abundance of west Atlantic bluefin tuna as estmated by CAGEAN. (Numerical abundance in thousands and biomass in metric tons. NE means not estimable.)

YEAR	TOTAL NUMERICAL ABUNDANCE	TOTAL BIOMASS	TOTAL CATCH BIOMASS	TOTAL EXP. BIOMASS	ANNUAL SURPLUS PROD.
1970	795.	39670.	5148.	19187.	5937.
1971	920.	40459.	5996.	20826.	6184.
1972	936.	40646.	4516.	22555.	5796.
1973	985.	41927.	4858.	25522.	6823.
1974	978.	43892.	5558.	28579.	5545.
1975	921.	43879.	4571.	28961.	5202.
1976	836.	44510.	5610.	29475.	5840.
1977	747.	44740.	5869.	28745.	4930.
1978	671.	43800.	5367.	26700.	3647.
1979	578.	42080.	5470.	24056.	3521.
1980	515.	40131.	5930.	21739.	2598.
1981	418.	36799.	5489.	18603.	1072.
1982	366.	32383.	1370.	15385.	1600.
1983	326.	32613.	2400.	16357.	1319.
1984	286.	31533.	2047.	16879.	745.
1985	259.	30231.	2428.	17638.	NE

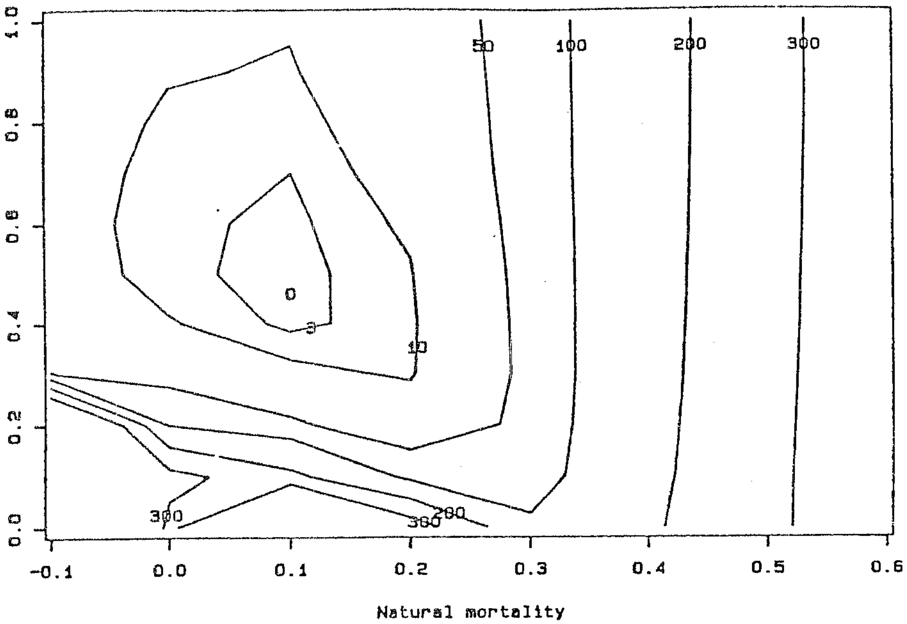

Fig. 1 Sum of squares surface produced by CAL using correct data. The "0" indicates the true values, M=0.1 and FF=0.46. The isocline labelled "3" is the 95% confidence region.

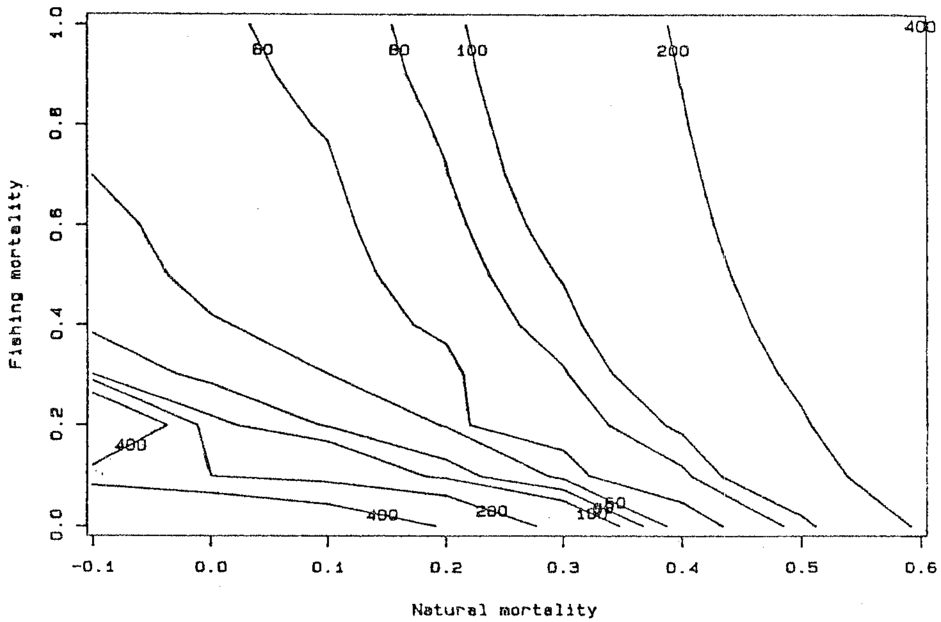

Fig. 2 Sum of squares surface produced by CAL using observed data. The minimum occurred at FF=0. The numbers on the isoclines are not directly comparable to those in Fig. 1 because CAL scales the sums of squares such that min=0 and max=999.

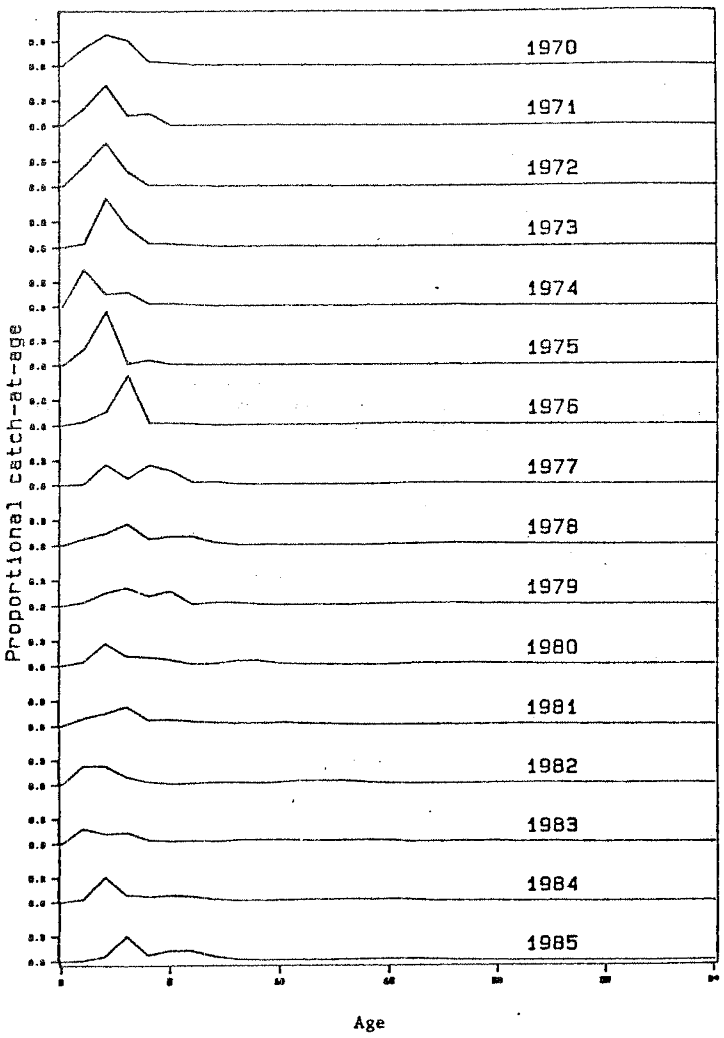

Fig. 3 Age composition of the catch as proportions of the total catch each year.

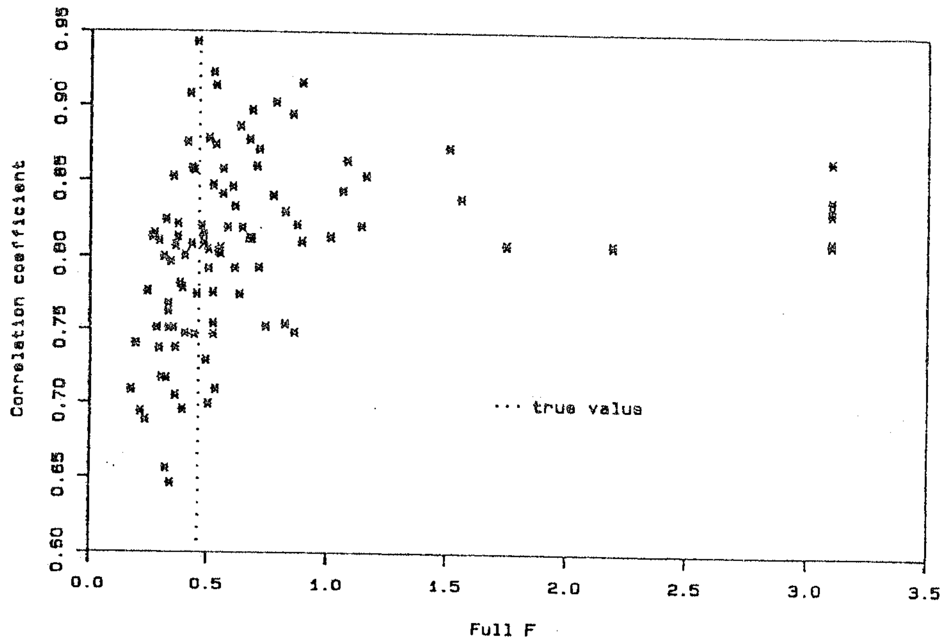

Fig. 4 Scatterplot of full F and total correlation coefficients estimated in 100 Monte Carlo runs of CAL. The true FF value is 0.46. See text for explanation of how the data were generated.

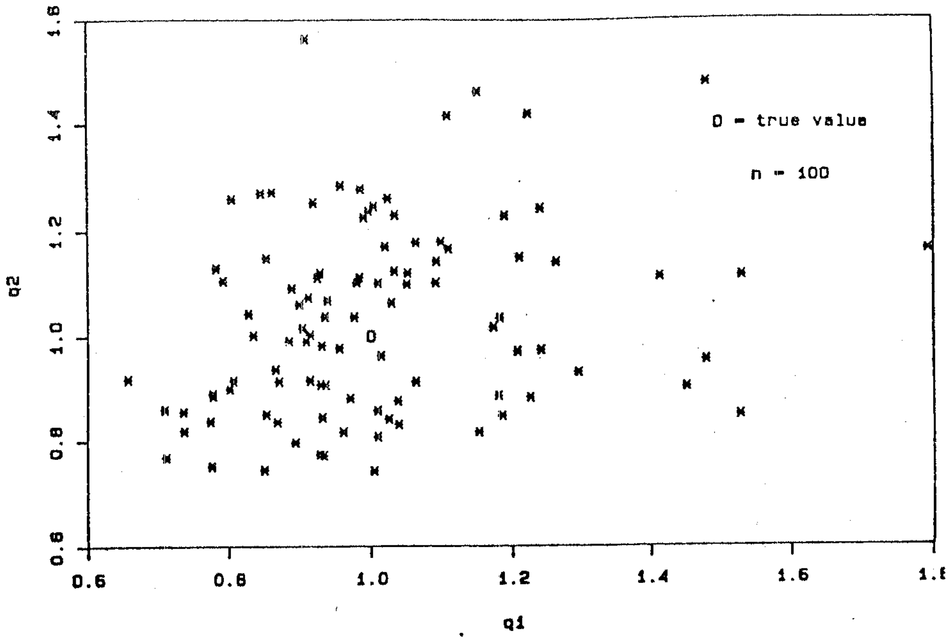

Fig. 5 Scatterplot of catchability coefficients estimated in Monte Carlo trials of CAL. The first abundance index is for ages 6 to 9 and the second index is for ages 10 to 30.

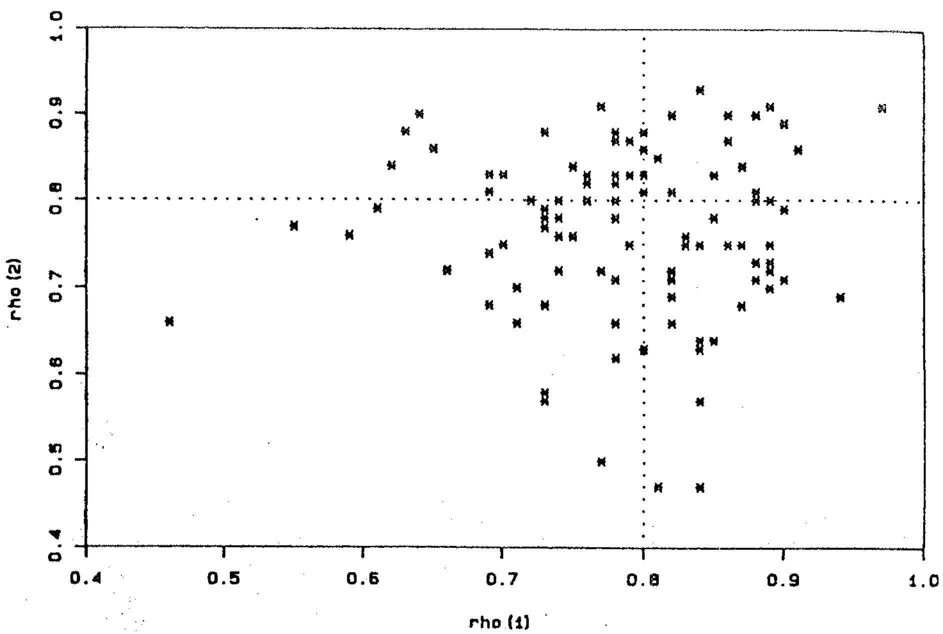

Fig. 6 Partial correlation of the two abundance indices to the numbers-at-age estimated by CAL. The dotted lines at 0.8 represent the criterion for accepting or rejecting abundance indices.

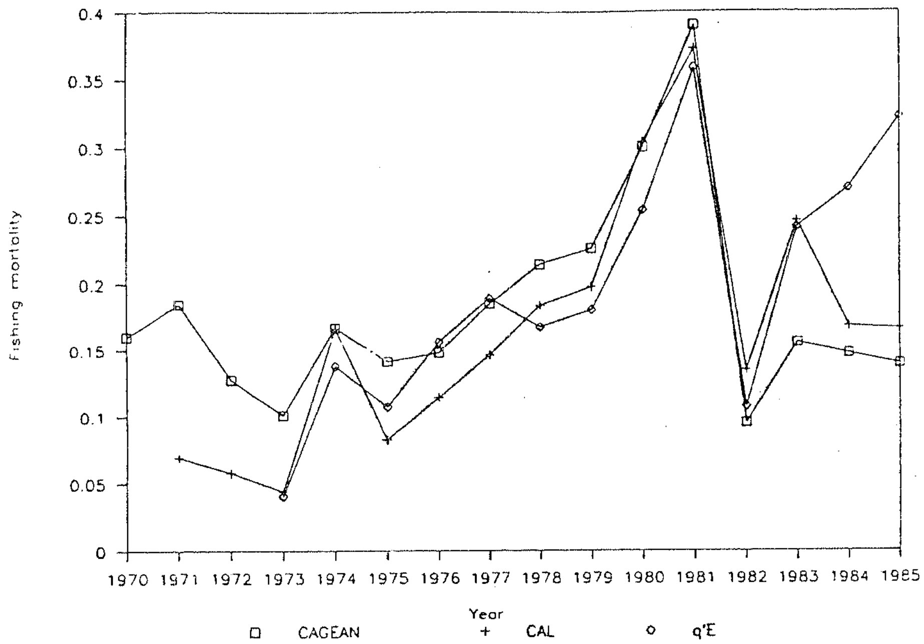

Fig. 7 Fishing mortality for ages 16 to 30 as estimated by CAGEAN (squares) and CAL (plus signs). The CAL estimates were multiplied by 0.6 to scale them to the CAGEAN estimates. The effort data (triangles), which were used to fit CAGEAN, are multiplied by the CAGEAN catchability coefficient (q') not the CAL catchability coefficient.

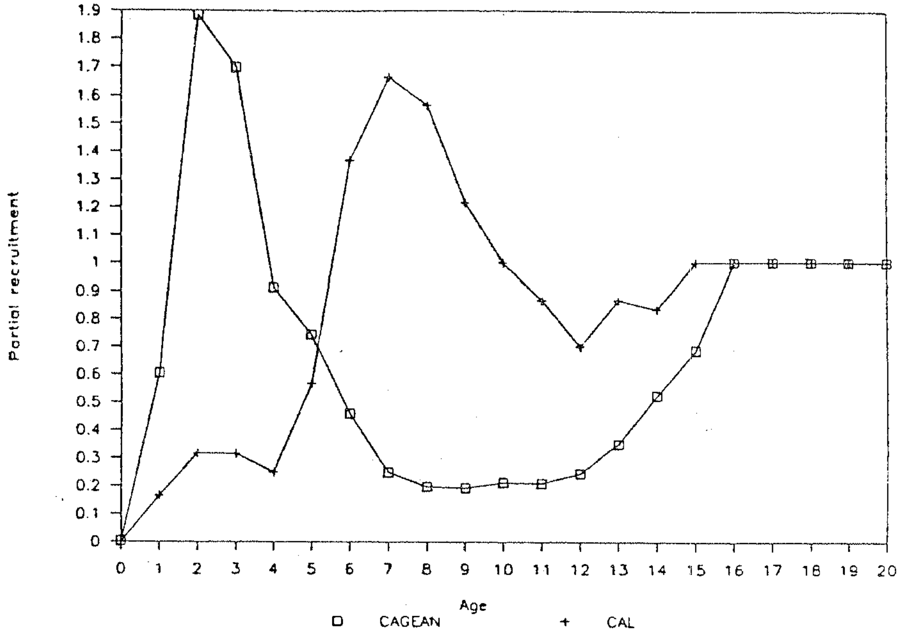

Fig. 8 Partial recruitment vectors estimated by CAGEAN (squ[illegible]) and assumed in the terminal year for CAL (plus signs). The CAL va[illegible]es were divided by 0.6 to scale them to the CAGEAN estimates. The selectivities for ages 20 to 30 (not shown) are all 1.0.

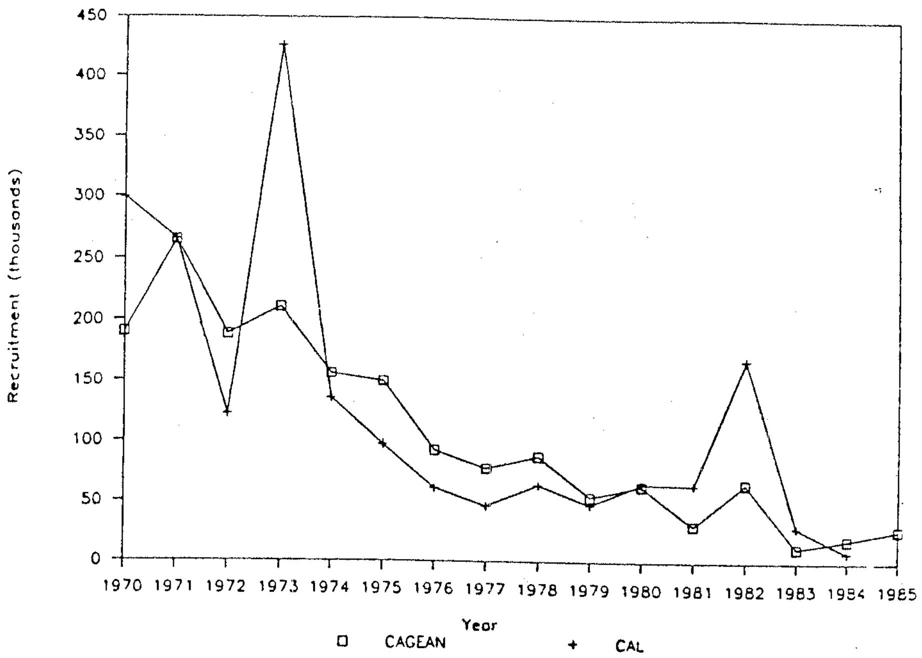

Fig. 9 Number of age zero bluefin tuna as estimated by CAGEAN (squares) and CAL (plus signs).