International Council for the Exploration of the Sea

ICES C.M. 1991/D:31
Statistics Cttee, Ref H

Using Simulation to Quantify Uncertainty in Sequential Population Analysis (SPA) and Derived Statistics, with Application to the North Atlantic Swordfish Fishery

Victor R. Restrepo[1], Joseph E. Powers[2], Stephen C. Turner[2], and John M. Hoenig[3]

[1] University of Miami, Rosenstiel School of Marine and Atmospheric Science, 4600 Rickenbacker Causeway, Miami, FL 33149, USA

[2] National Marine Fisheries Service, Southeast Fisheries Center, 75 Virginia Beach Drive, Miami, FL 33149, USA

[3] Department of Fisheries and Oceans, Science Branch, P.O. Box 5667, St. John's, NF A1C 5X1, Canada

Abstract. - Monte Carlo simulation is used to quantify the uncertainty in the results of sequential population analysis and in related derived statistics. Probability density functions describe the measured or perceived uncertainty in the inputs to the assessment model. Pseudo-data sets are then generated repeatedly from these distributions and used as inputs to the assessment model to examine the variability in the resulting parameter estimates. The approach quantifies uncertainty in the assessment of the swordfish stock in the North Atlantic Ocean, particularly in the estimates of population size, current fishing mortality, values of various reference fishing mortalities (e.g., $F_{0.1}$), and catch regulations necessary to achieve various objectives.

Fishery managers have long recognized the dangers of accepting parameter estimates without consideration of the inherent variability in the estimates of fish stock status and related parameters. Early strategies for dealing with this were quite simple, such as backing off from the estimate of the fishing mortality giving the maximum yield (F_{max}) to a more conservative value. Sensitivity analyses, in which the effects of various perturbations of the inputs are observed, have commonly been employed to obtain impressions of the probable bounds on the errors (e.g., Pope 1972; Pope and Garrod 1975). Recently, various authors have used the delta or Taylor's series method (see Kendall and Stuart 1977 pp. 246-248) to obtain analytical expressions or numerical solutions for the variances and covariances of outputs from simple sequential population analyses (SPAs) (Saila et al. 1985; Sampson 1987; Prager and MacCall 1988; Kimura 1989). These solutions tend to be complex, only asymptotically valid, and highly model specific. They have only been worked out for the simplest of the sequential population analysis models and some simple quota-setting procedures (e.g., Pope 1983).

It is possible to measure how well the fitted population estimates correspond to trends in indices of abundance when calibration procedures are applied to sequential population analyses. Estimates of standard errors of population obtained in this way do not reflect all of the variability in the inputs. For example, similar trends in population abundance may be obtained when two very different values of natural mortality rate are assumed as inputs; despite the similar trends (and hence correlation with the abundance indices), there may be large differences in the absolute estimates of abundance.

The term uncertainty will be used broadly here to refer to any variability or error that arises from the stock assessment process. Uncertainty can enter into an assessment in various ways. There may be uncertainties in the values of the inputs, e.g., the total catch may be estimated with error. Also, the formulation of the assessment model may be subject to uncertainty, and the analyst may make data-dependent decisions during the analysis which are subject to error. The degree to which these sources of error are incorporated into the analyses will determine the perceived uncertainty in the overall assessment results. If all sources of error are not accounted for properly, then estimates of the uncertainty in the assessment results may be too small.

Monte Carlo simulation is a convenient tool for studying a model's outputs given different types and levels of error in the model's inputs (e.g., Restrepo and Fox 1988). In a sensitivity analysis framework, Pope and Gray (1983), and Rivard (1983), used a Monte Carlo approach to study the relative contribution of various inputs to the overall uncertainty in total allowable catch (TAC) estimates obtained from calibrated SPAs. In this paper, we present a generalized method, also based on Monte Carlo simulation, for accounting for the uncertainty in assessment results, including the parameters directly estimated from the SPAs as well as derived statistics used to set management targets and allowable catches. We then illustrate how the method can be applied with an example for swordfish (*Xiphias gladius*) in the North Atlantic Ocean.

Quantifying Uncertainty by Simulation

Suppose the only uncertainty in the inputs to an assessment model concerned the value of the instantaneous natural mortality rate, M, and that it was felt that M could be anywhere is the interval from 0.15 to 0.25 yr^{-1} with equal likelihood. One could compute the assessment model results for a large number of uniformly spaced values of M in this interval (say, 100) and make histograms of the results. This, then, would represent the feelings about the relative likelihood of the estimated output taking on various values. If not all values of M were believed to be equally likely, then one could weight the 100 outputs by the probability associated with the corresponding inputs.

The above procedure becomes awkward when there are a number of inputs subject to uncertainty because the number of combinations of input parameter values becomes very large. An alternative is to use a Monte Carlo approach in which values of the inputs are drawn randomly from uncertainty distributions. A sufficiently large number of plausible input data sets are thus generated and used to compute the assessment model results such that the distributions of the estimated outputs are clearly defined. This may involve several hundred or several thousand runs, depending on the types of data and models used (in our work we found that 500 to 1000 data sets were necessary to obtain stable results).

A typical assessment of a fish stock using sequential population analysis involves three levels of analysis: First, data are prepared for the SPA. This usually involves estimating and ageing the annual catch, and computing indices of abundance for calibration. Second, the SPA itself is carried out (it is also frequently termed "VPA", for Virtual Population Analysis). In many cases several SPAs are carried out in order to examine the goodness of fits of the input data to alternative model formulations or simply to examine the sensitivity of the results to the alternative formulations. Third, derived statistics are computed. These are commonly so-called reference points (F_{max}, $F_{0.1}$), and forward projections of stock status and catches under alternative management actions.

It is intuitively simple to see how the Monte Carlo approach can be used to characterize the uncertainty in the entire analysis process, starting with the raw data collected for the first step in the above procedure. For instance, the total annual catches and their proportions at age can be obtained by resampling the original data that lead to the catch estimates, through a non-parametric bootstrap (Efron 1982). These bootstrapped catches would then be used to calibrate the SPAs, whose results, in turn, would affect the values of projected future catches.

In practice, however, the time and computer resources required to carry out such a large-scale simulation makes it more practical to derive the input uncertainty distributions from parametric statistical analyses of data (this would involve assuming a distribution type for the inputs and estimating their mean and variance). Often, the distributions for some of the inputs will not be based on a rigorous statistical treatment of the data, but rather will represent personal feelings about the likelihood of the inputs taking on particular values (this is probably more true of the natural mortality rate, M, which is usually assumed and not estimated). The outputs would then represent the analyst's personal uncertainty in the assessment results.

The above approach can be generalized to allow for uncertainty in the sequential population analysis model formulation as well. Suppose one feels that there is a 70% chance that the fishing mortality rate in the last year does not decline with age after a fully recruited age (this is often known as a "flat-topped" partial recruitment curve), and a 30% chance that it does ("dome-shaped" partial recruitment). Then one could conduct 70% of the simulations with an SPA that assumes the flat-topped curve and 30% with the dome-shaped curve. The resulting combination of outputs would reflect the feelings of uncertainty about the SPA model formulation. Similarly, the approach can also account for uncertainty concerning data-dependent decision making. For example, if several abundance indices are available, one might subject each index to a preliminary test to decide whether the index is acceptable for calibrating the sequential population analysis, e.g., via analysis of residuals. One can repeat this decision making process for each of the simulated data sets and thus account for the uncertainty associated with screening indices.

Application to Swordfish in the North Atlantic Ocean

Assessment Procedure. Swordfish in the North Atlantic Ocean are assessed by the International Commission for the Conservation of Atlantic Tunas (ICCAT). The assessment procedure is continually changing as experience is gained. The procedure below was used for the 1989 assessment (ICCAT 1990).

Nine age groups were recognized in the commercial catch, ages 1 to 9+ (where 9+ means age 9 and above). There were 11 years of catch-at-age data from 1978 to 1988. Fleets from the United States, Japan, and Spain account for most of the catch. Eleven abundance indices were available based on fleet-specific catch rates from the longline fisheries (ICCAT 1990).

Details of this assessment of the stock are presented in ICCAT (1990). Briefly, the procedure used was as follows: (1) A separable virtual population analysis, SVPA, (Pope and Shepherd 1982) was computed in order to obtain estimates of the age-effects or partial recruitment. Data from 1983 to 1988 were used for this under the assumption that the selectivity pattern remained stable during that period. For that analysis the terminal fishing mortality was 0.2 yr^{-1} and selectivity for the oldest age group was 3.0. (2) Gavaris' (1988) approach to sequential population analysis (ADAPT) was then used for calibration, with each abundance index used separately. A weighting factor for each index was obtained by setting the weight for the ith index equal to the reciprocal of the mean squared error after calibrating with the index. In performing the calibrations, ages 5 and above were assumed to be fully recruited ($S_a = 1.0$ for a = 5, 6, ..., 9+) and the partial recruitment for the other ages was as determined from the SVPA (i.e., from step 1). (3) The set of weights computed for the abundance indices were then rescaled so that they summed to unity. (4) The weights were then used in recalibrating the ADAPT SPA using all of the abundance indices at once. In doing so, the following constraints were used: S_1 was taken from the separable virtual population analysis, S_2 through S_5 were directly estimated through calibration, and S_6 through S_{9+} were set equal to the estimated S_5. The objective function used in the calibration is to minimize the weighted sum of the squared deviations from the predicted abundance indices, RSS:

$$RSS = \min \sum_i \sum_y W_i (I_{iy} - \hat{I}_{iy})^2,$$

where the subscripts i and y refer to the index and year, respectively, the W_i are the weighting factors from step (2) above, and I_{iy} and $\hat{I}_{iy}$ are the observed and predicted values of the indices, respectively.

After the fishing mortalities and population sizes were computed by the sequential population analysis, the values of F_{max} and $F_{0.1}$ (Gulland and Boerema 1973) were calculated from yield per recruit computations. Data from the terminal year (i.e., the most recent year available) were used to project the catch in the current year and then project the catch for the next year. For this, recruitment in the current year and the following year were assumed by ICCAT to be equal to the long term mean recruitment obtained in the sequential population analysis. The projections were made for a variety of fishing mortalities, specifically $F_{0.1}$, F_{max}, and $F_{status\ quo}$. The value of the spawning potential ratio was also computed. This biological reference point is defined as the current spawning biomass per recruit divided by the potential spawning biomass per recruit in the absence of fishing. It has been suggested that many of the fish stocks which have collapsed have had spawning potential ratios less than 0.2 (see Brown 1990 and Goodyear 1990 for a discussion).

Specification of Uncertainty in the Inputs. One thousand simulated data sets were analyzed using a version of ADAPT written in FORTRAN 77 (available from the authors). The formulation of the problem was made to mimic the 1989 ICCAT assessment for North Atlantic swordfish. However, we emphasize that the uncertainties in the inputs specified below are our *ad hoc* choices and, although realistic, are intended mainly for illustrative purposes.

Natural mortality. Uncertainty in the natural mortality rate (M) was specified as a uniformly distributed random variable in the interval from 0.1 to 0.3 per year. The value of 0.2 used by ICCAT (1990) is at the center of this range and the choice of a uniform distribution places equal confidence in all values in the interval.

Catch-at-age. Total annual catches were assumed to be lognormally distributed with a coefficient of variation of 10% and expected value equal to those in the assessment. A coefficient of variation of 10% indicates that the catches are known with high precision.

The proportions of the total catch in any year that make up each age component were assumed to follow a multinomial distribution with expected values equal to the observed proportions and sample size equal to 1% of the annual catch. This model for the uncertainty was purely heuristic rather than based on measured variances.

Abundance (CPUE) indices. The eleven available indices from the longline fisheries were also assumed to be lognormally distributed with a coefficient of variation of 10%. We chose a value of 10% as a rough approximation for all indices in all years. However, there is no reason why each index could not have a different coefficient of variation for each year depending on the amount of data available.

Results and Discussion

The simulations gave rise to 1000 sets of age- and year-specific fishing mortality rates and population sizes. We computed the coefficient of variation of these sets of estimates for each age-year combination (Figures 1a and 1b). As expected, the coefficients of variation were highest in the most recent (terminal) year, 1988. Also, the age groups which form the bulk of the catch (ages 3 - 5) were the best determined. It is interesting to note that the coefficients of variation of fishing mortality rates for ages 8 and 9 were consistently lower than those for preceding ages. This is due to the manner in which the estimates for ages 8 and 9 were determined: it was assumed that $F_{8y} = F_{9y}$ (subscripts refer to age and year respectively), and these were computed as a pooled average of fishing mortalities for ages 5 to 7. Thus, the uncertainty in the estimates of fishing mortality for the last two age groups is solely a function of the uncertainties in the estimates for ages 5 to 7. This underscores the fact that the simulation results are conditional not only on the input uncertainty distributions but on the formulation of the model being fitted as well.

The median recruitment (age 1) from the simulations increased over time (Figure 2). However, the 95% confidence bands, defined by the 2.5th and 97.5th percentiles of the 1000 estimates, are quite wide. The confidence bands provided by the Taylor series approximation for a single run with the actual data are much narrower than the ones obtained by the Monte Carlo approach. The former confidence bands indicate there is no uncertainty in the results for the converged part of the sequential population analysis in contrast to the simulation results. This is because the Taylor series results are conditional on the natural mortality rate, catch at age, etc. being known exactly whereas the simulation accounts for uncertainty in these inputs. For this reason, we believe the simulation results are more realistic.

Note that there appears to be very little interannual recruitment variability in the time series (Figure 2). This is probably due in part to the fact that fish ages were estimated from lengths by inverting the Gompertz growth equation and this tends to blur the age groups.

The population of fish age 5 and above appears to have declined rather steadily over time while the weighted fishing mortality rate appears to have increased (medians, Figures 3a and 3b). Here, weighted fishing mortality is defined to be the mean of the fishing mortality estimates for ages 5 through 9+ computed with weights proportional to the estimated population size at age. Again, the confidence bands are very wide.

It should be noted that for each run the estimates of fishing mortality, F_{ay}, and population size, N_{ay}, are highly correlated not only with each other but also with the value of natural mortality, M, used in the simulation run. For this reason, it is appropriate to examine trends in an estimated quantity one run at a time. We computed the ratio of the weighted fishing mortality in a given year y to the weighted F in the base year (taken to be 1978 in this example) for each simulation run (Figure 4). The distribution of the fishing mortality ratio in 1979 was centered around 1.0; the ratio in 1986, 1987 and 1988 was greater than 1.0 in 100% of the runs thus clearly indicating that fishing mortality has increased. This result is not obvious from examination of Figure 3b and illustrates how the Monte Carlo approach lends itself to hypothesis testing very easily.

Of course, the goals of an assessment are not restricted to estimating population sizes and mortality rates. Interest is often centered on catch projections and quotas, effort regulations, and risk analyses. For swordfish assessments, it is useful to contrast the estimated current level of fishing mortality against reference points such as $F_{0.1}$ and F_{max}. The uncertainty in such comparisons (e.g., the ratio of current F to $F_{0.1}$) can easily be quantified using the Monte Carlo procedure.

For each simulation run, we computed the multiplier that would be necessary to bring the estimated vector of age-specific fishing mortalities in the terminal year to the $F_{0.1}$ and F_{max} levels (Figure 5). For the computations, we used the run-specific natural mortality rate and the weight at age relationships used by ICCAT in the 1989 assessment. No uncertainty was specified for weight relationships although this could easily be added if appropriate information were available. From Figure 5, it is evident that , to achieve the $F_{0.1}$ goal, fishing mortality must be cut to about 25% of its current value (whatever that may be). With respect to F_{max}, it appears that fishing mortality must be cut by around 50% (Figure 5). Note, however, that this conclusion is considerably less certain than that for $F_{0.1}$ as evidenced by the fact that the distribution of multipliers is more spread out for F_{max} than it is for $F_{0.1}$.

We also computed 1000 projected catches in weight for 1989 with fishing mortality equal to that in 1988. We then projected the catch for 1990 with fishing mortality set at the midpoint between the fishing mortality in 1988 and $F_{0.1}$ (Figure 6). This is a method for gradually reducing fishing mortality to minimize the short-term impact that decreased landings have on fishermen (see Pelletier and Laurec 1990 for a discussion). Recruitments for 1989 and 1990 were drawn randomly from the empirical distribution of recruitments estimated from 1978 through 1987 on each iteration. If the fishing mortality does not change in 1989 from the level in 1988, catches are likely to be somewhere around the 1988 yield of approximately 18,000 mt. The 1990 yields are likely to be around 11,000 to 13,000 mt.

Using the Monte Carlo results, it is equally simple to obtain distributions of catches for fishing at other exploitation levels or to obtain distributions of fishing mortalities for fixed catch quotas. Similarly, the distribution of other projected variables, such as the spawning potential ratio that results from various catch and fishing mortality options, can be computed. In doing so, it is important to have the values of the inputs used in calibrating the SPAs (e.g. natural mortality) stored in each iteration, so that the projection computations use the same values.

Conclusions

Monte Carlo simulation has long been regarded as a very useful quantitative tool, especially for sensitivity analysis (e.g., Pope and Gray 1983 and Rivard 1983). It is also quite useful for studying the properties of specific assessment procedures (e.g., Kimura 1989; Mohn 1983). We believe that the Monte Carlo simulation approach we present is not only a versatile and intuitive method to quantify uncertainty in assessment results, but in many cases it may also be the only practical way to incorporate some types of input uncertainty which are not estimated statistically. Because the estimated uncertainties in the model outputs are conditional on what is known and what is assumed about the inputs, failure to acknowledge possible sources of uncertainty in a realistic manner may lead to overly optimistic views of the uncertainties in the model outputs. The Monte Carlo approach forces one to examine the nature and magnitudes of the uncertainties in the inputs and in the model formulation, and it allows one to study how uncertainties are propagated through the assessment and into the projections ultimately used for management recommendations.

Perhaps, one of the most encouraging aspects of our experience with application of the Monte Carlo method is the high precision with which we appear to be estimating <u>relative</u> statistics. Whereas, stock sizes are estimated with a large degree of error, current stock size relative to a previous level is not. Similarly, the effort required to generate a desired level of catch, relative to the terminal year's effort, is estimated with quite useful precision. Indeed, it is these kinds of relative statistics that are most useful for management.

Acknowledgments

Partial support for this study was provided through the Cooperative Institute for Marine and Atmospheric Studies by National Oceanic and Atmospheric Administration Cooperative Agreement NA90-RAH-0075. Additional support was provided through the Canada Atlantic Fisheries Adjustment Program (Northern Cod Science Program). We thank Nicholas Payton for programming assistance and James Baird for helpful suggestions.

Citations

Brown, B.
1990. Use of spawning stock size considerations in providing fishery management advice in the North Atlantic - a brief review. ICCAT Coll. Vol. Sci. Pap. 32(2):498-506. (SCRS/89/102)

Efron, B.
1982. The jackknife, the bootstrap and other resampling plans. Soc. Ind. Appl. Math, Philadelphia.

Gavaris, S.
1988. An adaptive framework for the estimation of population size. CAFSAC Res. Doc. 88/29.

Goodyear, C.P.
1990. Spawning stock biomass per recruit: the biological basis for a fisheries management tool. ICCAT Coll. Vol. Sci. Pap. 32(2):487-497. (SCRS/89/82)

Gulland, J. A. and L.K. Boerema.
1973. Scientific advice on catch levels. Fish. Bull., U.S. 71:325-335.

ICCAT (International Commission for the Conservation of Atlantic Tunas).
1990. Appendix 10 to Annex 8 *in* Report for Biennial Period, 1988-89, Part II (1989), English version, Madrid, Spain.

Kendall, M. and A. Stuart.
1977. The advanced theory of statistics, Volume 1, 4th ed. Macmillan, New York.

Kimura, D. K.
1989. Variability, tuning, and simulation for the Doubleday-Deriso catch-at-age model. Can. J. Fish. Aquat. Sci. 46:941-949.

Mohn, R.K.
1983. Effects of error in catch and effort data on tuning cohort analysis, with a postcript on logistic production models. Can. Spec. Publ. Fish. Aquat. Sci. No. 66. p 141-150.

Pelletier, D. and A. Laurec.
1990. Toward more efficient adaptive TAC policies with error-prone data. NAFO SCR. Doc. 90/100. (ser. no. N1835)

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

Pope, J. G.
1983. Analogies to the status quo TACs: Their nature and variance. Can. Spec. Publ. Fish. Aquat. Sci. No. 66. p 99-113.

Pope, J. G., and D. J. Garrod.
1975. Sources of error in catch and effort quota regulations with particular reference to variations in the catchability coefficient. Int. Commn. Northwest Atl. Fish. Res. Bull. 11: 17-30.

Pope, J. G, and D. Gray.
1983. An investigation of the relationship between the precision of assessment data and the precision of total allowable catches. Can. Spec. Publ. Fish. Aquat. Sci. No. 66. p 151-157.

Pope, J. G., and J. G. Shepherd.
1982 . A simple method for the consistent interpretation of catch-at-age data. J. Cons. int. Explor. Mer 40:176-184.

Prager, M. H. and A. MacCall.
1988. Sensitivities and variances of virtual population analysis as applied to the mackerel, *Scomber japonicus*. Can. J. Fish. Aquat. Sci. 45:539-547.

Restrepo, V.R. and W.W. Fox, Jr.
1988. Parameter uncertainty and simple yield-per-recruit analysis. Trans. Amer. Fish. Soc. 117:282-289.

Rivard, D.
1983. Effects of systematic, analytical, and sampling errors on catch estimates: A sensitivity analysis. Can. Spec. Publ. Fish. Aquat. Sci. No. 66. p 114-129.

Saila, S. B., E. Lorda, and H. A. Walker.
1985. The analysis of parameter error propagation in simple fishery models. Mar. Res. Econ. 1(3):235-246.

Sampson, D. B.
1987. Variance estimators for virtual population analysis. J. Cons. int. Explor. Mer 43:149-158.

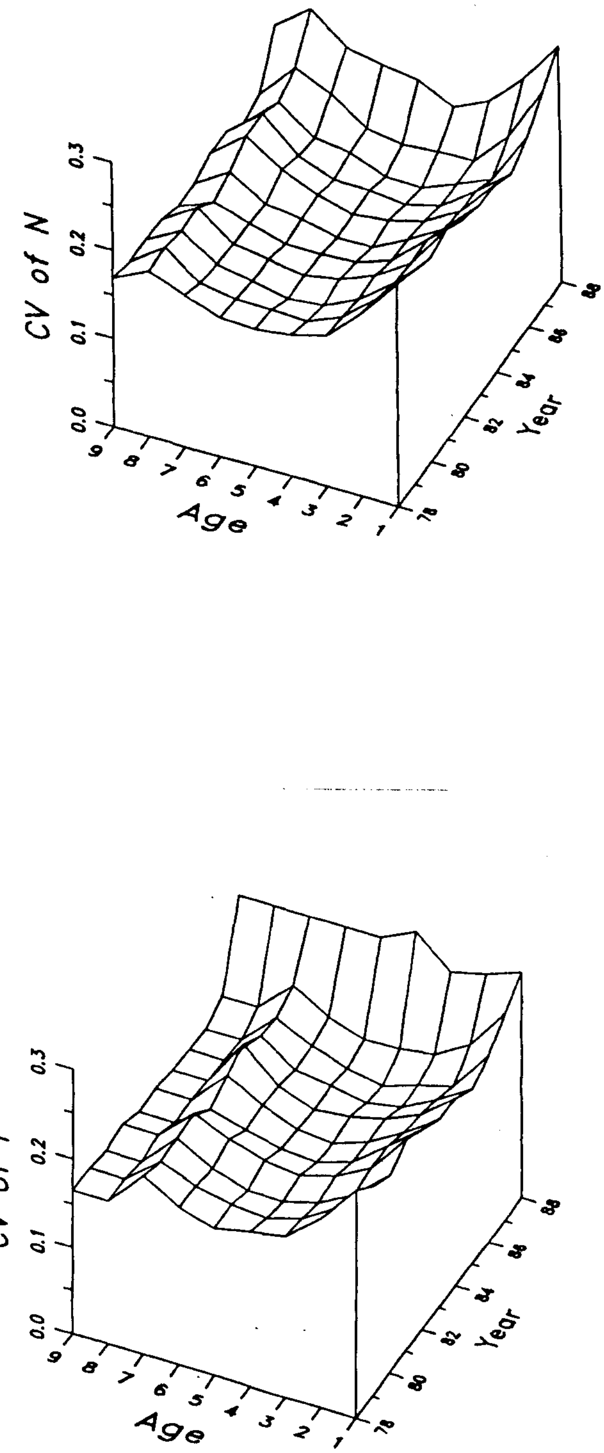

Figure 1. Coefficients of variation of outputs from the sequential population analyses of simulated swordfish data sets. a) age- and year-specific estimates of population numbers; b) age- and year-specific estimates of instantaneous fishing mortality.

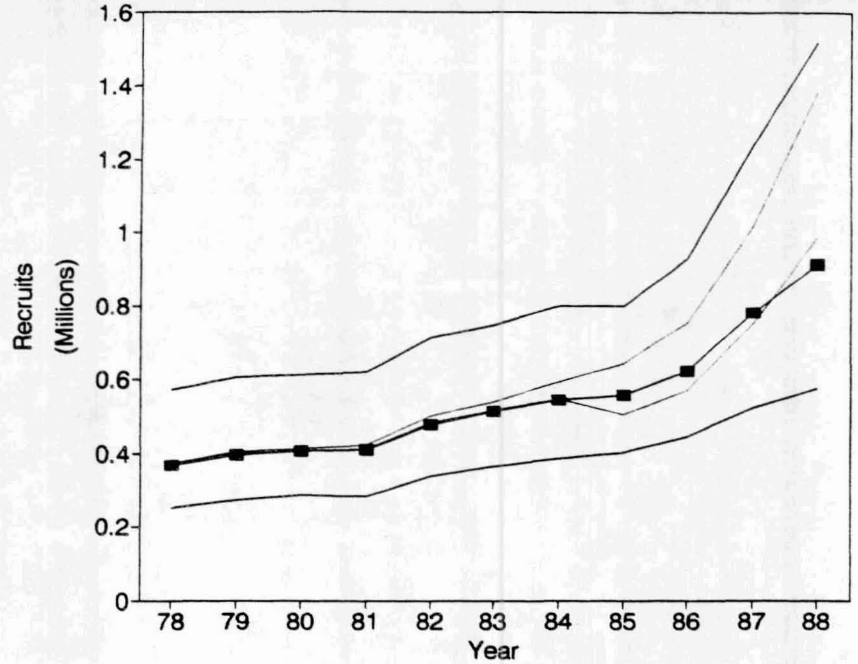

Figure 2. Distribution of recruitment estimates, by year, from the sequential population analyses. Outer lines are from the Monte Carlo simulations and show 95% confidence bands (determined as the 2.5th and 97.5th percentiles of the distribution resulting from 1000 simulations). Inner pair of lines shows the confidence bands obtained from a single run of the ADAPT program using the actual data. Line with symbols gives the median estimate for each year from the simulations.

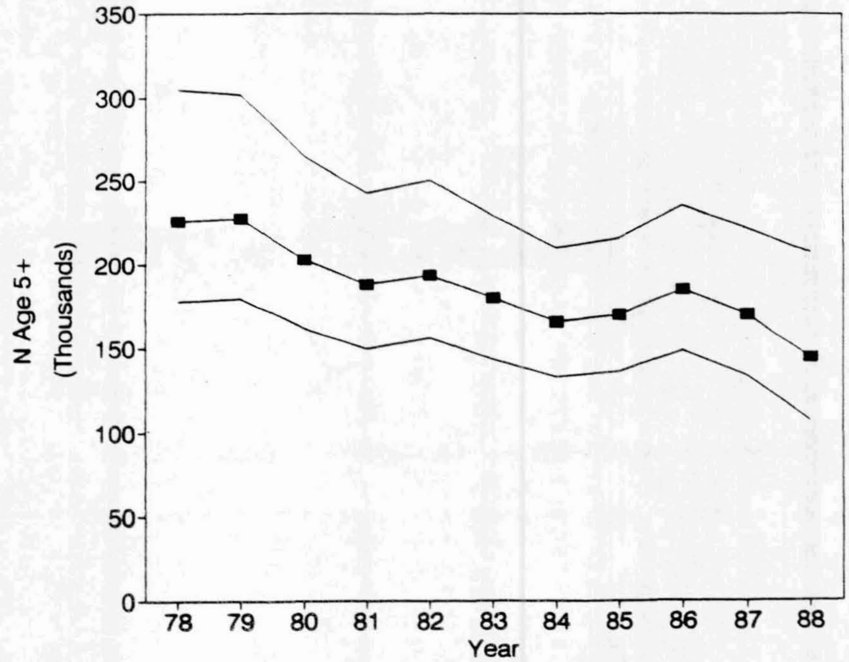

Figure 3. Medians, 2.5th percentiles and 97.5th percentiles of the output distributions from the Monte Carlo simulations. a) distribution of the estimates of the size of the population of fish aged 5 and above

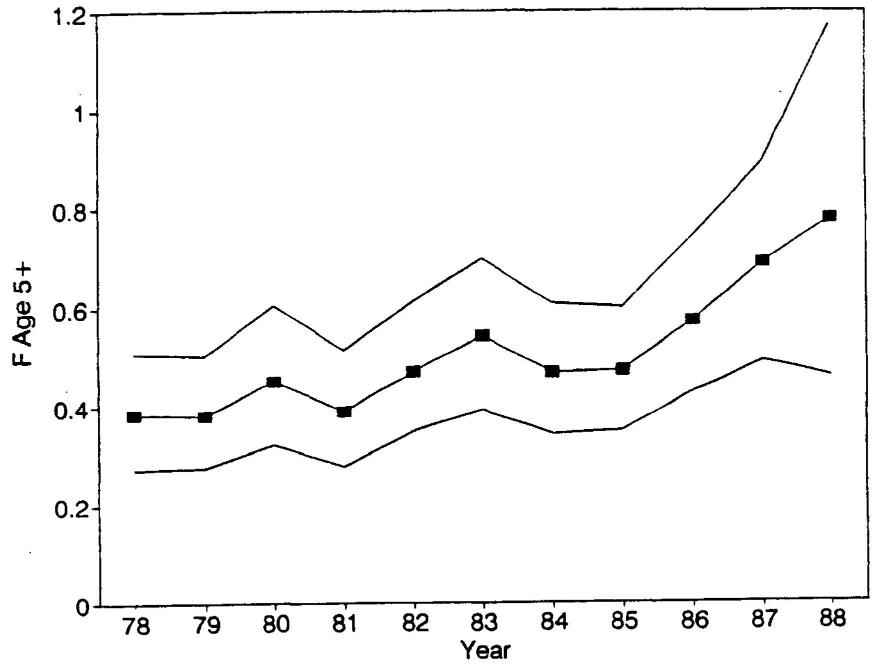

Figure 3. Medians, 2.5th percentiles and 97.5th percentiles of the output distributions from the Monte Carlo simulations. b) distribution of the estimates of the fishing mortality for fish aged 5 and above.

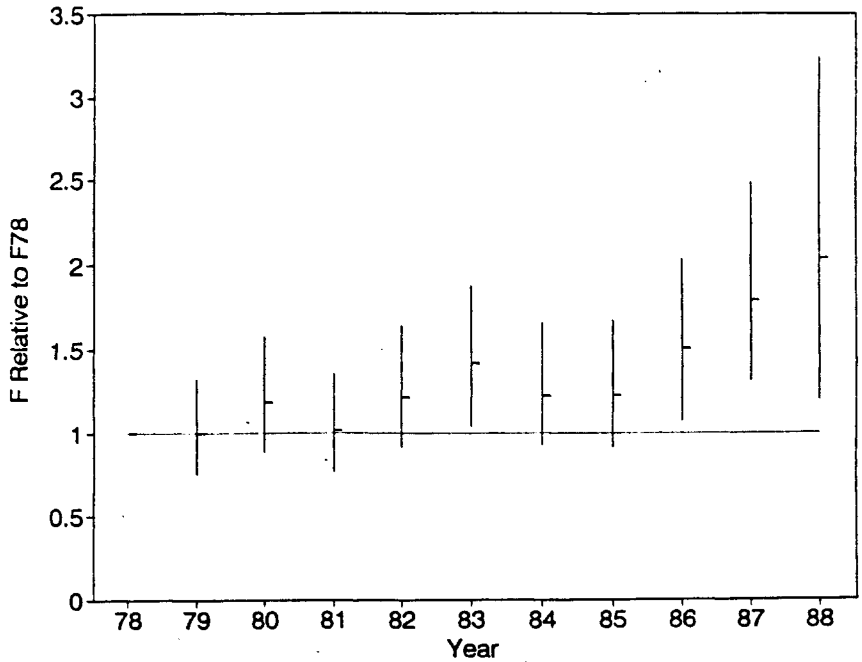

Figure 4. Distribution of the ratio of fishing mortality in year y to that in 1978 as a function of the year. Vertical bars indicate 95% confidence intervals based on percentiles; horizontal bars represent the median ratio.

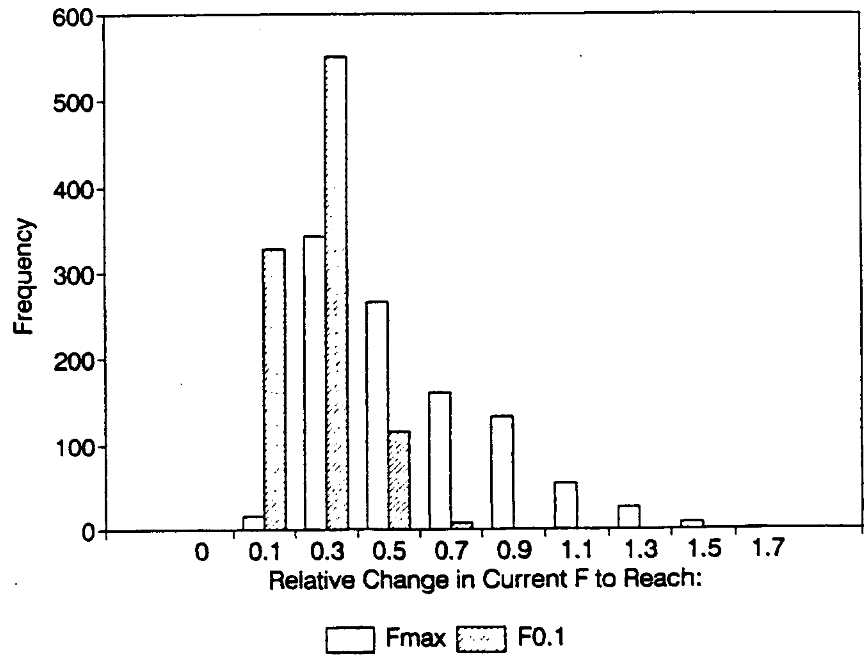

Figure 5. Multipliers necessary to bring the vector of age-specific fishing mortalities in the terminal year to the $F_{0.1}$ and F_{max} levels, for 1000 simulated data sets.

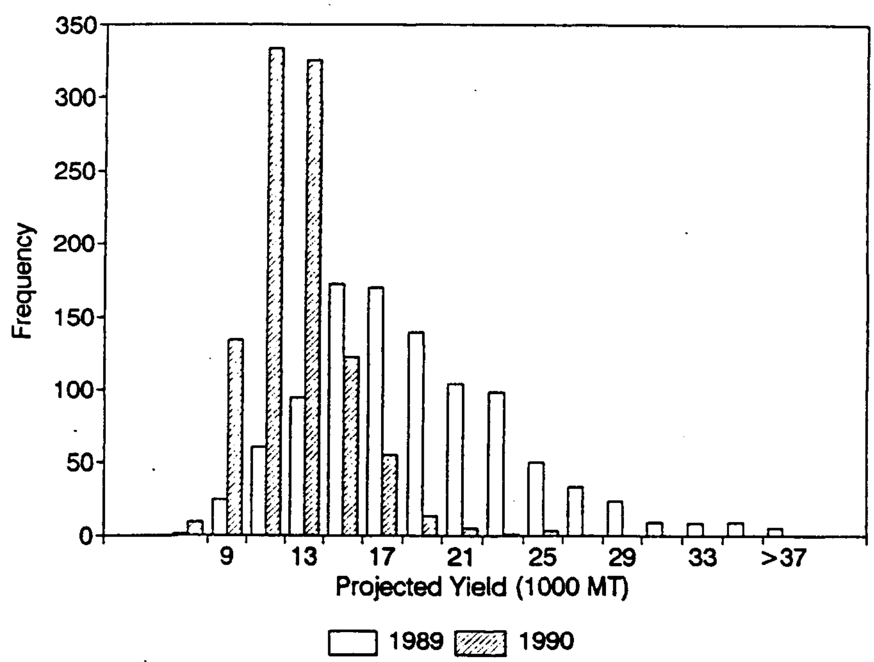

Figure 6. Distribution of estimated catches in 1989 when the fishing mortality is kept the same as in 1988 (open bars) and distribution of estimated catches in 1990 when the fishing mortality is equal to the midpoint between the fishing mortality in 1988 and $F_{0.1}$, assuming fishing mortality in 1989 was the same as in 1988 (cross-hatched bars).