NAFO Sci. Coun. Studies, 16: 59–77

Consistency of Some Northwest Atlantic Groundfish Stock Assessments

Alan Sinclair and Dominique Gascon
Fisheries Research Division, Department of Fisheries and Oceans
Maurice Lamontagne Institute, C.P. 1000, Mont-Joli, Quebec G5H 3Z4

Robert O'Boyle
Marine Fish Division, Department of Fisheries and Oceans
Bedford Institute of Oceanography, P. O. Box 1006
Dartmouth, Nova Scotia B2Y 4A2

Denis Rivard
Fisheries Research Branch, Biological Sciences Directorate
Department of Fisheries and Oceans, Ottawa, Ontario K1A 0E6

and

Stratis Gavaris
Marine Fish Division, Department of Fisheries and Oceans
Biological Station, St. Andrews, New Brunswick E0G 2X0

Abstract

The consistency of population estimates from seven northwest Atlantic groundfish stock assessments was investigated using a combined case study and simulation approach. The stocks investigated were Div. 2J+3KL cod, Subdiv. 3Pn + Div. 4RS cod, Div. 4T and Subdiv. 4Vn cod, Subdiv. 4Vs and Div. 4W cod, Div. 4TVW haddock, Div. 4X haddock, and Div. 5Z haddock. For each stock a series of assessments were performed using an objective and automated calibration technique (ADAPT). The assessments contained progressively shorter time series of input data and yielded several estimates of the same populations. The variability of the resulting estimates of the same population was investigated in terms of both range and trend when compared to those obtained from the assessment with the longest data series (the reference). For several stocks there was a tendency for the annual estimates to be higher than the reference estimates. Different formulations of the calibration model were attempted to eliminate this trend in selected case studies. Simulations of model error and statistical errors were also used to investigate possible causes for the observed trends. The tendency for annual assessments to overestimate the reference populations estimates was reproduced in cases of catch misreporting and misspecification of natural mortality in the presence of a trend in fishing mortality.

Introduction

Virtual population analysis (VPA) (Gulland, MS 1965) and cohort analysis (Pope, 1972) have been used extensively for estimating fish stock size for management purposes. The method basically consists of adding up the catches of a cohort of fish while adjusting for non-fishing or natural mortality (M) during the life of the cohort (Ulltang, 1977). An estimate of the number of surviving fish in the last year of the time series is required to begin the process. We have called this process sequential population analysis (SPA). These estimates can be derived by calibrating the analysis with an independent index of stock size. Calibration consists of choosing the set of survivor estimates that produces the best match between the SPA population estimates and the index. While the SPA estimates from the most recent time are highly sensitive to the assumed number of survivors, those from earlier years are not, provided that fishing mortality is high enough (Pope, 1972). Thus the population estimates are said to converge to values that are insensitive to the input values. After doing assessments for several years it is possible to compare the most recent estimates of the populations in years gone by to those that were obtained annually in previous assessments. This is what we have called retrospective analysis.

A working group was formed in 1986 by the Canadian Atlantic Fisheries Scientific Advisory Committee

(CAFSAC) to investigate the consistency of the Northwest Atlantic finfish stock assessments which used SPA. It was noted that, for several stocks, the yearly assessments generated more optimistic population estimates than those from the reference year (i.e. the most recent assessment) (Gascon, MS 1988, Anon., MS 1987). The pattern was age-structured with the deviations from the current assessment increasing with age. However, the assessments of the day employed a wide variety of *ad hoc* calibration techniques, too many to allow the systematic examination of the assessment deviations in relation to assessment method.

Thus, a more objective calibration framework was developed and this has become the main analytical tool of recent Atlantic Canadian stock assessments. We like to call it the adaptive framework, or ADAPT for short (Gavaris, MS 1988a). With ADAPT one treats the independent index as observed values and SPA is used as a model to produce predicted values. Functional relationships between the observations and the model results are defined, usually in the form of linear relationships on an age-by-age basis. The calibration process consists of defining an objective function, usually a sum of squares of the residuals between observed and predicted values, and then using non-linear techniques to choose the set of input parameters for SPA and the regression coefficients that minimize the objective function. The residuals may be treated in different ways to account for scale, relative error of the observation and their distribution. Two common treatments are a log transformation and standardization by the inverse standard error of the observation. The method addresses many of the technical problems noted in other *ad hoc* calibration procedures previously used (e.g. the basis for choosing the best estimate, determination of functional relationships, appropriate treatment of errors).

CAFSAC then directed its Statistics, Sampling and Surveys Subcommittee to evaluate the use of retrospective analysis as a tool for measuring the accuracy of past stock size estimates. A workshop was held in Halifax, N.S. in February 1989. Two main issues were considered; first, how reliable are the reference estimates (those obtained from the most recent SPA for the complete time series) and the assessed (all other) stock size estimates. Second, what is the best way to do a retrospective analysis? The discussions centred on population size estimates and did not consider catch projections. This was done in order to focus attention on the SPA. Population estimates were important to the accuracy of catch projections, but other factors such as target fishing mortality, weights-at-age and partial recruitment in the projection years were also important. The subject areas addressed by the four working groups of the workshop are reported and discussed in this paper.

Previous Use of Retrospective Analysis

SPA is a composite of at least three independent models: (1) a model relating population numbers, catch and M (the Baranov catch equation), here while M is an important parameter it is usually assumed constant across ages and years due to difficulties in estimation, (2) a model about age specific susceptibility (or relative catchabilities) to the fishery; the partial recruitment, and (3) a model relating other independent indices of stock size to those obtained from model 1.

The problem with model 1 is that one has n observations to estimate n+1 independent parameters, the $n+1^{th}$ parameter being fishing mortality (F) or population numbers in the last year. Therefore, model 3 is used to estimate the absolute value of F or population numbers in the last year, whereas model 2 is used to partition F amongst ages. The latter model may not be required if age specific indices are available, although it is usually included in one form or another to estimate terminal F of historical cohorts.

In spite of widespread use of SPA, it has received little attention of either theorists of field biologists, with perhaps the exception of the so-called tuning techniques (see review in Anon., MS 1988a) used to relate the three models. The use of "Retrospective Analysis" in assessing the performance of SPA is examined in this section of the report.

Gascon (MS 1988) examined the ratios between average F estimated in the final year of SPA, with the Fs in the converged part of the most recent SPA, for several gadoid stocks assessed at CAFSAC, NAFO and ICES. It was found that the range was quite large (x0.4 and x2.0 of converged F), but also that there was a definite bias toward underestimating F (6 overestimates *versus* 44 underestimates after 3 years) and therefore overestimating population size.

Rivard (1981) did a similar analysis but on a shorter data set. No bias was visible from his data, however, since he was working on a much shorter time series (in the unconverged part of the SPA), any bias may not have yet become apparent. Rivard and Foy (1987) undertook the most comprehensive study of errors in catch projections when they attempted to partition errors amongst the various sources of input. In all cases, initial stock size in the projections was the most important source of error, followed by partial recruitment and target fishing mortality, and weights-at-age in projected years. When variation in stock size was estimated by means of retrospective analysis, they found an average absolute difference of 42.7% in the stock sizes estimated for 1980 and of 49.8% for 1982 *versus* the reference estimates (of 1985 assessments). Again, there was an evident bias toward overestimating stock sizes (26 overestimates *versus* 9 underestimates), and amongst the underestimates, 4 occurred for stocks which had a poor or inappropriate database (i.e. Div. 3NO cod, Div. 4RST redfish). Population estimates for herring stocks were extremely imprecise (98% absolute error) and unbiased (4 under/3 over). For gadoids and pleuronectids, bias toward overestimating was systematic (19 over/3 under).

Anon. (MS 1984) similar analyses described the errors in ICES assessments and have been produced systematically in some assessments (Anon., 1985, MS 1988b), but owing to the difficulty of their interpretation, these forms of analysis have tended to disappear recently.

Pope and Gray (1983) examined the precision of catch projection of North sea groundfish stocks using Monte Carlo simulations. The relative importance of the sources of variation (i.e. fishing effort, recruitment and catch-at-age) varied from stock to stock. Coefficients of variation were smaller when F in the projection year was set equal to F in the previous year (i.e. the *status quo*), than when F was set to a specific target, since in the first case, errors in F tended to cancel out. Brander (1987) compared nominal catches to *status quo* projections made by various ICES working groups. He found a 14% error for one year projections and a 21% error for two year projections with no bias. However, nominal catches may be taken at different Fs than the one used in projections, and this would add another source of variability. These studies may be of limited relevance in the context of Atlantic Canadian stock assessments where *status quo* TACs are not used. Rather, projections are made at a specific target fishing mortality ($F_{0.1}$) and thus the precision and accuracy of population estimates are of greater importance.

In SPA, there are three major assumptions postulated or required: (1) that catch estimates are unbiased, (2) that catches and population numbers are related according to model 1, and (3) that M is known (and error free). We briefly examine some studies that attempted to assess the impact of biases in these assumptions.

Catch-at-age

Catch-at-age is usually the only measured parameter that intervenes in the SPA *sensu stricto*, and the variation (in the population or due to sampling) can be dealt with by standard statistical techniques. Pope (1972) and Sampson (1987) have derived variance formulae for population numbers when the variances in catch and F are given.

In addition, systematic errors can be introduced by misreporting, discarding or unrepresentative sampling. Mesnil (1980) was unable to assess the effect of such biases, but he expressed serious doubt about the usefulness of the technique under such circumstances. Sampson (1988) has shown that convergence (i.e. stability in the initial estimate of cohort size) was not maintained when errors in catch-at-age were added simultaneously to errors in other parameters.

Model errors

Ulltang (1977) extended the catch equation to incorporate migration, but migration rates are nearly impossible to measure in a systematic manner, therefore they are not considered in most assessments. Sims (1982) has examined analytically and by simulations, the assumption of the model that mortality occurs as an exponential function. He found moderately large departures (in the order of 20%) in the worst case scenarios (fishing restricted to the first or last month of the year), under very high values of M and F. The bias was usually much lower under more usual circumstances.

Natural mortality

M for groundfish is likely to be variable, depending on fish age, the abundance of predators and prey, and other factors. However, M is usually not estimated on a routine basis in SPA. Attempts have been made to estimate M due to predation in the North Sea (Anon., MS 1989) but this was based on several years of extensive stomach contents analysis and on the existence of fisheries which sample both predator and prey populations at juvenile and adult life history stages. This latter condition is not present in Atlantic Canadian fisheries. Nevertheless, the North Sea model has indicated substantial variation in predation mortality and has yielded some interesting conclusions. Specifically, it was suggested that increased mesh size in cod fisheries from 80 mm to 120 mm would actually decrease yield in other fisheries by increasing the predation of juveniles by cod and whiting (Anon., MS 1989). In other cases a "reasonable" value for M is assumed (0.2 or 0.3). Consequently, it is an additional assumption in the model, rather than a real parameter, and it is better treated as such.

Agger *et al*. (1973), Ulltang (1977), Mesnil (1980), Sims (1984), Sampson (1988) and Hilden (1988) have discussed various aspects of the effects of M on SPA. All agree that higher M leads to higher estimates of initial population numbers, and that errors in M could yield substantial errors in the estimate of population numbers. Sims (1984) found that on 10 and 20 year spans, errors of M (mean = 0.18) of +56%, +11%, –11% and –56% yielded errors in population numbers of +91%, +13%, –11% and –42%, and +261%, +26%, –20% and –62% respectively. Vetter (1988) thoroughly reviewed the methods of estimating M and the effects of the assumptions about it on fishery models. He concluded that the effects of errors in M were complex and were dependent on the errors in other parameters; in general, the effects were generally a function of the relative values of F and M. He provided strong evidence that M varied both from age to age, and from year to year.

If a bias in M exists, the effects will not be uniform throughout the VPA. The most recent years are all composed of incomplete cohorts, which will suffer smaller cumulative effects of the errors in M, than the more ancient, complete cohorts.

So far, no one has attempted a full analysis of the effects of M on the assessment process (Vetter, 1988), from the SPA to the tuning, projections and yield-per-recruit. A higher M would result in a higher estimate of historical population size, thus changing the slopes of the tuning relationships, yielding a lower (at least relative to the past) current estimate of population size, lower projected catches, but a higher F_{max} (Vetter, 1988). These effects may or may not cancel out, and it is difficult to determine at the present time.

Report of Group 1

This working group carried out case studies on seven stocks; Div. 2J+3KL cod (Baird and Bishop, MS 1989), Subdiv. 3Pn+Div. 4RS cod (Fréchet, MS 1988), Div. 4T and Subdiv. 4Vn cod (Chouinard and Sinclair, MS 1988), Subdiv. 4Vs and Div. 4W cod (Fanning *et al.*, MS 1988), Div. 4TVW haddock (Zwanenburg and Fanning, MS 1988), Div. 4X haddock (O'Boyle *et al.*, MS 1988), and Div. 5Z haddock (Gavaris, MS 1988b). The retrospective analysis for each stock commenced by first re-doing the 1988 assessment. This analysis served as the reference to which all subsequent runs were compared. Next, the assessments were conducted with progressively shorter catch and abundance index datasets. The population estimates from each assessment were compared to the reference estimates. Based on previous observations that systematic deviations from the reference estimates varied with age, comparisons were again done by age groups. The years used and the age groupings in the analyses are given in Table 1.

An index of deviation (D_{ijta}) was calculated as the ratio between the population estimates (N_{ijta}) for age-group i, stock j, population year t, and assessment year a, and the corresponding reference estimate ($N_{ijt,ref}$). The distribution of these ratios were examined among stocks and age groups. The index was calculated as:

$$D_{ijta} = \frac{N_{ijta}}{N_{ijt,ref}}$$

Plots of population size by age-class from the retrospective assessments indicate that of the seven stocks examined, Div. 4RS+Subdiv. 3Pn cod and Div. 5Z haddock had consistent patterns of estimates from one year to the next for all age groups; that is to say that by dropping one year of data, there was no tendency for the subsequent estimates of the same population to increase (Fig. 1). The pattern for Div. 4T and Subdiv. 4Vn cod was also consistent for the partially recruited ages. For Div. 2J+3KL cod, Subdiv. 4Vs and Div. 4W cod, and Div. 4TVW haddock there was a distinct tendency for the estimates in the assessment year to be higher than those in the following years. This was also the case for Div. 4T and Subdiv. 4Vn cod and Div. 4X haddock in the two older age groups.

Box and Whisker plots (Tukey, 1977) of D_{ijta} in Fig. 2 give the distribution of these ratios for all stocks by age-group. For Div. 2J+3KL cod and Subdiv. 3Pn+Div. 4RS cod the deviations for all age groups were the smallest and the ratios close to 1. The largest deviations were found for Div. 4VW haddock and Div. 5Z haddock. When compared among age-groups (Fig. 3), it was

TABLE 1. Data characteristics of groundfish stocks used in retrospective analyses. For the abundance index used, RV indicates the results of research vessel surveys while CPUE refers to commercial catch-per-unit-effort.

Stock	Abundance index used	Time span: SPA	Time span: Retro-spective analysis	Age groups in comparisons: Recruits	Partially recruited	Fully recruited
			Cod			
2J3KL	RV/CPUE	1978-88	1983-86	4	5-8	9-13
3Pn4RS	CPUE	1974-87	1980-85	4	5-9	10-15
4TVn	RV/CPUE	1971-87	1978-85	3-4	5-9	10-15
4VsW	RV	1971-87	1979-85	3	4-6	7-15
			Haddock			
4TVW	RV	1970-87	1980-85	1-3	4-6	7-11
4X	RV	1970-87	1980-85	1-3	4-6	7-11
5Z	RV	1963-87	1982-85	1	2-3	4-8

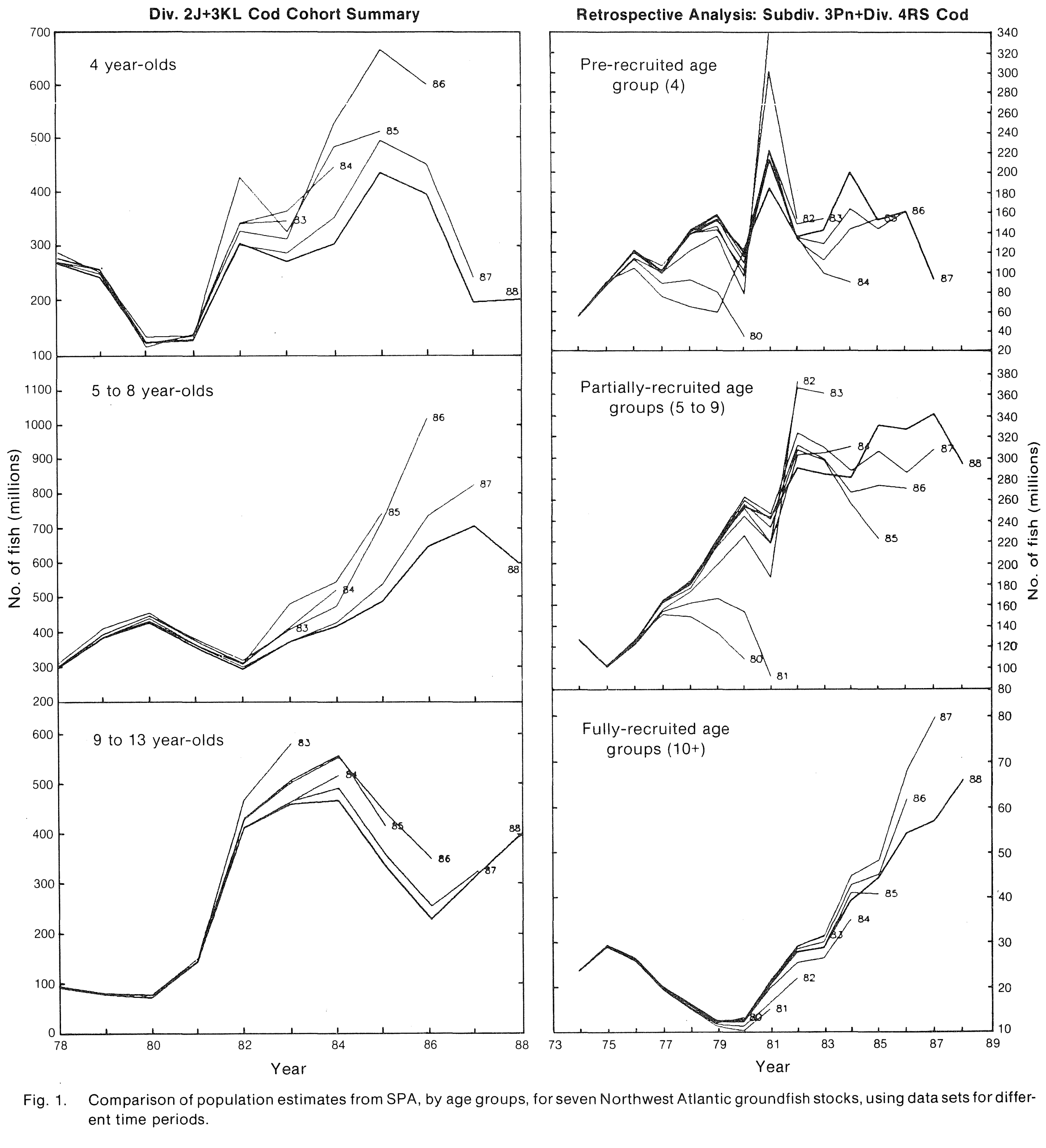

Fig. 1. Comparison of population estimates from SPA, by age groups, for seven Northwest Atlantic groundfish stocks, using data sets for different time periods.

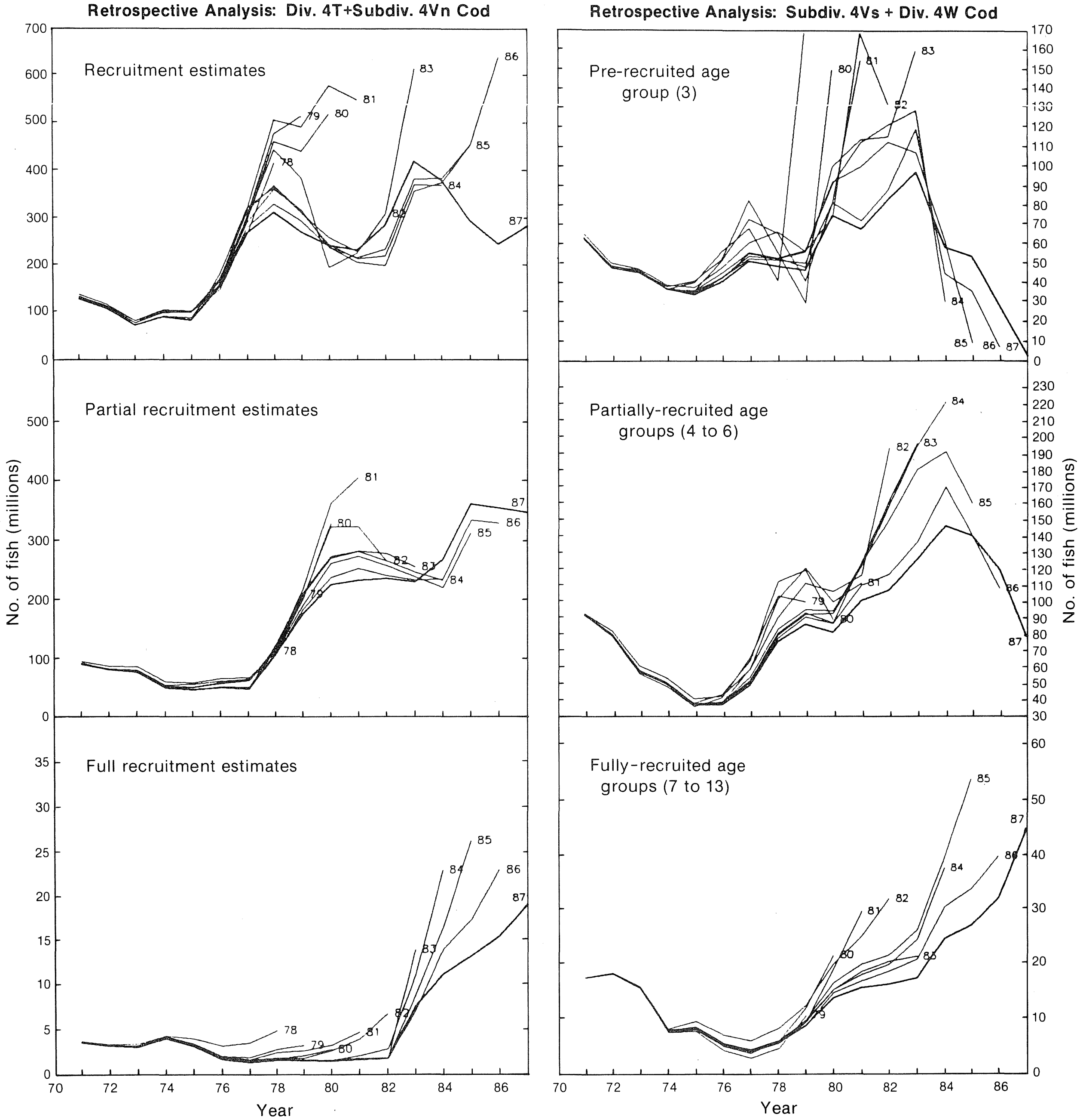

Fig. 1. (continued).

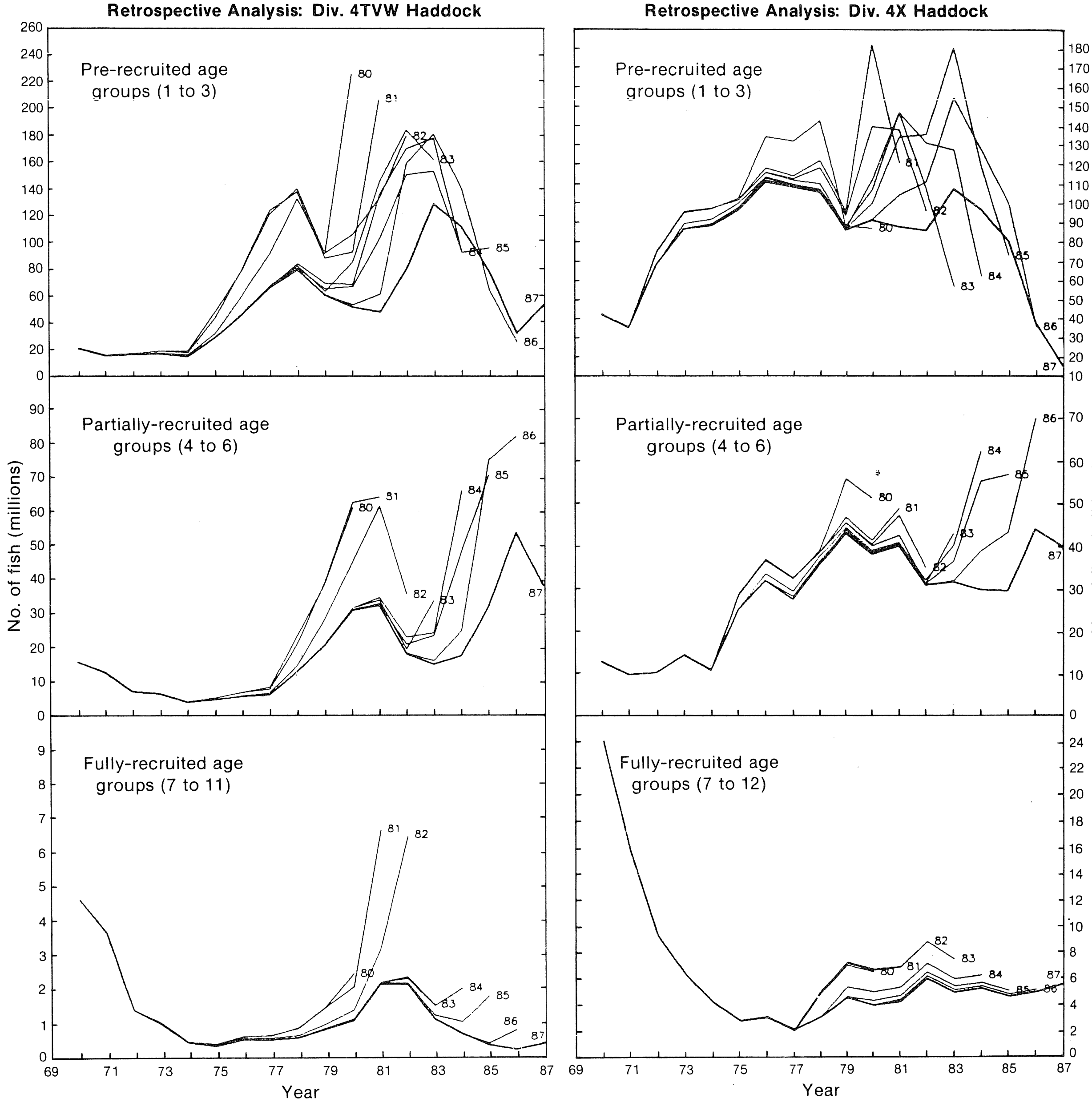

Fig. 1. (continued).

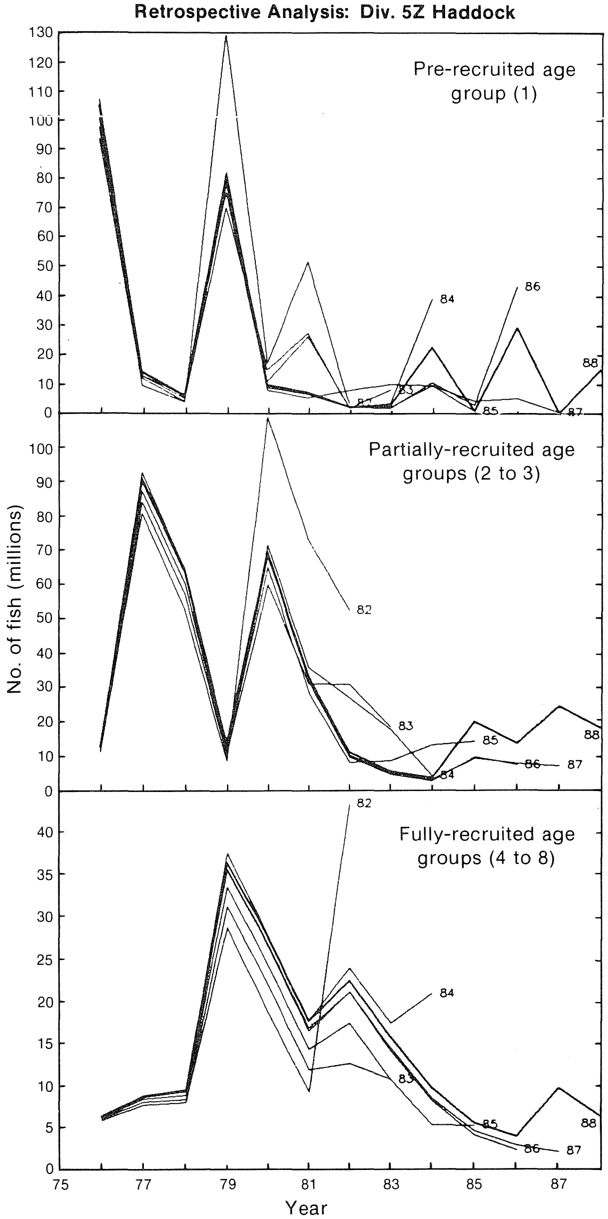

Fig. 1. (continued).

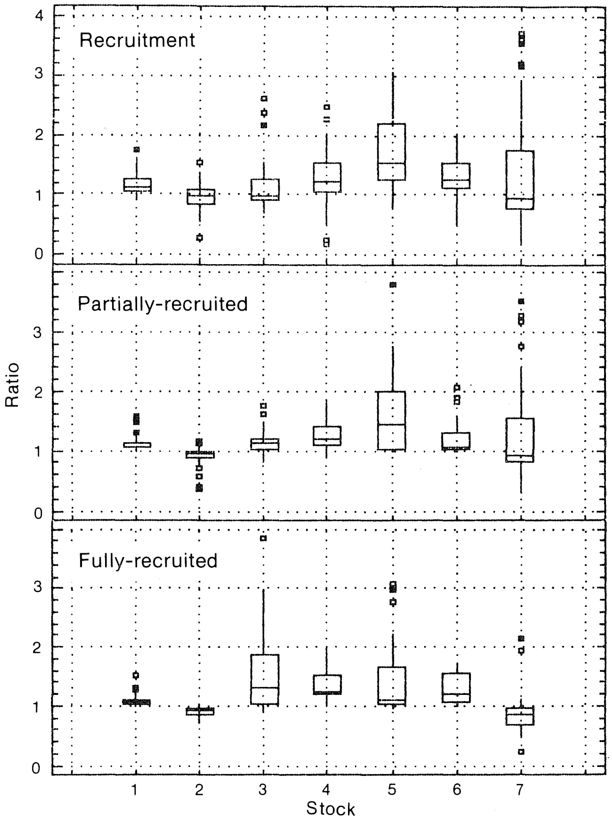

Fig. 2. Box and Whisker plots of the distribution of the ratios between assessed and reference population estimates for seven Northwest Atlantic groundfish stocks. The stocks are 1 — Div. 2J+3KL cod, 2 — Subdiv. 3Pn+Div. 4RS cod, 3 — Div. 4T+Subdiv. 4Vn cod, 4 — Subdiv. 4Vs+Div. 4W cod, 5 — Div. 4TVW haddock, 6 — Div. 4X haddock, 7 — Div. 5Z haddock.

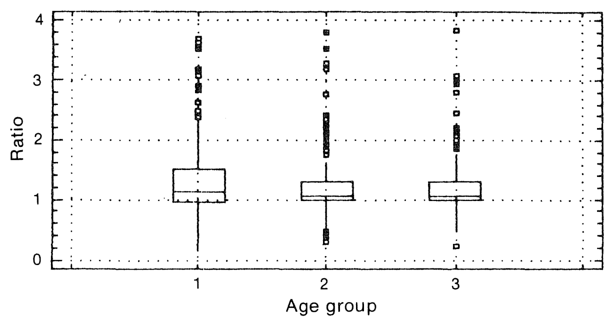

Fig. 3. The distributions of ratios between assessed and reference population estimates for all seven stocks combined, but separated by age groups (1 — recruitment, 2 — partially-recruited, 3 — fully-recruited).

apparent that over 75% of the population estimates were greater than the reference estimates. The deviations were slightly higher for the recruitment estimates than for the other two ages. When the ratios are plotted against the lag between the assessment year and the population year, one can see the effect of convergence of the SPAs (Fig. 4). After a lag of 3 years, 50% of the population estimates for partially and fully recruited ages were within ± 10% of the reference estimates. However, for estimates of recruitment, the same degree of convergence was not evident until the lag was 5 to 6 years. For the population estimates from the last year of the assessment series (i.e. lag = 0) the range of ratios was 0.3 to 3.9 and the middle 50% of the ratios were between 1.0–1.8. This indicated considerable variation in population estimates as more data were added to the assessments.

The group considered whether the source of abundance index data, either from a research survey or commercial catch rates, could be related to the pattern of population estimates. In the assessment of Div. 2J+3KL cod both indices are used in the calibration. Two series of assessments were thus conducted using survey and commercial catch rates separately in the calibration. The results are presented in terms of fully recrutied fishing mortality (Table 2). The calibrations with survey indices alone gave consistently higher fishing mortalities, and thus lower population estimates, than calibration with the commercial index although the two converged to virtually the same reference estimate by 1983. The commercial catch-rate series generated assessed F values 40–70% lower than the reference. The research survey index generated F values 30–40% larger than the current estimates in the 1983–85 period, but similar estimates in the last 2 years. A significant change in the pattern occurred in 1986, the year of what now appears to be an anomalously high survey estimate (Baird and Bishop; MS 1989).

How general this was for other stocks could not be determined due to time constaints. However, a similar comparison could be made for other stocks where both research survey and CPUE time series are available.

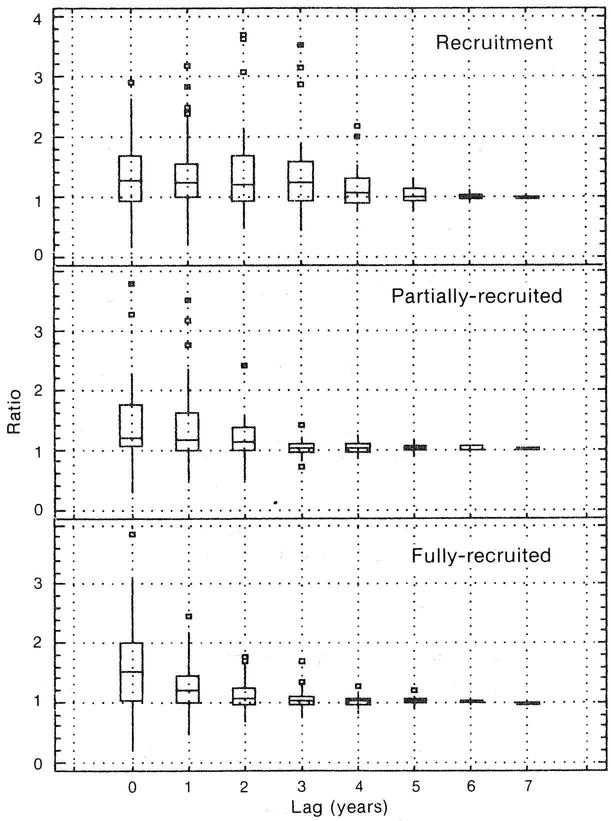

Fig. 4. Distribution of ratios between reference and assessed population estimates as a function of the lag between the last year in the SPA and the population year.

Report of Group 2

This group developed alternative formulations within ADAPT in an attempt to eliminate systematic deviations in the retrospective analyses. The alternative formulations which were considered are those related to: (1) structural changes in the underlying population dynamic model, (2) the formulation of the objective function for the minimization, or (3) the choice of the index (or combination of indices) for the calibration. The list of alternative formulations that were investigated is as follows:

- allowing temporal changes in catchabilities for the commercial fleet;
- allowing exploitation patterns to be dome-shaped rather than assuming full recruitment for older ages;
- allowing M to be estimated within the framework (i.e. M as a parameter) or assuming that M is something other than 0.2;
- using alternative indices of abundance or a different combination of them (relative weighting);
- using alternative formulations of the objective function (e.g. logarithmic transformations, weighting by inverse of standard error for indices, etc.);
- using age disaggregated indices from the commercial fleet rather than a single global index.

TABLE 2. Comparison of terminal Fs (age 13) obtained for the Div. 2J+3KL cod stock when either research vessel or commercial catch-per-unit-effort (CPUE) indices were used to calibrate the SPA.

	Calibration with RV			Calibration with CPUE		
Year	Reference	Assessed	Deviation (%)	Reference	Assessed	Deviation (%)
1983	0.472	0.627	32.8	0.470	0.202	-57.0
1984	0.507	0.707	39.4	0.475	0.174	-63.3
1985	0.545	0.758	39.1	0.552	0.147	-73.4
1986	0.484	0.480	-0.8	0.422	0.165	-60.9
1987	0.552	0.516	-6.5	0.361	0.223	-38.2
1988	0.566			0.305		

It was not possible to extensively explore the area of "alternative indices" within the time available. Similarly, the question of stock definition and its implication for the various formulations was not addressed. The questions of relative weighting of multiple indices and of using multiple indices in a single formulation or in separate formulations of the adaptive framework (combine estimates within or combine them after) were not considered extensively.

Cod in the southern Gulf of St. Lawrence (Div. 4T+Subdiv. 4Vn) and haddock on the eastern Scotian Shelf (Div. 4TVW) were used as case studies. The purpose of the exercise was to compare the retrospective analysis obtained by Group 1 with the retrospective view obtained from the "new" formulation. The aim was to find a formulation that provided the "most consistent" analysis. **Consistency, while a desirable property, must not be confounded with the "truth" and there is no guarantee that the "most consistent formulation" corresponds to the "truth".**

Cod in Div. 4T and Subdiv. 4Vn

The results for this stock are summarized in two figures, the ratios of assessed and reference estimates (Fig. 5) and a comparison of absolute population estimates from the different options (Fig. 6).

Research index-at-age using log transformations. In this formulation, only the research index-at-age was considered for the calibration. Rather than weighting the residuals by the inverse of standard error as was done in the original analysis, a logarithmic transformation was applied to the residuals. The results were comparable to those of the original analysis both in terms of deviations and absolute estimates. This was not surprising since the original formulation (Chouinard and Sinclair, MS 1988) gave very little weight (one-ninth) to the commercial catch rates. The estimates for partially recruited ages were generally less consistent and more dispersed using the RV data alone. The interannual changes in estimates of recruitment were less using the RV data alone.

Age-disaggregated commercial catch rates. In this formulation, only commercial catch rates-at-age (from otter trawlers) were used for the calibration. This formulation lead to a systematic underestimation of recruitment and partially recruited ages relative to the reference (Fig. 5). However, for the fully recruited ages, the stock size estimated each year was closer to the reference than those based on the surveys. Consequently, the commercial catch rates might be better for estimating the fully recruited ages but the research survey information provides a more satisfactory index for estimating recruitment and partially recruited ages.

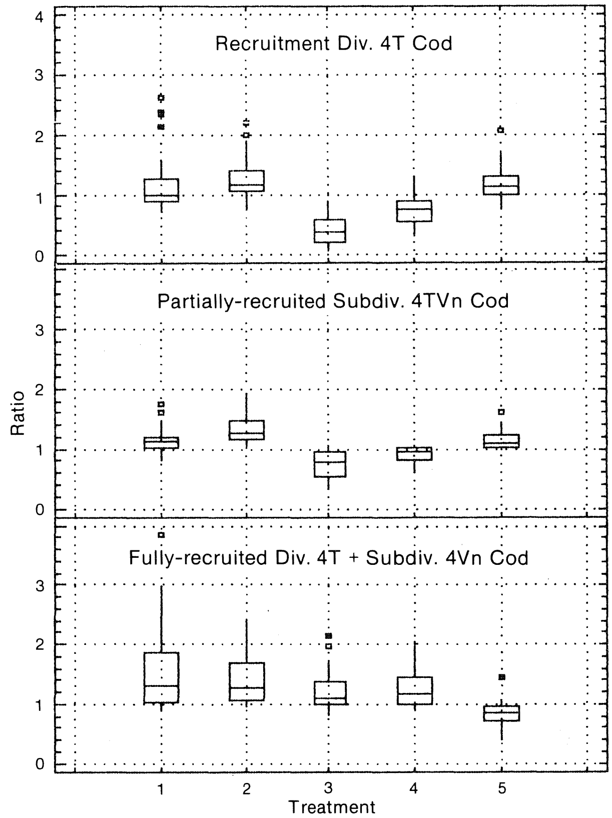

Fig. 5. Effects of changes in the ADAPT formulation on the ratios between assessed and reference population estimates for Div. 4T+Subdiv. 4Vn cod. Treatments were: 1 — the standard assessment, 2 — calibration with RV data only, 3 — calibration with CPUE data only, 4 — calibrated with RV and CPUE data and assuming a 5% annual increase in efficiency of the commercial fishery, 5 — calibrated with RV data only and forcing a dome-shaped PR pattern.

Time-varying catchability for the commercial fleet. In this formulation, the following modifications were used:

— logarithmic transformation was applied to the residuals;

— commercial catch rates were disaggregated by age (this implies a relative weighting of 1:1 for the commercial index and the research index);

— the catchability coefficients (q) for the commercial fleet were arbitrarily assumed to be time dependent, increasing at 5% per year.

The population estimates for fully recruited ages were systematically higher than the reference with this formulation but to a lesser extent than the results obtained using a formulation based on surveys only. The opposite was true for the partially recruited and recruitment estimates. This was similar to the pattern obtained when only the CPUE index was used. This option produced the lowest absolute population estimates.

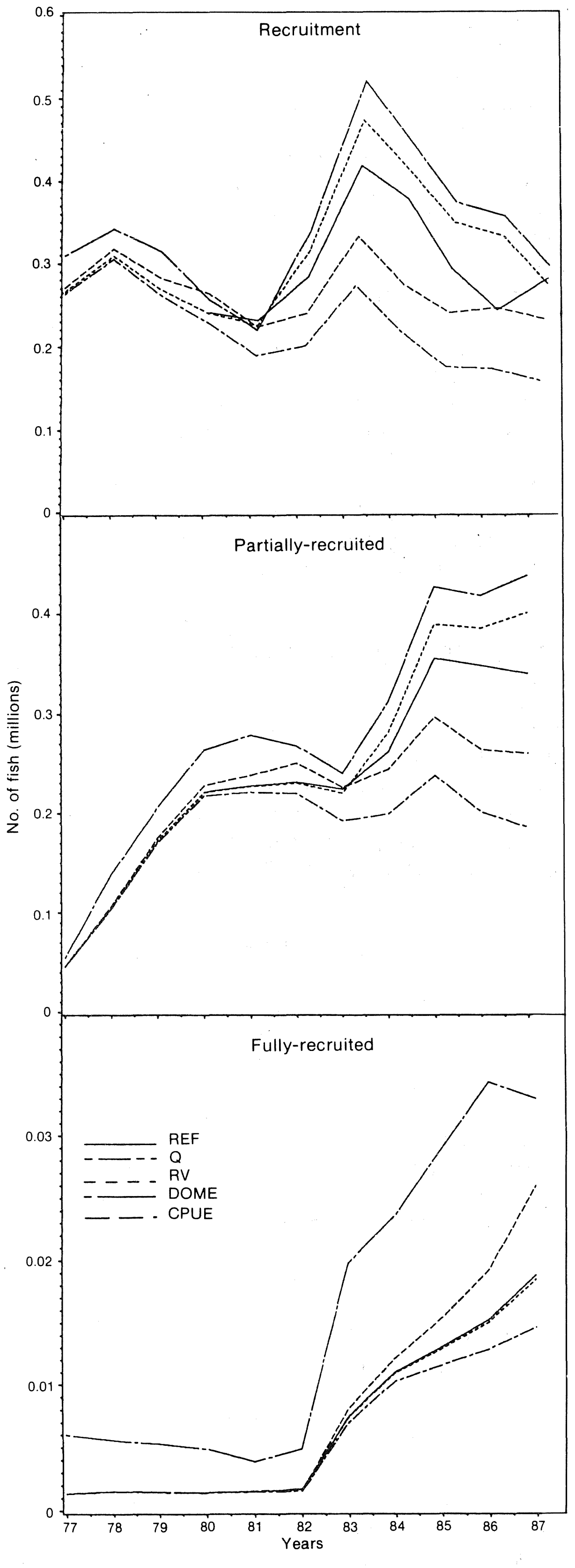

Fig. 6. Effects of changes of the ADAPT formulation on the absolute population estimates of Div. 4T+ Subdiv. 4Vn cod.

Forcing exploitation patterns to be dome-shaped. Input parameters for VPA are usually given as F for all ages in the last year and for the oldest ages for all other years. It has been our experience that there are insufficient data to estimate all these parameters. Instead, the oldest age Fs are estimated by assuming a relationship between F at younger ages in the same year and that at the oldest age. In the original formulation for this stock, F at the oldest age was assumed to be equal to the mean (weighted by population numbers) fully recruited F (ages 9 and 10), a so called flat-topped recruitment pattern. An alternative formulation was used where the pattern was assumed to be dome-shaped, that was where the oldest age F was set at 25% the fully recruited F. Only the research vessel index was used for calibration. This formulation produced a much more consistent retrospective pattern for the oldest age group and a slight improvement for the recruitment estimates. This option produced the highest absolute population estimates.

Defining M as a parameter. When M was introduced as an additional parameter, it was estimated to be 0.37. However, all values of the correlation matrix of the parameters became relatively large, which would indicate that while there is some information in the data to estimate M, there is insufficient information to simultaneously estimate M and F (through the survivors and the calibration coefficients).

Haddock in Div. 4TVW

Forcing exploitation patterns to be dome-shaped. F on the oldest age-group was set at 50% of the fully recruited F in each year. This formulation did not improve the retrospective analysis for any age group (Fig. 7).

Defining M as a parameter. When M was introduced as an additional parameter it was estimated to be 0.27. The standard error was relatively small, i.e. 0.09, but all values of the correlation matrix of the parameters were relatively large. Again, as was the case for Div. 4T+Subdiv. 4Vn cod, there was insufficient information to estimate simultaneously M and the other parameters.

Assuming M = 0.4. There was no consistency of the retrospective analysis when M was assumed to be 0.4. The analysis could not be completed for 1986 and 1987.

Assuming M is age-specific. In this formulation, M was assumed to take the following arbitrary values for each age:

Age	1	2	3	4	5	6	7	8	9	10	11
M	.9	.54	.36	.18	.18	.18	.18	.18	.18	.27	.324

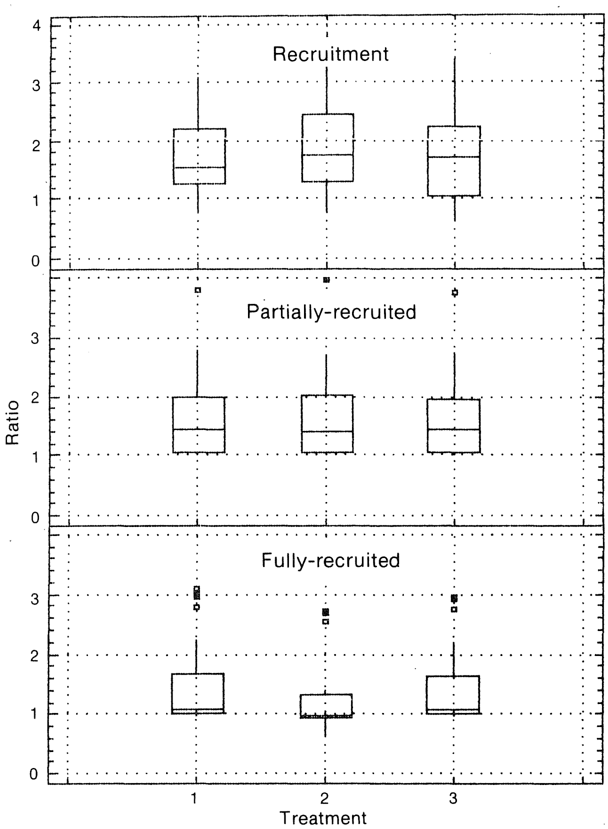

Fig. 7. Effects of changes in the ADAPT formulation on the ratios between assessed and reference population estimates for Div. 4TVW haddock. Treatments were : 1 — the standard assessment, 2 — calibration with a dome-shaped PR, 3 — calibrated with age dependent M.

With this pattern, no improvement in the consistency of the retrospective analysis could be detected (Fig. 7).

General comments

Adding parameters in an attempt to develop new/alternative formulations often led to an over-specification of the model. This was apparent by examination of the correlations among parameter estimates, a useful diagnostic tool available in ADAPT. Under these conditions, there was insufficient contrast in the data to allow estimation of all parameters simultaneously. This observation is consistent with the general experience in stock assessments, for example the inability to estimate F on the oldest age groups. For one of the two stocks for which dome-shaped partial recruitment was assumed, the retrospective analysis provided a more consistent picture than the formulation which assumes full recruitment for the oldest age. However, by assuming a dome, the results imply an abundance of older fish in the population that are not found either by the research surveys or the commercial fishery. In addition, when a dome-shaped pattern was assumed there was no complete convergence of estimates of abundance of the oldest age groups.

Recently, it was observed for many cod stocks of the Northwest Atlantic that a flat-topped partial recruitment pattern produced catchability estimates for the RVs which increased with age. From our understanding of the operation of survey trawls, stable if not declining catchabilities would be expected through the older ages. It was also observed that forcing the exploitation patterns to be dome-shaped brings the RV catchability estimates more in line with their expected pattern. The presence of a dome in the partial recruitment coefficients is not surprising because many of the gears which make the bulk of the commercial catch have a dome-shaped selectivity pattern (e.g. cod traps, gill nets, etc.). As a consequence of this, a dome-shaped partial recruitment has been assumed recently for the assessment of many cod stocks in the Northwest Atlantic, both within NAFO and CAFSAC. As indicated by the results of our simulations, the introduction of a dome could reduce, or perhaps even eliminate in certain cases, the systematic patterns observed in the retrospective analysis of many stocks.

The question of relative weighting of multiple indices in a single formulation and of using multiple indices in separate formulations of ADAPT (combine estimates within or combine them after) were only addressed in two new formulations for the Div. 4T+Subdiv. 4Vn cod stock. The commercial catch rates provided a more consistent estimation for the fully recruited ages but the research survey information proved more consistent for estimating recruitment and partially recruited ages. This should be explored further.

Report of Group 3

The general approach of this group was to use simulated data to investigate the effects of model misspecification on population estimates, both reference and in the assessment year. Input data for an assessment were generated with the following set of parameters: initial numbers-at-age, recruitment, fully recruited F, partial recruitment, and M-at-age. Population numbers (N) at age i and year t were then projected using the standard equation:

$$N_{i+1,t+1} = N_{i,t} e^{-(M_{i,t}+F_{i,t})}$$

Catch-at-age i and in year t was also generated using the Baranov catch equation:

$$C_{i,t} = \frac{F_{i,t} N_{i,t} (1 - e^{-(M_{i,t}+F_{i,t})})}{M_{i,t}+F_{i,t}}$$

The simulated population numbers were used as the index of abundance for calibration. For the analyses, either the input data (catch-at-age or index of abundance) or the parameters used in the SPA (M, partial recruitment) were adjusted to mimic systematic errors in the analytical models. The specific deviations investigated were:

- Differences between assumed and actual M;
- Changes in the catchability of the survey;
- Errors in partial recruitment assumptions (domed- or flat-topped);
- Misreporting of catches.

Populations with two age spans were used, ages 1 to 5+ and 1 to 10+. Catches and population numbers for older ages were combined. The formulations of ADAPT had the following in common:

Parameters estimated

- Survivors in the last year, either 1 to 4 or 1 to 8 depending on the age span of the simulated population;
- Slopes (ks) for the calibration relationships (to beginning of the year);

The relationships used the population estimates from SPA to predict the index, in the form:

$$\hat{I}_{i,t} = k_i N_{i,t.}$$

Structure imposed

- For the short age span, F at age 5+ in the final year was set equal to age 4. This was consistent with the partial recruitment used to generate the numbers. F on the oldest age was set equal to age 4,
- For the long age span F at age 9+ in the final year was set equal to the mean of ages 5–7, as was F on the oldest age,
- M was assumed to be 0.2 in all cases,
- Error in catch-at-age was assumed to be negligible and thus the abundance index was treated as the dependent variable in the calibration regressions. (see report of group 4 for details).

Differences between assumed and actual M-at-age

Four scenarios of deviations of M were investigated: true M declining, true M U-shaped, true M lower than assumed, and true M higher than assumed. In all cases M was assumed to be 0.2. The actual values used are given in Table 3.

The effects of these deviations between the true and assumed M on the resulting population estimates are demonstrated by the estimated slopes of the calibration relationships (Table 4). Since the true population abundance was used as the calibration index for each age, the true calibration slopes were 1.0. Higher estimated slopes indicated an underestimate of the true population (on average) while slopes less than 1.0 indicated an overestimate of the true population.

In the case with declining M, the older age groups were correctly estimated (slope = 1.0) since the assumed and true M values matched (Table 4). The SPA underestimated the true abundance for younger ages because the true M was higher than the assumed value. The population estimates were lower than the true values when the true M was higher than the assumed value (M = 0.3 and U-shaped), while the opposite was true when the true M was lower than that assumed (M = 0.1). The residuals in the calibration were negligible (less than 0.5%) meaning that the differences between the estimated population sizes and the index were almost totally explained by the calibration regressions. This was probably because the deviations between true and assumed M as well as the fully recruited F were constant for all years. Consequently, there was no divergence between the reference and assessed estimates and there was no retrospective pattern.

This was verified by using different M values while generating the simulated data. In a 10 year simulation with M = 0.2 for 5 years and M = 0.4 for 5 years, the calibration produced residuals and these were autocorrelated with time.

The group also investigated whether a trend in F accompanied by a misspecification of M could generate a retrospective pattern similar to that observed by group 1. Population and catch numbers were generated using M = 0.1 and M = 0.3 and either continuously increasing or decreasing F for a 20 year period. Trends

TABLE 3. Natural mortalities-at-age used to generate simulated data used by working group 3.

Age	True M			
	Decline	U-Shaped	M = 0.1	M = 0.3
1	1.00	0.50	0.10	0.30
2	0.80	0.40	0.10	0.30
3	0.60	0.30	0.10	0.30
4	0.40	0.20	0.10	0.30
5	0.20	0.10	0.10	0.30
6	0.20	0.10	0.10	0.30
7	0.20	0.20	0.10	0.30
8	0.20	0.30	0.10	0.30
9	0.20	0.40	0.10	0.30
10	0.20	0.50	0.10	0.30

TABLE 4. Slopes of calibration relationships as determined for population simulations of 20 years where various errors in M occurred. M was always assumed to be 0.2 while the true M is indicated in the table.

Age	Slopes of calibration relationships			
	M decline	M U-shaped	M = 0.1	M = 0.3
1	6.61	1.87	0.38	1.95
2	3.08	1.39	0.42	1.78
3	1.74	1.15	0.45	1.64
4	1.20	1.05	0.49	1.53
5	1.00	1.05	0.50	1.49
6	1.00	1.19	0.50	1.49
7	1.00	1.40	0.50	1.50
8	1.00	1.53	0.50	1.50

of F were from 0.05 to 1.00 in steps of 0.05. The data were analyzed assuming an M of 0.20. Examination of the deviations between the reference and assessed population estimates indicated retrospective patterns (Fig. 8). When M was underestimated and F decreased, there was a tendency for the assessed values to exceed the reference values. The same is true when M was overestimated and F increased. It was reported by Lapointe, *et al.* (1989) that such a situation created spurious trends in recruitment estimates. Many groundfish stocks in the Northwest Atlantic have experienced lower F since the extension of fisheries jurisdiction in 1977.

Changes in catchability of the index

As mentioned above, the simulated true population numbers were used as the calibration index. Thus, the true values were used in a manner analogous to having an abundance survey with a catchability of one. A change in catchability was simulated by changing the index either by year or by age before carrying out the calibrations. When the calibration index was doubled after 5 years this generated a discontinuous pattern in the residuals of the calibration regressions with respect to time. If the calibration index was adjusted in an age dependent manner the calibration regressions were perfect (no residuals) and the calibration slopes accurately estimated the simulated catchabilities.

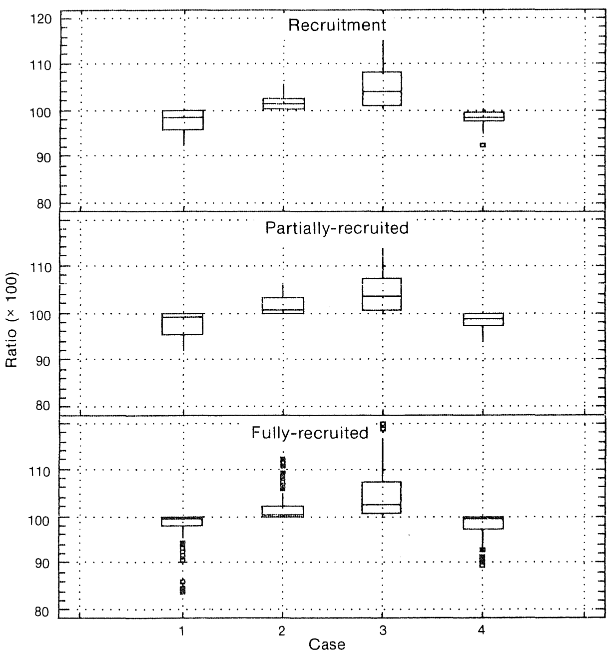

Fig. 8. Distributions of ratios between assessed and reference population estimates for simulated data where M is misspecified and there is a trend in F. In all cases M = 0.2 was used in SPA. In cases 1 and 2 the true M was 0.3 and in cases 3 and 4 the true M was 0.1. In cases 1 and 3 F increased, while F decreased in cases 2 and 4.

Partial recruitment

In this case the true population was generated with a dome-shaped partial recruitment (PR) and analyzed with a flat-topped PR (Table 5). The results indicated no deviation between the reference and assessed estimates. However, the population estimates were consistently lower than the true values. The resulting estimates of F indicated an increasing trend with age (Table 5). In a standard assessment, such a trend might be interpreted as a continually increasing PR to the fishery rather than the real dome-shaped pattern. The estimated ks also increased with age (Table 5). The largest errors were for the oldest ages and there was a convergence of the estimated and true population values toward younger ages. However, such a pattern of k in a standard assessment might be interpreted as an increasing trend in catchability to the survey. Only if a dome-shaped PR was assumed or if F was high (close to 1.0) did a dome appear in the F matrix. The lack of residuals in the fits made diagnosis of the misspecification of PR very difficult. Based on these observations, the interpretation of a dome-shaped pattern in a F matrix is unclear.

The same catch-at-age matrix was analyzed with sequential VPA (Pope and Shepherd, 1982) and the dome-shaped PR was detected.

Misreporting

In this case a simulation of annual assessments was used on a population with a 5 year age-span over a 30 year period. Two cases were examined: in one misreporting began after 10 years, in the second misreporting occurred throughout the period. The first simulation began with 10 years of perfect data. Beginning in year 11 only half of the actual catch-at-age was used, thus simulating a situation where only half of the catch was reported. An assessment was done each year and catches for the next year were projected at F = 0.2. The fishery then caught twice the TAC, and the catch-at-age necessary to achieve this was estimated. But for

TABLE 5. Fishing mortality and calibration constants (k) estimated using catch-at-age generated with a dome-shaped partial recruitment but assuming a flat-topped partial recruitment.

	Fishing mortality		Calibration slope (k)	
Age	Estimated	True	Estimated	True
1	0.01	0.04	1.48	1.00
2	0.03	0.08	1.49	1.00
3	0.09	0.12	1.51	1.00
4	0.23	0.16	1.56	1.00
5	0.37	0.20	1.70	1.00
6	0.46	0.20	2.02	1.00
7	0.49	0.16	2.61	1.00
8	0.53	0.12	3.63	1.00
9	0.54	0.08		
10	0.42	0.04		

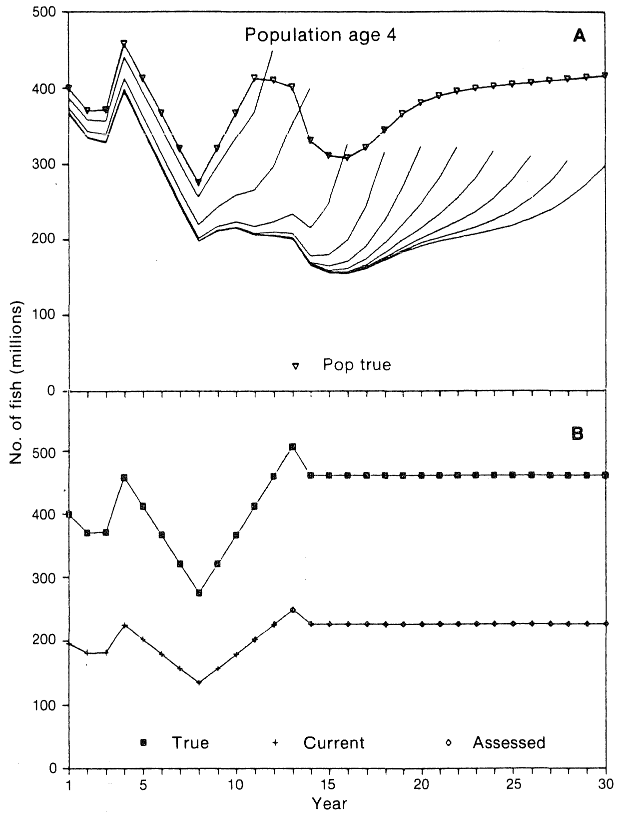

Fig. 9. Population estimates for two simulated misreporting situations with (**A**) only half the catch was reported after year 10 and (**B**) only half the catch was reported for the entire time series from year 1.

the next year's assessment only half the catch numbers were used since only half the catch was reported. The annual assessments of population size were consistently higher than the reference estimates when misreporting began at year 11 (Fig. 9). The assessed values were higher than the true values in the years immediately following the beginning of misreporting and the assessed estimates were always closer to the truth. However, the assessments consistently indicated increasing trends in population size contrary to the actual trends. This pattern in the retrospective analysis was similar to that observed for several cases studied by Group 1.

When misreporting always occurred the reference and assessed estimates were equal but half the real values. Thus the retrospective pattern described above was due to a change in simulated reporting practices.

Report of Group 4

There are two fundamentally different types of data available for the estimation of stock size, commercial catch-at-age and information from abundance indices. It is currently popular to use the catch-at-age with cohort analysis to estimate the population. An alternative approach is to use the abundance index scaled by age specific catchabilities to estimate the population. The choice of approach may be based on the relative uncertainty in the two types of data. This working group used simulated data to investigate the effects of error misspecification on the resulting population estimates.

Three formulations of ADAPT were considered to take into account the relative importance of the uncertainty in these data.

Model I	No-catch-error; assumes that the error in the catch-at-age can be ignored
Model II	No-index-error; assumes that the error in the abundance index can be ignored
Model III	Full-error; account for the error in both types of data in *ad hoc* sequential manner.

Model I is similar to the standard assessment approach. The cohort equations (Pope, 1972) were used to generate the population matrix. ADAPT was used to estimate the survivors in the final year and the calibration constants between the population estimates and the abundance index. The calibration criterion used was to minimize the residuals between the observed abundance indices and those predicted from the population estimates. The survivors for the oldest age were calculated by assuming that the F on that age was equal to the F on fully recruited ages. This assumption was consistent with the way the simulated data were produced.

Model II calculated the population as the product of the abundance indices and age specific calibration constants. The catch-at-age was then predicted with the estimated population by estimating the total annual mortality (Z) for each cohort, subtracting M to obtain annual F and using this to estimate annual catch-at-age using the Baranov equation. ADAPT was used to estimate the calibration constants using the criterion of minimizing the residuals between the observed and predicted catch-at-age.

Model III used both methods to sequentially estimate the population and produced two estimates of the population. A single population matrix was generated and the resulting true catch-at-age and abundance index were calculated for the given exploitation pattern and the index catchability.

The catch-at-age and index were decomposed into proportions and total numbers before lognormal error similar to that calculated for these stocks was added. Ten data sets were then generated for each of the three data classes described below. Each data class was analyzed with each model.

Data Class I	True catch-at-age, abundance index with error
Data Class II	True abundance index, catch-at-age with error
Data Class III	Both data types with error

The results are summarized in the tablulation below. Model I performed well even when the data violated model assumptions, i.e. there was error in the catch-at-age. There was also a slight tendency to underestimate the population size in the most recent years when there was error in the catch-at-age. Model II performed poorly. When the model assumptions were met, i.e. no error in the index, the estimates were unbiased but highly variable. When there was error in the index, the estimates from this Model were severely biased. Model III incorporated the populations from Models I and II. The poor performance of Model II was reflected in the results of Model III. Consequently, Model III did not perform as well as Model I.

	Data Class		
Model	I	II	III
I	good	acceptable	acceptable
II	biased severely	unbiased but highly variable	biased severely
III	poor	acceptable	poor

It appears that random errors in catch-at-age of a magnitude similar to what has been calculated for several finfish stocks can safely be ignored. That is, the performance of Model I was not severely degraded when there were errors in the catch-at-age. All models performed relatively well when there was no error in the index. Random error in the catch-at-age is easily handled if the abundance index is precise and accurate. However, attempts to estimate the survivors for the oldest age with Model I were unsuccessful with the level of error used in these simulations.

The coefficients of variation of the survivor estimates from Model I and 10 replications of Data Class III ranged from 34 to 55% (Table 6). A comparison of these to the model estimates showed that the model underestimated standard error by approximately 30%. This could be due to the model misspecification, i.e. the model estimates assumed that there was no error in the catch-at-age. Further simulations matching Model I and Data Class I would provide a clearer answer.

TABLE 6. Comparison of coefficient of variation for the replications and model estimates.

	Coefficient of variation	
Age	Replication estimates	Model estimates
1	0.55	0.45
2	0.39	0.33
3	0.39	0.28
4	0.40	0.28
5	0.34	0.26
6	0.51	0.32
7	0.48	0.33
8	0.42	0.35

It is noteworthy that Model I, the standard model now used, performed well if errors were random.

Summary and Conclusions

The workshop concentrated on the use of retrospective analysis for examining the consistency and accuracy of populations estimates derived from VPA. However, previous studies have shown that stock size estimates are not the sole source of uncertainties for catch projections. For instance, Rivard and Foy (1987) estimated that while population estimates might account for as much as two-thirds of the error in catch projections for certain groundfish stocks of the Northwest Atlantic, the errors in PR coefficients, weights-at-age and recruitment were not negligible. It was also noted that some of the parameters for the catch projection were negatively correlated with the population estimates and that, because of this, the errors arising from them could serve to cancel some of the errors associated with other sources when catch projections were made. Consequently, one must be careful interpreting the results presented here, as the uncertainties in the population estimates do not translate into errors in catch projections through a simple formula. The error propagation in catch projections has not been examined at this meeting and further investigations are warranted.

Most of the groundwork for this meeting was laid by the CAFSAC Working Group on the Accuracy of Analytical Assessments in 1986. Since then significant progress has been made in several areas as recommended by that Working Group in an unpublished report. This includes:

— the development of more objective assessment approaches;

— simulation studies of the effects of input on the converged part of SPA;

— quantification of the sources of variability in SPA.

Indeed the conclusions of Group 1 are consistent with the earlier observations that there was a tendency to overestimate the converged population estimates for the stocks investigated. We were unable to determine a common factor responsible for this pattern. Some improvement in the retrospective analysis was attained for Div. 4T+Subdiv. 4Vn cod if a dome-shaped PR was used for assigning F to the oldest ages. However, this was not the case for Div. 4TVW haddock. A misspecification of M in concert with a trend in F produced the desired pattern. It is generally accepted that Northwest Atlantic groundfish stocks have undergone a reduction in F since the extension of fisheries jurisdiction in 1977. Simulations of catch misreporting also generated the desired pattern of estimates if the reporting practices changed from full to partial reporting. However, these findings need study before conclusions may be drawn regarding probable causes in the case studies.

Given the present state of assessment methodology, it was generally concluded that the precision of stock size estimates in the final year will only be as good as the precision of the abundance index. For current abundance surveys age-by-age estimates have coefficients of variation of 30% or more. Coefficients of variation for aggregated age groups are less. While estimates of commercial catch rates have lower coefficients of variation, there is considerable uncertainty about changes in catchability and for several stocks such indices are currently not available. In addition Group 4 found, following simulations of realistic errors in basic input data, that coefficients of variation of 35% or more on population estimates should be expected. Thus, the development of more reliable abundance indices is called for if fisheries management plans require more precise population estimates.

The question of accuracy is another matter altogether. Here assumptions must be made about several important parameters. M is commonly assumed to be fixed through time and at age. However, there is a dearth of information regarding the dynamics of M. Data are usually inadequate to estimate the F (or survivors) of the oldest age-classes. However, faulty assumptions regarding the F on older fish can significantly bias estimates of population size and F in the past (Table 3). Other assessment methods such as separable VPA (Pope and Shepherd, 1982) may be useful for estimating the PR of older fish. Nominal catch data are often taken at face value when in fact there are no programs in place to measure accuracy of those data on a routine basis and there are substantial allegations that the reported values are far from the truth, particularly for the Scotian Shelf and Georges Bank area. The simulation studies presented here have indicated that population estimates may be biased in complex ways by faulty assumptions of these parameters. Furthermore, it was seen that in some cases the yearly assessed population sizes may be closer to the truth than the reference estimates from the most recent assessment. The use of diagnostic plots (residuals) may be useful in detecting deviations between assumed parameters and reality, however further simulation studies are needed. Thus it is clear that **estimates of population size from the converged part of the SPA do not necessarily represent the true population size for those years.**

Participants were asked to address three questions at this Special Session:

1. **Are retrospective analyses worth doing?**

   It was agreed that analyses of the type carried out here are useful for two purposes. First they indicate the variability, both in terms of range and direction, of population estimates depending on the number of years of data used. They are also useful for examining different assessment formulations which may improve the consistency of estimates. Once an improvement is attained, the implications of the new formulation in terms of population dynamics and biology of the resource should be investigated.

2. **How should they be done?**

   The general approach taken here was to repeat the assessments using a common formulation and sequentially dropping years of data. Then the variability of estimates of the same population were examined in relation to the time span of the analysis. Other factors not considered were trends in calibration coefficients and the interpretation of residuals. While not available for this meeting, the development of objective measures of variability, including direction, of the estimates is warranted.

3. **Can we use retrospective analyses to improve assessments?**

   To the extent that retrospective analyses generate questions and promote investigations, one may say that they may improve assessments. They do indicate the sources of variability in population estimates. However, this exercise alone will not necessarily give more reliable estimates of true population size.

Acknowledgements

This report presents the results of a workshop held by the Statistics, Sampling and Surveys Subcommittee of the Canadian Atlantic Fisheries Advisory Committee, and the meeting was attended by several Atlantic Canadian stock assessment scientists (Annex 1). While the authors coordinated the individual working groups and presented background material, the participants carried out many of the analyses and participated in the drafting and review of the report. Without this cooperative effort, the work would not have been possible. We also thank F. Savoie for help in preparing the figures and F. Melanson for helping to type the manuscript.

References

AGGER, P., I. BOËTIUS, and H. LASSEN. 1973. Error in virtual population analysis: the effect of uncertainties in the natural mortality coefficient. *ICES J. Cons.*, **35**(1): 93.

ANON. MS 1984. Report of the Working group on methods of fish stock assessments. *ICES C.M. Doc.*, No. Assess. 19, 56 p.

MS 1985. Report of the North Sea flatfish Working group. *ICES C.M. Doc.*, No. Assess. 13, 131 p.

MS 1987. Advice on the management of groundfish stocks in 1988. *CAFSAC Adv. Doc.* No. 16, 82 p.

MS 1988a. Report of the workshop on methods of fish stock assessments, Reykjavik, 6-12 July 1988. *ICES C.M. Doc.*, No. Assess. 26, 117 p.

MS 1988b. Report of the North Sea roundfish Working group. *ICES C.M. Doc.*, No. Assess. 21, 248 p.

MS 1989. Report of the multispecies assessment working group. *ICES C.M. Doc.*, No. Assess. 20, 176 p.

BAIRD, J. W., and C. A. BISHOP. MS 1989. Assessment of the cod stock in NAFO Divisions 2J and 3KL. *CAFSAC Res. Doc.*, No. 6, 102 p.

BRANDER, K. 1987. How well do working groups predict catches? *ICES J. Cons.*, **43**: 245-252.

CHOUINARD, G. A., and A. F. SINCLAIR. MS 1988. Assessment of the 4T and 4Vn (Jan. - Apr.) cod stock for 1988. *CAFSAC Res. Doc.*, No. 28, 47 p.

FANNING, L. P., S. J. SMITH, K. C. T. ZWANENBURG, and D. BEANLANDS. MS 1988. Assessment of 4VsW cod in 1987. *CAFSAC Res. Doc.*, No. 47, 57 p.

FRÉCHET, A. MS 1988. Evaluation du stock de morue de la subdivision 3Pn et des divisions 4R et 4S de l'OPANO en 1987. *CAFSAC Res. Doc.*, No. 24, 66 p.

GASCON, D. MS 1988. Catch projections. *In*: Collected papers on stocks assessment methods, 135-146, D. Rivard (ed.). *CAFSAC Res. Doc.*, No. 61, 167 p.

GAVARIS, S. MS 1988a. An adaptive framework for the estimation of population size. *CAFSAC Res. Doc.*, No. 29, 12 p.

MS 1988b. Assessment of haddock in NAFO Division 5Z. *CAFSAC Res. Doc.*, No. 64, 36 p.

GULLAND, J. A. MS 1965. Estimation of mortality rates. *ICES CM Doc.*, No. 3, 39 p.

HILDEN, M. 1988. Error in perception of stock and recruitment due to wrong choices of natural mortality rate in Virtual Population Analysis. *ICES J. Cons.*, **44**: 123-134.

LAPOINTE, M. F., R. M. PETERMAN, and A. D. MacCALL. 1989. Trends in fishing mortality rate along with errors in natural mortality rate can cause spurious time trends in fish stock abundances estimated by virtual population analysis (VPA). *Can. J. Fish. Aquat. Sci.*, **46**: 2129-2139.

MESNIL, B. 1980. Théorie et pratique de l'analyse de cohortes. *Rev. Trav. Inst. Pêches Marit.*, **44**: 119-155.

O'BOYLE R., J. SIMON, and K. FRANK. MS 1988. An evaluation of the population dynamics of 4X haddock during 1962-88 with yield projections to 1989. *CAFSAC Res. Doc.*, No. 72, 59 p.

POPE, J. G. 1972. An investigation of the accuracy of Virtual Population Analysis using Cohort Analysis. *ICNAF Res. Bull.*, **9**: 65-74.

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. *In*: Sampling commercial catches of marine fishes and invertebrates, W. G. Doubleday and D. Rivard (eds.). *Can. Spec. Publ. Fish. Aquat. Sci.*, **66**: 151-157.

POPE, J. G., and J. G. SHEPHERD. 1982. A simple method for the consistent interpretation of catch-at-age data. *ICES J. Cons.*, **40**: 176-184.

RIVARD, D. 1981. Catch projection and their relation to sampling error of research surveys. *In*: Bottom trawl surveys, W. G. Doubleday and D. Rivard (eds.). *Can. Spec. Publ. Fish. Aquat. Sci.*, **58**: 93-109.

RIVARD, D., and M. G. FOY. 1987. An analysis of errors in catch projections for Canadian Atlantic fish stocks. *Can. J. Fish. Aquat. Sci.*, **44**: 967-981.

SAMPSON, D. B. 1987. Variance estimator for virtual population analysis. *ICES J. Cons.*, **43**: 149-158.

1988. The stability of virtual population analysis cohort size estimates. *ICES J. Cons.*, **44**: 135-142.

SIMS, S. E. 1982. The effect of unevenly distributed catches on stock-size estimates using Virtual Population Analysis (Cohort Analysis). *ICES J. Cons.*, **40**: 47-52.

1984. An analysis of the effect of errors in the natural mortality rate and stock-size estimates using virtual population analysis (cohort analysis). *ICES J. Cons.*, **41**: 149-154.

TUKEY, J. W. 1977. Exploratory data analysis. Addison-Wesley Publishing Company. Reading Mass. 688 p.

ULLTANG, O. 1977. Sources of errors in and limitation of virtual population analysis (Cohort Analysis). *ICES J. Cons.*, **37**: 249-260.

VETTER, E. F. 1988. Estimation of natural mortality in fish stocks: a review. *Fish. Bull.*, **86**: 25-43.

ZWANENBURG, K., and P. FANNING. MS 1988. Haddock in 4TVW — population status and catch projections for 1989. *CAFSAC Res. Doc.*, No. 76, 34 p.

Annex 1

Participants

Name	Canadian Maritime Region
A. Sinclair (Chairman, Leader Group 3)	Gulf
I. McQuinn	Quebec
C. Annand	Scotia Fundy, Dartmouth
P. Fanning	Scotia Fundy, Dartmouth
K. Zwanenburg	Scotia Fundy, Dartmouth
C. Bishop	Newfoundland
W. R. Bowering	Newfoundland
D. B. Atkinson	Newfoundland
W. B. Brodie	Newfoundland
D. Power	Newfoundland
K. Brander	Scotia Fundy, Dartmouth
R. O'Boyle (Leader Group 1)	Scotia Fundy, Dartmouth
R. Mohn	Scotia Fundy, Halifax
G. Chouinard	Gulf
D. Rivard (Leader Group 2)	Headquarters, Ottawa
S. Gavaris (Leader Group 4)	Scotia Fundy, St. Andrews
D. Gascon	Quebec
E. Laberge	Quebec
A. Frechet	Quebec
G. Nielsen	Gulf
D. Clay	Gulf