Canadian Atlantic Fisheries Scientific Advisory Committee

Comité scientifique consultatif des pêches canadiennes dans l'Atlantique

CAFSAC Research Document 91/40

CSCPCA Document de recherche 91/40

Risk analysis of 4VsW cod

by

R. Mohn
Marine Fish Division
Department of Fisheries and Oceans
P.O. Box 1006
Dartmouth, Nova Scotia
B2Y 4A2

[1]This series documents the scientific basis for fisheries management advice in Atlantic Canada. As such, it addresses the issues of the day in the time frames required and the Research Documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Research Documents are produced in the official language in which they are provided to the Secretariat by the author.

[1]Cette série documente les bases scientifiques des conseils de gestion des pêches sur la côte atlantique du Canada. Comme telle, elle couvre les problèmes actuels selon les échéanciers voulus et les Documents de recherche qu'elle contient ne doivent pas être considérés comme des énoncés finals sur les sujets traités mais plutôt comme des rapports d'étape sur les études en cours.

Les Documents de recherche sont publiés dans la langue officielle utilisée par les auteurs dans le manuscrit envoyé au secrétariat.

Abstract.

A risk analysis is based on the recent assessment of 4VsW cod, from which starting values and estimates of variability are taken. Three year projections are replicated with stochastic inputs to estimate the distributions of biomass, yield and fishing mortality, F. The results are displayed as one and two dimensional integrated distributions and indicate a three year 30 kt TAC strategy (the current management plan) would probably decrease the biomass marginally in three years. However, the biomass is relatively difficult to predict as it is poorly estimated at the start of the projection and also recruitment is poorly determined. Sensitivity analyses on bias and propagation of errors are also presented which suggest that the system is well behaved.

Résumé

On procède à une analyse de risques en se fondant sur l'évaluation récente du stock de morue de 4VsW, d'où proviennent les valeurs de départ et les estimations de variabilité. Les projections sur trois ans sont reprises sous forme d'analyses stochastiques dans le but d'estimer les distributions de la biomasse, du rendement et de la mortalité F. On obtient des distributions unidimensionnelles et bidimensionnelles intégrées, révélant qu'une stratégie fixant le TPA à 30 kt (ce qui correspond au plan de gestion actuel) aurait vraisemblablement pour effet de réduire légèrement la biomasse en trois ans. Toutefois, cette biomasse est difficile à prédire, en raison des imperfections contenues dans son estimation initiale et également dans le calcul du recrutement. On présente aussi des analyses de sensibilité sur les biais et sur la propagation des erreurs. Elles révèlent que le système opère de manière satisfaisante.

Introduction.

Increasing interest has recently been expressed in vehicles with which to inform our clients of the precision of our advice. The recent NAFO workshop (Management under Uncertainties Related to Biology and Assessments, with Case Studies of some North American Fisheries, Halifax, September, 1990), the October, 1990 SSSS meeting and 1990 ICES Roundfish Working Group have all dealt with the application of risk analysis to fisheries management, and various forms of risk and sensitivity analysis were presented. In the following we attempt to develop these analytical methods further using 4VsW cod as a case study.

The basic risk analysis is straight forward, though somewhat repetitious, to perform. The projection model is conceptualized as a dynamical system with state variables (biomass, catch, F, etc.) which are determined from starting values, parameters (m, weight at age, etc.) and a model (the catch equation and exponential survivorship). A number of the parameters or starting values are simulated as distributions. Multiple runs are performed with parameters drawn from the respective distributions and the resultant state variables are are examined for bias and dispersion. There exist more sophisticated techniques, for example the Fourier Amplitude Sensitivity Test (e.g. Schaibly and Shuler. 1973) can greatly reduce the number of repetitions required to form estimates. As well as identifying 'sensitive' parameters the sensitivity analysis results can be used to assess the overall validity of the projection (Miller 1974) using an error propagation model which is conceptually similar to risk analysis.

Methods.

An age structured population, based on the 4VsW cod assessment (Fanning & MacEachern 1990) is projected forward and removals determined. The assessment provides weights at age, selectivity and recruitment indices and a starting population. We have chosen ages 1 to 9 where the oldest age is not modeled as a plus group. Where possible, uncertainties are estimated from the assessment. Table 1 contains the geometric means and standard deviations of the logged series (which approximates a coefficient of variation (CV) for small variances) for numbers at age, fishing mortality and selectivity. Other uncertainties are simply assigned (e.g. the CV for natural mortality). The problem of partitioning the variance into 'signal' and 'noise' is not scientifically addressed and was subjectively assigned as described below. Starting values are also from Fanning & MacEachern (1990) and are given in Table 2. The starting values do not match the means in Table 1 as they represent the best estimates of the January 1990 population. The projections are run for 3 years and catch, biomass, average F and maximum or fully recruited F are stored for subsequent analysis. Either a catch (TAC) or a target F may be specified. If catches are used, they are solved for iteratively to an accuracy of 100 t with the constraint that fully recruited F does not exceed 2. Any of six input elements may be given an uncertainty: starting numbers, recruitment, weight at age, selectivity, natural mortality or the TAC. The variation applied to the TAC is chosen from a distribution biased so that over-runs are likely to be of a greater magnitude than under-runs. This is done by halving the difference between under-runs and the nominal TAC as they are drawn from a lognormal distribution. Otherwise, for the other parameters the noise is lognormal. The noise is added before each replicate. Thus the natural mortality with noise added at each age remains unchanged throughout the 3 year projection. A series of test runs were performed with the noise added each year for the age specific parameters (m, selectivity, weights) but the results differed by only a few per cent which does not justify the increased complexity.

Typically 1000 or 5000 replicates are performed with the latter number requiring about 10 minutes on a microcomputer.

The results of the runs are displayed in tabular and graphic formats. A base run with a TAC of 35 kt for three years is used as an example and the means and coefficients of variation are compiled and an example may be seen in Table 3. Two indices of sensitivity are presented. The first is a relative sensitivity $R_{i,j}$ which is in essence a local derivative of the state variable x_i with respect to a parameter p_j. (See Table 4).

$$R_{i,j} = p_j \, \Delta x_i \, / \, x_i \, \Delta p_j$$

Or equivalently, what is the percentage change in variable i due to a percentage change in parameter j. A value of 100% indicates that a 10% change in parameter resulted in a 10% change in the state variable. Values greater than 100% reflect increasing sensitivity. To form these estimates the specified parameters are increased by 10% with all others remaining at the base levels. One exception to this scenario is the selectivity as it would not make sense at a value of 1.10 for a selectivity.. For the selectivity all values are increased by 10% and then truncated to unit magnitude. This has the effect of shifting the selectivity to a younger age.

The second index of system performance is the sensitivity to error propagation. If parameter j is drawn from a distribution having a given (10%) coefficient of variation, what will be the coefficient variation of state variable i. As well as the CV of the state variable, the resulting changes in the mean are also presented.(Table 5) The mean will be affected as the system has non-linear elements and because the noise is lognormal and as standard deviation increases so does the mean. If μ and σ are the mean and standard deviation of a normal distribution, then the mean of the lognormal is:

$$E[X] = \exp(\mu + \sigma^2/2)$$

Again the results have been scaled so that 100% indicates a unit derivative.

The main form of graphical representation is two dimensional distributions having biomass as one axis and fishing mortality or catch as the second. The appropriate state variables are aggregated into a rectangular grid to approximate a probability distribution. Then the distributions are integrated and displayed as contours (Figures 1 and 2). The distribution is integrated from the most probable cell to the least , however, instead of the the usual cumulative distribution which integrates from the lower end of the range to the upper (Figures 3-5). Appendix A contains more details of this integration scheme. This has the affect of accumulating the cells so that the most probable states may be identified. For example, once 50% of the distribution was found those cells could be shaded black. After 25% more of the distribution was accumulated the relevant area could be shaded grey and the would enclose the area of 75% of the most probable outcomes, etc.

Parameter estimation.

Numbers at age and F's at age were taken from Fanning & MacEachern (1990) to produce estimates of the variation for risk estimation. Logarithms were taken of the values and the mean and standard deviations estimated over the 18 years of data available (Table 1) The means of the logged values were then antilogged to give GM estimates. Neither the numbers nor F's at age for age 1 were given for the most recent year so values of 90,000 and .001 were assumed to fill the matrices. Also any values of F less than .001 in their data were rounded up to this level. The selectivities were derived from the F table by dividing each year's by its maximum value. The standard deviations for numbers at age increased with age from about 0.4 to about 0.7 by age 9. The standard deviations of the logged F's and selectivities showed very high variations for younger ages. Starting values for the projection were not taken from these mean values but from values given in the text of Fanning & MacEachern (1990) and are reproduced in Table 2.

Coefficients of variation for the recruitment are estimated at 0.38 in Table 1 from which we set a value of 0.4 for use in the risk analysis. The older ages have higher CV's in Table 1. These values may be compared to the CV's given in Fanning & MacEachern (1990, Table 17a) for the parameters estimating numbers at age from their ADAPT run. For ages 4 to 9 the values are 0.44, 0.34, 0.31, 0.29, 0.38 and 0.38 respectively. The difference in the two series may be ascribed to the partitioning of variance into signal and noise. Strong year classes will cause variation but are not noise. A value of 0.4 was assigned to the CV for initial numbers at age. Weights at age are assumed to be fairly easy to estimate and are assigned a CV of 0.1. As mentioned above, given the absence of any data, a CV of 0.2 has (conservatively) been assigned to the natural mortality. The CV's for the selectivities are large, particularly for younger animals. A fair proportion of this variation will be in response to year class strengths and changes in management regulations and markets over the 18 year data period. A relatively low value of 10% is assigned for selectivity. Although this value is lower than might seem appropriate it was chosen because higher values produced very bizarre selectivity patterns which caused outliers in projections. The CV for the TAC's was arbitrarily set at 20%. As mentioned in Methods, the lognormal distribution of errors for TAC's is adjusted such that under-runs are less likely to be large.

Results.

Table 4 contains the resultant sensitivities to 10% perturbations in 6 parameters. The starting values from Table 2 are used and a 35 kt TAC is set for all three years. These results do not have any stochastic elements active. As would be expected a 10% increase in the TAC, results in approximately a 10% increase in catch in each year of the 3 year projection and hence relative sensitivities on the order of 100%. The biomass shows a weak negative response and the average F went up considerably, especially by the 3rd year. The next block in this table shows that the system is insensitive to a 10% increase in recruitment. The system shows very similar sensitivities to an increase to either the initial numbers or weights at age as would be expected because they have a similar effect on biomass. Because the catch is limited by a TAC and solved for iteratively, it is not affected. The 10% increase in natural mortality has a cumulative affect on biomass which in turn requires harder fishing to catch the quota. The shift in selectivity to younger ages requires that both average and maximum, or fully recruited, F be monitored. The influence on average F is in the range of 20% while the maximum F is around 50%. In general terms most sensitivities fall into three classes: those much less than 100% which reflect an insensitivity of that variable to the specific parameter, those near 100% which are proportionally affected, and those greater than 100% which show the nonlinearity between F and catch.

Table 5 shows the sensitivity of the response to adding lognormal variation to various parameters, error propagation. The first block in this table is the effects of adding a 10% CV to the TAC, subject to the adjustment described under Methods. These results are based on 500 replicates. After three years the catch and F display a slight increase in their means, on the order of half the response in the CV. The response of the CV's was slightly less than 100% which reflects the adjustment for underruns and the F response is slightly over 100% because of the non-linear relationship between F and C. The addition of a CV to the recruitment series has only a small cumulative effect in the projection. And as in Table 4 the CV has a very similar affect when added to either the initial numbers or to the weights at age. The system is relatively insensitive to variance in the natural mortality. Noise added to the selectivity pattern has a considerable affect on the resultant F_{Max}. This is because of the interaction of lognormal noise and the traditional normalization applied to selectivity. If a single age gets a large value from the addition of noise all others are scaled downward and F_{Max} must be increased to compensate. Table 5 suggests that the sensitivity to noise does not have any unpleasant non-linearities or amplifications. It also shows that, excepting the sensitivity of F_{Max} to noise in the selectivity, bias is not greatly influenced by lognormal noise being added to the parameters.

Figures 1 and 2 are the integrated probability distributions from year 3 of the base run with a 35 kt TAC and 40% noise on initial numbers and recruitment, 20% on TAC's and natural mortality and 10% on weights at age and selectivity. Figure 1 indicates that after 3 years the biomass will be between 180 and 280 kt with about a 50% probability. This may be compared to Figures 3 and 4 which show similar information displaying the marginal distribution. Such distributions and their integrals are of the form used in Hoenig et al. (1990) to present results, although the variables of the distributions are not the same. Unfortunately Figures 1 and 3-4 were derived from separate runs so a point by point comparison is not supported. Figure 2 displays the biomass versus F distribution from the same data as Figure 1. Reference values of $F_{0.1}$ and F_{sq} are superimposed. $F_{0.1}$ is seen to be well within the cloud from the 35 kt strategy and the probability of exceeding F_{sq} is low. Figure 5 shows this and the high degree to which the distribution is skewed. The fine structure in Figures 1 and 2 does not have significance and is artifactual from the aggregation scale and to a lesser degree from the limited number of iterations used in producing the data.

Figures 6 and 7 show the 50% contours for a range of constant TAC strategies. In Figure 6 the increasing TAC moves the contour up and to the left. For 20 and 35 kt the principal axes of contours are parallel to the figure's axes, reflecting that the variables are independent. The contour for the 65 kt strategy is tilted about 45° showing a high correlation between the biomass and the catch. The reason for this is that the system has become resource limited after 3 years under this strategy. This is because the fishing effort is limited to an F of 2, the capacity of the fleet, and resource availability limits the catch. The contour for a 50 kt TAC reflects a combination of independent and dependent modes. That is some of the runs were limited by the TAC and some were limited by the resource. Figure 7 shows a projection of the same data set data with F replacing Yield as the Y-axis. In this plot the two regimes of the 50 kt TAC strategy are clearly shown as separate domains.

Discussion.

We have presented sensitivity and risk analysis for the 4VsW cod stock. The sensitivity analysis of the 3 year projections was rather encouraging in that no strong (much greater than 100%) sensitivities are revealed. These results were for both traditional sensitivity and the propagation of variances. The risk analysis focused on three state variables, biomass, yield and F. Two dimensional projections of this 3 dimensional space were presented to aid in interpretation. A 35 kt TAC would appear to reasonable strategy. Even so the biomass is still difficult to predict for year 3. The 50% range of the most probable values ranges from 180 to 280 kt.

We have not dealt with the definition of risk. Francis (1990) discusses risk as the probability of 'something bad' happening in a given time under a given policy. For most fish stocks 'something bad' is hard to quantify. We do not have stock-recruit relationships so that recruitment overfishing, or the probability of stock collapse, cannot be defined. Because yield per recruit relationships are generally smooth functions, a critical level of growth overfishing cannot be determined. The situation for growth overfishing is further confounded by the economics of harvesting. Our results did show an interesting bifurcation when the system becomes resource limited. A 5 or 10% probability of entering the resource limited state might provide a reasonable threshold and definition of 'something bad'.

Our results are hampered by the amount of computation required and the requisite communication among the various programs to produce the final output. A major constraint was the coarseness of the aggregation of the replicate runs for integration and display. Minor changes of parameters could not be undertaken casually. Fortunately, the sensitivity analysis suggests that minor changes to parameters will not greatly affect results.

The above results, which are for predictions into the near future, should not be extrapolated to longer time frames. This is because the recruits were drawn from a stationary distribution and no stock-recruit relationship was used. Because recruitment dominates projections of more than a few years, both deterministic and stochastic aspects need to be known.

The variances used in the above risk analysis are hopefully of the right magnitude. The recruitment series and starting numbers should be estimable with proper statistical methods. Natural mortality and its distribution on the other hand will remain elusive. We used only lognormal distributions but it is reasonable to expect that other distributions may also be implied upon more detailed analysis of the data. More work is also required concerning how to package the results to render them assimilable by our clients.

References.

Fanning, L.P. and W.J. MacEachern. Stock status of 4VsW cod in 1989 using a Half-year SPA Formulation. CAFSAC Res. Doc 90/88.

Francis, R.I.C.C. 1990. Risk analysis in fishery management. NAFO SCR Doc. 90/93. 18p.

Hoenig,J.M., J.W. Baird and V.R. Restrepo. 1990. Describing the Risk Associated with a Given Quota when the Stock Assessment is based on Sequential Population Analysis. CAFSAC Res. Doc. 90/75. 13p.

Miller, D.R., 1974. Sensitivity analysis and validation of simulation models. J. Theor. Biol. 48:345-360.

Schaibly, J.H. and K.E. Shuler. 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. II. Applications. J. Chem. Phys. 59:3879-3888.

Table 1. Geometric mean (GM) and standard deviation (SD) of log for numbers, F's and selectivities at age

Age	Numbers *10^3		F		Selectivity	
	GM	SD	GM	SD	GM	SD
1	85554	.38	.003	1.448	.005	1.113
2	69634	.38	.006	2.359	.011	2.008
3	53385	.41	.032	1.855	.057	1.579
4	35858	.37	.151	.786	.267	.561
5	24419	.44	.334	.520	.592	.318
6	14002	.58	.426	.391	.755	.262
7	7425	.67	.447	.372	.793	.223
8	3509	.71	.461	.433	.817	.182
9	1640	.71	.400	.444	.708	.361

Table 2. Starting values used in risk analysis.

Age	Weight	Numbers *10^3	Selectivity
1	.07	77.	.0001
2	.28	63.	.001
3	.59	60	.01
4	.95	48.8	.16
5	1.29	20.7	.53
6	1.69	15.1	.81
7	2.20	6.9	1.00
8	2.62	10.9	1.00
9	3.32	2.2	1.00

Table 3. Sample means and coefficients of variation from the base run.

	Mean	CV	mean	CV	mean	CV
Biomass	228.	.17	239.	.19	242.	.20
Catch	36.8	.16	36.9	.16	36.8	.16
F	.252	.32	.249	.38	.254	.44

Table 4 Relative sensitives from base run. Relative sensitivity of state variables in percent resulting from a 10% change in specified parameters.

	Biomass	Catch	Average F	Maximum F
TAC				
Year 1	0	91	123	
2	–17	100	169	
3	–28	96	196	
Recruitment Series				
Year 1	3	0	0	
2	12	0	0	
3	28	0	–1	
Initial Numbers				
Year 1	97	–1	–109	
2	10.5	0	–137	
3	100	–2	–157	
Weights at age				
Year 1	100	–1	–109	
2	117	0	–138	
3	128	–2	–158	
Natural Mortality				
Year 1	0	0	11	
2	–21	1	40	
3	–42	–4	69	
Selectivity				
Year 1	0	0	-15	-47
2	-1	0	-22	-53
3	-4	1	-25	-57

Table 5. Error propagation from a (nominal) 10% CV in specified parameters. The first 4 columns are the shift in the mean and the second 4 are the CV 's of the state variables.

	Change in mean (%)				CV (%)			
	B	C	F_{Max}	F_{Ave}	B	C	F_{Max}	F_{Ave}
TAC								
Year 1	0	29	37		0	79	99	
2	–5	26	46		14	79	105	
3	–8	17	42		17	81	106	
Recruitment Series								
Year 1	0	0	0		3	0	0	
2	1	0	0		10	0	0	
3	1	0	0		18	0	0	
Initial Numbers								
Year 1	5	0	–3		37	0	55	
2	5	0	–3		41	1	66	
3	5	–2	–5		41	1	75	
Weights at age								
Year 1	4	0	–1		38	0	54	
2	6	0	–2		43	1	67	
3	6	–2	–3		48	1	82	
Natural Mortality								
Year 1	0	0	0		0	0	5	
2	–1	0	1		8	0	16	
3	–2	–1	1		16	1	29	
Selectivity								
Year 1	0	0	86	1	0	0	68	21
2	0	0	83	–1	1	1	69	25
3	0	–1	82	–2	4	1	73	25

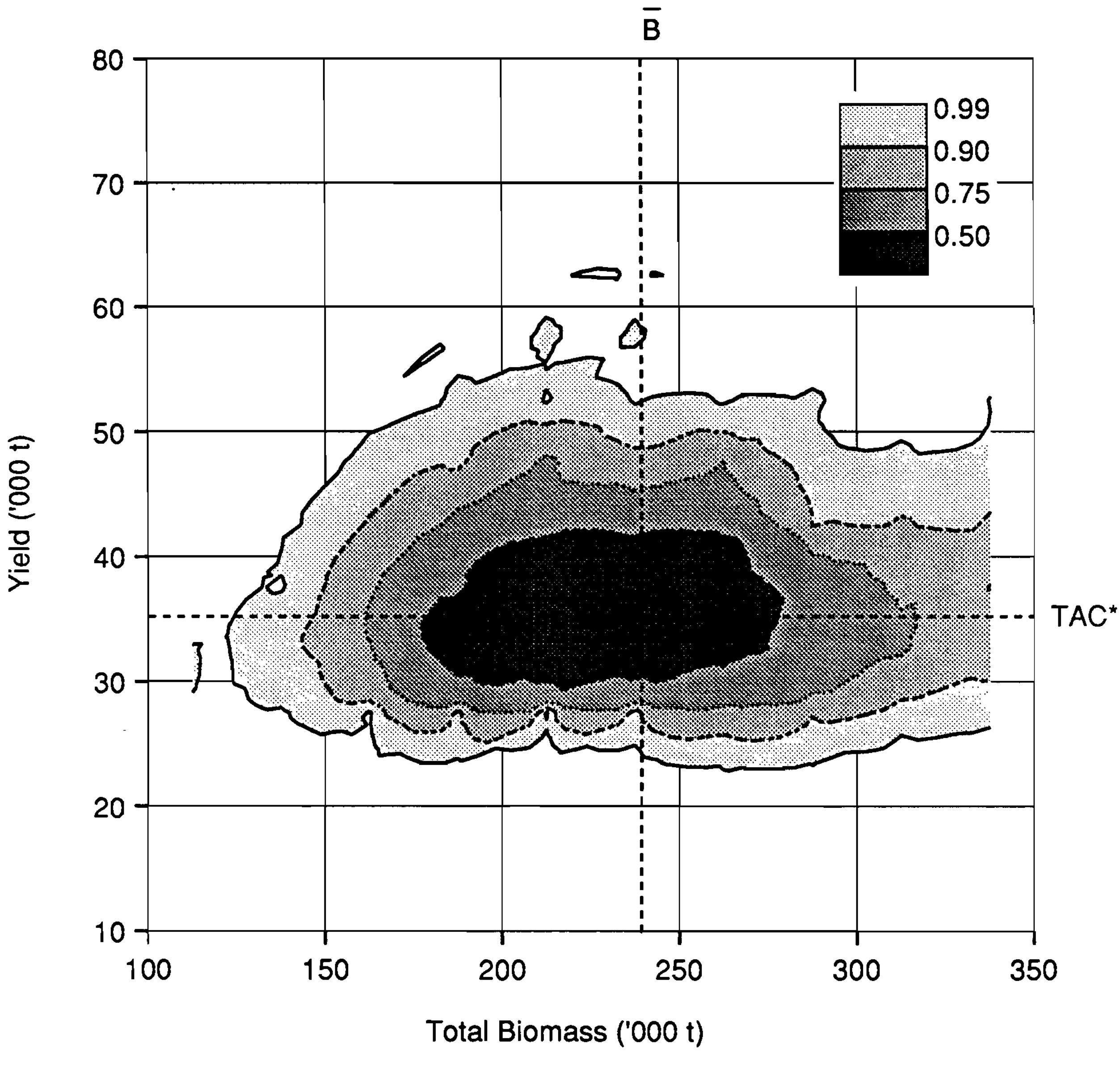

Figure 1. Sample integrated probability distribution of Yield vs biomass in 3rd year of projection.

*For reference average biomass and TAC are shown in dashed lines.

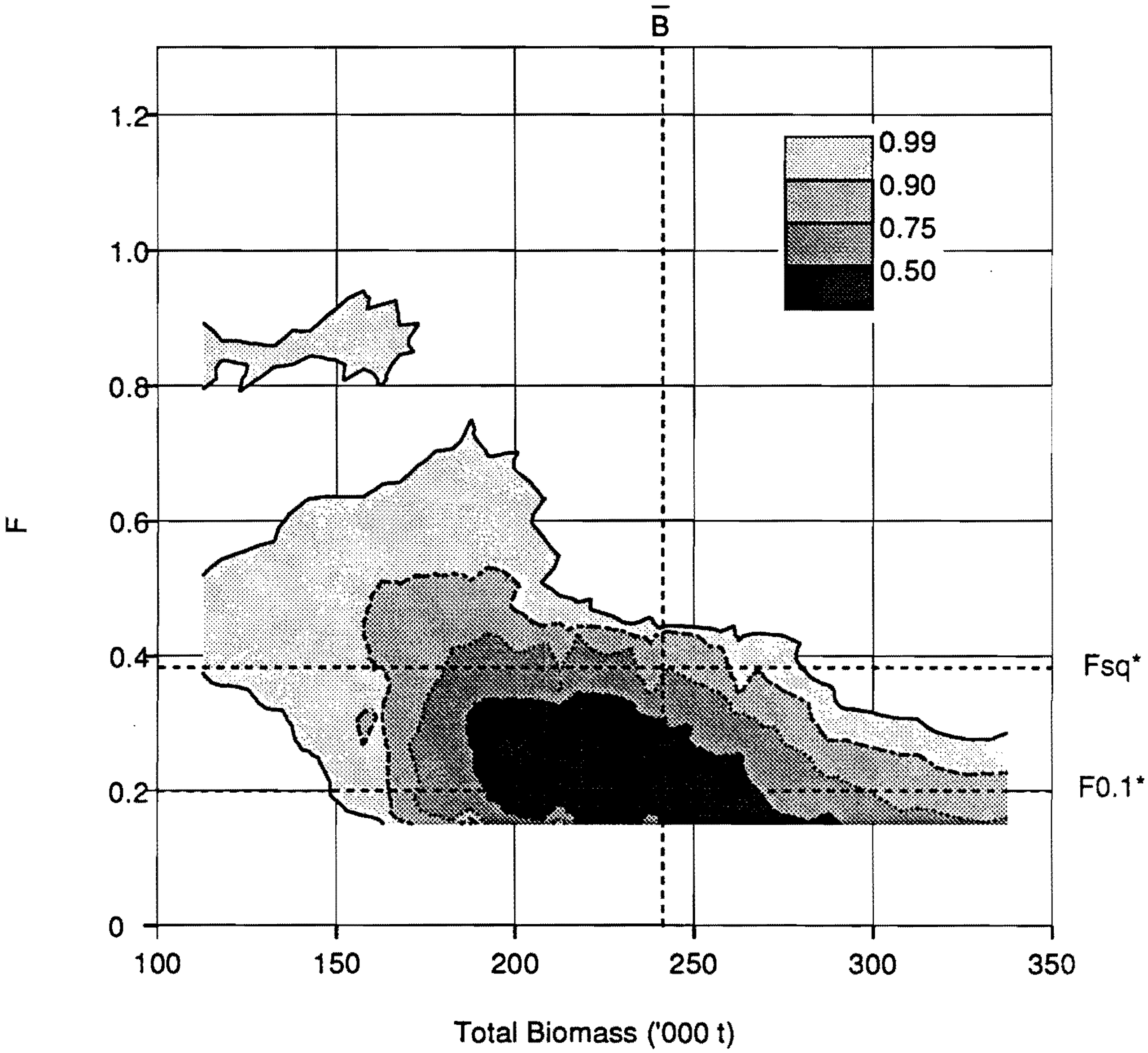

Figure 2. Sample integrated probability distribution of F vs biomass in 3rd year of projection.

*For reference the average biomass, status quo F and F0.1 are shown.

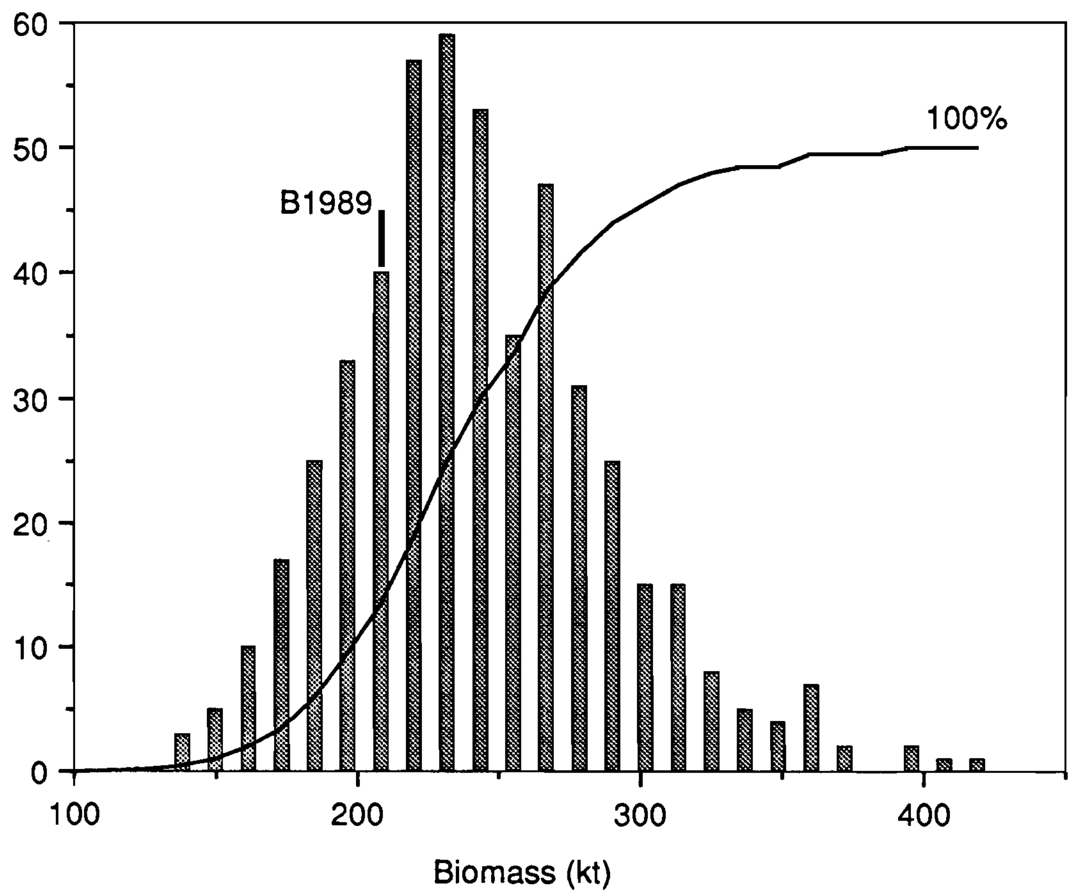

Figure 3 . Frequency and cumulative distributions of total biomass after 3 years of 35 kt strategy. The 1989 biomass is given for reference

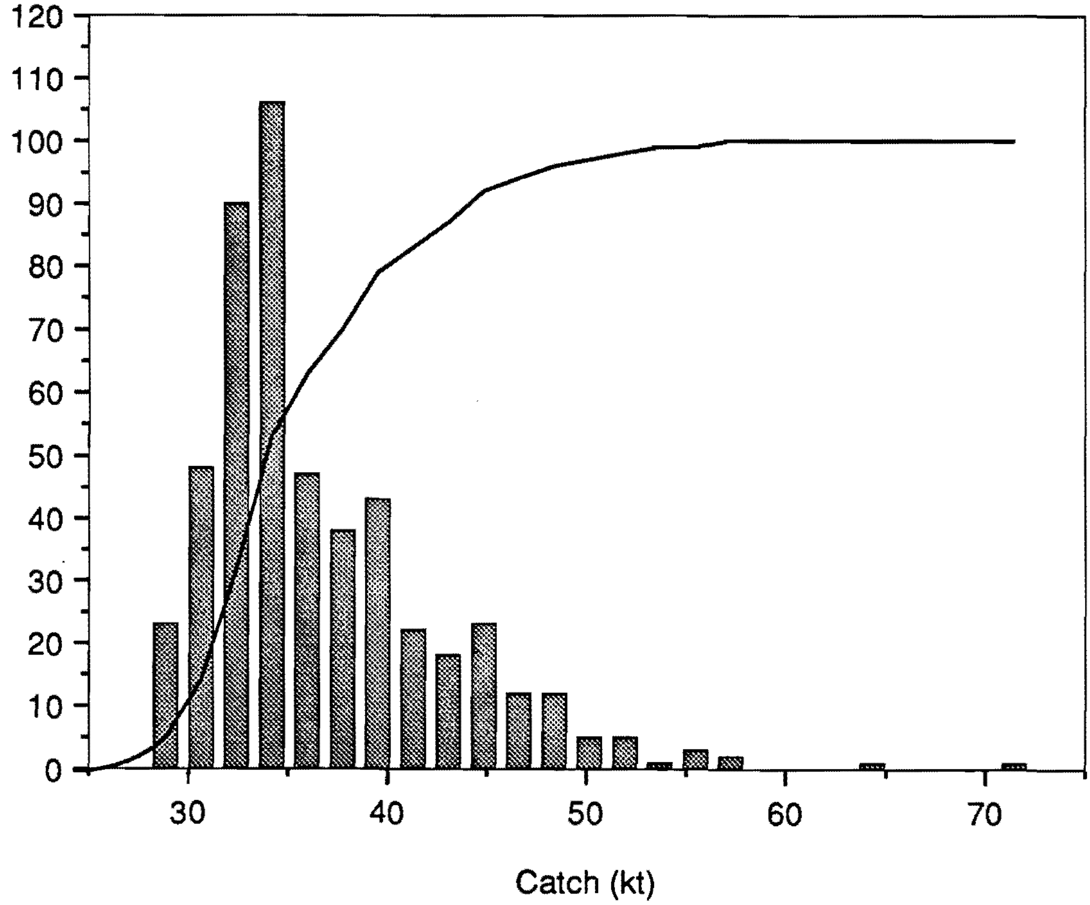

Figure 4 . Frequency and cumulative distributions of catch after 3 years of 35 kt strategy.

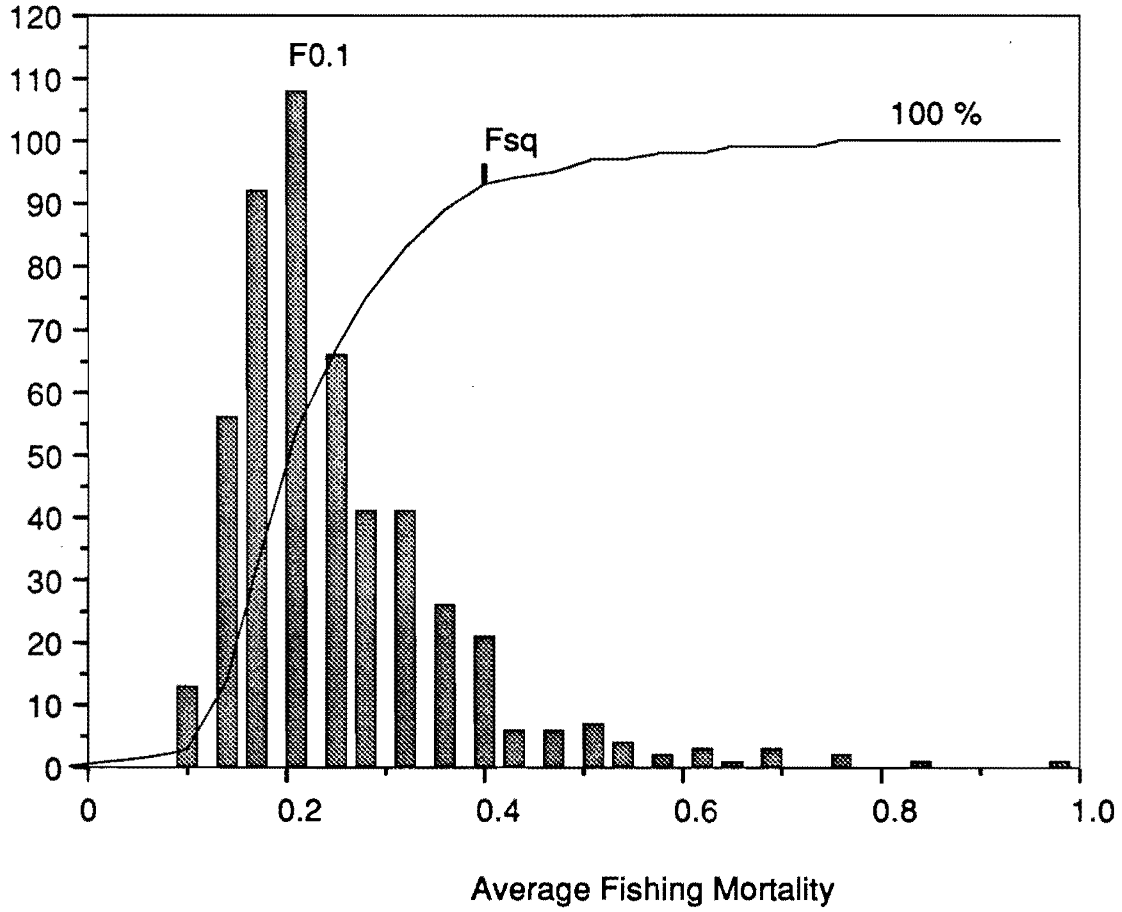

Figure 5. Frequency and cumulative distributions of fishing mortality after 3 years of 35 kt strategy. Reference F0.1 and Fsq are given.

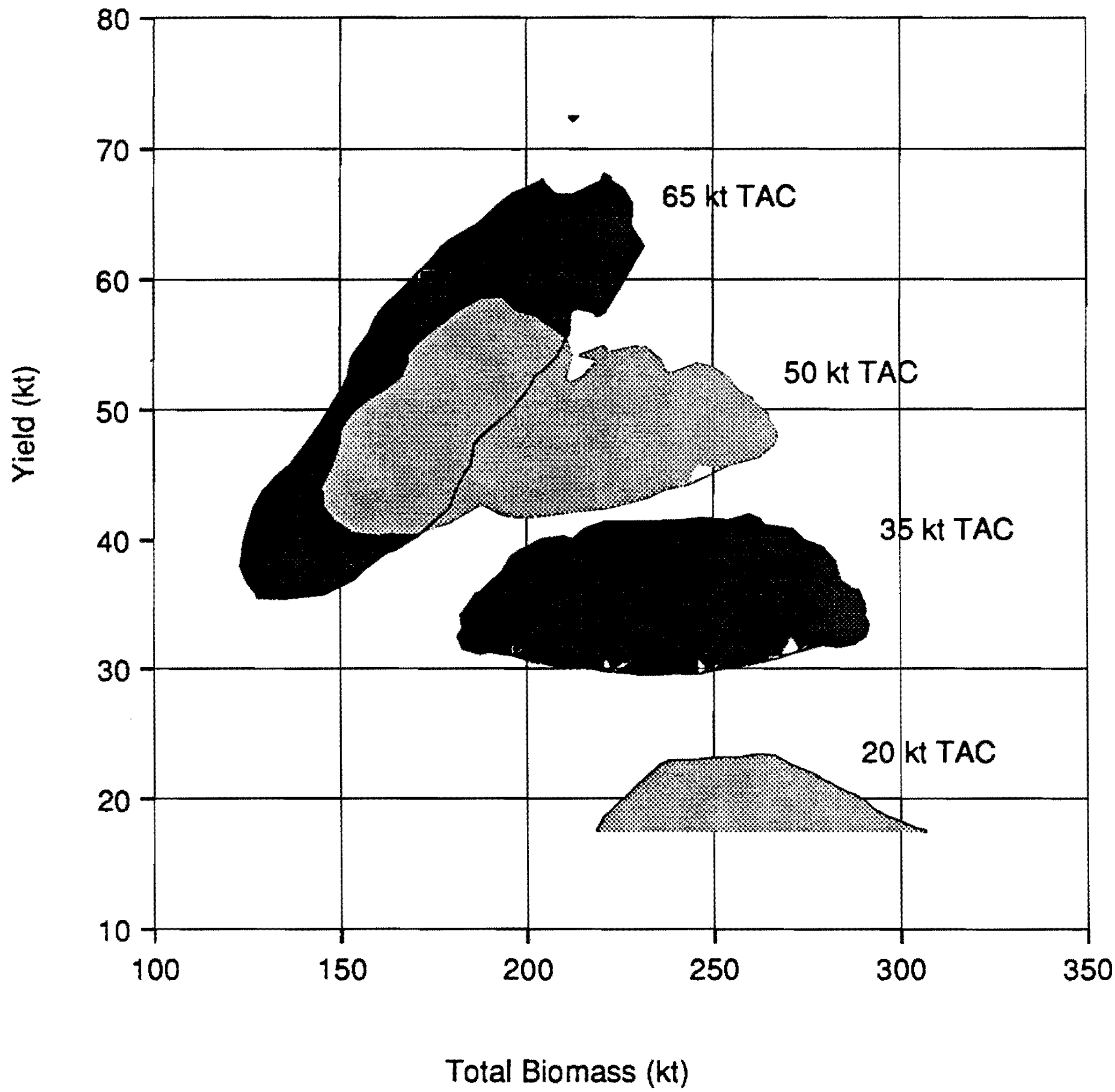

Figure 6. 50% contours after 3 years for 4 levels of TAC

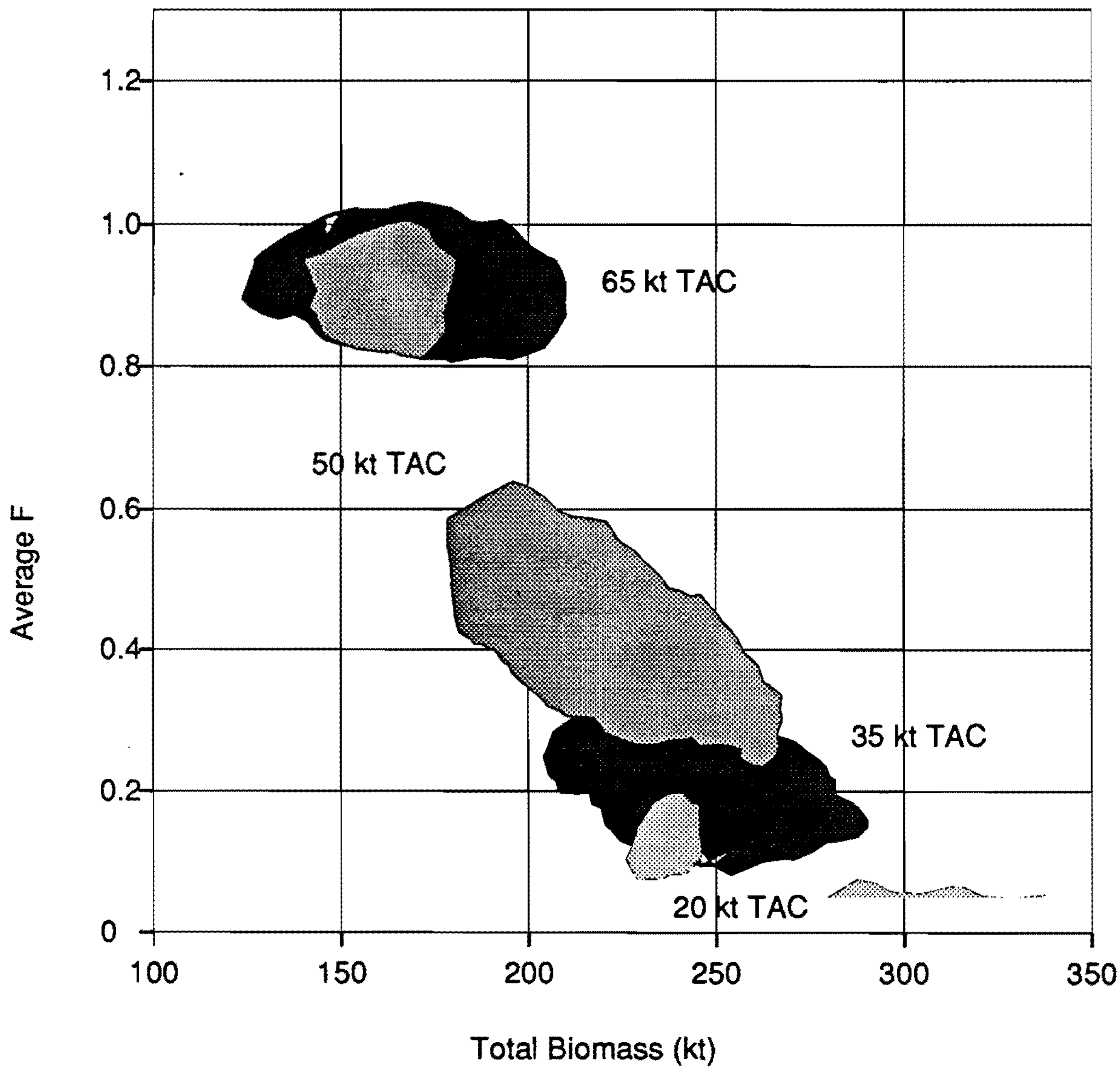

Figure 7 . 50% contours after 3 years for 4 levels of TAC.

Appendix A. Integration scheme.

The two-dimensional contour plots in this paper were produced by integrating two-dimensional discrete frequency distributions from most probable to least probable cell. This is done to identify the most probable region of final states from a given policy. The details of an integration of this sort in one dimension may be seen in Figures A1-3 and in Table A1. The data in this example is a frequency histogram distributed over 20 cells with a total of 100 events (Figure A1). The cells are sorted form the largest to the smallest (3rd column of Table A1) The three most probable cells are label 16, 13 and 11 in Figure A1 and the first three integrated (accumulated) values (0.16, 0.29 and 0.40) are identified in Figure A2. The integrated values are inverted by subtraction from 1 and the results are shown in Figure A3. This figure also shows the 50% contour level. The 4 cells which are equal to or larger than 50% are shaded. The shaded cells correspond to the darkest shaded areas in Figures 1 and 2 of the text. In either case, the shaded portion is the area of the most probable 50% of final states as defined by the relevant frequency histogram.

Table A1. Example data (columns 1 and 2) which are then sorted, (columns 3 and 4) integrated (column 5) and inverted (column 6).

Cell #	Freq.	Sorted	Cell #	Integrate	Invert
1	0	16	11	16	84
2	2	13	9	29	71
3	0	11	12	40	60
4	1	10	10	50	50
5	2	9	8	59	41
6	5	7	7	66	34
7	7	7	13	73	27
8	9	6	14	79	21
9	13	5	6	84	16
10	10	4	15	88	12
11	16	3	17	91	9
12	11	2	2	93	7
13	7	2	5	95	5
14	6	2	16	97	3
15	4	1	4	98	2
16	2	1	19	99	1
17	3	1	20	100	0
18	0	0	1	100	0
19	1	0	3	100	0
20	1	0	18	100	0

Figure A1. Frequency Distribution

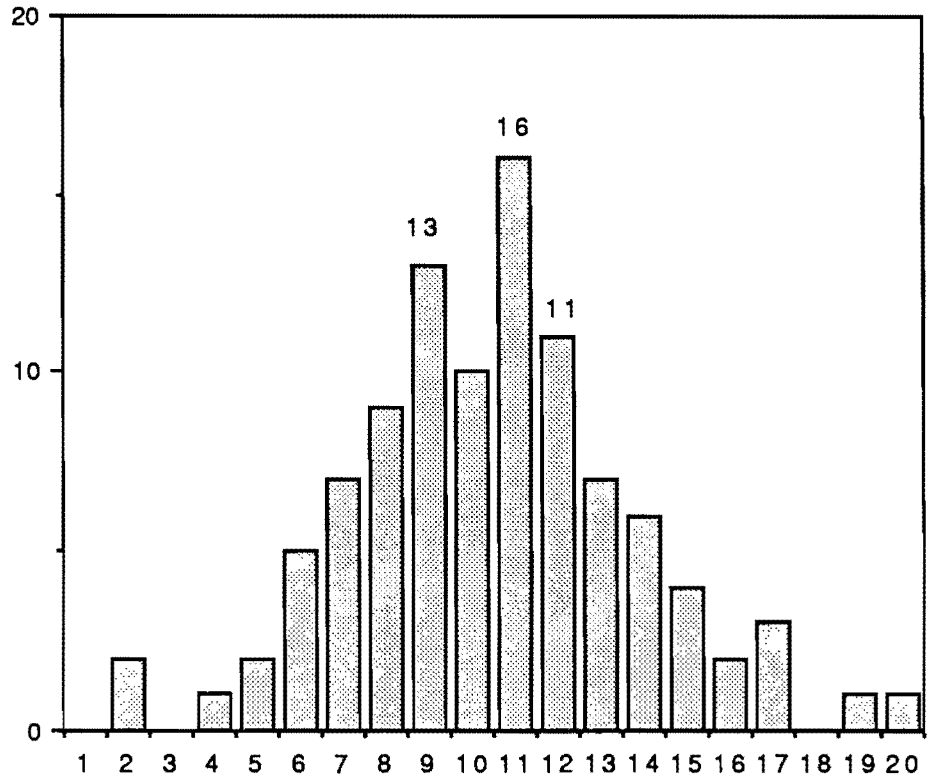

Figure A2. Integrate from most frequent down

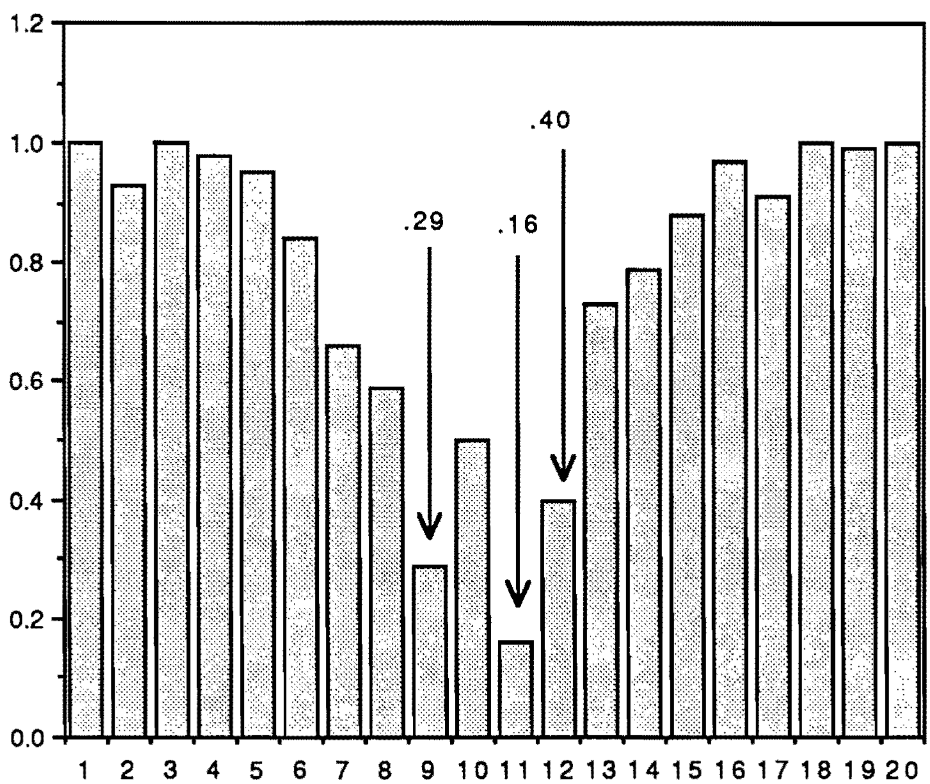

Figure A3. Inverted, integrated distribution with 50% contour level.