Northwest Atlantic Fisheries Organization

Serial No. N1848

NAFO SCR Doc. 90/112

SCIENTIFIC COUNCIL MEETING - SEPTEMBER 1990

SPECIAL SESSION ON MANAGEMENT UNDER UNCERTAINTIES
RELATED TO BIOLOGY AND ASSESSMENTS

Uncertainty Associated With Cod-Capelin Interactions: How Much is Too Much?

by

Peter A. Shelton, Lenore Fahrig and Russell B. Millar

Department of Fisheries and Oceans, P. O. Box 5667
St. John's, Newfoundland, Canada A1C 5X1

Abstract

Previous analyses have shown that cod annual length increments off Newfoundland and Labrador are positively related to temperature anomaly and 1/(cod biomass). It has also been shown that replacing 1/(cod biomass) with (capelin abundance index)/(cod biomass) results in a much weaker fit. A randomization test supported the conclusion that there was insufficient evidence to reject the null hypothesis that cod length increments are independent of capelin abundance. In this paper the probability of type II error (not rejecting the null hypothesis when it should have been) in the earlier analyses is examined by simulating fake length increment data from a growth model with error and carrying out similar fits to those attempted on the real data. It is argued that the cost of type II error could be large if it precipitates an increase in commercial exploitation of capelin.

Introduction

The influence of population density, the physical environment and interaction with other species (food organisms, competitors and predators) have important influences on population growth rate through such processes as body growth, reproduction and survival. Where such effects can be detected and the relationships defining them are reasonably stationary (unlikely), they may have importance in yield projections and resource management in general. For example, if the annual length increment of cod is inversely related to cod density, then, for rebuilding stocks, yield projections that do not take this into account may be over-optimistic.

In a previous analysis of cod growth off Newfoundland and Labrador, Akenhead et al. (1982) calculated length increments by subtracting individual cohort mean lengths from the value for the succeeding year for length at age data obtained for age 4 to 8 fish from the inshore cod trap fishery in Div 2J3KL. They found no significant relationship between annual length increment averaged across ages and indices of temperature and capelin abundance for the region. Elsewhere Beacham (1983) found that growth rates for the southern Gulf of St. Lawrence cod aged 5 to 8 were positively correlated with water temperature and inversely correlated with stock biomass. Magnússon and Pálsson (1989) report some evidence of a delayed decrease in cod growth rate during a period of low capelin abundance off Iceland, and for the Barent Sea the conventional wisdom is that cod growth rates were lower following the decline of the capelin in the late 1970's and early 1980's.

In a re-analysis of data for the Newfoundland-Labrador region, Millar and Myers (1990) found highly significant relationships between cod growth, cod biomass and temperature. Fahrig et al. (unpublished ms) developed a multispecies version of this model by replacing the 1/cod term with capelin/cod, where capelin refers to a capelin index of abundance (Fig. 5), assembled from several sources and considered to be the best representation of the past history for Divs. 2J3KL (Carscadden pers. comm.). They found that the multispecies cod growth model had a higher residual sums of squares than the single species growth model. A randomization test was performed in which the fit obtained from the multispecies model was compared to 100 fits to randomly permuted capelin biomass data. Of the fits to permuted data, 68% were as good or better than that obtained by fitting the multispecies model to the correctly sequenced capelin biomass data. The authors therefore concluded that they were unable to reject H_0, the null hypothesis, that capelin abundance influenced cod growth rate.

Two explanations exist. Either there really is no relationship, or error obscures the relationship. Cod off Newfoundland and Labrador do feed extensively on capelin and appear to have limited ability to compensate for a low abundance of capelin by eating relatively more other species (Lilly 1989). It therefore seems likely that capelin abundance could influence cod growth over some range of capelin abundance. We therefore decided to examine the potential role of error in obscuring the relationship. Recently Peterman (1990) has drawn attention to type II error problems in fisheries research and management. Type II error occurs when H_0 is not rejected when it should have been. This may be a consequence of small sample size and/or large error. Type II error may be costly and Peterman (1990) recommends steps such as Monte Carlo simulation to explore the circumstances under which Type II error may occur. As an example he cites de la Mare (1984) who used simulation to show that variability in catch rates could obscure trends in the abundance of certain whale stocks. In this study we follow this approach and use simulation to examine whether error could obscure a relationship in the existing time series. We then explore the influence of more data and more accurate data on the ability to distinguish a relationship.

The model

Of the several environmentally dependent von Bertalanffy models presented and discussed by Millar and Myers (1990) we chose to concentrate on their Model 2 because this obtained the best fit to their data. In this model, L_∞ is a function of the environment, in this case relative food abundance, modelled as 1/cod biomass, and temperature anomaly. Cod biomass was calculated by Millar and Myers (1990) from Adapt estimates of numbers at age for 3+ fish in Divs. 2J3KL and weights at age from the commercial catch in the same area. The temperature anomaly data was derived from bottom measurements taken at Station 27 close to St John's. Length data came from survey trawls. We concentrate on the fit to the length data for Div. 2J which is surveyed in fall.

The Millar/Myers Model 2 is

$$L_t = L_\infty \ (1\text{-a}\, e^{-tk}) \qquad t = 1 \text{ or } 2 \qquad (1)$$

$$L_t = (L_{t-1} + (L_\infty(1 - b1\ (1/cod)\text{-}b2\ tanom) - L_{t-1})\ (1\text{-}e^{-k})) \quad t>2 \quad (2)$$

where L_t is the expected length at age t, cod is the cod biomass, tanom is the temperature anomaly, and b1, b2, a, k and L_∞ are parameters that have to be estimated. The observed length at age t (L'_t) is related to the expected length at age t in the following manner

$$L'_t = L_t + \varepsilon \qquad (3)$$

where ε is normally distributed error with mean = 0 and variance = σ^2. The estimation is performed by nonlinear weighted (by number of fish aged) least-squares (derivative-fee (DUD), SAS 1985). The original data from Div. 2J used by Millar and Myers (1990) to fit this model are shown in Figs. 1 to 4. The fit for Div. 2J showed b1 and b2 to be highly significant and explained 51% of the residual sums of squares from a conventional three parameter (a, k, L_∞) von Bertalanffy model fit (Millar and Myers 1990). The multispecies version of the model replaces equation (2) with

$$L_t = (L_{t-1} + (L_\infty(1 - b1\ (capelin/cod)\text{-}b2\ tanom) - L_{t-1})\ (1\text{-}e^{-k})) \qquad (4)$$

in which capelin is the index of capelin abundance shown in Fig. 5.

The simulation

In order to examine the probability of detecting a capelin effect in the existing data and the extent to which more data or more accurate data would increase statistical power, a simulation model was developed in which error is apportioned among various sources in some detail. As a standard it was assumed the the estimate of cod abundance in NAFO Div. 2J3KL has a coefficient of variation or cv = 0.3, that the capelin index has a cv = 0.4 and that the temperature anomaly has a cv = 0.3. The cv's for cod and capelin may be considered upper bounds while the cv for temperature anomaly is arbitrary. Analysis of the within year cv of length at age for Div. 2J fall surveys (Fig. 6) shows that this seldom exceeds 0.1, and this was introduced as sampling error for length at age. Error in age determination *per se* was not modelled.

Data for a standard run were simulated in the following manner:
(i) A time series of fake data, equal in length to the actual data, of cod biomass, capelin biomass index and temperature anomaly were drawn randomly from normal distributions with mean and variance estimated from the data in Figs. 3-5.
(ii) A fake number of fish aged in each age class was drawn randomly

from a poisson distribution with the expected value, λ = mean number of fish aged in each age class estimated from the data in Fig. 2.
(iii) The fake data were then used to generate expected lengths at age using equations (1) and the multispecies version of (2), with values of L_∞=102.939 , a=0.896 , k=0.115, b1=-0.002 and b2=-0.476, the parameter estimates obtained for the fit of the multispecies model to data for Div. 2J (Fahrig et al. unpublished).

Fitting the fake data was simulated as follows:
(i) Mean length at age was simulated by taking the fake number of fish aged and multiplying the fake expected length of each fish by an error drawn randomly from a normal distribution with mean=1 and cv=0.1, reflecting the variability shown in Fig. 6, and then calculating the mean length at age.
(ii) Error in the biomass of cod, the capelin abundance index and the temperature anomaly were simulated by multiplying the individual fake values by an error drawn randomly from normal distributions with mean =1 and cv = 0.3 for cod, cv = 0.4 for capelin and cv = 0.3 for temperature anomaly for the "standard" run.
(iii) Error not accounted for above was introduced as random additive error following equation (3), so that the average simulated weighted residual sums of squares (wrss) from 100 simulations with the single species model equalled that obtained by Millar and Myers (1990).
(iv) The values generated by steps (i) - (iii) were then used to re-estimate the original parameters using the non-linear fitting procedure described above, and to determine the simulated wrss.

This procedure was repeated 100 times to provide a frequency distribution of wrss. Ratkowsky (1983) recommends that between 500 and 1000 random data sets be used in simulations of non-linear models but this was not practical in the present study because of excessively long run-times, and the results must therefore be treated with due caution. After 100 runs with the fit performed on the actual capelin time series used in the multispecies model to generate length at age values, the procedure was repeated a further 100 times with fit data generated from a second, and therefore independent, random, capelin biomass time series drawn from the same distribution as the first. This is in keeping with approach of Fahrig et al. (unpublished ms) who compared the fit obtained with the actual data to the fit with randomly permuted capelin biomass index. The probability of type II error (β) was assessed directly from the two sets of 100 runs by enumerating the number of runs with actual capelin that had wrss values greater than the fifth highest wrss value for the runs with random capelin.

In order to have the same mean wrss value obtained for the fit of the single species model to the actual data, about 40 000, the additive error ε was assigned a value of σ = 2.2 cm. This is variability not accounted for by the model or by error in cod biomass, temperature anomaly and within-year variability in length at age in the population.

To determine whether a capelin effect could be expected to be detected in the original data or by improving the data, 12 simulations in 6 pairs were carried out:

1. Standard, actual capelin:- replicates the real data set with additive process error σ=2.2, and with error cv's on cod biomass, capelin abundance index and temperature anomaly of

0.3, 0.4, and 0.3 respectively, as well as a cv of 0.1 on length at age. These cv's were considered to be upper bounds. The λ for the number of fish aged in each age class was set equal to the mean from the survey data.

2. Standard, random capelin:- As above but with a second, random capelin time series used in the fit.
3. Accurate, actual capelin:- cv's on cod biomass, capelin abundance index and temperature anomaly all reduced to 0.2.
4. Accurate, random capelin:- as in 3 but with random capelin used in the fit.
5. Accurate, intense, actual capelin:- as in 3 but with the mean number aged per age class increased by a factor of 5.
6. Accurate, intense, random capelin:- as in 5 but with random capelin used in the fit.
7. Long, actual capelin:- as in 1 but with the time horizon increased by a further 16 years.
8. Long, random capelin:- as in 7 but with random capelin used in the fit.
9. Long, accurate, intense, actual capelin:- as in 8 but with mean number of fish aged per age class increased by a factor of 5.
10. Long, accurate, intense, random capelin:- as in 11 but with random capelin used in the fit.
11. Standard, zero additive error, actual capelin:- as in 1 but with ε=0.
12. Standard, zero additive error, random capelin:- as in 13 but with random capelin.

Results and discussion

Because of error, the combination of observations of the real world (in this case cod biomass, capelin abundance, temperature anomaly, and mean length of cod at age over several years), is just one realization from a large set of possible combinations of values. If this realised combination of values does not provide a significantly better fit to the model than that obtained by replacing capelin in the fit data with random capelin, (i.e. independent of the measured cod length at age), then the H_0 of no capelin effect cannot be rejected. This was the approach adopted by Fahrig et al. (unpublished ms) for the available data for Divs. 2J3KL. The simulation results reported here attempt to determine the probability of not rejecting H_0 when it should have been (type II error), and to examine whether better data would improve the probability of detecting a capelin effect. The study assumes that the multispecies version of the Millar/Myers cod growth model is the true system and uses it to generate fake data with error which is not explained by the model, which is observed with further error. Fitting the model with no error returns the original parameter values and wrss=0 (a useful test that the program, but not necessarily the logic, is free of bugs).

Simulation 1 ("standard") is an attempt to replicate the situation pertaining to Div. 2J, with upper estimates of the error on cod biomass, capelin abundance and temperature anomaly (Fig. 7). When the wrss value is compared to that obtained from simulation 2 (random capelin), the results are not encouraging. There is a 0.96 probability of making type II error (Table 1). Further simulations explored whether the probability of making type II error could be reduced by having more or better data (Table 1, Figs 8-12). In simulation 3 ("accurate") the cv's of the error on cod biomass, capelin abundance and temperature anomaly have been set at what may be considered lower estimates, and the probability of type II error is reduced to 0.92. If, in addition, the mean number of fish aged in each age class is increased by a factor of 5 (simulation 5, "accurate intense"), the probability of type II error is further reduced to 0.87. If only the length of the data series is increased (a further 16 years data added, "long"), the probability of type II error is no lower than in the original data (simulation 7). For the "long, accurate, intense" runs (simulation 9) the probability of type II error is 0.95. Finally, a run was carried out with the same error used for the "standard" run but with $\varepsilon = 0$ ("standard, zero additive error", simulation 11), giving a probability of type II error of 0.93.

The conclusion from the simulations is that the error, or that part of the data that cannot be explained by the model, is such that there is not sufficient statistical power to reliably detect a capelin effect. The importance of multispecies effects in fisheries management is often emphasized (e.g. Kerr and Ryder 1989, Pope 1989). If the outcome obtained in this study is found to generally hold in other studies of multispecies effects, then the expectation that multispecies models of the kind discussed here can improve over single species models will not be fulfilled. Error will not allow multispecies effects to be reliably detected in the data. Where an important resource such as cod may be influenced by management action, such as increased harvests of capelin, a conservative attitude should be adopted even when attempts to detect an effect of capelin on cod fail, because the expectation is that they will fail as a result of error.

References

Beacham, T. D. 1983. Growth and maturity of Atlantic cod (*Gadus morhua*) in the southern Gulf of St Lawrence. Can. Tech. Rep. Fish. Aquat. Sci. 1142, 35p.

De la Mare, W. K. 1984, On the power of catch per unit effort series to detect declines in whale stocks. Rep. Int. Whaling Comm. 34:655-662.

Fahrig, L., Shelton, P. A. and Millar, R.B. unpublished ms. An approach to Newfoundland Boreal multispecies modelling. Working paper presented at the ICES Multispecies Working Group Special Meeting held in Bergen, Norway, 23-27 April 1990. 18p.

Kerr, S. R. and Ryder, R.A. 1989. Current approaches to multispecies analyses of marine fisheries. Can. J. Fish. Aquat. Sci. 46:528-534.

Lilly, G. R. 1989. Inter-annual variability in predation by Atlantic cod (*Gadus morhua*) on capelin (*Mallotus villosus*) and other prey off southern Labrador and northeastern Newfoundland in autumn. ICES 1989 MSM Symp./No. 17, 28p.

Magnússon, K. G. and Pálsson, O. K. 1989. Predator-prey interactions of cod and capelin in Icelandic waters. Paper presented at the ICES 1989 MSM Symp.

Millar, R. B. and Myers, R. A. 1990 Modelling environmentally induced change in size and age for Atlantic Canada cod stocks. CAFSAC Res. Doc. No. __ , __ p.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research management. Can. J. Fish. Aquat. Sci. 47:2-15.

Pope, J.G. 1989. Multispecies extensions to age-structured assessment models. Am. Fish. Soc. Symp. 6:102-111.

Ratkowsky, D. A. 1983. Nonlinear regression modeling. A Unified Practical Approach. Marcel Dekker, INC., New York, p276.

SAS. 1985. User's Guide : Statistics, Version 5 Edition. SAS Institute Inc., Cary, NC, USA. 956p

Table 1

Minimum, maximum, mean and standard deviation of the weighted residual sums of squares obtained in the simulation runs. In each couplet, the first set of values refers to the runs with the actual fake capelin in the model and the second set to the runs with random capelin. The probability of type II error (β) obtained from a comparison of the runs with actual and random capelin is given. The sd's for runs 11 and 12 are not given because the distribution is skewed.

run	min	max	mean	sd	β
1	30854	65317	44112	7494	
2	28857	59300	43478	6959	0.96
3	26794	51704	39068	5916	
4	28481	59847	42079	6704	0.92
5	108146	275305	180597	27188	
6	142103	265609	192928	28464	0.87
7	97581	152818	118595	9697	
8	96557	163935	123156	13718	0.97
9	390789	669142	493581	46876	
10	369922	692853	527750	58804	0.95
11	14096	73007	32626		
12	15347	93308	38530		0.93

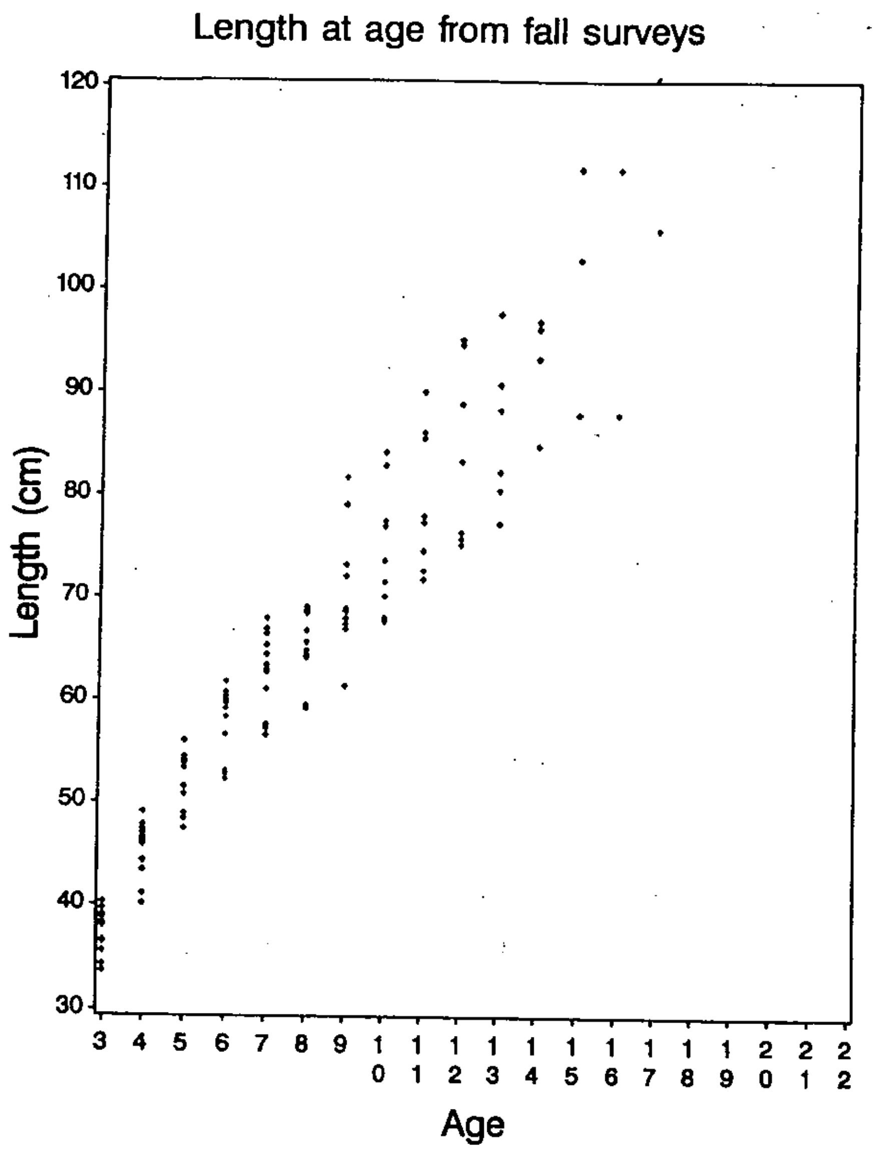

Fig. 1. Length at age in fall surveys conducted in NAFO Div. 2J, 1976-87.

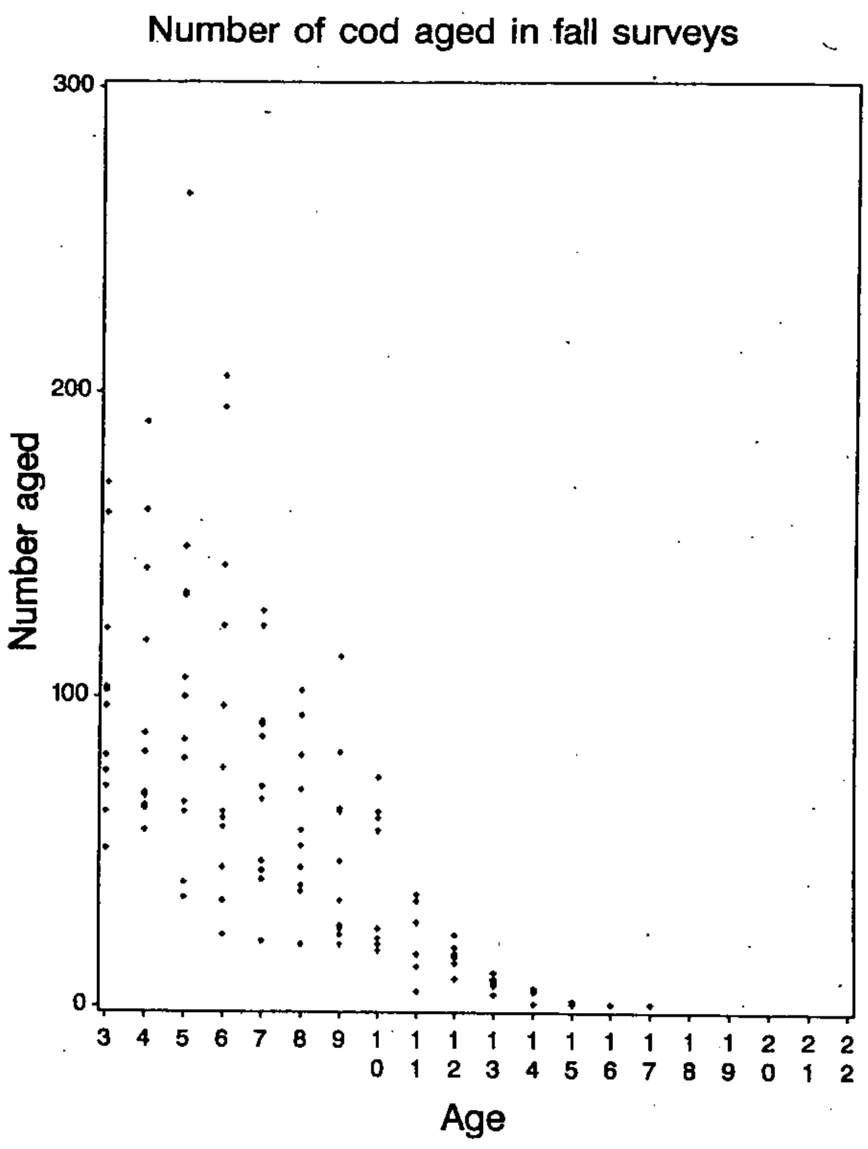

Fig. 2. Number of cod aged from samples collected in fall surveys in NAFO Div. 2J, 1976-87.

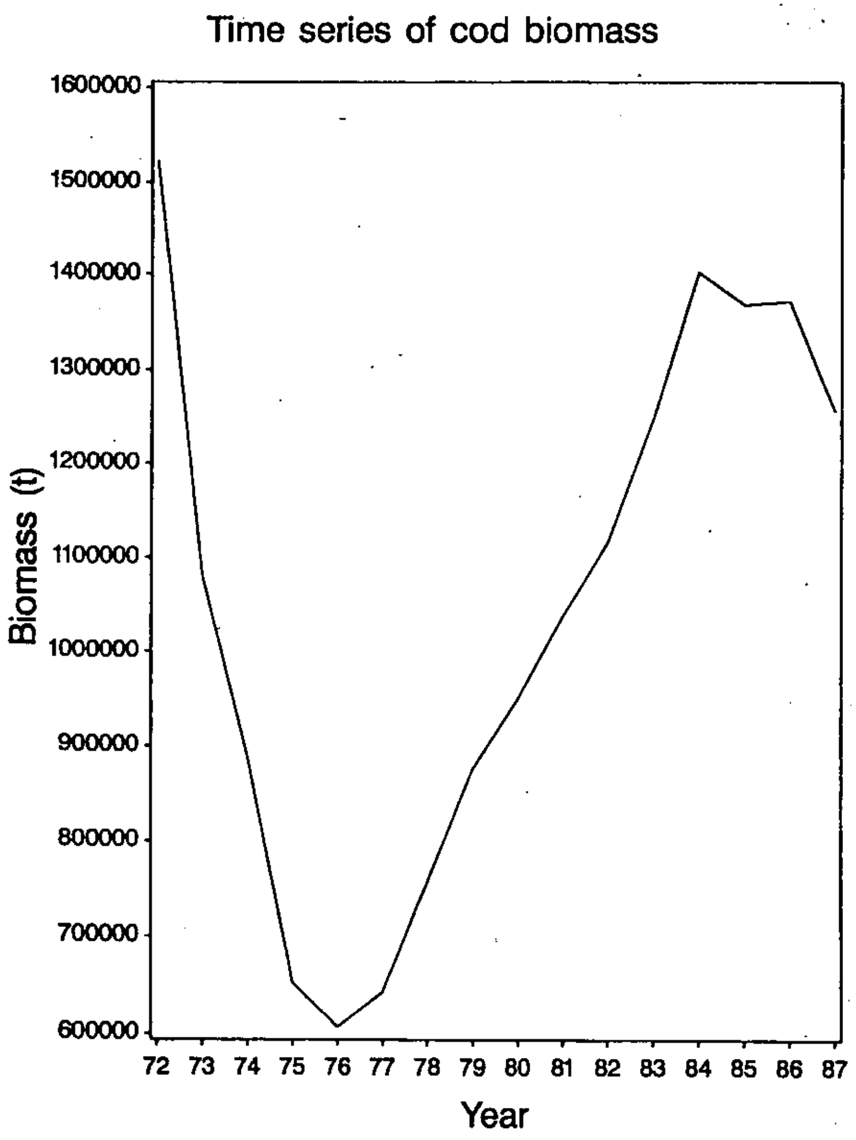

Fig. 3. Time series of 3+ cod biomass for NAFO Div. 2J3KL calculated using Adapt estimates of numbers at age and mean weight at age from the catch.

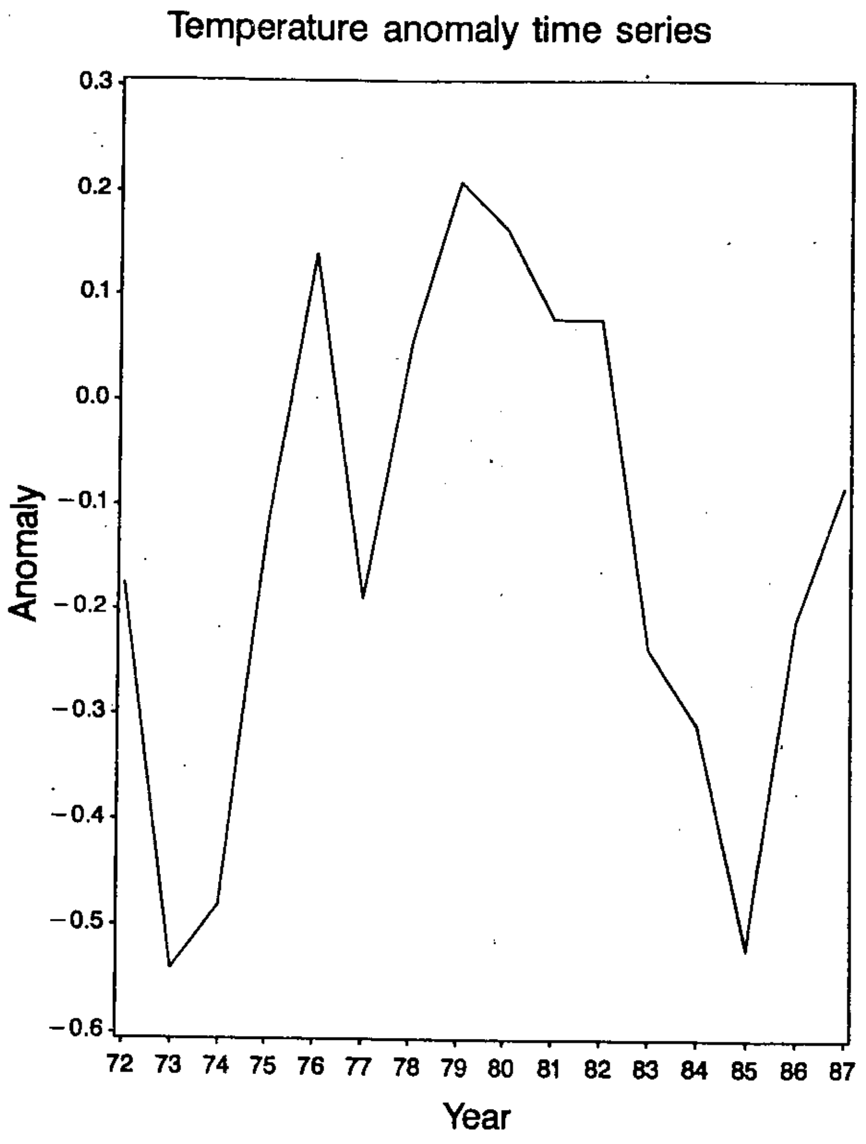

Fig. 4. Annual mean temperature anomaly time series for Station 27.

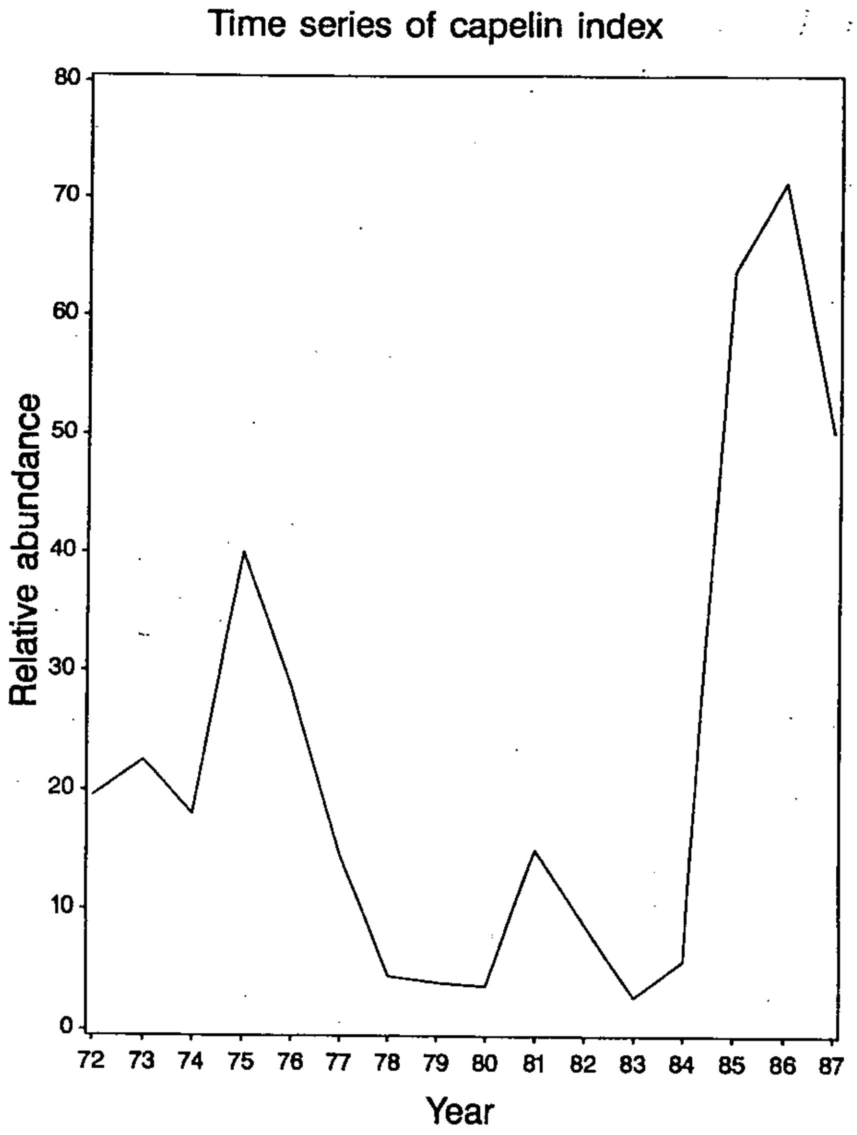

Fig. 5. Capelin abundance index assembled from several sources.

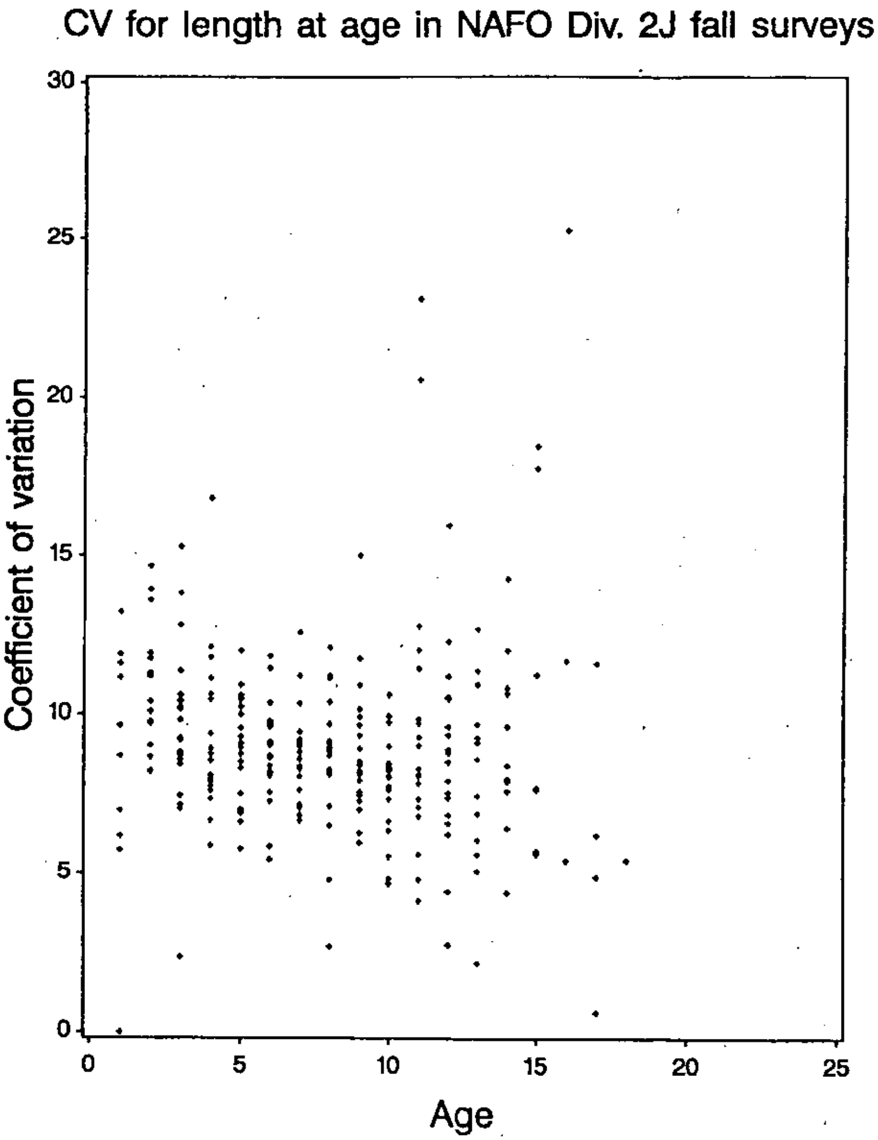

Fig. 6. Within-year coefficient of variation in length at age.

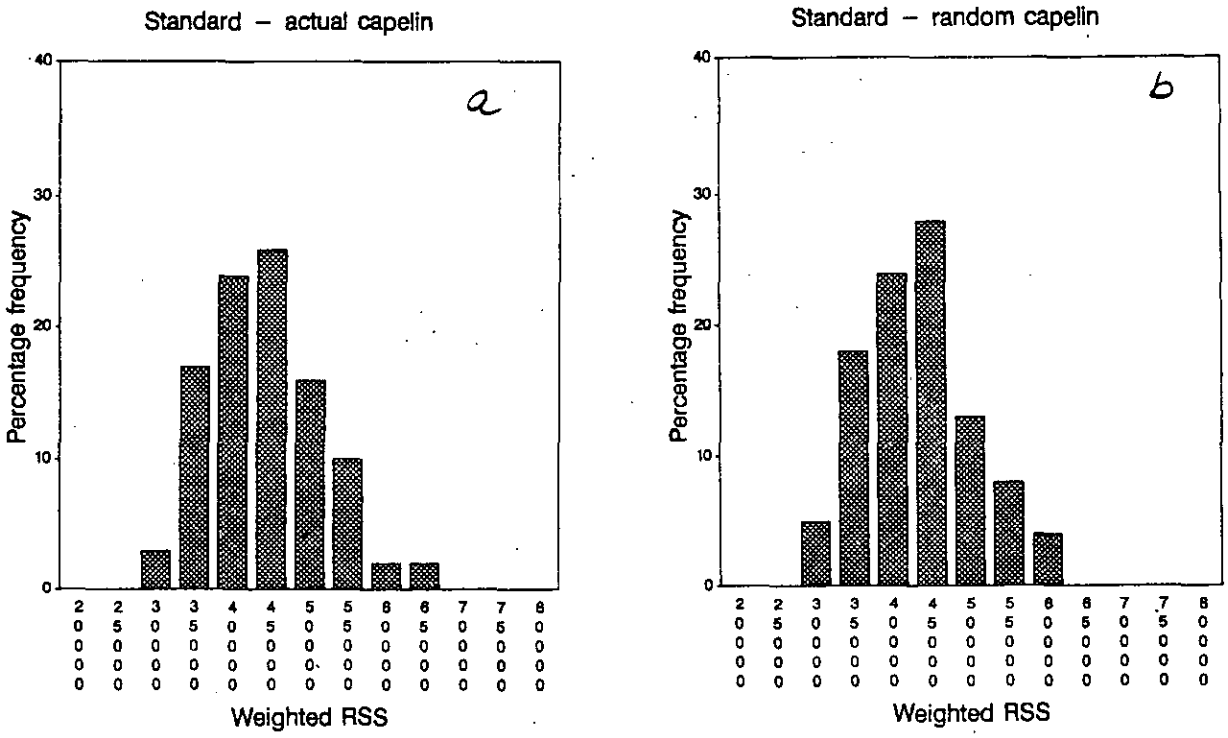

Fig. 7. a. Simulation 1. Standard, actual capelin:- replicates the real data set with additive process error σ=2.2, and with error cv's on cod biomass, capelin abundance index and temperature anomaly of 0.3, 0.4, and 0.3 respectively as well as a cv of 0.1 on length at age. These cv's were considered to be upper bounds. The lambda for the number of fish aged in each age class was set equal to the mean lambda from the survey data.

b. Simulation 2. Standard, random capelin:- Asin simulation 2 but with a second, random capelin time series independent of cod growth used in the fit.

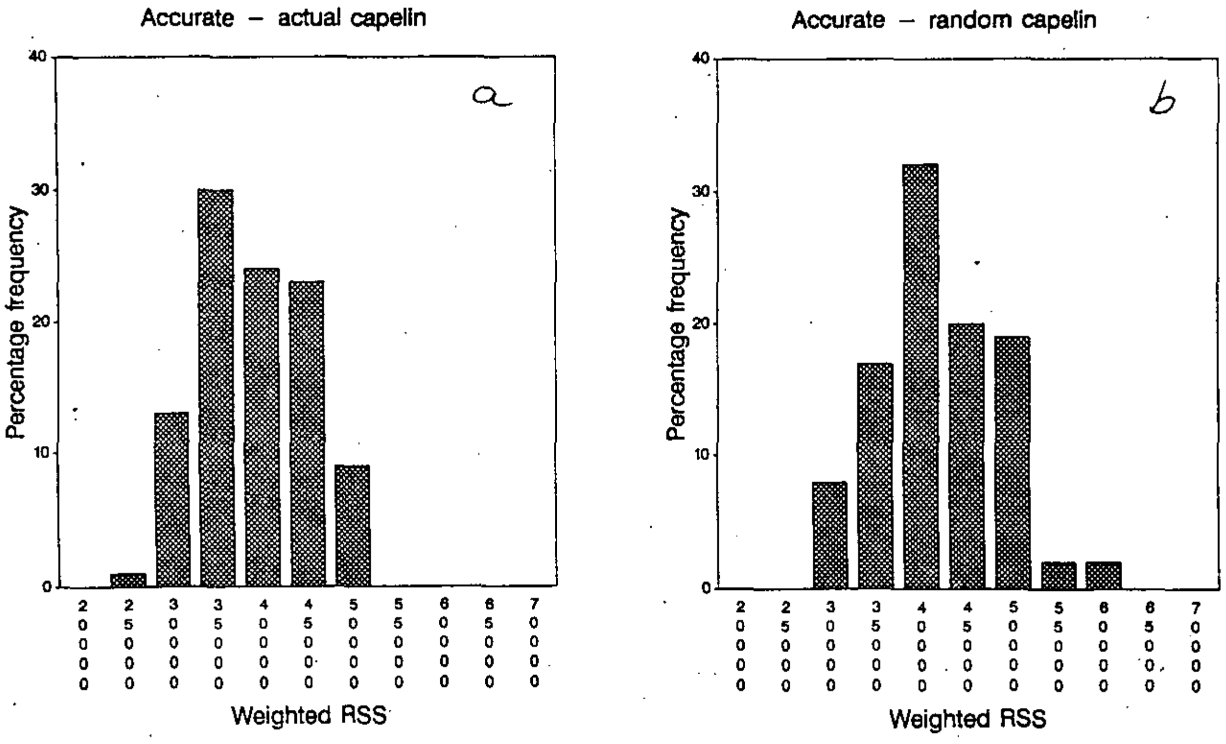

Fig. 8. a. Simulation 3. Accurate, actual capelin:- cv's on cod biomass, capelin abundance index and temperature anomaly all reduced to 0.2.

b. Simulation 4. Accurate, random capelin:- as in simulation 3 but with random capelin used in the fit.

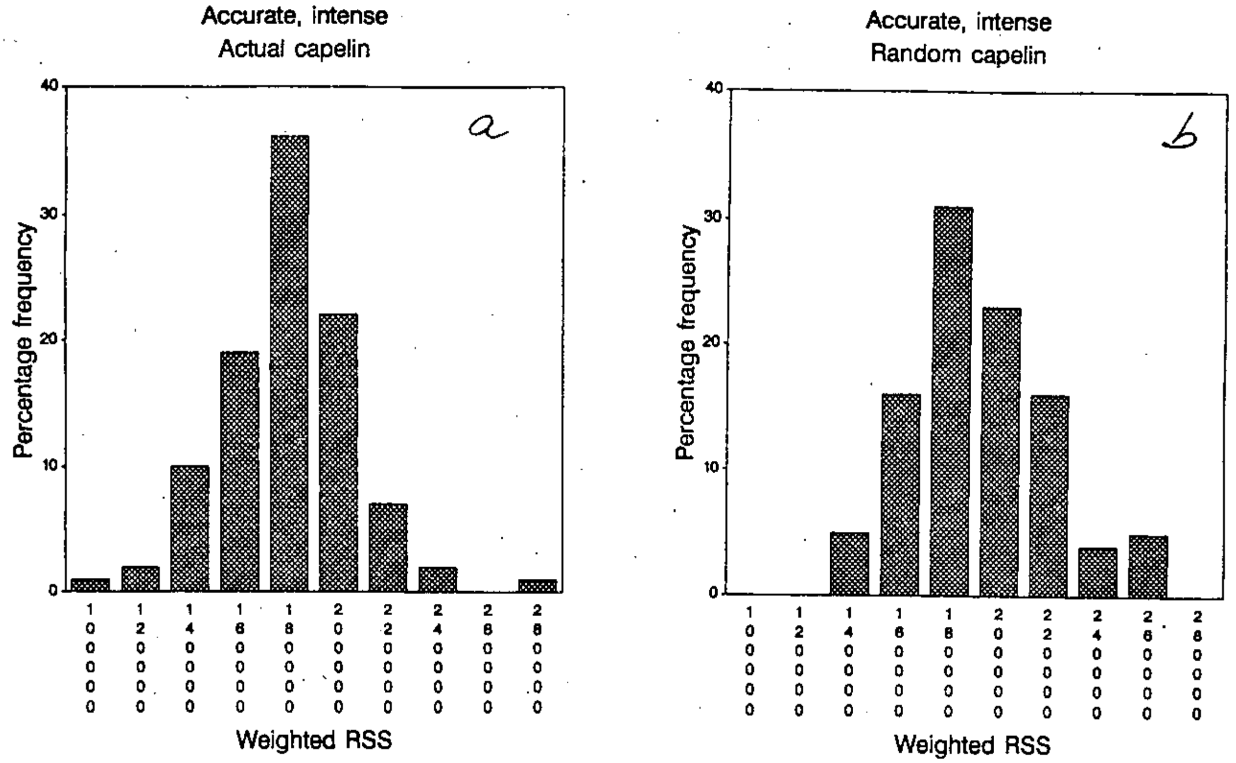

Fig. 9. a. Simulation 5. Accurate, intense, actual capelin:- as in simulation 3 but with the mean number aged per age class increased by a factor of 5.
b. Simulation 6. Accurate, intense, random capelin:- as in simulation 5 but with random capelin used in the fit.

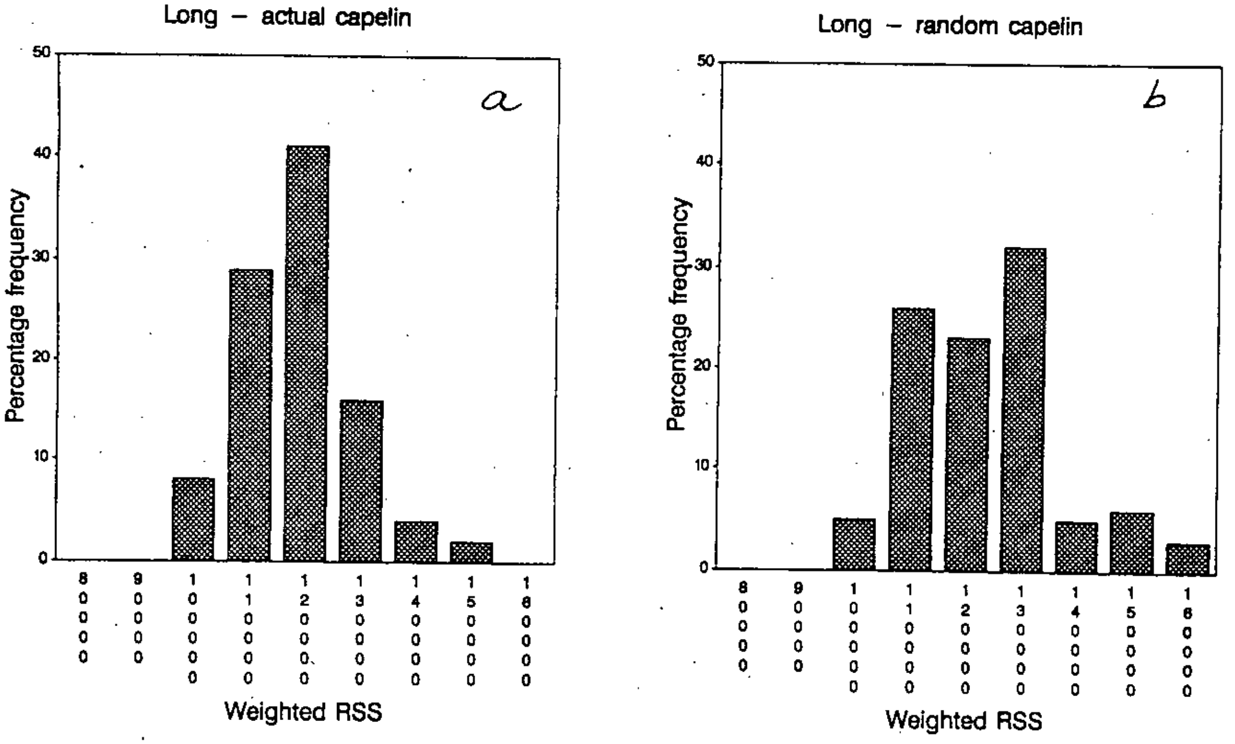

Fig. 10. a. Simulation 7. Long, actual capelin:- as in simulation 1 but with the time horizon increased by a further 16 years.
b. Simulation 8. Long, random capelin:- as in simulation 7 but with random capelin used in the fit.

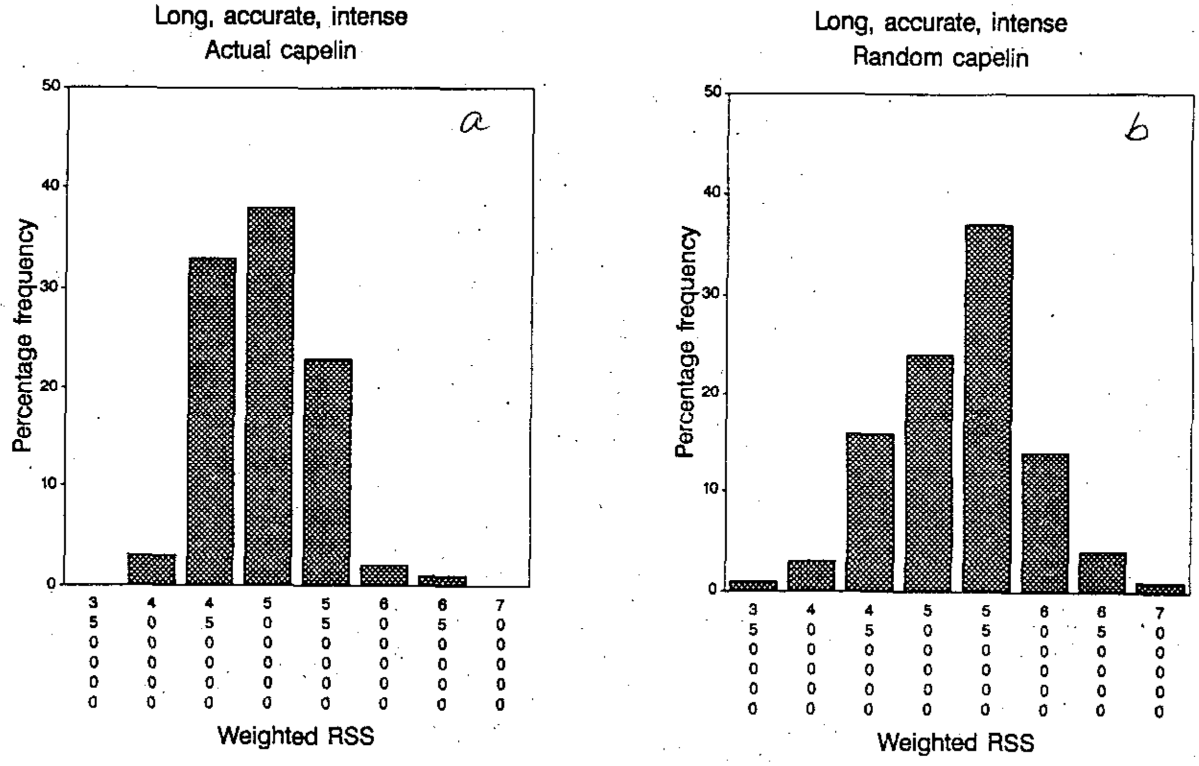

Fig 11. a. Simulation 11. Long, accurate, intense, actual capelin:- as in simulation 8 but with mean number of fish aged per age class increased by a factor of 5.
b. Simulation 12. Long, accurate, intense, random capelin:- as in simulation 11 but with random capelin used in the fit.

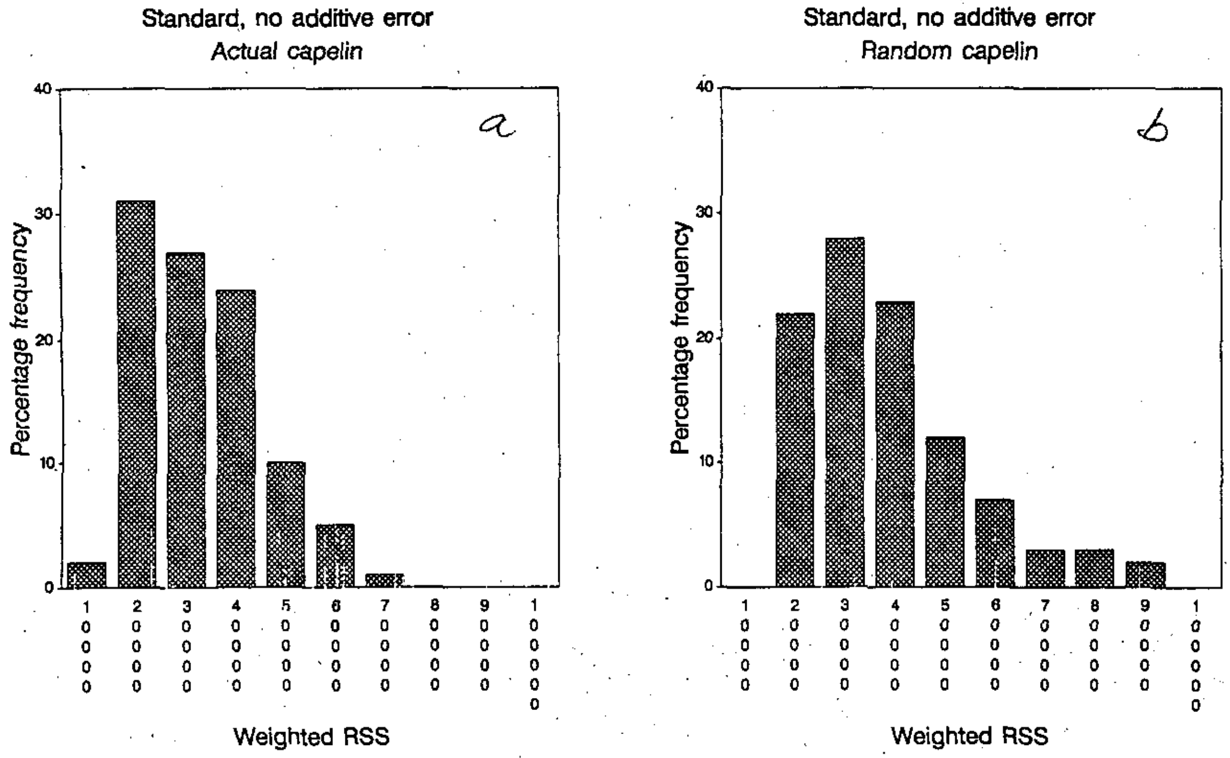

Fig. 12. a. Simulation 13. Standard, zero additive error, actual capelin:- as in simulation1 but with $\varepsilon=0$.
b. Simulation 14. Standard, zero additive error, random capelin:- as in simulation *13 but with random capelin.*