NAFO Sci. Coun. Studies, 16: 13–19

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

Peter A. Shelton, Lenore Fahrig and Russell B. Millar
Department of Fisheries and Oceans, Science Branch
P. O. Box 5667, St. John's, Newfoundland, Canada A1C 5X1

Abstract

Previous analyses have shown that annual length increments of cod off Newfoundland and Labrador are positively related to temperature anomaly and to 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 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 NAFO Div. 2J+3KL. 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) reported some evidence of a delayed decrease in cod growth rate during a period of low capelin abundance off Iceland, and for the Barents Sea the conventional wisdom is that cod growth rates were lower following the decline of the capelin in the late-1970s and early-1980s.

In a reanalysis of data for the Newfoundland-Labrador region, Millar and Myers (MS 1990) found highly significant relationships between cod growth, cod biomass and temperature. Fahrig, *et al*. (unpublished manuscript)[1] 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, assembled from several sources and considered to be the best representation of the past history for Div. 2J+3KL (J. Carscadden, Northwest Atlantic Fisheries Centre, St. John's, Newfoundland, 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 of 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_o, the null hypothesis, that capelin abundance did not influence cod growth rate.

Two explanations exist. Either there really is no relationship, or error obscures the relationship. Cod off

---

[1] Fahrig, L., P. A. Shelton and R. B. Millar. 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, 18 p.

Newfoundland and Labrador do feed extensively on capelin and they appear to have limited ability to compensate for a low abundance of capelin by eating relatively more of 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_o 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 this approach is followed and simulation is used to examine whether error could obscure a relationship in the existing time series. The influence is then explored of more data and more accurate data on the ability to distinguish a relationship.

The Model

Of several environmentally dependent von Bertalanffy models presented and discussed by Millar and Myers (MS 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 (MS 1990) from ADAPT estimates of numbers-at-age for 3+ fish in Div. 2J+3KL and weights-at-age from the commercial catch in the same area. The temperature anomaly data was derived from bottom measurements taken at the standard hydrographic Station 27 close to St. John's. Length data was from survey trawls. We concentrate on the fit to the length data for Div. 2J which is surveyed in autumn.

The Millar/Myers Model 2 is:

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

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

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:

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

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 (MS 1990) to fit this model are shown in Fig. 1 to 4. The fit for Div. 2J showed b1 and b2 to be highly significant and explained 51% of the residual sums of square from a conventional three parameter (a, k, L_∞) von Bertalanffy model fit (Millar and Myers, MS 1990). The multispecies version of the model replaces equation (2) with

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

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 that the

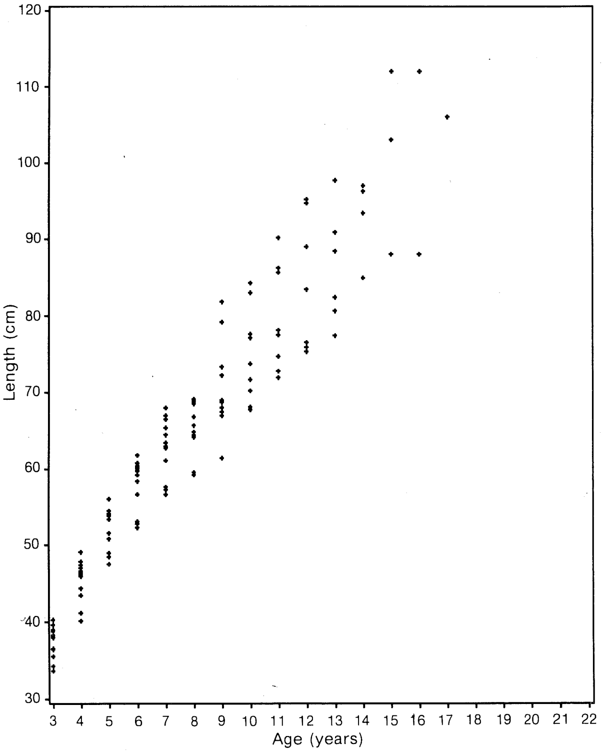

Fig. 1. Length-at-age in autumn surveys conducted in NAFO Div. 2J, 1976–87.

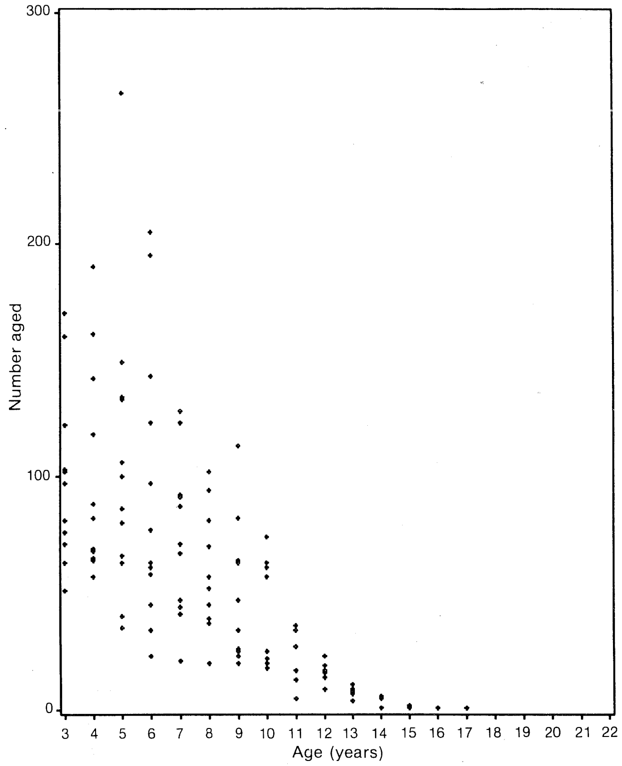

Fig. 2. Number of cod aged from samples collected in autumn surveys in NAFO Div. 2J, 1976–87.

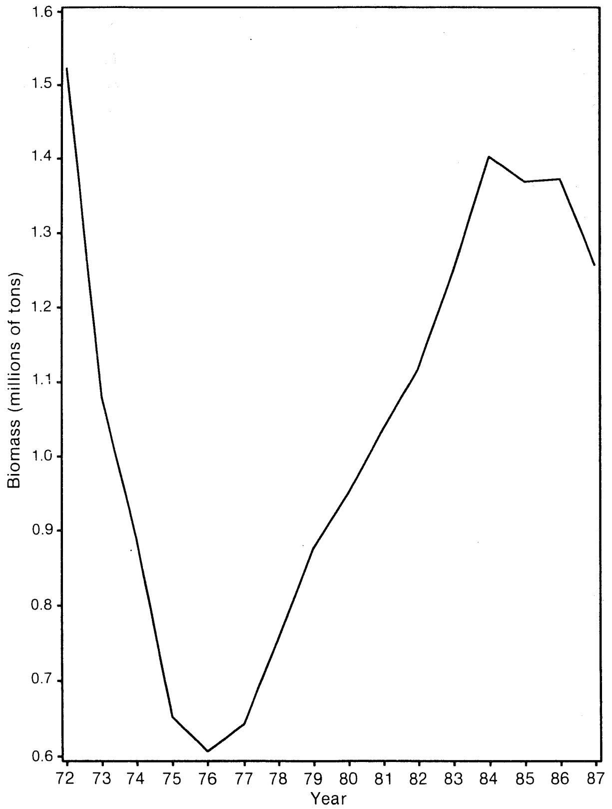

Fig. 3. Time series of age 3+ cod biomass for NAFO Div. 2J+3KL calculated using ADAPT estimates of numbers-at-age and mean weight-at-age from the catch.

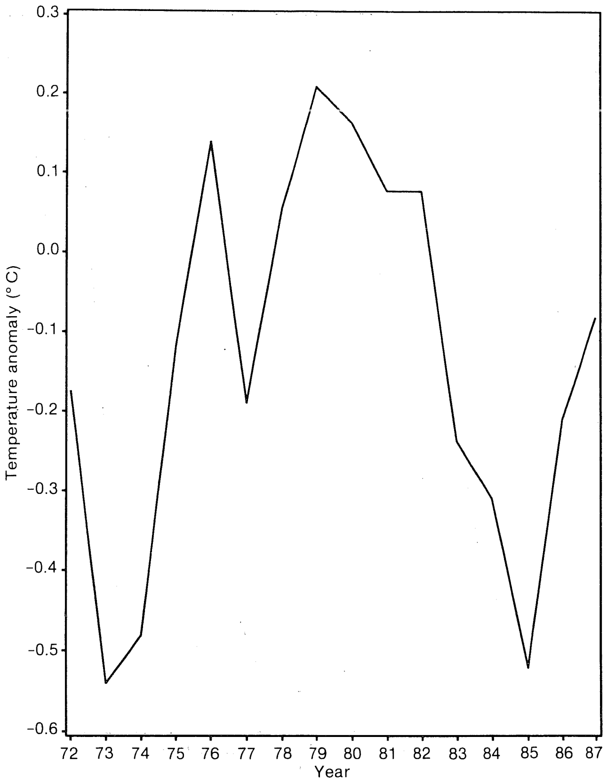

Fig. 4. Annual mean temperature anomaly time series for standard Hydrographic Station 27, off St. John's, Newfoundland.

estimate of cod abundance in Div. 2J+3KL 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 CVs 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 autumn 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 modeled.

Data for a standard run were simulated in the following manner:

i) A time series of simulated 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 Fig. 3–5.

ii) The 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 simulated data were then used to generate expected lengths-at-age using equations (1) and the multispecies version of (2), with values of $L_\infty = 102.939$, $a = 0.896$, $k = 0.115$, $b1 = -0.002$ and $b2 = -0.476$, the parameter estimates obtained for the fit of the multispecies model to data for Div. 2J (Fahrig *et al.*, unpublished manuscript)[1].

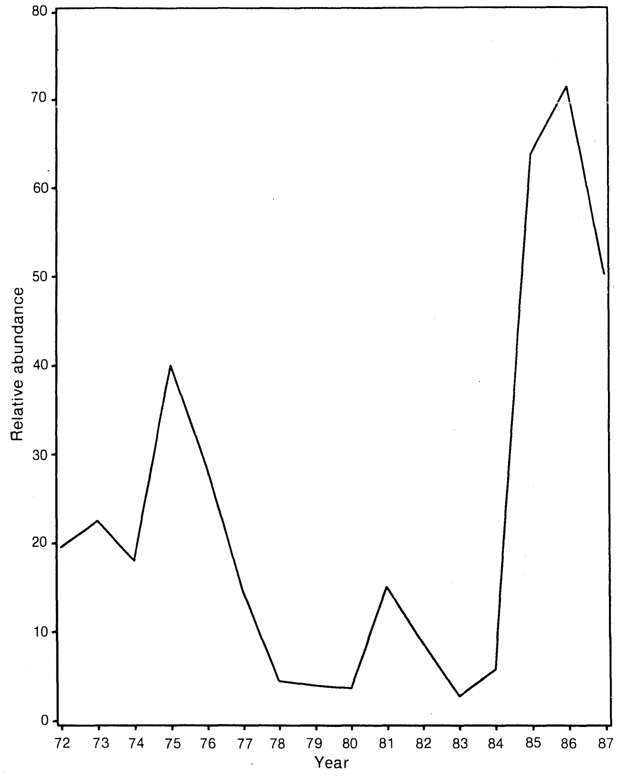

Fig. 5. Capelin abundance index assembled from several sources.

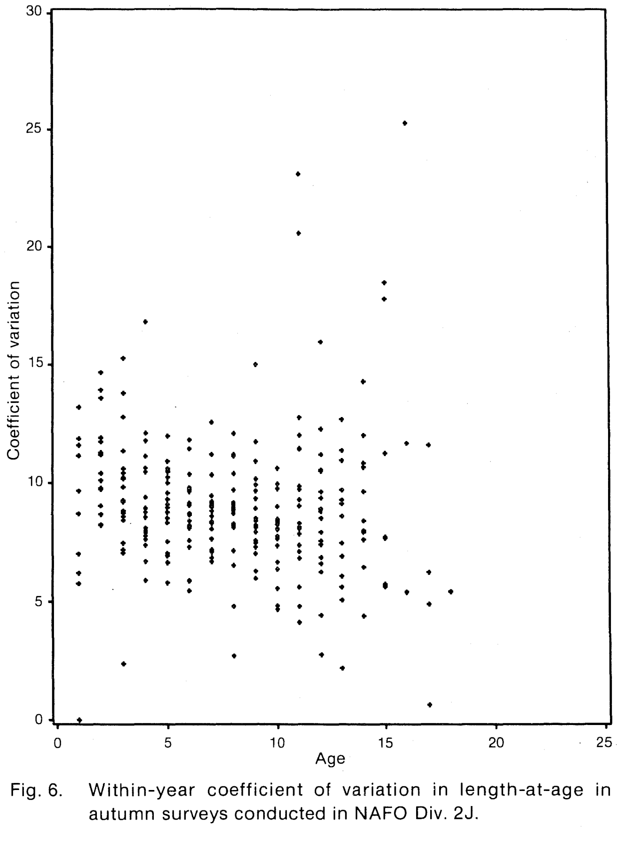

Fig. 6. Within-year coefficient of variation in length-at-age in autumn surveys conducted in NAFO Div. 2J.

Fitting the simulated data was carried out as follows:

i) Mean length-at-age was simulated by taking the simulated number of fish aged and multiplying the simulated 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 simulated 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 (MS 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 1,000 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 should 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. The null hypothesis, H_o: there is no difference between actual and random capelin fit WRSS, and the alternative hypothesis, H_a: the WRSS from the fit with actual capelin is less than the WRSS with random capelin, were then tested using the two sets of 100 simulated WRSS values. The level of significance of α was set at 0.05, thereby defining a critical WRSS value in the distribution obtained from 100 runs with random capelin. The probability of Type II error (β) was then assessed directly by finding the proportion of the 100 runs with actual capelin that had WRSS values greater than the critical (fifth lowest) WRSS value in the 100 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 $\sigma = 2.2$ cm. This is variability that is 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 data, 12 simulations in 6 pairs were carried out:

1. Standard, actual capelin: replicates the real data set with additive process error $\sigma = 2.2$, and with error CVs 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 CVs 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: CVs 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 8 but with random capelin used in the fit.
11. Standard, zero additive error, actual capelin: as in 1 but with $\varepsilon = 0$.
12. Standard, zero additive error, random capelin: as in 1 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 realized 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. second capelin time series 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 manuscript)[1] for the available data for Div. 2J+3KL. 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 simulated 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). In Simulation 3 ("accurate") the CVs 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 of data added; "long"), the probability of Type II error is no lower than in the original data (Simulation 7).

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 simulated 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 standard deviations for runs 11 and 12 were not estimated because the distribution is skewed.

Run	Minimum	Maximum	Mean	Standard deviation	β
1	30,854	65,317	44,112	7,494	
2	28,857	59,300	43,478	6,959	0.96
3	26,794	51,704	39,068	5,916	
4	28,481	59,847	42,079	6,704	0.92
5	108,146	275,305	180,597	27,188	
6	142,103	265,609	192,928	28,464	0.87
7	97,581	152,818	118,595	9,697	
8	96,557	163,935	123,156	13,718	0.97
9	390,789	669,142	493,581	46,876	
10	369,922	692,853	527,750	58,804	0.95
11	14,096	73,007	32,626		
12	15,347	93,308	38,530		0.93

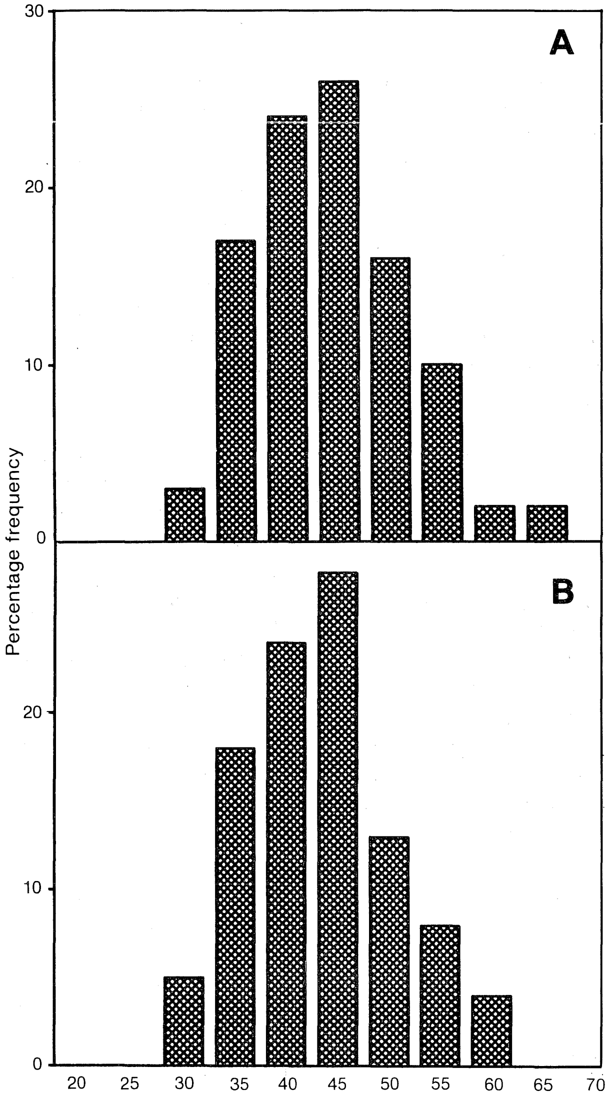

Fig. 7. (**A**) Simulation 1. Standard, actual capelin: replicates the real data set with additive process error $\sigma = 2.2$, and with error CVs on cod biomass, capelin abundance index and temperature anomaly of 0.3, 0.4, and 0.3 respectively, as well as a CV of 0.1 on length-at-age. These CVs were 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: as in Simulation 1 but with a second, random capelin time series independent of cod growth used in the fit.

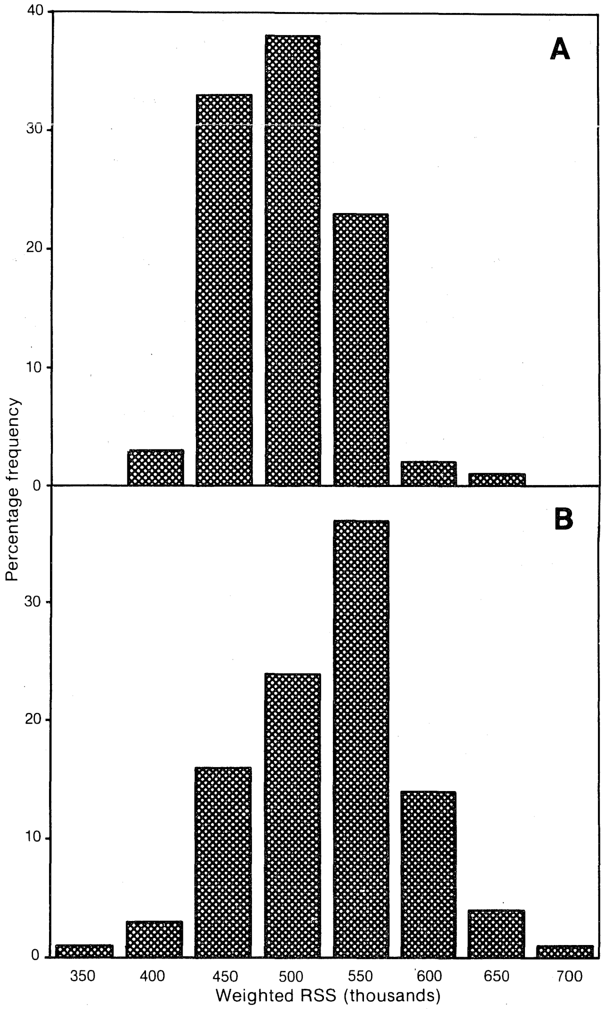

Fig. 8. (**A**) Simulation 11. Long, accurate, intense, actual capelin: as in Simulation 8 but with number of fish aged per age-class increased by factor 5. (**B**) Simulation 12. Long, accurate, intense, random capelin: as in Simulation 11 but with random capelin used in the fit.

For the "long, accurate, intense" runs (Simulation 9, Fig. 8) 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, under the conditions of the assumed model and error structure, and with α set at 0.05, a capelin effect cannot be reliably detected. Surprisingly, even substantial improvements to the data failed to have a large influence on β. Greater statistical power ($1-\beta$) can be obtained by increasing the value of α from the arbitrary traditional value of 0.05, and this may be appropriate for some fishery problems where being unable to reject H_0 has important management implications (Peterman, 1990). For example, failure to reject H_0 that capelin abundance has no effect on cod growth could lead to increased harvests of capelin followed by a decline in cod yield. At the very least, fisheries managers should be made aware of the statistical power of analyses on which important management decisions are based.

The importance of multispecies effects in fisheries managment is often emphasized (e.g. Kerr and Ryder, 1989; Pope, 1989) but there is little in the literature to indicate whether or not these effects can be reliably detected in available data. Some *a priori* analyses before new data collection or *a posteriori* analyses after failure to reject H_o, as advocated by Peterman (1990), may do much to place the analysis of multispecies effects on a more realistic footing.

References

AKENHEAD, S. A., J. CARSCADDEN, H. LEAR, G. R. LILLY, and R. WELLS. 1982. Cod-capelin interactions off north-eastern Newfoundland and Labrador. *In*: Multispecies approaches to fisheries management advice, M. C. Mercer (ed.). *Can. Spec. Publ. Fish. Aquat. Sci.*, **59**: 141–148.

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**: 35 p.

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.

KERR, S. R., and R. A. RYDER. 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.*, Doc. No. 17, 28 p.

MAGNÚSSON, K. G., and O. K. PÁLSSON. 1989. Predator-prey interactions of cod and capelin in Icelandic waters. *ICES 1989 MSM Symp.*, Doc. No. 11, 21 p.

MILLAR, R. B., and R. A. MYERS. MS 1990. Modelling environmentally induced change in size and age for Atlantic Canada cod stocks. *CAFSAC Res. Doc.*, No. 48, 14 p.

PETERMAN, R. M. 1990. Statistical power analysis can improve fisheries management. *Can. J. Fish. Aquat. Sci.*, **47**: 2–15.

POPE, J. G. 1989. Multispecies extensions to age-structured assessment models. *Amer. Fish. Soc. Symp.*, **6**: 102–111.

RATKOWSKY, D. A. 1983. Nonlinear regression modelling. A Unified Practical Approach. Marcel Dekker, Inc., New York, p. 276.

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