ICES STATUTORY MEETING 1993

C.M. 1993/d:11
Sess.P

HOW TO REDUCE THE IMPACT OF MODEL UNCERTAINTY ON ASSESSMENTS AND ADVICE

by

Jake Rice
Department of Fisheries & Oceans
Pacific Biological Station
Nanaimo, B.C.
Canada V9R 5K6

ABSTRACT

Assessments and projections can have uncertainty because data are noisy or unrepresentative, and because model formulations do not capture the biological processes correctly. When scientists quantify risk, they lose none of the uncertainty due to data sources. Moreover, they may add further model uncertainty because they must specify the error structure of the models, as well as the biological processes.

Non-parametric density estimation methods estimate the probability density function (pdf) of attributes from data, without specifying a particular functional relationship or error distribution. Non-parametric density estimation methods can forecast the pdf of features like recruitment or weight-at-age from biological and/or physical influences. They can also capture environmental variability in tuning indices without having to specify the relationship between the index and the environmental factor (say, survey estimates and water temperatures). The quantitative results can be displayed as probability distributions or ogives.

When models are well motivated, or model parameters are of direct interest, model-based methods should be used to quantify uncertainty. When there are reasons to question any particular model formulation, non-parametric methods are useful for representing the true uncertainty in the advice and the range possible outcomes of management actions. From my experiences using distributions and ogives to present results to managers, I will draw some generalizations about the effectiveness of each type of display, and how each can influence decisions.

INTRODUCTION

Both the models we use to represent fish populations and fisheries and the data we use to parameterize the models determine the quantitative advice we give on the status and management of fish stocks. Both the models and the data are imperfect. The imperfections are major components of the uncertainty in advice and the risk in management options.

Data can be made less imperfect through improved survey and sampling designs, through collection of more data; and through better tools for making measurements. Nonetheless data will never be perfect. When scientists estimate risk or uncertainty, the error or variance in the data will contribute to the estimates.

The models which use the data are intended to represent actual biological processes. At best the models simplify and may misspecify the real processes. Moreover, the quantitative models must not only specify the functional relationship between entities (say, stock and recruitment), but the models must specify the form of the variability around that functional relationship (the error term). Models can be improved through research on biological and fisheries processes, and through approaches to model selection which render poor models implausible in comparison to good models.

The first strategy is limited because good fisheries research takes time. The second strategy is limited by the quality and quantity of data available to discriminate among alternative models. Even good data are often inconclusive or contradictory (e.g. Walters 1985, Richards 1991). Therefore, even many widely used models have been chosen for mathematical convenience rather than for any strong theoretical or empirical justification (e.g. Ricker 1954). Now, the computing power available routinely to most fisheries scientists has decreased our need for the mathematical "convenience" of simple functional relationships in fisheries models.

Estimating model parameters from poor data can make even good models go bad (Walters and Ludwig 1981, Schnute & Hilborn in press). Conversely, why put good data into a model which misspecifies functional relationships or error distributions? We would be adding additional inaccuracy to the level of uncertainty already present in the data themselves. In those cases we can apply quantitative tools which use the data directly, without interposing between the data and the advice functional relationships or error terms in which we have little faith. Nonparametric density estimation methods are one class of such tools (Silverman 1986).

ANALYTICAL METHODS

Evans and Rice (1988) describe one simple nonparametric density estimation in detail; a kernel estimator based on the Cauchy distribution. The kernel can use the same data as functional models; say pairs of historic ssbs and recruitments. The kernel estimates the probability density function (pdf) for recruitment given a new ssb (forecast or estimated) by weighting each historic recruitment by the similarity of the historic ssb which produced it to the new ssb. The weighting function is:

$$w_i = 1 / [1+(x_i / D)^2]$$

D is a tuning parameter estimated through crossvalidation (Bowman, et al. 1984). The x_i's are the differences between the "new" ssb and each of the i historic ssbs. The w_i's are scaled to sum to 1.0. The scaled weights are, themselves, the estimated pdf. Naturally, any pair of variables can be substituted for stock and recruitment, and users can smooth the estimated pdf further, if desired.

Evans and Rice (1988) report results of simulations which show that the kernel estimator performed as well or better (smaller average error, much better worst case performance) than traditional parametric approaches to stock - recruitment forecasting, except under ideal conditions. The "ideal" conditions were that the true functional form of the stock - recruit relationship was specified correctly, AND the error term was specified correctly, AND the noise (contribution of the error term to the simulated data) was small compared to the signal [contribution of the functional relation to the simulated data]. Rice and Evans (1988) and Rice (1993) illustrate applications of the Cauchy kernel to developing advice on rebuilding fish stocks and on fish - habitat interactions.

APPLICATIONS TO RISK AND UNCERTAINTY

PROBLEM 1: Selecting a rebuilding strategy.

To forecast the effects of a management strategy, fisheries scientists often must forecast expected recruitment levels. For example, in 1977 Canada adopted a management strategy to rebuild the cod stock in NAFO Division 2J3KL. Decision makers needed projections of the trajectory the stock would take under different F's. Although concerns about natural variation in recruitment were prominent in the ICNAF advice, there were no estimates of risk or uncertainty associated with the rebuilding forecasts. The stock and recruit data available at that time did not allow identification of an appropriate stock recruit function, because there was inadequate information on the curvature and descending limb of the function, if either existed (Fig 1). Therefore, ICNAF assumed annual recruitment would be the average of cohort estimates from 1962-1972 (ICNAF 1977).

Rice and Evans (1988) used the kernel estimator to investigate six questions about rebuilding cod stock. Their answers included estimates of the range of outcomes possible given the data which were available in 1977. Suppose the explicit objective had been to achieve an SSB of 1.0 million metric tonnes (mmt),within 10 years. If recruitment are chosen without consideration of SSB (i.e. if the kernel parameter D is made large relative to the range of historical SSB's), the risk of failing to achieve that objective with F=0.16 is negligible (Fig 2a). If the kernel parameter is tuned, the pdf of recruitment does vary with SSB and the probability of failing to achieve that objective is 0.30 (Fig 2a). Moreover, the approach allowed estimation of the full pdf of future SSB for different values of F, so the probability of the stock achieving various targets under different management regimes could be evaluated (Fig 2b). Interesting, when the work was done in 1986 the kernel algorithm estimated the median value for actual average fishing mortality between 1977 and 1984 to be 0.37. This value was substantially higher than VPA estimates at that time, but fairly close to current VPA estimates of F in those years. Not only did the data-based method provide estimates of uncertainty directly, it also estimated the central moment well, compared to model-based methods.

PROBLEM 2: Age Composition of Pacific Hake Catch

Canada and the US both fish the migratory stock of Pacific Hake (*Merluccius productus*). In a "typical year" the stock spawns in February off Southern California and northern Mexico, then migrates north. Older fish migrate further, so the Canadian harvest, taken from June to September, has an older age composition than the US harvest. In discussions concerning allocation of catch between the countries, it has been argued that tonne for the tonne the Canadian fishery places the stock at higher risk, because it has a greater impact on spawning biomass, and therefore on future recruitment.

Hake year-classes vary greatly. The influence of SSB on recruitment is weak. The data do not lend themselves to parametric analyses; again neither the peak of the dome nor the shape of the limbs can be determined with confidence (Swartzman, et al. 1983; Fig 3). We incorporated a Cauchy kernel estimator of recruitment from SSB into a simulation model of the hake population and fishery. We used the model to explore the effects of altering the age composition of the combined Canada - US harvest on future recruitment, future SSB's, and future yields. Scenarios ranged from age compositions younger than the present US catch to age compositions older then present Canadian catch (Table 1; data from Dorn and Methot 1992).

The cyclic nature of hake recruitment is present in results of all simulations (Fig 4a), as is the skewed and bimodal pattern in the frequency of occurrence of cohort sizes (Fig 4b). Some catch is foregone with the older age composition (Fig 5a), but SSB is actually higher (Fig 5b), and catches are more stable (Table 2).

Figures 4 and 5 are averages of 200 runs, and do not reflect uncertainty and risk. The hake stock is managed with a strategy which reduces exploitation significantly whenever the SSB falls below a critical level. Probability of reaching the critical level is available directly from the simulations. Suppose we identify the critical level as 0.225 mmt (the value used in 1992). With an older age composition of catch, the critical level is hit less often than with a younger age composition. One can examine the uncertainty of any features presented as averages in earlier figures. For example, catches and ssb's after 20 or 40 years are highly skewed for all age compositions (Fig 6a,b).

PROBLEM 3: Hydroacoustic Index of Capelin Abundance

Annual hydroacoustic surveys of capelin abundance are conducted in the Northwest Atlantic. Oceanographic conditions influence capelin distribution and aggregation patterns (Carscadden et al. 1989). To use the survey results as either an absolute or relative index of abundance in an assessment without accounting for oceanographic influences would include unnecessary variance and uncertainty. However, there is no theoretical justification for any specific functional relationship between temperature and abundance. Inspection of the data suggests the relationship may not be smooth and continuous (Fig 7).

Exploratory regression analyses produced statistically significant fits to the data (Fig 7). However, model estimates were consistently biased at temperatures above 2°C and below 0°C. Because of the aggregation of capelin, there also would be major problems specifying the error term in regression-based models. If both the central moments and the uncertainty estimates of a model are unreliable, it is unlikely to estimate the risk of various harvesting options accurately.

Relating hydroacoustic estimates of biomass to bottom temperature with the Cauchy-based kernel accounted for over half the variance in the hydroacoustic estimates (Fig 8). The ogives of abundance at different temperatures correspond well with the patterns in the original data. For very cold water (-0.5°C) only low capelin densities are likely. The pdf changes little until temperature reaches +1.5°C. Over the next 2°C, the upper limb of the ogive changes substantially, however the ogives remain nearly flat between 2 and 70 units of capelin. This reflects the schooling nature of capelin. Even in preferred temperatures one does not expect an "average" abundance of capelin. It is the likelihood of encountering high abundance which is varying.

From survey data and maps of bottom temperatures, one can construct a grid of ogives for the area. Algorithms then can estimate abundance indices through resampling from the ogives or using other strategies. Repeated estimation will represent a realistic pdf of actual abundance. The pdf may be quite skewed and irregular, but can be used directly to estimate the uncertainties and risks of various management options.

PROBLEM 4: Trawl Index of Shrimp Abundance

Shrimp stocks in British Columbia are surveyed with trawls (Boutillier 1992). Bathometry gives significant spatial structure to the shrimp populations. Physical oceanography also influences distribution and abundance. No functional relationships have been developed to account for the spatial or temperature influences on the trawl samples. The management strategy for shrimp stocks includes advising closure of fisheries when biomass estimates fall below a critical level set for individual grounds.

A multidimensional extension of the kernel estimator was used to estimate the pdf of abundance for a grid of sample points. When constructing the pdf of abundance at a given ("grid") position, influence of each sample point was weighted by the similarity of each sample point to the "grid" point on three factors: temperature, depth, and isotropic (positional) space (Evans et al. ms). Crossvalidation was used to select Cauchy parameters simultaneously for all 3 features. Differences in the ratios of the width of the tuning window to the range of observations for each feature provided the differential weighting of each feature when estimating the pdf of abundance (Table 3).

The ogive mapping procedure produced a pdf of abundance at each point on a spatial grid over the area. The annual abundance estimate, and the uncertainty around that estimate, were estimated through repeated sampling from the grid of pdfs. That overall pdf of abundance provides a direct estimate of the probability that the stock lies below the critical level. The pdfs are estimated without having to develop overall parametric relationships between shrimp abundance and temperature, nor detailed parametric models of the spatial distribution of shrimp on each fishing ground. Parametric versions of such models, and particularly parameterized error terms and uncertainties of such models, have not been produced.

DISCUSSION

DATA BASED OR MODEL BASED RISK ESTIMATES?

Consider problems like estimating how risk to the stock varies with age composition of hake catches. It may be possible to develop model-based representations of recruitment dynamics and use these to estimate risk. It would require finding normalizing transformations, obtaining stable estimates of error distributions, and many other steps. The best models could only be approximate ones. Moreover, given the data, it would be extremely difficult to reject many different representations of the functional relationships and errors of stock and recruitment, yet at least some representations would have to be wrong.

A careful advisor would run forecasts with all the models and error structures which are plausible given the data (cf Richards 1991). As a result the range of possible outcomes would increase, compared to results of any single model. The data already contain substantial uncertainly about the future dynamics of the stock. Proper use of model-based approaches can only add uncertainty, and science advisors would assign less certainty and greater risk to any management option.

We can use the data directly to produce the necessary forecasts and uncertainty estimates. These can be related directly to the decisions managers must make. What additional purpose is served by addition of weakly supported models?

WHAT DO WE REALLY WANT TO KNOW ABOUT RISK?

As we focus on assessing risk and uncertainty in fisheries advice and management, many decisions depend more on the shape of the tails of a distribution than on its central moments. Managers may wish to keep unpleasant events rare, or enhance infrequent opportunities.

As the capelin example showed, model-based estimates can be especially poor away from the average condition. Moreover, the mean condition may be unlikely. It can be argued that more sophisticated models, allowing for aggregated targets and complex patterns of errors, address concerns about entire distributions rather than just central moments. Again, one has only the data to guide the proper treatment of error. If the data often are inadequate to discriminate among alternative functional relationships, is it likely they would be better able to identify the proper error structure? Data-based methods aren't a cure-all, but they can be helpful.

PRESENTATION OF RESULTS

In this paper I have usually presented results as cumulative frequency distributions (cfd) rather than as pdf's. Both contain the same information. Which manner of presentation conveys the key information more effectively?

In my experience, effectiveness depends on the nature of the message. PDF's are more familiar to most scientists, but with many audiences either mode of presentation requires explanation. If the core message deals with the most likely event, or the average condition, pdfs seem to focus attention better on that point. If the core message is the probability of an extreme event, ogives focus attention on the shape of the distribution.

The difference is clear in an example I fudged from the hake forecasting model. Consider cumulative catch over 20 years. From the pdf (Fig 9b), the mode of test 2 is clearly at higher catches than the mode of test 1. We are used to making decisions based on differences in the peak of a pdf. Many would infer fishing patterns producing test 2 are better. However, due to the skew and spread of the distribution of forecasted catches, such a conclusion is unwarranted. The cfd (Fig 9a) shows that the mode is not representative of the "average" condition. In fact, the medians and means differ very little. Do we conclude the tests produce the same catches? The cfd shows the scenarios do differ: Test 1 has a higher likelihood of producing the larger catches. Although producing large catches is not generally viewed as a risk, the communication issue is appropriate. Both graphs show the same information, of course. Nonetheless, ogives can highlight the tails of the distributions. When managers select among options to avoid risks, or to encourage unlikely events, ogives may be a particularly effective communication tool.

LITERATURE CITED

Boutillier, J.A., W. R. Harling and D. E. Young. 1990. W. E. RICKER Shrimp survey 88-S-1, West Coast of Vancouver Island, April 28-May 8, 1988. Can. Data Rep. Fish. Aquat. Sci. 808: 40 p.

Carscadden, J.E., K.T. Frank, and D.S. Miller. 1989. Capelin (*Mallotus villosus*) spawning on the Southeast Shoal: influence of physical factors past and present. Can. J. Fish. Aquat. Sci. 46: 1743-1754.

Dorn, M.W. and R.D. Methot. 1992 Status of th coastal Pacific whiting resource in 1992. *in* Pacific Fishery Management Council, Status of the Pacific Coast groundfish fishery through 1992 and recommended acceptable biological catches in 1993, p. A1-A58. (Document prepared for the Council and its advisory entities) Pacific Fisheries Management Council, Metro Center, Suite 420, 2000 SW First Ave. Portland, OR.

Evans, G.T and J.C. Rice. 1988. Predicting recruitment from stock size without the mediation of a functional relation. J. Cons. int. Exp. Mer 44: 111-122.

Evans, G.T., J.C.Rice, and S.A. Akenhead. ms. The probability distribution for abundance of fish, and its pattern in space.

ICNAF. 1977. Redbook 1977: Standing Committee on Research and Statistics. Proceedings of special meeting December 1976 and annual meeting May-June 1977. Dartmouth N.S. Canada

Rice, J.C., in press. Forecasting abundance from habitat measures using nonparametric density estimation methods. Can. J. Fish Aquat. Sci. 50.xxx-xxx.

Rice, J.C. and G.T. Evans. 1988. Tools for embracing uncertainty in the management of the cod fishery in NAFO divisions 2J+3KL. J. Cons. int. Exp. Mer 45:73-81.

Richards, L.J. 1991. Use of contradictory data sources in stock assessments. Fish. Res. 11:225-238.

Ricker, W.E. 1954. Stock and recruitment. J. Fish. res. Bd. Can. 11:559-623.

Schnute, J.T. and R. Hilborn. in press. Analysis of contradictory data sources in fish stock assessment. Can. J. Fish. Aquat. Sci. xx:xxx-xxx.

Swartzman, G.L., W.M. Getz, R.C. Francis, R.T. Haar, and K. Rose. 1983. A management analysis of the Pacific whiting (*Merluccius productus*) fishery using an age-structured stochastic recruitment model. Can. J. Fish. Aquat. Sci. 40:524-539.

Walters, C.J. 1985. Bias in the estimation of functional relationships form time series data. Can. J. Fish. Aquat. Sci. 42:147-149.

Walters, C.J. and D.L. Ludwig. 1981. Effects of measurement errors on the assessment of stock-recruitment relationships. Can. J. Fish. Aquat. Sci. 38:704-710.

FIGURE LEGENDS

Fig. 1. Scatter plot of recruits (#'s at age 4) vs. ssb for cod in NAFO Div. 2J3KL from 1962 - 1977 (see Rice and Evans 1988).

Fig 2. Ogives of SSB estimated from 200 independent simulations, starting with 1977 conditions and proceeding 10 years: a) F= 0.16, using a very large value for "D" (no stock-recruit relationship) or "D" tuned with data in Fig. 1; b) tuned "D" and F increasing from 0.16 to 0.52 in steps of 0.03.

Fig 3. Scatter plot of recruits (#'s at age 2) vs ssb for Pacific hake, using data from Dorn and Methot 1992.

Fig 4. Recruitments of Pacific Hake stock and fishery over 100 years; population parameters from table 1. a) Time series of recruits averaged over 200 independent simulations. b) Ln frequency of occurrence (y-axis) of cohorts of various sizes (x-axis); from 200 simulations, each of 100 years.

Fig 5. As in Fig 4a, but for a) catch, and b) SSB.

Fig 6. Ogive of a) catches and b) SSb, after 20 years, from 200 simulations of Pacific hake stock and fishery (Fig 4),

Fig. 7. Scatter plot of capelin density (arbitrary hydroacoustic units) vs bottom temperature, from survey of NAFO Div. 3LNO in 1987 with linear and polynomial regression lines.

Fig 8. Ogives of density of capelin for 4 trial values of bottom temperature; data from Fig. 7, and tuned Cauchy algorithm.

Fig 9. Forecast cumulative catches of hake over 20 years, for two test fishing scenarios (details are artificial, to produce desired pattern of frequencies for example). a) Results displayed as ogive (cfd), b) results displayed as pdf.

TABLE 1. Population parameters input to hake simulation model.

AGE	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Data from the 1992 assessment														
POPAC	0.678	0.528	0.042	0.476	0.475	0.054	0	0.454	0.008	0.002	0.002	0.159	0.005	0.067
NMORT	0.237	0.237	0.237	0.237	0.237	0.237	0.237	0.237	0.237	0.237	0.237	0.237	0.237	0.237
POPWT	0.278	0.371	0.448	0.531	0.575	0.628	0.666	0.678	0.694	0.75	0.778	0.88	0.967	1.047
MATURE	0.19	0.64	0.77	0.82	0.93	0.97	1	1	1	1	1	1	1	1
USSLCT	0.04	0.15	0.42	0.73	0.92	0.98	1	0.99	0.97	0.86	0.58	0.24	0.06	0.01
CANSLCT	0	0.51	0.56	0.61	0.67	0.74	0.8	0.88	0.95	1	0.94	0.68	0.32	0.1
IPROPFEM	0.48	0.501	0.512	0.52	0.524	0.526	0.529	0.536	0.539	0.544	0.553	0.561	0.568	0.575
Experimental catch at age vectors (selectivities)														
Assessment	0.037	0.150	0.417	0.676	0.828	0.904	0.916	0.928	0.960	0.910	0.751	0.410	0.163	0.047
Younger	0.053	0.213	0.514	0.763	0.860	0.878	0.887	0.878	0.860	0.763	0.514	0.213	0.053	0.009
Older	0.036	0.145	0.403	0.654	0.801	0.875	0.886	0.851	0.919	0.967	0.909	0.658	0.310	0.097

POPAC = Initial populaion vector (billions)
NMORT = Natural mortality rate
POPWT = Population weight at age (Kg)
MATURE = Proportion of sexually mature females
USSLCT = U.S. fishery selectivity at age.
CANSLCT = Canadian fishery selectivity at age.

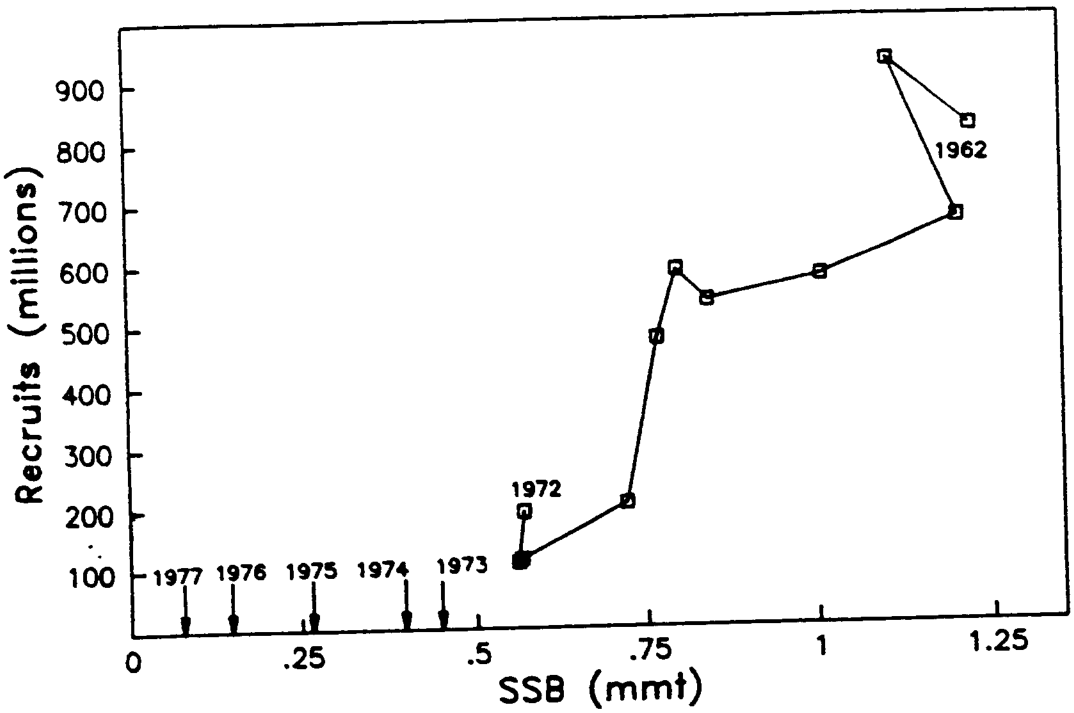
Assessment = Total fishery selectivity at age.
Younger = Fishing selectivity for younger fish
Older = Fishing selectivity for older fish

TABLE 2.

Over 200 simulations of hake model, tabulation of the number of years that the SSB was below the critical value used by managers to indicate need to apply conservative exploitaiton rates.

Years	Control Catch	Younger Catch	Older Catch
0	0	0	1
25	26	25	25
50	28	22	29
75	21	27	23
100	21	19	15
125	22	16	18
150	22	21	26
175	15	19	23
200	45	51	39

FIG 1



Recruits (millions)
900
800
700
600
500
400
300
200
100
1962
1972
1977
1976
1975
1974
1973
0
.25
.5
.75
1
1.25
SSB (mmt)

Projected SSB of Cod After 10 Years
F=0.16

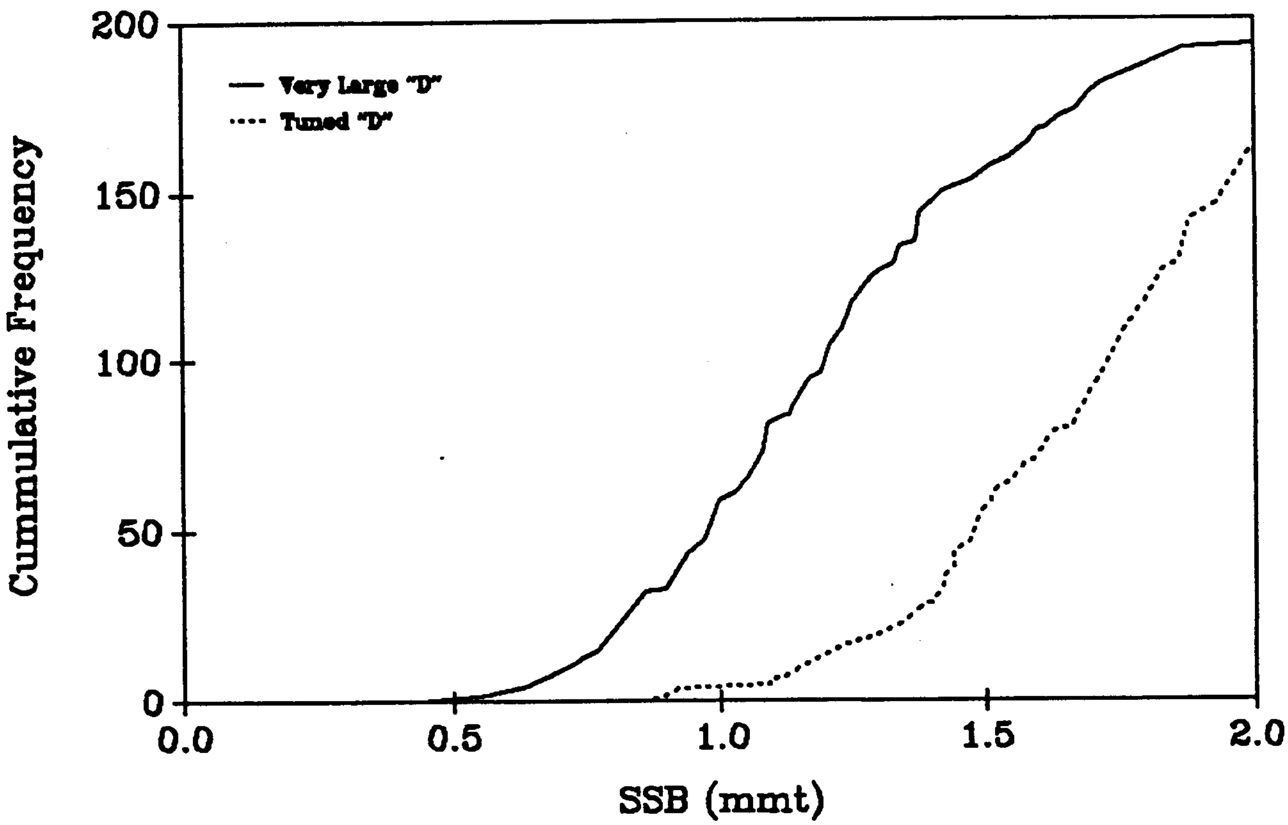


FIG 2A

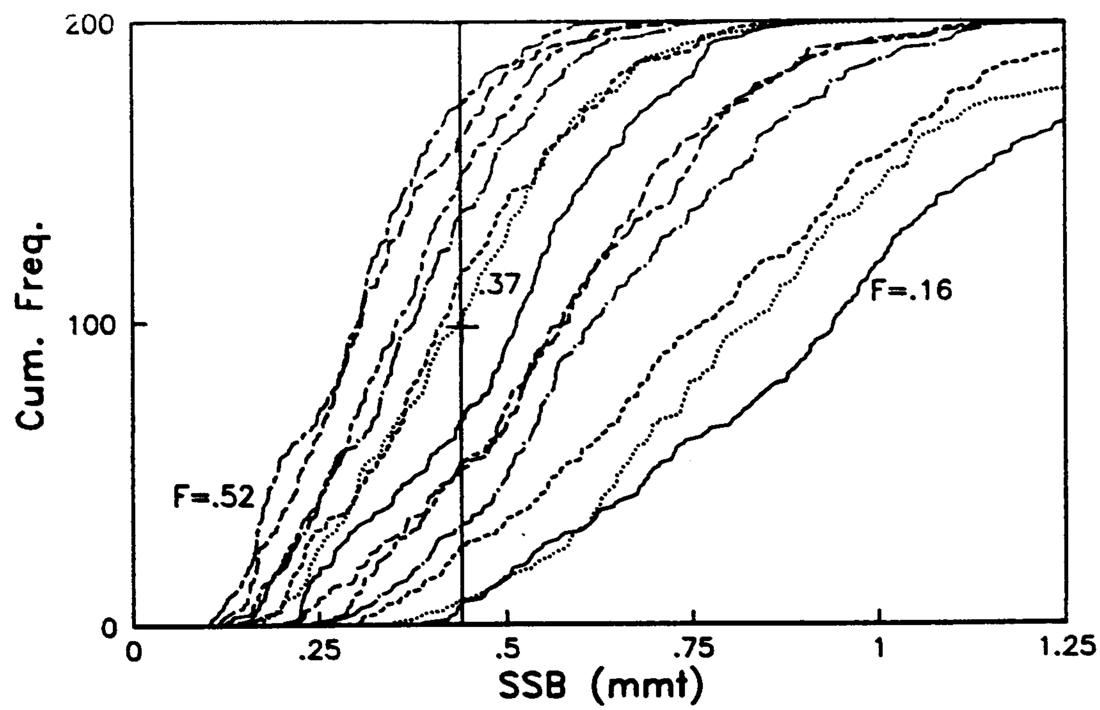


FIG 2B

FIG 3

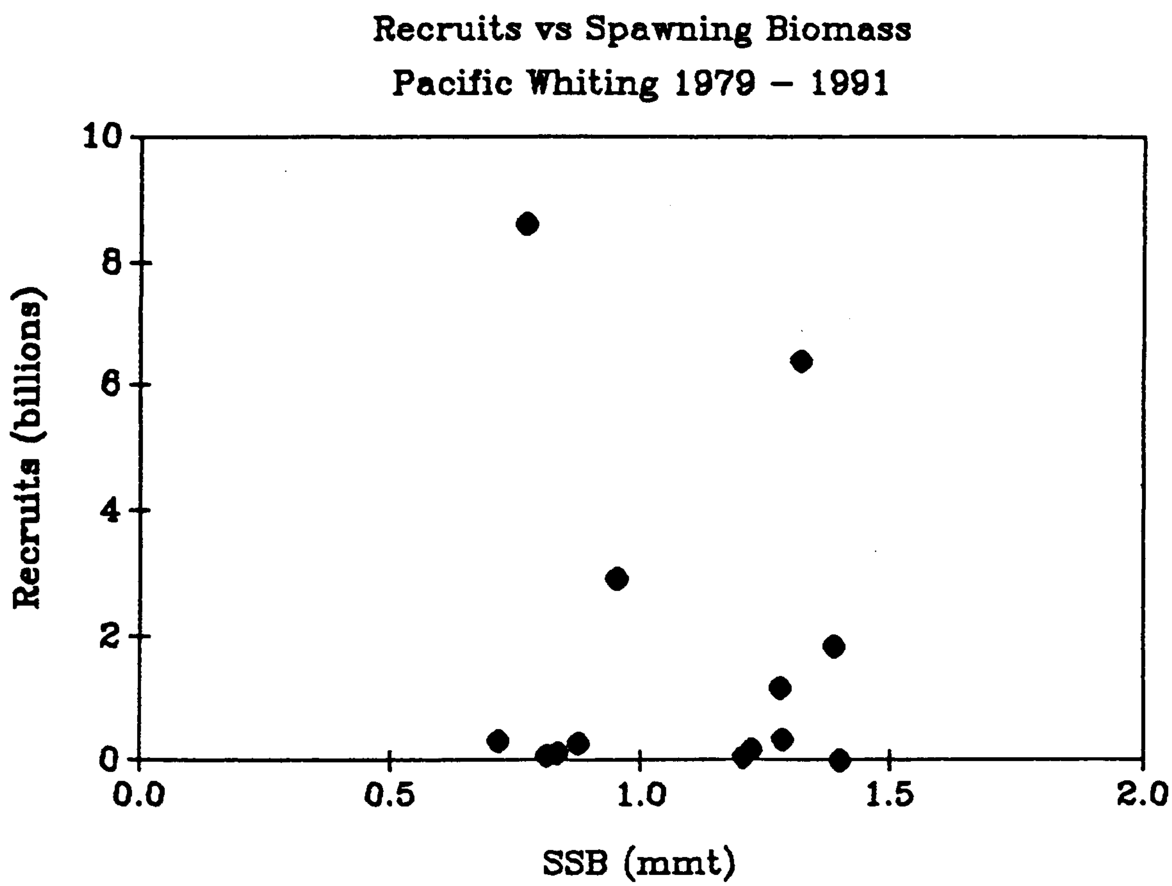

Recruits vs Spawning Biomass
Pacific Whiting 1979 - 1991
Recruits (billions)
10
8
6
4
2
0
0.0
0.5
1.0
1.5
2.0
SSB (mmt)

FIG 4A

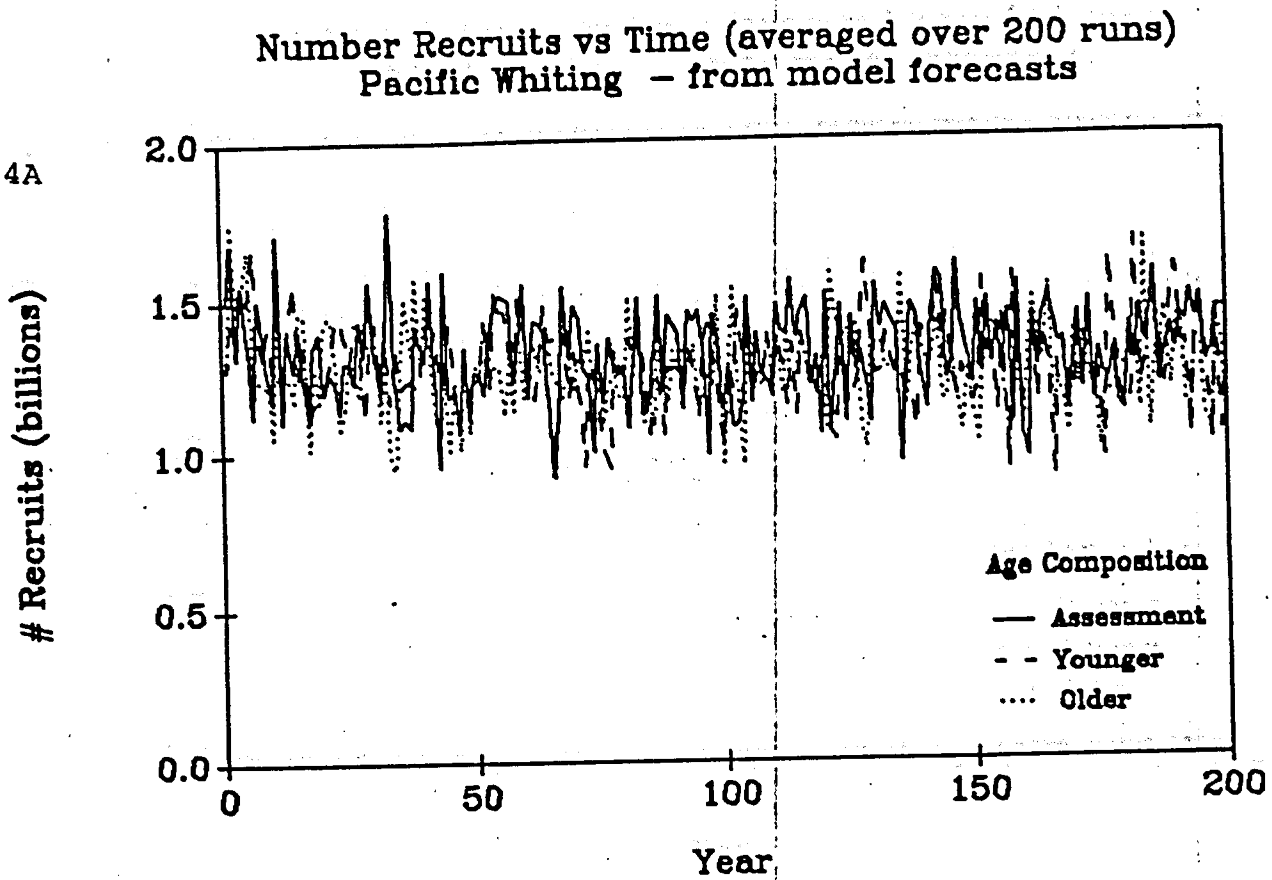

Number Recruits vs Time (averaged over 200 runs)
Pacific Whiting – from model forecasts
Recruits (billions)
2.0
1.5
1.0
0.5
0.0
0
50
100
150
200
Year
Age Composition
Assessment
Younger
Older

FIG 4B

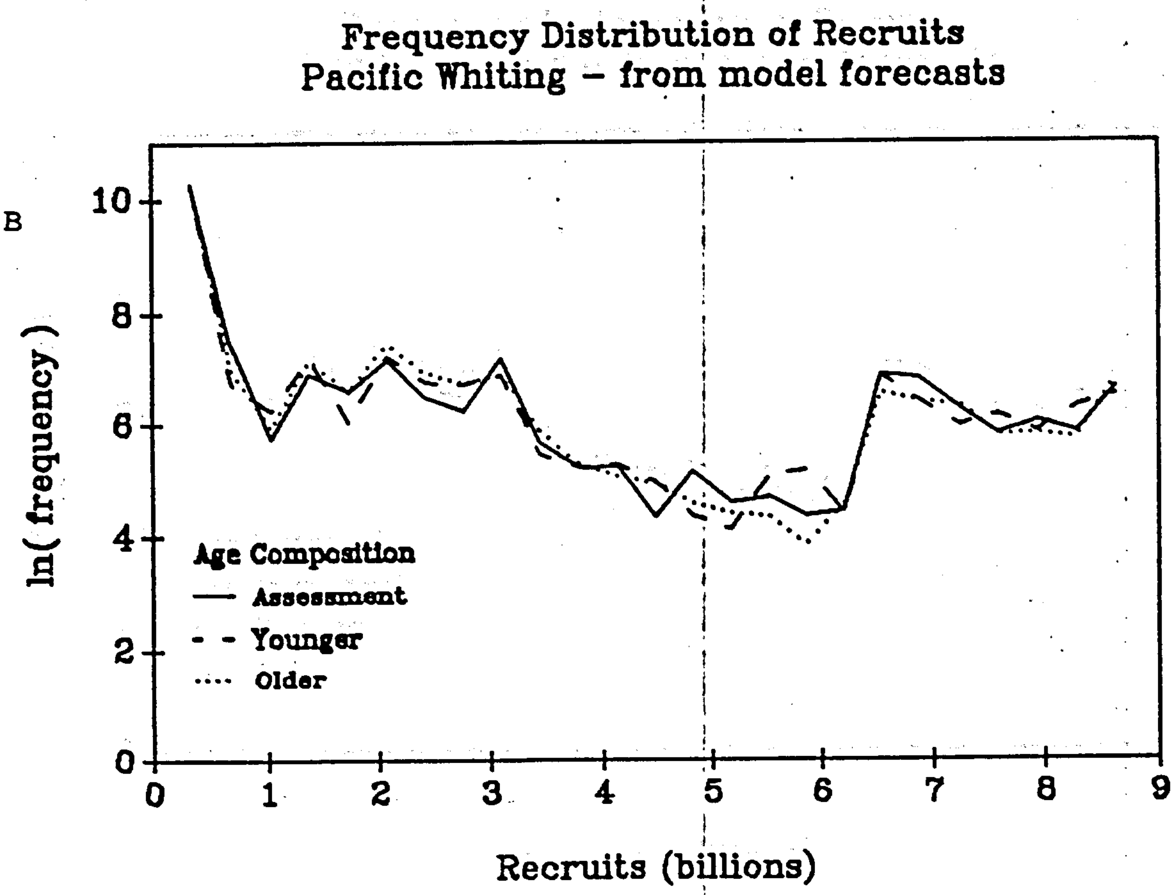

Frequency Distribution of Recruits
Pacific Whiting – from model forecasts
ln(frequency)
10
8
6
4
2
0
0
1
2
3
4
5
6
7
8
9
Recruits (billions)
Age Composition
Assessment
Younger
Older

FIG 5A

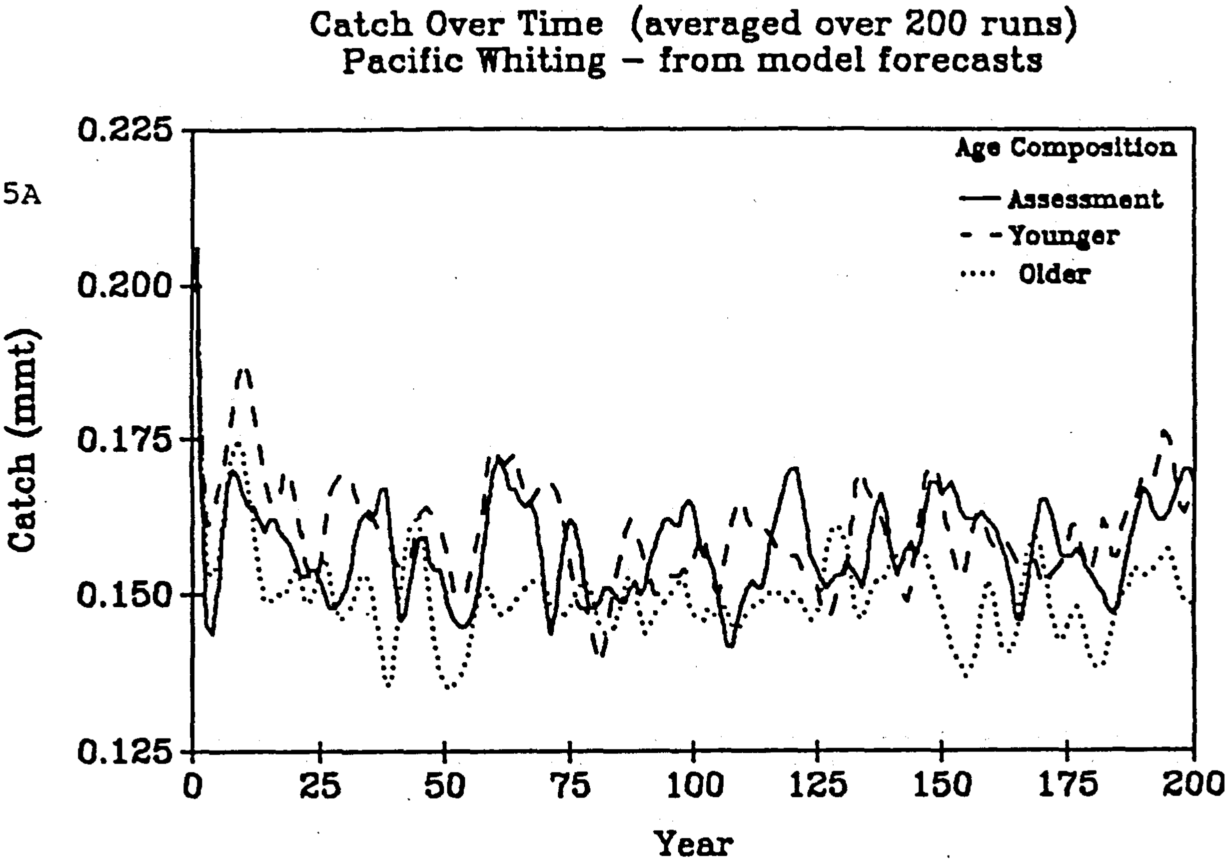
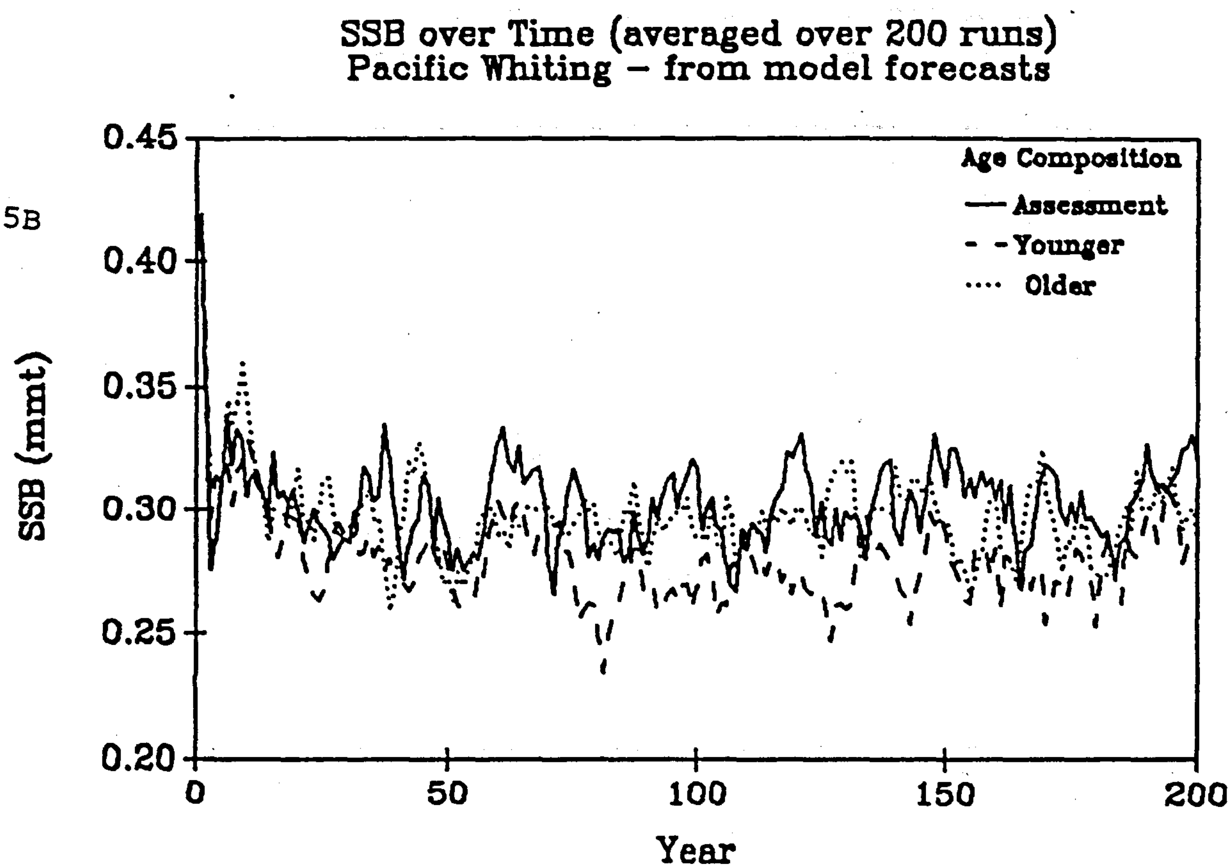
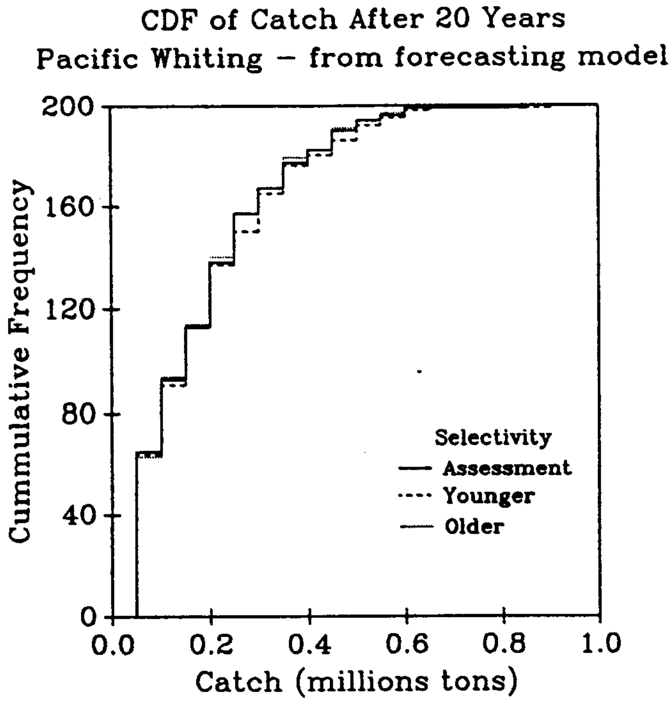

Catch Over Time (averaged over 200 runs)
Pacific Whiting – from model forecasts
Age Composition
Assessment
Younger
Older
Catch (mmt)
0.225
0.200
0.175
0.150
0.125
0
25
50
75
100
125
150
175
200
Year

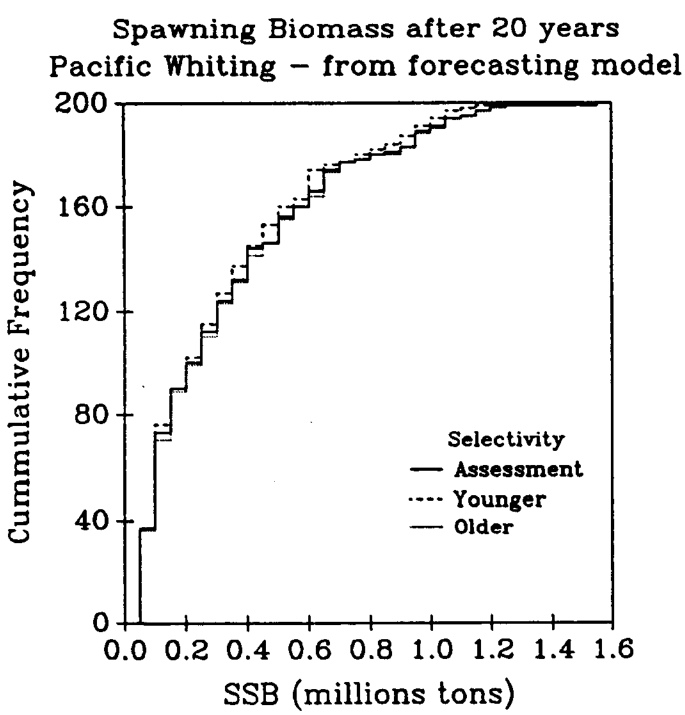
FIG 5B


SSB over Time (averaged over 200 runs)
Pacific Whiting – from model forecasts
Age Composition
Assessment
Younger
Older
SSB (mmt)
0.45
0.40
0.35
0.30
0.25
0.20
0
50
100
150
200
Year



CDF of Catch After 20 Years
Pacific Whiting – from forecasting model
Cummulative Frequency
Catch (millions tons)
Selectivity
Assessment
Younger
Older

FIG 6A



Spawning Biomass after 20 years
Pacific Whiting – from forecasting model
Cummulative Frequency
SSB (millions tons)
Selectivity
Assessment
Younger
Older

FIG 6B

Capelin Density vs. Bottom Temperature – 1987

FIG 7

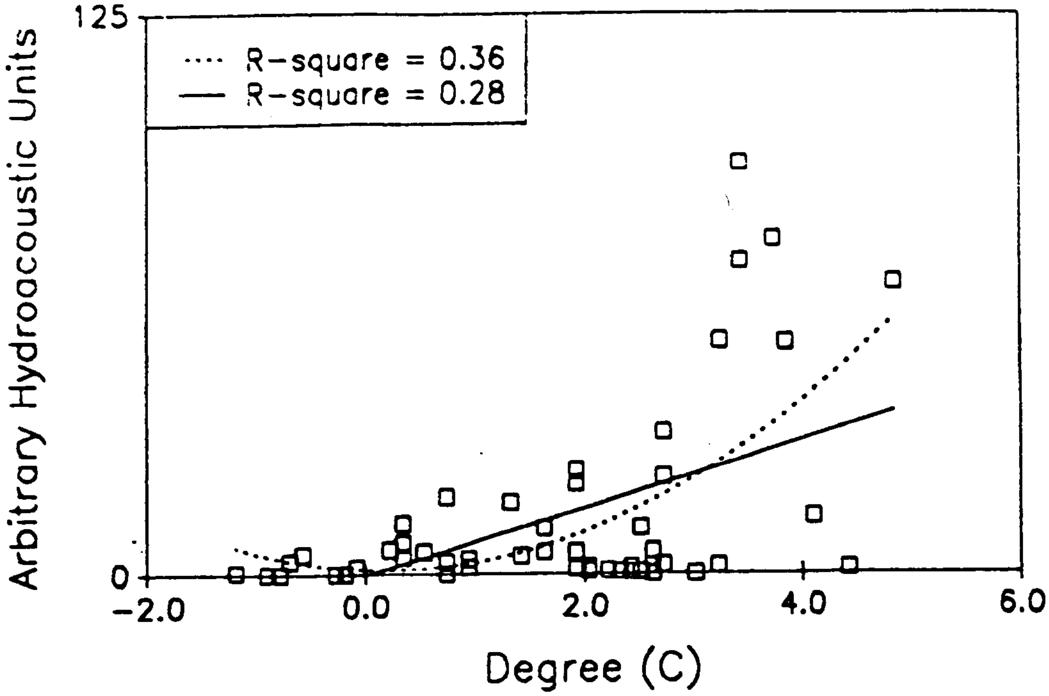


FIG 8

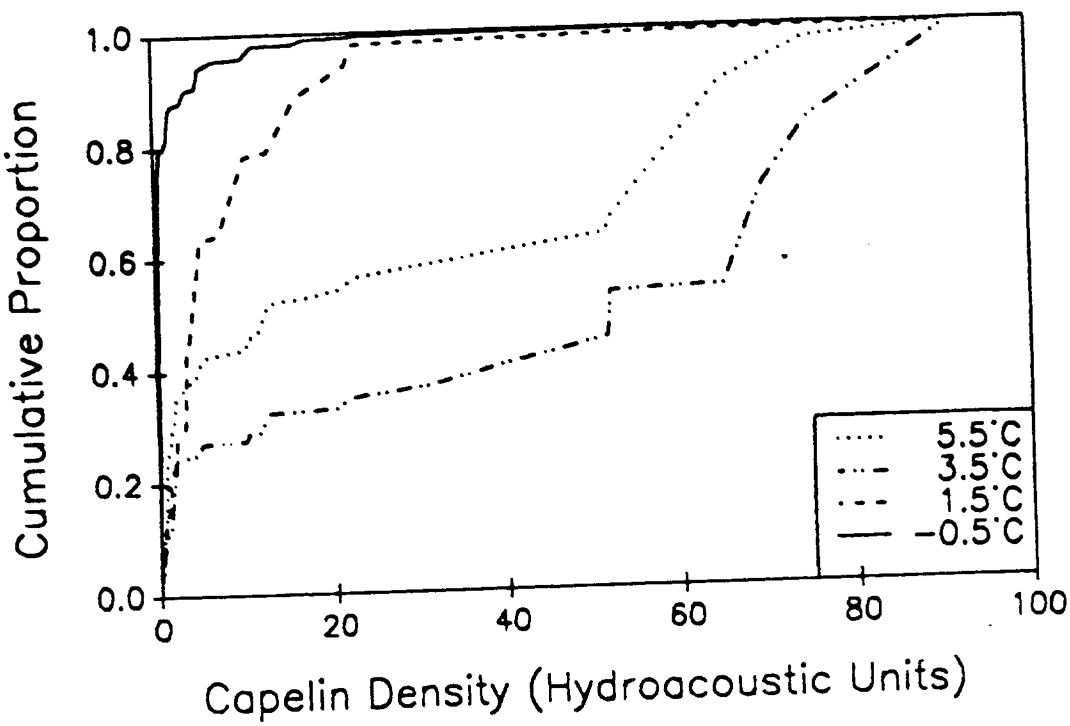


Fig. 9a

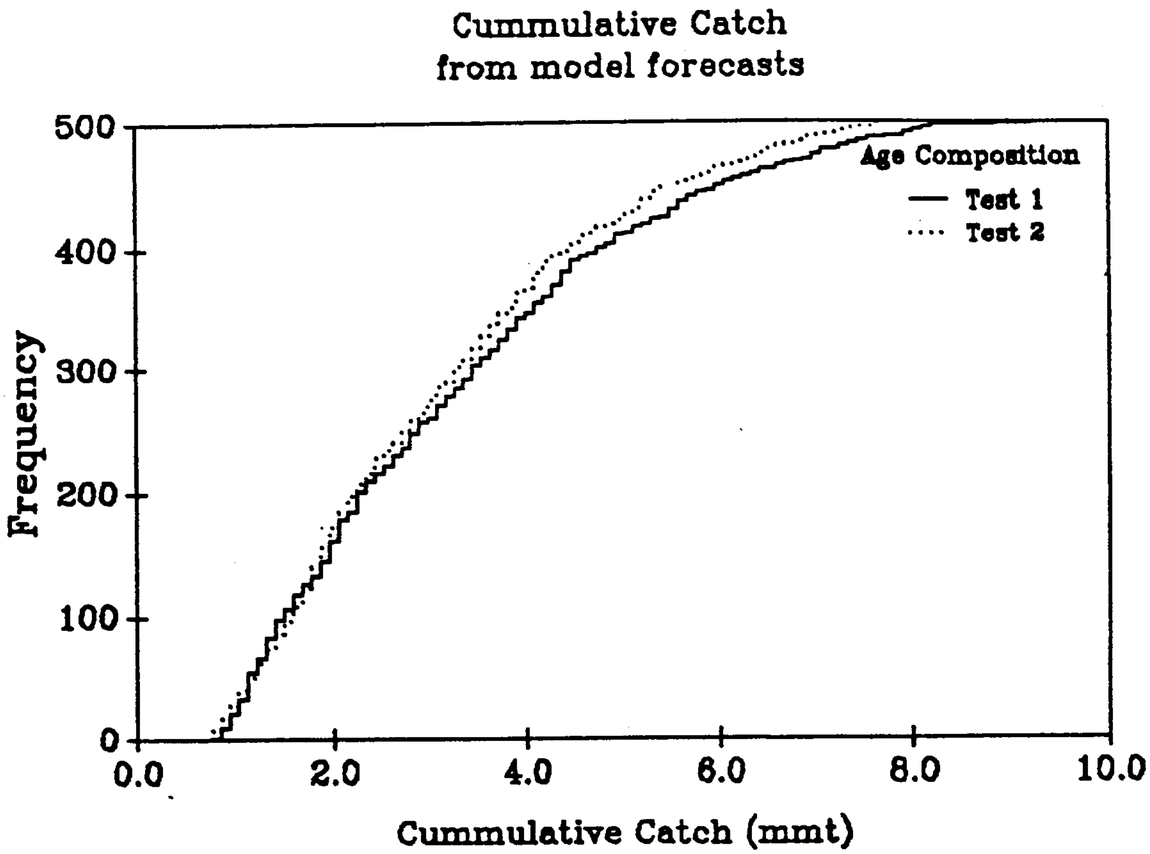
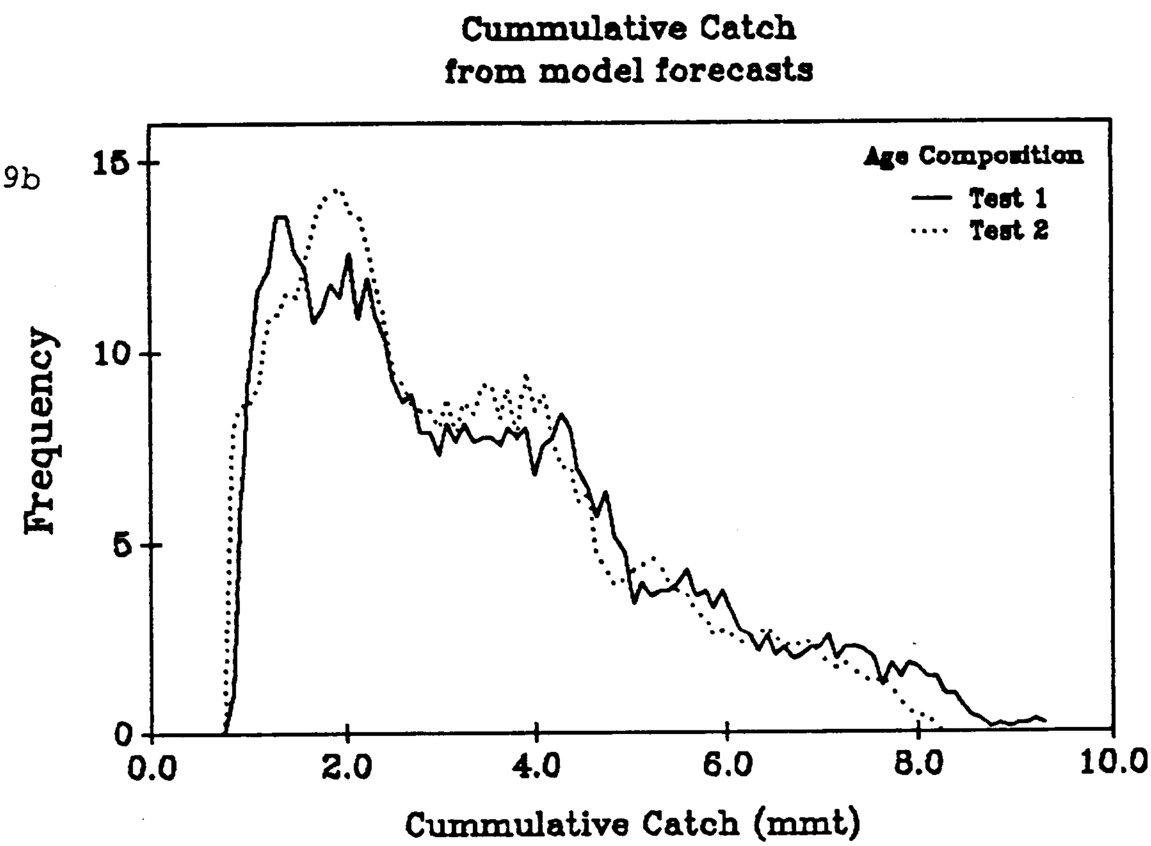

Cummulative Catch
from model forecasts
Frequency
Cummulative Catch (mmt)
Age Composition
Test 1
Test 2
0
100
200
300
400
500
0.0
2.0
4.0
6.0
8.0
10.0

Fig. 9b


Cummulative Catch
from model forecasts
Frequency
Cummulative Catch (mmt)
Age Composition
Test 1
Test 2
0
5
10
15
0.0
2.0
4.0
6.0
8.0
10.0