Northwest Atlantic Fisheries Organization

Serial No. N4003 **NAFO SCR Doc. 98/102**

SCIENTIFIC COUNCIL MEETING - SEPTEMBER 1998

Variations in Maturity of Haddock in the Barents Sea in Relation to Year-class Strength, Age, Size, Sex and Area

by

Knut Korsbrekke

Institute of Marine Research, P.O. Box 1870 Nordnes,
5024 BERGEN, NORWAY

ABSTRACT

Data from the Norwegian Barents Sea bottom trawl surveys in February from 1985 through 1996 are used to analyse variations in haddock (*Melanogrammus aeglefinus* L.) maturity. Several link functions to estimate maturity ogives are compared. Due to the clustering of samples, both the goodness of fit and various other test statistics are calculated using an effective sampling size lower than the total number of samples. Length is used as a continuous explanatory variable while year, age, sex and area are class variables in two different models. A strong year effect seems somehow to be related to the very large 1990 year-class. That is: when abundance is high, the proportion mature at length is reduced and due to the reduced growth at this high abundance the proportions mature at age show an even stronger reduction. Both sex and area show significant effects. Males mature at younger ages and shorter lengths than females and there is a tendency for proportions mature to be lower in the eastern part of the Barents Sea.

INTRODUCTION

Haddock is the second most important of the commercially exploited groundfish species in the Barents Sea. In the years covered by the analyses in this paper, the landings have ranged from 26 thousand tons in 1990 to a high of 173 thousand tons in 1996. The age 3 recruitment estimates (ICES CM 1998/Assess:2) show a few strong year-classes with most year-classes more than one order of magnitude lower. A few year-classes, mostly neighbouring the strongest, seem to be around 20-50 percent of a strong year-class.

The main spawning areas are on the continental slopes in the western part of the Barents Sea and spawning takes place mostly towards the end of April. (Solemdal *et al.*, 1989, Solemdal *et al.*, 1997).

Strong fluctuations in recruitment seem to be common for most haddock populations. Several authors have reported changes in maturity at length related to changes in growth. Reduced growth (usually coinciding with high abundance) increases the length at first spawning (Becham 1983, Kovtsova 1993, Templeman *et al.*, 1978, Tomosova, 1983).

MATERIALS AND METHODS

Data from the Norwegian Barents Sea bottom trawl survey in the years 1985 to 1996 are used in this study. The survey is a combined acoustic and bottom trawl survey for demersal fish and is conducted annually from the end of January to the beginning of March. The main aim of the survey is to map the spatial distributions and obtain indices of abundance for the most important commercially exploited demersal fish species in the Barents Sea. The target species are cod (*Gadus morhua*), haddock (*Melanogrammus aeglefinus*), golden redfish (*Sebastes marinus*), beaked redfish (*Sebastes mentella*) and Greenland halibut (*Reinhardtius hippoglossoides*). A description of the survey can be found in (Jakobsen *et al.*, 1997).

The survey area was expanded in 1993 and because of this change to the time series, only data from the central region (A, B, C and D), which had good coverage in all years, were used (see Fig. 1).

The effort and equipment employed in these years have changed (Jakobsen *et al.*, 1997). The biggest change was the switch from bobbins to rockhopper ground gear in 1989. Another change was a reduction of mesh size in the codend from 40 to 22 mm in 1994 that introduced an increase in the abundance indices of small fish including haddock. In this work however, it is assumed that this change in mesh size had no effect on the observed proportions of mature haddock.

Proportions mature are modelled as a response probability using a link function (Nelder and Wedderburn, 1972):

$$g(p) = \alpha + \mathbf{b}'\mathbf{x}$$

The link functions used in these analyses are the logit (1), probit (2) and complementary log-log (3) link functions:

$$g(p) = \log\left(\frac{p}{1-p}\right) \tag{1}$$

$$g(p) = \Phi^{-1}(p) \tag{2}$$

where $\Phi^{-1}(p)$ is the inverse of the standard cumulative distribution function

$$\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{\frac{-z^2}{2}} dz$$

$$g(p) = \log(-\log(1-p)) \tag{3}$$

Samples for age, sex and maturity are clustered samples from a limited number of trawl stations. It has been shown that samples from one station show higher similarity compared to samples from neighbouring stations (i.e. there is intra-haul correlation). This has been shown for parameters such as mean weight or length, but also for stomach content data (Pennington and Voelstad, 1992, Bogstad *et al.*, 1995). This means that increasing the within station number of samples yields less new information than the sample size normally would indicate. This could lead to non-significant effects wrongly being included in the model. To avoid this problem the yearly sample size was reduced to:

$$N_{eff} = e^{\frac{\log(N)+\log(n)}{2}} \tag{4}$$

where the effective sampling size N_{eff} is a function of the total number of samples *(N)* and the number of trawl stations with sampling of age, sex, maturation and size *(n)*. N_{eff} and N are shown in the following table:

Year	**85**	**86**	**87**	**88**	**89**	**90**	**91**	**92**	**93**	**94**	**95**	**96**
N	296	701	454	503	861	399	437	595	894	905	1011	798
N_{eff}	43	108	83	96	132	62	67	116	176	186	211	227

Samples are stratified in 5 cm length intervals and there were 5 samples in each interval. This was changed to 2 samples in each interval in 1993 and to 1 sample in each length interval in 1996. The number of stations with age sampling has been increased from 2 in each stratum to all stations in 1996. The highest (theoretical) degree of clustering would occur if all samples in one station were equal. Then one sample from each station would suffice and $N_{eff}=n$. The degree of clustering may vary from year to year and from station to station. The work by Pennington and Voelstad (1992) suggests that the effective sample sizes used in this study is on the conservative side.
The weighting factors used in the estimation of age-length keys (see Jakobsen et al. 1997 and also Morgan and Hoenig (1997) for a discussion on the handling of length stratified samples) are adjusted (multiplied by a constant) so that the sum of the weights equals N_{eff}.

The explanatory variables used in the analysis are described in the following table:

Year	class variable	1985-1996

Length	continuous variable	length in cm
Age	class variable	age classified from otholitt readings
Sex	class variable	female or male
Area	class variable	West (A, B or C) or east (D)
St. log(VPA)	continuous variable	Log population numbers standardised within age groups (mean=0)
Growth	continuous variable	increase in mean length since previous survey by age and sex

The following second order interactions were also included:

	Length	Age	Sex	Area	St. log VPA	Growth
Length		*	*	*		
Age	*		*	*	*	*
Sex	*	*		*	*	*
Area	*	*	*		*	*
St. log(VPA)		*	*	*		
Growth		*	*	*		

The third order interactions were st. log(VPA)*age*sex, growth*age*sex, st. log(VPA)*sex*area and growth*sex*area. The standardised log(VPA) was calculated as log population numbers by year and age standardised within age (by subtracting the mean value). The estimated population numbers used was taken from ICES CM 1998 / Assess:2.

All explanatory variables or their interactions were included or excluded using a stepwise method. At any step an adjusted chi-square statistic was calculated for all effects not already included in the model. The effect with the largest χ^2 value was then included if it was significant at the 0.05 level. The new model was then fitted and non-significant effects were removed.

The inclusion of a year effect is interesting from a model fitting perspective, but can be less satisfying when looking for underlying mechanisms causing changes in the proportions mature. Such mechanisms could be related to changes in year-class strength, growth conditions or geographic distribution. The material was therefore analysed both with and without the year effect.

RESULTS

The use of the logit link function (or logistic regression) performed slightly better than the other link functions, but without changing any of the conclusions on which effects best described the variations in observed maturity. Not an unexpected result as long as the functions define curves of quite similar shape and as long as the functional differences are much smaller than the variance in the model. The choice of a logistic link function seems to be preferred for several species (Chen *et al.*, 1994, Munger *et al.*, 1994).

When including a year effect the following regression gave the best fit to the data:

$$\log\left(\frac{p}{1-p}\right) = I + l \cdot L + Y_i + A_j + B_{s,a} + s_{i,j} \cdot C_a \qquad (5)$$

where

I	intercept
l	length in cm
L	length parameter
Y_i	year effect in year i
A_j	age effect for age j
$B_{s,a}$	interaction effect of sex s in area a
$s_{i,j}$	standardised log population numbers in year i age j
C_a	log population numbers parameter for area a

Results without year effect:

$$\log\left(\frac{p}{1-p}\right) = I + l \cdot L + A_j + B_S + C_a + s_{i,j} \cdot D \quad (6)$$

where

I	intercept
l	length in cm
L	length parameter
A_j	age effect for age j
B_s	effect of sex s
C_a	effect of area a
$s_{i,j}$	standardised log population numbers in year i age j
D	log population numbers parameter

Model (5) with year effect resulted in: r^2=0.4975, while model (6) resulted in r^2=0.4326. The relatively low degree of variation explained by the models should be viewed in light of the reduced effective sampling size. Using full sample size model (5) estimation resulted in r^2=0.6197. The Hosmer-Lemeshow goodness of fit test did not produce significant results (no reason to reject the model including the choice of the logit as link function). The maximum likelihood estimates of the parameters are shown in tables 1 and 2.

(5) and (6) represent "final" runs in a long series of including different combinations of effects.

The computer software used allowed for "forcing" effects into the model and handling the remaining effects in a stepwise manner. This allowed for handling one of the problems with stepwise regressions: a stepwise inclusion allows for inclusion of only one effect at a time while two other effects may contribute more when both are included in a model. No such problem seemed to occur in these analyses and most interactions seemed to perform slightly worse than the simple effects.

As expected the individual length was the most important explanatory variable. The growth in the population varies between years. In Fig. 2 mean lengths at age for the cohorts 1988-1990 are shown together with mean length at age for all the years. The period 1988-1990 shows an increase in growth, but as seen from the length by cohort plots the 1988 year-class was able to maintain a higher growth rate than the younger ones through the decline in growth conditions in the 1990s. The most obvious cause for this would be the year-class strength increase.

The other continuous variable used in the models was the standardised log(VPA). Log population numbers are shown on Fig. 3. The bottom left plot shows the impact of the strong 1983 and 1990 year-classes. The overall higher overall abundance towards the end of the period is causing the most of the standardized log population numbers for a cohort seems to increase with age (bottom right).

Predicted proportions mature at age by area and sex for the years 1990 and 1996 are shown in Fig. 4 and 5 for comparison purposes. Curves connect minimum and maximum observed length.

The standardized log(VPA) numbers and growth are the only explanatory variables not observed directly during the surveys. The growth rate was calculated as observed differences in mean length since the previous survey, while the standardised log(VPA) numbers can be viewed as being independent of the survey. The last assumption is only an approximation due to the fact that for the last 2-3 years, the VPA estimates have not converged.

The abundance (standardised log(VPA) numbers) showed significant effects in both (5) and (6) and it is believed that the use of a year effect in (5) further reduced the level of "noise" so that the interaction with area was significant. The abundance effect in (5) and (6) could be replaced with a growth effect (without interaction), but this gave a poorer fit.

The year effects of model (5) are presented on figure 6 together with the age effect in model (5) and age effect in model (6). The age effects are quite similar and both effects have a remarkable trend. Increasing the age one year shifts the maturity ogive to the left with close to a constant value.

DISCUSSION

The Barents Sea as a habitat for haddock shows large fluctuations in the environmental conditions. These fluctuations together with the effects of varying competition (density dependent effects) are causing the large variations in growth and size seen in Fig. 2. The abundance of haddock was at a medium to low level in the late-1980s and the increase in growth should be more related to the increasing temperatures and prey abundance than to any density dependent effects. It is reasonable to believe that the good growth conditions were the major cause for the very good 1990 recruitment. The very large 1990 year-class together with reduced temperatures lead to the decline in growth observed in the later years. Length (or size) is the dominant factor determining maturity in haddock and with the observed variations it is clear that the size information also contains information on overall growth conditions including environmental factors and density dependent effects.

The estimated models in this work predict a reduction in proportions matures with up to 25% if the length is reduced with 10 cm. This means that the changes in mean lengths itself explains quite a lot of the observed variation in maturity.

Both models showed a remarkable age effect. Keeping other effects constant the effect of being 1 year older shifts the maturity ogive to the left (on a length axis) and there is a higher probability of maturing. This 1 year increase effect is close to a constant for proportions in the vicinity of 0.5.

The strong year effect in model (5) could not be replaced with any of the other explanatory variables or interactions between these. The effect start out at a high level up to 1990 and shows a decrease since then. As mentioned earlier this coincides with the increased abundance caused by the 1990 year-class and a reduction in temperatures. Note however that the temperatures were at very low levels in 1986-1987 when the year effect was high. Further insight in the mechanisms that is causing this year to year trend could be gained by including temperature effects and other environmentally linked effects in the modelling.

The growth effect did not perform very well compared with the other effects included in the models. As pointed out earlier the variation in growth are already partially embedded in the observed length. And it addition it should be mentioned that these analyses have produced models that predict the probability of sampling mature fish as opposed to immature. High growth would typically indicate a higher probability for maturing, but only so for the first time spawners. Repeat spawners have reduced growth and the assumption of a linear effect of growth rate would be violated. The proportions of repeat spawners would to some extent depend on the overall fishing mortality and the fishing pattern.

The interaction effect between age and growth rate did not perform very well either. Such an interaction is equal to allowing different growth effects between different age groups. As long as both the length at age and also the proportions repeat spawners vary, the assumption of a linear effect can again be violated.

The interaction between standardised log population numbers and area indicated that the difference between areas was larger in times with higher abundance. That is: at higher abundance with poorer growth conditions the longer distance from the eastern area to the spawning areas in the west could be of more importance than in years with better conditions.

In model (5) the interaction between area and sex suggests that the distance to the spawning areas is of importance (the area effect is larger for males than for females). As the spawning grounds are in the west the difference observed between areas are quite reasonable. Any spawning migration starting more than two months before spawning will lead to an increase in the observed proportions mature in the west. The difference between the sexes could then be caused by males migrating to the spawning grounds earlier than the females.

As for other interactions the interactions of sex with other explanatory variables performed worse than the first order effect used in (6). One can however not rule out the possibility that such effects could be included in more refined models.
Models that estimate or predicts proportions mature are of importance in assessment work. The models presented in this study will only be able to predict maturation based on predictions of the explanatory variables. They should be quite useful as a tool for improving the current years estimation of maturity and spawning biomass.

REFERENCES

BECHAM, T. D. 1983. Variability in size and age at sexual maturity of haddock (*Melanogrammus aeglefinus*) on the Scotian Shelf in the Northwest Atlantic Ocean. *Can. Tech. Rep. Fish. Aquat. Sci.*, No. 1168. 1983.

BOGSTAD, B., M. PENNINGTON, and J. H. VOELSTAD. 1995. Cost-efficient survey designs for estimating food consumption by fish. *Fisheries Research. Amsterdam*, vol. 23, no. 1-2, pp. 37-46.

CHEN, Y., and J. E. PALOHEIMO. 1994. Estimating fish length and age at 50% maturity using a logistic type model. *Aquat. Sci.*, Vol. 56, No. 3, pp. 206-219. 1994.

HOSMER, D.W. Jr., and S. LE,ESHOW. 1989. Applied Logistic Regression. New York: John Wiley & Sons, Inc.

ICES CM 1998 / Assess:2. Report of the Arctic Fisheries Working Group.

JAKOBSEN, T., K. KORSBREKKE, S. MEHL, and O. NAKKEN. 1997. Norwegian combined acoustic and bottom trawl surveys for demersal fish in the Barents Sea during winter. *ICES C.M. Doc.*, No. Y:17.

KOVTSOVA, M. V. 1993. Growth rate and maturation of Arct-Norwegian haddock in 1987-1990. *ICES C.M. Doc.*, No. G:58

MORGAN, M. J., and J. M. HOENIG. 1997. Estimating maturity-at-age from length stratified sampling. *J. Northwest Atl. Fish. Sci.*, **21**: 51-63.

MUNGER, C. R., G. R. WILDE, and B. J. FOLLIS. 1994. Flathead catfish age and size at maturation in Texas. *N. Am. J. Fish. Manage.*, **14**(2): 403-408.

NELDER, J. A., and R. W. M. WEDDERBURN. 1972. Generalized Linear Models. *Journal of the Royal Statistical Society*, Series A, 135, pp. 761-768.

PENNINGTON, M., and J. H. VOELSTAD. 1991. Assessing the effect of intra-haul correlation and variable density on population estimates from marine surveys. *ICES C.M. Doc.*, No. D:14.

SOLEMDAL, P., T. KNUTSEN, and H. BJOERKE. 1989. Spawning areas and spawning period of the North-East Arctic haddock (*Melanogrammus aeglefinus* L.). HELP Havforskningsinst. Egg-larveprogram. no. 25. 1989.

SOLEMDAL, P., N. MUKHINA, T. KNUTSEN, H. BJOERKE, and P. FOSSUM. 1997. Maturation, spawning and egg drift of Arcto-Norwegian haddock (*Melanogrammus aeglefinus*). Presented at the Fisheries Society of the British Isles Annual Symposium. Galway (Ireland), 8-11 July 1997.

TEMPLEMAN, W., V. M. HODDER, and R. WELLS. 1978. Sexual maturity and spawning in haddock, *Melanogrammus aeglefinus*, of the southern Grand Bank. *Res. Bull. ICNAF*, **13**: 53-65.

TORMOSOVA, I. D. 1983. Variation in the age at maturity of the North Sea haddock, *Melanogrammus aeglefinus* (Gadidae). *J. Ichtyol.*, 23(3): 68-74.

TABLE 1. Summary statistics for the logistic regression with year effect included.

	DF	Parameter estimate	Standard deviation	Chi-square statistic	P
I	1	-6.9641	1.1588	36.1163	0.0001
L	1	0.1264	0.0161	61.2678	0.0001
Y_{85}	1	2.0560	0.7490	7.5342	0.0061
Y_{86}	1	3.2755	0.4820	46.1852	0.0001
Y_{87}	1	2.8711	0.4970	33.3733	0.0001
Y_{88}	1	2.8707	0.4317	44.2211	0.0001
Y_{89}	1	2.8767	0.4347	43.7849	0.0001
Y_{90}	1	3.0901	0.5600	30.4437	0.0001
Y_{91}	1	2.4383	0.7118	11.7356	0.0006
Y_{92}	1	1.7862	0.5745	9.6678	0.0019
Y_{93}	1	1.7472	0.4600	14.4298	0.0001
Y_{94}	1	1.1251	0.3960	8.0739	0.0045
Y_{95}	1	1.0126	0.3327	9.2647	0.0023
Y_{96}	0	0.0000			
A_3	1	-4.1927	0.7749	29.2739	0.0001
A_4	1	-3.0200	0.6713	20.2412	0.0001
A_5	1	-2.0662	0.6178	11.1876	0.0008
A_6	1	-0.9400	0.5902	2.5368	0.1112
A_7	0	0.0000			
$B_{females,west}$	1	0.4949	0.3293	2.2583	0.1329
$B_{females,east}$	1	-0.2291	0.2816	0.6622	0.4158
$B_{males,west}$	1	1.4430	0.3196	20.3797	0.0001
$B_{males,east}$	0	0.0000			
C_{west}	1	0.2187	0.1335	2.6844	0.1013
C_{east}	1	0.4064	0.1682	5.8371	0.0157

TABLE 2. Summary statistics for the logistic regression without year effect.

	DF	Parameter estimate	Standard deviation	Chi-square statistic	P
I	1	-2.6308	0.8648	9.2539	0.0023
L	1	0.0993	0.0124	64.1220	0.0001
A_3	1	-4.7508	0.6426	54.6563	0.0001
A_4	1	-3.4869	0.5723	37.1179	0.0001
A_5	1	-2.5586	0.5550	21.2496	0.0001
A_6	1	-1.4442	0.5581	6.6955	0.0097
A_7	0	0.0000			
$B_{females}$	1	-0.6233	0.1590	15.3631	0.0001
B_{males}	0	0.0000			
C_{west}	1	1.1130	0.1646	45.7103	0.0001
C_{east}	0	0.0000			
D	1	-0.2444	0.0748	10.6678	0.0011

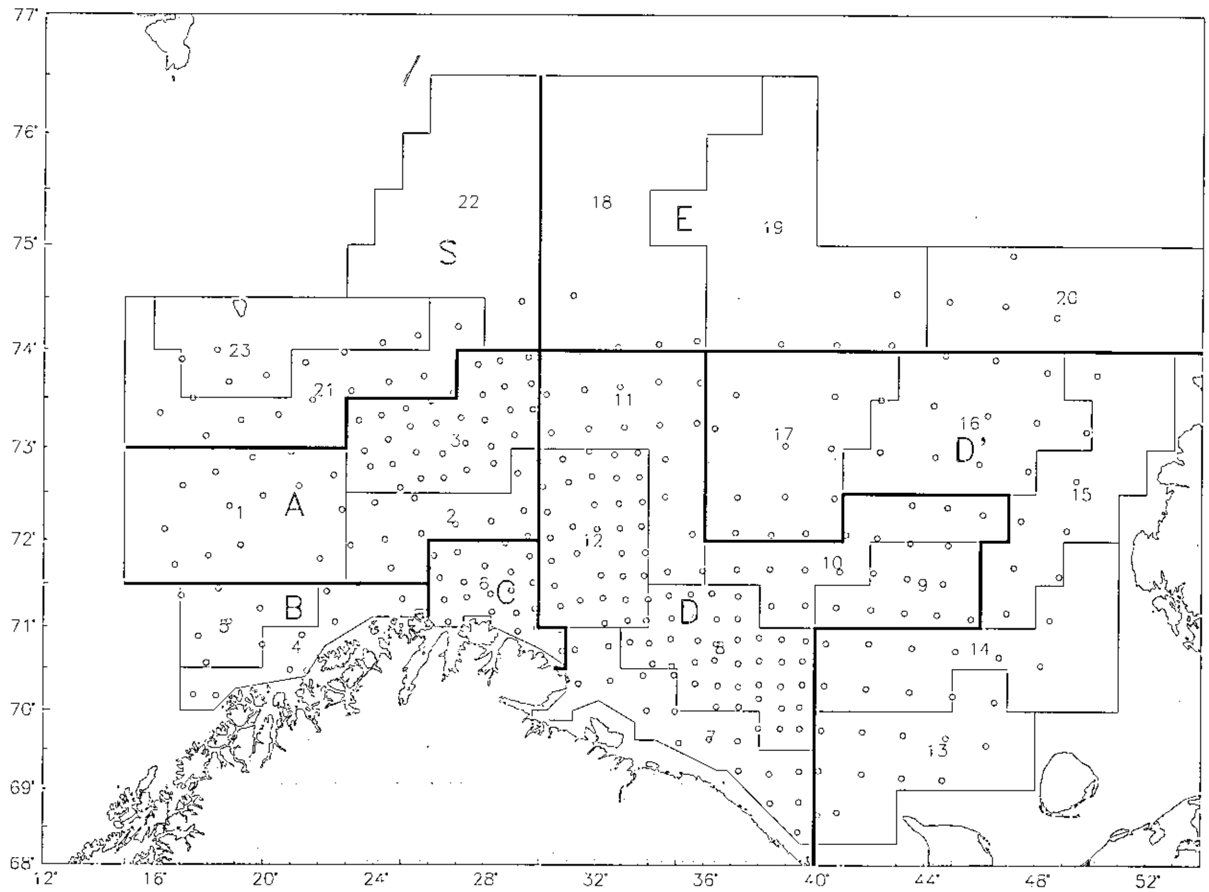

Fig. 1. Strata (numbers) and sub areas (letters) used in the bottom trawl survey. Trawl stations winter 1996 are shown.

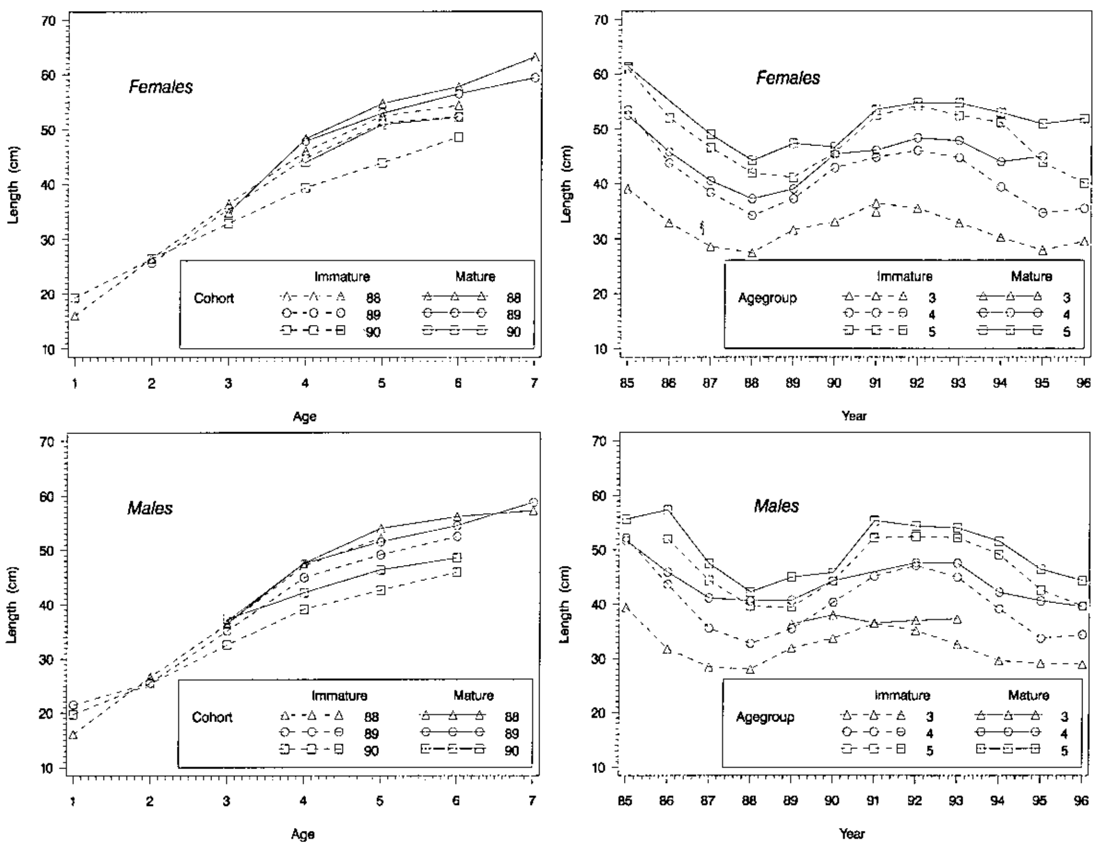

Fig. 2. Left: Mean length at age for the 1988, 1989 and 1990 year-class. Mean length of females (top) and males (bottom) for age groups 1-7 for both immature and mature fish are shown. Right: Mean length at ages 3-5 for females (top) and males (bottom) both immature and mature fish.

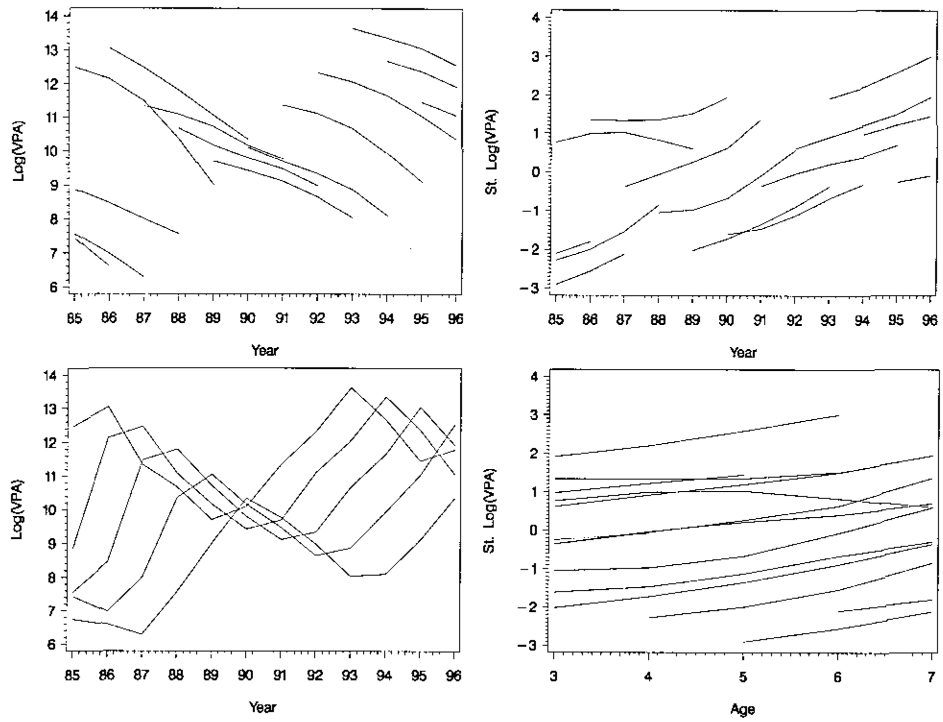

Fig. 3. Log population numbers (Log(VPA)) by cohort for the years 1985-1996 (upper left), Log(VPA) for ages 3-7 for the years 1985-1996 (bottom left), standardised Log(VPA) by cohort for the years 1985-1996 (upper right) and standardised Log(VPA) by cohort for the age groups 3 to 7.

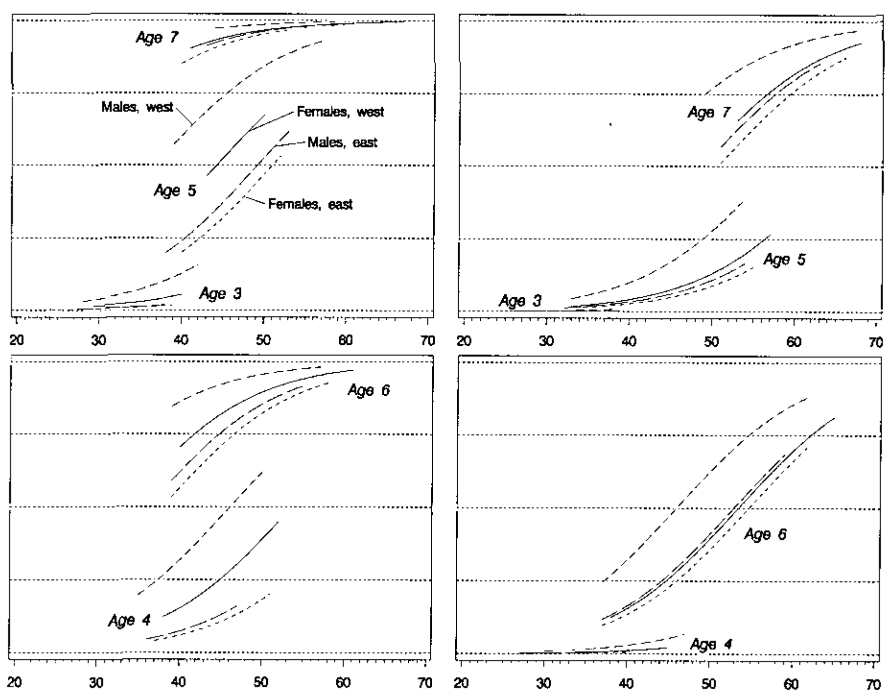

Fig. 4. Modelled maturity ogives with year effect (model 5) for 1990 (left) and 1996 (right). Curves for age groups 3 to 7 are shown for each area and sex.

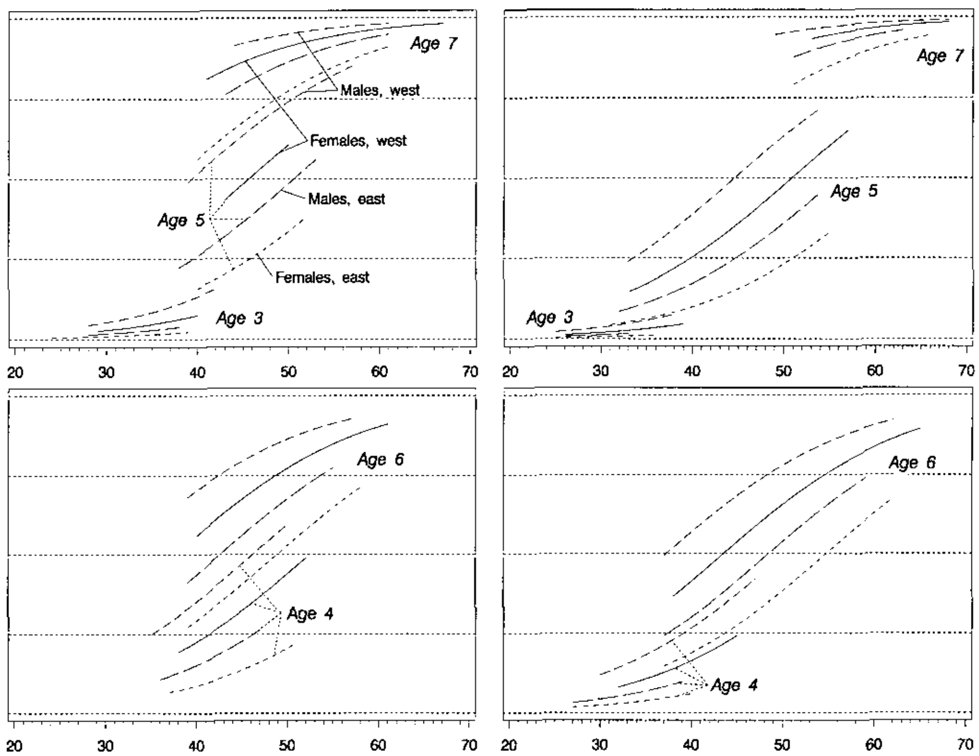

Fig. 5. Modelled maturity ogives without year effect (model 6) for 1990 (left) and 1996 (right). Curves for age groups 3 to 7 are shown for each area and sex.

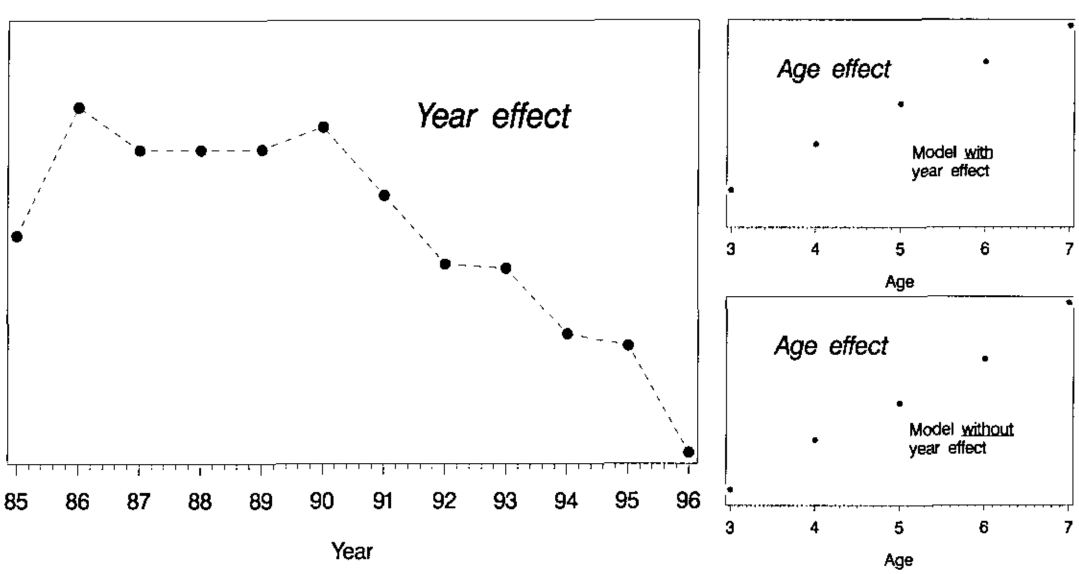

Fig. 6. Year effect parameters shown together with the age effect parameters (model 5) together with the age effects for the model without year effect (model 6).