ICES Annual Science Conference
Assessment Methods
C.M. 1997/DD:10

An assessment of a long-lived redfish species, *Sebastes marinus*, in Boreal waters

by
Gunnar Stefánsson
and
Þorsteinn Sigurðsson

Marine Research Institute, P.O. BOX 1390
Skúlagata 4, 121 Reykjavík
Iceland

ABSTRACT

An assessment method for fish without age-based data is presented. The method is tested on cod (*Gadus morhua* in Icelandic waters) and applied to a long-lived species of redfish (*Sebastes marinus*) in the East Greenland-Iceland-Faeroe area. The method utilizes survey indices and length distributions from survey and catch data. Test results show that although the estimation of initial conditions is problematic, the model adequately reproduces subsequent relative levels of recruitment, biomass and fishing mortality. The method appears to be adequate for use with redfish species which have considerable contrasts in the data sets of abundance and length distributions but considerable care needs to be exercised in its use due to the confounding of various variables when only survey indices and length distributions are used for assessments.

Introduction

Assessments of redfish tend to be problematic *in lieu* of their longevity, the lack of adequate age readings and various problems associated with stock identification. Thus, although ICES assessments for Division Va were at times age-based (NWWG 1992 and ACFM 1992), such methods were later abandoned in light of the lack of reliability of age readings and inconsistency of those compared to information on recruitment variability based on length distributions (NWWG 1995). In recent years these methods have been replaced by CPUE-based methods (NWWG 1997) which give some indications of the relative state of the stock.

In addition to the assessment work, parallel work has been initiated by ICES methods groups on the development and evaluation of methods for assessing fish stocks with varying levels of available data (MWG 1995).

This paper describes results from testing and applying one such model, which combines information from different sources into a single objective function. In order to test the model, it is first applied to cod (*Gadus morhua*) in Icelandic waters, where there is a detailed and robust age-based assessment available for comparative purposes.

Data sets

The following describes in detail the available data for golden redfish, *S. marinus*, which is the primary target of the assessment procedure. Analogous data sets for cod, used for testing the procedure, will be described in passing.

Landings data

Landings data are taken in accordance to official statistic as reported to the Icelandic Directorate of Fisheries for cod but for *S. marinus* data from the NWWG were used since catches of *S. marinus* and *S. mentella* are combined in official statistic. The division of catches between those two species are therefore based on indirect methods using log-books and samples taken by the Marine Research Institute.

Length distributions

Survey

The Icelandic groundfish survey (Pálsson *et al.1989*) includes length measurements of cod and *S. marinus*. The survey was planned with a special attention to cod and haddock and therefore some important distribution areas of other stocks may not be well covered. However, the survey seems to reflect the status of many other stocks and is used for that purpose (see MRI 1997a) .

The resulting length distributions for *S. marinus* for the period 1985-1997 are given in Fig. 1.

Figure 1. *S. Marinus* in div. Va. Survey length distributions.

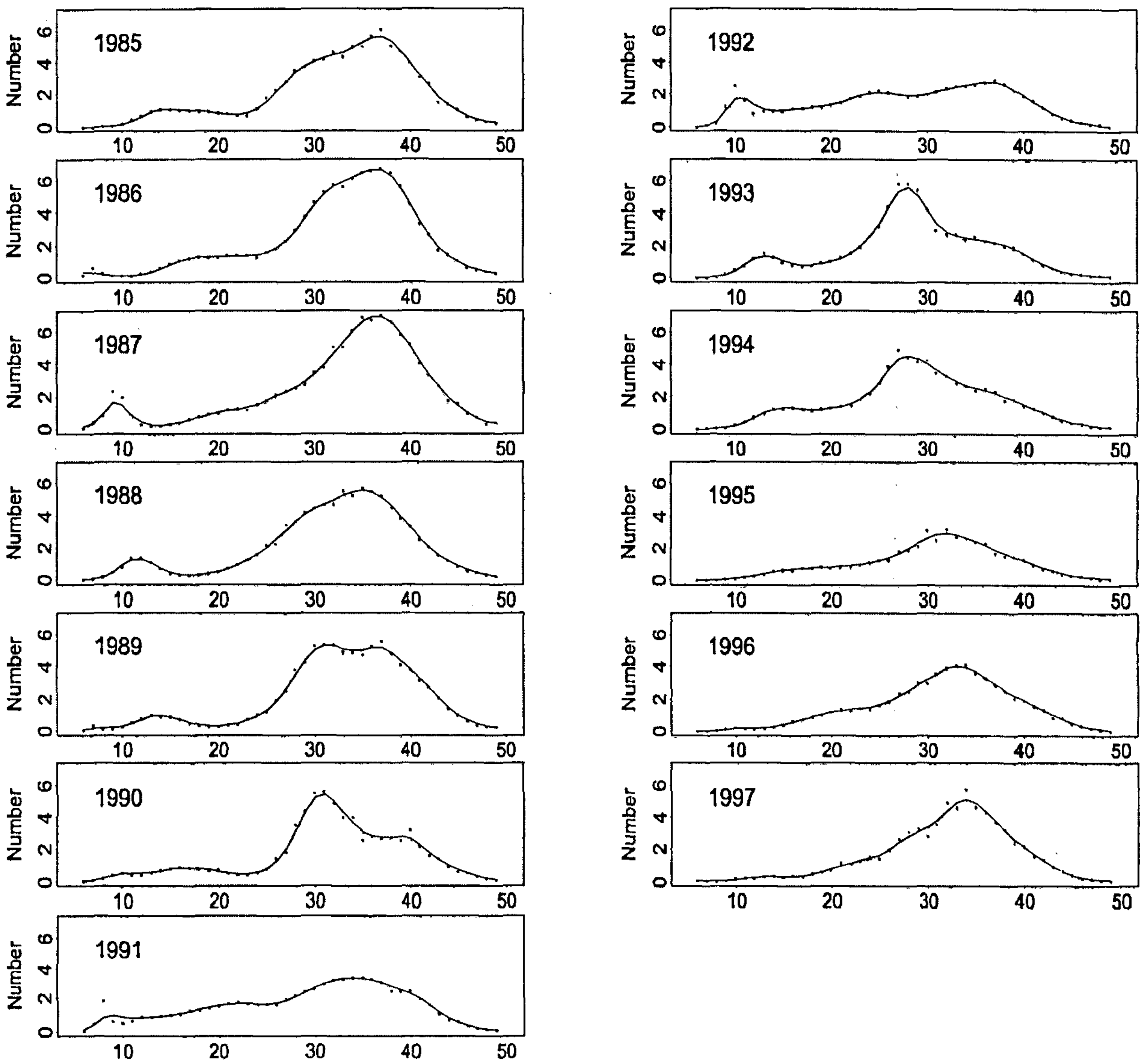
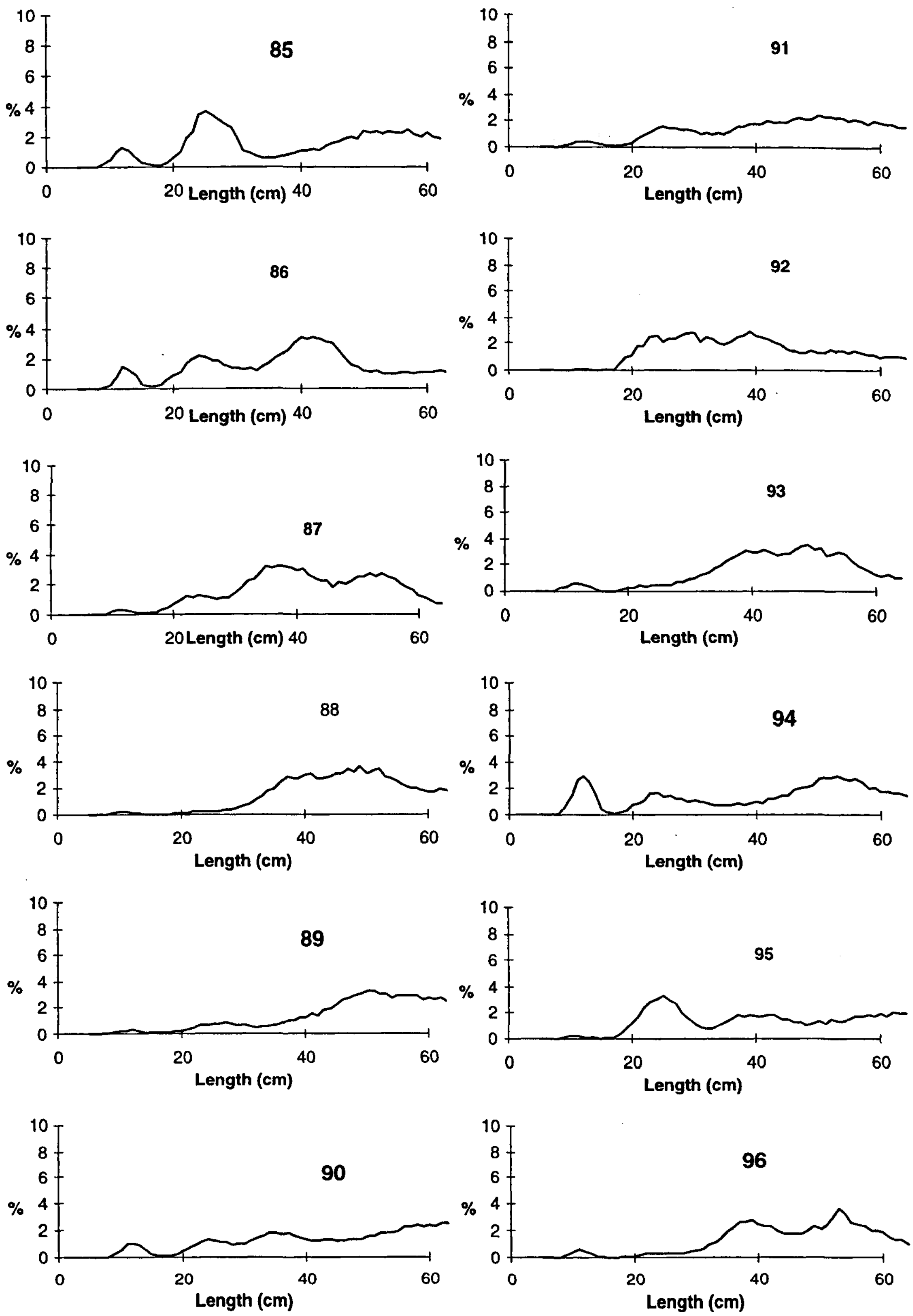


Figure 2. Iceland cod. Survey length distributions.



Length distributions for both species indicate that recruitment is highly variable. A large yearclass constitutes a clear, consistent peak in the length distributions. It appears, therefore, that it should be possible to use the length distributions to elucidate a recruitment signal. For cod, the indices of one year fish from the survey have shown to be reliable estimate of stock size as 3 year old fish as backcalculated form the VPA.

For redfish in particular, peaks in the length distributions seem to indicate a relatively strong 1985 and 1990 yearclasses, which is first seen as 1 year old and is clearly seen in several subsequent years. However, between these abundant yearclasses there appear to be years with almost no recruitment.

These data also indicate that this redfish species is a slow growing species, with annual growth of approximately 2-2.5 cm per year after the first year. It also seems that annual growth from O group to 2 year old is variable, where the 1990 year class has approximate 1 cm greater average length than e.g. the 1985 yearclass.

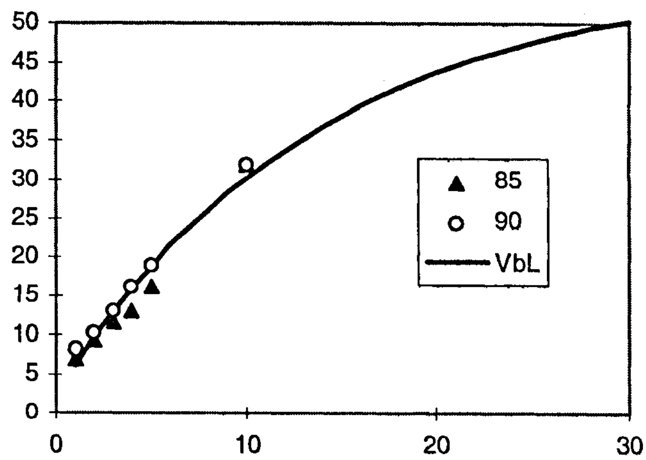
Age determinations

No reliable age determinations exist at present for the redfish stock but preliminary indications from otolith age readings are that the age distribution in the fishable stock includes roughly 9-20 year old fish (unpubl. data). Age determinations are available for cod, but will not be used directly in this paper. In stead the results from age-based assessments will be used for comparative purposes when testing the present assessment procedure.

Length at age

In spite of the lack of age determinations it is clear from available length distributions that a reasonable growth model should provide predicted length at age in the vicinity of 7-8 cm at age 1, with increments of approximately 2-2.5 cm per year for the following 5 years. Also, from existing (unpublished) age readings data, it seems that a 30-35 cm redfish can be 10-12 year old, which indicates an average growth of approximately 2.4 cm annually from age 1.

Figure 3. Von Bertalanffy curve for redfish along with data. Also assumed in the fitting process is that the length at age 10 is 32 cm and at age 50 to be 55 cm.



It is not at all clear how this growth should be interpolated but the Von Bertalanffy growth function ($L_\infty*(1-e^{(-Kl*(t-t0))})$) can be fitted to the peaks for the 1985 and 1991 yearclasses, resulting in the parameters estimates given in the table and fitted curve shown in Fig. 3.

As a further step, growth constants could in principle be estimated as a part of the model.

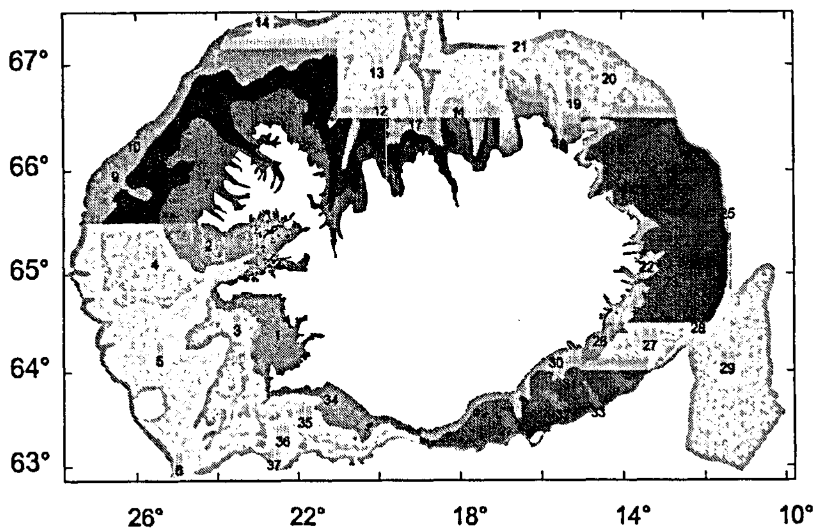
Length-weight relationship

The length-weight relationship $w=\alpha l^{\beta}$ where estimated, based on all available data for the period 1990-1996, both from from survey and catch data. The coefficients are estimated as $\alpha=0.01064$ and $\beta=3.0893$.

Survey abundance data

The length distribution of numbers caught in the groundfish survey is integrated for each station separately in order to obtain groundfish survey biomass measure, U_{sy} for each year, y and station, s (MRI 1997 b). In the process, the length distributions are weighted with the commercial selection pattern so that these indices correspond to a fishable biomass measure by station. In effect the selection pattern for the survey is assumed to be constant over the fishable sizes of redfish and a simple integration procedure is used to come up with a measure of fishable biomass at each station. The resulting abundance by station is then integrated across stations by assuming each station to reflect a specified area. This results in a survey index describing the relative fishable biomass per year.

Figure 4. Stratification for calculation of survey abundance.

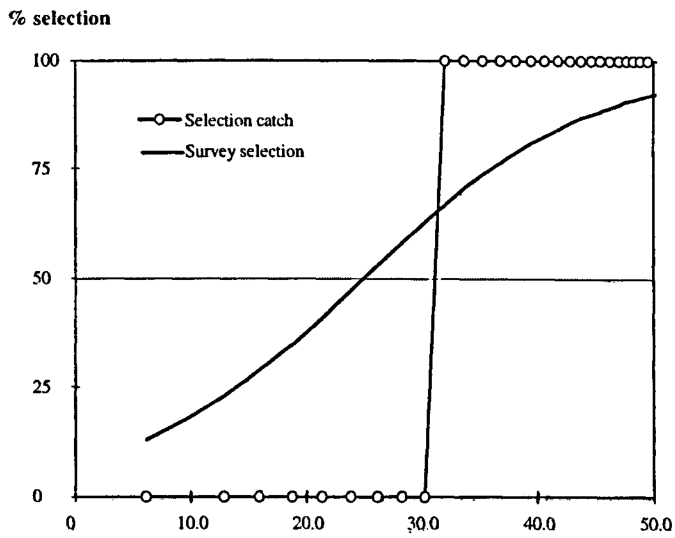


The model

As usual, denote by N_{ay} the population numbers at age a at the beginning of year y, F_{ay} the fishing mortality rate affecting fish of age a in year y and let M denote the natural mortality rate, assumed constant across years and ages. The basic stock equations are as usual given by $N_{a+1,y+1}=N_{ay}\exp(-Z_{ay})$ where $Z_{ay}=F_{ay}+M$ denotes the total mortality rate.

Let R_y denote the recruitment in year y, so that for the first age group, a=0, $N_{0y}=R_y$. If it is assumed that the population at the start of exploitation is a virgin stock which has had constant recruitment, R_0, for a number of years, then the stock size in the initial year is given by $N_{a,0}=R_0\exp(-aM)$ since a yearclass of age a is assumed to have been originally of constant size R_0, but subsequently reduced by a constant natural mortality for a years.

Figure 5. Selection patterns assumed.



For a stock which does not start out at a virgin level, M in the previous equation needs to be replaced with an assumption of an average total mortality up to the given age. This certainly applies to the cod examples. If this is not used, the initial stock biomass levels will be much to high for a stock such as the cod.

Length at age and weight at length are given in the preceding sections. For a given fishing mortality pattern and stock numbers, this data can be used to predict landings.

In order to relate landings in tonnes to an overall average fishing mortality, a selection pattern needs t o be assumed or estimated. For the sake of parsimony, a simple model for the selection pattern is taken as length based, with the relative selection at length l given with $S_l = 1/(1+\exp(-K^C*(l-L^C_{50})))$. In this equation L^C_{50} denotes the length at 50% selection. Naturally, K^C and L^C_{50} are unknown parameters in this equation and need to be estimated or assumed.

Given this selection pattern and length at age, the fishing mortality at age is given by $F_{ay}=F_yS_{l_a}$ where l_a is the length at age a.

Similarly, a survey selection pattern, $S^S_l = 1/(1+\exp(-K^S*(l-L^S_{50})))$ can be defined in order to link the total numbers at age to the numbers observed in the survey.

For given values of annual recruitment, R_y, the fishing mortality rate in each year, F_y, natural mortality, M, selection parameters, K^C and L^C_{50}, it is clear that a complete stock projection can be made for all years. Adding the length-weight relationship allows for the prediction of landings in each years. Conversely, given the landings, the fishing mortality in each year is determined, leaving only the selection parameters, recruitment and natural mortality undetermined.

Natural mortality will in all cases be taken to be fixed at a predetermined value.

Different assumptions can now be used to complete the estimation or prediction of stock sizes. A simple model is to assume constant recruitment throughout and to estimate this average recruitment level based e.g. on the trend in survey abundance. Alternatively, survey length distributions could be used to estimate the relative abundance of the different yearclasses.

Some important additional model components include the fishable biomass given by $B_y=\Sigma_a w_a S_a N_{ay}$ and the predicted catch given by the catch equation so that the landings are given by $Y_y=\Sigma_a w_a F_y S_a N_{ay}(1-\exp(-Z_{ay}))/Z_{ay}$.

Linking the model to data

Given the above population model, a set of assumed recruitment levels will result in a fishable biomass trajectory, B_y. Any CPUE series can in principle be linked to this trajectory through a simple catchability model. However, it is considerably more interesting to link survey data to the model. This can be either done by setting up a survey selection pattern, $S^S_l = 1/(1+\exp(-K^S*(l-L^S_{50})))$ and using this selection pattern to generate a "survey biomass": $B^S_y=\Sigma_l w_l S^S_l N_{ly}$ which links directly to the total abundance data from the survey or by generating an index of fishable biomass from the survey. The latter approach has been taken in this paper. Hence, the model is used to generate a fishable biomass based on $B^C_y=\Sigma_l w_l S^C_l N_{ly}$ where $S^C_l = 1/(1+\exp(-K^C*(l-L^C_{50})))$ is the selection pattern for the commercial fishery.

The survey abundance index of the fishable biomass, U_y, is assumed to be linked to the population biomass through a catchability, so that $U_y=qB_y$.

On log-scale this implies that $\ln(U_y)=\ln(q)+\ln(B_y)$ and it is therefore clear that under the assumption of lognormal errors, the catchability will always be estimated through $\ln(q)=(1/T)\Sigma_y\ln(U_y/B_y)$ where T denotes the total number of survey years. Hence the estimation of catchability does not require any special minimisation. This results in a natural definition of the sum of squared errors for the abundance series: $D^U=\Sigma_a(\ln(U_y)-\ln(qB_y))^2$.

Although a log-transform is often used in this fashion, the initial model will be taken without the log-transform. In this case the definition of the sum of squared errors for the abundance series becomes $D^U=\Sigma_a(U_y-qB_y)^2$ where the point estimate of q is now $\Sigma_y U_y/\Sigma B_y$.

Adding the growth model and assuming that all fish in each age group have the same length results in a certain population length distribution for each year. These length distributions can be based on the survey by using the numbers $S^S_a N_{ay}$ in length group l_a or based on the landings by using the numbers $S^C_a N_{ay}$ in the length group.

Naturally, assuming these discrete steps in the modeled growth curve may result in a rather bizarre histogram. Several approaches can be taken to alleviate this problem. For example, the cumulative distribution function (c.d.f., D_{lyg}) can be used and compared with the empirical distribution function (e.d.f., E_{lyg}) based on the survey (g=S) or commercial (g=C) length distributions for each year (y). The actual comparison can then be based on the integrated squared difference:$D^L=\Sigma_{lyg}(D_{lyg}-E_{lyg})^2$ where the summation is based on a fixed set of integer length values encompassing the observed measurements for each data set. This one possibility of several for comparing c.d.f.'s, whereas several such possibilities exist (the Cramèr von Mises statistic, Anderson-Darling statistic *etc.*, *cf.* Durbin, 1973)

An alternative approach is to use a smoother to smooth the steps in the length distribution. In this case it is possible to compare directly the length distributions (*i.e.* the probability density function, p.d.f.) rather than the cumulative version. The smoother is taken here simply as a description of a Gaussian distribution of each age group around its mean length, *i.e.* a parameter is used to describe the standard deviation of the population length distribution in each age group. Details of how the p.d.f. is linked to the model can make a considerable difference to the results. Examples of measures of the difference between a modeled p.d.f. and data are discussed *e.g.* by MacDonald and Pitcher (1979) in the context of length measurements and by McGullagh and Nelder (1989) in the more general context of generalized linear models.

Examples of measures of difference include for example $D^L=\Sigma_{lyg}(p_{lyg}-\pi_{lyg})^2$ and $D^L=\Sigma_{lyg}(p_{lyg}-\pi_{lyg})^2/\pi_{lyg}(1-\pi_{lyg})$, where p and π denote the observed and true (*i.e.* model) proportions, respectively.

Finally, it is possible that certain model parameter values will result in population trajectories which cannot sustain the historical catches. Thus there is a further error term defined by the difference between the observed catch (Y^O) and modeled catch (Y^M), given e.g. by $D^Y=\Sigma_y(\ln Y^O{}_y-\ln Y^M{}_y)^2$.

The total error to be made as small as possible is now given by $D= D^U+D^L+D^Y$, although in some cases weights may be attached to each component.

It should be noted that all applications of the method have yielded results where $D^Y=0$. It should also be noted that using an arbitrary set of growth curves, selection pattern and recruitment levels is highly likely to lead to spurious results.

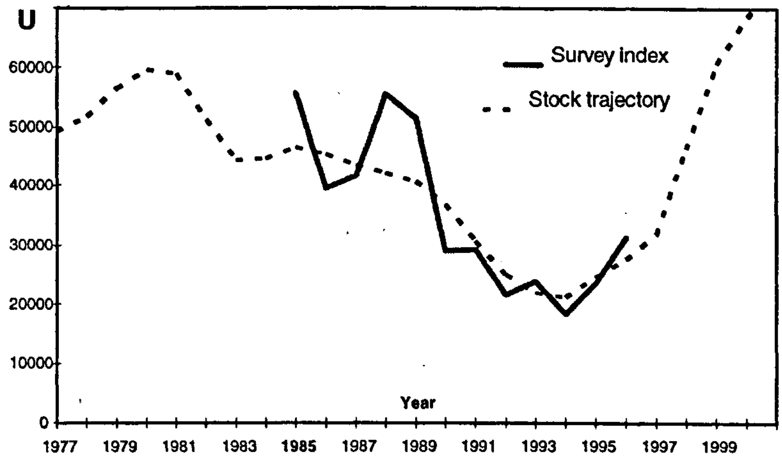
Initial values

The model is started with a set of default values for all components. This includes the estimated von Bertalanffy and assumed selection parameters mentioned above.

	Parameter	Initial value
von B	t_0	-0.60
	K	0.07
	L_∞	56.7
Survey sel	K^S	0.1
	$L^S{}_{50}$	25
Comm. sel	K^C	10
	$L^C{}_{50}$	31
Nat. mort.	M	0.05

Model testing

Figure 6. Iceland cod. Fitted stock trajectory along with survey data.

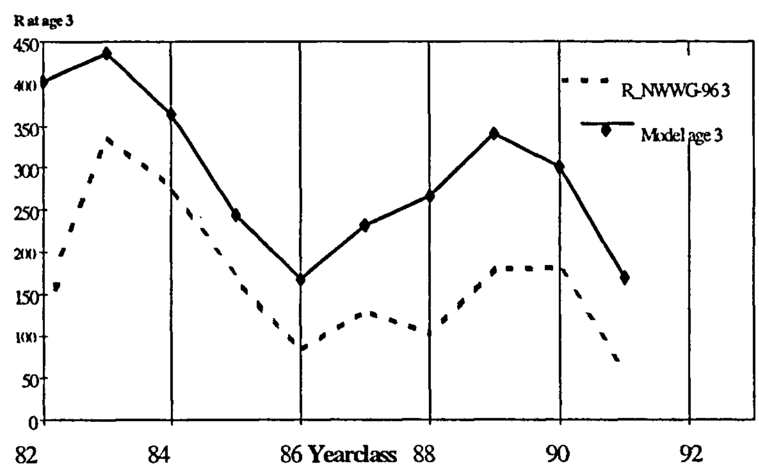


The model was applied to the Icelandic cod (*Gadus morhua*) for comparison with the standard methods of estimating the state of the stock. The input parameters corresponding to those used for the redfish were defined as shown in the textable above.

The stock abundance trend along with the fitted stock trajectory is given in Fig. 6. It is seen that the observed decline is tracked overall, although the individual years can of course deviate somewhat.

In this figure, the future projections of the stock are given with a very low fishing mortality of around $F_{0.1}$ or about 0.2. Hence the projected stock size is seen to increase dramatically.

Figure 7. Cod recruitment as estimated from redfish model and VPA (1996 assessment).



The recruitment estimates as obtained from applying the redfish model and from the 1996 NWWG report are given in Fig. 7. It is seen that the new method gives the same overall trend, although the level is a bit off.

This conclusion depends somewhat on the time period considered. Notably, there are restrictive assumptions on the first year and the yearclasses which are not observed from age 1 onwards are very poorly estimated.

Application to *S. marinus*

A sequence of models were fit to different subsets of the *S. marinus* data. In order to verify the estimability and consistency of different parameters, the initial model (0) consists simply of fixing all parameters at default values and only estimating a single average recruitment level. In order to avoid confounding due to growth model specifications and selection in the catches, the first runs are based on the length distributions for redfish below 25cm along with the survey index.

Some special attention has to be given to the weighting factors for *e.g.* the length distributions of juveniles *etc.* It is clear that if the full length distribution is used as a minimization criterion, then an error in the growth function for the oldest fish may adversely affect mortality and recruitment estimates, since the abundance of old fish is a function of all these factors. Therefore, while emphasizing recruitment (and fixing growth), only the recruiting portion of the length distribution is used. Further, the two components of the sum of squared error (SSE), *i.e.* from the length distribution and the survey distribution are scaled so that the give (roughly) the same contribution to the overall SSE.

A sequence of more complex models are now estimated, starting from a single recruitment level (0), then considering three different recruitment levels, one for the initial stock, one level for the typical (poor) yearclass and a multiplier to describe how much larger the good 1980, 1985 and 1990 yearclasses are compared to the low ones. Model (2) then goes on to estimate each yearclass separately.

The following table shows the results from these 3 estimations. It is seen that the largest single SSE-decrease is observed in the length distributions when going from a single average recruitment level to 3 levels. It is also seen that there is much less pronounced SSE decrease when going to estimating each recruiting yearclass separately.

Table 2. Sums of squares and some parameter values.

Model	SSE(L)	SSE(U)	Avg R	mult	low	equil	recruitment
0	517	8104	188	1	188	188	Average recr only
1	89	6765	83	2.7	57	245	3 recruitment levels
2	78	4529	83			252	All recs estimated
3	262	1622	110				Heavy wt to U
4	57	6955	81				Heavy wt to L

The last two lines in the table provide the summary of results obtained when the length and index information is given unequal weight (each is multiplied by 100). It is clear that each term can be made much smaller than in the equal-weights case, and this is particularly true of the survey index.

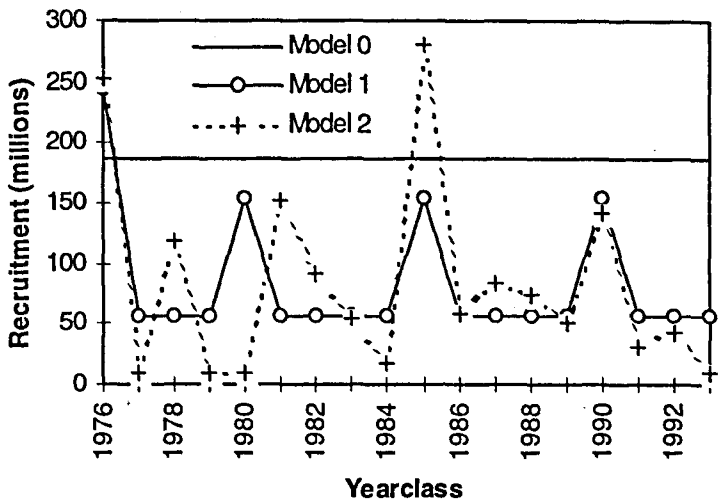
Taken at face value this seems to indicate that the two criteria, lengths and indices, are giving contradictory indications as to what the point estimates should be.

A more realistic interpretation is, however, that model misspecification or lack of model flexibility does not allow both criteria to be minimized well enough simultaneously.

It would therefore seem clear that there is information in each data set which can be used to provide some more flexibility in the model than is already used. In particular it may be possible to estimate some growth parameters (possibly some annual variation in these), or even the standard deviation of length at age.

On the other hand it is also clear that it is highly dubious whether this model can really be used to estimate every single yearclass strength. The reason for this is that the length distributions do not really provide the discriminatory power needed to decide whether *e.g.* the 1980 yearclass is large and surrounding ones are weak or whether e.g. the 1978 and 1981 yearclasses are above average and the 1980 yearclass is weak. The data does, however, have the power to detect that there is stronger recruitment sometime around 1980 than is the norm.

Figure 8. *S. marinus* recruitment at age 1 as estimated by the model



In principle the different models can be compared using F-tests, or even χ^2-tests. It is not at all clear, however, what the various degrees of freedom should be in such comparisons. Although the survey index SSE may reasonably be assumed to have degrees of freedom equal to the number of years, the same is not true of the contribution from the length distributions nor of the sum of these two components.

Discussion

The preliminary results presented here indicate that it is feasible to use an age-based population model to estimate population sizes for fairly long-lived fish stocks using only length-based and aggregated abundance data. Validations of the model indicate that the results for cod are fairly consistent with VPA results, but the method can also be used in the case where there are no age-readings available and such data are not used in the model.

Further work should concentrate on completing the model definition, particularly the optimization criteria. The model should then be used for estimation in such a fashion as to obtain point estimates for estimable quantities but quantities which are hard to estimate should be left at assumed levels.

It is quite likely that the inclusion of Bayesian priors would enhance the estimation procedure considerably, particularly since this automatically resolves some of the finer estimation details in a consistent fashion.

More immediate future work should include detailed tests of the effects of different definitions of the sums of squared errors as well as the possibility of estimating more parameters.

References

ACFM 1992. Report of the Advisory Committee on Fishery Management. ICES Coop. Res. Rep. 193.

Durbin, J. 1973. Distribution theory for tests based on the sample distribution function. Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM). Pennsylvania.

MacDonald, P. D. M. and Pitcher, T. J. (1979). Age-groups from size-frequency data : a versatile and efficient method of analyzing distribution mixtures. J. Fish. Res. Board Can. 36 987-1001

McCullagh, P. and Nelder, J. A. 1989. Generalized linear models. Chapman and Hall. London. 511pp.

MWG 1995. Report of the Working Group on Methods of Fish Stock Assessments. ICES C.M. 1995/Assess:11.

MRI 1997 a. State of Marine Stocks in Icelandic Waters 1996/97. Prosects for the Quota year 1997/98. Hafrannsóknastofnun, fjölrit nr 56 (in Icelandic with english summary).170 p.

MRI 1997 b. The Icelandic groundfish survey. Marine Research Institute Technical Report.

NWWG 1992. Report of the North Western Working Group. ICES C. M. 1992/Assess.

NWWG 1995. Report of the North Western Working Group. ICES C. M. 1995/Assess.

NWWG 1997. Report of the North Western Working Group. ICES C. M. 1997/Assess.

Pálsson Ó.K., Jónsson, E., Schopka, S.A., Stefánsson, G, Steinarsson, B.Æ., (1989). Icelandic groundfish survey data used to improve precision in stock assessments. Journal of Northwest Atlantic Fisheries Science, 9, 53-72.