COOPERATIVE RESEARCH REPORT

NO. 133

REPORT OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENTS

https://doi.org/10.17895/ices.pub.7935

ISBN 978-87-7482-607-1

ISSN 2707-7144

International Council for the Exploration of the Sea
Palægade 2-4, 1261 Copenhagen K
Denmark

September, 1985

"The Delphi Method" (by M.Hilden)

See J Zuboy (1981).

TABLE OF CONTENTS

ctd.

Table of Contents (ctd) Page

-o-o-o-

REPORT OF THE WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENTS

1. INTRODUCTION

1.1 Participants

E Aro	Finland
D W Armstrong	U.K. (Scotland)
F A v. Beek	Netherlands
H Caswell	USA
R Chevalier	France
H P Cornus	Federal Republic of Germany
G Gudmundsson	Iceland
T Helgason	Iceland
M P Hildén	Finland
T Jakobsen	Norway
P Kanneworff	Denmark
A Kristiansen	Denmark (Faroes)
A Laurec	France
J-P Lussiaà-Berdou	Canada
J J Maguire	Canada
J Modin	Sweden
R Mohn	Canada
S A Murawski	USA
J G Pope	U.K. (England & Wales)
D Rivard	Canada
S A Schopka	Iceland
J G Shepherd (Chairman)	U.K. (England & Wales)

Mr K Hoydal, ICES Statistician, also participated in the meeting.

1.2 Terms of Reference

At the Statutory Meeting in 1983 it was decided (C.Res.1983/2:8:16) that the Working Group on Methods of Fish Stock Assessments (Chairman: Dr J G Shepherd) should meet at ICES headquarters from 11-15 June 1984 to:

(i) propose methods for estimating recruitment in the short term,

(ii) propose simple methods for computing TACs,

(iii) evaluate and make recommendations on the use of regressions in fish stock assessments,

(iv) reconsider ways to calculate biological reference points.

1.3 Working Papers

Working papers were available on topics (i) to (iii), and these are listed in Appendix A. Where the material has not been published elsewhere, the content of these has, where appropriate, been summarised in this report. The reports of the *ad hoc* Working Group on the Use of Effort Data in Assessments (1981) and the Working Group on Methods of Fish Stock Assessments (1983) have now been reprinted in the ICES 'Cooperative Research Report' series (Anon., 1984a).

1.4 Notation

The Working Group adhered as far as possible to the standard notation used previously; an updated summary is given in Appendix B.

1.5 Work carried out

The Working Group was able to consider topics (i) to (iii) in some detail, and the results of this work are reported in Sections 4, 2 and 3 of the report. No written comments had been received on the question of biological reference points (topic (iv)), although a verbal account of the discussion which took place at the 1983 Statutory Meeting was given. This indicated that some participants had not been entirely convinced of the utility of the reference points F_{high} and F_{low} proposed in 1983, and that there had perhaps been some confusion concerning the manner in which these should be used, and what function they could reasonably be expected to perform. A brief recapitulation of the purpose and method of use is therefore included in Section 5.3 of this report, which also contains a short account of recent work in the USA extending the use of these concepts.

Reports of further work on the other topics considered in 1983 were also received and are reported in Section 5.1 and 5.2.

2. SIMPLER METHODS FOR CATCH FORECASTING

2.1 Background

It has been clear for some years that the full analytical assessment procedure (usually VPA plus a catch forecast) for short-term predictions is a rather circular procedure.

VPA is really little more than a transformation of a fairly large and expensive data set (many years of catch-at-age data) into an alternative representation (fishing mortalities and population numbers). The process of catch forecasting recombines these derived quantities, and sums over ages. This process is time-consuming and occasionally contentious, and various efforts have been made to lay bare the essentials of the procedure, and see if and how it may be simplified.

The work of Pope (1983) on ANOVA TACs (discussed in a little detail in Section 2.2 below) was a considerable advance, since he was able to construct an alternative method for the analysis and use of catch-at-age data, which does not depend on any assumption concerning natural mortality, and is only weakly dependent on the estimated level of terminal fishing mortality.

Approaching the problem from the other end, various people have experimented with empirical methods based on time-series analysis (Mendelsohn, 1981; Stocker and Hilborn, 1981). Work aimed at bridging the gap between age-structured and stock/production models (see e.g., Deriso, 1980) is also relevant, although the assumption of deterministic recruitment is usually inadequate unless recruitment is very stable or exploitation rates are low.

More recently Shepherd (1984) has shown how a catch forecast may be constructed from time series of catch and recruitment data only, provided certain rather crude approximations are made. This formalises explicitly the rather obvious fact that a forecast catch is composed of a contribution from the survivors of the old stock, and a contribution from new recruits. It therefore provided a biologically-based framework for the construction of time-series based methods.

The justification for, and performance of some of these simpler methods are discussed below. Further discussion and test results may be found in the Working Paper No.2 (Pope, 1984).

Most of the simpler methods aim to estimate catch if fishing mortality is maintained at its most recent level, the <u>status quo</u> catch (SQC) as defined by Pope (1983). Shepherd (1984) discusses how SQC estimates may be used to derive TACs for various management objectives. It is essential to stress that the calculation of an SQC does <u>not</u> imply that a TAC should be set at that level. The SQC is merely the central ingredient of any catch forecast. Terms such as "short-cut TACs" sometimes applied to these methods are merely convenient but loose terminology, and should not deter managers from setting TACs well above or below the SQC when they have adequate justification.

2.2 The ANOVA Method

The VPA prognosis technique of predicting future catch levels has two steps. The first is fitting the data to a model, a descriptive step. The second is estimating the future catch, a predictive step. The usual approach to the descriptive phase is done by tuning a VPA, a process which may be fairly subjective. The technique called ANOVA TAC (Analysis of Variance, Total Allowable Catch) allows one to estimate a future catch in an objective fashion from basically the same data as are currently used (i.e., catch-at-age, recruitment indices and effort data). ANOVA TAC is comewhat a misnomer as its output is not a TAC but rather an anticipated catch level and the method is not strictly ANOVA. The method is described in Pope (1983), which is based upon some of the results in Pope and Shepherd (1982).

Catch-at-age data may be simply described as being the resultant of three effects. First, an effect which is a function of age and is usually associated with selectivity, S(a). Secondly, an effect over time, generally years, which is a fishing mortality, F(y). Finally, an effect working diagonally in age and time due to the cohort size. By taking the log of the ratio of successive catches down a cohort, the year class strength effect is removed leaving age and year terms as the principle determinants of the transformed data.

$$D(y,a) = \ell n \left\{ C(y+1, a+1) / C(y,a) \right\} \qquad 2.2.1$$

The remaining year effects $\alpha(y)$ and age effects $\beta(a)$ may be thought of as analogs to the marginal description in a 2-way analysis of variance.

$$D(y,a) = \alpha(y) + \beta(a) + \mu + \varepsilon \qquad 2.2.2$$

without explicit expression of the interaction term. (For a simple separable fishery this will be small but will be larger if selection has changed with time (Anon., 1984, Part II)).

The α's are then related to effort data in order to predict future year effects. A linear regression is performed between the α's and the expression

$$\ell n \left\{ E(y+1)^{\frac{1}{2}} / E(y)^{3/2} \right\} \qquad 2.2.3$$

leading to a relationship of the form

$$\alpha(y) = a + 1/2 \, \ell n \left[E(y + 1)/E(y)^{3} \right] \qquad 2.2.4$$

Because the α's must sum to zero a can be determined by summing 2.2.4 and the predicted year effects become

$$\alpha(t+1) = \alpha(t) = a - \ln[E(t)] \qquad 2.2.5$$

with the assumption that $F(y+2) = F(y+1) = F(y)$, i.e., the status quo assumption. Because it has been assumed that effort and fishing mortality are proportional, this is equivalent to assuming a constant effort level over the period of prediction.

The other parameters from the ANOVA model, the age effect β which results from partial recruitment effects, and μ may be combined to predict catches from catch-at-age data for the terminal year.

$$C(t+1,j+1) = C(t,y) \quad \exp\{\alpha(t) + \beta(j) + \mu + \sigma_1^2/2\} \qquad 2.2.6$$

where σ_1^2 is the correction for the log-transformation of the data and σ is estimated from the mean square error from the ANOVA. The division by 2 is an approximation and would better be replaced by the correction given by Pennington (1983).

In order to obtain estimates for the recruitment to the exploited stock for years t and t+1, a relationship is established between indices of recruitment, probably from survey data, $R(y,a)$ and catch and effort data. The proposed method for doing this is to assume a proportionality coefficient $r(a)$ which relates the catch at age to the effort applied to the incoming recruits. The actual form chosen by Pope (1983) for estimating this coefficient is

$$\bar{r}(a) = \left\{\prod_{y=1}^{t-1} C(y,a)/E(y)\ R(y,a)\right\}^{1/t-1} \exp \sigma_2{}^2/2 \qquad 2.2.7$$

where $\sigma_2{}^2/2$ is a correction similar to $\sigma_1 2/2$ in Equation 2.2.6 and is estimated from Var $(C(y,a)/E(y)\ R(y,a))$.

We may now write the estimated catch for the recruits to the year t+2 as

$$C(t+2,a) = E(t)\ R\ (t+2,a)\ \bar{r}\ (a) \qquad a = 1 \text{ to } 3 \qquad 2.2.8$$

Recalling that the older age classes' contributions are found by 2.2.6, the entire catch for the year t+2 is predicted. The catches at age may be multiplied by weight at age to give the anticipated yield.

The ANOVA TAC may be thought of as being closely related to Separable VPA, but instead of estimating fishing mortality and population from the catch data, the descriptive phase mentioned above, it is used to directly estimate a future catch. The prediction requires effort and recruitment information. The ANOVA TAC is a simplification of the usual VPA approach in that it is less subjective and does not require tuning, although it has the same data requirements. The reason for its development was to investigate the relationship between data quality and the quality of predicted catch levels; it could, however, be regarded as a short-cut method. The short-cut TACs described below require much less data (normally catch at age is not required). Therefore, the ANOVA TAC is in some sense intermediate between the short-cut and the usual VPA approach (Pope, 1984). To date it has not been tested on a complete real data set. Pope (1983) tests it with real catch-at-age and recruitment data, but with fishing effort derived from fishing mortality from separable VPA which would assure a good relationship between F and E.

2.3 The SHOT Method (Shepherd's Hang Over TAC)

Shepherd (1984) proposed a simple method for making estimates of catch forecasts when few data other than annual catches are available. Th method calculates the expected catch level under the assumption that no change in fishing mortality will occur in the year(s) covered by the forecast (i.e., status quo is maintained). Variants of the method allow one to incorporate information on recruitment and stock size (e.g., catch rate index), if such information is available.

The exploited biomass at the beginning of year n + 1 is given by the previous biomass, as modified by catches, stock production due to recruitment during the year P(n), and the effects of growth and natural mortality

$$B(n + 1) = B(n) - Y(n) + P(n) + (G - M)\,\bar{B}\,(n) \qquad (1)$$

Shepherd (1984) argues that (G - M) is a small fraction of one and suggests that the last term of Equation (1) can be neglected. Effectively, this assumes that growth in weight of exploited fish roughly cancels losses due to natural mortality; we note that this may not in fact be a good approximation if fishing mortality varies over a wide range. When $\tilde{F}(n+1) = \tilde{F}(n) = \tilde{F}$, Equation (1) can be used to show that the catch in year n + 1 is given by

$$Y(n + 1) = (1 - \tilde{F})\,Y(n) + \tilde{F}\,P(n) \qquad (2)$$

where $\tilde{F}$ is used to denote a yield/biomass ratio.

In other words, the status quo catch is just a weighted average of last year's catch and the production due to new recruits. While Equation (2) provides an estimate of catch only one year ahead, the formula can be repeatedly applied to provide estimates further ahead if required. Such an extension of the forecasting horizon will require additional assumptions about Y and F in the intervening year(s).

The estimation of P(n) may be carried out in many possible ways, depending upon the type of data available. Specifically, Shepherd (1984) considers the following cases:

A. If only catch information is available, the status quo catch can be defined in its simplest form as

$$Y\,(n + 1) = Y_{sq} = (1 - \tilde{F})\,Y(n) + \tilde{F}\,\bar{Y} \qquad (3)$$

where $\bar{Y}$ is the average catch over a number of years. This result relies on the assumption that recruitment is near average. It will therefore fail to give sufficiently conservative results for a declining stock. The relative weights $\tilde{F}$ and $(1 - \tilde{F})$ depend on the level of fishing mortality, but the dependence of the estimated status quo catch on $\tilde{F}$ is quite weak because the result must lie somewhere between Y(n) and $\bar{Y}$. Equation (3) is closely related to an autoregressive model. However, Equation (3) arises from explicit assumptions regarding stock dynamics and target fishing mortalities (i.e., status quo) and, consequently, we know under which conditions the model is likely to apply. A purely empirical autoregressive model (i.e., one derived solely from the inspection of autocorrelations) would have provided no information on these conditions. Finally, it is noted that as one forecasts further and further into the future, the estimated status quo catch reduces to using the average catch.

B. If an index of recruitment $r(n)$ is available, then the status quo catch can be defined as

$$Y(n+1) = Ysq = (1 - \tilde{F})\, Y(n) + \tilde{F}\,\frac{\bar{P}}{\bar{r}}\, r(n) \qquad (4)$$

where $\bar{r}$ is the mean index of recruitment. If F is assumed constant over a number of years, the coefficients of Equation (4) can be found by regression of $Y(n+1)$ on $Y(n)$ and $r(n)$. A zero intercept is indicated by the model, but may not give a predictor with optimal properties.

C. If both an index of recruitment and an index of stock size are available, Shepherd (1984) shows that the status quo catch can be defined as

$$Ysq = (1 - \tilde{F})\, Y(n) + \frac{r(n)}{\bar{r}} \left\{ \tilde{F}\,\bar{Y} + Y(n)\,\frac{(1 - R)}{(n - 1)} \right\} \qquad (5)$$

where R is the ratio between the initial and the final stock size (in practice, R can be calculated from the stock size index). Equation (5) relies on the assumption that the index of recruitment $\{r(n)\}$ is directly proportional to the production due to recruits $\{P(n)\}$. Clearly, if recruitment in year n is near average and if stock size is near average, Equation (5) reduces to Equation (3).

To use these models, some estimate of $\tilde{F}(n)$ is required, which estimate is often hard to acquire. Simply guessing a value (or a range of values) may be tolerable in desperate cases. However, Brander (pers.comm.) has pointed out that when F has remained reasonably constant for a number of years, the regression of $Y(n+1)$ on $Y(n)$ can be used to obtain an estimate of F. As suggested by Equation (3), the slope of the regression is just $(1 - \tilde{F})$, and the intercept is related to average yield and production.

2.4 The DROP and DOPE Methods

Deriso (1980) suggested an equilibrium model which forecasts relative yields from biomass estimates in previous years and a stock/recruitment relation. The approach was examined further by Roff (1983), who tested its usefulness as a short-term predictor. Pope (1984) extended the approach to include recruitment field data (DROP method, after Deriso, Roff, Pope) and also provided a variant which uses catch in numbers as additional input (DOPE method after Deriso, Pope).

Deriso's method is appealing since it approximates the behaviour of an age-structured model of a more complicated structure. This is achieved by the inclusion of the Brody growth coefficient as well as survival fractions. The original approach assumed constant annual mortalities. However, the DROP and DOPE versions allow for fractional adjustments of fishing mortalities, i.e., a status quo catch can be defined.

The Deriso, DROP and DOPE family can be derived from a Brody growth equation formulated as Ford-Walford relations, i.e.,

$$W(a+1) = (1+\rho)\, W(a) - \rho\, W(a-1) = \rho\, W(a) + (1-\rho)\, W_\infty \qquad 2.4.3$$

It should be noted that the Brody equation lacks an inflection point and therefore strictly only applies to older fish when applied to weight data. For several stocks, age at recruitment is not sufficiently high to meet this criterion.

Modifying Deriso's Equation 2, one may describe the exploitable biomass B(y+1) during year y+1 as

$$B(y+1) = B(y)\ (1+\rho)\ \exp(-Z) - B(y-1)\ \rho\ \exp\ (-2Z) + R(y+1) \qquad 2.4.1$$

where R(y+1) is the recruiting biomass,

ρ is the Brody growth weight coefficient

exp(-Z) is the survival proportion.

Thus, the forecast B(y+1) is determined by the growth weighted average of surviving biomasses in the two preceding years and by the recruitment biomass. Adult survival is assumed constant from year to year. The first term on the right hand side of the above equation shows that if B(y) was high, then B(y+1) should be higher. This is adjusted by a 'hangover' effect in the second term, which allows for a high growth rate of the adults provided B(y) was high due to recruitment rather than a large stock size. Similarly, a high ρ results in a high growth potential, i.e., there is a relatively large proportion of fast growing adults.

The DROP method is based on Equation 2.4.1 and can be further transformed into a yield function. Since the catch is Y(y)=F(y) B(y), then

$$Y(y+1) = Y(y)\ (1+\rho)\ \exp(-Z)\ \frac{F(y+1)}{F(y)}$$

$$- Y(y-1)\ \rho\ \exp\ (-2Z)\ \frac{F(y+1)}{F(y-1)} + R(y+1)\ F(y+1)$$

If a status quo situation for the year y+1 is desired, then set F(y+1) = pF(y) and the DROP equation will reduce to

$$Y(t+1) = p\ Y(t)\ (1+\rho)\ \exp(-Z) - p\ Y(t-1)\ \frac{F(t)}{F(t-1)}\ \rho\ \exp(-2Z)$$

$$+ p\ R(t+1)\ F(t) \qquad 2.4.2$$

Equation 2.4.1 and hence Equation 2.4.2 could be parameterised by using biological data to estimate ρ exp(-Z) and F(t). Alternatively, plausible guesses might be made or the parameters could be fitted by making a multiple regression of Y(y) on Y(y-1), Y(y-2) and R(y).

In practice, some composite methods might prove most satisfactory such as estimating or guessing ρ and exp (-Z) and then regressing

$$Y(y) - Y(y-1)\ (1+\rho)\ \exp(-Z) + Y(y-2)\ \rho\ \exp(-2Z)$$

against the recruitment index for year y.

The DOPE method can be derived from the right-hand expression in Equation 2.4.3 (see p. 6). In short, stock in numbers at age and year is multiplied into the expression,and ages beyond recruitment are summed up. The exploitable biomass (i.e., including recruitment) then becomes

$$B(y+1) = B(y)\ \rho\ \exp(-Z) + \left[\sum_{a>r} N(a,y)\right]\ (1-\rho)W_\infty\ \exp(-Z) + R(y+1)$$

Since $\sum_{a>r} N(a,y) = B(y)/\overline{W}(y)$ where $\overline{W}(y)$ is the average weight of catchable fish in year y, the equation will be:

$$B(y+1) = B(y) \left[\rho + (1-\rho)\, W_\infty/\bar{W}(y)\right] \exp(-Z) + R(y+1) \qquad 2.4.4$$

and the dependence of B(y+1) on the average weight of fish in year y is made more apparent.

A yield function describing a status quo situation can be obtained in a similar way as for Equation 2.4.2. As the catch in numbers $C(y) = Y(y)/\bar{W}(y)$, the following formula is derived:

$$Y(y+1) = PY(y)\, \rho \exp(-Z) + PC(y)\, (1-\rho)\, W_\infty \exp(-Z) + P\, F(y)\, R(y+1) \qquad 2.4.5$$

Equation 2.4.5 might be fitted as a multiple regression or by biological estimation of the various constants. It should be stressed that the possibility to independently estimate constants within a model enhances the probability of a good approximation.

The DOPE method has a possible advantage over the DROP method, since (if size compositions are available), catch in numbers or mean weight is easier to estimate than a ratio of catch rates.

Both approaches would probably benefit if selection effects S(a) could be included, e.g., $S(a+1)\, W(a+1) = \rho\, S(a)\, W(a) +$ constant. This is the case in real life.

2.5 The Use of Time-Series Methods

Several recent papers have considered the applicability of time-series analysis methods for the forecasting of trends in various fisheries data (Boudreault *et al.*, 1977; Saila *et al.*, 1980; Mendelssohn, 1981; Kirkley *et al.*, 1982; Fogarty, 1984 (MS)). Specifically, these papers have employed the ARIMA (Autoregressive-Integrated Moving Average) models developed principally by Box and Jenkins (1976). In general, ARIMA models are a flexible and powerful class of linear stochastic difference equation predictors. ARIMA models are generally based on predicting a value in a time-series based on a linear combination of its own past values, past errors (called shocks or interventions), and past values of other time series that may explain the objective (in the multivariate case).

With respect to the problem of catch prediction (forecasting), ARIMA models may be applied based on a univariate time series (simply a time history of yield to the fishery). Alternatively, multivariate time-series models (called transfer function models) may be employed. The multivariate models employ some influential variable in the forecasting procedure (such as using explanatory environmental variables in recruitment forecasts). In this case the most likely candidate is data on an index of recruitment.

The general 'rule-of-thumb' for time-series analyses under ARIMA procedures is that the number of observations should be more than about 30 and less than 2 000. With thousands the method requires much computer time and memory. With fewer than 30 observations the parameters are generally not estimated very well, although the number required depends crucially on the noise (unexplained variability), which is usually high in this context.

Time series in fisheries data of greatest length are generally historical catch (landings) statistics, fishing effort, and environmental observations (e.g., temperature data). Unfortunately, time series of recruitment sampling are generally too short to be considered reliable for use in such models.

A recent paper by Fogarty (1984 (MS)) illustrates the use of univariate and transfer function (multivariate) time-series procedures for forecasting American lobster landings. The time series of landings data available for analysis was 1928-1981 (1982 data were reserved for comparison with predicted landings). The univariate time-series models resulted in predicted 1982 yields that were within 4% of the observed catch (Figure 2.5.1). A transfer function (multivariate) time-series model was fitted with lagged sea water temperature as an independent variable. A significant temperature effect at a time lag of 6 years (approximately the length of time from spawning to recruitment to the fishery) resulted in a reduction in residual variance of approximately 13% relative to the corresponding univariate model (Figure 2.5.2).

Other catch projections that have been made successfully with ARIMA procedures are those that predict monthly changes that typically follow some harmonic trend (Mendelssohn, 1981; Fogarty, 1984 (MS)).

2.6 The Performance of the Simpler Methods

In the time available to the Working Group it was not possible to fully test the various short-cut methods. Nevertheless, a start was made on testing in order to focus on what further work would be needed. Two approaches were adopted. These were:

(a) Tests of prediction methods in simulated data sets.

(b) Tests of predictions on real stocks.

2.6.1 Tests on simulated data sets

The performance of the SHOT, DROP and DOPE methods was tested on data from 3 simulated stocks.

Stocks were simulated based upon the recent structure of the North Sea haddock. As pointed out (Pope, 1984), the choice of this fish should tend to favour the DROP and DOPE methods over the SHOT method.

Data were generated for a period of 26 years for various assumptions on range of recruitment and variation in fishing effort from year to year as follows.

Conventional predictive regressions were used, since minimal variability is a practically desirable property of TAC estimates.

Stock 1: High recruitment variation
Low fishing effort variation

Stock 2: Moderate recruitment variation
Low fishing effort variation

Stock 3: Low recruitment variation
High fishing effort variation.

Implementation of the various methods consisted of carrying out the following multiple regressions for each stock:

Stock 1:

SHOT: Y(t+1) on Y(t), R(1, t+1)

DROP: Y(t+1) on Y(t), Y(t-1), R(2, t+1)

DOPE: Y(t+1) on Y(t), C(t), R(2, t+1)

Stock 2:

SHOT: Y(t+1) on Y(t), R(2,t+1)

DROP: Y(t+1) on Y(t), Y(t-1), R(2,t+1)

DOPE: Y(t+1) on Y(t), C(t) , R(2,t+1)

Stock 3:

SHOT: Y(t+1) on Y(t), R(2,t+1)

DROP: Y(t+1) on Y(t), Y(t-1), R(2,t+1)

DOPE: Y(t+1) on Y(t), C(t), R(2,t+1)

To be as fair as possible to each method, three recruitment series were adopted. (The numbers of recruits in year t+1 at ages 0, 1 and 2.) For each method and stock the fit to the full data set was made and the recruitment series which effectively minimised the residual variation chosen (minor deviations from this rule were allowed to give consistency). In practice, the 2 year olds were used for all but the SHOT method on Stock 1.

Having made this choice, the data from the first 16 years were fitted and used to forecast the *status quo* catch in the 17th year. The data from the first 17 years were then fitted and used to predict the *status quo* catch in year 18 and so on. This is an approach used by Stocker and Hilborn (1981). The predictions from the three methods for the three simulated stocks are shown in Tables 2.6.1, 2.6.2 and 2.6.3. Also shown are the mean squared deviations of the projected results (D) from the true values and the statistics $1-D/\sigma^2$, which Stocker and Hilborn suggest as similar to a coefficient of determination.

For Stock 1 with extremely variable recruitment (the largest recruitment was 500 times the smallest) and steadily declining effort, the SHOT method did not perform very well ($1-D/\sigma^2 = .35$). The DROP and DOPE methods, however, performed more creditably with $1-D/\sigma^2$ being .59 and .78, respectively. The DOPE method thus predicted 78% of the variation in yield over the 10 prediction years. This is in line with the theoretical arguments of Pope (1984). Predictions for all the methods appear to improve for the later years, and this may be a function of the number of years for which fitting data were available. This might possibly argue for the choice of biologically sensible coefficient values rather than multiple regression generated coefficient values when time series are short. This is a point which might be addressed by further simulations.

For Stock 2 with less variable recruitment and steadily declining effort, the methods all performed quite well with $1-D/\sigma^2$ of .67, .78 and .76, respectively, thus supporting the argument that very variable recruitment would have a more serious effect on the SHOT than the DROP and DOPE methods.

For Stock 3 with still less variable recruitment but with erratic but largely trendless effort data, the SHOT method performed better than the other two methods. The reason that it should be better than the DROP method is presumably that the latter method ignores an effort correction to the second term. The reason for the improvement over the DOPE method is, however, less clear.

From the above work the following tentative conclusions emerge:

(1) All methods reduced the variation compared to using just the mean catch, on three quite exacting stocks.

(2) If the variability of recruitment is very high, the DOPE and DROP methods are likely to perform better than the SHOT method.

(3) If most of the variation in yield is due to short-term fluctuation in effort, then the SHOT method seems to be preferable to the DOPE or DROP.

Clearly, these conclusions need to be examined much more exhaustively on real and simulated data. Equally clearly we need to consider how best to construct predictors. This might be done on biological grounds or by multiple regression techniques.

2.6.2 Short-cut methods applied to real stocks

Georges Bank scallop data

Yield data are available for this data set from 1953 to 1983. Catch at age data and hence VPA recruitment estimates are available from 1972.

Using the yield data set only, two models were fitted:

$$Y(t+1) = a + b\,Y(t)$$

and $Y(t+1) = a + b\,Y(t) + c\,Y(t-1)$

Figure 2.6.1 shows the plot of $Y(t+1)$ on $Y(t)$.

The results of the fishing on years 1953-82 were used to predict the yield in 1983, and the results of the fishing on years 1953-81 were used to predict the yield in 1982. Results are shown in Table 2.6.4.

The second data set was more detailed in that it contained recruitment and catch in number from a cohort analysis for the period 1972-83. Table 2.6.4 summarises the results. It was found that $R(3,t)$ performed much better than $R(2,t)$. This observation is consistent with the low partial recruitment observed for the youngest age class in the fishery.

The results suggest that given an appropriate and reliable recruitment index the SHOT, DROP and DOPE methods should greatly improve predictions over a simpled lagged correlation. This is not surprising for a recruitment-dominated fishery such as this. The results also suggest that the R^2 of regressions is not useful in discriminating amongst methods.

Baltic herring data

The SHOT, DROP and DOPE methods were used on the Baltic herring of area 29NE. The data series available was short (8 years), but when the models were fitted to the first 7 years of data with multiple regression, all the methods produced catch forecasts not differing more than ±10% of the observed value. However, the regression coefficient obtained from the multiple regression could not be given any meaningful interpretation, since the regression coefficients of recruitment consistently come out negative.

The age at which recruitment is assumed to occur (i.e., the lag used for the recruitment series) needs to be selected with great care, and further tests may be necessary for these data.

It is difficult, therefore, to draw any consistent conclusions from this data set. For short time series it would seem sensible in future to fit biologically sensible coefficients based on an appreciation of population characteristics.

The ICES Industrial Fisheries Working Group (Anon., 1984b) has used this method with some success.

2.6.3 General conclusions

The short-cut methods (SHOT, DROP and DOPE) appear to be useful in forecasting catch levels where other methods are unavailable. Their use would seem reasonable, therefore, where data are sparse. Clearly, much more exhaustive tests on simulated and real data are indicated before the most appropriate method for a stock can be defined and before the most appropriate method of coefficient estimation can be identified. It would, therefore, be premature to recommend that such methods should be used very widely, but the preliminary results are promising, and the Working Group looks forward to examining the results of further tests in the future.

3. THE APPLICATION OF LINEAR REGRESSION

3.1 Introduction

3.1.1 Background

Linear regression is widely used in fisheries research (and elsewhere) both to examine and test for the existence of relationships, for the appropriateness of various models, and as the basis for making predictions.

If high quality data on a well-defined relationship are available, there is usually little difficulty in deciding upon suitable parameters for the relationship. However, when the data are (as is usual) subject to appreciable variability, the determination of the most appropriate representation of the relationship becomes a little more difficult. This happens because various factors need to be taken into account, including:

1) The purpose for which the relationship is required (prediction, estimation of parameters *per se*, test of dependence, etc.);

2) The source and nature of the variability (measurement error, structural variability, form of probability distribution etc.);

3) The origin of past and future observations (whether drawn randomly from a probability distribution, or controlled in some way).

There is a variety of methods and formulae appropriate to various circumstances. The statistical methodology is described and discussed in a comprehensible way in the book by Sprent (1969), and the various standard forms are summarised by Pope and Shanks (1982) and also by Seim and Saether (1983) amongst many others. Ricker (1973) recommended the use of geometric mean functional regression, but see the comments of Pope and Shanks (1982), and further comments below.

The results obtained from the different methods and formulae generally differ appreciably only when the evidence for a linear relationship is weak (small correlation coefficient), or the method used is seriously inappropriate, or a prediction is made well outside the range of the data. Given care and common sense, the results from the various methods usually agree as well as can be expected, particularly when one takes account of their confidence limits, which will almost always be wide when discrepancies occur.

Thus, estimates obtained from the optimal strictly method for a particular situation may not differ significantly from those obtained from other methods, which, whilst not optimal or statistically justifiable, may nevertheless be reasonable and fairly robust. This includes both non-parametric methods (e.g., use of median slope for regression through the origin), which the Working Group did not have time to discuss in detail, or the results which might be obtained by a reasonable man with a ruler.

However, the Working Group recognised that even quite minor discrepancies can cause doubt, confusion and argument, and that there is a need for established and objective methods whose results can be checked and reproduced exactly if necessary. This report, therefore, contains a brief discussion of the main points which need to be considered before attempting to analyse a given set of data, and a guide to the appropriateness of various methods. The number of possible combinations of circumstances is enormous, and it was not possible to construct a complete guide. We hope, however, that people following our recommendations may be steered away from serious error.

Where the purpose of the analysis and the nature of the data are quite clear, one should naturally use the appropriate method. The Working Group recognised, however, that there may still be cases where there is doubt, and therefore proposed the use of a particular method in such cases which is to some extent central and robust, and unlikely to lead to serious error.

This proposal is to some extent conditioned by the philosophy that all methods of analysis and prediction are thinkable, that what is required is an understanding of their properties, and that it is not always necessary to seek and use an optimal method, if a satisfactory one will suffice.

The remainder of Section 3 of this report, therefore, contains a brief survey of considerations affecting the use of linear regression methods, and what to do about them; this is followed by an account of recent work on methods for deriving confidence limits for the calibration problem (which is quite a common application in fisheries research), and a description of further work to be undertaken to obtain a more complete solution. The results of analysis, by various methods of example data sets, are described, and finally some practical recommendations are made.

3.1.2 The purpose of linear regression

3.1.2.1 Prediction

A regression may be required simply in order to make a prediction of one (dependent) variable from a given value of another (independent) one; this may be based on some assumption or belief in a causal relationship (hence the nomenclature) or it may simply be empirical.

Various predictors are possible, and their properties may be described (*inter alia*) by such quantities as bias, variance, mean square prediction error etc. The parameters of a line fitted and designed to be used for prediction usually differ from those of a line fitted in order to describe an underlying relationship (see below), because amongst other things it may make use of prior information about the probability distribution of the observations.

The classical result in the regression problem, in which the slope of the line is estimated as

$$\hat{g} = S_{xy}/S_{xx}$$

is the solution of several problems, including that of making an unbiased estimate of the expected value of y for a given value of x, taking account of such information as is contained in the data about the probability distribution from which the observations are drawn, under certain assumptions (normality, homoscedasticity, etc.). It is also (coincidentally) the solution to the problem of making an unbiassed estimate of the expected value of y, if the independent variable (x) is measured without error, whether or not the observations are from a random sample (under similar conditions). It is, therefore, a very useful predictor, and widely used as such. It is, however, not necessarily optimal, since this depends on the precise objective of the prediction, and thus on the loss function, which is taken to apply.

A loss function is a construct which defines precisely the desirability of various properties of a predictor, including bias, variance, and other factors which derive from economic, social or any other considerations. The minimum of a loss function therefore defines precisely what is meant by optimal, and this depends entirely on the application.

Such functions are rather difficult to construct in practice, and finding optimal predictors for them may be even more difficult. However, it is known that for a particular loss function, mean square prediction error (the sum of variance and the square of bias) which is a useful measure of the final accuracy of a prediction, the optimal predictor differs from the classical one. This matter is discussed by Harding (Working Paper No.7), and at some length by Copas (1983).

For the present purpose it suffices to note that prediction is a particular use of linear modelling, and that the methods required for prediction may differ quite considerably from those for other purposes, and also according to the exigencies of the task in hand.

3.1.2.2 Calibration

A special case of the prediction problem arises when one has past data on the relationship between a precise and imprecise measurement of (hopefully) the same thing, and wishes to use a future imprecise observation to predict the true (precise) value. This is known as the calibration problem, and has generated controversy among statisticians. A detailed account of the problem, and some new results for its resolution, both by Harding, were available to the Working Group (Working Paper No.7).

Many problems in fisheries research fall almost into this category - the estimation of year class strength from a survey index, based on the 'calibration' against VPA estimates of past year class strengths, is an example. This is not quite the classical case, because the VPA results cannot be considered to be measured with negligible error (variability) since they are essentially just summations of catch-at-age data, which are certainly subject to appreciable sampling error.

In order to evaluate the methods available for the treatment of such data, the Working Group assumed that VPA results could be treated as precise, so that a few example calculations could be performed.

3.1.2.3 Functional regression

Regression may also be used to determine the parameters of a model relationship, either because these are of intrinsic interest, or as a basis for prediction in some circumstances. Maximum likelihood techniques are often used for this purpose (although other methods are also applicable).

The theory is well developed for the common case, where there exists a strictly deterministic underlying relationship, but where both variables are subject to independent errors (Lindley (1947) see also Pope and Shanks, 1982 and Davies and Goldsmith, 1976). Provided that some information can be provided about the 'error' variance of one variable, or their ratio, explicit maximum likelihood solutions can be obtained.

Assume the (exact) functional relationship

$$\eta = \alpha + \beta\xi$$

The variables η and ξ are measured with some error:

$$x_i = \xi_i + \delta_i$$

$$y_i = \eta_i + \epsilon_i$$

It is assumed that δ and ϵ are normally distributed, independent with zero mean and variances:

$$\mathrm{var}\,(\delta) = \sigma_x^{\,2}, \quad \mathrm{var}\,(\epsilon) = \sigma_y^{\,2}$$

Now we assume that the ratio:

$$\lambda = \sigma_y^{\,2}/\sigma_x^{\,2}$$

is known. An estimate g of the slope β is constructed as follows:

Let

$$p = (S_{yy} - \lambda S_{xx})/2S_{xy}$$

Then

$$g = p + \sqrt{p^2 + \lambda} \qquad \ldots\ldots\ldots\ldots\ 3.1$$

where

$$S_{xy} = \sum_i (x_i - \bar{x})(y_i - \bar{y})$$

$$S_{xx} = \sum_i (x_i - \bar{x})^2$$

$$S_{yy} = \sum_i (y_i - \bar{y})^2$$

An estimate h of the intercept α is then:

$$h = \bar{y} - g\bar{x}$$

If λ is infinite ('error' variance associated with x is negligible), this reduces to the classical result for regression of y on x

$$\hat{g} = S_{xy}/S_{xx} \qquad \ldots\ldots\ldots\ldots\ 3.2$$

whilst if λ is zero ('error' variance associated with y is negligible), this reduces to the classical result for regression of x on y

$$\hat{g} = S_{yy}/S_{xy} \qquad \ldots\ldots\ldots\ldots\ 3.3$$

Note that a little care is needed in taking the limit $\lambda \longrightarrow \infty$, and that here and elsewhere we always have g as a measure of the slope dy/dx, not vice versa, so that g has the dimensions of $[y]/[x]$.

All other results for intermediate (and of course inherently positive) values of λ give results for g intermediate between the limits set by the two classical values given above.

The family of results expressed by Equation 3.1 above, regarding λ as a parameter, is very useful because most of the classical results are members of the family, as follows:

Value of λ	Standard result
∞	Classical y on x
1	Major axis method
S_{yy}/S_{xx}	Reduced major axis method (geometric mean functional regression)
0	Classical x on y

(See also Seim and Saether, 1983, and Pope and Shanks, 1983).

With an appropriate choice of λ (which must be specified and cannot be determined from the data (see Copas, 1972)), equation 3.1, therefore, supplies the solution to a variety of problems of both the predictive and functional type. (In general, of course, λ should be specified by examining the estimated error variances, and not simply as a conventional value).

This is a remarkable result, which may perhaps be exploited in cases where there is doubt about the exact nature of the problem, since the results of all the standard methods lie between the limits set by Equations 3.2 and 3.3, and except when the correlation coefficient is small (in which case any predictor will have wide confidence limits), the result is not very sensitive to the choice of λ.

This is illustrated in Figure 3.1, which shows (in non-dimensional units) the estimates produced by Equation 3.1 for a range of values of λ and several values of the correlation coefficient. It is clear that the major part of the variation of g occurs over a range of values of λ within about a factor of ten either way of the value $\lambda = (S_{yy}/S_{xx})$. Thus, even an approximate estimate of λ may be sufficient to lead to an acceptably precise estimate of g.

Equation 3.1 of course provides only a point estimate of the gradient. Some information on the confidence region of an estimate is also needed. The theory for this case is not yet fully worked out, so far as we know, but work is in progress, and some preliminary results are available. It is believed that the method of mapping of the likelihood function used by Harding can be generalised to the case $\lambda \neq 0$, and this would provide a thorough and practically useful presentation of the information (it should be possible to construct plots similar to that of Figure 3.1 for any data set). Work along these lines is in progress (E F Harding, pers.comm.) (see also Section 3.3).

Meanwhile, the results of Lindley and El-Sayad (1968) suggest that the distribution of $\hat{g}$ is approximately log-normal, with a logarithmic standard deviation (approximately a coefficient of variation) given by:

$$\left\{(1 - r^2)/n\, r^2\right\}^{\frac{1}{2}} \quad \text{.....................} \quad 3.4$$

Confidence limits have also been estimated by Creasy (1956) (see also Davies and Goldsmith, 1976) but involve more complicated calculations. Since in the near future likelihood mapping should provide a better solution, the Working Group suggests that substitution of Equation 3.4 in standard formula for the standard error of a prediction should provide an adequate interim estimate for estimation of the confidence region.

It should be noted that the results given above all generalise immediately to allow for

(a) heteroscedasticity (but not non-normality), by calculating weighted sums of products with weights given by the inverse variances for each data point;

(b) forcing the regression through the origin by computing the sums of products of deviations from the origin, rather than the mean.

It is also of interest that the result from Equation 3.1 conforms with two of the practical criteria set out by Ricker (1975), and recently re-stated by Schnute (1984), namely, that it is invariant under the exchange of x and y (it is of course necessary to use the reciprocal of the dimensional quantity λ, and the result is the reciprocal of the original estimate as required), and also invariant under a change of scale or units (provided of course that the consequential change to λ is also made).

The properties of the predictor obtained by using Equation 3.1 as an estimate of the slope are not well known, but almost certainly intermediate between those of the classical predictor and calibration methods. The likelihood surface will provide adequate information in due course, and estimates of bias and variance would not be particularly useful unless one knew the loss function for the problem (especially since the variance may not be well defined, and the distribution not normal). Thus, such a predictor cannot be regarded as <u>optimal</u> for any known problem, but it may be <u>adequate</u> for many.

3.2 Likelihood Methods and Confidence Limits

A working paper in four sections by E F Harding "Considerations concerning linear regression" was made available to the Working Group. The final version of this will be published in due course. Meanwhile a brief summary of the contents is given here.

The first two sections dealt with the "Calibration problem": two variables, X and Y, are connected by a linear relationship. Y is observed with great accuracy, but X with substantial measurement errors. Given a new observation of X we want to predict Y. This is different from ordinary linear regression, where the errors are associated with the dependent variable. A solution to this problem is presented in Harding's second section with numerical examples of its application. This solution is appropriate in cases where the observations of the quantity which we want to predict are much more accurate than the observations used for the prediction. An example of the use of this method of likelihood mapping is given below (Section 3.4).

The third paper is concerned with "Prediction and shrinkage", and contains, among other things, comments on a paper on the subject by Copas (1983). Generally, the most straightforward predictor of the dependent variable where a linear regression relationship exists is not the "best". What the "best" predictor is depends on the purpose for which the predictor is needed. A formal solution of this can be obtained by specifying a loss function. Commonly, little effort is spent on this, and a sensible criterion in many cases is the prediction mean square error, $E(\hat{Y} - Y)^2$, where Y is the variable which we want to predict, and $\hat{Y}$ is the predictor. This predicor is usually biassed, i.e., $E(\hat{Y} - Y) \neq 0$. The bias, however, will not be large compared with the magnitude of the prediction errors. Unbiassed predictors usually have larger variances than the minimum MSE predictor. In the absence of detailed information about the loss associated with prediction errors, the mean square error criterion strikes a sensible balance between bias and variance.

The fourth paper presents a general discussion of the problem of predicting Y given an observation of X, when both X and Y are random variables. For this purpose a likelihood function approach is adopted, and information about accuracy could be presented by graphs of the partial likelihood function. Various complications which are likely to arise in practice and can be treated objectively are discussed. These include non-linearity, non-normality, and the case where variances are not constant. Some of these problems can be analysed by use of statistical packages (such as GLIM), and some suggestions about this are provided. The paper concludes with a section on post-fit diagnostics where advice is given on the examination of residuals and treatment of outliers.

The Working Group considers the analysis of E F Harding of this important class of problems of great value and hopes that this fruitful contact between fisheries research scientists and academic research in statistical methodology will be maintained.

3.3 Future Work

G Gudmundsson described first attempts at the Department of Applied Mathematics at the University of Iceland to estimate a model of two random variables such as indices from 0-group surveys and recruitment assessment on the following lines:

Use the transformations suggested by Box and Cox (1964) to transform both variables to approximately normal distributions. Assume that the transformed variables constitute a bivariable normal distribution. Attempts to estimate the necessary parameters by maximum likelihood methods were unsuccessful; the model is too general for the small number of observations available.

Using a more restricted model (such as that for the functional regression problem), it should be possible to extend the method used by Harding to the more general problems, and thus compute confidence regions for the solutions to a wide range of problems.

The Department of Applied Mathematics is prepared to continue this work in collaboration with the Marine Research Institute, taking into account the suggestions of Harding's working paper. The plan is to work through a couple of actual examples and make programs developed for this purpose available to the members of the Working Group.

3.4 Results of Worked Numerical Examples

The above formulae (3.1 to 3.3) were used to evaluate the relationship between 0-group indices and VPA recruitment values. Data were taken from Working Paper No.8, Figures 7 and 8. The data are given in Table 3.4.1, and the correlation coefficient is 0.87.

In the first data set, the ξ-variable corresponds to USA spring survey catch per tow at age 2 of mackerel, and the η-variable gives the stock size at the same age determined from cohort analysis. Visual inspection shows that a power curve fits the data better than a linear curve. So actually, the model used is of the type:

$$\eta = \alpha \; \xi^{\beta}$$

Now λ corresponds to the ratio of the variances of the log-transformed variables. Therefore scaling of the variables is irrelevant. The following estimates of the power β and the factor α were obtained:

λ	Power	Factor
0	0.42	740
0.1	0.35	840
1.0	0.30	930
10.0	0.29	945
∞	0.28	950

The ratio $\lambda = 0$, as said before, corresponds to assuming no error in the VPA values, whereas for $\lambda = \infty$ we assume the survey indices to be exact. Presumably, the variation in the survey index is larger than that in the VPA values. Hence $\lambda = 0.1$ seems closer to reality than $\lambda = 10$.

The second experiment relates USA autumn survey catch per tow at age 0 for Gulf of Maine silver hake (ξ-variable) with year class size at age 1 from VPA (η-variable). Following the study quoted, a linear relationship was assumed, and scaling matters. So, in the following table λ is replaced by

$$\lambda^{*} = \lambda \cdot \left(\frac{\bar{x}}{\bar{y}} \right)^2$$

In this example, the calculated values are:

λ^{*}	Slope	Intercept
0	33	55
0.1	30	64
1.0	25	85
10.0	23	92
∞	22	94

Again, $\lambda^{*} = 0.1$ might be closest to reality.

Due to computational limitations, no standard deviations or test-indicators were evaluated.

The original model

$$\eta = \alpha + \beta \xi$$

was assumed to have no (structural) error, which may be unrealistic for some applications. If there is such an error, the problem is changed, and other methods of solution might be required, although Pope and Shanks (1982) suggest that the solution to the problem of bivariate structural relationship is the same as that given above.

The results of the calculations on the Northwest Atlantic mackerel stock given above are illustrated in Figure 3.4.1, together with the contours of equal likelihood for the calibration problem computed using a program supplied by E F Harding. It is clear that the likelihood contours are centred on the result for $\lambda = 0$, as they should, and that within the range of the observed data, all three lines shown fall within the range for a likelihood ratio greater than 0.8. This corresponds approximately to the 50% confidence region, so it is clear that predictions from any of these lines would not differ significantly, even at the 50% level.

3.5 General Advice and Recommendations

Since there exist a number of approaches for fitting a straight line to observations, the Working Group suggested the following practical guidelines, which can be applied to any model:

1. Plot the observations

   A scattergram of the observations provides an indication of the degree of linearity of the data and of the distribution of variability. Carefully examine the extreme observations (apply the 'thumb' technique) as these will be 'influential' in the estimation of the model; determine which sub-set of the observations are to be utilised. If this inspection leads to the rejection of some observations, apply the same 'rejection criteria' to all observations remaining in the sub-set. If data look useless, do not proceed further (e.g., the presence of a single extreme point which is known (expected) to be erroneous would be influential and may lead to spurious results).

2. Transformations

   Homoscedasticity is a desirable property: the methods discussed here/above assume homoscedasticity. If the scattergram shows evidence of heteroscedasticity, then consider transforming them so as to make the distribution of observations homoscedastic. The log-log transformation often provides satisfactory results and stablises the variance of measurement errors. However, a transformation may destroy other desirable features. An alternative is to calculate weighted sums of products, using weights proportional to the reciprocal of the estimated variance of the data points.

3. Choice of the model

   A number of candidates may exist, at least theoretically, e.g., straight lines (possibly forced through the origin), curvilinear models (power laws, polynomials, etc.), multiple regression, etc. If a linear model seems desirable, identify which type of regression line is required (predictive regression, calibration, or functional regression). If a functional regression is chosen, the ratio of error variances λ should be estimated from whatever knowledge is available on the nature and source of the data.

   It should be noted that the geometric mean functional regression is a maximum likelihood solution only when λ is known to be S_{yy}/S_{xx} - which is not likely to be very often. However, λ is in general a dimensional quantity, so that it is wrong to assume $\lambda = 1$ unless both x and y are similar quantities, having similar errors, measured in similar units. If no information is available, the choice $\lambda = S_{yy}/S_{xx}$ may be the best last resort of the desperate, since it at least uses a natural scaling and leads to a central estimate. Such a last resort should, however, be avoided if at all possible, and no further justification can be given for it than expedience. It is of course a 'central' estimate, and therefore unlikely to be far wrong.

4. Acceptability of the model

In practice, acceptability may be defined in terms of "plausibility" of the model and in terms of statistical significance (do the data support the model?). For example, a negative slope may not correspond to a "plausible" model and thus could be rejected even if statistically significant. "Statistical" acceptability is cast in terms of deviations from underlying assumptions (e.g., regarding the error structure) and in terms of precision of the estimates. For instance, we should always:

- examine the significance of the regression coefficients; in defining the model, prefer parsimony (for example, if intercept is not significant, use a regression through origin); use non-significant coefficients with caution;
- inspect the distribution of the residuals (histogram); test for normality of the residuals, including skewness and kurtosis, if there is a sufficient number of observations (Cox and Hinkley, 1974); identify outliers;
- test for lack-of-fit (Draper and Smith, 1966) or systematic trends in the residuals; such test may suggest that another form of the model would be more appropriate;
- calculate the confidence intervals of the predictions.

5. Extrapolation

Results and predictions from regressions are supported by data (evidence) only within the range covered by the data observed. Extrapolation outside this range *may* be valid *if* the model chosen is correct. Such an extrapolation, however, is based largely on *faith* in the model, and statistical methods can provide no justification for it.

6. Retransformation

Note that when a transformation has been applied, a correction will normally be required if unbiassed estimates of the original quantities are required (this is well-known for the logarithmic transformation, see e.g., Granger and Newbold (1977)). This will not usually be necessary unless an unbiassed predictor is required. No correction is required for quantiles (such as the median).

4. FORECASTING OF RECRUITMENT

4.1 Background

Recruitment forecasting represents a key element in stock assessment and prediction. In analytical stock assessments, recruitment estimates are combined with retrospective (e.g., VPA) analyses to derive starting conditions for predicting the impacts of various management options for the forthcoming fishing season. If a significant proportion of a species is likely to be composed of new recruits (e.g., North Sea cod), accurate predictions of recruitment strength are critical to the development of proper management measures.

The consequences resulting from overestimating recruitment would be excessive fishing mortality rates. The consequences of underestimating recruitment would be less serious, since in practice the management option could essentially self-correct in year N + 1 by increasing the fishing rate, once the true character of the year class recruiting in year N has been verified. The consequences of "bad" recruitment forecasts (either low or high) could in theory be modelled by a "loss function" (penalty function) incorporating the potential trade-offs due to yield per recruit considerations, the dependence of fishing mortality rate on the difference between projected and actual recruitment, and socio-economics (e.g., are recruits proportionally more valuable if harvested in year N + 1 than in year N?). However, this is difficult to do and is rarely feasible in practice.

In general, recruitment forecasts for assessment purposes are made in one of two circumstances:

(a) when reliable field sampling data are not available (i.e., for correlation type estimates),

(b) when field sampling (particularly young fish surveys) are available.

In the former circumstance, various methods that have been or could be used to generate working estimate include (amongst others):

(1) use of the previous year's recruitment value from VPA (analogy to 'the best forecast of tomorrow's weather is today's'),

(2) geometric mean or some quantile of the recruitment series,

(3) mean of the last N years recruitment data,

(4) use of spawner/recruit relationships,

(5) Delphi (group consensus) methods by those responsible for recruitment forecasts.

When field sampling data are available, recruitment forecasts have generally been based on results of young fish surveys (0-group, 1-group) correlated with VPA, appropriately time-lagged. Results of egg- and larval surveys have generally been of little use for recruitment prediction probably due to the substantial variability in survival rate between these surveys and the recruitment age, and the high measurement errors sometimes associated with these surveys.

The Working Group considered various methods for recruitment prediction, both from correlation type analyses and based on various assumptions based on previous recruitment patterns. Statistical models and potential errors associated with regression analyses of survey/VPA were also evaluated (see also Section 3.4).

4.2 The Use of Quantiles

When little information on an incoming year class is available, prediction may be based on the properties of the series of estimates in previous years. Predictions based on quantiles can be justified on the basis of the expected frequency with which the prediction will be too high or too low (e.g., using the median, or 50th percentile, as a prediction should overestimate recruitment as often as it underestimates it).

Some additional insight into the use of the median, and into the dangers of using the expected value, follows from the typically log-normal distribution of recruitment series (Hennemuth et al., 1980; Garrod, 1983). For the log-normal distribution, the median is equal to the geometric mean, which is given approximately by

$$GM(x) = \bar{x} - \frac{\sigma_x^2}{2\bar{x}} \qquad \text{.........} \qquad (4.2.1)$$

(This relationship also gives an approximation to the geometric-arithmetic mean relationship in any distribution.)

Thus, the median of a log-normal distribution is lower than the mean by an amount which is proportional to the variance and inversely proportional to the mean. Using the arithmetic mean as a prediction of recruitment lends excess weight to rare large values; it will overestimate recruitment most of the time. This danger is greatest for stocks with a high variance in recruitment.

In some cases a more conservative prediction (i.e., a lower quantile) may be appropriate: if it appears that recruitment is likely to be lower than usual (see Section 4.8.1), or if one is especially concerned to avoid overestimation of recruitment. From Equation (4.2.1) it is apparent that a lower quantile, corresponding to a figure even further below the arithmetic mean than is the median, can be thought of as arising from a formula like (4.2.1) but with extra weight assigned to the variance term. This is typical of decisions made under uncertainty (Keeney and Raiffa, 1976), in which the loss functions involve the expected outcome discounted by the variance in that outcome. A Working Group attempting to decide on a quantile for the prediction of recruitment might wish to consider the mean-variance trade-off explicitly.

4.3 The Use of Stock/Recruitment Relationships

Stock/recruitment relationships can in principle be used for forecasting recruitment, but the variability of the data is usually too great to produce useful estimates of recruitment (Cushing, 1973). However, in a stock for which some form of a stock/recruitment relationship has been demonstrated (or is suspected), it could be used not only to make a prediction, but also to obtain a suggestion of which quantile (see Section 4.2) should be used. This is effectively a semi-quantitative method (see Section 4.8), which might be particularly useful, if the stock is low and there are reasons to suspect that a recruitment failure is possible. Such a procedure would motivate the use of lower quantiles rather than the median for the recruitment forecast.

One way in which even rather variable stock-recruitment data can be used to obtain working estimates has been suggested by Getz and Swartzman (1981) and Swartzman et al. (1983). The idea is basically to divide the observed stock sizes and the corresponding recruitment into intervals of stock sizes. The number of intervals is dependent on the nature of the stock-recruitment data and their grouping in the stock recruit plane. Each interval corresponds to different levels of recruitment with a given probability.

The probabilities can be regarded as elements, which represent a Markov-type stock-recruitment transition matrix (Getz and Swartzman, 1981).

4.4 The Use of Survey Indices

Pre-recruit survey indices generally provide the most useful method for forecasting recruitment - indeed, this is usually the only technique providing useful forecasting ability. (Commercial cpue of young fish is regarded as survey data for present purposes.)

Pre-recruit indices are usually rather imprecise, with large, often highly skewed sampling variability, of approximately log-normal form, and a coefficient of variation rarely less than 30%.

Estimates of past recruitment from VPA are usually relatively precise, but since they are essentially just the sum of several years' catch-at-age data for a cohort, they are clearly not entirely free of sampling error. VPA estimates are probably biassed, because of incorrect assumptions about natural mortality on the younger ages, and perhaps also if catches are under-reported and discarding is substantial, but this should not be particularly serious, since such biasses will be reversed when the recruitment estimates are fed into a catch forecast, provided this is done consistently.

It is often noted that extreme high and low values of survey indices are not followed by equally extreme values of recruitment when these are subsequently determined. This would be a natural consequence of high variability in the survey index, but there are also biological and practical reasons to suspect that such an effect could also be systematic. Among these are the possibility of density-dependent mortality subsequent to the determination of the index, changes of vulnerability and/or distribution of the young fish at different levels of abundance. It is, therefore, quite plausible that a non-linear model may be required to adequately describe the relationship (a power law is an obvious candidate), although the most plausible null hypothesis probably remains strict proportionality (i.e., a linear relationship through the origin, corresponding to a power law with an exponent of one).

Since recruitment also tends to be log-normally distributed, one has the happy situation where a logarithmic transformation of both variables is likely to lead to several desirable properties at once, namely approximate normality and homoscedasticity of the variability (including errors), normality of the distribution of the observations, and a plausible linear relationship. In most cases this should normally be a satisfactory treatment of the data. If the slope (power) determined is not significantly different to one, however, the use of the simpler model with the power forced to be one is probably preferable. This implies a straight line through the origin for the untransformed data, having a slope given by the geometric mean ratio y/x in the data.

If VPA results can be considered as good estimates of recruitment in the sense that the error they introduce is limited compared to that associated with research survey indices, estimating a forthcoming recruitment from a survey index can be identified as a calibration problem as described by Harding (Work.pap.No.7).This would imply the use of the regression line of survey index on VPA recruitment using maximum likelihood estimators. The main advantage of such an approach is related to the construction of confidence intervals, which can even be considered to apply if other estimators than the maximum likelihood estimators are finally chosen to get the point estimation.

Maximum likelihood estimator (M.L.E) is in fact far from being the only conceivable estimator. Whatever estimator is used, its statistical properties and the practical implications in terms of management decisions must be known. From this second point of view it is especially relevant to note that the different regression lines lead to more or less optimistic estimates of recruitment. If risk-adverse management strategies are to be preferred (which is not a general recommendation), it would be appropriate to choose between the considered estimators the more pessimistic one. In general, this would imply using different estimators when recruitment is above or below the average.

From a statistical point of view it must be recalled that M.L.E. are not generally unbiassed. It must also be remembered that they do not take into account the *a priori* likelihood of the different possible levels of recruitment, as described by the observed (and usually skewed) distribution of past recruitments. Paying attention to this distribution would give an argument to use the regression line of VPA recruitment on survey indices. This line would give recruitment indices systematically closer to the observed mean recruitments, being more pessimistic for large recruitments and more optimistic for poor ones than the previous regression line, derived from the calibration problem. It could, therefore, exacerbate the problems of managing a stock where recruitment is in decline, and for this reason would probably not be a generally acceptable procedure. When the likely errors in the VPA recruitment are not negligible, the problem is not the pure calibration case, and the theory is not well-established. Bearing in mind the results reported in Section 3.4, however, it seems likely that the differences between alternative estimators will be small compared with their confidence intervals. The functional regression (with λ specified) seems the most plausible, and provides a sensible central estimate.

Pre-recruit index data can be badly affected by changes in survey strategy (or fishing pattern, if commercial data are used). Data free of such problems should be used if possible. It must be stressed that even the most careful statistical treatment cannot compensate for bad data. The confidence intervals of any prediction made should be carefully assessed, and if very wide would indicate using a robust estimate, abandoning the attempt at prediction, or using the result simply as an indication for a semi-quantitative method (see Section 4.8).

Some technical points concerning the construction and precision of survey indices in general (see Section 4.7) are of course relevant to this problem.

It could be tempting to try to improve the predictive power of this method by including explanatory variables (such as temperature, salinity, oxygen content, etc.) along with the survey index in a multiple regression analysis. This should be done with great caution and only when the use of environmental variables is justified by a causative model. When predictions are made, the relative importance of the survey vs the environmental variables should be assessed. If the environmental factors are dominant the prediction should be used with great care, since such relationships often break down at unpredictable intervals. This is discussed in Section 4.5 below.

4.5 The Use of Other Explanatory Variables

Striking coincidences are often noted between outstandingly good or poor year classes and environmental conditions. In the absence of suitable survey indices, it may, therefore, be tempting to attempt to use environmental data as a basis for prediction.

This usually involves the use of a multiple linear regression model. Considerable care is needed in such analyses because many such attempts have explained past data quite well, but failed to give useful predictions. There are many possible reasons for this. The relationship may not be linear (there may be an optimum temperature, for example), or the biological effects of environmental factors may be different under different conditions (e.g., different stock sizes). Since there are usually many possible explanatory variables to choose from, the model may have insufficient degrees of freedom if too many are included (in this case the considerations discussed by Copas become relevant).

In general, therefore, the application of multiple regression techniques to what are usually short and noisy time series must be regarded with considerable scepticism, particularly since the explanatory value of such models is usually quite low, and their predictive value even less. In favourable circumstances, however, they may be of use, and might be a useful input to the semi-quantitative methods discussed in Section 4.8.

4.6 The Use of Time-Series Analysis Methods for Recruitment Forecasting

Comments on the application of time-series analysis models, given in Section 2.5, are generally applicable to the case of recruitment forecasting. In particular, the utility of multivariate (transfer function) models to predict recruitment levels from young fish surveys appears low due to the short time series generally available and the high level of variability. Alternatively, the use of univariate time-series models, combined with certain intervention terms (known as shocks to the time series) appears promising, particularly where the VPA time series is greater than about 25 years. Recent preliminary application of the technique to the Georges Bank (USA) haddock stock (time series 40+ years long) has shown promise, particularly when intervention terms have been used to account for changing fishing patterns over the time period (initial equilibrium, followed by an extremely large year class, and subsequent recruitment overfishing).

Further work on application of these techniques to the recruitment estimation problem is warranted, particularly where recruitment time series are relatively long.

4.7 The Construction of Survey Indices

The general level of within stratum variance of logarithmic transformed catch data found within groundfish survey results appears to favour a logarithmic raising approach as advocated in Pennington (1983). It was, however, reported that recent studies at Lowestoft (Hunton, 1984) suggest that where hauls have been repeated at the same positions over time (and only in this case) much of the within-strata variation (about ½) can be ascribed to systematic station effects. This serves to reduce the log variance and thus makes the argument for raising, using re-transformed log means, somewhat less compelling. In some tests there seemed little to choose between the logarithmic and linear raising of catches.

If station effects are systematic, the multiplicative models provide an alternative method for estimation of the time (year) effects. However, these are sensitive to changes of distribution whilst linear raising is not.

These results suggest that when station positions are fixed, the statistical properties of linear raising estimates (efficiency, etc.) may be less poor than previously thought, and the method has the considerable virtue of simplicity.

4.8 The Use of Semi-Quantitative Methods

4.8.1 Combination methods

Quantitative data may often be available from field sampling programmes, but the data may be considered unreliable for the purposes of making recruitment projections directly. Such data sets may include egg-, larval- or O-group surveys, where sampling variability is quite high, or survival rate between the field observations and recruitment to the fishery is highly variable (i.e., the strength of the incoming year class may not yet be determined). In such cases, it may be advisable to combine these 'anedoctal' observations of the potential recruitment strength with the quantile approach. For example, an O-group survey index may not be useful for recruitment projections, but the index itself may be designated as 'low', 'medium' or 'high', based on some objective or subjective criteria. Recruitment projections could be based on an association between the category designated from the sampling data, and the quantile distribution of the recruitment time series:

O-Group Index		Recruitment Estimate
'Low'	-	Lower quartile
'Medium'	-	Median
'High'	-	Upper quartile

This procedure allows the use of some observational data that may be evidence for relative recruitment strength, but which is unreliable for forecasting directly.

4.8.2 Decision analysis technique

Various operation research techniques have been developed to support the decision-making process given uncertain, imperfect, or non-quantifiable data (Moder and Elmaghraby, 1978). In particular, the Delphi technique (Zuboy, 1981) has been employed to arrive at group consensus regarding the level of some quantitative (but unknown) variable or to prioritise alternative actions. The Delphi technique was developed during World War II (as was much of the field of operations research) as a tool for such decision analysis problems. The process of achieving group consensus based on Delphi follows a formal structure involving initial anonymous estimates by the various 'experts' assembled. Those giving extreme values about the median are asked to support their positions (usually by written comment), followed by another round of balloting. The process may continue until some stopping criterion has been reached. Such a formal procedure may not be practical for the Working Group environment of ICES (e.g., to estimate recruitment level, given unreliable or non-existant sampling data). Nevertheless, a form of the Delphi procedure does in fact operate in order to arrive at a reasonable and appropriate recruitment level. The use of more systematic methods for this application may be helpful with regard to documenting the process by which a group consensus was reached.

5. OTHER TOPICS

5.1 Separable VPA

5.1.1 New version of separable VPA

A new version of the separable VPA method has been produced at the Fisheries Laboratory, Lowestoft, and is described by Stevens (1984). This permits the user to apply weighting factors (which may either be specified or determined automatically) to the residuals for each pair of age groups, and also to apply specified weights to the residuals for each pair of years.

This removes the need to eliminate data for poorly sampled age groups or years from the data set, allows for appropriate account to be taken of relatively precise and noisy data, and removes the need for the ad hoc extension procedures used previously.

The new method is, therefore, somewhat more versatile and easier to apply than the previous version, and gives very similar results on equivalent problems.

5.1.2 Statistical tests of common restrictions in fish stock assessment

A working paper was presented on the subject by Gudmundsson (1984). The main contention of the paper is that fish stock assessment should aim to comply with good statistical practice, which is is widely accepted in other fields of research. A succinct and authoritative description of such methodology is given in the introduction to 'Theoretical Statistics' by Cox and Hinkley (1974). Commonly applied restrictions and simplifications can to some extent be tested by comparison with estimates based on more general specifications. This includes the assumption of separable age and year effects, no variation with age above full recruitment, proportionality with a given index of fishing effort and no immigration. In least squares estimation it is necessary to avoid that most of the residual variance is confined to a particular age, year or year class. Serial correlation may indicate that a model is mis-specified.

Numerical examples of the application of some specification tests and residual diagnostics are presented with data on cod, haddock and whiting from the North Sea and cod from the Northeast Arctic and Iceland. The hypothesis that residuals are normally distributed tends to be rejected by Kurtosis tests, when models are fitted to logarithms of catch ratios. This is avoided by using untransformed observations, weighted according to age and year class. The hypothesis of separability was accepted for some stocks and rejected for others. The hypothesis of constant fishing mortality from the age of full recruitment is accepted, but the hypothesis of proportionality with a given effort index is rejected in all cases which were investigated.

It may happen that an assumption in a model (a hypothesis) can be rejected on grounds of statistical significance, but its inadequacy may nevertheless be of little practical significance. A penetrating and useful discussion of this question is given by Cox (1977).

5.2 Analysis of Catch-at-Age and Groundfish Survey CPUE Data

A method for the joint analysis of catch at age and cpue or groundfish survey indices was presented by Pope and Shepherd (1984). This is based on a plausible model which does not depend on VPA, and is a development of the method of Collie and Sissenwine (1982). Such methods are closely related to the 'Survivors' method of Doubleday (1981) and Rivard (1980), and utilise the same data.

Preliminary results are promising, but validity of the assumptions made by the method has not yet been subjected to statistical tests. Comparative analysis of the same data using the 'Survivors' method suggests that this may give more stable results, and a description of the most recent version of this method would be useful.

The method described by Pope and Shepherd (1984) does not assume separability of the fishing mortality pattern, although the work reported by Gudmundsson (above) and by the Working Group last year (Anon.,1984a)suggests that this may often be acceptable. Such an additional assumption would be likely to stabilise the results of such methods appreciably.

5.3 The Calculation of Biological Reference Points

5.3.1 General

The Working Group confirmed its previous conclusion that long-term goals for fishery management cannot be deduced from Y/R analyses alone, since such goals involve social and economic factors in addition to purely biological considerations. Furthermore, the Y/R analysis takes no account of the possibility of stock collapse under high fishing mortality.

Analysis of yield and biomass per recruit may, however, be used to determine 'biological reference points', which may assist in understanding the consequences of various management strategies, since they may serve as navigational markers or signposts. Among these are F_{max} and $F_{0.1}$, and also the quantities F_{low}, F_{med} and F_{high} introduced in 1983 (Anon., 1984a).

5.3.2 Purpose of F_{low}, F_{med} and F_{high}

These new biological reference points were introduced because fishing at various intensities not only causes variations of Y/R, but also of biomass per recruit. Indeed, increased fishing mortality invariably causes a monotonic decrease of biomass per recruit.

To persist at any level of fishing mortality, a stock must on average produce recruitment per unit biomass equal to the reciprocal of the biomass per recruit at that level of F.

Examination of stock and recruitment data, as described in Section 5.3 of Anon., 1984a(Pt II), enables levels of recruitment per unit biomass, which have been often, regularly, or rarely exceeded by a stock, to be determined. The corresponding levels of fishing mortality are designated F_{low}, F_{med} and F_{high}. The meaning of these quantities is, therefore, that:

(i) At levels of F below F_{low} there is plenty of evidence that the stock can produce sufficient recruitment per unit biomass to sustain itself.

(ii) At levels of F in the vicinity of F_{med} the evidence that the stock can produce sufficient recruitment per unit biomass to sustain itself is equivocal.

(iii) At levels of F above F_{high} there is little evidence that the stock can produce sufficient recruitment per unit biomass to sustain itself.

Levels of F above F_{high}, therefore, correspond to unknown territory in which the prudent would venture with great caution. With luck, of course, it may turn out that the stock is capable of producing higher recruitment per unit biomass than anything observed before. There is, however, no guarantee that this will be so.

Thus, the quantity of F_{high} is not an estimate of a fishing mortality at which a stock will collapse. It is, however, an estimate of a level of F above which the risk of collapse should be taken seriously.

Finally, it must be remembered that recruitment may well be affected systematically by environmental and ecological factors (including multispecies interactions). Thus, maintaining F below F_{high} or F_{med} provides no guarantee of persistence. It does, however, mean that a collapse could not reasonably be attributed to mis-management (overfishing) - whereas if F is allowed to rise above F_{high} there would be a *prima facie* case that overfishing was at least a contributing factor.

5.3.3 Use of F_{low}, F_{med} and F_{high}

The Working Group received a verbal account of a comment that "the North Sea herring stock had collapsed whilst remaining between the limits set by F_{low} and F_{high}". This comment is difficult to understand, because the analysis presented in Anon. 1984a(Pt II, Section 5.3 and Figure 5.3.4) indicates that F_{high} is about 0.86, and the fishing mortality on this stock exceeded that level for 8 out of the 9 years between 1967 and 1975 (ICES, Doc. C.M.1983/Assess:9).

The data points for stock and recruitment naturally and inevitably fall mostly between the lines drawn to correspond to F_{high} and F_{low} (Anon., 1984a,Pt II, Figure 5.3.4.B), simply because that is the way the lines are determined in the first place. This, therefore, constitutes no test of the method. The purpose is to determine levels of fishing mortality which relate approximately to different levels of risk, and the transfer from estimates of biomass and recruitment to estimates of fishing mortality (via biomass per recruit) is an essential part of the process. It may therefore be that the comment reported above is due to a misunderstanding of the way in which the method is intended to be applied.

5.3.4 The use of $F_{0.1}$

The Working Group was informed that its recommendations concerning the the use of $F_{0.1}$ had been found to a little ambiguous, and would benefit from clarification. The Working Group, therefore, emphasises that:

(a) $F_{0.1}$ is a measure of fishing mortality at which "high yields may probably be taken without unnecessary expenditure of effort" (Anon., 1984a,Pt II, Section 5.4). Its derivation has nothing to do with stock collapse, and there is no reason to assume that exploitation at or below $F_{0.1}$ will safeguard against a collapse of a stock.

(b) As stated in the Editor's Note on p.88 of Anon. 1984a in the statement that "if $F_{0.1}$ is to be adopted at all as a biological reference point, it should be used always and not only when F_{max} does not exist", the word "always" is intended to mean regularly, for stocks for which it is considered to be a suitable reference point - not for all stocks. The recommendation is intended to discourage switching from $F_{0.1}$ to F_{max}, simply because this causes confusion, and perhaps unnecessary variability in the scientific advice about the consequences of various management measures.

(c) The (relatively) favourable comments about $F_{0.1}$ made in Anon. 1984a(Pt II) refer only to its use as an alternative to F_{max} and should not be constructed as advocating the use of $F_{0.1}$ as the principal biological reference point, or as endorsing the use of biological reference points as targets for management.

At low values of F the yield per recruit curve is sensitive to the value of natural mortality assumed. The position of $F_{0.1}$ and F_{max} on the curve, therefore, might in some cases depend to a great extent on the choice of M by the Working Group. In those cases, and when there is no or little biological justification for the chosen value of M, biological reference points such as $F_{0.1}$ and F_{max} should be considered with great caution.

The Working Group therefore examined the sensitivity of F_{max} and $F_{0.1}$ to the choice of M, for two standard data sets used by North Sea Assessment Working Groups. The results are given in the text table below.

SOLE North Sea

x Input from 1983 report

M	$F_{0.1}$	F_{max}
.05	.205	.362
.10	.310	.573
.15	.395	.921
.20	.493	not found
.30	.722	not found

COD North Sea

* Input from pred.table 1983 report

M	$F_{0.1}$	F_{max}
.10	.107	.168
.15	.137	.210
.20	.164	.252
.30	.220	.344
.50	.354	.632
.70	.542	not found

These results are plotted in Figure 5.3.3. It is notable that both $F_{0.1}$ and F_{max} are more sensitive to M than M itself and that $F_{0.1}$ is (as would be expected) less sensitive to variations of M than is F_{max}.

5.3.5 The use of M as a biological reference point

There has been some discussion in connection with the management of the Western mackerel stock about the relative merits of using $F_{0.1}$ or M as a target for management. The Working Group draws attention to the statement in Section 5.3.4 (c) above. The discussion has, however, implicitly related to the possibility of collapse through stock and recruitment failure. It has been claimed that general production models tend to indicate that collapse is likely when F becomes more than a few times M (see e.g., Shepherd, 1982) or twice F_{msy} (Schaefer, 1954), but such conclusions are based largely on implicit treatment of the stock and recruitment relationship which may not bear a close examination. Several stocks have been exploited for many years at several times M without collapsing, and the critical level of fishing mortality clearly depends on several aspects of the life history and the nature of the fishery.

The only biological reference points designed to relate directly to the probability of recruitment failure known to the Working Group are the F_{low}, F_{med} and F_{high} measures introduced in 1983 (Anon., 1984) and further discussed above. F_{high} should naturally be treated as a level to be avoided, rather than as a target of management.

5.3.6 Recent extensions of the biomass per recruit approach

The biological reference points of F_{high}, F_{low} and F_{med} imply certain levels of spawning stock biomass per recruit relative to the virgin population. The objective of maintaining SSB/R as a proportion of that for the unexploited population can be addressed not only by manipulation of the fishing mortality rate, but also by the fishing pattern (age at first capture). Fig. 5.3.1 is a plot of the isopleths of the percentage of virgin SSB/R at various combinations of age at first capture and fishing mortality rates for Georges Bank haddock.Clearly, some fixed objective of the SSB/R level could be achieved by combined manipulation of the two variables.In Fig. 5.3.2 stock/recruitment data for the stock are plotted along with three lines relating to the percentage of virgin SSB/R under fishing patterns previously exhibited in the fishery (age at first capture of 2). The 10% level (relating to F = 0.9) appears excessive, since virtually all points in the lower left-hand corner of the graph lie below the line. The slope of this line (R/B) is an index of survival. If most points of very low spawning stock are below the line, then the chances of stock recovery will be lower than if a line with lower slope (higher survival and thus greater SSB/R) was followed. In this example, the 10% line appears roughly equivalent to F_{high}. F_{med} is achieved at a level of about 25% of virgin SSB/R.

By developing these types of analyses for a variety of species, we may be in the position of developing general guidelines for target SSB/R percentages to be maintained for stocks when no spawner-recruit data are available.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1 Simpler Methods for Catch Forecasting

1) Simpler methods for making catch forecasts can be constructed and, if chosen with care, are capable of estimating status quo catches with useful precision, and modest data requirements.

2) Where catch-at-age data are available, the ANOVA method may be useful. Otherwise even simpler methods (e.g., SHOT, DROP, DOPE) are necessary, and the following conclusions relate to these.

3) Such forecasts normally require time series of both catch data and a reliable index of recruitment (where new recruits are a significant fraction of the catch). Lack of recruitment data results in a serious loss of precision except for stocks subject to a low total mortality or relatively constant recruitment. The predictive value in the absence of recruitment data is likely to be low enough to make such methods unattractive except in desperate cases (e.g., as an alternative to average catches).

4) The methods presently available may all be regarded as particular choices of Autoregressive Moving Average (ARIMA) models, with the number and choice of terms, and the acceptable values of the coefficients, heavily constrained by the biological models.

5) Of the methods tested, the SHOT method usually performs significantly better than even simpler methods. Slightly more complex models such as the DROP and DOPE methods generally perform only a little better, except on stocks with highly variable recruitment where the improvement may be appreciable.

6) All the methods are based on simple dynamic models for exploited biomass, and the forecast for status quo catch may be viewed as forecasts of biomass, and thus also catch rate (in stocks where cpue is related to stock size).

7) The incorporation of cpue or other stock size index data into the framework of these simpler models needs to be explored. At present, this can only be done at the expense of considerable elaboration, and with only modest success.

8) The methods chosen focus on *status quo* catches simply because these can be estimated most easily and precisely. There is no implication that the maintenance of the *status quo* is desirable on biological grounds. The SQC estimate may be increased or decreased by any desired amount if the management goals indicate that such increases or decreases are desirable.

9) The simple methods provide little information on the state of the stock or the adequacy of the recruitment index. They may be useful for short-term forecasting, when more complete data are lacking, but are of little use in evaluating long-term consequences of management, for which more complete data are required. In no circumstances should they be used to justify abandonment of existing sampling programmes.

10) More research on the properties of these and similar methods is required before their use can be wholeheartedly endorsed.

11) The consequences of successful management at or near *status quo* levels for the utility of these methods should be examined.

6.2 Linear Regression

12) Various methods for carrying out linear regression are available, and most of the well-known ones can be shown to be maximum likelihood methods under appropriate assumptions.

13) The results obtained from the various methods differ appreciably only when the data are rather poor (i.e., the predictive value of the regression is rather low anyway).

14) Which of the methods is most appropriate (optimal) in a particular case depends on a variety of factors, notably:

   (a) the purpose of the calculation (prediction, determination of relationship, calibration),

   (b) the nature of the data (nature and magnitude of variability in both dependent and independent variables; whether or not the data were obtained by random sampling or in a controlled fashion),

   (c) the type of model (whether or not a finite intercept is *a priori* plausible, whether a curvilinear model is required).

15) The choice of an optimal model for prediction is particularly difficult, because it depends strongly on the precise nature and purpose of the prediction required (perhaps summarised by a loss function).

16) When it is quite clear that a predictive or functional regression is required, the appropriate method should be used. In cases of doubt, a method based on fitting the relationship (functional regression) is unlikely to lead to serious error, and may provide an acceptable (though not optimal) prediction if required.

17) The confidence region of any prediction made should be assessed. Likelihood mapping techniques are available and programs will soon be available, and these provide an excellent method of doing this. Conventional formulae for standard errors of a prediction provide a simple approximate alternative.

18) Appropriate treatment of the data and choice of appropriate models are more important than agonising over estimation procedures. All data should be plotted and examined for outliers. The evidence for and likelihood of heteroscedasticity (non-constancy of error variance) should be considered, and an appropriate transformation or weighted calculation should be carried out. The plausibility of various models (finite intercept, power law, etc.) should be considered. Residuals should be examined and tested. Where a computed intercept is not significantly different from zero, and a zero intercept is plausible, a regression through the origin will probably be more robust. More specific recommendations are given in Section 3.5.

19) Extrapolation outside the range of the data is likely to be erroneous unless the model is correct. If extrapolation is done, the result is largely based on faith in the model, and statistical methods provide no justification.

20) If a 'conservative' prediction (deviating less from the mean than other methods) is required for any reason, the classical regression of the quantity to be predicted on the other variate may be used. Such a prediction may not however be 'conservative' in other senses (e.g., so far as the fish stock is concerned), since it would for example be likely to overestimate recruitment when it is poor.

6.3 Forecasting of Recruitment

21) When no explanatory or predictive information is available, the median of past values is a more robust estimate of a typical value than the mean. If the data are log-normally distributed, the geometric mean is an approximation to the median.

22) If there is any reason to suppose that recruitment may be poor, or if one wishes to err on the side of caution, the lower quartile is a reasonably robust estimate of a typical low value.

23) There is no objection to forecasting recruitment from stock/recruitment data, but the predictive value is likely to be low. In the absence of other information, this procedure would be more cautious than using a typical value (e.g., the median) if the stock is low and there is any reason to suspect that recruitment failure is possible.

24) The only technique for forecasting recruitment with useful predictive value is usually the use of pre-recruit indices (though there are some exceptions to this rule). The quality of such indices is usually better if they are based on 1- or 2-group fish than on 0-group fish.

25) The data on both recruitment and recruit index are likely to be subject to variability which is approximately log-normal. A logarithmic transformation of both variables may therefore be useful. There may be various reasons to suspect that a non-linear relationship (possibly of power law form) might be appropriate. The slope of the regression should be forced to be one if it is not significantly different from one (implying a straight line through the origin whose slope is the geometric mean ratio of the data).

26) The usual considerations concerning linear regression apply. This case is (approximately) an example of the calibration problem, so the most appropriate method is to use the regression of survey index on VPA recruitment. If the variability of the recruitment estimates is substantial, the choice of the best method is not certain,but the use of maximum likelihood functional regression is likely to provide an acceptable solution. The appropriate value of λ should be chosen with care (see Section 3.2).

27) Multiple regression methods for forecasting recruitment are likely to be misleading if applied uncritically, and may have little predictive ability. No explanatory variables should be used for which a plausible mechanism cannot be constructed. The number of explanatory variables should be much less than the number of data points. The plausibility of a multiple linear model should be carefully considered before such methods are used - an alternative, more complex, model may be much more appropriate.

28) Time-series methods for forecasting recruitment have not usually been very successful, because the series are too short to give good results on data with high variability. If time-series models are fitted, they should be parsimonious (use few parameters), and they should be tested for predictive ability. Recruitment series often appear to be serially correlated, but statistical tests usually fail to reject the null hypothesis of no correlation.

29) Examination of survey data when positions are fixed usually shows a very strong positional effect, which may be misinterpreted as variability (and will appear as variability if survey station positions are not fixed). Multiplicative models generally explain much of the variance and can be used to set an upper limit on the variability of the survey data. Indexes constructed by simple summation are robust against a change of spatial distribution (whilst those from multiplicative models are not) and they place most weight on the most abundant catches. The statistical properties (efficiency, etc.) of such indices seem to be less poor than previously assumed.

6.4 Other Topics

30) Tests of various common models used in the interpretation of catch and age data show that they can often be rejected on statistical grounds. In particular, the assumption that fishing mortality is proportional to effort can be tested, and may not be consistent with the data. The practical consequences of any such imperfections in the models, however, need to be examined.

31) Least squares methods for the joint analysis of catch at age and survey on cpue data continue to give promising results, but a fully operational version is not yet available for the ICES area. It appears that the inclusion of additional constraints such as a separate model for fishing mortality would lead to a more robust method, but the computational consequences have not yet been determined.

32) Biological reference points such as F_{low}, F_{med} and F_{high} have been found useful in some investigations, and it has also been noted that the fishing mortality on North Sea herring exceeded F_{high} for 8 out of 9 years preceding the collapse of the stock.

33) The use of biological reference points such as F_{max} and $F_{0.1}$ as targets of management ignores some important social and economic factors. Biological reference points should be used as signposts and navigational markers, not necessarily as targets in their own right.

6.5 General

34) The reports of the Working Group should continue to be distributed to all members of Assessment Working Groups, and published in the Cooperative Research Report series, to make them accessible to a wider readership.

35) The Working Group considers that suitable topics for consideration at its next meeting would be:

1) Sensitivity of assessment techniques to assumptions concerning natural mortality.
2) Effect of discarding on assessment calculations, especially mesh assessments.
3) Advances in simpler methods of assessment (especially those based on the use of size composition).

36) The interval of one year between Working Group meetings has been found to be a little short to permit necessary research to be carried out, and the Working Group suggests that its next meeting should be held in November 1985.

ACKNOWLEDGEMENT

The Working Group acknowledges with thanks the assistance of Dr E. F. Harding, who made available a preliminary account of his work prior to publication.

7. REFERENCES

7.1 General

Anon. 1984. Reports of the ad hoc Working Group on the Use of Effort Data in Assessments, and the Working Group on the Methods of Fish Stock Assessments. ICES, Coop.Res.Rep., No.129.

Anon. 1984b. Report of the Industrial Fisheries Working Group. ICES, Doc. C.M.1984/Assess:9.

Collie, J S and M P Sissenwine. 1982. Estimating population size from relative abundance data measured with error. ICES, Doc. C.M.1982/G:34.

Cox, D R. 1977. The role of significance tests. Scandinavian Journal of Statistics, 4(2):49-70.

Cox, D R and D V Hinkley 1974. Theoretical statistics. Chapman and Hall.

Doubleday, W G. 1981. A method for estimating the abundance of survivors of an exploited fish population using commercial catch-at-age and research vessel abundance indices. Can.Spec.Publ.Fish. Aquat. Sci., 58:164-178.

Granger, C W J and P Newbold. 1977. Forecasting economic time series. Academic Press.

Gudmundsson, G. 1984. Statistical tests of common restrictions in fish stock assessment. (Unpubl. MS, Working Paper, No.11.)

Pope, J G and J G Shepherd. 1984. On the integrated analysis of catch-at-age and groundfish survey or cpue data. ICES, Doc. C.M.1984/G:16 (Working Paper, No.9).

Rivard, D. 1980. APL programs for stock assessment. Can.Tech.Rep.Fish. Aquat.Sci., 53, 103 pp.

Schaefer, M B. 1954. Some aspects of the dynamics of populations important to the management of commercial marine fisheries. Inter.Amer. Trop.Tuna Comm. Bull., 1(2):25-56.

Shepherd, J G. 1982. A family of general production models for exploited populations. Math.Biosciences, 59:77-93.

Stevens, S M. 1984. A method for weighting residuals in separable VPA. ICES, Doc. C.M.1984/D:4.

7.2 Simpler Methods for TACs

Anon. 1983. Report of the Working Group on Methods of Fish Stock Assessments. Doc. C.M.1983/Assess:17 (also appears in ICES Coop.Res. Rep., No.129, issued 1984).

Box, G E P and G M Jenkins. 1976. Time series analysis: forecasting and control. Rev.Ed. Holden-Day, San Francisco. 575 pp.

Boudreault, F-R, J-N Dupont et C Sylvain. 1977. Modèles linéaires de prédiction des débarquements de homard au Iles de la Madeleine (Golfe du St. Laurent). J.Fish.Res. Bd Can., 34:379-383.

Brody, S. 1927. Growth rates. Univ.Missouri Agric.Exped.St.Bull., No.97.

Brody, S. 1945. Bioenergetics and growth. Reinhold Publ.Corp., New York, N.Y. 1023 pp. (See also Ricker (1975), p.219.).

Deriso, R B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Can.J.Fish.Aquat.Sci., 37:268-282.

Fogarty, M J.1984 (MS). Temperature/yield relationships for the Maine American lobster (Homarus americanus) fishery: a time series analysis approach. (Unpubl. MS, Northeast Fish.Center, Woods Hole Lab., USA.)

Kirkley, J E, M Pennington and B E Brown. 1982. A short-term forecasting approach for analysing the effects of harvesting quotas: application to the Georges Bank yellowtail flounder (Limanda ferruginea) fishery. J.Cons.int.Explor.Mer, 40:173-175.

Mendelssohn, R. 1981. Using Box-Jenkin's model to forecast fishery dynamics: identification, estimation and checking. Fish.Bull., 78(4):887-896.

Pennington, M. 1983. Efficient estimators of abundance for fish and plankton surveys. Biometrics, 39:281-286.

Ponomarenko, V P. 1974. Prediction equations of total annual catch of the northeast Arctic cod and haddock stocks. App. to Rep. of the Northeast Arctic Fisheries WG, 1974. ICES, Doc. C.M.1974/F:7.

Pope, J G. 1983. Analogies to the status quo TACs: their nature and variance. Can.Spec.Publ.Fish.Aquat.Sci., 66:99-113. (Working Paper, No.4.)

Pope, J G. Short-cut and status quo TACs: an overview. (Working Paper, No.1).

Pope, J G. 1984. The performance of short-cut methods for catch forecasts. ICES, Doc. C.M.1984/D:3. (Working Paper, No.2)

Pope, J G and J G Shepherd. 1982. A simple method for the consistent interpretation of catch-at-age data. J.Cons.int.Explor.Mer, 40:176-184.

Pope, J G and D Gray. 1983. An investigation of the relationship between the precision of assessment data and the precision of total allowable catch. Can.Spec.Publ.Fish.Aquat.Sci., 66:151-157.

Ricker, W E. 1975. Computation and interpretation of biological statistics of fish populations. Bull.Fish.Res. Bd Can., 191, 382 pp.

Roff, D A. 1983. Analysis of catch/effort data: a comparison of three methods. Can.J.Fish.Aquat.Sci., 40:1496-1506.

Saila, S B, M Wigbout and R J Lermit. 1980. Comparison of some time series models for the analysis of fisheries data. J.Cons.int.Explor.Mer, 39(1):44-52.

Shepherd, J G. Status quo catch estimation and its use in fisheries management. ICES, Doc. C.M.1984/G:5. (Working Paper, No.3.)

Shepherd, J G. (MS). A time-dependent stock production model for fish stock assessment and short-term forecasts. (Unpubl.MS) (Working Paper, No.5.)

Schnute, J. 1977. Improved estimates for the Schaefer production model: theoretical considerations. J.Fish.Res. Bd Can., 34:583-603.

Stocker, M and R Hilborn. 1981. Short-term forecasting in marine fish stocks. Can.J.Fish.Aquat.Sci., 38:1247-1254.

Walter, G G. 1975. Graphical methods for estimating parameters in simple models of fisheries. J.Fish.Res.Bd Can., 32:2163-2168.

Walter, G G. 1976. Non-equilibrium regulation of fisheries. Int.Comm. NW Atlant.Fish. Sel.Paper 1:129-140.

Walter, G G. 1978. A surplus yield model incorporating recruitment and applied to a stock of Atlantic mackerel. J.Fish.Res. Bd Can., 35:229-234.

7.3 Linear Regression

Box, G E P and D R Cox. 1964. An analysis of transformations. J.R.Stat. Soc.B, pp.211-252.

Copas, J B. 1972. The likelihood surface in the linear functional relationship problem. J.Roy.Stat., Soc. B34:274-278.

Creasy, M A. 1956. Confidence limits for the gradient in the linear functional relationship. J.Roy.Stat.,Soc. B18:65-69.

Davies, O L and P L Goldsmith. 1976. Statistical methods in research and production (4th edit.). Longmans, London.

Draper and Smith. 1966. Applied regression analysis. J Wiley (New York).

Harding, E F. Considerations concerning the application of linear regression. (Unpubl.MS; Working Paper, No.7).

Hunter, W G and W F Lamboy. 1981. A Bayesian analysis of the linear calibration problem. Technometrics, 23:323-328, plus discussion pp. 329-350.

Jolicoeur, P. 1975. Linear regressions in fisheries research: some comments. J.Fish.Res. Bd Can., 32:1491-1494.

Lindley, D V. 1947. Regression lines and linear functional relationships. Proc.Roy.Stat., Soc. B9:218-244.

Lindley, D V and G N El-Sayad. 1968. The Bayesian estimation of a linear functional relationship. J.Roy.Stat. Soc. B30:190-2o2.

Pope, J A and A M Shanks. 1982. Fitting relationships in fishery research. ICES, Doc. C.M.1982/D:9.

Ricker, W E. 1973. Linear regressions in fisheries research. J.Fish.Res. Bd Can., 30:409-434.

Ricker, W E. 1975. A note concerning Prof. Jolicoeur's comments. J.Fish. Res. Bd Can., 32:1494-1498.

Seim, E and B-E Seither. 1983. On rethinking allometry: which regression model to use. J.Theor.Biol. 104:161-168.

Sprent, P. 1969. Models in regression and related topics. Methuen, London.

Sprent, P and G R Dolby. 1980. Reponse to query: the geometric mean functional relationship. Biometrics, 36(1):547-550.

7.4 Recruitment Prediction

Anderson, E D. 1977. Assessment of the Northwest Atlantic mackerel stock. ICES, Doc. C.M.1977/H:40, 26 pp.

Brown, B E and R C Hennemuth. 1971. Prediction of yellowtail flounder population size from pre-recruit catches. Int.Comm. Northw. Atlantic Fish Redbook 1971 (III):221-228.

Clark, S H. 1979. Application of bottom-trawl survey data to fish stock assessment. Fisheries (A Bulletin of the American Fisheries Society), 4(3):9-15.

Corten, A. 1980a. The use of larval abundance data for estimating the stock size of North Sea herring (II). ICES, Doc. C.M.1980/H:37.

Corten, A. 1980b. Report of ICES Young Fish Survey 1980: herring data. ICES, Doc. C.M.1980/H:35.

Cushing, D H. 1973. The dependence of recruitment on parent stock in different groups of fishes. J.Cons.int.Explor.Mer, 33(3):340-362.

Garrod, D J. 1983. On the variability of year class strength. ICES, J.Cons.int.Explor.Mer, 41:63-66.

Getz, W M and G L Swartzman. 1981. A probability transition matrix model for yield estimation in fisheries with highly variable recruitment. Can.J.Fish.Aquat.Sci., 38:847-855.

Hennemuth, R C, J E Palmer and B E Brown. 1980. A statistical distribution of recruitment in eighteen selected stocks. J.Northw.Atlantic Fish.Sci., 1:101-111.

Hunton, J K. 1984. Temporal and spatial variation in abundance estimates from Groundfish Surveys. ICES, Doc. C.M.1984/G:19.

Keeney, R and H Raiffa. 1976. Decisions with multiple objectives: preferences and value trade offs. Wiley, New York.

Mesnil, B. 1983. Indices d'abondance des juveniles de poissons plats devant les côtes françaises de la Manche est et Mer du Nord. Résultats des campagnes D.Y.F.S. 1977 à 1982. ICES, Doc. C.M.1983/G:55.

Moder, J J and S E Elmaghraby (Eds.). 1978. Handbook of operations research. Van Nostrand, Reinold, New York. 622 pp.

Murawski, S A. (MS) Methods for forecasting recruitment. (Unpubl. MS - Working Paper, No.8).

Saville, A and D Schnack. 1981. Overview - some thoughts on the current status of studies of fish egg and larval distribution and abundance. Rapp.P.-v. Réun.Cons.Int.Explor.Mer, 178:153-157.

Shepherd, J G. 1982. A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. J.Cons.int.Explor.Mer, 40(1):67-75.

Sissenwine, M P. 1984. Why do fish populations vary? Dahlen Workshop on Exploitation of Marine Communities, paper 1. (in press.)

Swartzman, G L, W M Getz, R C Francis, 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.

Zuboy, J. 1981. A new tool for fishery managers: The Delphi technique. North Amer. Journ. of Fisheries Management, 1:55-59.

Table 2.6.1 Estimates of *Status Quo* Catch for Simulated Data from Stock 1

Year	True *Status Quo* Catch	Estimates using		
		SHOT	DROP	DOPE
17	1 234	397	2 084	1 858
18	2 274	3 245	1 390	2 346
19	1 990	1 420	2 318	1 566
20	1 017	1 214	1 194	822
21	630	721	355	497
22	742	804	533	655
23	961	835	974	955
24	1 221	1 396	1 151	1 141
25	930	823	995	722
26	527	666	593	421
D		209 652	133 460	69 856
$1-D/\sigma^2$		.35	.59	.78
r.m.s pred. error	49% (CV)	40%	32%	23%

Table 2.6.2 Estimates of *Status Quo* Catch for Simulated Catch from Stock 2

Year	True *Status Quo* Catch	Estimates using		
		SHOT	DROP	DOPE
17	99	72	74	108
18	129	134	134	125
19	107	91	103	84
20	71	74	76	67
21	65	67	62	66
22	80	75	74	78
23	98	85	87	92
24	104	99	102	97
25	92	87	91	79
26	73	70	72	66
D		127.6	86.3	95.0
$1-D/\sigma^2$		.67	.78	.76
r.m.s pred.error	21% (CV)	12%	10%	11%

Table 2.6.3 Estimates of Status Quo Catch for Simulated Data from Stock 3

Year	True Status Quo Catch	Estimates using		
		SHOT	DROP	DOPE
17	99	81	87	64
18	154	150	120	147
19	127	105	100	102
20	106	119	107	122
21	88	102	116	103
22	81	96	97	99
23	90	93	94	91
24	112	112	103	112
25	89	102	100	102
26	80	91	95	90
D		171.3	282.0	273.1
$1-D/\sigma^2$		.68	.48	.50
r.m.s pred.error	23% (CV)	13%	16%	16%

Table 2.6.4 Prediction of Georges Bank Scallop

Data	Year	Y(t+1) true	Y(t+1) predicted	r^2 from regression	Years fitted	Independent Variables	Name
Using 1953-1983 Yield Data	1982	1 041	1 570	.35	27	Y(t)	
			1 702	.33	27	Y(t), Y(t-1)	
	1983	625	1 041	.40	28	Y(t)	
			959	.40	28	Y(t), Y(t-1)	
Using 1972-1983 Yield and VPA	1982	1 041	1 620	.40	10	Y(t)	
			1 245	.25	10	Y(t), R(3,t)	SHOT
			1 279	.25	10	Y(t), Y(t-1), R(3,t)	DROP
			1 382	.24	10	Y(t), C(t), R(3,t)	DOPE
	1983	625	1 108	.51	11	Y(t)	
			544	.26	11	Y(t), R(3,t)	SHOT
			553	.26	11	Y(t), Y(t-1), R(3,t)	DROP
			518	.26	11	Y(t), C(t), R(3,t)	DOPE
			1 141	.51	11	Y(t), C(t), R(2,t)	DOPE

Table 3.4.1 Data for worked example.
Read from Fig. 7 of Clark (1979) (NW Atlantic Mackerel).

VPA Recruits	Survey Index	Log VPA Recruits	Log Survey Index
800	1	6.68461	0
1500	1.5	7.31322	4.05465E-01
1550	5	7.34601	1.60944
1200	6	7.09008	1.79176
1300	7	7.17012	1.94591
2300	12.5	7.74066	2.52573
2200	13.5	7.69621	2.60269
2300	22	7.74066	3.09104

Figure 2.5.1 Comparison of the observed (———) and fitted (-----) lobster yield series for the period 1928-81. Data fitted by univariate ARIMA method (Fogarty, 1984, MS).

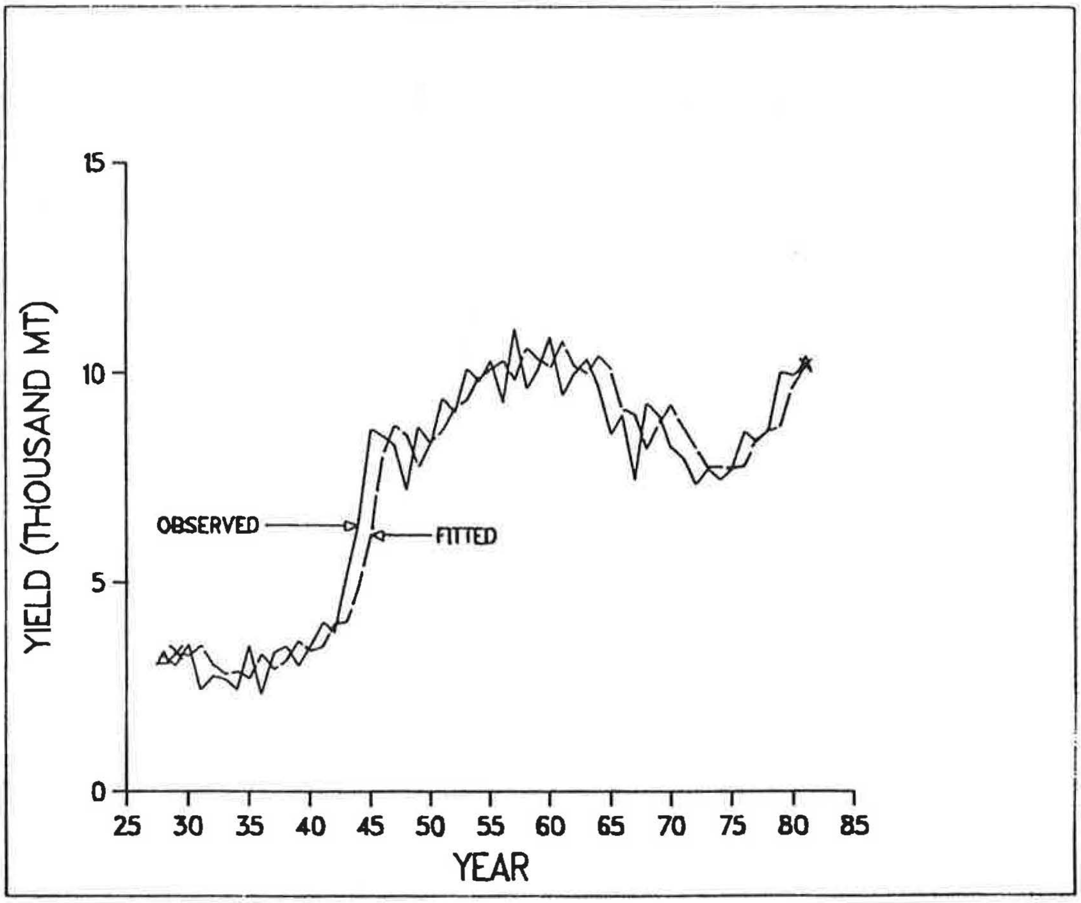


Figure 2.5.2 Comparison of the observed (———) fitted univariate (-·-·--) and fitted transfer function (—— --) series for the period 1945-81, for American lobster (Fogarty, 1984, MS).

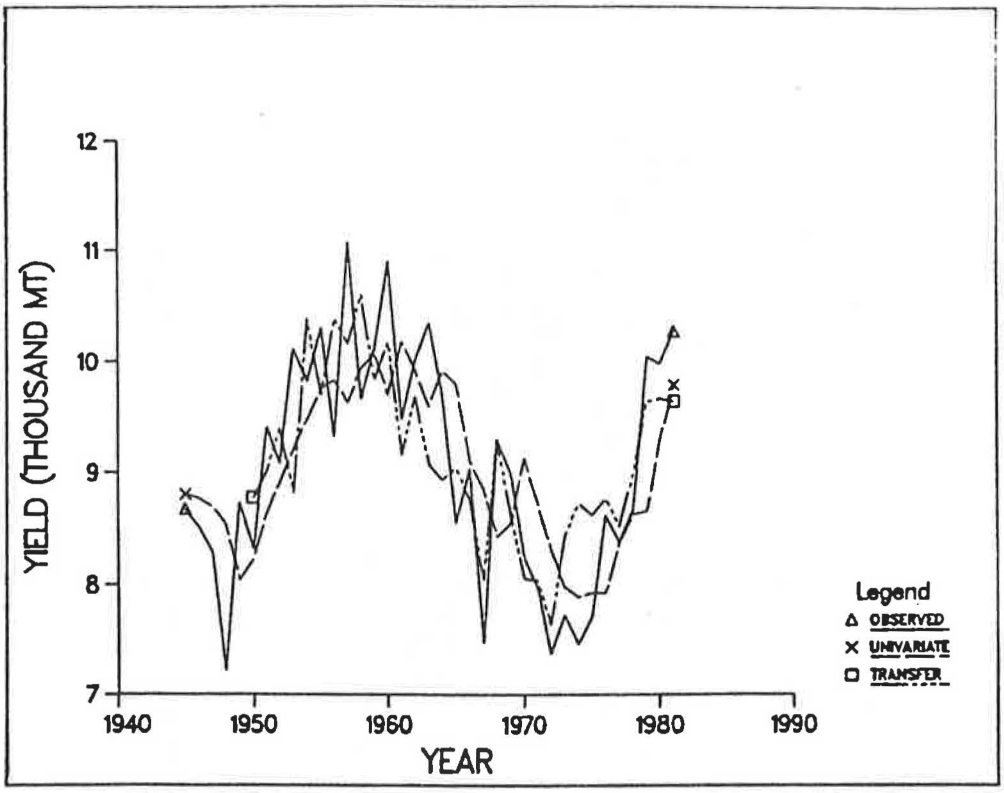


Figure 2.6.1 Plot Y(t+1) on Y(t) for Georges Bank scallops.

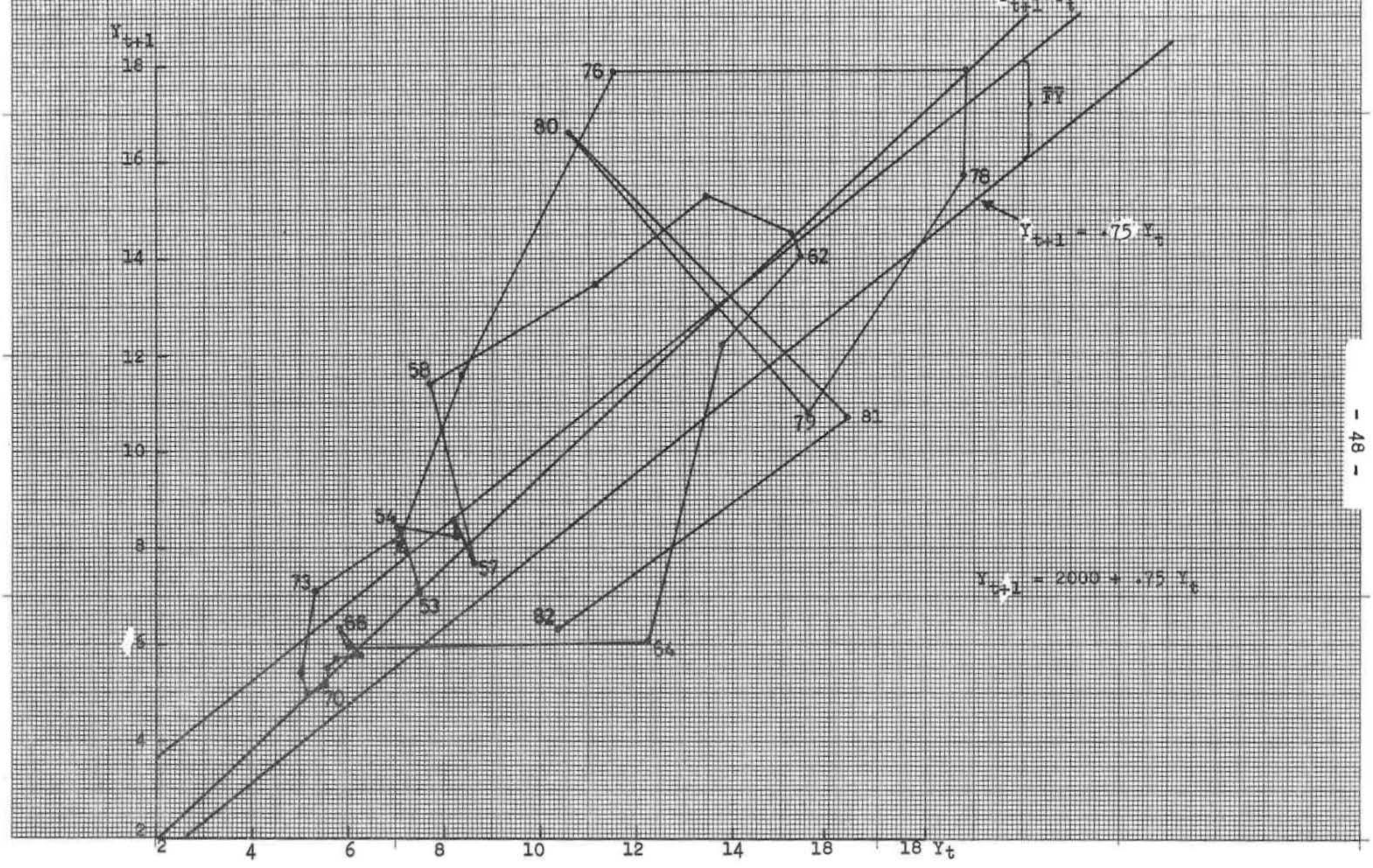


Figure 3.1 Regression analysis. (For a detailed explanation, see p.16.)

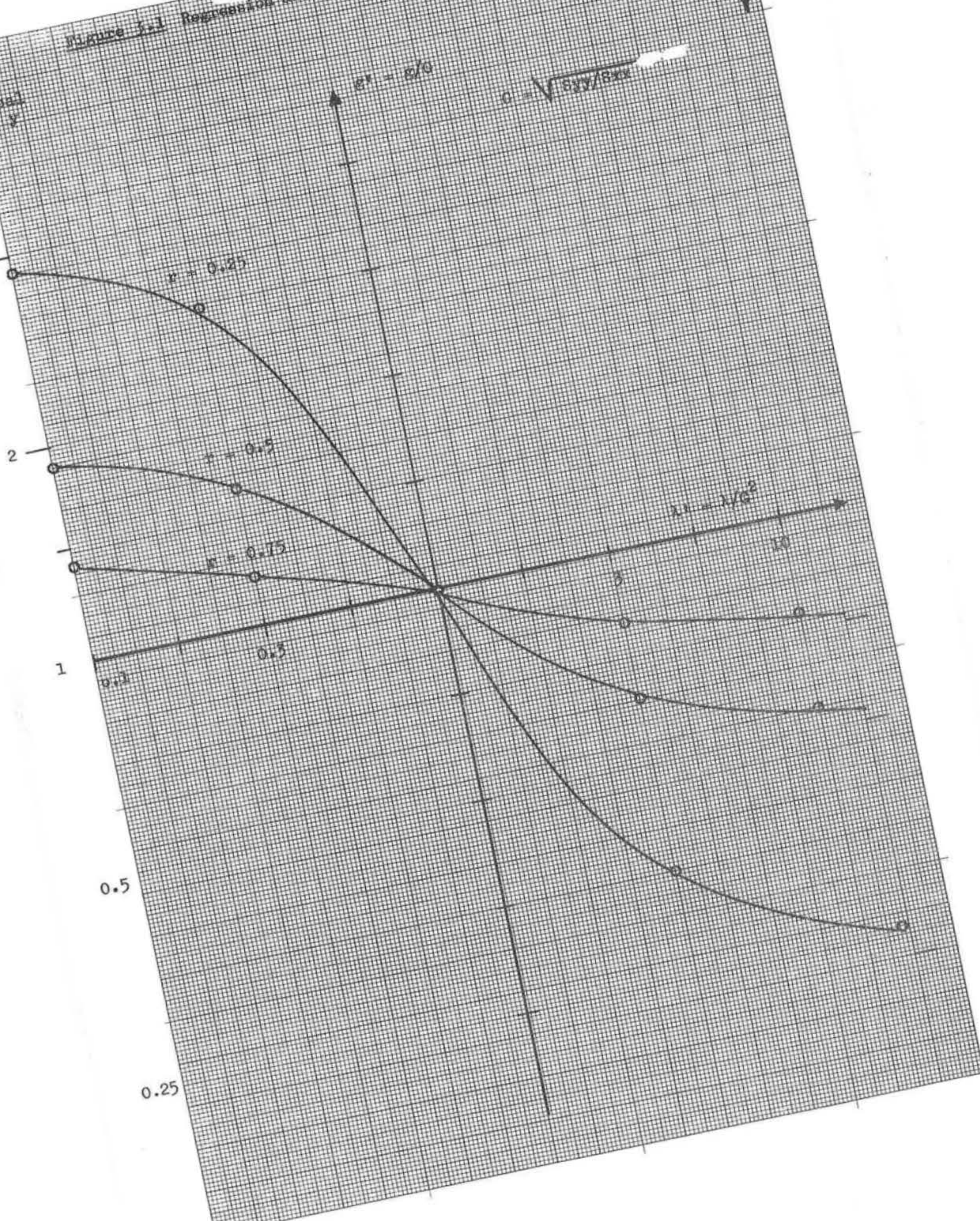


Figure 3.4.1 NW-Atlantic MACKEREL.

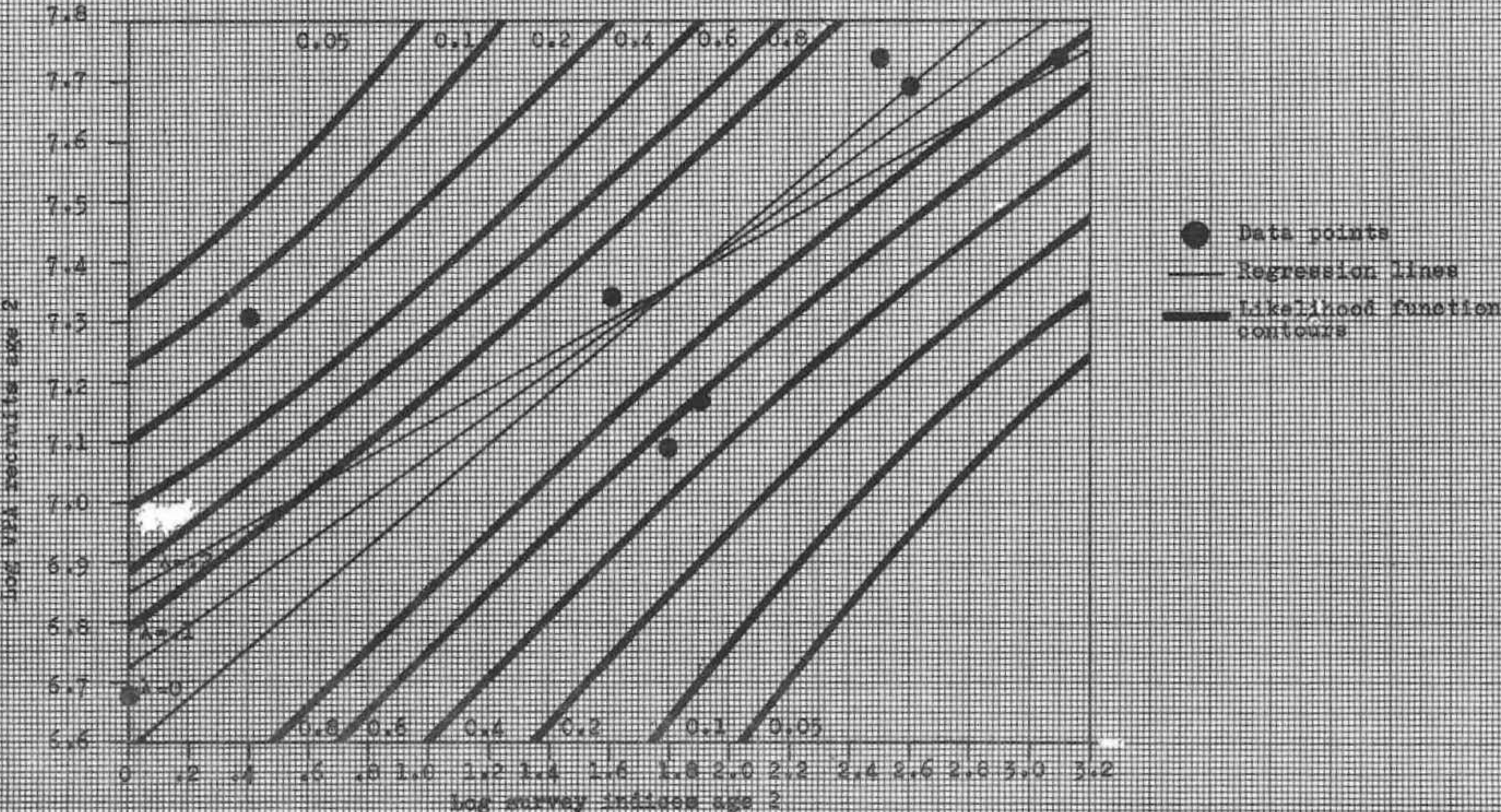


Figure 5.3.1 Spawning stock biomass per recruit calculations for Georges Bank HADDOCK. Values are expressed as a percentage of virgin spawning stock biomass/recruit and are a function of F and age at first capture.

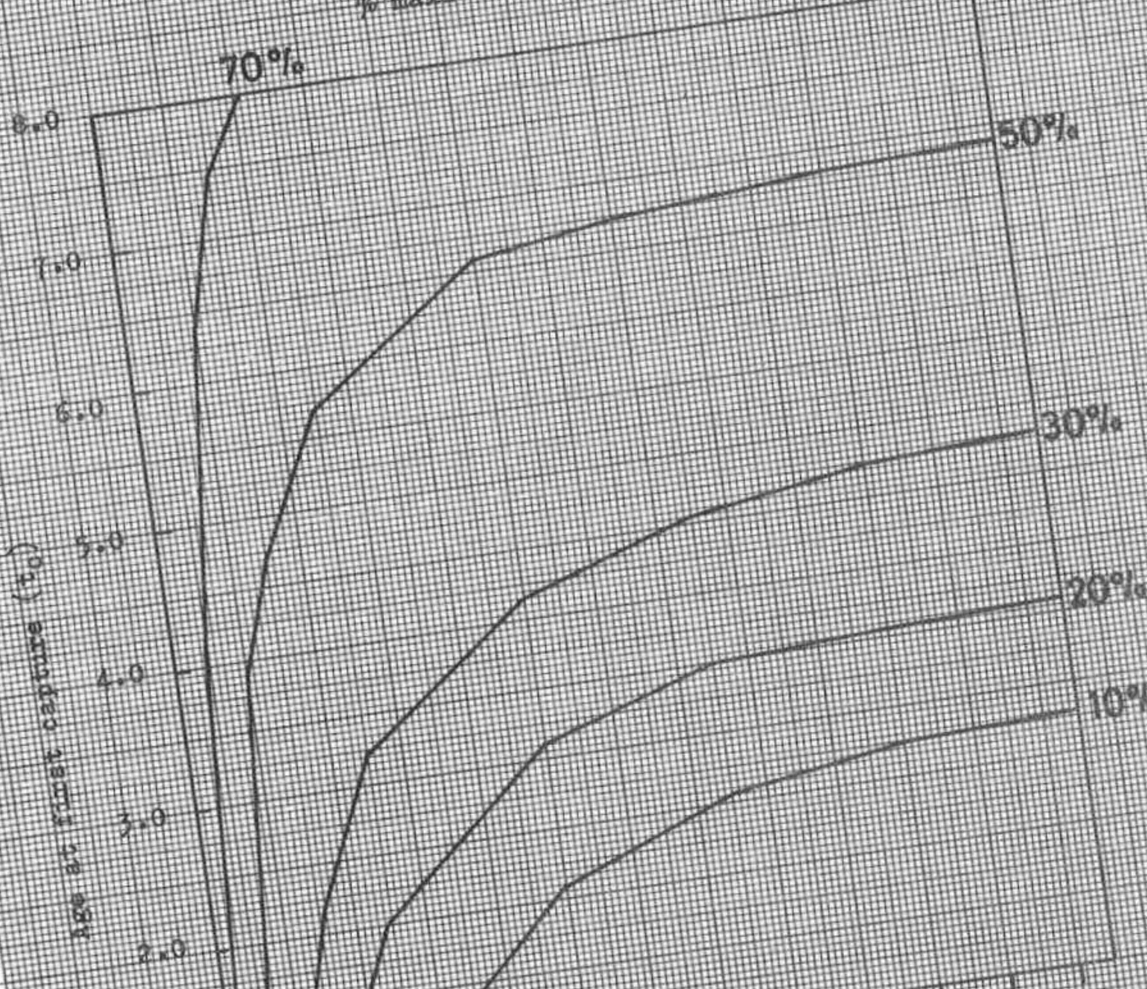


Figure 5.3.2 Stock/recruitment plot for Georges Bank HADDOCK, with 'replacement lines' for various proportions of virgin spawning stock biomass/recruit.

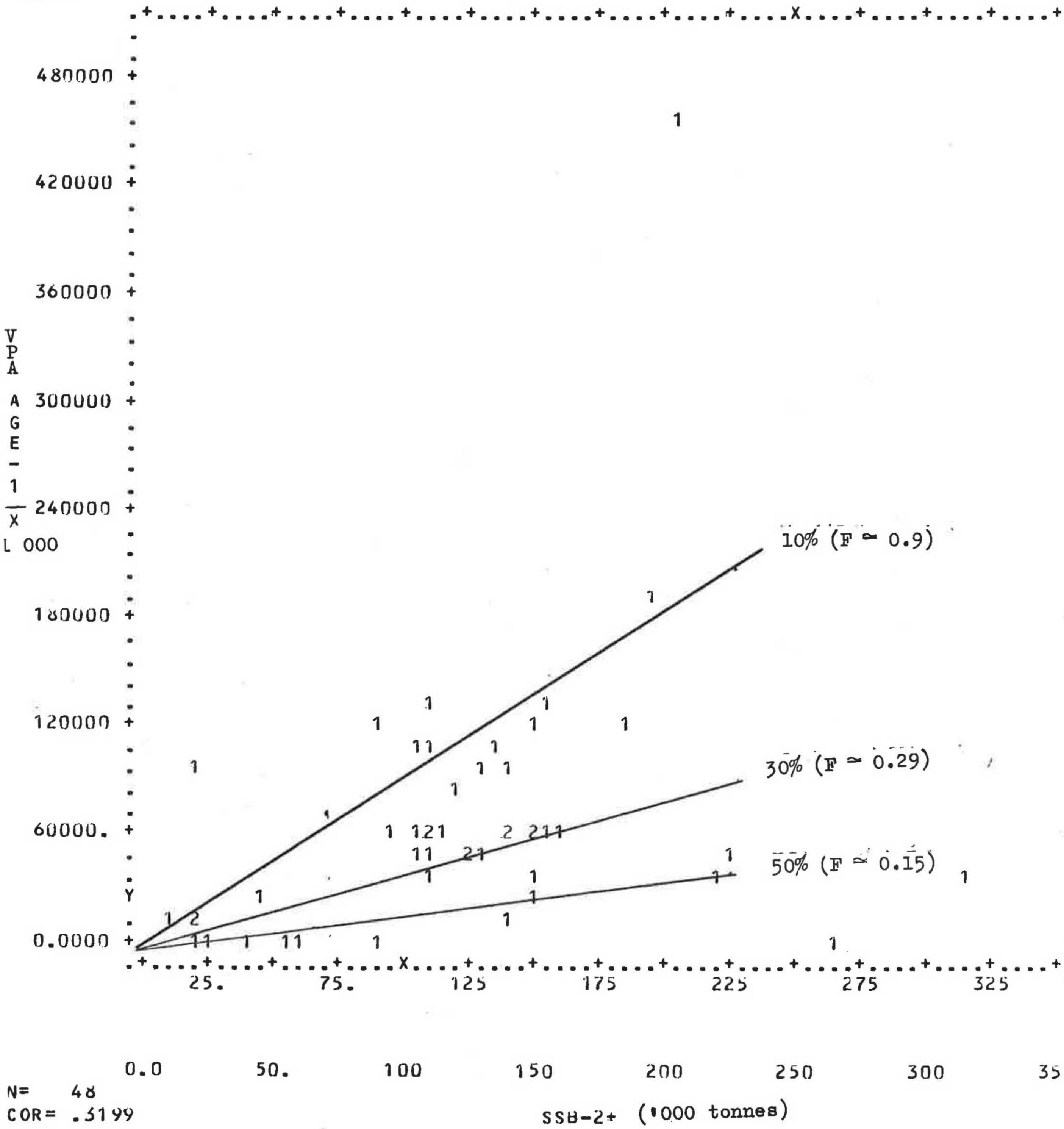
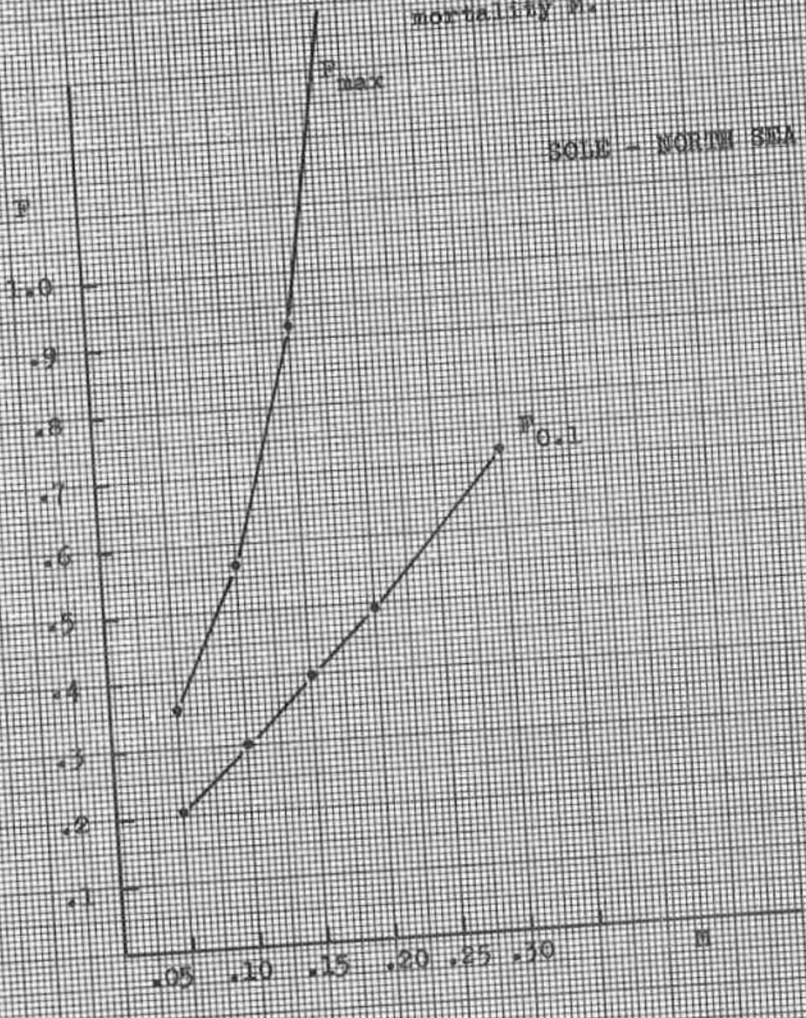
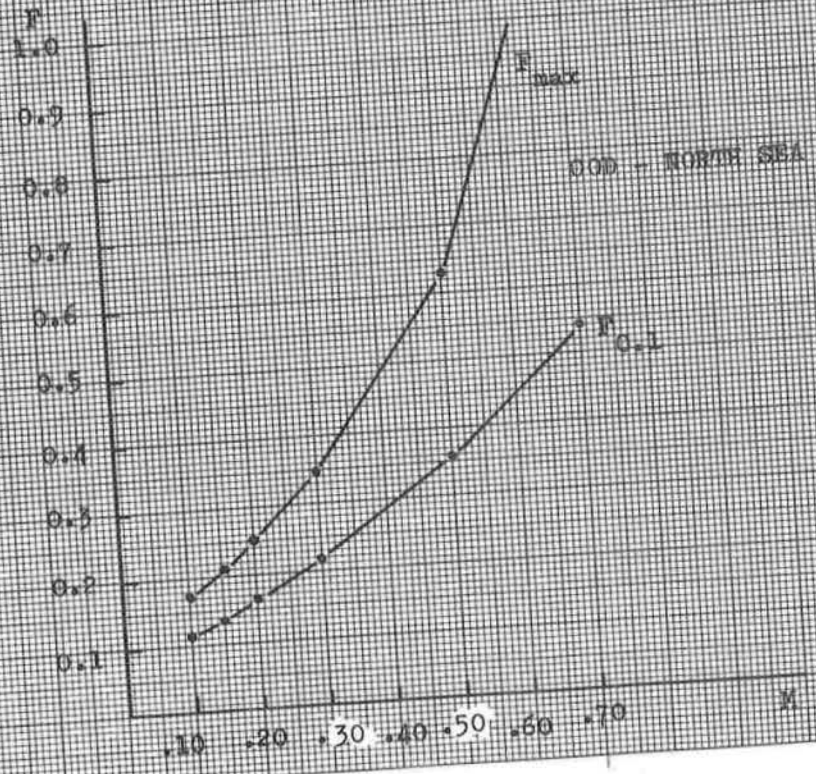


Figure 5.3.3 The dependence of biological reference points, F_{max} and $F_{0.1}$, of assumptions on natural mortality M.





APPENDIX A: Working Papers for 1984 Meeting

Simpler Methods for Catch Forecasts

1) J G Pope. Short-cut and *status quo* TACs: an overview (Unpubl.MS).

2) J G Pope. The peformance of short-cut methods for catch forecasts. ICES, Doc. C.M.1984/D:3.

3) J G Shepherd. *Status quo* catch estimation and its use in fisheries management. ICES, Doc. C.M.1984/G:5.

4) J G Pope. Analogies to the *status quo* TACs: their nature and variance. Can. Spec. Publ. Fish Aquat. Sci., 66: 99-113 (1983).

5) J G Shepherd. A time-dependent stock production model for fish stock assessment and short-term forecasts. (Unpubl.MS).

Linear Regression

6) J G Shepherd. Undated letter sent to WG members, Autumn 1983.

7) E F Harding. Considerations concerning the application of linear regression (Unpubl.MS).

10) J A Pope and A M Shanks. Fitting relationships in fishery research. ICES Doc. C.M.1982/D:9.

Methods for Forecasting Recruitment

8) S A Murawski. Methods for forecasting recruitment (Unpubl.MS).

Other Topics

9) J G Pope and J G Shepherd. On the integrated analysis of catch-at-age and groundfish survey or cpue data. ICES Doc. C.M.1984/G:16.

11) G Gudmundsson. Statistical tests of common restrictions in fish stock assessment (Unpubl.MS).

APPENDIX B: STANDARD NOTATION

(Note: Other minor usages are defined in the text)

Suffices and Indices

y	indicates	year
f	"	fleet
a	"	age group
t	"	last (terminal) year
g	"	oldest (greatest) age group
$	"	summation over all possible values of index (usually fleets)
#	"	summation over all fleets having effort data
@	"	an average (usually over years)
*	"	a reference value

Quantities

C (y,f,a)	Catch in number
E (y,f)	Fishing effort
F (y,f,a)	Fishing mortality
F_S (y,f)	Separable estimate of overall fishing mortality
q	Catchability coefficient (in P = qE)
Y	Yield in weight
W	Weight of an individual fish
B	Biomass
P	Population number (also fishing power)
E	Fishing effort
U	Yield or landings per unit of effort
C	Catch in numbers of fish (including discards)
N	Stock in numbers of fish
F	Instantaneous fishing mortality rate
Z	Instantaneous total mortality rate
M	Instantaneous natural mortality rate
S	Selection coefficient defined as the relative fishing mortality (over age)
R	Recruitment

APPENDIC C:

ICES WORKING GROUP ON METHODS OF FISH STOCK ASSESSMENT

SUMMARY OF TOPICS

		1981*	1983	1984	1985
1.	Application of Separable VPA		M	r	
2.	Simpler methods for TACs			M	p
3.	Measures of overall fishing mortality				
4.	Use of Effort data in assessments	M	M	r	
5.	Need for two-sex assessments				
6.	Computation and use of yield-per-recruit		M	m	
7.	Inclusion of discards in assessments				p
8.	Methods for estimation of recruitment			M	
9.	Density dependence (growth, mortality etc.)				
10.	Linear regression in assessments			M	
11.	Effect of age-dependent natural mortality				p

M = Major topic, m = minor topic, r = reprise, p = proposed

* Meeting of ICES Working Group on Use of Effort Data in Assessments

Indication of spine colours

Reports of the Advisory Committee on Fishery Management	Red
Reports of the Advisory Committee on Marine Pollution	Yellow
Fish Assessment Reports	Grey
Pollution Studies	Green
Others	Black

-o-o-o-