Advisory Committee on Fishery Management

ICES CM 1995/Assess:15

REPORT OF THE WORKING GROUP ON LONG-TERM MANAGEMENT MEASURES

Lowestoft, United Kingdom
4–12 April 1995

This report is not to be quoted without prior consultation with the General Secretary. The document is a report of an expert group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Palægade 2–4 DK-1261 Copenhagen K Denmark

TABLE OF CONTENTS

Section Page

1 INTRODUCTION

1.1 Participants

Abaunza, P.	Spain
Babayan, V.	Russia
Christensen, S.	Denmark
Darby, C.	UK(E&W)
Davis, B.	Canada
Gabriel, W.	USA
Jakobsen, T.	Norway
Kell, L.	UK(E&W)
Magnusson, K.	Iceland
Marchal, P.	UK(E&W)
O'Brien, C.	UK(E&W)
Pastoors, M. (part-time)	Netherlands
Pereiro, J.	Spain
Rackham, B.	UK(E&W)
Restrepo, V.	USA
Rijnsdorp, A.	Netherlands
Schweder, T. (part-time)	Norway
Shotton, R.	FAO
Sigurdsson, T.	Iceland
Sinclair, A.	Canada
Stokes, K. (Chairman)	UK(E&W)
Thompson, G.	USA

1.2 Terms of Reference

At the 82nd Statutory Meeting it was decided that the Working Group on Long-Term Management Measures (Chairman: Dr T.K. Stokes, UK) will meet in Lowestoft, UK from 4–12 April 1995 to:

a) using examples relevant to the ICES area develop further, methods for assessing the effects of technical conservation measures in different fisheries systems, taking account, as appropriate, of spatial and multi-species factors;

b) demonstrate the framework(s) for evaluating management strategies for fisheries systems (including MBALs), using North Sea plaice as an example. Suggest specific ways in which the results from such studies might be incorporated into the advice given by ACFM;

c) advise on the data (and quality) requirements needed to provide advice on the effects of technical conservation measures. In particular, advise on the feasibility of providing advice for widely distributed multinational fisheries and fish stocks;

d) define focus areas for further development of multi-species/multifleet assessment models for future work by the Working Group.

1.3 Acknowledgements

The Working Group would like to thank all at the Fisheries Laboratory, Lowestoft, for the excellent facilities and help provided and for their hospitality.

1.4 Outline and Introductory Remarks on Terms of Reference

In 1994, the Working Group (Anon, 1994a) distinguished between strategies, which define a general approach to achieving objectives, and tactics, which are the detailed measures by which strategies are implemented. Although a separation was made between the evaluation of strategies and of tactics, it was clearly stated that the evaluation of tactics should take place within the context of well defined strategies.

The Working Group stated that the evaluation of management strategies is performed most effectively in the context of entire management procedures (i.e, the combination of a particular assessment procedure plus particular control laws and their implementation). The Group considered that evaluation of management measures through simulation studies promised enhanced insights but recognised that results are dependent upon the characteristics of the simulated system. Simulations should thus not be viewed as providing predictions, but as a tool for comparison of the relative performance of alternative strategies applied to particular fisheries systems.

The Working Group described an approach to the evaluation of alternative strategies (see Fig. 1.4) that essentially relies upon scenario modelling, that is the construction of plausible underlying system models, the simulation of both assessment and control procedures (with feedback to the underlying system) and the recording of performance statistics from both the underlying system and the perceived system. Scenario modelling should take account of the range of uncertainties in the underlying system, observation, assessment, control implementation, etc. The outputs should be defined so as to permit the comparison of the performance of different management procedures.

The Working Group noted that the system model should be a plausible representation of the structural dynamics and incorporate appropriate process noise (e.g., stochastic recruitment) and that "observations" from the system (e.g., simulated survey abundance data) must include suitable error obtained by simulating a measurement procedure which samples the underlying system. The assessment model may mis-specify the underlying system as described above. Deviance in implementing the specified controls may occur. Evaluation of the management procedure involves simulating the underlying system and observations from the system, assessing the state of the system based on those observations, making predictions and implementing the controls over a time period whilst monitoring performance indices and their statistics.

The approach advocated by the LTMM WG is well established in the resource management context and has been adopted in a variety of fisheries and regions (e.g. de la Mare, 1985, 1986; Donovan, 1989; Francis, 1992; Horwood, 1994; IWC, 1993; Powers and Restrepo, 1993; Punt, 1991,1995; Restrepo *et al*, 1992; Restrepo and Rosenberg, 1994; Sakuramoto and Tanaka, 1986; UNEP, 1992).

At this meeting, the Working Group has tried to be more integrative and not to make a clear distinction between evaluating tactics and objectives. The report follows a sequence through the analysis of basic data to consider their information content (especially with respect to spatial factors; Section 2), analyses of highly detailed fisheries data in order to better understand the nature of certain fisheries (also Section 2), the evaluation of management measures (strategies and tactics; Section 3) and, finally, how to incorporate results of analyses into the advice of ACFM (Section 5). A demonstration of an evaluation framework, applied to North Sea plaice, is given in Section 4 and recommendations are made in Section 7.

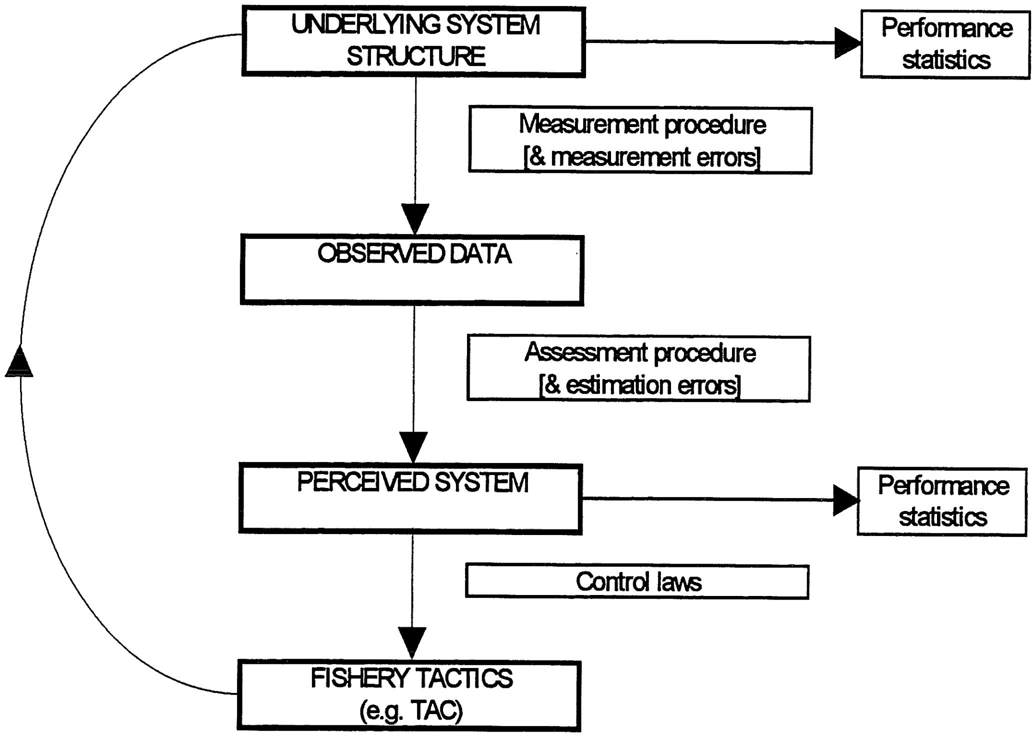


Figure 1.4 Flow chart for simulations used in the evaluation framework (after Anon, 1994a)

Section 2 of this report considers work in progress on understanding the details of different fisheries systems. The section addresses ToRs (a & c). Work on detailed analyses of English cod and plaice vessel trip data is described and inferences are drawn as to the feasibility of attempting to predict the detailed consequences of measures such as closed areas. Some strategic modelling which addresses the effects of closed areas of different sizes when different transport and effort reallocation rates are assumed is also presented. The underlying system model is not regarded as definitive and the results of the study should only be considered as demonstrative. The results make concrete, however, the common wisdom that to be effective, closed areas need to be large and that transport rate needs to be relatively small.

Investigations of the distributions of fish and fishermen are described together with methods for assessing how to manage mixed fisheries to achieve maximum revenue returns whilst maintaining the by-catch of protected species at acceptable levels. The discussion for the section focuses on the interaction between spatial and temporal data, model specification and the implications for providing advice on the effects of management measures.

Section 3 includes descriptions of a number of attempts to apply the scenario modelling approach to evaluate the performance of different management approaches. The examples all use very different programming methods and range from prototype procedures still in early development to highly developed multispecies, multifleet applications with carefully considered experimental designs. The examples include not only biological, but also economic, detail. In addition to the computer intensive approaches, an analytic approach to deriving a risk-averse long-term harvesting strategy is described. Such methodology offers a rigorous and formal approach that should be developed further in the future. Generally, progress in the area of evaluating management strategies has been encouraging. The Working Group has served a useful purpose as a forum for groups working in this field and should continue to serve this function in the future.

Section 4 addresses ToR (b). A single species, multifleet, underlying system model for North Sea plaice is developed and used within an evaluation framework described in section 3.2.3. The evaluation uses the standard ICES assessment routine (XSA) and compares the performance of a number of different management procedures applied to the underlying system with feedback control. Again, the underlying system model is not regarded as definitive and the results of the study should only be considered as demonstrative. Work on plaice could now be continued, drawing on a number of sources - the basic data analyses and detailed spatial models described in Section 2, the evaluation procedures described in Section 3 and 4 and other biological (and economic) inputs. Overall, the fisheries for North Sea plaice could provide an ideal place to start work on a comprehensive evaluation of various management measures.

The essential message of Section 5 is that this Working Group can supply interpretable summaries and diagrams, in a consistent form, as a basis for the comparison and evaluation of management measures. Whether or not ACFM would wish to incorporate all or any such summaries and diagrams into its report is debatable. Nevertheless, they should be suitable for interpretation and consequent provision of management advice.

Section 6 discusses the future of the Working Group and possibilities for conducting well founded evaluation work within the ICES structure. No conclusions are drawn but recommendations are made as to the future work plan and chairmanship.

2 METHODS FOR ASSESSING THE EFFECTS OF TECHNICAL CONSERVATION MEASURES

2.1 Introduction

The use of technical measures such as closed areas has become an increasingly important management tactic within the ICES area. Catch databases disaggregated by species, age, quarter, fleet and ICES statistical rectangle have, for example, been used as input to elaborate deterministic models in attempts to assess (i.e.,., to forecast or "predict") the effects of measures such as area closures or mesh increases. National data collection programs have not, however, been generally designed to provide catch-at-age or other data at the fine-grained level required for such work. More importantly, little or no work has been undertaken to assess the feasibility of such approaches given the supporting knowledge and information content of data.

In this section, examples of work relating to the fundamentals of prediction and monitoring are presented. These range from analyses of vessel trip data, to strategic models to aid in understanding the interaction of closed area size and fish transport rate, and to detailed simulation models of North Sea plaice taking account of fish movements.

2.2 Examples

2.2.1 Modelling and predicting catchability-at-age using English vessel trip data

WD5 and WD13 present an approach taken to investigate the relationship between catchability and its variance for two species (cod and plaice) in the North Sea and then to assess the performance of various predictive models for catchability based on variables and factors which are controllable by management actions. The approach was illustrated specifically for English data obtained from individual vessel trips for which market samples were available. This approach might be used for any species and fleets.

For each species separately, the modelling of a fleet's catchability-at-age in year y, q(y,a,f), was assumed to have the form:

$$q(y,a,f) = F_p(y,a,f) / E(y,f)$$

where a fleet's partial F-at-age, $F_p(y,a,f) = C(y,a,f)*F(y,a)/C(y,a)$, was determined through a VPA but the effort exerted by a fleet in a particular year, E(y,f), was assumed to be known for survey data but to be unknown for commercial data. Initially, fleet catchabilities-at-age were assumed to have log-normally distributed errors but this was shown to be an untenable assumption when applied to survey data collected by the English Groundfish Survey (EGFS). In the case of English commercial vessel trip data, there was the additional complication of a need to determine a suitable effort function before analysis of q(y,a,f) could begin. Therefore, the analyses presented were undertaken using $F_p(y,a,f)$ and an estimate for $var\{F_p(y,a,f)\}$, based on a combination of the variance due to ALK (age-length key) sampling and the variance due to length sampling, together with the estimation of a suitable effort function.

For each species, the partial F-at-age was shown to be consistent with an assumption of constant coefficient of variation (CV):

$$\mathrm{var}\{F_p(y,a,f)\} = \Phi^2 [E\{F_p(y,a,f)\}]^2$$

where E denotes expectation and Φ denotes the coefficient of variation. Further, $F_p(y,a,f)$ followed the multiplicative model:

$$F_p(y,a,f) = g_1(\text{variables related to effort})\, g_2(\text{concomitant variables})\, \varepsilon$$

where g_1, g_2 were functions specific to the stock and ε are independently identically distributed with $E\{\varepsilon\}=1$ such that:

$$\ln E\{F_p(y,a,f)\} = f_1(\text{variables related to effort}) + f_2(\text{concomitant variables})$$

The terms f_1, f_2 are linear functions determined through modelling. Estimation of the linear functions f_1 and f_2 was investigated by quasi-likelihood estimation with the assumption of constant CV and logarithmic link. Initial model building was restricted to consideration of those variables thought to influence the effort function f_1, followed by an investigation of whether or not any interactions existed between the variables incorporated into the effort function f_1. Finally, the specification of the concomitant variable function f_2 was investigated. Selection of variables in a (forward) sequential manner on the basis of their contribution to the deviance reduction was adopted. Variable selection on the basis of backward elimination from a maximal model was not feasible because of problems caused by the presence of multicollinearity in the data sets for each year.

From the analyses presented on English data, it appears that the effort formulation with an age effect may be estimated. Furthermore, it would appear that little or no information relating to the provision of advice at spatial scales for the two fishery stocks (cod and plaice) is estimable in a meaningful way from the available English data using this technique (without the imposition of external constraints). However, this is in direct contrast to an analysis of the STCF database of catch and effort presented in Anon(1994b) where spatial effects were postulated by rectangle by fleet by quarter by species. To reconcile this apparent disparity between the results presented in Anon(1994b) for plaice and those presented in WD13, it may only be possible to estimate a suitable effort formulation using the market samples of a particular country. However, once the effort formulation has been determined as precisely as possible for each country in a fishery, the catch data in the STCF database may then (possibly) be used to investigate potential spatial effects.

Using the relevant parameter estimates for the effort formulation specific to a particular year and stock, the consequences of applying a previously estimated/predicted effort function to data subsequently collected was investigated. Effort predictions for year #n based on analyses of data from year #{n-1} or year #{n-2} or may be consistently under-estimated or biased. Changes in the underlying distribution of vessel trawl types may, and probably will, influence the effort formulation and its estimation. While future predictions for year #n must be made with the best available information, continuous monitoring of a fishery would seem advisable if one is to detect changes sooner rather than later.

2.2.2 How well does fishery distribution reflect fish distribution?

Commercial fisheries are a potential source of information on fish distribution for seasons and areas without research surveys. How well fisheries map fish distribution can depend on the degree of competition among fishers. If interference competition occurs in fisheries, then fishing effort is expected to be distributed among areas so as to equalise catch rates among areas (e.g. Gillis *et al.* 1993). In this case, fish density may be better mapped by spatial variation in fishing effort than by spatial variation in catch rates.

These questions were examined using fishery logbook and research vessel data for September in the southern gulf of St. Lawrence. Logbook data for cod-directed trips by otter trawlers and seiners were aggregated by 10' grids of latitude and longitude. Survey catch rates of commercial-sized cod were interpolated to the same 10' grid using kriging. Stratum means were also compared. Data were available for 6 years: 1986, 1988-92. Fishing effort in September 1987 was too small to include in the analysis.

Comparisons of weekly changes in effort with weekly changes in catch rates failed to reveal evidence of competition. A Leslie analysis modelling catch rates in terms of cumulative catch also revealed no evidence of exploitation competition.

Neither effort nor commercial catch rates were closely related to the survey estimates of fish density at either the 10' grid or stratum scale ($R^2 \le 0.05$). The spatial pattern of effort was more closely correlated between years than were the spatial patterns of survey and commercial catch rates. Examination of fishery and fish distribution maps indicated that effort was deployed to the same areas each year and the occurrence in some years of high concentrations of cod in areas not fished. At this time of year, effort was deployed near to home ports and spatial coverage of the fishery did not extend to many areas occupied by the stock.

The spatial distribution of effort and commercial catch rates in the southern Gulf of St. Lawrence in September did not map cod distribution as seen by the research survey. In September, cod are dispersed over the feeding grounds and catch rates are at the annual low. Considerations such as marginal costs appear to dominate the fishing behaviour and effort is deployed in traditional patterns near home ports. Fishery behaviour and the relation between fishery and fish distribution needs to be examined at other times of year when fish are more aggregated and fishing effort is more intense.

2.2.3 Linear programming applied to spatially and temporally disaggregated multispecies landings and revenue data

Linear programming methods can be applied to spatially and temporally disaggregated multispecies landings and revenue data, to determine which combination of areas, seasons, and gear types maximise revenue from non-restricted species relative to impact on restricted species (Logan, pers. comm.; Mayo, pers. comm.; Figure 2.2.3). When landings of some species are severely restricted by total allowable catch levels, this technique allows the identification of fisheries which generate most of their revenue from non-restricted species (compared to the fishing mortality they exert on restricted species as e.g., by-catch) and which in sum would be expected to catch less than the TACs for the restricted species. This potentially enables those fisheries to continue, although major fisheries targeting restricted species may be closed. To prevent the increased targeting on restricted species in fisheries where those species were primarily minor by-catch, landings of restricted species may be prohibited. Implementation of this type of approach is equivalent to separate TACs for each area/season/gear combination. Without restrictions on new entrants to open area/season/gear combinations, currently viable fisheries may deteriorate. Ratios of revenue generated by non-restricted species to total landings of restricted species would be expected to change as stocks of restricted species rebuilt. Ratios also do not include discards of restricted species, which may change the results of the linear programming model as well. These characteristics, combined with uncertainty about how effort from closed fisheries would be reallocated, would require flexible adaptive management measures as management proceeded.

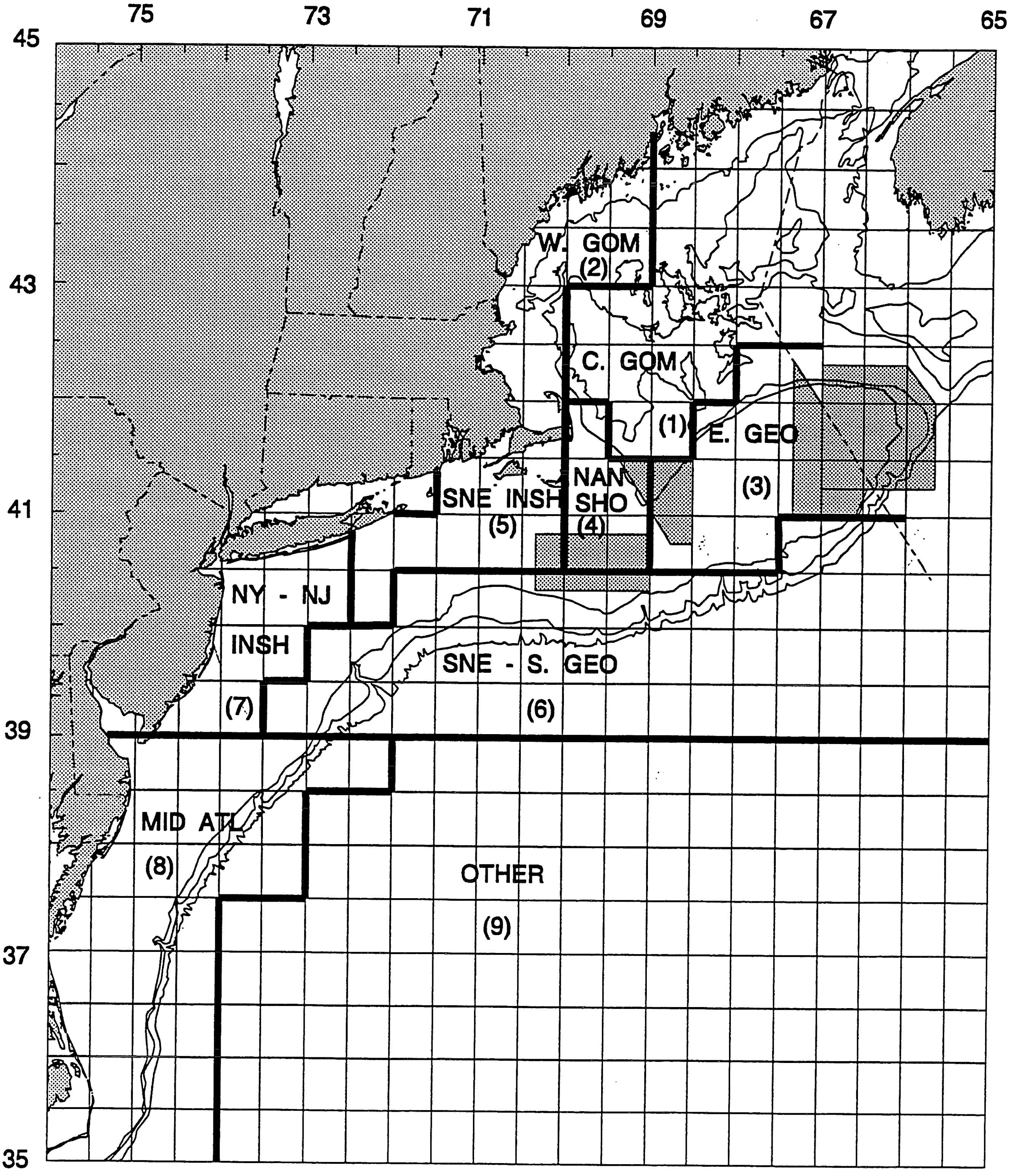
2.2.4 SIMP - Simulating the spatial dynamics of North Sea plaice

2.2.4.1 Model description

As a part of a bio-economic simulation model of the flatfish fishery in the North Sea (Janssen *et al.* 1994), developed by the National Institute for Coastal and Marine Management, the Netherlands Institute for Agricultural Economics and the Netherlands Institute for Fisheries Research, a biological simulation model was developed that described the spatial dynamics of plaice capturing both the seasonal and ontogenetical changes in distributions (WD 10; Rijnsdorp and Pastoors, 1994). The population model is based on principles of recruitment, growth, mortality (natural and fishing), migration and dispersion. The population is composed of six size classes: a pre-recruit size class, a discard size class representing the undersized fish which are caught but discarded, and four size classes representing commercial market categories. The current version of the model is deterministic and does not include density-dependent feedback mechanisms such as growth or stock-recruitment. It allows for one fishing fleet only.

Figure 2.2.3 (a) Example of fishing areas defined from fine-scale landings data summaries, cluster analyses of survey and commercial data, and management/industry input. (b) Example of linear programming results which identify feasible fisheries (area/seasonal/gear combinations) which have high rates of revenue from landings of non-restricted species relative to removals of restricted species (Atlantic cod, haddock and yellowtail flounder), constrained by total allowable catches for restricted species over all allowable area/seasonal/gear combinations. (c) Example summary of all open fisheries by area, quarter and gear type.

(a)



Fishing regions, quarter-degree squares and existing closure areas (shaded)

Continued

Figure 2.2.3 (continued)

OPEN areas
Redefined areas and fluke trawl separate
17:11 Tuesday, March 28, 1995

(b)

AREA=CENGOM

OBS	QTR	GEAR	GBCOD	GBYT	HADDOCK	SNEYT	GOMCOD	CHYT	REVRAT
1	Q1	DREDGE	0.0	0	0	0	0.0	1.0	3242564.00
2	Q1	HOOK	21053.0	0	4789	0	23396.5	49238.5	24.95
3	Q1	LMESHGIL	2788.0	0	5199	0	39215.5	47202.5	10.52
4	Q1	LMESHTRW	154765.0	298	118862	0	365735.0	639660.0	8.02
5	Q1	OTHER	0.0	0	0	0	50.5	50.5	394834.51
6	Q2	DREDGE	35.0	660	0	0	0.0	695.0	1041.02
7	Q2	HOOK	2396.0	0	0	0	6236.5	8632.5	24.99
8	Q2	OTHER	0.0	0	0	0	12.0	12.0	1749185.67
9	Q2	SMESHTRW	1332.5	0	100	0	2655.0	4087.5	11.68
10	Q3	DREDGE	47.0	222	0	0	1.5	270.5	1206.20
11	Q3	HOOK	17948.5	7	1316	0	19682.5	38954.0	9.62
12	Q3	LMESHTRW	243937.5	2834	145511	0	486751.0	879033.5	5.16
13	Q3	OTHER	0.0	0	0	0	0.0	1.0	22074673.00
14	Q3	SMESHTRW	160.0	0	0	0	0.0	160.0	1129.83
15	Q4	DREDGE	543.5	310	0	0	0.0	853.5	5119.51
16	Q4	HOOK	23684.0	0	1987	0	40355.5	66026.5	6.83
17	Q4	LMESHGIL	10633.5	0	4946	0	174080.0	189659.5	3.77
18	Q4	LMESHTRW	202938.0	1514	102289	0	404083.0	710824.0	6.53
19	Q4	OTHER	667.0	0	7	0	0.0	674.0	33723.39
20	Q4	SMESHTRW	0.0	0	250	0	598.0	848.0	9.77

AREA=EGEOBK

OBS	QTR	GEAR	GBCOD	GBYT	HADDOCK	SNEYT	GOMCOD	CHYT	REVRAT
21	Q1	DREDGE	13637.5	94965	380	8395	0.0	117377.5	81.58
22	Q1	FLUKETRW	40.0	115	5	0	0.0	160.0	105.20
23	Q1	HOOK	0.0	0	0	0	0.0	1.0	17917.00
24	Q1	LMESHGIL	0.0	0	0	0	199.5	199.5	38.55
25	Q1	OTHER	0.0	0	0	0	0.0	1.0	150019.00
26	Q2	DREDGE	16528.5	115255	7655	3485	0.0	142923.5	64.64
27	Q2	FLUKETRW	0.0	0	0	0	0.0	1.0	18942.00
28	Q2	HOOK	546.5	0	11	0	0.0	557.5	122.44
29	Q2	LMESHGIL	930.5	0	0	0	37.5	968.0	16.79
30	Q2	OTHER	0.0	0	0	0	0.0	1.0	161892.00
31	Q2	SMESHTRW	16373.5	1686	2565	0	0.0	20624.5	6.18
32	Q3	DREDGE	8515.0	165960	0	20095	0.0	194570.0	54.26
33	Q3	FLUKETRW	0.0	0	0	0	0.0	1.0	1720.00
34	Q3	HOOK	92.5	0	0	0	0.0	92.5	274.52
35	Q3	LMESHGIL	13000.0	0	0	0	0.0	13000.0	0.11
36	Q3	OTHER	0.0	0	0	0	0.0	1.0	771403.00
37	Q3	SMESHTRW	2909.0	1716	74	0	0.0	4699.0	249.22
38	Q4	DREDGE	6933.0	80550	240	2785	0.0	90508.0	71.51
39	Q4	FLUKETRW	0.0	0	0	0	0.0	1.0	58.00
40	Q4	HOOK	285.5	0	72	0	1869.5	2227.0	35.90
41	Q4	LMESHTRW	967508.0	517319	66175	1731	5250.0	1557983.0	1.73
42	Q4	OTHER	0.0	0	0	0	0.0	1.0	509550.00

Continued

Figure 2.2.3 (continued)

(c)

	CENGOM		WESGOM		EGEOBK		NANTSHO		SNEINS		SNESGB		MIDATL		NYNJINSH		OTHER	
HOOKS																	X	
DREDGES																	X	X
																	X	X
TRAWLS, LG MESH																	X	X
																		X
TRAWLS, SM MESH	X				X												X	X
						X											X	X
TRAWLS, FLUKE	X	X	X															X
	X	X	X	X													X	X
GILLNETS ,SM MESH		X			X	X				X	X	X					X	
	X	X			X	X	X	X	X		X	X	X	X		X	X	X
GILLNETS , LG MESH											X							
						X							X					
OTHER																		

Key input parameters of the model include growth rate (von Bertalanffy's Linf, K), monthly migration vectors describing the speed and direction of migration by month and by ICES rectangle, the relationship between migration speed and fish size, the initial population composition, the spatial distribution of the recruitment over the nursery grounds, catchability parameters by size class (assumed to be independent of size in initial runs), and the distribution of fishing effort by quarter and ICES rectangle.

The model is written in FORTRAN using the simulation environment SENECA. Input parameters are presented to the model as external files.

The model can be used to carry out cohort simulations on a per recruit basis and full population simulations with annual recruitment. The model was parameterised on the observed growth rate, the results of tagging experiments and the observed exploitation pattern from the VPA. An example of the output of a cohort simulation is given in Table 2.2.4.1.1. The estimation procedure of the catchability coefficient, assumed to be independent of size, is illustrated in Fig.2.2.4.1.1.

Table 2.2.4.1.1 Results of a cohort simulation run. Simulated 'true' population numbers and fishing mortality rates are compared to the values perceived by a VPA of the simulated landings. F-l and F-d refer to the partial fishing mortality rate of the landings and discard fractions. The input parameters of the simulation were: Linf=43.7cm, K=0.30, q=1.07E-07, m1=0.005, m2=0.04, m3=0.05, m4=0.129, m5=0.0136, m6=0.205

	Simulation results						VPA Perception	
Age	Popn. 1 January numbers	Z	F-d	F-l	Discards	Landings	F-l	Popn numbers
1	1000000	0.426	0.276	0.041	224500	33280	0.083	442978
2	652900	0.707	0.390	0.212	182700	99200	0.336	365490
3	322000	0.767	0.258	0.404	58120	90780	0.518	235331
4	149600	0.794	0.155	0.536	15960	55310	0.614	126255
5	67630	0.800	0.089	0.608	4153	28310	0.655	61610
6	30400	0.781	0.050	0.631	1052	13310	0.657	28882
7	13920	0.740	0.027	0.614	265	6035	0.628	13533
8	6640	0.705	0.014	0.590	67	2813	0.597	6541
9	3281	0.649	0.007	0.544	17	1314	0.547	3257
10	1715	0.619	0.003	0.518	4	662	0.519	1708
11	924	0.594	0.002	0.495	1	345	0.496	922
12	510	0.577	0.001	0.479	0	186	0.479	510
13	286	0.570	0.000	0.473	0	103	0.473	286
14	162	0.556	0.000	0.459	0	57	0.459	162
15	93	0.557	0.000	0.458	0	33	0.458	93
16	53							53

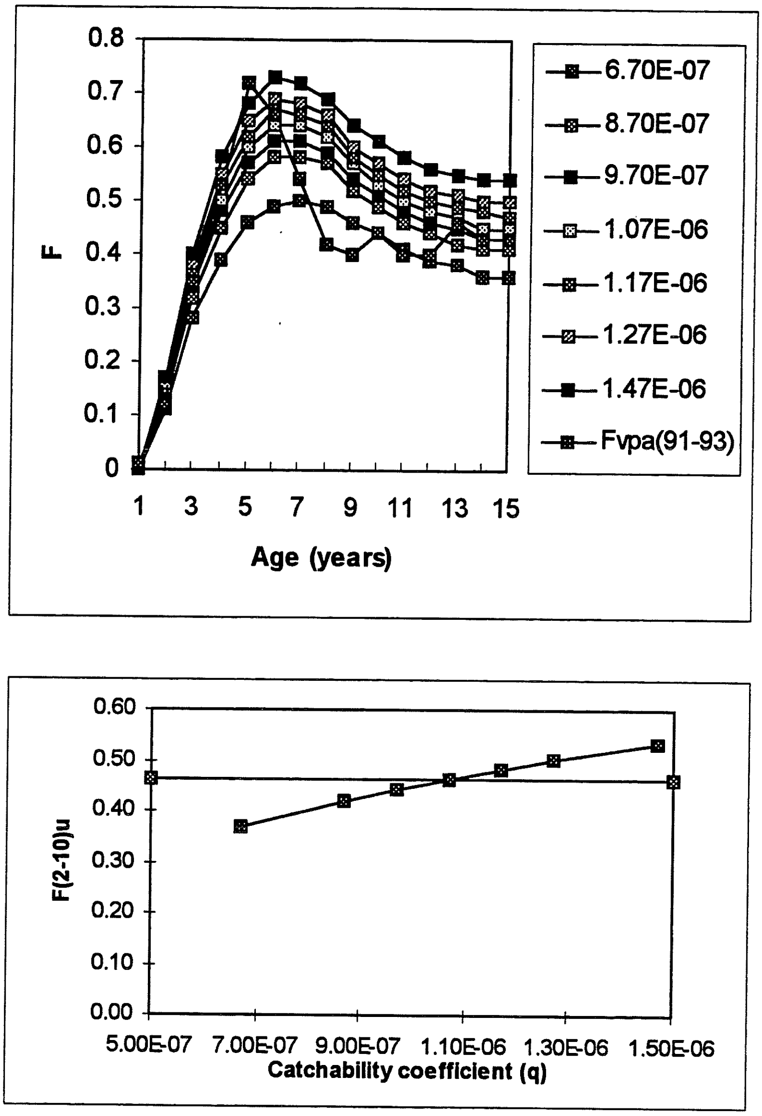


Figure 2.2.4.1.1 Simulated exploitation patterns at various levels of the catchability coefficient (q) and the VPA-estimated pattern (upper panel). At a q=1.07E-7 the F(2-10)u from the VPA and the simulations are equal (lower panel).

The simulated 'true' population numbers at the end of each year, numbers landed and discarded, and fishing and discard mortality rates are compared to the estimates perceived by VPA of the numbers landed. Due to the high discarding (59%), the recruitment to the fisheries perceived by the VPA as the number of 1-yr olds is only 44% of the true recruitment.

Performance statistics from the simulation are compared to observed values in Table 2.2.4.1.2. The simulated percentage discards is somewhat higher than the average percentage observed in 41 trips of commercial beam trawlers (van Beek, 1990). The Y/R is somewhat high compared to the values of the latest stock assessment (Anon, 1995), whereas the simulated SSB/R is somewhat too low. The simulated growth and exploitation pattern compares favourably with the observed values, although the observed exploitation pattern shows a more distinct peak at the younger age groups (Fig.2.2.4.1.1). Since parameterisation has not been optimised, the results should be considered as preliminary.

Table 2.2.4.1.2 Comparison of performance statistics of two cohort simulation runs and the observed values. Simulation 1 assumed catchability to be constant for size classes 2 to 6 (q=1.07E-7), whereas simulation-assumed a reduced q=0.75E-7 for the discard size class. % discards refers to the percentage of discards of the total catch numbers, Y/R and SSB/R refer to the yield and spawning stock biomass per recruit.

	Simulation 1	Simulation 2	Observation	Source
% discards	59.4	51.8	50%	van Beek (1990)
F(2-10)u	0.463	0.464	0.458	ICES (1995)
Yield/R (g)	294	283	219	ICES (1995)
SSB/R (g)	396	491	536	ICES (1995)

2.2.4.2 Sensitivity analysis

The sensitivity analysis presented in WD9 focused on the effects of migration rate, growth rate and catchability.

Both growth rate and migration rate influenced the level as well as the shape of the exploitation pattern (Fig.2.2.4.2.1). At low rates of growth or migration, the exploitation pattern was flat. At increasing rates of fishing mortality and growth the pattern became dome shaped. If migration rate further increased the exploitation pattern flattened again. When growth rate increased the percentage discards decreases and the corresponding percentage recruitment and the yield and SSB per recruit increases (see Table 2.2.4.2.1). Also the mean fishing mortality rate of the landings increases. Consistent results were obtained as a response to an increase in migration rate. Both of these results can be explained by the faster movement of plaice into the exploitable size classes.

Table 2.2.4.2.1 Effect of the growth rate (K) and migration speed (MF) on the various performance statistics: % discards, % recruits, Yield and SSB. The migration speed is varied by employing a multiplication factor MF to the standard migration speeds of 8.0E-03, 4.0E-02, 1.29E-01, 1.36E-01, 2.08E-01 of the various size classes respectively. F(2–10)u is the average F across ages 2–10.

Run#	K	% discards	% recruits	Yield	SSB	F(2–10)u
3	0.16	73.6	29.9	67.9	123.9	0.276
12	0.25	63.9	40.1	107.8	192.5	0.423
13	0.30	59.4	44.7	128.9	228.0	0.463
0	0.40	51.5	52.6	169.3	295.5	0.506

Run#	K	% discards	% recruits	Yield	SSB	F(2–10)u
33	0.25	54.7	51.3	30.3	56.9	0.361
29	0.50	56.5	48.4	67.9	123.9	0.402
30	0.75	58.1	46.3	103.5	185.2	0.429
13	1.00	59.4	44.7	128.9	228.0	0.463
31	1.50	61.4	42.3	130.4	230.5	0.530
32	2.00	62.8	40.6	150.9	264.8	0.557
34	4.00	65.8	37.4	170.9	298.1	0.565

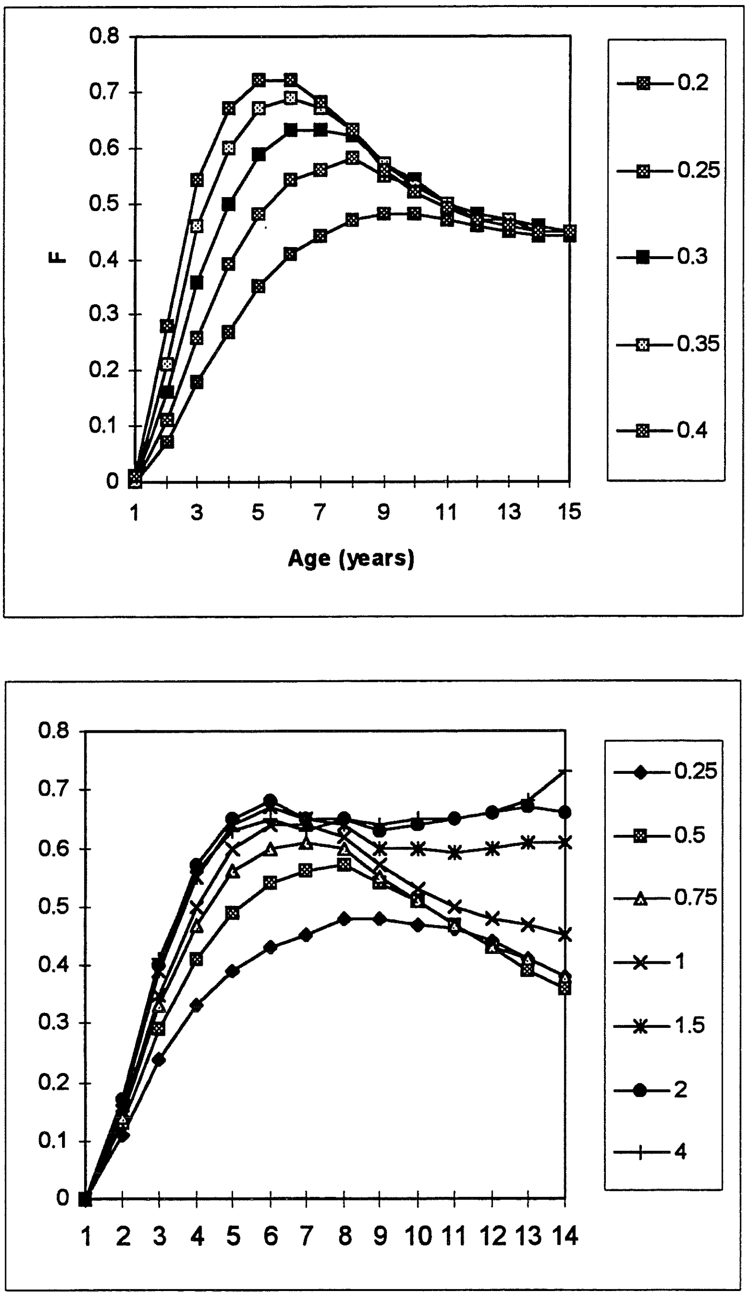


Figure 2.2.4.2.1 Simulated exploitation pattern as a function of growth rate (K; upper panel) and the migration speed (MF; lower panel).

Exploring the effects of added random variation on the migration vectors, as a sensitivity analysis, both on speed and direction suggested that the % discards, % recruits and Y/R and SSB/R were rather insensitive. A 10% random variance to the monthly migration vectors affected the performance statistics by less than 1%.

The various exploratory runs indicated that the simulated discard percentage was generally higher than the mean value observed. The discard percentage was highly sensitive for variations in the catchability of the discard size class. A reduction of the catchability of this size class by 25% yielded a discard percentage of 52%, which is close to the observed mean value. The corresponding performance statistics of this run also improved the correspondence of the Y/R an SSB/R statistics (Table 2.2.4.1.2).

2.2.4.3 Application: Closed areas (Plaice Box)

North Sea flatfish fisheries are characterised by substantial discarding of undersized fish. In order to reduce discarding a closed area - the "plaice box" - was established in 1989. The biological basis of this measure was a deterministic model employing an observed distribution of age groups and percentage of undersized fish by ICES rectangle and by quarter (Anon, 1987; Anon, 1993). The model predicted a 25% increase in recruitment to the fisheries when the box was closed for all discarding fleets in the 2nd and 3rd quarter. The predicted gain was revised downward to 8% in order to take account of the continued fishing by exemption fleets and the increased fishing in the plaice box area in the 4th quarter (Anon, 1994b). Despite the expected gain in recruitment, the realised yield and SSB decreased considerably in recent years (Anon, 1995). The perceived lack of effect, however, may be due to the observed decrease in juvenile growth which occurred in the second half of the 1980s (Rijnsdorp & van Leeuwen, 1994). The decrease in growth may have extended the period during which the fish are exposed to discarding and natural mortality. The extent to which changes in growth rate interfered with the effect of the plaice box could not be studied using the original models but have been explored using SIMP.

Two series of simulation runs were carried out for K values between 0.16 to 0.40 and for a situation with the effort distribution representative for 1991 under the assumption of no-box closure (base-line) and a full closure of the plaice box for all fleets during the whole year (plaice-box).

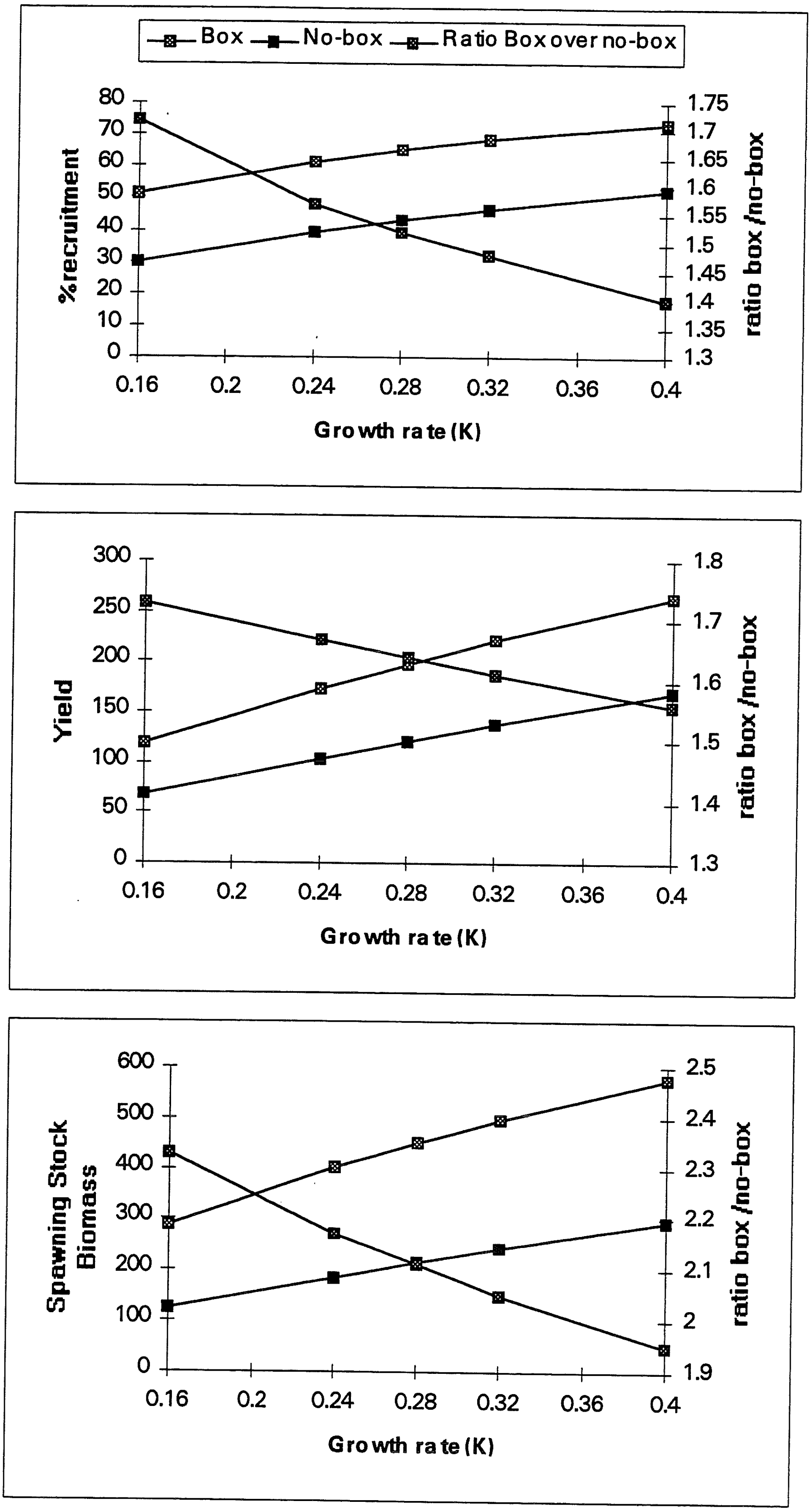


Figure 2.2.4.3.1 Simulated effects of a complete closure of the plaice box (whole year, all fleets) for various levels of growth rate (K). The recent decrease in plaice growth corresponded with a decrease in K of approximately 0.30 to 0.25. The relative effect of the plaice box is predicted to increase at reduced growth rates.

The base-line model was run with the 1991 effort distribution of the international fleets as used previously by the Study Group on the Plaice Box (Anon, 1994b). Although the statistical analysis used violated some of its assumptions (see WD13; Section 2.2.1), the resulting fleet correction factors in 1989 and 1991 were in agreement and the estimated distribution of international fishing effort may be considered realistic. This effort distribution approximates the distribution with no box although the plaice box was already in operation during the 2nd and 3rd quarter in that year. The data could nevertheless be used because the plaice box did not fully match the borders of the ICES rectangles and allowed fishing in 8 of the 10 plaice box rectangles. A study of the micro-distribution of Dutch beam trawlers showed that heavy fishing indeed occurred along the edges of the box (Fig 9 in Anon, 1994b).

In the plaice box run it was assumed that no fishing effort occurred in the plaice box rectangles during the whole year and that the plaice box effort was reallocated to the bordering rectangles. Hence, this run indicates the maximum effect of closing the box. The results of the simulation show that % Recruitment, Yield and SSB increase with growth rate (Fig. 2.2.4.3.1). The recently observed decrease in growth rate corresponds to an approximate decrease in K from 0.30 to 0.25, the % recruitment is reduced by approximately 10%, and the Yield and SSB by approximately 16%. Establishment of the plaice box gives a substantial increase in both % recruitment, Yield and SSB. The relative effect of the plaice box effect increases when the growth rate declines, as indicated by the ratio of the box statistic over the no-box statistic.

2.2.4.4 Discussion

Although the first model explorations have yielded encouraging results, it is recognised that the model made a number of simplifications and that model explorations have only just started. The simplifications and their likely effect on model performance will now be discussed.

The population comprised of a small number of size classes which resulted in a greater variability in growth. Although this effect has been reduced by allowing growth to occur by quarter only, increasing the number of size classes may be expected to improve the performance of the model.

The model includes both migration and dispersion components. The latter, however, was not modelled independently but resulted from the derivation of the migration vectors in the X- and Y-direction. The dispersion, therefore, will differ between areas in the North Sea depending on the direction and length of the migration vectors and may not necessarily lead to realistic results. Likewise, the gradual process of dispersion of juvenile fish from the coastal nursery grounds to the offshore areas was modelled by assuming that the smallest size classes followed the same annual migration cycle as the larger size classes, but with a much lower migration speed. Although the model may be improved by separately modelling dispersion and offshore movement of juveniles, the simulated ontogenetic change in distribution (Rijnsdorp and Pastoors, 1994), as well as the simulated exploitation patterns and percentage discards, suggest that dispersion is modelled with the right order of magnitude.

The lack of stock structure (e.g., homing behaviour) results in a gradual dispersion of a cohort over the sea. As a result, the surviving fish will tend to concentrate in the areas which are fished less intensively. This may not be realistic if fish show homing behaviour which gives a more restricted spatial distribution.

There is some evidence that the assumption of a size independent catchability does not hold in plaice, due to seasonal differences in catchability which are related to behavioural differences in the spawning period (Rijnsdorp 1993). Future improvements of the model may include the specification of a seasonal catchability coefficient. Also, the simulation did not take account of differences in biology between the sexes which show substantial differences in growth rate and exploitation pattern. The change in sex ratio with size, and hence the catchability and growth parameter, will be a function of the rate of exploitation. It is difficult to envisage the quantitative effects on the performance of the model.

In its present version the model is deterministic and does not take account of variance and uncertainty in the input parameters. The simulation environment SENECA offers, however, the possibility to include stochasticity and to explore the sensitivity of the performance statistics according to the fractional factorial design suggested in Section 3.3.

A further development may be to include feedback mechanisms such as density-dependent growth and a stock-recruitment relationship and to link it with models such as those presented in section 3, to construct a spatially disaggregated management evaluation model.

The high spatial resolution makes the model particularly suitable to study technical measures such as closed areas or seasons. The results of the exploration of the effects of the plaice box are preliminary because the input parameters of the model have not been optimised and the plaice box only restricts the fishing of part of the fleet. The simulations nevertheless provide relevant insight in that a decrease in growth rate will reduce recruitment and yield to the fisheries and thus may (partly) explain the recently observed decrease in these statistics.

The results so far obtained suggest that the model may provide a useful tool in exploring the effects of the plaice box in relation to changes in growth rate and to various levels of fishing effort of the exemption fleet. These question should be addressed taking account of the uncertainty in input parameters and alternative strategies of effort reallocation.

2.2.5 The interaction between closed area size and fish transport rate

Per recruit models can be used from a strategic perspective to evaluate the relative importance of different technical conservation measures and biological characteristics such as movement rates. Although models of this type can be relatively difficult to validate, they provide an easy way to consider effects of closed areas and associated reallocation of fishery effort when combined with fish movement rates. In one example case (Rago, pers. comm., Figure 2.2.5), it was shown that when most of the population was vulnerable to the fishery (i.e., unprotected by the closed area), the reallocation of effort to the open area had an important effect on spawning stock biomass per recruit. When most of the population was protected by the closed area, spawning stock biomass per recruit was very sensitive to the effect of transport rate. Stochastic features could be added to this type of model, as well as additional age, size or season-specific components to the transport rates. This approach is not computationally intensive and focuses on the importance of estimates or assumptions related to movement rates of both fish and fishermen.

2.3 Discussion and Conclusions

The theme common to this section can be generalised as a disaggregated model of catch, effort, and abundance:

$$C = qEN$$

which becomes

$$C_{a,y,f,r,q} = q_{a,y,f,r,q} E_{y,f,r,q} N_{a,y,r,q}$$

where the subscript a denotes age, y denotes year, f denotes fleet, r denotes region, and q denotes quarter.

However, F and N as estimated from a typical VPA only provide information on age and year components. The degree to which catch can be resolved to age, fleet, region, and quarter and to which effort can be resolved to region and quarter varies widely between national and international databases. Consequently, the degree to which $\overline{N}$ may be estimated on fine temporal and spatial scales is limited. Mismatches between estimators (e.g. partial F by fleet and quarter) and VPA results (e.g. annual beginning of year N) can lead to model misspecification.

Analyses of catch rates based on biomass summed over age may also be performed where the aim is to investigate overall F, rather than selection pattern.

Presentations helped in the identification of components of sampling schemes and databases which might need to be improved in order to increase confidence in the advice that might be given.

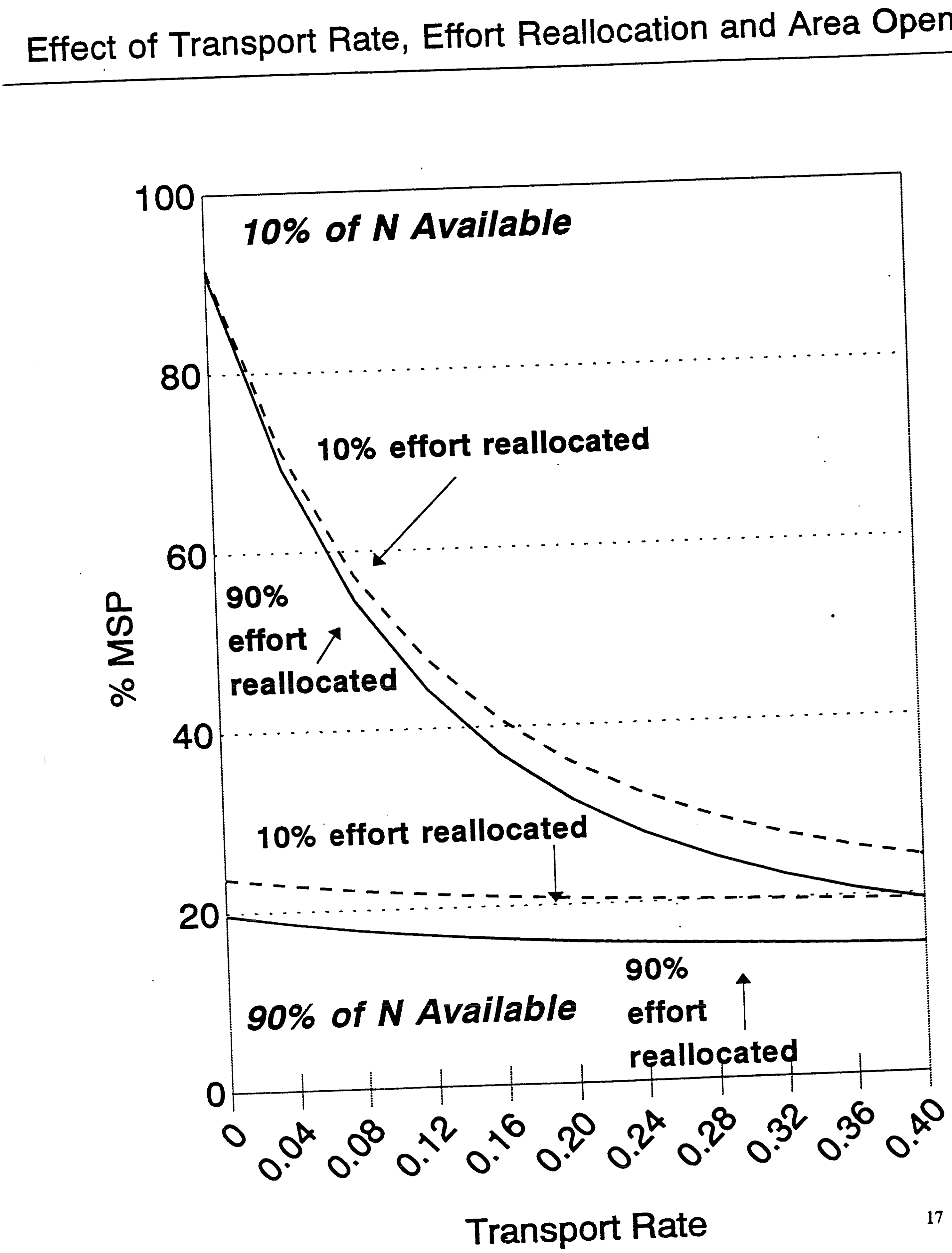
Additional information may be available from auxiliary data. One potential approach to developing information on $N_{a,y,r,q}$ is to evaluate patterns in distribution from research survey data. Estimates of $N_{a,y}$ from VPA could be modified by survey-based proportionality factors, with attendant variance estimates. However, the spatial scale of these patterns may be coarser than the spatial scale at which the fisheries operate potentially leading to highly variable estimates of local q. Alternatively, VPAs can be carried out for sub-populations including mixing (e.g. Quinn *et al.*, 1990). Market sampling data may provide an additional approach: even if market sampling data were biased (e.g., market sampling data reflect landings, collection on certain weekdays and from certain types of vessels), such data may, through suitable pictorial representation, contain information for the monitoring of fish stocks.

Visual inspection of graphical displays can reveal discernible patterns in fish stocks due to space and time. While this allows monitoring of effects, it does not by itself permit prediction of the effects on fish stocks of changes in space and time. Inference about changes in management strategy are difficult to quantify in an objective way. For example, mismatches among overlays of survey, landings and effort data may indicate the need to identify additional mechanisms operating in the fishery such as traditional behaviour patterns or economic factors. Failures to identify these mechanisms may lead to differences in predicted versus observed management effects in the fishery.

Model building is difficult at the best of times. All assumptions of the modelling process must be checked and their validity assessed prior to, during and following detailed analyses. Otherwise, models may be produced that are appealing because they are suggestive of an intuitive interpretation but violate the assumptions under which they were built. A systematic review of levels of resolution associated with each component assessment data set will reveal which model formulations will be feasible to describe historical patterns. Simple equilibrium models which incorporate spatial and behavioural components as control variables provide a strategic perspective. Dynamic models which incorporate detailed spatial dynamics in the model structure provide insight on alternative mechanisms to explain observed effects of management measures. The two approaches are complementary in a system of planning and evaluation of management strategies and tactics.

For some fisheries, data by vessel trip may be too fine-grained, and aggregated data like that in the STCF database might be used, for example in the context of an area-based VPA. The STCF database is a potentially

Figure 2.2.5 Example of spawning stock biomass per recruit obtained for southern New England yellowtail flounder as a function of alternative levels of stock protection by area closure, proportion of effort reallocated out of a proposed closed area and into the open area, and transport rate (Rago, pers. comm.).



valuable data source. Pictorial descriptions derived from the database would provide the foundation for the formulation of quantitative models. It is difficult to define a priori specific analyses without investigating the characteristics of the data directly, e.g., quality and inherent variability in the data. This would include the evaluation of whether the current resolution of reported data (e.g., effort measures) was appropriate for the models being contemplated. It would also indicate whether changes in the type of data collected and/or reported would be appropriate, either retrospectively or in the future. At present, the database contains information only as far as 1991. Recommendations to update the database would be contingent on reanalysis.

Although effects of historically implemented management regimes on fishery systems are confounded with other aspects of system behaviour (without carefully designed experiments), it could be of interest to evaluate trends in fishery system reactions to management regimes. This hindsight approach should allow us to design programs for monitoring impacts of management regimes in the future, as well as to more realistically quantify management implementation error in simulations. In light of the importance of the STCF database, access to and reanalyses of this database is recommended. Again, reanalysis would indicate the scale at which to collect data for the evaluation of conservation tactics.

Complete historical data at high levels of resolution will not necessarily ensure accurate predictions of the future, especially in the face of changing management regimes.

3 THE EVALUATION OF MANAGEMENT MEASURES

3.1 Introduction

This section contains a number of examples of evaluation frameworks and specific implementations. The majority of the approaches are based on the scenario modelling described in Anon (1994a) and in section 1.4. Also included are examples of assessing the trade-off between short- and long-term catch variability using different controllers and an analytic solution to finding an objectively risk-averse long-term management strategy that takes account of both measurement and process error.

3.2 Example Frameworks and Applications for the Evaluation of Management Measures

3.2.1 Evaluating management measures for two populations with mixing

WD1 presents a simple framework for evaluating management measures for two populations with mixing between them. The framework used is that shown in Figure 1.4. The underlying system consists of two age-structured populations with fixed rates of movement between them, governed by stock-specific stock recruitment relationships. In the framework, this system is observed every two years in terms of catch and relative abundance data (with noise) and an age-structured production model is used to assess the system during the management period. During the assessment step, mixing is either ignored or accounted for with or without bias. TACs are then set according to a control law (F_{msy} or $F_{0.1}$) and the underlying system is updated.

The limited set of simulations conducted indicates that including mixing rates in the assessment can sometimes improve management performance. However, this is not always the case and results are evidently case-specific. WD1 also showed that implementation tactics can be extremely important for populations with mixing. The inability to fully implement management measures for one of the populations can lead to deterioration in management performance for the other.

3.2.2 Multi-annual strategies and MBAL

The European Commission has recognised the problems posed in setting appropriate annual TACs for single species. It advocated the introduction of multi-annual, and perhaps multi-species TACs as an aid to overcoming a cumbersome procedure; these changes would also be expected to give a greater degree of stability to the industry. Multi-annual TACs and multi-annual decision making were both reviewed by the EC's Scientific and Technical Committee for Fisheries and the concepts and mechanisms of application were further elaborated by the European Commission, resulting in some specific proposals for action by the Council of Ministers. From the scientific advisory side, in 1991, ACFM added to its traditional advice, based on biologically and economically sensible fishing reference points (F_{max}, $F_{0.1}$ etc.), the concept of a "Minimum Biologically Acceptable Level" or MBAL. This was defined as a stock size below which data indicate an increased probability of reduced recruitment. If a stock is below MBAL, or destined to go below MBAL, ACFM will, according to these new guidelines, strongly recommend management action to safeguard the stock. If a stock is considered safe then no recommendation is made and catch options are presented. The idea is that, if a stock is deemed to be safe from recruitment failure, then managers can decide on catch levels, taking account of socio-economic considerations not available to ACFM.

Some implications for management by the combination multi-annual TACs with MBALs (see WD 7), were examined in 200 simulations, each of 200 years, based upon the biology and exploitation of cod in the Celtic Sea (ICES divisions VIIf+g). Random variations in recruitment and estimated numbers at age were incorporated into the model. A constant catch was fixed for a number of years (varying from 1 to 10) and the limitations on the size and duration of such catches determined in order to maintain SSB above a specific threshold level. Figs. 3.2.2.1 and 3.2.2.2 illustrate the main finding of these simulations. A failure parameter, a_1 , was accumulated if the SSB fell below MBAL at least once, at any time during a 200 year simulation. It is shown that, in order to maintain the SSB above MBAL, target stock sizes will have to be considerably greater than the MBAL, e.g., eight times even when the catch is held constant for one year only (Fig. 3.2.2.1). A second failure parameter, a_2 , was accumulated each time the SSB fell below MBAL in 200 simulations of 200 years. Although this statistic allows more flexibility in the management, at the 5% probability threshold and when the catch is held constant for one year only, target stock sizes, still have to be three times above MBAL (Fig. 3.2.2.1). Consequently, multiannual strategies that have MBAL as an objective, are ineffective for the management of stocks such as Celtic Sea cod, which are subject to huge fluctuations in their biological parameter levels. This kind of strategy might, however, be applied to other stocks such as Celtic Sea flatfish, as their biological parameters are more constant.

3.2.3 "FiFi" - a spreadsheet based framework for evaluating management under uncertainty

WD 3 describes a framework developed at Lowestoft to evaluate management measures as advocated by the Working Group in 1994 (see Anon, 1994a; Fig. 1.4). The framework has come to be known as the "Fisheries Fire Engine" or "FiFi" for short. It has been developed using the commercially available spreadsheet Excel and a commercially available add-in, Crystal Ball or @RISK, that currently allows simulation models to be run in spreadsheets under Microsoft Windows. This approach has been made possible by implementing ICES assessment working group methods as Excel functions written in C++. A modular approach means that individual components can easily be changed to allow different management procedures to be simulated and compared against alternatives (either by changing the assessment procedure or the controls). The use of Excel simplifies development whilst compiled functions reduce execution time and allow complex models and methods to be implemented.

The methods written as C++ functions are XSA (see, e.g., Darby and Flatman,1994), the prediction program currently used by the North Sea Demersal Working Group (Anon, 1995) and a non-linear least squares estimation routine (Levenberg-Marquardt; modified from Press *et al*, 1992) to calculate the F vectors by fleet required to take a given quota. The remaining calculations are implemented as equations in spreadsheet cells in the normal way.

The model is contained in two spreadsheets corresponding to the underlying system and the perceived system, with a third spreadsheet containing all the parameter estimates and XSA controls.

Multi-species effects (technical and/or biological) could be implemented in the same form. Fleet-species interactions could be modelled by defining a catchabilty matrix over species, age and quarter for each fleet. Biological interactions, essentially estimates of M-at-age by species as a function of abundance of other species, could in principle be modelled using either the Shepherd steady-state model (Shepherd, 1984) or MSFOR (Sparre, 1984), implemented as functions within the FiFi DLL. The underlying age distributions could vary spatially and natural mortality, growth and maturity could be functions of the abundance of other species as well as of age or external variables.

As the model becomes more complicated it would be preferable to implement more of it in the form of functions contained in FiFi. A set of functions will be developed that can be used in a range of Windows applications and readily used to model a variety of scenarios without extensive additional programming. In principle, this should make modelling management under uncertainty available to a wider audience.

The code used in the Excel functions can also be used to produce a stand-alone application that would be suitable for implementing a particular, or limited, range of management protocols. The form of such an application would probably be based on Visual Basic code front-ending a database and outputting graphics and summaries to Excel or SAS for presentation. Alternatively, the C++ code could be sewn together to form a stand-alone DOS or Windows application.

A demonstration of FiFi, using North Sea plaice as an example, is described in section 4.

3.2.4 Evaluation of management strategies in the Greenland shrimp fishery

WD 6 includes a bioeconomic model of the Greenland shrimp (*Pandalus borealis*) in the Davis Strait. Monte Carlo simulations are applied to estimate the long term yield as the total NPV (net present value) of the resource rent of 10 years of fishing.

The underlying system model consists of a biological model including an age-structured production model with variable recruitment, and variable growth and natural mortality rates; and an economic system model relating costs and effort.

The perceived system model consists of an age-structured single species production model with constant M, and growth assumed to follow the von Bertalanffy growth equation. To simulate a sampling procedure the number-at-age is estimated as an unbiased random variate from a normal distribution with the true number of the underlying system being the mean.

Figure 3.2.2.1 Probability p_1 of a_1-type failures plotted against Q for values of F of 1-10 years, where an a_1-type failure was identified when the SSB fell below the MBAL at least once within 200 years over 200 runs.

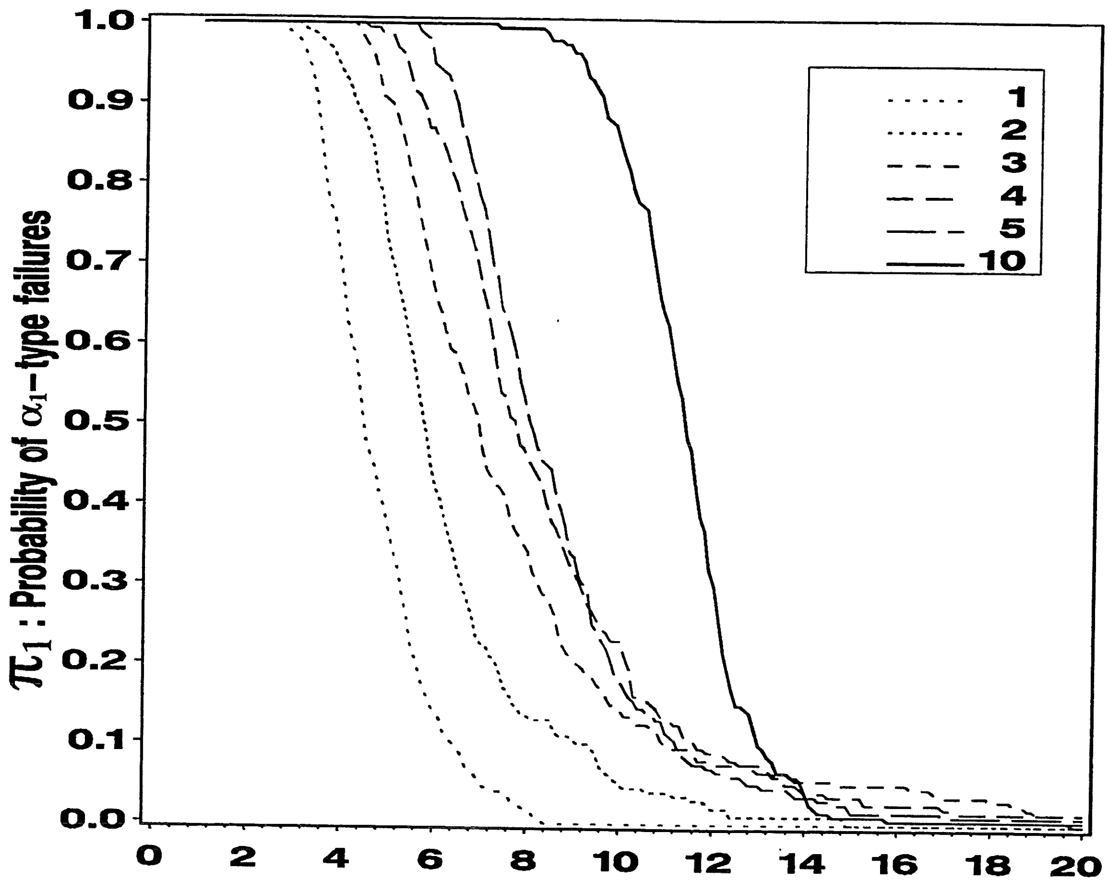
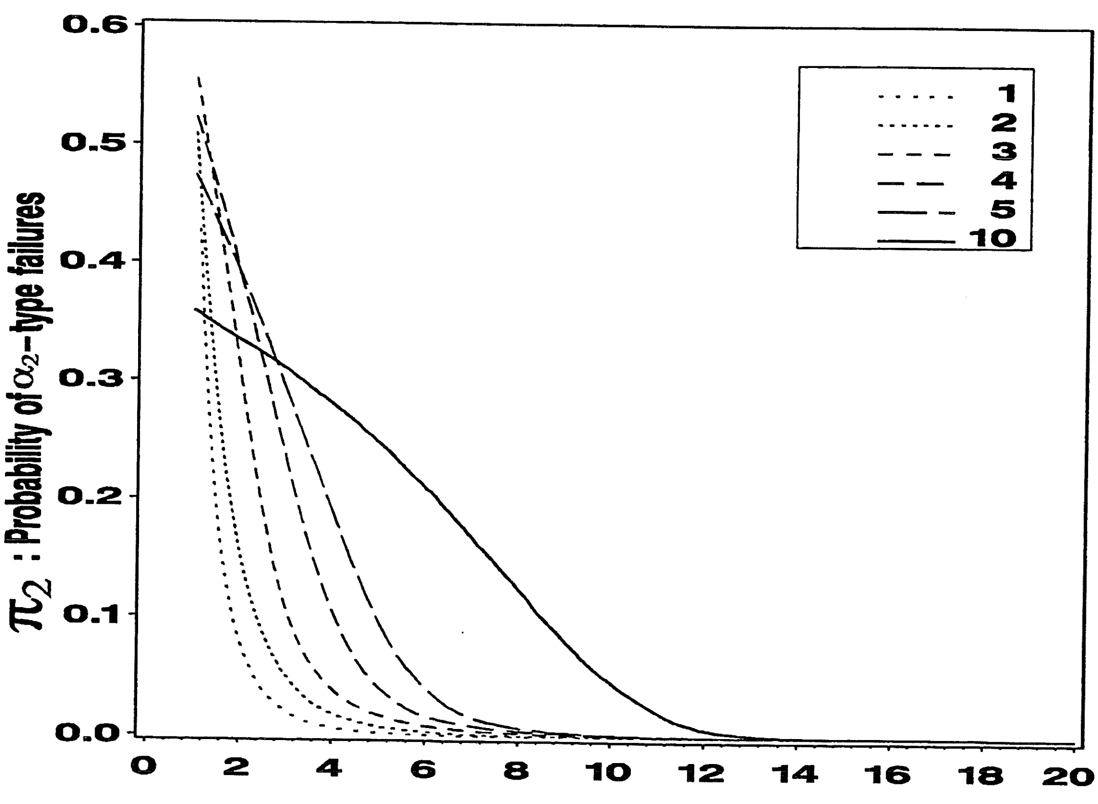


Figure 3.2.2.2 Probability p_2 of a_2-type failures plotted against Q for values of F of 1-10 years, where an a_2-type failure was identified each time the SSB fell below the MBAL at any time over 200 runs of 200 years each.



The performance of two management strategies was investigated: an effort regulating strategy and a strategy with total allowable catch limitations.

The assessment procedure was simulated by application of 3 different control laws: constant target F, Target F + MBAL and a control law also taking into account the catch level of the previous years of fishing.

The advice from the assessment procedure is used to update the underlying system directly or is passed on to a procedure simulating a decision making process. The decision making procedure gives equal weight to the biological advice and requests put forward by the fishery sector based on economic considerations.

Implementation errors are introduced to the analysis of the TAC strategies by simulating two different fleet behaviours by which highgrading of catches take place. One highgrading behaviour assumes that highgrading occurs as long as there is a positive resource rent, whereas the other highgrading behaviour assumes that fishing and highgrading takes place as long as merely the variable costs are covered.

The study suggests that the overall best performing management strategy is to regulate effort and reduce target F if SSB is below MBAL. If the management strategy is to regulate the fishery by means of TACs the control law taking the information on the catch level of the prior years of fishing should be applied.

If the biological advice is modified by a political process giving equal weight to requests put forward by the fishing fleet, the realised effort may be significantly higher than is suggested by the biological assessment. Also, if fleet behaviour includes highgrading of catches, the realised effort level will be higher than the one producing the maximum long-term resource rent. In both cases, the reduction in resource rent is mainly due to increased effort costs.

3.2.5 SCENARIO BARENTS SEA

WD 12 describes SCENARIO BARENTS SEA, a tool for evaluating fisheries management regimes. The paper is an incomplete draft of work in progress. The SCENARIO BARENTS SEA project is carried out at the Norwegian Computing Centre by statisticians. That institution is not part of the established system of fisheries science in Norway, and the project relies on input and competence from the fisheries science establishment. A group consisting of fisheries ecologists (Bergen), fisheries economics (Tromsø) and management (the ministry) meet with the project group twice a year. This way of organising the work in the project has certain advantages. It provides a neutral meeting ground for the three established traditions, where the dialogue is directed towards establishing a common understanding of the main features of the system. It does also, at least in theory, make the three groups come forward with explicit formulation of the management strategies that are currently used and other strategies that are considered, how assessment actually is carried out, and what should be the objectives for short and long term management of fisheries.

The project aims at establishing a framework for the evaluation of fisheries management regimes in the Barents Sea (and later on also in the Norwegian Sea, with marine mammals included) for cod, capelin and herring. The model is programmed in FORTRAN, and it consists of a simulation model that aims at giving a parsimonious environment for specifying scenarios both for the ecological system and the human system (fisheries and monitoring activities). It furthermore provides for various management regimes based on VPA-calculations for cod and herring, and a residual regime of capelin management (CapTool). The model is age and length specific, but not regional. Maturation, predation, growth etc. are estimated afresh, but can be altered in specific scenarios. The fishery is multi-fleet, with room for overfishing, discarding etc. and management is done by setting TACs.

Two different experiments are distinguished. The probing scenarios are used to investigate specific questions in isolation, to improve our understanding of various strategies and the model. The main emphasis is, however, on uncertainty scenarios. These are laid out with the aim of spanning a region of plausible states of nature (including the human element). Here, the simulation results, filtered by performance statistics, provide a basis for investigating the properties of the management regimes under study with respect to robustness and efficiency in a comparative setting. Uncertainties that are relatively well understood statistically, like that in the recruitment function, are modelled stochastically. Uncertainty that is less well understood, like that in our knowledge, is accounted for in the experimentation by specifying the set of uncertainty scenarios. These aim at spanning what the informed regard as the plausible region in state space. The uncertainty scenarios are specified by a set of categorical factors. There are often many factors, and the set of uncertainty scenarios is determined by combinatorial experimental design, to maximise information at limited computational costs. The strategy is to simulate a limited number of replicates (3–10) for each scenario, rather than to have many scenarios. The output from an uncertainty experiment is analysed by regression techniques. The robustness properties relate to how well the management regime handles scenarios that are particularly difficult, while efficiency relates to performance in the centre of the plausibility region.

At present, a set of uncertainty scenarios are being run; reports on these could not be given. Results of a few probing scenarios were reported.

3.2.6 Testing management procedures for oceanic redfish

WD 11 considers initial steps towards developing and testing management procedures for oceanic redfish and points to the need of starting a process like the one used by the IWC when developing and testing its Revised Management Procedure (RMP). This process will be an interaction between testing and modifying potential procedures to deal with a wide variety of robustness tests. A number of procedures should be considered, tested and their performances evaluated and compared.

The stock of oceanic redfish is a good candidate stock when investigating the feasibility of using management procedures based on fixed and well specified feedback rules and to construct testing schemes whereby various such procedures can be tested and their performance evaluated and compared by computer simulation. Fairly reliable biomass estimates can be obtained at regular intervals by hydroacoustic methods. These biomass estimates will form the input into the management procedure. Catches have only been taken for a short time and have been low compared to stock size; it is therefore possible to have a fair idea of the initial biomass and hence of carrying capacity. All this should make it relatively easy to devise simple management rules and to test them by simulation.

One possible procedure for managing this stock is described in WD 11. This procedure is of an empirical nature. The estimated relative rate of change in stock biomass (calculated from a time series of biomass estimates) is used to modify TACs by a prescribed feedback rule. In addition, the TAC is increased or decreased by a specified percentage at regular intervals, depending on whether the biomass estimate is above or below a prescribed target level. A procedure similar to this one was originally constructed to manage the harvest from whale stocks and underwent thorough testing by the IWC. Results from a few preliminary tests (all deterministic) are given. Further testing will be carried out in the near future. These tests should, for example, cover different population dynamics, effects of different types of errors (both process errors and observation error), etc.

The possible usefulness of generic testing and evaluating schemes like FiFi (see section 3.2.3) should be explained. It is likely to be suitable for initial evaluation, but eventually, case specific tests should be carried out before a procedure is implemented.

3.2.7 Assessing groundfish resources of the Celtic Sea VIIF+G with several métiers: a multiannual approach

Traditional advice provided by ACFM is based upon single-species TACs, calculated on a year-to-year basis, and various technical conservation measures. TACs are calculated every year and may be subject to adjustments several times a year. Annual variations in TACs make it hard for the fishing industry to make long-term plans. The European Commission has recognised the importance of this and has advocated the introduction of multi-annual TACs. Enforcing multi-annual strategies, i.e., setting a stable level of catches or fishing effort in advance for a given period, would allow the fishing industry and fishermen to make better plans for the future. TAEs (Total Allowable fishing Effort) have already been recommended by the European Commission. As an alternative to TACs, TAEs provide direct control of fishing mortality.

Traditional objectives in fisheries management, based upon the short- or long- term yield per recruit versus fishing mortality, generally aim to target one of the reference points F_{max}, $F_{0.1}$, F_{SQ} or F_{med}. Composite objectives based on varied compromises between F_{max} and F_{SQ} have been sought, and properties of each have been compared (Pelletier, 1991; Pelletier and Laurec, 1992). A similar approach is developed in WD8. However, in order to take into account the multifleet aspect of the fishery, levels of exploitation rate by métier, i.e. levels of fishing effort by métiers relative to the reference year (1991), are separately targeted.

The issue of technical interactions, involving several fleets and several species, is illustrated by reference to the demersal fisheries of the Celtic Sea. In particular, the resources of area VIIf+g, essentially made up five species (cod, whiting, sole, plaice and *Nephrops*), are mainly harvested by three shallow-water métiers: *Nephrops* otter-trawlers (métier 8), non-*Nephrops* otter-trawlers (métier 5) and beam-trawlers (métier 6).

The first task is the definition of an objective that this fishery system could achieve. WD8 focuses on maximising the yields in weight for each species. Ideally, fisheries operate at the Fmax value for each of the five species with the three métiers as variables. In general this is not feasible; it is, however, possible to try to find the optimal allocation of effort amongst métiers, relative to their reference value. A composite objective is sought, which would be a compromise between maximising yields (for all métiers) and leaving the exploitation rate at the value it had in the year of the assessment (for all métiers).

Multi-annual TAEs have been set every 5 years in order to (i) target a composite level of exploitation rate (related to the optimal level and to the *status quo* by parameter g); (ii) to minimise interannual variability in terms of fishing effort; and (iii) to minimise interannual variability in terms of catches. The parameter g has been varied between 0 and 1, and the relative performance of the assessments evaluated by considering several short- and long- term criteria; i.e. total values of landings, an index of effort variability and an index of catch variability.

An optimum is achieved when g is set to 0.5. This represents a minimisation of the response surface for short- and long- term interannual catch variability and short- and long- term value of total landings (Fig. 3.2.7.1a-f). This strategy does not lead to very high levels of inter-annual variation in exploitation rates. Time series of averaged (over 100 simulations) exploitation rate and catches by métiers for this strategy are shown in Fig. 3.2.7.2a-f. Exploitation rates by métiers decrease slowly towards their respective target value, which they have not completely reached after 20 years.

The relative change in the time series of catches by species and métiers is most spectacular for cod. Catches from *Nephrops* otter-trawlers exceed catches from non- *Nephrops* otter-trawlers after 10 years. Conversely, catches of *Nephrops* are almost constant. Whether decreasing or increasing, average catches and exploitation rates reach their equilibrium gradually and do not always achieve it within 20 years.

The g=0.5 strategy can be considered as successful. However, other strategies related to the parameters held constant have yet to be explored. Eventually, this study aims to explore new objectives, new strategies and new tools for fishery management. Multi-annual strategies, composite objectives and TAEs will be modelled for the Celtic Sea multi-fleet multi-species system.

3.2.8 Risk-averse implementation of a production model

WD2 presented a model based on a stochastic differential equation which is estimated through a Kalman filter and applied in a decision-theoretic context. The stochastic differential equation serves to introduce process error into the underlying surplus production model, which in this case is based on the Gompertz growth formula. One advantage of the stochastic differential equation approach in general is that it enables the form of the process error to be derived rather than assumed. For the particular model considered in WD2, the stochastic differential equation approach has the further advantage that the process error can be derived in closed form. This enables the probability density function (pdf) of stock size to be written for any point in time, past or future.

The Kalman filter can be thought of as a Bayesian updating of the estimated pdf of stock size. With the addition of each new datum in the time series, the Kalman filter applies Bayes' rule to the most recent pdf (the prior distribution) to estimate an improved (posterior) pdf. The prior distributions can be used to predict each succeeding datum, and the series of discrepancies between the predicted and observed values defines a likelihood in terms of the four basic parameters of the model (the Gompertz growth rate, the carrying capacity, the fishing mortality rate, and a parameter that scales the process error variance). The model parameters can be estimated by maximizing the likelihood, maximizing the posterior density, or taking the mean of the posterior distribution.

The posterior distribution of the Gompertz growth parameter also plays a role in defining a harvest rate that is formally risk averse (in the decision-theoretic sense). One way to implement a risk-averse management objective is to maximize the expectation of the logarithm of stationary yield (the "MELSY" harvest strategy). In the model described in WD2, the MELSY harvest rate is simply the harmonic mean of the pdf of the Gompertz growth parameter. A risk-averse catch recommendation can then be obtained by applying the MELSY harvest rate to the projected mean stock size at time *t* and integrating over time between the endpoints of the relevant harvest period.

Finally, WD2 shows how decision theory can also be used to address the possibility of parametric nonstationarity, that is, the occurrence of a discrete change in parameter values at some (unknown) point in the time series. Instead of framing the question in terms of which point is the *most likely* location of the *true* change point, a decision-theoretic approach would frame the question in terms of which *choice* of change points corresponds to the lowest risk (i.e., lowest expected loss of utility). In order to show how such an approach might work in principle, a hypothetical example is constructed in which an analytic solution is obtained. Presumably, of course, an empirical example would require a numerical solution.

3.2.9 Experimental design

To investigate the comparative merits of a set of management strategies with respect to robustness and efficiency, a number of uncertainty scenarios need to be run. Each of these scenarios should be regarded as an experimental run, with the experimental factors set according to the specification of that particular scenario. In medical, agricultural and other experimental fields, the statistical technique of experimental design has been used for many years. The main idea is that with several factors, the number of possible combination settings gets prohibitively large. To reduce this number, one uses combinatorial methods like fractional factorial designs (see, e.g., Box, Hunter and Hunter, 1978) to set up a balanced set of designs. By using such designs, the information content in the results is optimised, given the experimental effort. An inferior design, that unfortunately still prevailing in the experimental field of computer based model experimentation, i.e., scenario simulation, is to vary one factor at a time, with all other factors set at a reference level. Not only will this "one factor at a time" design prevent the estimation of interaction effects, but it also provides estimates of main effects that are inferior to estimates obtained from properly designed estimates of comparable experimental effort.

The design chosen does determine the design matrix, in the sense of linear models. This is not a coincidence. Linear models (and GLMs) are to a large degree devel-

Figure 3.2.7.1 a: Long term index of effort variability plotted against short term index of effort variability for values of g scaled between 0 and 1 with steps of 0.1. b: Long term index of catches variability plotted against short term index of catches variability for values of g scaled between 0 and 1 with steps of 0.1. c: Long term landings values (tonnes) plotted against short term landings values (tonnes) for values of g scaled between 0 and 1 with steps of 0.1. d: Short term index of effort variability plotted against short term landings values (tonnes) for values of g scaled between 0 and 1 with steps of 0.1. e: Short term index of effort variability plotted against short term index of catches variability for values of g scaled between 0 and 1 with steps of 0.1 f: Long term index of catches variability plotted against long term landings values (tonnes) for values of g scaled between 0 and 1 with steps of 0.1.

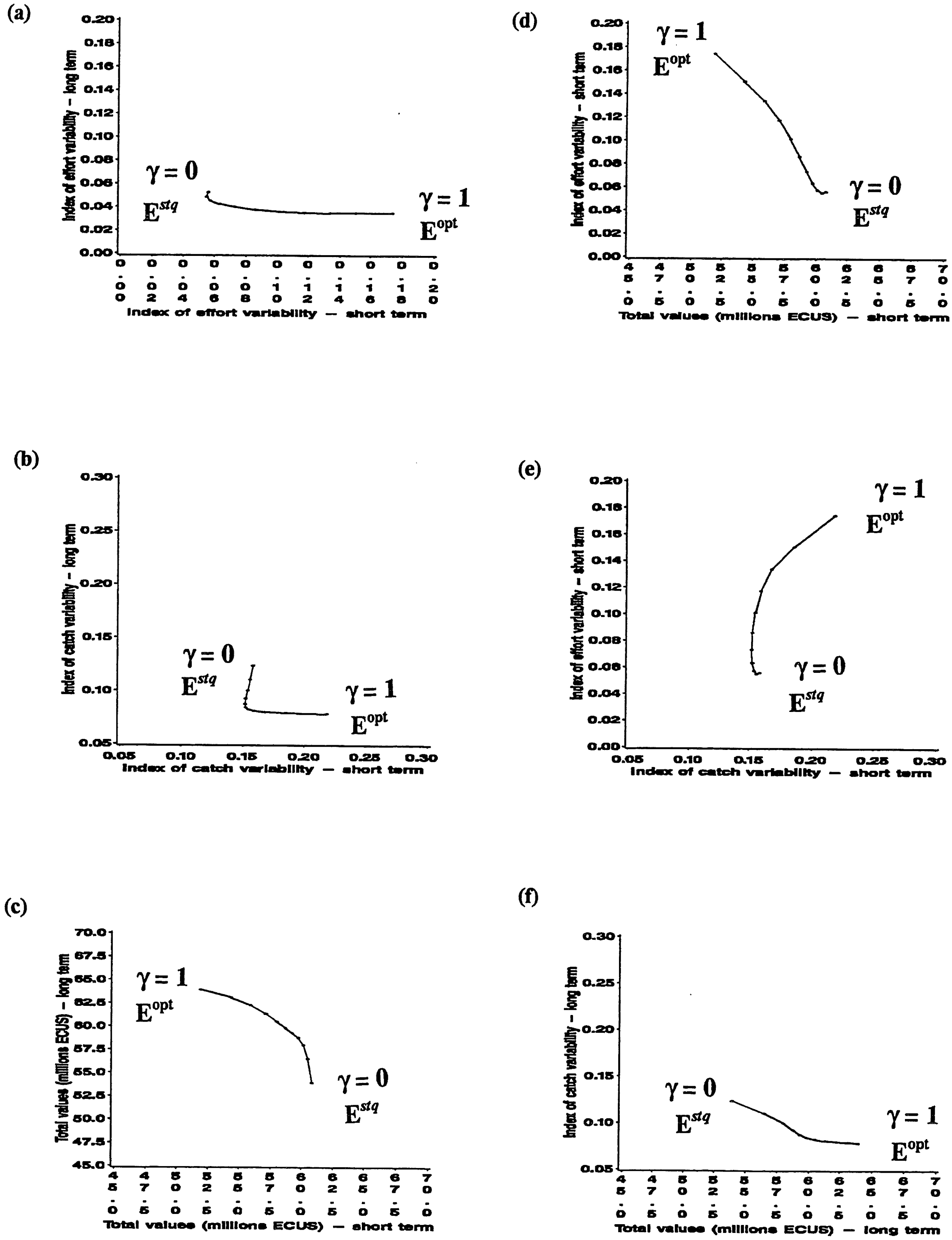
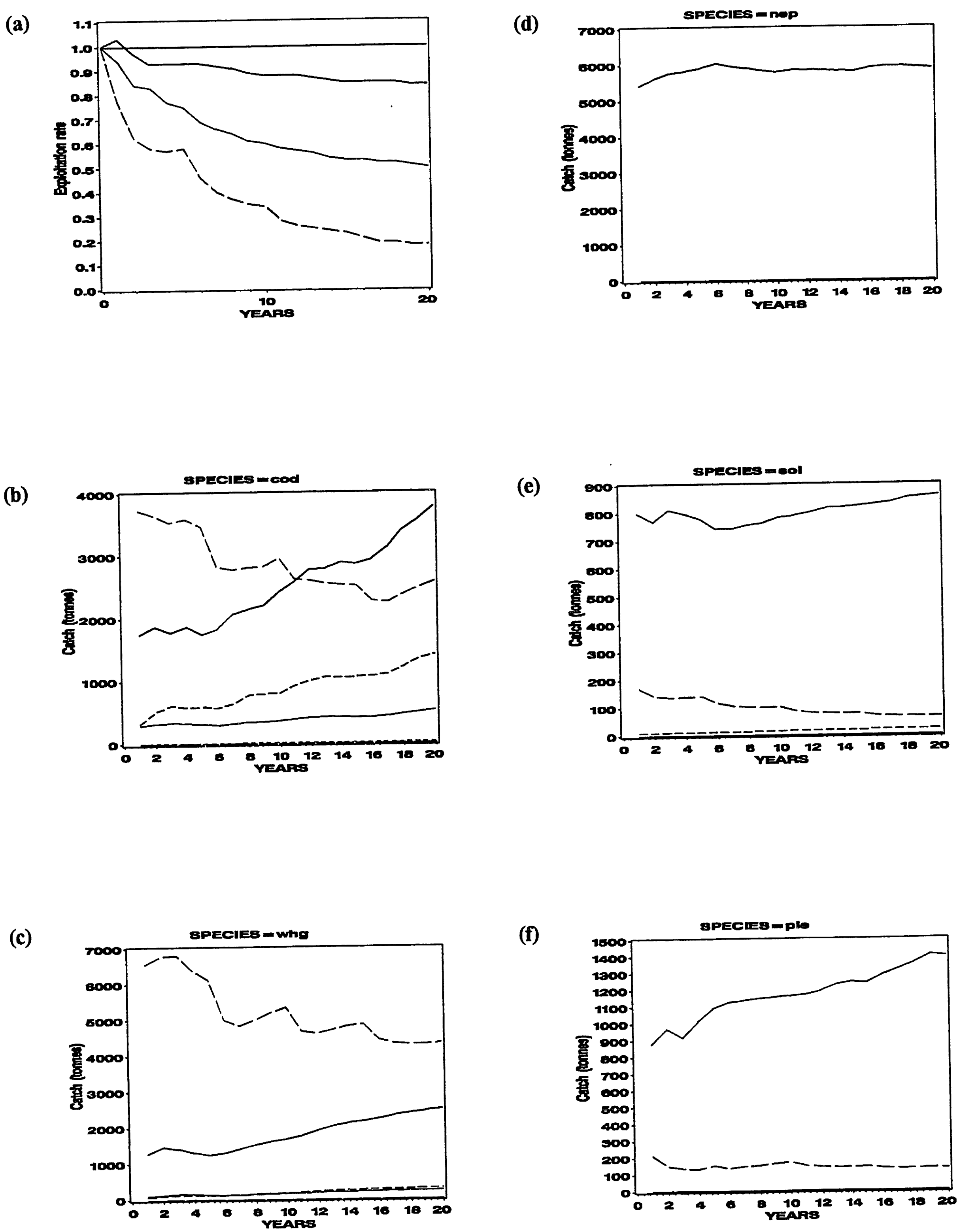


Figure 3.2.7.2 a: Exploitation rates by metier plotted over 20 years with g=0.5. b: Catches of cod (tonnes) by metier plotted over 20 years with g=0.5. c: Catches of whiting (tonnes) by metier plotted over 20 years with g=0.5. d: Catches of *nephrops* (tonnes) by metier plotted over 20 years with g=0.5. e: Catches of sole (tonnes) by metier plotted over 20 years with g=0.5. f: Catches of plaice (tonnes) by metier plotted over 20 years with g=0.5.



oped for the purpose of analysing experimental data. Regression techniques are thus the appropriate methodology for extracting information from the often large amount of output from a simulation experiment. The regression technique should include both the standard estimation and model fitting techniques, residual analysis and other diagnostic techniques.

It is not a prerequisite that the simulation model is stochastic. Designed experiments and analysis of the resulting output by regression, is just as efficient and valid for deterministic simulations. The residual errors with respect to a descriptive linear model will then consist of numerical effects and interaction effects that are not included in the model because they are regarded as being of less concern.

3.3 Discussion and Conclusions

All of the issues involved in the processes of stock assessment and fishery management are, of course, also relevant to the process of management measure evaluation. A number of these are discussed below.

3.3.1 Deterministic and Stochastic models

Several types of error were defined in last year's Working Group report, including process error, measurement (or observation) error, estimation error, implementation error, and model (or structural) error.

Of the above, process and measurement error are the two types considered most commonly in fishery models. By way of definition, a "stochastic" model will be distinguished here from a "deterministic" model by the presence (in the former) of explicit allowance for either process error or measurement error (this definition may be slightly different from that contained in some textbooks, but seemed useful for the purposes of the Working Group). Both deterministic and stochastic models can be useful tools in management measure evaluation. However, the particular uses to which they are put may make one type of model more appropriate than the other. For example, deterministic models may be especially useful in developing theory or strategies, in making very short-term projections, or in the formulation of stochastic models (as a limiting case or as a "skeletal" version). On the other hand, if parameters are to be estimated from data or non-hypothetical projections beyond the very short term are required, it is difficult to imagine an example (in fisheries assessment and management, at least), where a stochastic model would not be preferred. Most of the models described in this report are stochastic (the exceptions being those described in sections 2.2.4 and 3.2.6), generally incorporating both observation and process error.

Turning to estimation error, it may be noted that the model described in section 3.2.8 accounts for such error explicitly in defining an optimal harvest rate. Related to estimation error is the issue of parametric non-stationarity. An implicit assumption of most fishery models is that parameters remain constant with respect to time, or that they fluctuate randomly within a time-invariant distribution. Of course, it is possible for parameter values (or their distributions) to change systematically over time, for example through changes in the environment or through natural selection. The model described in section 3.2.8 considers how discrete changes in parameter values might be detected and estimated, while the model described in section 3.2.3 is flexible enough that either discrete or continuous changes in parameter values could potentially be examined.

The "FiFi" framework described in section 3.2.3 is a good example of a model that incorporates implementation error (see also Figure 1.4).

None of the working documents presented to this year's meeting of the Working group explicitly considered structural error. Of course, it is usually possible to imagine how differences in qualitative model characteristics could be described quantitatively (through the specification of additional parameters), so the question of whether a particular model allows for structural error may not always be answerable unambiguously.

3.3.2 Classical and Bayesian approaches

Most fishery models (including most of those discussed in this report) are developed within the framework of classical, or "frequentist" statistics. Inferences regarding parameter values are often made in the form of a classical hypothesis test, and stochastic results are often presented in the form of cumulative distribution functions (which give the probability of a random variable taking on a value less than or equal to X). Such models are based on a strong and successful historical tradition, and provide tools which can be fairly readily generalised to other approaches.

The introduction of Bayesian approaches provides another suite of options for the evaluation of management measures. One useful contrast to consider in this context is the difference between those ("frequentist") approaches to risk that limit their consideration to probabilities of alternative outcomes (e.g., different stock sizes at some future time) and those ("decision-theoretic") approaches that include both the probabilities of alternative outcomes and the utility associated with those outcomes (e.g., the yields associated with different stock sizes at some future time). Section 3.2.3 of this report describes a framework which tends to fall in the frequentist category, while the models described in sections 3.2.4 and 3.2.8 fall in the decision-theoretic category. It should be noted, of course, that frequentist analyses can often be generalized to the decision-theoretic level, and that decision-theoretic analyses can typically be scaled back to the frequentist level (see, e.g., section 4.1.2).

One weakness of the frequentist approach is that it is difficult to assign normative significance to the results. The situation is parallel, in some sense, to the classical hypothesis test. For example, the frequentist approach is sometimes used (e.g., section 4.1.2) to compute the fishing mortality rate at which the probability of stock biomass falling below some specified level (e.g., MBAL) reaches a critical value (e.g., 5%). However, this approach begs the question: why were MBAL and the critical probability levels set at their respective values, and what is suboptimal about allowing these critical values to be transgressed?

On the other hand, while the results stemming from a decision-theoretic approach have clear normative significance, the approach suffers from a very practical problem: its results are dependent on a definition of utility, which must be specified by someone. Scientists and managers alike may, to varying degrees, be reluctant to specify such a definition. One way in which the exercise might be made more comfortable for scientists would be to specify a number of utility definitions, leaving the final choice up to someone else.

Amongst the issues to consider in formulating a definition of utility are the time horizon to use and the degree of relative risk aversion to impose. For example, in considering the issue of time horizon, the model described in section 3.2.2 examined two alternatives, one in which the time horizon spanned years 1–5 into the future, and the other in which it spanned years 10–20 into the future. The analysis described in Section 3.2.7 explicitly considers the problem of estimating an optimal transition to a desirable equilibrium. At the far end of the spectrum, the model described in section 3.2.8 took the limit of the transitional probability distribution of yield as time approached infinity. In the area of relative risk aversion, the model described in section 3.2.8 again fell at the end of the spectrum, choosing a logarithmic utility function. The logarithmic utility function has a relative risk aversion of unity, as contrasted with the risk-neutral linear utility function, which has a relative risk aversion of zero.

In addition to suggesting the decision-theoretic approach to risk analysis, the Bayesian framework also provides a means of formally utilizing information not contained in the data typically used in a single stock assessment. For example, the fishing mortality rate associated with maximum sustainable yield (F_{MSY}) is typically difficult to estimate with much precision in a single stock assessment. However, before the assessment is even begun, it is likely that the analysis has a fairly well informed judgment as to what a reasonable value might be. For example, most assessment scientists would likely agree that F_{MSY} is probably not ten times greater than the natural mortality rate. When cast in the form of a prior distribution, Bayes' rule allows such prior knowledge to be incorporated into the analysis in a rigorous way. On the other hand, if such prior knowledge is largely or entirely subjective in nature, translating such knowledge into a prior distribution is not a trivial exercise, and scientists may balk at specifying parameter values that are not derived objectively. In any case, if the analyst is determined not to let prior knowledge influence the outcome of the analysis, or if there truly is no prior knowledge regarding the value of a particular parameter, a "noninformative" prior can often be identified.

3.3.3 Simple and complex models

The question of how much complexity to build into a model is a basic problem in evaluating fishery management measures. All else being equal, models should incorporate whatever features are necessary to symbolize adequately the phenomena being modelled. Thus, some modellers have a tendency to build more and more complexity into their models, for fear of omitting significant features. However, unlike some problems encountered in engineering, for example, the "correct" model of stock (or more generally, fishery) dynamics is never obvious, and attempts to frame the modelling exercise in terms of a search for such a model are almost certainly misplaced. Rather, any given model should be viewed simply as being more or less appropriate than another, based on the objectives of the modelling exercise, data availability, and other constraints (e.g., the time available for analysis).

The structure of any fishery model can always be made more "realistic" by incorporating additional behaviours, variables, or parameters that are perceived to symbolize some aspect of real fishery systems. In other words, when modeling real fishery systems, the analyst can be certain that the model structure chosen will be less complex than the system being modeled. However, even if this were not the case (i.e., even if the "correct" model structure were known), a deliberately simplified model could potentially outperform the correct one. The model described in section 3.2.1 presents an example of this, based on a hypothetical system of known structure. If a simplified model can sometimes outperform a more complex model even when the latter is structurally correct, the same result must certainly be possible when the more complex model is already far less complex than the real system being modeled.

Overparametrisation is another way to think about this problem. Just because additional parameters *can* be estimated does not mean that they *should* be, as the estimation error of all parameters may tend to increase with the number of parameters. Perhaps worse yet, a model may be overparametrised to such an extent that unique parameter estimates cannot be obtained at all. The analyses described in section 2.2.1 explore potential dangers of this sort.

Several factors relating to relative model complexity are discussed in section 2 of this report, for example spatial structure and fleet structure. Two additional factors are discussed below.

3.3.3.1 Age (or size) structure

One specific example of a complicating feature that a modeller may or may not choose to include is age (or size) structure. Certainly, real fish stocks exhibit both age and size structure (and genetic structure, sex structure, etc.). The appropriate question, though, is whether inclusion of such structure in a particular model is helpful, given the objectives of the modeling exercise, data availability, and other constraints. Of the models described in this report, only those discussed in section 3.2.6 and 3.2.8 do not include either age or size structure, and none of the models formally evaluate the tradeoffs between including and ignoring age or size structure.

3.3.3.2 Species interactions

A second specific example of a complicating feature that a modeler may or may not choose to include is species interactions. Again, it is clear that species in real fishery systems interact with one another. However, inclusion of this feature in the types of models typically used to evaluate fishery management measures is difficult. Of the frameworks discussed in this report, only that described in section 3.2.5 includes species interactions, and there the number of species is limited to three (section 3.2.3 also mentions species interactions as a possible future addition to the model described therein). None of the frameworks includes anything that could be called a full ecosystem model.

3.3.4 Summary

As shown in this report, there is a diversity of models that can be used to evaluate management measures. There is no consensus as to what approach is "best"; nor should there be consensus, for different approaches may be appropriate under different circumstances. Additionally, different approaches may be usefully conducted in parallel to examine the robustness of their results.

For the near future at least, the Working Group recommends that work continue among a diversity of approaches as the fledgling discipline of management measure evaluation begins to mature.

FiFi (see Section 3.2.3) demonstrates an approach to modelling and evaluating management measures using a simulation framework. FiFi is not a case specific implementation of a particular scenario, although it could be use in such a way, rather it shows that if the components of a fishery system are presented in a modular way, they may be integrated in a generic environment.

Typically, programs which implement particular assessment tools (methods) have been written as standard applications dependent on differing file input and output structures. Whilst this has allowed working groups to perform assessments, there are difficulties when trying to integrate such programs directly into a general framework to evaluate management measures.

FiFi shows how, if programs are written in a modular way and compiled as a Dynamic Link Library (DLL), routines can be called by standard Microsoft Windows applications.

The Working Group, through its activities, has identified useful components of a fishery system that can be programmed in a modular way. These could be integrated into a generic environment for comprehensive modelling and assessment.

Given the scope for complexity and the associated computational needs of management evaluation work, it would be useful to investigate criteria for determining appropriate types and levels of process, measurement, model, and implementation error for evaluation of specific problems. The scope for limiting the number of iterations and experimental runs using efficient experimental designs should be investigated.

4 EVALUATION OF MANAGEMENT MEASURES FOR NORTH SEA PLAICE

The Working Group has been asked to demonstrate a specific example of the evaluation of management measures using North Sea Plaice as an example (ToR (b)). Although data for this case study were obtained from the most recent assessment of the stock (Anon., 1995) the results described here are for illustration only. The working group is not suggesting specific advice for North Sea plaice. A more complete evaluation should be carefully planned and should, as stated in Section 5.1, examine the performance of various management measures with respect to specific objectives. The Working Group believes that such a comprehensive assessment cannot be thoroughly carried out, all the way from planning to analyses, during the span of a typical assessment meeting, e.g., one week. Much of the work should be accomplished inter-sessionally with co-operation among scientists knowledgeable about the system being modelled: its biological, economic, and behavioural components.

4.1 Management Evaluation Framework

Two approaches were used, one used an age-structured model which generated data necessary for VPA-based assessments (see WDs 3&4 and Section 3.2.3). The second used a general production framework (see WD 2 and Section 3.2.8). There are a very large number of possible management scenarios that could be tested with either of these methods. The working group focused, however, on three specific questions: the effect of alternative "true" stock-recruitment relationships, the use of a minimum biologically acceptable level (MBAL) in the control law, and the effect of observation error in the abundance indices. These questions were not examined thoroughly; the limited set of analyses here is for illustrative purposes only.

4.1.1 Age-Structured Model (FiFi)

Age-structured simulations were used to model a) the underlying system structure consisting of a plausible "true" plaice population and the commercial fishery; b) sampling procedures to obtain the necessary assessment data (catch-at-age, weight-at-age, abundance indices, etc.); c) an assessment procedure to estimate catch forecast parameters; and d) decisions on future fishery management tactics using fishery control laws (e.g., TAC's). The plaice assessment data from 1958-93 were used as the basis for the evaluation.

Process error was added to the projections of the underlying system's stock-recruitment relationship, natural mortality, and selectivities as described in WD 3. The 1993 start-of-year population was used for all forward projections. The observed catch-at-age, weight-at-age, and abundance indices-at-age were generated from the true values but with added measurement error. XSA was used for stock assessments, M was assumed to be constant across years. A basic simulation consisted of 50 replicates of a 21-year run into the future with annual assessments and feedback to the true underlying system *via* implementation of TAC's.

4.1.1.1 Evaluation

The experiment was designed to evaluate the performance of a control law with or without MBAL, the effect of alternative stock-recruitment hypotheses and the level of measurement error in the abundance indices.

The basic control law was based on the average of the perceived F_{med} and F_{SQ}. A subjectively chosen MBAL of 200,000 t was used to modify the control law in some simulations. This value was chosen in an attempt to provide some contrast in the performance of the control laws. Incorporation of MBAL into the control law was as follows: If the assessment estimated the SSB to be lower than MBAL, then the forecast catch was taken as the average of the catch at F_{med}, F_{SQ}, and the catch needed to take the population to MBAL in the next year. If the estimated population biomass was above MBAL, then the forecast F was the average of F_{med} and F_{SQ}.

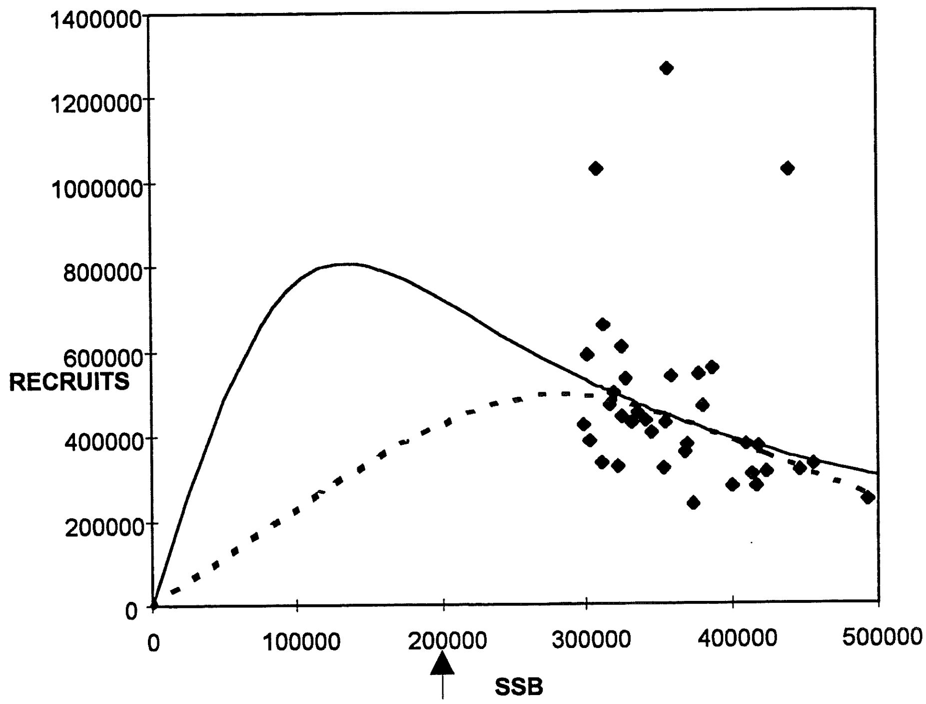


Figure 4.1.1.1 Two sample stock and recruitment relationships used in the North Sea plaice simulations. The arrow indicates the subjectively chosen MBAL for the control laws.

Two stock-recruitment relationships were used (Figure 4.1.1.1). Both were of the Ricker type although parameterised using Shepherd's equation (1982), one being more dome-shaped than the other. The scatter of historical points was to the right of the maxima in both cases.

Some runs were made with no measurement error on the assessment data. However, there was still process noise in the underlying system. These runs were contrasted with simulations where measurement error was added to the observed data on the scale of se = 0.5 (ln scale) on the indices from research vessel surveys and se = 0.1 on the commercial indices.

The number of runs required to investigate the effects of various factors one-at-a-time would be 2x2x2=8 for the MBAL/S-R/Noise levels. A fractional factorial design was therefore used to reduce the total number of simulations required, given that the simulations were computer intensive (see section 3.3). The following simulations were performed.

Run	Index Error	Stock/Recruitment	Use of MBAL
1	none	large dome	Yes
2	some	large dome	No
3	some	small dome	Yes
4	none	small dome	No

Design Matrix			
	0	1	1
	1	1	0
	1	0	1
	0	0	0

Such a design could be used to investigate main effects. Additional runs would be needed to test for interactions. For example, if the potential interaction of stock-recruitment relationship and use of MBAL in the control law was of interest. Two more runs would be needed with a design of:

Design Matrix			
	0	1	0
	0	0	1

The 50 duplicates for each simulation experiment were conducted such that they used the same sequence of random number deviates across experiments. This is known as a variance reduction technique in simulation and its objective is to reduce the number of runs required. In a thorough assessment, however, the number of runs required should be estimated empirically and would probably be greater than 50.

Results were analysed with ANOVA's. Year was included to account for trends induced by the starting conditions. However, year was not considered a controllable effect.

4.1.1.2 ANOVA results

The contrast among treatments in these illustrative runs was limited mainly because of the shape of the stock-recruitment relationship, the initial stock size, and the control laws applied. The population remained in a relatively stable region of the stock and recruitment space. The analyses are summarised in Tables 4.1.1.2.1 and 4.1.1.2.2.

For the true SSB, perceived SSB, and true catch, and once the year effects are taken into account, the shape of the stock-recruitment relationship had the greatest effect on the results. If the relationship was more dome-shaped, then the SSB's and catches were higher than if the relationship was flatter owing to the generally higher recruitment realisations (Figure 4.1.1.2).

In the case of the difference between true and perceived biomass, there was a relatively strong year effect, where the differences were the highest in the early years of the simulation but were then dampened out as more years were added to the series. This resulted from an apparent trend in catchabilities in the original assessment data as illustrated in WD3. The new (projected) data generated in the simulation did not have such a bias and the effect of the historical assessment data on the tuning decreased as new data were added in the simulations. After the year effect, the quality of assessment data had the greatest effect on the differences.

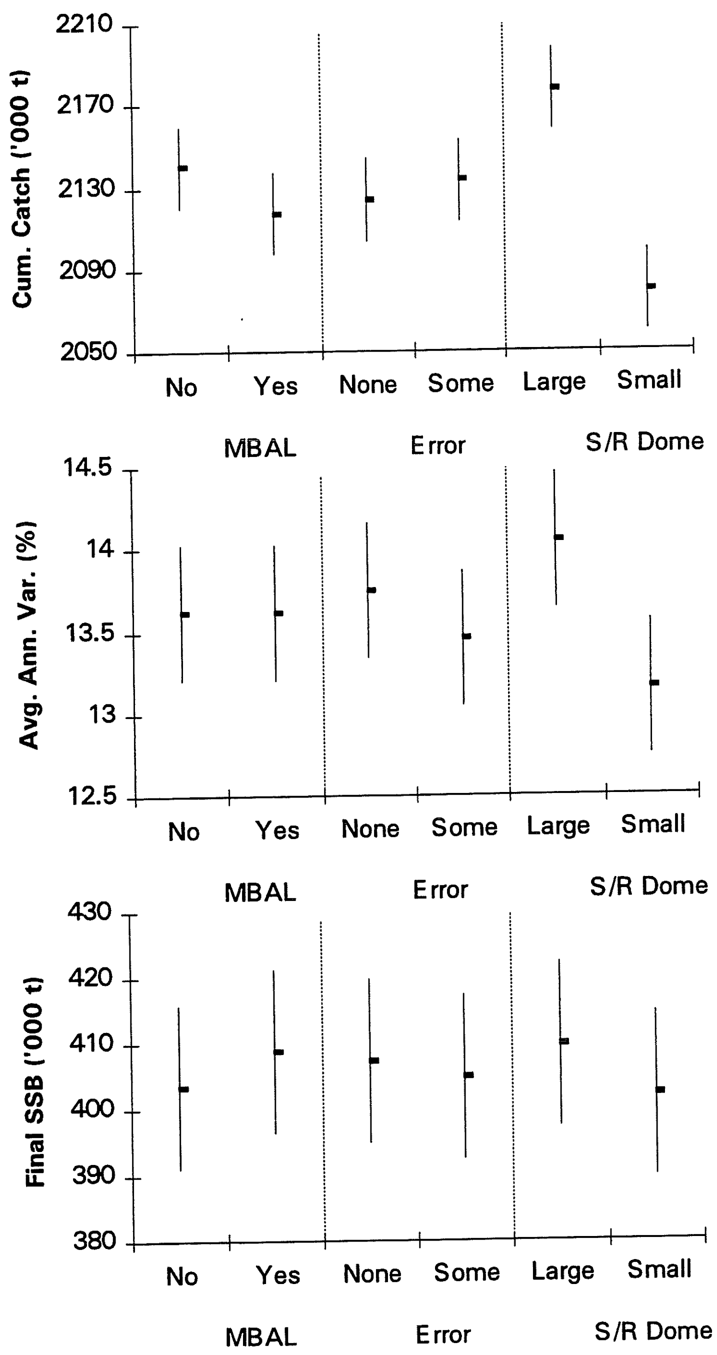


Figure 4.1.1.2 Least square mean estimates of cumulative catch (upper panel), average annual variance (middle panel) and final SSB by main effects in the simulations evaluating North Sea plaice management measures. Error bars show 2 standard errors.

Table 4.1.1.2.1 F-statistics from ANOVA analyses of true SSB, perceived SSB, true catch and the differences (perceived - true) SSB from simulation of North-Sea plaice.

	Performance Statistics			
Treatment	True SSB	Perceived SSB	True Catch	Per - True SSB
Index Error	2.08	9.89	1.05	40.11
Stock/Rec	54.31	30.55	119.53	0.01
Control Law	1.76	0.00	6.33	2.41
Year	165.47	129.54	90.58	44.58

Table 4.1.1.2.2 Parameter estimates from analyses summarised in table 4.1.1.2.1

	Performance Statistics			
Parameter Est.	True SSB	Perceived	True Catch	Per - True SSB
Grand Mean	400031	407635	99210	7604
No Index Error	2288	-6747	-438	-9035
Large Dome	11691	11861	4672	170
No MBAL	-2104	109	1075	-2213
1994	-135328	-196662	12285	-61333
1995	-120448	-185487	-14421	-65040
1996	-87592	-147306	-14436	-59714
1997	-49046	-97291	-15475	-48246
1998	-9629	-46983	-4645	-37353
1999	15398	-14867	532	-30264
2000	29608	-2102	10688	-31710
2001	21754	-3453	12092	-25207
2002	1922	-30248	14414	-32170
2003	-26257	-44664	10009	-18407
2004	-52083	-45438	7849	6645
2005	-70802	-53408	5541	17394
2006	-82394	-86240	3536	-3846
2007	-86874	-92045	-2805	-5172
2008	-78682	-92280	-4123	-13598
2009	-65376	-88065	-7556	-22689
2010	-47211	-75653	-7506	-28442
2011	-27128	-50389	-7985	-23261
2012	-11544	-35355	-4883	-23811
2013	-1830	-15719	-3864	-13889

These analyses can be extended to evaluate the effects of these factors on susceptibility of the stock to fall below MBAL as well as the probability that the population declines below MBAL but the low level is not perceived in the assessment, or the probability a population is perceived to be below MBAL by the assessment but is actually above MBAL. A qualitative inspection of these patterns does not indicate a clear difference in performance mainly because the population (true and perceived) was above the MBAL for 97% of the simulated years:

True SSB	Above MBAL		Below MBAL	
Perceived SSB	Above MBAL	Below MBAL	Above MBAL	Below MBAL
Run 1	1036	13	0	1
Run 2	1018	27	1	4
Run 3	1013	26	1	10
Run 4	1019	28	0	3

4.1.2 Production model

An alternative analysis was conducted using the model described in WD2 (see section 3.2.8). Time constraints did not permit this alternative analysis to parallel the experimental design of the age-structured (FiFi) analysis exactly, but the two analyses do have enough points in common to permit some comparison.

4.1.2.1 Methods

The Netherlands beam trawl survey data for 1985–1993 and the corresponding aggregate catch data (in weight) were used as inputs. The survey data were obtained as a time- and age-specific set of numerical abundance indices. The survey data were used to estimate a time series of age-aggregated biomass indices as follows: Numbers-at-age from the survey in each year were multiplied by the corresponding commercial weights to obtain a time- and age-specific matrix of survey biomass indices. This matrix was then summed across age.

The survey data were used to estimate an age-aggregated average survey catchability (q) as follows: The time- and age-specific survey biomass indices were divided by the corresponding catchabilities estimated from VPA to obtain a time- and age-specific matrix of survey biomass indices, which was then summed across age to get a time series of absolute survey biomass estimates. The time series of aggregate survey biomass indices was then regressed against the time series of absolute survey biomass estimates (with zero intercept). The resulting slope parameter was taken to represent q, with a value of 2.125×10^{-4}.

Survey sampling error (log scale) was set at a value of 0.5 to correspond with the age-structured (FiFi) analysis. Given this rather high value, the harvest sampling error (log scale) was set at the same level (0.5) in order to keep the harvest data from dominating the analysis. Parameters of the joint prior distribution of a (the Gompertz growth parameter) and s (a scale parameter proportionate to process error standard deviation) were set at the levels described in WD2.

Given the above inputs, estimates for the parameters a and s were obtained as the means of their respective marginal posterior distributions, while fishing mortality rate f and carrying capacity k were obtained as their respective maximum likelihood estimates conditional on a and q. This gave parameter values of a=0.187, f=0.398, k=2.528 x 10^6, and s=0.225. The mean of the posterior distribution for log stock size in the terminal time period (1993) was 12.59.

These values were then used to seed a forward projection spanning the years 1994-2014. A single realization of this time series was projected for each of two scenarios, one in which the standard deviation of the observation error on the (log) survey index remained at a level of 0.5 (*status quo* observation error), and one in which it was reduced to zero (no observation error).

For each of these two projections, stationary probability density functions (pdfs) of stock size and yield were computed analytically, following the methods described in WD2. Four constant-rate harvest strategies were explored for each projection: i) a "risk-averse" strategy of harvesting at the rate that maximizes the expectation of the logarithm of stationary yield (f_{MELSY}), ii) a "threshold" strategy of harvesting at the rate that results in a 5% probability of falling below a threshold (MBAL) of 200,000 t in the stationary pdf (f_T), iii) a "status quo" strategy of maintaining the fishing mortality rate at the level of 0.398 estimated for the baseline portion (1985-1993) of the time series (f_{SQ}), and iv) a "certainty-equivalent" strategy of harvesting at a rate equal to the value of parameter a as re-estimated from the entire (1985-2014) time series (f_a; the "certainty equivalent" designation arises from the fact that the f_{MELSY} rate converges on a as the degree of uncertainty surrounding the value of a goes to zero).

For each of the two projections and each of the four harvest strategies, four evaluation measures were computed: i) the expected stationary yield (ESY), ii) ESY relative to the maximum, iii) the expected log stationary yield (ELSY), and iv) ELSY relative to the maximum.

Finally, for each of the two projections, each of the four harvest strategies, and each of the four evaluation measures, two cases were considered: one in which the parameter values used to compute the evaluation measures were set at the levels estimated from the entire (1985-2014) time series, and one in which the parameter values used to compute the evaluation measures were set at the levels actually used to make the 1994-2014 projections. These two cases were labeled "estimated" and "true," respectively.

4.1.2.2 Results

Results are summarized in Table 4.1.2.2. It should be emphasized that these results are examples only, as parameter estimates are based on just a single pair of realizations with respect to projected stock sizes and catches for the years 1994-2014.

Nevertheless, in this example it is clear that both the risk-averse and certainty-equivalent harvest rates (f_{MELSY} and f_a, respectively) are well below either the threshold (f_T) or status quo (f_{SQ}) harvest rates regardless of the observation error level used for the forward projections. The risk-averse harvest rate is consistently lower than the certainty-equivalent rate, as would be expected from their respective definitions. The threshold and status quo harvest rates are approximately equal, purely by coincidence.

In terms of ESY, the risk-averse and certainty-equivalent strategies perform similarly when contrasted with the comparatively poor performance of the threshold and status quo strategies. When the "estimated" parameter values are used to compute ESY, the certainty-equivalent strategy outperforms the risk-averse strategy by definition. However, when the "true" parameter values are used to compute ESY, the relative performance of these two strategies depends on the observation error level.

In terms of ELSY, again the risk-averse and certainty-equivalent strategies outperform the threshold and status quo strategies, although the margin is not so great (being measured on a log scale). When the "estimated" parameter values are used to compute ELSY, the risk-averse strategy outperforms the certainty-equivalent strategy by definition. However, when the "true" parameter values are used to compute ELSY, the relative performance of these two strategies is again ambiguous, depending (in this example) on the observation error level.

Table 4.1.2.2 Results of production model analysis. f = fishing mortality rate (in a constant-rate strategy); $f_{MELSY} = f$ that maximizes expected log stationary yield; $f_T = f$ that results in a 5% chance of falling below threshold; $f_{SQ} = f$ estimated for baseline portion of time series; $f_a = f$ set equal to estimate of Gompertz growth parameter a; ESY = expected stationary yield; ELSY = expected log stationary yield; absolute = results are expressed in tons; relative = results are expressed relative to maximum

Error:	0.5		0.0	
Parameters	**Estimated**	**True**	**Estimated**	**True**
f_{MELSY}	0.136	N/A	0.202	N/A
f_T	0.395	N/A	0.389	N/A
f_{SQ}	0.398	N/A	0.398	N/A
f_a	0.168	N/A	0.247	N/A
ESY (Absolute)				
f_{MELSY}	191000	178000	150000	186000
f_T	119000	129000	135000	132000
f_{SQ}	11800	128000	133000	128000
f_a	195000	185000	153000	178000
ESY (Relative)				
f_{MELSY}	0.979	0.955	0.981	0.997
f_T	0.609	0.695	0.887	0.707
f_{SQ}	0.603	0.689	0.874	0.689
f_a	1.000	0.994	1.000	0.958
ELSY (Absolute)				
f_{MELSY}	11.942	12.021	11.686	12.063
f_T	11.105	11.703	11.417	11.720
f_{SQ}	11.090	11.694	11.394	11.694
f_a	11.918	12.061	11.665	12.024
ELSY (Relative)				
f_{MELSY}	1.000	0.996	1.000	1.000
f_T	0.930	0.970	0.977	0.971
f_{SQ}	0.929	0.969	0.975	0.969
f_a	0.998	1.000	0.998	0.996

4.2 Discussion and Conclusions

The fractional factorial experimental design for the North Sea plaice in Section 4.1.1 considered a limited set of performance questions for a single harvest control law, namely:

i) How is performance affected by two alternative stock-recruitment relationships in the underlying system?
ii) How is performance affected by the presence of observation errors in the tuning data (and, therefore, by variability in assessment results)?

iii) How is performance affected when the harvest control law includes a fixed MBAL of 200,000 t?

Clearly, a number of additional related questions could be asked as detailed below. Additional runs may also be required to investigate interactions.

Question (i) could be extended to include other plausible stock-recruitment relationships. In the runs performed, the two relationships chosen are very similar for SSB > 300,000 t, and it would be informative to explore functional forms that differ at such SSB levels.

Question (ii) could be expanded to investigate more fully the effect of different components of uncertainty (catches, abundance data), in terms of the shape of uncertainty distributions. The example assessment model, XSA, assumes lognormally distributed relative abundance observations. One could investigate the effect of alternative data distributions (e.g. gamma) as well as the effect of alternative assumed distributions (in XSA).

Question (iii) should be expanded to explore alternative MBALs. The 200,000 t level was chosen arbitrarily and a more comprehensive assessment should examine this critically. An important alternative is to explore MBAL definitions that depend on the assessment results, e.g., the SSB that corresponds to 50% of the maximum recruitment in a fitted stock-recruitment relationship (see e.g., Rosenberg *et al*, 1994; Thompson, 1994). It may also be important to compare management strategies that allow for MBAL to be redefined as data are accumulated (e.g., after each assessment) against strategies that use a fixed MBAL, set at the beginning of the management period.

The examples above illustrate alternatives to the three questions that were asked during this demonstration. In the planning of a fuller (or comprehensive) assessment for North Sea plaice, it would be necessary to identify additional questions to be explored. Such questions could be grouped into the following categories:

- Performance under alternative underlying biological systems. The underlying system could contain a spatial component with migration, such as that presented in WD 1 (see Section 3.2.1) or based on work presented in WDs 9 and 10 (see Section 2.2.4). In addition, species interactions could be explored although they are thought to be negligible for North Sea plaice.

- Performance under alternative observation/assessment systems. The assessment model can either ignore or account for spatial structure and migration, as illustrated in WD 1. In addition, various types of assessment models can be examined such as length-structured ones, production models, etc.

- Performance under alternative implementation tactics. Decisions can be either fully or imprecisely implemented, with or without bias. It is particularly important to hypothesise which types of management decisions can lead to deterioration in the quality of data collected from the system which would have a feed-back effect. For instance, spatial catch allocations may be difficult to implement and/or monitor, and may impact the proportion of discards through time. Also, fleets' behavioural characteristics often change in response to regulations and it is important to take them into account during implementation (see also Sections 2.2.3 and 2.2.4).

It is possible to extract or generate a variety of summary statistics to use as measures of performance, e.g. some percentile (such as the median) of SSB and/or catch at a particular time or over a time scale. Such statistics might be final SSB, Cumulative Catch, the number of times that SSB falls below a reference level (MBAL or, if the stock is being rebuilt, below the initial level), average annual variation (AAV) in catch or variation over different time scales to reflect different priorities. The latter time scales might be three years to reflect the high discount rates of fishermen, typically 30%, or longer reflecting the length of time needed to rebuild a stock.

Trajectories over time can be presented but are open to misinterpretation. Summary statistics such as cumulative yield or time-specific quantities such as continuing yield beyond a certain year are therefore preferable. All of these measures would be presented with associated distributions or distribution statistics.

The policy objectives under consideration such as optimising yield, economic return, employment or conservation of stocks will dictate the actual summary statistics that are of interest in each case.

To further develop the simulation approaches detailed in Section 3, it will be necessary to include details such as those discussed in Section 2. Whilst fractional factorial designs allow the information content of simulations to be optimised, they also allow response surfaces to be calculated over a range of levels for he design factors. It should be possible to use such response surfaces as a simplified way to implement sub-models within an evaluation framework.

The North Sea plaice data were also analyzed using a production model (section 4.1.2). The easiest point of comparison between the production model results and the results of section 4.1.1 is the probability of the stock falling below MBAL. In section 4.1.1, the results indicate only about a 3% chance of falling below MBAL under the *status quo* fishing mortality rate. In comparison, section 4.1.2 indicates that the fishing mortality rate corresponding to a 5% probability of falling below MBAL (in the stationary distribution) is very close to the *status quo* rate.

Central differences between the production model analysis and the FiFi analysis are that the former i) has a simpler structure (e.g., it omits age structure); ii) assumes zero implementation error; iii) examines only a single realization in its simulation of the years 1994-2014; iv) uses Bayesian techniques to estimate parameters; and v) examines a wider variety of control laws, including one based on an explicitly risk-averse utility function.

The production model's ability to address both frequentist and decision-theoretic formulations of risk (see section 3.3.2), its ability to obtain closed-form solutions for future probabilities of stock size and yield, and its adaptability to fisheries which lack age data commend its further development. The Working Group therefore recommends that development of this approach continue during the coming year.

5 INCORPORATION OF ADVICE ON MANAGEMENT STRATEGIES INTO THE ADVICE GIVEN BY ACFM

5.1 Introduction

The Working Group reiterates that management measures need to be evaluated using quantities which relate their performance to either inferred or, preferably, specified *objectives*, perhaps in the form of utility functions. It is unlikely that the various management commissions or national governments, to which ICES provides advice as a basis for decision making, will formalise objectives, though this should, perhaps, be further explored. There are, however, certain objectives which are appropriate to all fisheries systems. These would, for example, relate to conservation, yield levels, stability of yield and economic factors. It is possible for this Working Group to provide performance indicators for various management measures, evaluated for given fisheries, which reflect these generalised objectives. Specifically, the Working Group could provide performance indicators for conservation (e.g., SSB), yield (e.g., total catch or continuing catch after some specified time period), stability of yield (e.g. AAV - average annual variation in catch) and economic performance (e.g., NPV - net present value).

It is emphasised that the purpose of the evaluation procedures advocated by the Working Group is to compare, relative to some baseline, the performance of different assessment and/or management procedures conditional on an underlying system model. The advocated evaluation procedures do not in any sense attempt to predict the detailed trajectories of stocks in the short, medium or long-term.

The evaluation of the procedures tested involves the use of summary outputs which provide probabilistic information on quantities of interest such as SSB (true or perceived), yield, AAV, NPV. Deciding which management measures to implement is not simple but involves either a formalised or *ad hoc* trade-off, between the various objectives, reflected by the probabilistic summaries of the relevant performance indicators. Whilst there is a well developed field of multi-criteria decision making, actual resource management decisions tend to be accomplished by human integration. For this reason, readily interpretable summaries of a limited set of performance indicators are essential.

Note that the quantities relevant to decision making will not all necessarily be biological. The Working Group would not suggest that detailed economic analyses be conducted by ICES working groups but it may eventually be desirable for ICES to move in this direction in order that it can provide relevant, broad-based advice to its various "clients". In the near future, it may be desirable to include at least some simple economic indicator(s) in the summary outputs. The obvious candidate quantity is NPV; a statistic which is easy to calculate given simple economic inputs and which provides an indication of the trade-off in immediate and near-term costs and benefits.

For certain fisheries, the indicators suggested above (i.e., SSB, yield, AAV and NPV) might be sufficient. For particular cases, however, alternative or additional indicators might be appropriate. The Working Group does not believe that it would be easy or necessarily helpful to incorporate an over-simplified, standard summary output into the ACFM report, for any stocks/fisheries for which a "comprehensive assessment" is undertaken.

Perhaps the best way that the Working Group could facilitate the decision making process, and advisory function of ACFM would be to develop a standardised format for the presentation of performance indicators. In this way, it should be possible for both the Working Group and ACFM to become familiar with the interpretation of summary outputs from complex modelling exercises and for ACFM to convert these summaries into written, unambiguous advice. If a standardised format were to be developed, this could, of course, be incorporated into its report if ACFM so wishes.

The Working Group suggests that a number of the outputs in section 4 of this report are good models of how to summarise outputs. The Working Group further notes that those outputs are not without precedence and are modelled on outputs now commonly used in, for example, the International Whaling Commission, where they have been used by both scientists and administrators in the decision making process.

To help ACFM decide on the appropriateness of these outputs, or perhaps to suggest alternatives, the following section explains in detail how the plots are constructed and how they should be interpreted.

5.2 Examples

Figures 5.2.1 and 5.2.2 show two ways of elucidating the trade-off between objectives (reflected by certain performance statistics) when different management procedures are imposed on an underlying system. The outputs summarised in these figure are derived from the work reported in WD 4 rather than from the demonstration reported in this report (Section 4).

The performance statistics shown are final "true" SSB, average annual variation in catch (AAV), cumulative catch and the number of times the "true" SSB falls below MBAL. Note that this does not necessarily equate to the number of times that SSB is perceived to fall below MBAL.

Figure 5.2.1 shows median and percentile plots of the performance statistics for four control laws (F_{SQ}; F_{MED}; $f(F_{SQ}$, F_{MED} and $F_{MBAL})$; multi-annual version of $f(F_{SQ}$, F_{MED} and $F_{MBAL}))$. It is straightforward to compare the performance of the four control laws across the three statistics.

Figure 5.2.2 summarises the same information in a different way. The plots show the 95% confidence ellipses of the performance indices collected from all replicate runs of the scenario model. The trade-offs for the four control laws are clearly seen.

It is apparent that, in this example case, the F_{SQ} control results in a higher cumulative catch than all other controls but that this is traded-off against a lower final SSB and a greater number of occasions when the true SSB is below MBAL. Note that in this particular example, also based on North Sea plaice, MBAL is taken to be 300,000 t as used by ACFM (not 200,000 t as in section 4).

Figures such as 5.2.1 and 5.2.2 may contain information of two types: i) the relative performance of different management procedures on a range of relevant quantities (e.g., SSB, AAV) for a specified underlying system model, or ii) the performance of the same management procedure on a range of quantities for a range of assumptions in an underlying system model. The figures presented above are of the first type. They could be used to assess the relative performance, and implied trade-off in objectives, for a range of management procedures applied to a baseline underlying system. Figures of the second type would be used to test the robustness of management procedures to assumptions in the underlying system model. Figure 4.1.1.2 combines information of both types (i) and (ii) and is thus not included in this section. It's use would probably be confined to the Working Group to aid in the development and screening process.

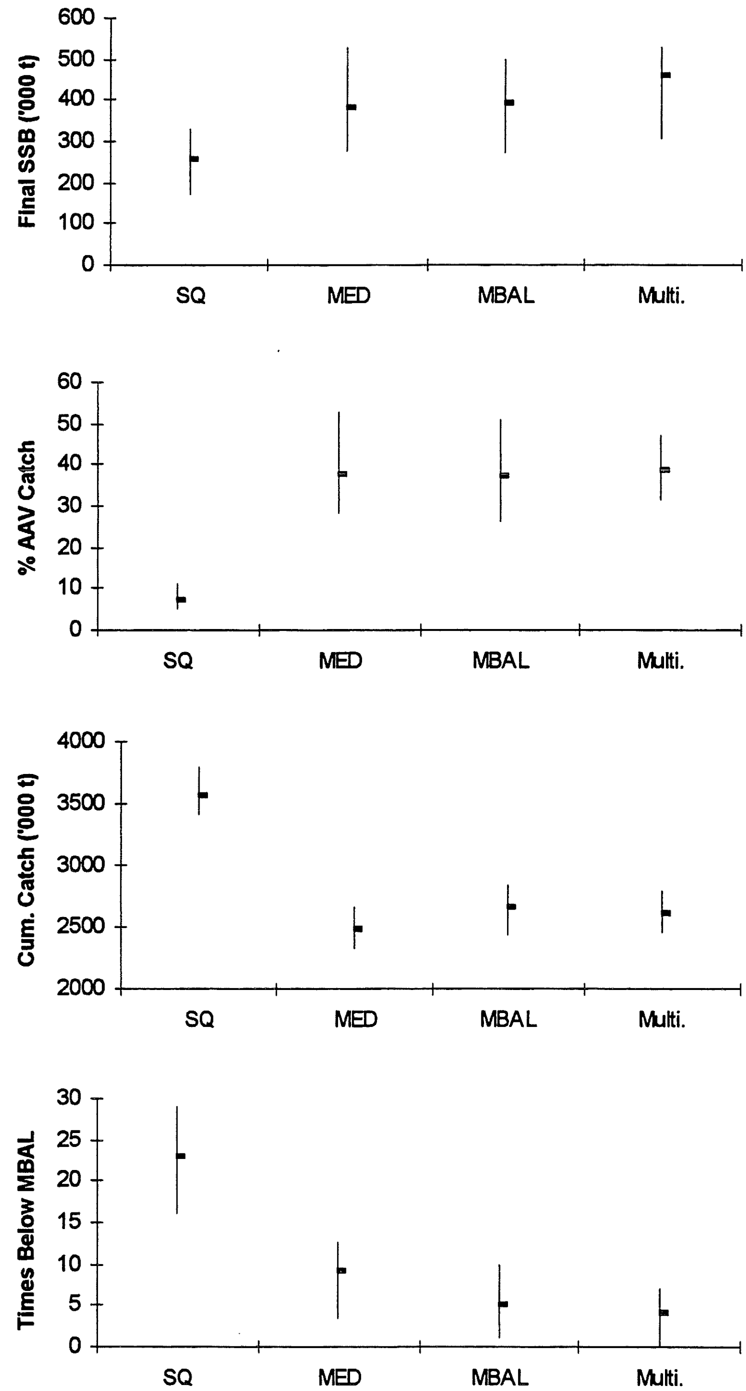
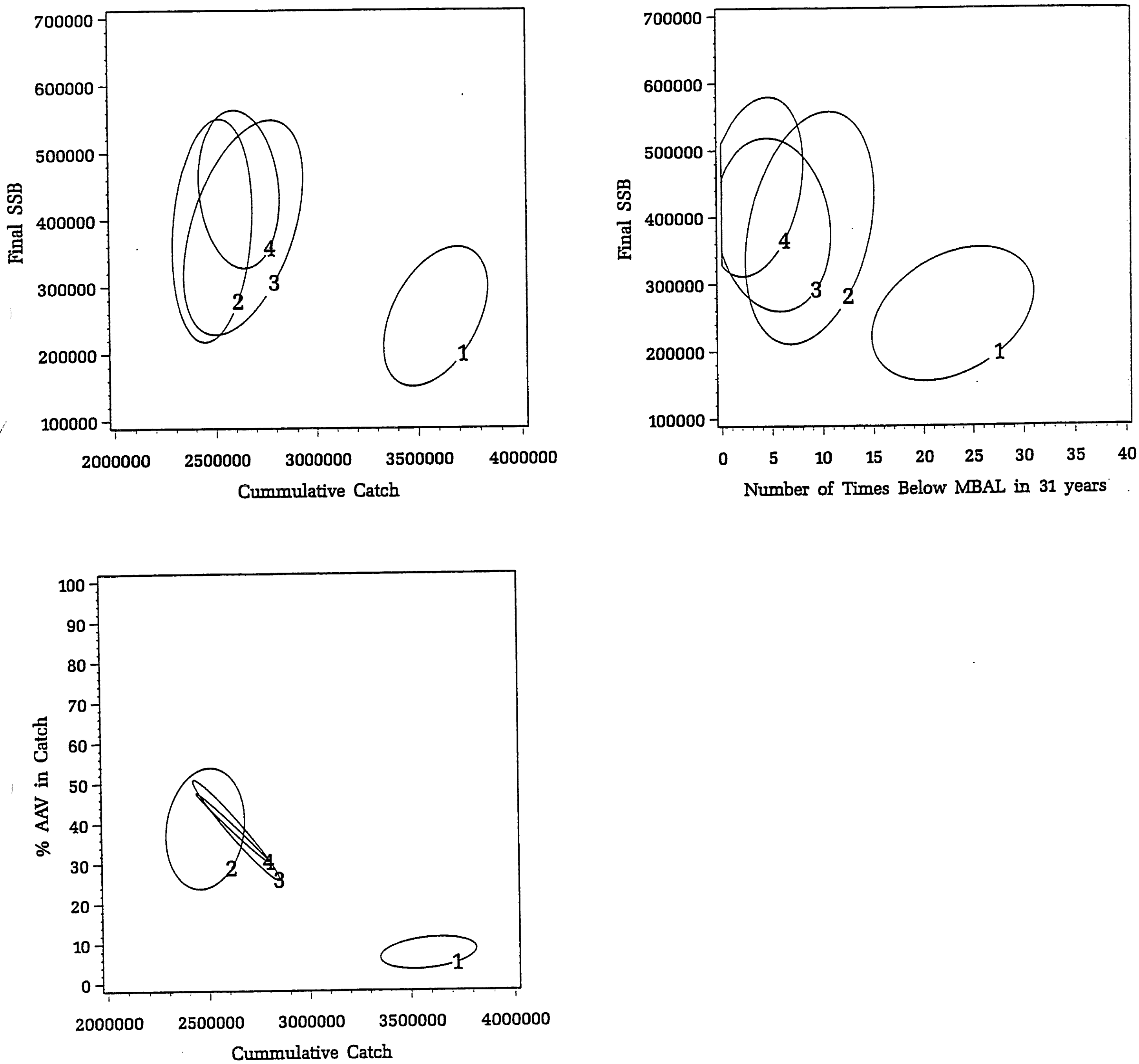


Figure 5.2.1 Median values of final SSB, cumulative catch, AAV catch, and number of times SSB < MBAL for simulations of 4 control laws for North Sea plaice. Bars show the 5th and 95th percentiles.

Figure 5.2.2 a: 95% confidence ellipsesof %AAV in catch *vs* Cumulative catch for four control laws, as described in the text. b: 95% confidence ellipses of Final SSB *vs* Cumulative catch. c: 95% confidence ellipses of Final SSB *vs* the number of times MBAL falls below the true SSB. Numbers 1-4 refer to four different control laws (see text).



6 ANY OTHER BUSINESS

6.1 Response to ACFM Questions on WG Activities

The Working Group noted a letter from the Chairman of ACFM asking for responses to questions concerning the work of the group. It was agreed that individual members would provide feedback to the Chairman who would arrange feedback to the Chairman of ACFM.

6.2 Request for Advice on the Setting Up of an ICES Gear Selectivity Database

The Working Group considered a proposal for creating an ICES database on gear selectivity. The Working Group noted that :

i) there is no standardised method for estimating selection parameters, which in some cases may even be based only on fitting a selection curve by eye. A database containing only the selection parameters without giving the basis for their estimation is therefore not satisfactory;

ii) there are numerous factors which affect gear selectivity, e.g. size of the catch, fishing area, vessel size, rigging of the gear, etc., and a database should contain the relevant information about such factors;

iii) recent developments in gear technology, e.g. introduction of grids in trawls, has made the calculation of selection parameters more complex. Disaggregated data (raw data from each haul) would give more flexibility in developing models for selectivity and in making statistical calculations.

Although the Working Group in principle supports the idea of an ICES Gear Selectivity Database, it is realised that considerable effort is needed to set up such a database and that this may interfere with other and possibly more important tasks of the Secretariat.

6.3 The Future of the LTMM WG and Other ICES Methodologically Oriented Working Groups

The Working Group discussed whether or not its work would be best served by continuing as a current, single working group or integrated with that of other working groups. At present, there are practical problems associated with the timing of the meetings and the availability of participants. Also, and more importantly, it is clear from the North Sea plaice work in this meeting (Section 4), the Barents Sea multispecies work outlined in section 3.2.5 or similar work in other fora, that in order properly to develop underlying system models and investigate the performance of management procedures, a wide range of contributors is required. It is also clear that such work cannot be started and finished within one meeting but requires considerable commitment, coordination and inter-sessional work.

The Working Group considered various options for combining with other ICES working groups - particularly the other methodologically-oriented ones. One suggestion was that the Multispecies and Long-Term Management Measures Working Groups could usefully combine. Another suggestion was that, at regular intervals, the Assessment Working Groups should conduct fuller assessments with enhanced participation. There was no general agreement as to what would be an ideal way forward but there was strong agreement that the work of this Working Group should be continued. Specifically, there was agreement that the evaluation of management measures must account for uncertainty and that it is important to consider fisheries systems rather than individual stocks.

The Working Group noted that the evaluation of fisheries systems is always case specific and requires a considerable amount of detailed analyses. The Assessment Working Groups are clearly already fully committed to the basic updating of annual single species assessments. It is unlikely, therefore, that adding the burden of more wide-ranging fisheries systems assessments is a sensible option. In order to conduct assessments of fisheries systems, multispecies aspects (both biological and technical) need to be considered as too do economic aspects. The work involved can be statistically, mathematically and computationally demanding but requires considerable biological and other non-technical expertise and input. For this reason, the Working Group would suggest the continuation of work addressing issues relevant to the management of fisheries systems.

The current Chairman of the Working Group finishes his term after this meeting. The Working Group discussed a possible replacement and agreed that the current Chairman would relay its views to ACFM.

With regard to the next meeting of the Working Group, two of the recommendations in Section 7 are of direct relevance. The remaining recommendations could all be refined into useful terms of reference in consultation with ACFM and the new Working Group Chairman.

7 RECOMMENDATIONS

The subject of MBALs and other suitable population thresholds is receiving considerable attention. Frameworks for the evaluation of management measures such as those demonstrated in this report provide an ideal setting for examining the performance of alternative MBAL definitions for particular stocks. The Working Group **recommends** that, at its next meeting, it consider the subject of MBAL definitions for particular stocks, with emphasis on evaluating their robustness and efficacy in terms of achieving specific management objectives.

The Working Group, through its activities, has identified useful components of a fishery system that can be programmed in a modular way. These could be integrated into a generic environment for comprehensive modelling and assessment. The Working Group **recommends** that additional generic tools be developed with a view to enhancing future flexibility. This will require inter-sessional contact between Working Group members.

A production model is described in Section 3.2.8. This model's ability to address both frequentist and decision-theoretic formulations of risk (see section 3.3.2), its ability to obtain closed-form solutions for future probabilities of stock size and yield, and its adaptability to fisheries which lack age data commend its further development. The Working Group therefore **recommends** that development of this approach continue during the coming year.

The Working Group noted that none of the working documents explicitly considered structural errors (i.e., model errors). Qualitatively different models can lead to quantitatively the same inferences. Therefore, the Working Group **recommends** further study of the equivalencies between differing model formulations and errors.

The STCF database is a valuable source of data which has not been fully utilised. The Working Group notes that access to these data are important to its future activities and therefore **recommends** that an investigation should be undertaken to assess the information content and its usefulness. Results of such analyse should be presented at the next Working Group meeting.

Given the scope for complexity and the associated computational needs of management evaluation work, it would be useful to investigate criteria for determining appropriate types and levels of process, measurement, model, and implementation error for evaluation of specific problems. The Working Group **recommends** the development of case studies which explore this issue.

The Working Group **recommends** that the scope for limiting the number of iterations and experimental runs, and the ability to generate response surfaces that may be used to generate added realism within simulation frameworks, should be further investigated using efficient experimental designs.

The Working Group **recommends** that simulation frameworks such as those in Section 3, be used to evaluate the effectiveness of area closures (e.g., the "plaice box"). Particular attention should be paid to the power of detecting gains or losses from such closures given data by model by assessment interactions.

8 REFERENCES AND WORKING DOCUMENTS

8.1 References

Anon., 1987, Report of the ad hoc meeting of the North Sea Flatfish Working Group. ICES C.M. 1987/Assess:14. 84 pp.

Anon.,1993, Report of the Working Group on Long-Term Management Measures. ICES C.M.1993/Assess:7

Anon.,1994, Report of the Working Group on Long-Term Management Measures. ICES C.M. 1994 /Assess:11

Anon., 1994b. Report of the Study Group on the Plaice Box. ICES C.M.1992/Assess:14. 52 pp.

Anon., 1995, Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak. ICES C.M.1995/Assess:8

van Beek, F.A., 1990, Discard sampling programme for the North Sea, Dutch participation. Internal Report Netherlands Institute for Fisheries Research, DEMVIS 90-303. 24 pp.

Box, G.E., Hunter, W.G. and S.Hunter, 1978, Statistics for Experimenters. Wiley, New York.

Darby, C.D. and S. Flatman, 1992, Virtual Population Analysis. Version 3.1 (Windows/Dos) User Guide Information Technology Series no. 1, Directorate of Fisheries Research, Lowestoft.

Donovan, G.P. (Ed.), 1989, The Comprehensive Assessment of Whale Stocks: the early years. *Rep. IWC*, Special Issue 11.

Francis, R.I.C.C., 1992, Use of risk analysis to assess fishery management strategies: a case study using orange roughy (Hoplostethus atlanticus) on the Chatham Rise, New Zealand. *Can. J. Fish. Aquat. Sci.* 49:922-930

Gillis, D.M., R.M. Peterman, and A.V. Tyler, 1993, Movement dynamics in a fishery: application of ideal free distribution to spatial allocation of effort. *Can. J. Fish. Aquat. Sci.* 50: 323-333.

Horwood, J.W. ,1994, Modelling of Fisheries Management Strategies. Privately published, 85 pp.

IWC, 1993, Report of the Scientific Committee. *Rep. Int. Whal. Commn.* 43:57-64

Janssen, G., W. Dol, M. Pastoors and A.D. Rijnsdorp, 1994, A bio-economic simulation model of flatfish fisheries in the North Sea. ICES C.M. 1994/T:19.

de la Mare, W.K., 1985, Simulation studies on management procedures. *Rep. Int. Whal. Commn.* 36:421-450

de la Mare, W.K., 1986, On the management of exploited whale populations. D.Phil thesis, University of York.

Pelletier, D., 1991, Towards more efficient TAC policies with error-prone data. NAFO Sci. Coun. Studies, 16: 153-163.

Pelletier, D. and A. Laurec, 1992. Management under uncertainty: defining strategies for reducing overexploitation. *ICES J. Mar. Sci.*, 49: 389-401.

Powers, J.E. and V. R. Restrepo, 1993, Evaluation of stock assessment research for Gulf of Mexico king mackerel: benefits and costs to management. *North Am. J. Fish. Management* 13: 15-26.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and B.P. Flannery, 1992, Numerical Recipes in C (2nd. Edition). Chapman and Hall, London, 994 pp.

Punt, A.E., 1991, Management Procedures for Cape Hake and Baleen Whale Resources. Ph.D. Thesis, University of Cape Town, 875 pp.

Punt, A.E., 1995, The performance of a production-model management procedure. *Fisheries Research* 21:349-374

Quinn, T.J., Deriso, R.B. and P.R. Neal, 1990, Migratory catch-age analysis. *Can. J. Fish. Aquat. Sci.*, 47:2315-2327

Restrepo, V. R., J. M. Hoenig, J.E. Powers, J. W. Baird and S. C. Turner, 1992, A simple simulation approach to risk and cost analysis with applications to swordfish and cod fisheries. *Fish. Bull. U.S.* 90: 736-748.

Restrepo, V. and A. Rosenberg, 1994, Evaluating the performance of harvest control law under uncertainty via simulation. WD 18 submitted to ICES Long-Term Management Measures Working Group (ICES C.M. 1994/Assess:11).

Rijnsdorp, A.D., 1993, Selection differentials of male and female North Sea plaice Pleuronectes platessa L. and changes in maturation and fecundity. p 19-39 In Stokes, T.K., McGlade, J.M. and R. Law (eds.), The Exploitation of Evolving Resources. Springer Verlag, Berlin.

Rijnsdorp, A.D. and P.I. van Leeuwen, 1994, Changes in growth of North Sea plaice since 1950 and its relation to density, eutrophication, beam trawl effort and temperature. ICES C.M. 1994/G:10.

Rijnsdorp, A.D. and M. Pastoors, 1994, A simulation model of the spatial dynamics of North Sea plaice (*Pleuronectes platessa* L.) based on tagging data. ICES C.M. 1994/Mini:16.

Rosenberg, A., Mace, P., Thompson, G., Darcy, G., Clark, W., Collie, J., Gabriel, W., Maccall, A., Methot, R., Powers, J., Restrepo, V., Wainwright, T, Botsford, L., Hoenig, J. and K. Stokes, 1994, Scientific review of Definitions of Overfishing in U.S. Fishery Management Plans. Document prepared for the National Marine Fisheries Service, 205 pp.

Sakuramoto, K. and S. Tanaka, 1986, A simulation study on management of whale stocks considering feedback systems. SC/38/O10 (unpublished paper presented to the IWC Scientific Committee).

Shepherd, J.G., 1982.

Shepherd, J.G.,1984, A promising method for the assessment of multispecies fisheries. ICES C.M. 1984/G:4.

Sparre, P., 1984, A computer program for estimation of food suitability coefficients from stomach content data and multispecies VPA. ICES C.M. 1984/G:25.

Thompson, G.G., 1994, A general diffusion model of stock-recruitment systems with stochastic mortality, ICES C.M. 1994/T:18

UNEP, 1992, Report of the Meeting of the Scientific Advisory Committee of the Marine Mammal Action Plan, Liege 27 November -1 December.

8.2 Working Documents

Working documents are listed sequentially by number and then in the form of a “taxonomy” for future use.

8.2.1 Sequential WD list

WD1 Punt, A. and V. Restrepo, Some effects of ignoring mixing when managing fish populations subject to ‘limited’ mixing.

WD2 Thompson, G., Risk averse implementation of a Kalman Filter model of stock dynamics with allowance for parametric nonstationarity.

WD3 Kell, L. and K. Stokes, A framework for evaluating management under uncertainty.

WD4 Kell, L., Stokes, K. and B. Rackham, Example runs and outputs from FiFi based on a prototype North Sea plaice implementation.

WD5 O'Brien, C., Kell, L., Rackham, B. and K. Stokes, The feasibility of modelling and predicting catchability-at-age using vessel trip data on North Sea cod for the years 1983-1992.

WD6 Christensen, S., Evaluation of management strategies in the Greenland shrimp fishery

WD7 Marchal, P. and J.Horwood, Multi-annual TACs and minimum biological levels.

WD8 Marchal, P., Assessing groundfish resources of the Celtic Sea VIIF+G with several métiers: a multiannual approach

WD9 Rijnsdorp, A.D., Sensitivity Analysis and Validation of a Simulation Model of Flatfish and the Flatfish Fishery in the North Sea

WD10 Rijnsdorp, A.D. and M.A. Pastoors, A Simulation model of the Spatial Dynamics of North Sea plaice (*Pleuronectes platessa L.)* based on tagging data

WD11 Magnusson, K., A preliminary note on possible management procedures for Oceanic Redfish *Sebastes mentella* and some deterministic simulations

WD12 Hagen, G., Hatlebakk, E. and T. Schweder, SCENARIO BARENTS SEA: A tool for evaluating fisheries management regimes

WD13 O'Brien, C., Modelling and predicting catchability-at-age using vessel trip data on North Sea plaice for the years 1983-1992.

8.2.2 "Taxonomy" of WDs

1) Deterministic versus stochastic

- A) Prediction versus monitoring
- B) Estimation versus exploration
- C) Implementation error
  - i) WD3 (Kell and Stokes)
  - ii) WD4 (Kell et al.)

2) Spatial heterogeneity versus spatial homogeneity

- A) Migration
  - i) WD1 (Punt and Restrepo)
  - ii) WD9 (Rijnsdorp)
  - iii) WD10 (Rijnsdorp and Pastoors)
- B) Spatial statistics (GIS, kriging)

3) Long-term versus short term

- A) Multi-annual TACs
  - i) WD7 (Marchal and Horwood)
  - ii) WD8 (Marchal)
- B) Stationary pdf versus transition pdf
- C) Equilibrium versus non-equilibrium

4) Simplicity versus complexity

- A) Overparameterisation
  - i) WD5 (O'Brien et al.)
  - ii) WD13 (O'Brien)
- B) Model misspecification
  - i) WD1 (Punt and Restrepo)
- C) Analytic versus numerical solution techniques
- D) Mean-variance/ accuracy-precision trade-offs

5) Age (or size) structure versus production

- A) Age (or size) structure
  - i) WD1 (Punt and Restrepo)
  - ii) WD3 (Kell and Stokes)
  - iii) WD4 (Kell et al.)
  - iv) WD5 (O'Brien et al.)
  - v) WD6 (Christensen)
  - vi) WD7 (Marchal and Horwood)
  - vii) WD8 (Marchal)
  - viii) WD9 (Rijnsdorp)
  - ix) WD10 (Rijnsdorp and Pastoors)
  - x) WD12 (Hagen et al.)
  - xi) WD13 (O'Brien)
- B) Production
  - i) WD 2 (Thompson)
  - ii) WD 11 (Magnusson)

6) Single species versus multiple species

- A) Potential use of multiple species
  - i) WD 3 (Kell and Stokes)
- B) Limited use of multiple species
  - i) WD 12 (Hagen *et al*)
- C) Full ecosystem models

7) Parametric stationarity versus non-stationarity

A) Discrete changes in parameter values
i) WD 2 (Thompson)

B) Continuous changes in parameter values (trends)
i) WD3 (Kell and Stokes)

8) Frequentist versus Bayesian

A) Probability versus risk
i) WD2 (Thompson)

B) Use of priors versus non-use thereof