1360

The effect of management choices on the sustainability and economic performance of a mixed fishery: a simulation study

S. B. M. Kraak, F. C. Buisman, M. Dickey-Collas, J. J. Poos, M. A. Pastoors, J. G. P. Smit, J. A. E. van Oostenbrugge, and N. Daan

Kraak, S. B. M., Buisman, F. C., Dickey-Collas, M., Poos, J. J., Pastoors, M. A., Smit, J. G. P., van Oostenbrugge, J. A. E., and Daan, N. 2008. The effect of management choices on the sustainability and economic performance of a mixed fishery: a simulation study. – ICES Journal of Marine Science, 65: 697–712.

Alternative management scenarios were evaluated in a simulation framework that mimicked the recent exploitation of sole and plaice in the North Sea. A large proportion of plaice is taken as bycatch of the beam trawl fleet targeting sole, yet current management of the two stocks assumes no interaction in their exploitation. The evaluation criteria included biological and economic sustainability, and stability in the management measures. The fishery was assumed to respond to management restrictions by dropping the least profitable trips. We investigated two contrasting management strategies, single-species total allowable catches, and effort regulation. Under the assumptions made, the latter strategy performed better. The results suggest that, given assessment error and bias, a strategy that accounts for the mixed nature of a fishery and that occasionally results in perceived underexploitation may work best. Stability in fishing mortality reinforces itself, through lower assessment bias, and management corrections become less frequent. The common assumption in many stock assessments in EC waters that fishing mortality in the most recent year should resemble the value obtained in previous years ("shrinkage") had a negative effect on the stability of control measures.

Keywords: economic performance, effort regulation, management strategy evaluation, mixed fishery, North Sea plaice, North Sea sole.

Received 14 December 2006; accepted 20 February 2008.

S. B. M. Kraak, M. Dickey-Collas, J. J. Poos, M. A. Pastoors, and N. Daan: Wageningen IMARES (Institute for Marine Resources and Ecosystem Studies), PO Box 68, 1970 AB IJmuiden, The Netherlands. F. C. Buisman, J. G. P. Smit, and J. A. E. van Oostenbrugge: Agricultural Economics Research Institute (LEI), PO Box 29703, 2502 LS Den Haag, The Netherlands. Correspondence to S. B. M. Kraak: tel: +31 (0)317 487177; fax: +31 (0)317 487326; e-mail: sarah.kraak@wur.nl.

Introduction

Management by total allowable catches (TACs) of species caught in mixed fisheries is problematic because the quota of different species may be exhausted at different rates. Fishers are therefore faced with a dilemma when the quota for one species is exhausted: stop fishing and underutilize the quota for other species, or continue fishing and discard or illegally land overquota fish (Daan, 1997). When they choose the latter option, the target fishing mortality (F) implied by the TAC will be exceeded, and the scientific basis for stock assessment and future management advice will be compromised if the stock assessment is based on official landings data alone (assuming that these equal the catches).

In the North Sea, a large proportion of the landings of plaice (*Pleuronectes platessa*) and sole (*Solea solea*) comes from the Dutch beam trawl fishery in the form of a mixed catch. However, exploitation of the two stocks is managed separately, and the TAC management does not account for their technical and economic interactions (Piet and Rice, 2004). Even the new EC management plan for North Sea plaice and sole (EC, 2007), which was adopted in 2007 and implemented for the first time in 2008, considers the TACs for the two species separately. Although the problem is well known, the solution is less obvious. An effort-management system is a possible alternative. Although the EC plan complements the two single-species TACs with effort limitations, it is currently not clear whether the fishery will be restricted by TACs or by effort limitations. Therefore, it is worth investigating the potential effects on the stocks and on the viability of the fisheries of a management system that restricts effort.

Simulations of exploited populations and management procedures provide insight into the sensitivities of a system to different management regimes, even if the ultimate management performance cannot be predicted (Butterworth and Punt, 1999; Sainsbury *et al.*, 2000; Kell *et al.*, 2002, 2005; Punt *et al.*, 2002; Ulrich *et al.*, 2002; Harwood and Stokes, 2003; Pastoors *et al.*, 2007). We developed a simulation framework containing a "true" population and a "perceived" population, as well as a feedback loop between the two. In the simulation, data collected each year from the true population serve as input to the annual stock assessments. As the data collection and the assessment procedures introduce error, the resulting perceived population deviates from the true population. It is the perceived population that drives the management decisions. These decisions are then imposed on the true population in the next year of the simulation.

Building on earlier work (Ulrich *et al.*, 2002; Machiels *et al.*, 2007; Pastoors *et al.*, 2007), we attempt to characterize the mixed North Sea flatfish fishery and to explore and understand the potential effects of alternative management regimes.

The simulation framework was based on Kell *et al.* (2002) and Kell and Bromley (2004), but with the addition of economic considerations. Although the economic impact of alternative management strategies has been analysed for various fisheries (e.g. Frost, 1997; Sutinen, 1999), simulations have rarely been used (de Wilde, 1999; Salz and Frost, 2000; Pascoe *et al.*, 2001; Ulrich *et al.*, 2002), because this requires economic data by fleet or preferably even by individual vessel and trip.

We report on two scenarios here, one reflecting current management by TACs, and one based on effort restrictions. The management goals for both scenarios were set according to the precautionary approach to fishery management, although the new EC management plan formulates different objectives, such as target fishing mortalities. The precautionary approach provides the framework for management advice delivered by the International Council for Exploration of the Sea (ICES, 2001). It requires that to have stocks and fisheries within safe biological limits, the probability should be high that the spawning-stock biomass [B (although the ICES jargon traditionally uses the symbol SSB for spawning-stock biomass, we use B for consistency, because B_{lim} and B_{pa} also refer specifically to spawning-stock biomass)] is above a limit value (called B_{lim}) below which recruitment may become impaired. In addition, the probability should be high that fishing mortality (F) is below a limit value (called F_{lim}) that will drive B to B_{lim}. Because of uncertainty in the annual estimates of B and F, ICES has defined more conservative (precautionary) operational reference points, B_{pa} and F_{pa} (the subscript pa standing for precautionary approach). When a stock is estimated to be above B_{pa}, the probability should be high that in reality it is above B_{lim}. Similarly, when F is estimated to be below F_{pa}, the probability should be high that in reality it is below F_{lim}. This management framework is therefore risk avoiding and does not target reference points such as MSY. In practice, F_{pa} is often used as a *de facto* target F, because the TAC advice is usually the catch that is forecast under F_{pa}. Therefore, the advice is simulated corresponding to F_{pa} and the TACs and effort restrictions are set accordingly. Note that, whereas in reality the advice is not always followed, in our simulations it is.

The objective of our study was to monitor the two management scenarios over annual time-steps, and to evaluate their performance with respect to biological (e.g. sustainable stock development), economic (e.g. net revenues), and management criteria (stability of management measures).

Methods

The model was implemented in the FishLab simulation framework [the version used here is held by the first author; the FishLab framework is no longer maintained, because it has been transferred to FLR (Kell *et al.*, 2007)], representing a set of dynamic link libraries (DLLs) that can be called from within Excel (Kell *et al.*, 2002, 2005). The equations are listed in Appendix A. Simulations were run in a "Monte Carlo" set-up ($\times 100$) to evaluate the variability in the final outcomes. The structure of the model is illustrated in the flow diagram (Figure 1).

Basic model

The model consisted of two main parts: the operating model (OM) simulated the true system, and the management procedure (MP) simulated the perceived system and associated management decisions. The OM represented two age-structured populations that mimicked North Sea sole and plaice, respectively. These populations were developing in annual time-steps from a starting population in 1957, given annual recruitment and mortality (F and natural mortality M). The MP simulated (i) annual observations taken from the populations, such as commercial catch-at-age data and tuning series, (ii) annual stock assessments by XSA (extended survivors analysis; Darby and Flatman, 1994) and associated catch forecasts, and (iii) annual management decisions depending on the scenario. The XSA method, a calibrated variant of virtual population analysis (VPA), is currently used to assess flatfish (ICES, 2007).

In the MP, the annual "target F" (the fishing mortality averaged over ages 2–8, i.e. F_{2-8}) was set according to the F_{pa} determined by ICES: 0.4 for sole and 0.3 for plaice (ICES, 2004). In the TAC scenario, the catch forecast under F_{pa} of each species was set as the respective TAC. In the effort scenario, the catch forecasts under F_{pa} were translated into an allowable effort in the economic submodel (see below). All options in the simulated assessment and forecast procedure matched those used by ICES (2004), and settings were maintained for the entire simulation period. This is contrary to common practice of ICES Working Groups, which may make small changes to the settings each year.

The simulation consisted of a historic part spanning the years from 1957 to 2002 and a projected part from 2003 to 2015. Therefore, 2002 was the first year in which an assessment was carried out leading to a management decision for the next year. From 2003, F was affected by the management decision made the year before, through a feedback loop. The technical interaction between sole and plaice in the Dutch beam trawl fishery was assumed to reflect the linkage between the two species in all fisheries. Neither catching nor discarding of undersized fish was assumed to take place. Appendix B gives a detailed description of the input data in the basic model.

Economic submodel

The economic submodel was constructed to calculate the annual removals by the fishery, as well as costs and revenues. Economic parameters were obtained from an analysis of logbook data and data on costs and revenues by the LEI panel, a sample representing 25% of the Dutch fleet. The price-elasticity parameters of sole and plaice were based on Nielsen (1999). For the effort scenario, the submodel also functioned as a means to calculate the management decision, i.e. the allowable effort, from the catch forecasts under F_{pa}. The calculations were done through Cobb–Douglas production functions, the derivation of whose parameters is explained below. The spawning-stock biomasses of the two species and the Dutch portions of their respective catch forecasts in each year served as input for the economic submodel, which then calculated the associated annual true catches (see below). The calculated catches (after raising them back from the Dutch to the international level) were output and fed back into the basic model, as realized removals in the OM. Revenues were calculated from the landed catches and the prices per kilogramme, the latter being dependent on the landings, but subject to price elasticity. The price-elasticity model causes revenues to fluctuate less than landings. The costs were calculated from the effort deployed. Equations and parameters are listed in Appendix A.

The economic submodel was designed as a short-term model to predict adjustments within the existing fleet in response to management policies. For individual vessels, responses would consist of adjustments in seasonal or spatial effort allocation, in the use

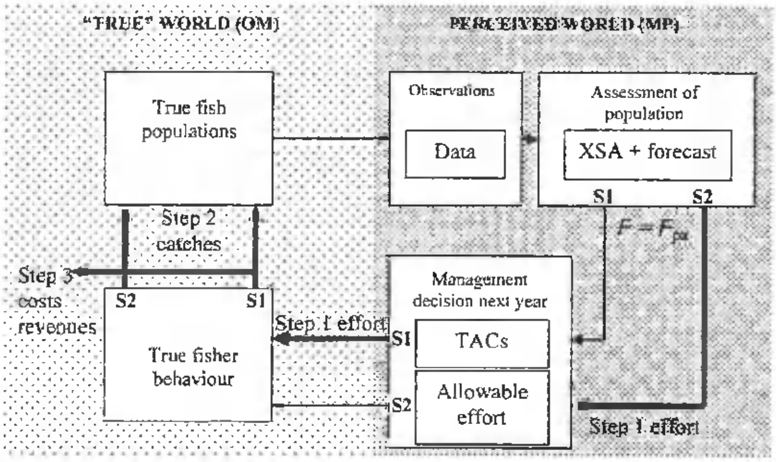

Figure 1. Flow diagram of the simulation. The thick arrows mark where calculations in the economic submodel are carried out, denoted by steps 1, 2, and 3, as explained in the main text. S1 and S2 denote scenarios 1 and 2, respectively.

of gears, or in the number of effective sea-days. Fishers were assumed to maximize their net revenues per unit of the restricting factor. In the TAC scenario, the restricting factor was taken to be the weighted value of the individual transferable quotas (ITQs) of sole and plaice, a value calculated by multiplying the sole and plaice ITQs with fixed average prices, then summing the products. Hence, fishers were assumed to maximize their net revenues, divided by the weighted value of the ITQs they used during that trip. In the effort-restriction scenario, fishers were assumed to maximize their net revenues per day at sea. The basic assumption was that with a management restriction on the fishery, the least efficient trips (those with lowest net revenues per unit of the restricting factor) would be dropped first.

To calculate the parameters for the submodel, we used the Dutch reported landings and effort data and corresponding economic data by fishing trip for 2002, but to reduce the number of records in the database, individual trips were grouped by vessel, month, and gear. These vessel–month–gear groups were sorted according to descending efficiency, i.e. descending net revenues per unit of the restricting factor. Subsequently, all records were cumulated, and landings were regressed against fishing effort. This relationship between effort and landings of each species led to a different production function for each management scenario, because fishers were expected to follow a different strategy according to the type of restriction.

The exponential regression of landings on fishing effort was designed as a classical Cobb–Douglas production function of the form

$$L_i = \alpha_i B_i E^{\beta}, \tag{1}$$

where L denotes landings, E the fishing effort in horsepower-days (hp-days), B the spawning-stock biomass, i the species, and α and β are constants ($\beta < 1$). The actual equations and the estimated parameter values are listed in Appendix A. An important feature of this production function is that the rate of increase in landings decreases as effort increases. This reflects the assumption that if the fishery is restricted, the least efficient trips would be dropped first.

One assumption underlying the economic submodel is that the efficiency of month–vessel–gear groups is not distributed randomly, and that fishers may decide to cancel specific month–vessel–gear combinations to maximize net revenues. An analysis of variance of the net revenues per hp-day was performed to test this hypothesis. Selected explanatory variables were month, gear, interaction of month and gear, available hp-days per year, vessel, hp-group of vessel, interaction of gear and hp-group, and interaction of hp-group and available hp-days. For the entire fleet, 56% of the variance could be explained by these variables (Table 1), vessel, hp-group, and gear being the most important. For individual hp-groups (Table 2), the variance explained varied between 53% for Euro-cutters (engine power <300 hp) and 79% for vessels >2000 hp, and the most important explanatory variables were vessel and month. These results make it plausible that the industry should be able to cancel the least efficient trips when they faced a restriction. Therefore, net revenues would decrease less than proportionally with the restriction, by concentrating trips in the most efficient seasons, and by trading quota or hp-days from less-efficient vessels to more-efficient ones. In the long term, this may cause less-efficient vessels to be withdrawn from the fishery.

The variability in net revenues and catches of plaice and sole among groups of trips is probably less than among individual trips. This means that we may be underestimating the potential for fishers to optimize their behaviour, and hence the curvature of the production functions for the two species.

Scenarios

Scenario 1: TAC management

In this scenario, the two species are managed independently by single-species TACs. We make the assumption that the fisheries primarily target sole (the most valuable species by a factor of 4–5 in terms of price per kilogramme), and that fishing continues until the sole TAC has been fully taken, irrespective of the plaice TAC. In other words, exploitation of sole determines the behaviour of the fleets, resulting in under- or overexploitation of the plaice TAC. We assume that overquota catch of plaice is not landed, and therefore not accounted for in the assessment,

Table 1. Analysis of variance for the whole fleet: test of between-subjects effects with the dependent variable being net revenues per hp-day.

Source	Type 1 sum of squares	d.f.	Mean square	F	Significance
Corrected model	524 751[a]	60	8 702.3	9.95	<0.01
Intercept	209 706	1	209 706.0	239.80	<0.01
hp-group	235 545	5	47 109.1	53.87	<0.01
Gear	126 310	2	63 155.0	72.22	<0.01
Month	41 056.9	1	3732.4	4.26	<0.01
Gear * hp-group	106 866	6	17 811.0	20.36	<0.01
Gear * month	88 309.4	2	4014.0	4.59	<0.01
hp-group * hp- days	17 023.8	6	2837.3	3.24	<0.01
Vessel	463 240	55	8407.2	9.61	<0.01
Error	409 483	468	874.4	–	–
Total	955 206	528	–	–	–
Corrected total	934 235	528	–	–	–

[a] $r^2 = 0.562$ (adjusted $r^2 = 0.505$).

leading to a discrepancy between the true catch and the perceived catch (=landings). This scenario is thought to approximately reflect the current situation, although it clearly is a simplification: not all fleets target sole, and a small part of the Dutch beam trawl fleet even exploits the area north of 55°N, where sole are virtually absent and the target species is plaice.

The plaice catch taken is calculated from the effort required to deplete the sole TAC (see below), and may be below or exceed the plaice TAC. In both cases, the calculated plaice catch is fed back into the OM as the true catch. If the plaice catch exceeds the TAC, the landings (=perceived catch) in the MP are equal to the TAC, otherwise landings are equal to the true catch. The age distribution of the landings is assumed to be equal to the age distribution of the true catch.

Scenario 2: effort management

For this scenario, management is based on the allowable effort. We decided to simulate a management rule where the allowable effort is determined by the species for which the catch forecast imposes the most severe effort restriction. Hence, this effort is the lower of the two values estimated to yield the predicted catch under the target F for each species. Of course, managers may decide to apply a different type of effort-management regime, e.g. one where the allowable effort is determined by the catch forecast of the most valuable species, or of the species that needs least restriction. We did not run simulations of such management regimes.

The respective catches taken with the allowable effort serve directly as input for the true catch in the OM, as well as for the landings in the MP. In contrast to control by TACs, where discarding of overquota catch is implied under full compliance, if the quotas do not match, compliance under effort control implies that all fish caught may be landed.

Calculations

Because the parameters of the economic submodel are based on the Dutch fishery only, it was necessary to scale the forecast catches down to the Dutch portions of the total international catches. Similarly, the catches subsequently taken with Dutch effort had to be scaled up again to arrive at international catches. The proportions of the international sole and plaice catches taken by the Dutch fishery were assumed to be constant at 74% and 45%, respectively (averages over the period 1995–2002; ICES, 2004).

The calculation of annual "true" catches and economic results was performed in three steps.

(i) Using the inverse production functions (listed in Appendix A), effort was calculated from the spawning-stock biomass (B) and the Dutch part of the catch forecast under the target F. In scenario 1, the true effort equals the effort required to deplete the sole TAC, using the true B in the calculation. In scenario 2, the true effort equals the allowed effort (the lower estimate of the effort required to deplete the forecast catch of the two species), using the perceived B to mimic a management process in which only estimates are available, but not the true values.

(ii) True Dutch catches taken with the effort determined in step (i) were calculated, using the true B in the production functions listed in Appendix A. In scenario 1, only the plaice catch had to be calculated, because the sole catch was set equal to the TAC. The true Dutch catches were then raised to true international catches.

(iii) Prices, costs, revenues, and profits were calculated from Dutch landings and effort, using the costs and revenues functions listed in Appendix A. In scenario 2, landings were always equal to catches, whereas in scenario 1, potential overquota catches for plaice were not landed.

Uncertainty

Uncertainty was explicitly incorporated into the simulation framework, acknowledging the presence of a variety of sources (Rosenberg and Restrepo, 1994). In the simulation, this uncertainty caused the perception to deviate from the true world. Sources of error include: process error, attributable to natural variation in dynamic processes (e.g. recruitment); measurement error, generated when collecting observations from the populations; estimation error, arising from estimating parameters of the dynamic process during the assessment process; model error, because the true complexity of the dynamics can never be captured; and implementation error, attributable to imperfect implementation, e.g. when TACs are exceeded. In our case, model error was minimal because the equations used to construct the simulated populations were the same as those used to assess them. Variation among Monte Carlo runs came from two sources: random variability in recruitment around a Ricker relationship, and sampling error in the generation of the tuning series. Because the time-series of stock and recruitment data since 1957 do not provide evidence for one particular stock–recruitment relationship, the choice of a Ricker relationship was arbitrary, and we did not test robustness to alternative hypotheses (Machiels *et al.*, 2007; Pastoors *et al.*, 2007). Considering that recruitment variation is the major source of process error, we did not include additional forms, such as variability in weight-at-age or selectivity (see Pastoors *et al.*, 2007).

Table 2. Analyses of variance for hp-groups: test of between-subjects effects with the dependent variable being net revenues per hp-day.

hp-group	Source	Type 1 sum of squares	d.f.	Mean square	F	Significance
0 – 260	Corrected model	5 061 141.5[b]	263	19 243.884	5.098	< 0.001
	Intercept	414 947.950	1	414 947.950	109.934	< 0.001
	Gear	233 166.043	2	116 583.021	30.887	< 0.001
	Month	124 604.387	11	11 327.672	3.001	0.001
	Gear * month	102 830.578	22	4810.481	1.274	0.178
	hp-days	14 199.940	1	14 199.940	3.762	0.053
	Vessel	4 583 340.5	227	20 190.927	5.349	< 0.001
	Error	4 000 976.7	1060	3774.506	–	–
	Total	9 477 066.2	1324	–	–	–
	Corrected total	963.378	1323	–	–	–
261 – 300	Corrected model	21 890.617[d]	203	107.836	10.254	< 0.001
	Intercept	27 472.774	1	27 472.774	2612.356	< 0.001
	Gear	1279.602	2	639.801	60.838	< 0.001
	Month	4493.482	11	408.498	38.844	< 0.001
	Gear * month	2830.191	22	128.645	12.233	< 0.001
	hp-days	546.835	1	546.835	51.998	< 0.001
	Vessel	12 740.507	167	76.290	7.254	< 0.001
	Error	19 749.935	1878	10.516	–	–
	Total	69 113.326	2082	–	–	–
	Corrected total	41 640.552	2081	–	–	–
301 – 800	Corrected model	574.644[e]	32	17.958	4.957	< 0.001
	Intercept	145.440	1	145.440	40.150	< 0.001
	Gear	16.036	1	16.036	4.427	0.037
	Month	170.649	11	15.514	4.238	< 0.001
	Gear * month	37.386	7	5.341	1.474	0.181
	hp-days	27.998	1	27.998	7.729	0.006
	Vessel	322.575	12	26.881	7.421	< 0.001
	Error	492.649	136	3.622	–	–
	Total	1212.733	169	–	–	–
	Corrected total	1067.293	168	–	–	–
801 – 1500	Corrected model	154.004[f]	35	4.400	5.228	< 0.001
	Intercept	141.502	1	141.502	168.114	< 0.001
	Gear	50.556	1	50.556	60.064	< 0.001
	Month	27.557	11	2.505	2.976	0.002
	Gear * month	22.274	11	2.025	2.406	0.010
	hp-days	0.209	1	0.209	0.248	0.619
	Vessel	53.407	11	4.855	5.768	0.000
	Error	93.430	111	0.842	–	–
	Total	388.936	147	–	–	–
	Corrected total	247.433	146	–	–	–
1501 – 2000	Corrected model	738.795[c]	98	7.539	27.928	< 0.001
	Intercept	1600.308	1	1600.308	5928.560	< 0.001
	Gear	50.058	1	50.058	185.445	< 0.001
	Month	369.071	11	33.552	124.298	< 0.001
	Gear * month	10.119	5	2.024	7.497	< 0.001
	hp-days	13.296	1	23.296	86.302	< 0.001
	Vessel	286.251	80	3.578	13.256	< 0.001
	Error	224.583	832	0.270	–	–
	Total	2563.686	931	–	–	–
	Corrected total	963.378	930	–	–	–

Continued

Table 2. *Continued*

hp-group	Source	Type 1 sum of squares	d.f.	Mean square	F	Significance
>2000	Corrected model	607.188[a]	68	8.929	30.274	<0.001
	Intercept	934.379	1	944.379	3 201.857	<0.001
	Gear	27.691	1	27.691	93.885	<0.001
	Month	210.317	11	19.120	64.824	<0.001
	Gear * month	0.168	2	9.309E−02	0.316	0.729
	hp-days	3.001	1	3.001	10.176	0.002
	Vessel	365.993	53	6.906	23.413	<0.001
	Error	166.645	565	0.295	–	–
	Total	1718 213	634	–	–	–
	Corrected total	773.834	633	–	–	–

[a] $r^2 = 0.785$ (adjusted $r^2 = 0.759$).
[b] $r^2 = 0$ 0.558 (adjusted $r^2 = 0$ 0.449).
[c] $r^2 = 0$ 0.767 (adjusted $r^2 = 0$ 0.739).
[d] $r^2 = 0$ 0.526 (adjusted $r^2 = 0$ 0.474).
[e] $r^2 = 0.538$ (adjusted $r^2 = 0$ 0.430).
[f] $r^2 = 0$ 0.622 (adjusted $r^2 = 0$ 0.503).

Results

Scenarios

In both scenarios, the plaice stock recovered and generally stayed above $B_{\rm lim}$, whereas the sole stock generally remained above $B_{\rm lim}$ (Figure 2). The performance therefore conformed to the management goal of the precautionary approach, where the advice should ascertain that stocks would not drop below these limit reference points. In scenario 1, however, the initial rise of the spawning stock was followed by a decline, in contrast to a continuing rise in scenario 2. Moreover, F could rise steeply in scenario 1 when stock sizes were already decreasing, requiring drastic reductions subsequently. In contrast, F remained stable in scenario 2. This stability closer to the target F resulted in greater population growth than in scenario 1, and discrepancies between the perceived and true populations were smaller.

In scenario 1, there was overquota fishing of plaice throughout the simulation period in at least some Monte Carlo runs (Figure 3), resulting in the perceived catch (=landings) being lower than the true catch. This is the consequence of target Fs for sole and plaice not corresponding to similar effort levels (with the assumption that the sole TACs are adhered to), leading to conflicting management objectives. Dutch effort increased massively, but it declined towards the end of the simulation period (Figure 4).

In scenario 2, the effort restriction was more often (except the first year) determined by the plaice assessment than by that for sole (Figure 5). Despite the effort restriction alternating between being plaice- or sole-determined, the resulting effort level was fairly constant over the years (Figure 4). The conflicting level of effort required to meet the target Fs of the two species necessarily led to perceived "underexploitation" of one of the species (often alternating between the species among years). This term is used here to mean exploitation below the target catch forecast, whereas "overexploitation" means exploitation exceeding the target forecast. Formally, there should not be overexploitation in scenario 2, because the lower of two efforts was always used. However, there was sometimes overexploitation because the MP calculated allowable effort based on the predicted spawning-stock biomass, whereas the true spawning stock was sometimes larger, leading to higher catches than predicted.

In scenario 1, simulated prices first decreased and then rose, whereas they dropped first and then stabilized in scenario 2 (Figure 4). Net revenues followed the development of the landings (mirroring the development of the prices): in scenario 1, they increased but then declined, whereas they continued to rise in scenario 2 (Figure 4).

Comparative evaluation

We selected a range of evaluation criteria (Table 3) to review and compare the results of the two scenarios. In terms of the sustainable exploitation of the stocks, both management regimes generally led to avoidance of limit reference points. Scenario 2 performed slightly, but significantly, better than scenario 1: for sole, scenario 2 never violated $B_{\rm lim}$, whereas scenario 1 did so occasionally. For plaice, scenario 2 violated $B_{\rm lim}$ less often than scenario 1. Mean landings did not differ significantly, but the annual variation in sole landings was lower in scenario 1, and in plaice landings lower in scenario 2. From an economic perspective, scenario 2 appeared to perform better than scenario 1, with significantly higher net revenues in both the short and the long term. These higher net revenues seem to be due to lower variable costs at lower effort. Although net revenues over the whole period were similar, the net revenues in scenario 2 did not decline during the latter part of the simulation period, suggesting that if the simulation period had been prolonged, scenario 2 would have also yielded higher net revenues over the whole period. Economic stability between years was similar in the two scenarios. With respect to management, scenario 2 was more stable, requiring effort restrictions to vary annually by only 4% on average, whereas TACs varied annually by 14% and 12% for sole and plaice, respectively.

Shrinkage

In scenario 1, the perception of stock development lagged several years behind their true development (Figure 2). This can be seen more easily when the true population and the perceived population estimates from each Monte Carlo run are compared (Figure 6a). Because of this lag, there was a cyclical alternation of underestimation and overestimation.

We investigated whether this lag was caused by shrinkage, a technical setting in the XSA assessment model affecting the estimated F in the most recent years. The setting used traditionally in routine ICES assessments of North Sea plaice and sole is

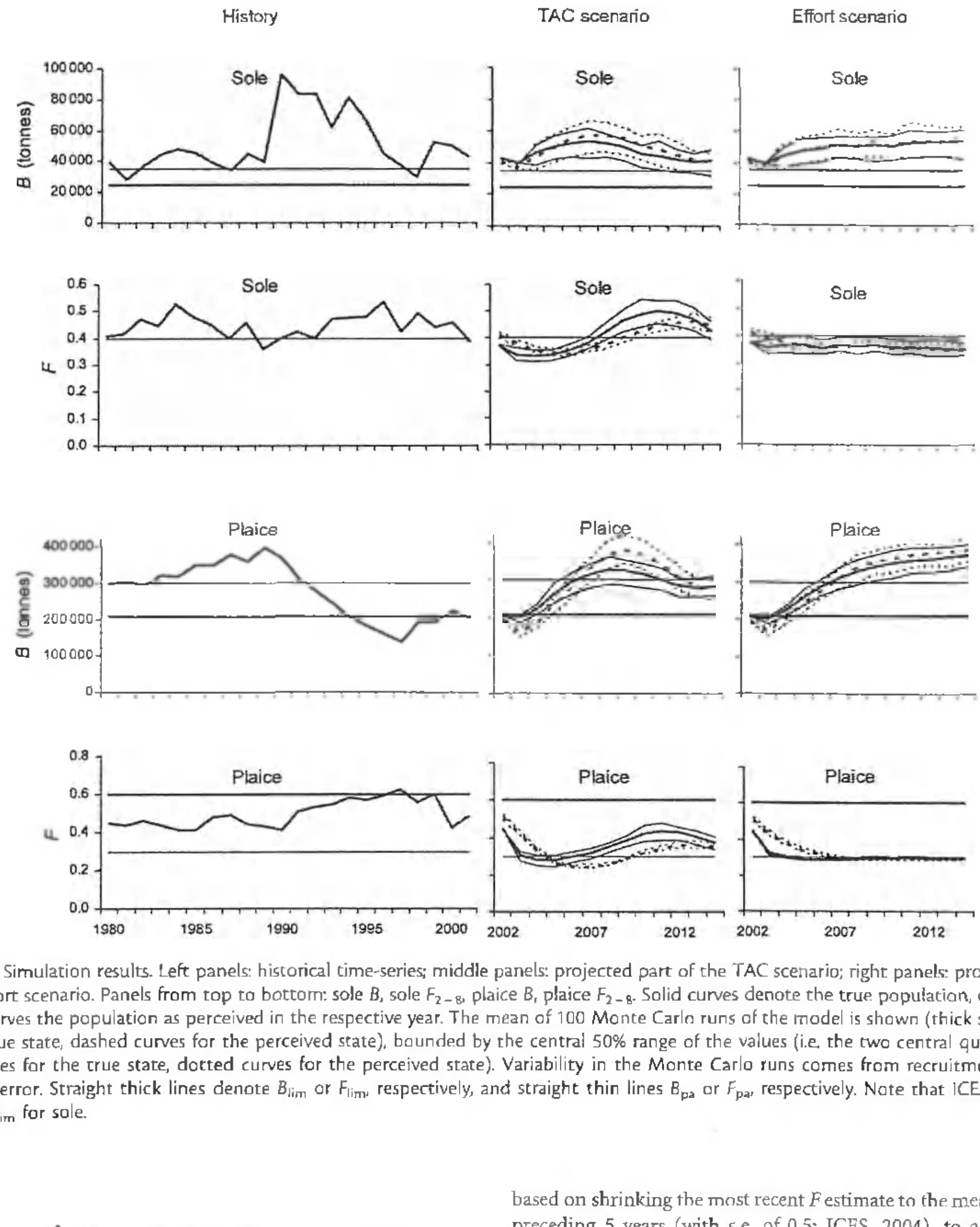

Figure 2. Simulation results. Left panels: historical time-series; middle panels: projected part of the TAC scenario; right panels: projected part of the effort scenario. Panels from top to bottom: sole B, sole F_{2-8}, plaice B, plaice F_{2-8}. Solid curves denote the true population, dashed and dotted curves the population as perceived in the respective year. The mean of 100 Monte Carlo runs of the model is shown (thick solid curves for the true state, dashed curves for the perceived state), bounded by the central 50% range of the values (i.e. the two central quartiles; thin solid curves for the true state, dotted curves for the perceived state). Variability in the Monte Carlo runs comes from recruitment and sampling error. Straight thick lines denote B_{lim} or F_{lim}, respectively, and straight thin lines B_{pa} or F_{pa}, respectively. Note that ICES has not defined F_{lim} for sole.

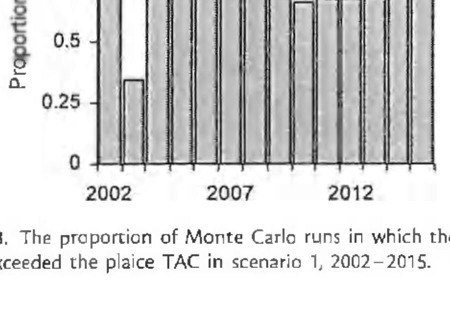

Figure 3. The proportion of Monte Carlo runs in which the plaice catch exceeded the plaice TAC in scenario 1, 2002–2015.

based on shrinking the most recent F estimate to the mean over the preceding 5 years (with s.e. of 0.5; ICES, 2004), to account for uncertainties in the non-converged part of the VPA. The choice of a specific degree of shrinkage reflects a trade-off between uncertainty and bias in the estimate of F. Therefore, the consequence of strong shrinkage is that actually realized changes in F in recent years are partly overruled by values of F in the more distant past. For comparison, we carried out 100 Monte Carlo runs with identical settings, except that F was shrunk less strongly towards the recent average (over 3 years with s.e. = 2.0). In this case (Figure 6b), the discrepancy between perceived and true population size was reduced markedly and the cycles disappeared. Therefore, the lag in perception of the true population development and the cyclical alternation of under- and overestimation appear to be caused by the strong shrinkage setting. Under low

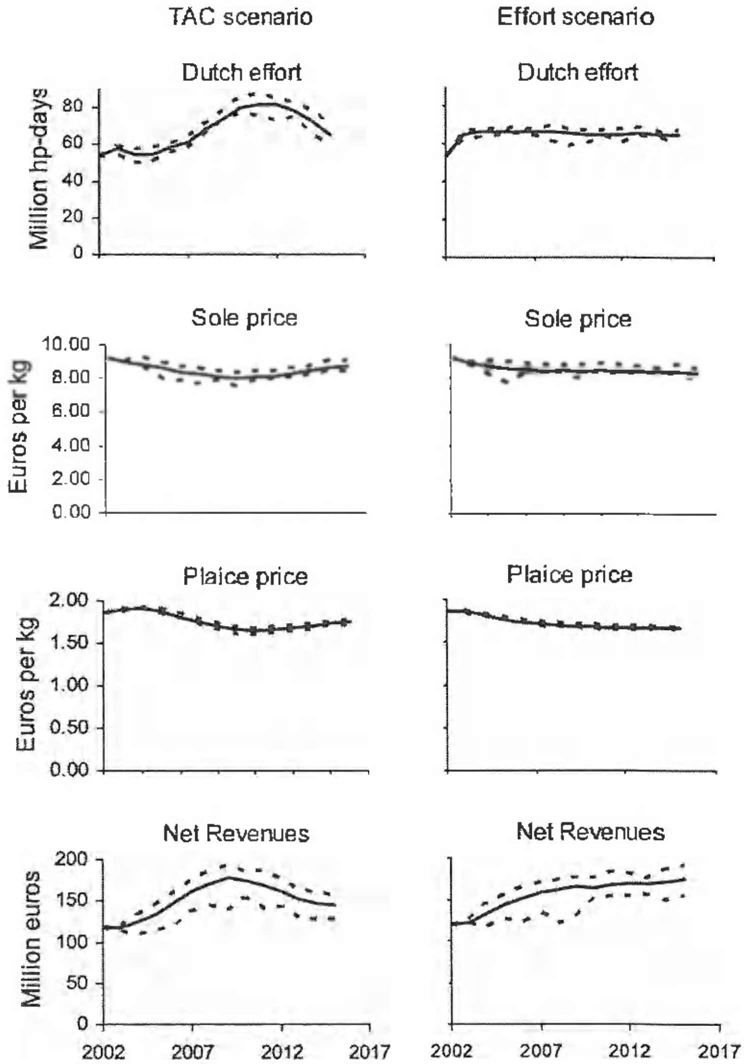

Figure 4. Simulation results. Left panels: projected part of the TAC scenario; right panels: projected part of the effort scenario. Panels from top to bottom: Dutch effort, Dutch sole price, Dutch plaice price, Dutch net revenues. Solid curves: mean; dotted curves: upper and lower quartiles.

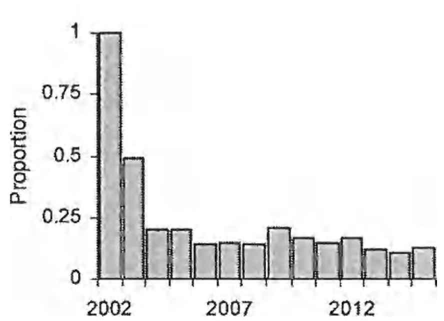

Figure 5. The proportion of Monte Carlo runs in which the sole catch forecast determined the effort restriction in scenario 2, 2002–2015.

shrinkage, spawning stock and F of the two species remain more stable (Figure 7). Although scenario 2 was run with the same strong shrinkage as scenario 1, there was no indication of cycles, because management stabilizes effort directly, and if F does not vary over time, shrinking to the mean has no additional effect.

Discussion

Assumptions

The construction of the true population was based on similar assumptions and the same equations as those used in the assessment of the perceived population. This may easily lead to overconfidence in the results, and must be acknowledged as a possible weakness in this type of analysis.

Moreover, the true natural mortality, maturity-at-age, and historical weights in the OM are assumed to be known in the MP, which is, of course, never the case. The same applies to the true

Table 3. Performance with respect to biological, economic, and management evaluation criteria for the two scenarios.

Parameter	Scenario 1 TAC	Scenario 2 effort restriction
Biological criteria for sole		
Frequency $F > F_{pa}$ (target F; %)	51.6	**15.2**
Frequency $B < B_{lim}$ (%)	1.9	**0.0**
Mean (true) landings ('000 t)	266	258
Mean absolute difference in landings between two consecutive years (%)	**14.3**	16.5
Biological criteria for plaice	–	–
Year of recovery ($\geq$75% of runs with $B > B_{lim}$)	2004	2004
Frequency $F > F_{pa}$ (target F; %)	74.2	**36.0**
Frequency $B < B_{lim}$ from 2004 on (%)	4.5	**1.8**
Mean (true) landings ('000 t)	1130	1209
Mean absolute difference in landings between two consecutive years (%)	11.5	**7.2**
Economic criteria	–	–
Mean net short-term revenues (2003–2005; €million)	378	**400**
Mean net long-term revenues (2013–2015; €million)	444	**514**
Overall mean net revenues (2003–2015; €million)	1984	2049
Mean absolute difference in net revenues between two consecutive years (%)	9.3	9.2
Management criteria	–	–
Mean absolute difference in management measure between consecutive years (%)	Sole TAC: 14.3 Plaice TAC: 12.3	Allowable effort: **3.6**

Emboldened scores are significantly better than the score in the alternative scenario (t-test assuming equal variances, $p < 0.05$; proportions are arcsine-square-root transformed; no Bonferroni correction).

catch composition. The implicit but unrealistic assumption here is that the market-sampling programme gives precise estimates of the age composition of the landed catch. Another assumption is that the age composition is the same in the landed and the discarded overquota catches. This is an unrealistic assumption, because fishers may high-grade and selectively discard small fish of low value when approaching exhaustion of their quota. Consequently, the potential sources of variation and bias are underestimated in the approach, and the results may be overoptimistic with regards to how well the MP is able to monitor the true development of stocks. Moreover, there are other simplifications. For instance, historical recruitment series of sole and plaice appear to be correlated, whereas the projected recruitment of the two species in the simulations was uncorrelated.

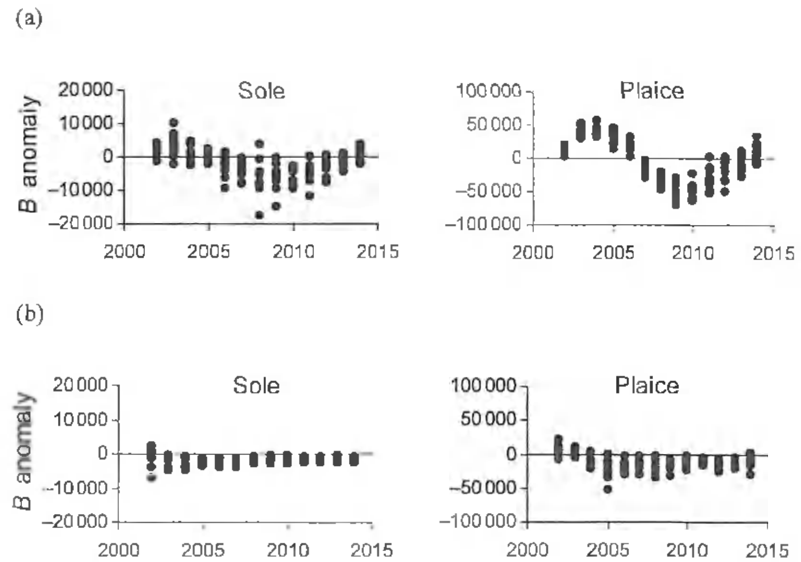

Figure 6. The difference between true and perceived estimates of B of 10 Monte Carlo runs of scenario 1. Left panels, sole; right panels, plaice. Positive values mean that the perceived population is underestimated compared with the true population. Variability in the Monte Carlo runs comes from recruitment and sampling error. (a) With strong shrinkage (over 5 years, s.e. = 0.5). (b) With weak shrinkage (over 3 years, s.e. = 2.0).

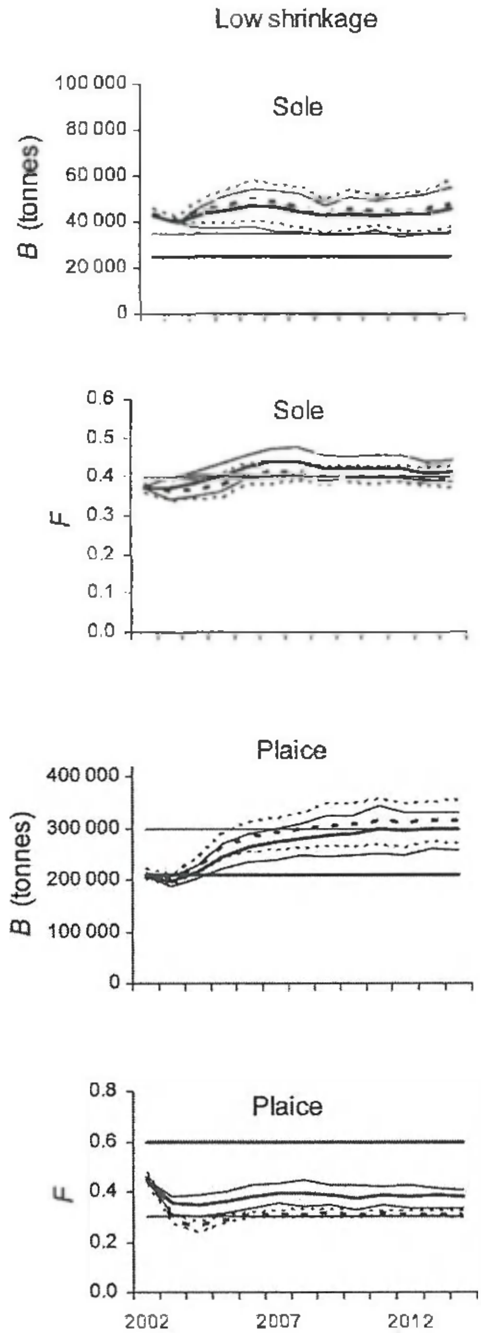

Figure 7. Scenario 1 with weak shrinkage (over 3 years, s.e. = 2.0). See legend of Figure 2 for full explanation.

An important restriction of this comparative analysis of two management scenarios is that many of the underlying data are based on the present behaviour of the fleets, which is a response to the present management system (single-species TACs). For instance, catch rates may change when single-species TACs are replaced by effort restrictions, because fishers will gain additional freedom to target individual species, and this may affect the direction in which a fishery develops. We have assumed that the fishery will not change from a sole-targeted one, but this cannot be guaranteed. In fact, the assumption that the sole TAC drives the plaice catch is a simplification, because a small part of the Dutch fleet fishes in an area where sole catches are negligible. Moreover, under effort restrictions, fishers may also maximize their profit by improving effort characteristics that are not restricted and increase the catchability (the so-called "technology creep" phenomenon; Pascoe *et al.*, 2001; Ulrich *et al.*, 2002; Rijnsdorp *et al.*, 2006). When this happens, parameters of the production functions change over time, and effective catches for a given amount of effort will differ from expected catches.

Another assumption relates to the behaviour of fishers with regard to high-grading. Under the existing single-species quota regime, catches may be higher than landings because of high-grading and discarding of overquota catches. Under effort restrictions, there is no incentive for high-grading and overquota discarding, so landings-per-unit-of-effort may increase. However, the production functions are based on (historical) landings data only, whereas they are used in the model for generating catch data. In reality, catches are likely to be bigger than the landings, which could cause underestimation of modelled catches.

The discarding of undersized plaice has been ignored in this analysis. Because of the different body shapes of sole and plaice, the selectivity of the prescribed 80 mm mesh size is very different for the two species. The 50% retention length in these nets is 27 cm for sole (minimum landing size 24 cm), but 18 cm for plaice (minimum landing size 27 cm). This implies that large quantities of undersized plaice are caught and discarded. Discard estimates were not available at the time of this study, but have recently been included in the plaice stock assessment even though their reliability is questionable (ICES, 2005). By neglecting the discarded fraction, the implicit assumption is that only fish above the respective minimum landing sizes are caught. Kell and Bromley (2004) pointed out that taking management decisions for North Sea plaice based on simulations without accounting for discarding leads to incorrect perceptions of stock development. However, the available estimates are based on crude extrapolations, and their incorporation might easily lead to unreliable results. Dickey-Collas *et al.* (2007) found that the inclusion of noisy discard estimates may make population trends hard to track with XSA. Although we acknowledge that assessments may be improved, this should not devalue the comparative value of the two scenarios assessed without these complications.

We also assumed full compliance to the management measures (allowing for legal discarding of overquota plaice catch in scenario 1). In reality, compliance with TACs may be a function of the discrepancies between the TACs set in consecutive years. Without directed research into fishers' responses to management measures, it is problematic to incorporate implementation bias into a model.

The technical interaction between sole and plaice in the Dutch beam trawl fishery is taken as a proxy for the linkage between the two species in all fisheries. This assumption imposes friction with another assumption that Dutch catches can be multiplied by fixed but different factors to arrive at the respective international catches (1.35 for sole, 2.25 for plaice). The friction arises because the different multipliers imply that, internationally, more plaice is caught than can be accounted for by the mixed-fishery interaction.

We further assumed that future management decisions were based on fishing at a target F that equals the respective F_{pa} (for both species in the TAC scenario, and for the species requiring stronger effort-restriction in the effort-regulation scenario). In reality, the EC management plan for North Sea plaice and sole adopted in 2007 and implemented for the first time in 2008 prescribes target *Fs* that are reduced by 10% each year until the final objectives of $F = 0.3$ for plaice and $F = 0.2$ for sole will have been reached. With this plan becoming operational, the stocks should recover more quickly and lower *Fs* would be realized. However, because our study was carried out before the new management plan existed, we did not incorporate any adjustments of target *F* in the simulations, which may have led to a more pessimistic development than the new plan would allow for. An earlier version of the current management plan has been evaluated by Machiels *et al.* (2007) and the Scientific, Technical, and Economic Committee for Fisheries (STECF, 2006), through simulation studies similar to this one.

Ulrich *et al.* (2002) undertook a comparable study on the mixed-flatfish fishery of the North Sea, also including economics. They modelled individual fleets, included discarding of undersized fish, and assumed that *F* is proportional to effort. Similarly, the simulation study that evaluated the draft EC management plan (Machiels *et al.*, 2007) assumed a proportional relation between effort and *F*. However, this assumption seemed problematic in the evaluation by STECF (2006). To our knowledge, a proportional relationship between effort and *F* has not been convincingly demonstrated. On the contrary, the relationship varies in response to technical, environmental, and behavioural factors (Rijnsdorp *et al.*, 2006). The last authors noted that through optimization behaviour of fishers, fishing mortality can be affected less than proportionally by effort reductions. We therefore believe that a production function based on the idea that a fishery faced with management restrictions will cancel the least profitable trips first is a more realistic approach. However, although we introduced economic data to mimic changes in fleet behaviour in response to management, the approach is still simplistic. In practice, species compositions of the catch may change outside the simulated range in response to the management strategy chosen.

Another limitation is the lack of sensitivity tests to evaluate the robustness of our conclusions by varying the various assumptions. It is clear that there is a tension between the level of detail that may seem desirable and the level that can be provided when specifying, conditioning, and validating OMs in management evaluation studies (Pastoors *et al.*, 2007).

To put all these assumptions into perspective, we stress that scenario 1 is supposed to reflect largely the management system as it has been in place until and including 2007. The outcomes of our simulations suggest that this system would guarantee that both stocks are managed effectively in avoiding the limit reference points, and that the TAC system is a safeguard against overexploitation. However, this statement contrasts sharply with the recent perceptions of the developments in both stocks, suggesting that scenario 1 is too optimistic as a representation of the recent regime.

Shrinkage

For many demersal stocks in northern Europe, the use of strong shrinkage of *F* towards the recent 5-year mean has been a common practice until recently. For North Sea plaice and sole, this practice has changed recently (ICES, 2007), when the results of the simulations reported here were made available to the assessment working group in an earlier draft (Kraak *et al.*, 2004). Our analysis shows that if *F* varies markedly, as in scenario 1, the use of strong shrinkage induces cyclical developments in the stock. This instability is caused by the perception of the stock being always out of phase with the true state. Shepherd (1999) commented that shrinking *F* to the recent mean might cause conflict with strong signals in the surveys, particularly if catch data are inaccurate. However, such a conclusion misses the point that even if the catch is well sampled and there are no strong signals in surveys, shrinkage may introduce a greater bias than the uncertainty originating from the unconverged part of the VPA that it is trying to avoid. Importantly, the strength of the shrinkage chosen reflects a trade-off between bias and uncertainty (Dickey-Collas *et al.*, 2007). In our simulations, many sources of uncertainty were absent (as explained above), but in ICES assessments, total uncertainty may be substantial.

The bias has marked effects on the advice and corresponding management measures. During a period with a strong trend in *F*, shrinkage results in a discrepancy in the perception of the true state of the stock (see also Dickey-Collas *et al.*, 2007). Therefore, the advice, and the measures taken, will be inappropriate relative to the reference points aimed for, and always lag behind, leading to overshoots and undershoots. These problems may explain the retrospective bias often observed in assessments. They may also lead to unwelcome and unnecessary variations in TACs, so undermining the credibility of the scientific advice (by causing large retrospective change) and prohibiting implementation of timely and stable management measures. As recovery plans or other conservative measures are often associated with strong reductions in *F*, as intended by the adopted EC management plan for North Sea plaice and sole, the perception of success of such plans would be jeopardized if shrinkage of *F* to the mean were applied.

Comparison of the two scenarios

This simulation exercise was undertaken to explore the methodology and to gain insight into the question of how alternative management scenarios may differentially affect stock development, the fishery, its economy, and its management. The projections are not meant to be viewed as quantitative stock forecasts or predictions to be used in North Sea flatfish management (Pastoors *et al.*, 2007). Rather, the two scenarios represent a qualitative comparison of two management systems for a particular set of simplifying assumptions. The conclusion as to which management scenario performs best should be viewed as conditional on many restrictive assumptions, because the conclusions may not hold if some of these were relaxed or changed.

Based on the evaluation criteria chosen, the effort-management scenario appears to be more suitable for managing the beam trawl fishery than the TAC scenario, because it results in a more positive biological development of both stocks, in higher short-term and long-term economic profits, and greater stability of annual management measures.

Although both scenarios generally led to sustainable exploitation (i.e. they kept *F* below F_{lim}, and the spawning stock above B_{lim}), effort management did so slightly better and led generally to larger stocks and lower effort. This is because the effort-restriction implemented was tuned to the species most in need of restriction, so acknowledging the mixed-species nature of the fishery. This often resulted in catches of the other species that were lower than might be taken with the single-species target F (F_{pa}). This, in turn, was favourable for the development of that stock, and was paid back later. However, other effort-management regimes can be envisaged and very different outcomes expected. For instance, a system aimed specifically at sustaining the fishery of the most valuable species could easily lead to overexploitation of the other.

Scenario 1, mimicking the current situation, performed worse because the sole catch determined effort, whereby the plaice stock was often overexploited and some plaice catches could not be landed and did not contribute to revenues. Still, this scenario is probably overoptimistic as a reflection of the current management regime, in view of the negative development of the stocks observed recently under the current management regime.

The development of the stocks was more stable under scenario 2, leading to lower assessment bias and error (lower discrepancy between perceived and true states of both stocks). Assessment bias caused by shrinkage reinforces instability. Conversely, the lack of assessment bias reinforces stability and makes the advice less sensitive to shrinkage. This should, in turn, lead to the setting of more appropriate measures (in this case, allowed effort).

To conclude, many of the differences between the scenario outcomes are attributable to the way the management regimes dealt with the conflicting target *F*s for sole and plaice and the inconsistent fleet effort required to execute these *F*s. Our study suggests that a management strategy that occasionally results in perceived underexploitation of the stocks may work best, given the existence of assessment error and bias. Our results also suggest that stability in fishing mortality reinforces itself, because assessment bias is lower under greater stability, and corrections become less necessary.

Acknowledgements

This study was financed by the Department of Fisheries of the Dutch Ministry of Agriculture, Nature and Food Quality (LNV) under contract Ond/2002-1/6c/1. We thank Laurie Kell for playing a large role in developing the simulation models and the concept of the paper, and Clara Ulrich, Trevor Hutton, and anonymous referees for their detailed comments on earlier versions of the manuscript.

References

Butterworth, D. S., and Punt, A. E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985–998.

Daan, N. 1997. TAC management in North Sea flatfish fisheries. Journal of Sea Research, 37: 321–341.

Darby, C. D., and Flatman, S. 1994. Virtual Population Analysis: Version 3.1 User Guide. MAFF Directorate of Fisheries Research, IT Report 1. 85 pp.

de Wilde, W. 1999. Measurement of economic impacts of fishery management decisions, The Dutch case (associated with the Klepper & Lasch work). EAFE Conference 1999, Dublin.

Dickey-Collas, M., Pastoors, M. A., and van Keeken, O. A. 2007. Precisely wrong or vaguely right: simulations of noisy discard data and trends in fishing effort being included in the stock assessment of North Sea plaice. ICES Journal of Marine Science, 64: 1641–1649.

EC. 2007. Council Regulation (EC) No 676/2007 of 11 June 2007 establishing a multiannual plan for fisheries exploiting stocks of plaice and sole in the North Sea. http://eur-lex.europa.eu/LexUriServ/site/en/oj/2007/l_157/l_15720070619en00010006.pdf.

Frost, H. 1997. Transition from quota to effort regulation. South Jutland University Press, DIFER Report Series 23/97. Report with Financial Assistance from the EU Commission DG XIV (DG14.A-94/03).

Harwood, J., and Stokes, K. 2003. Coping with uncertainty in ecological advice: lessons from fisheries. Trends in Ecology and Evolution, 18: 617–622.

ICES. 2001. Report of the ICES Advisory Committee on Fishery Management. ICES Cooperative Research Report, 242.

ICES. 2003. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) Copenhagen, Denmark, 11–20 June 2002. ICES Document CM 2003/ACFM: 02.

ICES. 2004. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) Boulogne sur Mer, France, 9–18 September 2003. ICES Document CM 2004/ACFM: 07.

ICES. 2005. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) Bergen, Norway, 7–16 September 2004. ICES Document CM 2005/ACFM: 07.

ICES. 2007. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK), ICES Headquarters, 5–14 September 2006. ACFM: 35. http://www.ices.dk/reports/ACFM/2006/WGNSSK/directory.asp.

Kell, L. T., and Bromley, P. J. 2004. Implications for current management advice for North Sea plaice (*Pleuronectes platessa* L.). 2. Increased biological realism in recruitment, growth, density-dependent sexual maturation and the impact of sexual dimorphism and fishery discards. Journal of Sea Research, 51: 301–312.

Kell, L. T., Kirkwood, G., Pilling, G., Abaunza, P., Aps, R., van Beek, F. A., Biseau, A. *et al.* 2002. Analysis of possibilities of limiting the annual fluctuations in TACs (MATES. Final Report FISH/2001/02/02). CEFAS, Lowestoft, UK.

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J-M., Garcia, D., Hillary, R., Jardim, E. *et al.* 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640–646.

Kell, L. T., Pilling, G. M., Kirkwood, G., Pastoors, M. A., Mesnil, B., Korsbrekke, K., Abaunza, P. *et al.* 2005. An evaluation of the implicit management procedure used for some ICES roundfish stocks. ICES Journal of Marine Science, 62: 750–759.

Kraak, S. B. M., Buisman, F. C., Dickey-Collas, M., Poos, J. J., Pastoors, M. A., Smit, J. P. G., Daan, N. *et al.* 2004. How can we manage mixed fisheries? A simulation study of the effect of management choices on the sustainability and economic performance of a mixed fishery. ICES Document CM 2004/FF: 11.

Machiels, M. A. M., Kraak, S. B. M., and Beek, F. A. van 2007. Evaluation of a management plan as proposed by the European Commission in 2006 for fisheries exploiting stocks of plaice and sole in the North Sea. Wageningen IMARES Report, C011/07. IJmuiden, the Netherlands. http://old-stecf.jrc.it/meetings/sgeca/0605/imares.pdf.

Nielsen, M. 1999. EU Seafood Markets—Integration and Demand, Working Paper No13/1999, SJFI, Copenhagen.

Pascoe, S., Andersen, J., and de Wilde, W. 2001. The impact of management regulation on the technical efficiency of vessels in the Dutch beam trawl fishery. European Review of Agricultural Economics, 28: 187–206.

Pastoors, M. A., Poos, J. J., Kraak, S. B. M., and Machiels, M. A. M. 2007. Validating management simulation models and implications for communicating results to stakeholders. ICES Journal of Marine Science, 64: 818–824.

Piet, G. J., and Rice, J. C. 2004. Performance of precautionary reference points in providing management advice on North Sea fish Stocks. ICES Journal of Marine Science, 61: 1305–1312.

Punt, A. E., Smith, A. D. M., and Cui, G. 2002. Evaluation of management tools for Australia's South East fishery. 1. Modelling the South East fishery taking into account technical interactions. Marine and Freshwater Research, 53: 615–629.

Rijnsdorp, A. D., Daan, N., and Dekker, W. 2006. Partial fishing mortality per fishing trip: a useful indicator of effective fishing effort in mixed demersal fisheries. ICES Journal of Marine Science, 63: 556–566.

Rosenberg, A. A., and Restrepo, V. R. 1994. Uncertainty and risk evaluation in stock assessment advice for U.S. marine fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 51: 2715–2720.

Sainsbury, K. J., Punt, A. E., and Smith, A. D. M. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES Journal of Marine Science, 57: 731–741.

Salz, P., and Frost, H. 2000. Model for economic interpretation of the ACFM advice (EIAA). *In* Proceedings of the XIIth Annual Conference of the European Association of Fisheries Economists (EAFE). Ed. by E. Lindebo, and N. Vestergaard. University of Southern Denmark.

Shepherd, J. G. 1999. Extended survivors analysis: an improved method for the analysis of catch at age data and abundance indices. ICES Journal of Marine Science, 56: 584–591.

STECF. 2006. Commission staff working paper, impact assessment: plaice and sole long-term management. STECF Sub-group SGECA-SGRST-06-05, Brussels, 26–29 September 2006. http://stecf.jrc.ec.europa.eu/20.

Sutinen, J. G. 1999. What works well and why: evidence from fishery-management experiences in OECD countries. ICES Journal of Marine Science, 56: 1051–1058.

Ulrich, C., Pascoe, S., Sparre, P. J., de Wilde, J-W., and Marchal, P. 2002. Influence of trends in fishing power on bioeconomics in the North Sea flatfish fishery regulated by catches or by effort quotas. Canadian Journal of Fisheries and Aquatic Sciences, 59: 829–843.

Appendix A: equations and symbols used in the model

Equations

Population dynamics

$$N_{i,a+1,y+1} = N_{i,a,y}e^{-Z_{i,a,y}} \tag{2}$$

$$N_{i,p,y} = N_{i,p-1,y-1}e^{-Z_{i,p-1,y-1}} + N_{i,p,y-1}e^{-Z_{i,p,y-1}} \tag{3}$$

Mortality rates

$$Z_{i,a,y} = F_{i,a,y} + M_{i,a,y} \tag{4}$$

$$F_{i,a,y} = S_{i,a,y}f_{i,y} \tag{5}$$

Catch equation

$$C_{i,a,y} = N_{i,a,y}\frac{F_{i,a,y}}{Z_{i,a,y}}(1 - e^{-Z_{i,a,y}}) \tag{6}$$

Stock–recruitment relationship (Ricker)

$$N_{i,r,y} = (\alpha_i B_{i,y-r}e^{-\beta_i B_{i,y-r}})e^{\varepsilon_{i,y}-\sigma^2/2} \tag{7}$$

$$\varepsilon_{i,y} \sim N(0, \sigma_i^2) \tag{8}$$

$$\sigma_i^2 = \ln(\mathrm{CV}_i^2 + 1) \tag{9}$$

Catch per unit effort models

$$U'_{i,f,a,y} = \frac{U_{i,f,a,y}}{A_{i,f,a,y}} \tag{10}$$

$$A_{i,f,a,y} = \frac{(e^{-\alpha_f Z_{i,a,y}} - e^{\beta_f Z_{i,a,y}})}{(\beta_f - \alpha_f)Z_{i,a,y}} \tag{11}$$

$$U'_{i,f,a,y} = q_{i,f,a}N_{i,a,y}e^{N(0,\varphi^2_{i,f,a})-\varphi^2_{i,f,a}/2} \tag{12}$$

Yield

$$Y_{i,y} = \sum_{a=r}^{p} C_{i,a,y}W_{C,i,a,y} \tag{13}$$

Spawning-stock biomass

$$B_{i,y} = \sum_{a=r}^{p} N_{i,a,y}W_{S,i,a,y}O_{i,a,y} \tag{14}$$

Single-species TAC scenario

$$C_{s,y} = Q_{s,y} \tag{15}$$

$$E_y = \alpha_0(M_{s,y})^{-\beta_0}Q_{s,y}^{\beta_0} \tag{16}$$

$$C_{p,y} = \alpha_{1,p}M_{p,y}E_y^{\beta_{1,p}} \tag{17}$$

Effort-restriction scenario

$$E_y = \mathrm{MIN}\Big[\Big(\alpha_{0,s}(M_{s,y})^{-\beta_0^s}Q_{s,y}^{\beta_0^s}\Big), \Big(\alpha_{0,p}(M_{p,y})^{-\beta_0^p}Q_{p,y}^{\beta_0^p}\Big)\Big] \tag{18}$$

$$C_{s,y} = \alpha_s M_{s,y}E_y^{\beta_s} \tag{19}$$

$$C_{p,y} = \alpha_p M_{p,y}E_y^{\beta_p} \tag{20}$$

Economic variables

$$P_{i,y} = P_i^0\left(\frac{L_{i,y}}{L_i^0}\right)^{e_i} \tag{21}$$

$$R_{O,y} = R_O^0\left(\frac{E_y}{E^0}\right) \tag{22}$$

$$R_{T,y} = \sum(L_{i,y}P_{i,y}) + R_{O,y} \tag{23}$$

$$C_{V,y} = \alpha_c + \beta_c E_y 1000 \tag{24}$$

$$R_{N,y} = R_{T,y} - C_{V,y} \tag{25}.$$

Symbols used in equations

	Parameter	Definition
Population dynamics	$N_{i,a,y}$	Numbers of fish of species i of age a at the start of year y
	p	Age of the plus group; here, $p = 15$ years
	$Z_{i,a,y}$	Total mortality of species i at age a in year y
Mortality rates	$M_{i,a,y}$	Natural mortality of species i at age a in year y
	$F_{f,a,y}$	Fishing mortality of species i at age a in year y
	$f_{f,y}$	Year effect of fishing mortality of species i in year y
	$s_{i,a,y}$	Selection pattern of species i at age a in year y
Catch equation	$C_{i,a,y}$	Catch in numbers of species i at age a in year y
Stock–recruitment relationship	r	Age at first recruitment to the fishery; here, $r = 1$ year
	$B_{i,y}$	Spawning-stock biomass of species i in year y
	α_i, β_i	Constants; for i = sole $\alpha_i = 5.1055$, $\beta_i = 0.0000168$, and for i = plaice $\alpha_i = 3.81$, $\beta_i = 0.00000331$ (estimated as explained in Appendix B)
	$\varepsilon_{i,y}$	Recruitment residual of species i in year y
	σ_i	Standard error of recruitment residuals of species i
	CV_i	Coefficient of variation of recruitment; for i = sole $CV_i = 0.5$, and for i=plaice $CV_i = 0.35$ (estimated as explained in Appendix B)
Catch per unit effort models	$U_{i,f,a,y}$	Cpue of species i by fleet f at age a in year y
	$U'_{i,f,a,y}$	Cpue of species i by fleet f at age a adjusted to start of year y
	$A_{i,f,a,y}$	Averaging factor of species i for fleet f relating the population abundance-at-age a during the time at which the catch was taken to the population abundance-at-age a at the beginning of the year y
	$q_{i,f,a}$	Catchability, relationship between cpue and numbers of species i for fleet f at age a (source explained in Appendix B)
	α_f	Start of the period of fishing of fleet f, at 0.66 of the year
	β_f	End of the fishing period of fleet f, at 0.75 of the year
	$\varphi_{i,f,a}$	Standard error of cpue residuals of species i of fleet f at age a (source explained in Appendix B)
Yield	$W_{C,i,a,y}$	Body mass in the catch of species i at age a in year y
	$Y_{i,y}$	Total catch mass of species i of all ages in year y
Spawning-stock biomass	$W_{S,i,a,y}$	Body mass in the stock of species i at age a in year y
	$O_{i,a,y}$	Proportion mature of species i at age a in year y
Single-species TAC scenario and effort-restriction scenario	E_y	Fishing effort in million hp-days in year y
	$M_{i,y}$	B of species i in year y relative to the B of that species in the reference year (2002)
	$Q_{i,y}$	Catch forecast (tonnes) of species i for year y, using the equation for yield Y
	$C_{i,y}$	Catch (tonnes) of species i in year y
	s (subscript)	Sole
	p (subscript)	Plaice
	α_0, β_0, $\alpha_{0,s}$, $\alpha_{0,p}$, β^s_0, β^p_0, α_s, β_s, α_p, β_p	Constants; $\alpha_0 = 0.0009786$, $\beta_0 = 1.177$, $\alpha_{0,s} = 0.00011698$, $\alpha_{0,p} = 0.00001882$, $\beta^s_0 = 1.4057$, $\beta^p_0 = 1.4578$, $\alpha_s = 628.637$, $\beta_s = 0.7105$, $\alpha_p = 916.432$ and $\beta_p = 0.8496$ in scenario 1, $\alpha_p = 1747.43$ and $\beta_p = 0.6853$ in scenario 2 (estimated from regressions as explained in the main text)
Economic variables	$P_{i,y}$	Dutch price in euros per kg of species i in year y
	$C_{V,y}$	Dutch variable costs in 1000 euros in year y
	α_c, β_c	Constants; $\alpha_c = -1.629$ and $\beta_c = 1.4355$ in scenario 1, $\alpha_c = 0.4715$ and $\beta_c = 1.4021$ in scenario 2 (estimated from regressions, see main text)
	$R_{T,y}$	Dutch total revenues in thousand euros in year y

Continued

Continued

Parameter	Definition
$R_{N,y}$	Dutch net revenues in thousand euros in year y
$L_{i,y}$	Dutch landings of species i (tonnes) in year y
L_i^0	Dutch landings of species i (tonnes) in reference year (2002); $L_s^0 = 10\,611$ and L_p^0 26 977 in scenario 1, $L_s^0 = 10\,611$ and L_p^0 26 668 in scenario 2 (from LEI data)
P_i^0	Dutch average price in euros per kg of species i in reference year (2002); $P_s^0 = 9.26$ and $P_p^0 = 1.86$ (from LEI data)
e_i	Price elasticity of species i; $e_s = -0.3$ and $e_p = -0.2$ (after Nielsen, 1999)
$R_{O,y}$	Dutch revenues from other species in thousand euros in year y
R_O^0	Dutch revenues from other species in reference year (2002) in thousand euros; $R_o^0 = 46\,385$ (from LEI data)
E^0	Dutch fishing effort in reference year (2002) in million hp-days; $E^0 = 53.569$ in scenario 1, $E^0 = 53.346$ in scenario 2 (from LEI data)

Appendix B: input data

The input data required for the OM are:

(i) Initial (1957) population numbers-at-age ($N_{i,a,y}$); from ICES (2004);

(ii) Natural mortality-at-age M ($M_{i,a,y}$); assumed constant (ICES, 2004);

(iii) Maturity-at-age ($O_{i,a,y}$); assumed constant (ICES, 2004);

(iv) Mean weight-at-age of fish in the stock ($W_{S,i,a,y}$) and in the catch ($W_{C,i,a,y}$); for the period 1957–2001 from ICES (2004); for the years 2002–2015 the running average over the previous 3 years;

(v) Selection pattern (relative F-at-age, $S_{i,a,y}$); for 1957–2001 absolute F-at-age values from ICES (2003); for 2002–2015 assumed constant as the average over the period 1997–2001;

(vi) Recruits ($N_{i,1,y}$); for 1957–2001 from ICES (2004); for 2002–2015 according to a Ricker model with a lognormal error distribution (determined by regressing recruits on SSB based on 1957–2001 data; standard deviations chosen through visual inspection of the resulting variation). Recruitment in the period 2002–2015 is a source of variability between simulation runs.

(vii) Yield ($Y_{i,y}$); for 1957–2002 from ICES (2004); from 2003 on the yield is determined by the management decision set in the previous year.

From these input data, the following quantities are calculated for both stocks by year: catch-at-age ($C_{i,a,y}$), population numbers-at-age ($N_{i,a,y}$), F-at-age ($F_{i,a,y}$), F_{2-8} ($F_{i,2-8,y}$), and spawning-stock biomass ($B_{i,y}$).

The input required for the MP (as required by XSA) is as follows.

(i) Perceived catch-at-age; equal to true catch-at-age in the OM for sole, and for plaice if the true catch did not exceed the TAC; otherwise equal to the TAC with the same age composition;

(ii) Perceived M; equal to true M in the OM;

(iii) Perceived maturity-at-age; equal to true maturity in the OM;

(iv) Perceived weight-at-age of fish in the stock and in the catch; averages of true values of the previous 3 years in the OM.

In addition, for each species, two catch per unit effort (cpue) series ($U'_{i,f,a,y}$) are generated for tuning, from true N-at-age, true M-at-age, true F-at-age, and catchabilities ($q_{i,f,a}$, see below), without a power model. Both series are set to take place in late summer, taking into account the proportion of the year the fish have been exploited (survey starts at 0.66 and ends at 0.75). Cpue series 1 commences in 1984 and involves ages 1–9, and series 2 commences in 1982 and involves ages 1–3. Catchabilities $q_{i,f,a}$ and their standard errors $\varphi_{i,f,a}$ are given in Table B1. These values are taken from existing research vessel survey tuning series (ICES, 2003), except for ages 2–9 of sole series 1, which are from a commercial fleet (ICES, 2004). The generation of tuning series with random error contributes a second source of variability between simulation runs.

In the MP, the annual assessment is performed by XSA, using the perceived data, with the settings for both stocks as used in ICES (2004) for plaice. For runs with low shrinkage, the s.e. has been set at 2 (as opposed to 0.5 for default runs) over 3 years (as opposed to 5 years for the default runs). The assessment generates values for perceived F-at-age, perceived N-at-age, perceived F_{2-8}, and perceived B up to the last data year. Short-term forecasts were run, under the assumption that F_{2-8} in the intermediate year equalled F_{2-8} in the previous year, and with the respective target

Table B1. Catchability ($q_{i,f,a}$) and standard error ($\varphi_{i,f,a}$) of cpue series 1 and 2 for sole and plaice.

Age a	Sole as species i				Plaice as species i			
	$q_{i,1,a}$	$\varphi_{i,1,a}$	$q_{i,2,a}$	$\varphi_{i,2,a}$	$q_{i,1,a}$	$\varphi_{i,1,a}$	$q_{i,2,a}$	$\varphi_{i,2,a}$
1	8.89	0.29	3.94	0.18	7.29	0.54	2.47	0.37
2	6.22	0.50	4.92	0.31	7.70	0.31	3.62	0.43
3	5.27	0.24	5.57	0.50	8.62	0.24	4.93	0.49
4	5.11	0.22	–	–	9.48	0.21	–	–
5	5.09	0.22	–	–	10.13	0.22	–	–
6	5.27	0.19	–	–	10.45	0.28	–	–
7	5.31	0.26	–	–	10.73	0.29	–	–
8	5.31	0.27	–	–	10.76	0.34	–	–
9	5.31	0.20	–	–	11.08	0.38	–	–

Fs (0.4 and 0.3 for sole and plaice, respectively) for the TAC year. The selection pattern in the TAC year and the intermediate year was assumed to be equal to the average over the previous 3 years. These forecasts generate a predicted *B* for the start of the TAC year and 1 year later and the forecast catch for the TAC year. Note that in reality in 2004 the age range for *F* as well as the pa reference points changed (ICES, 2005); we used the values before that change (ICES, 2004).

doi:10.1093/icesjms/fsn045

The effect of management choices on the sustainability and economic performance of a mixed fishery: a simulation study

S. B. M. Kraak, F. C. Buisman, M. Dickey-Collas, J. J. Poos, M. A. Pastoors, J. G. P. Smit, J. A. E. van Oostenbrugge, and N. Daan

Kraak, S. B. M., Buisman, F. C., Dickey-Collas, M., Poos, J. J., Pastoors, M. A., Smit, J. G. P., van Oostenbrugge, J. A. E., and Daan, N. 2008. The effect of management choices on the sustainability and economic performance of a mixed fishery: a simulation study. – ICES Journal of Marine Science, 65: 697–712.

Alternative management scenarios were evaluated in a simulation framework that mimicked the recent exploitation of sole and plaice in the North Sea. A large proportion of plaice is taken as bycatch of the beam trawl fleet targeting sole, yet current management of the two stocks assumes no interaction in their exploitation. The evaluation criteria included biological and economic sustainability, and stability in the management measures. The fishery was assumed to respond to management restrictions by dropping the least profitable trips. We investigated two contrasting management strategies, single-species total allowable catches, and effort regulation. Under the assumptions made, the latter strategy performed better. The results suggest that, given assessment error and bias, a strategy that accounts for the mixed nature of a fishery and that occasionally results in perceived underexploitation may work best. Stability in fishing mortality reinforces itself, through lower assessment bias, and management corrections become less frequent. The common assumption in many stock assessments in EC waters that fishing mortality in the most recent year should resemble the value obtained in previous years ("shrinkage") had a negative effect on the stability of control measures.

Keywords: economic performance, effort regulation, management strategy evaluation, mixed fishery, North Sea plaice, North Sea sole.

Received 14 December 2006; accepted 20 February 2008.

S. B. M. Kraak, M. Dickey-Collas, J. J. Poos, M. A. Pastoors, and N. Daan: Wageningen IMARES (Institute for Marine Resources and Ecosystem Studies), PO Box 68, 1970 AB IJmuiden, The Netherlands. F. C. Buisman, J. G. P. Smit, and J. A. E. van Oostenbrugge: Agricultural Economics Research Institute (LEI), PO Box 29703, 2502 LS Den Haag, The Netherlands. Correspondence to S. B. M. Kraak: tel: +31 (0)317 487177; fax: +31 (0)317 487326; e-mail: sarah.kraak@wur.nl.

Introduction

Management by total allowable catches (TACs) of species caught in mixed fisheries is problematic because the quota of different species may be exhausted at different rates. Fishers are therefore faced with a dilemma when the quota for one species is exhausted: stop fishing and underutilize the quota for other species, or continue fishing and discard or illegally land overquota fish (Daan, 1997). When they choose the latter option, the target fishing mortality (F) implied by the TAC will be exceeded, and the scientific basis for stock assessment and future management advice will be compromised if the stock assessment is based on official landings data alone (assuming that these equal the catches).

In the North Sea, a large proportion of the landings of plaice (*Pleuronectes platessa*) and sole (*Solea solea*) comes from the Dutch beam trawl fishery in the form of a mixed catch. However, exploitation of the two stocks is managed separately, and the TAC management does not account for their technical and economic interactions (Piet and Rice, 2004). Even the new EC management plan for North Sea plaice and sole (EC, 2007), which was adopted in 2007 and implemented for the first time in 2008, considers the TACs for the two species separately. Although the problem is well known, the solution is less obvious. An effort-management system is a possible alternative. Although the EC plan complements the two single-species TACs with effort limitations, it is currently not clear whether the fishery will be restricted by TACs or by effort limitations. Therefore, it is worth investigating the potential effects on the stocks and on the viability of the fisheries of a management system that restricts effort.

Simulations of exploited populations and management procedures provide insight into the sensitivities of a system to different management regimes, even if the ultimate management performance cannot be predicted (Butterworth and Punt, 1999; Sainsbury *et al.*, 2000; Kell *et al.*, 2002, 2005; Punt *et al.*, 2002; Ulrich *et al.*, 2002; Harwood and Stokes, 2003; Pastoors *et al.*, 2007). We developed a simulation framework containing a "true" population and a "perceived" population, as well as a feedback loop between the two. In the simulation, data collected each year from the true population serve as input to the annual stock assessments. As the data collection and the assessment procedures introduce error, the resulting perceived population deviates from the true population. It is the perceived population that drives the management decisions. These decisions are then imposed on the true population in the next year of the simulation.

Building on earlier work (Ulrich *et al.*, 2002; Machiels *et al.*, 2007; Pastoors *et al.*, 2007), we attempt to characterize the mixed North Sea flatfish fishery and to explore and understand the potential effects of alternative management regimes.

The simulation framework was based on Kell *et al.* (2002) and Kell and Bromley (2004), but with the addition of economic considerations. Although the economic impact of alternative management strategies has been analysed for various fisheries (e.g. Frost, 1997; Sutinen, 1999), simulations have rarely been used (de Wilde, 1999; Salz and Frost, 2000; Pascoe *et al.*, 2001; Ulrich *et al.*, 2002), because this requires economic data by fleet or preferably even by individual vessel and trip.

We report on two scenarios here, one reflecting current management by TACs, and one based on effort restrictions. The management goals for both scenarios were set according to the precautionary approach to fishery management, although the new EC management plan formulates different objectives, such as target fishing mortalities. The precautionary approach provides the framework for management advice delivered by the International Council for Exploration of the Sea (ICES, 2001). It requires that to have stocks and fisheries within safe biological limits, the probability should be high that the spawning-stock biomass [B (although the ICES jargon traditionally uses the symbol SSB for spawning-stock biomass, we use B for consistency, because B_{lim} and B_{pa} also refer specifically to spawning-stock biomass)] is above a limit value (called B_{lim}) below which recruitment may become impaired. In addition, the probability should be high that fishing mortality (F) is below a limit value (called F_{lim}) that will drive B to B_{lim}. Because of uncertainty in the annual estimates of B and F, ICES has defined more conservative (precautionary) operational reference points, B_{pa} and F_{pa} (the subscript pa standing for precautionary approach). When a stock is estimated to be above B_{pa}, the probability should be high that in reality it is above B_{lim}. Similarly, when F is estimated to be below F_{pa}, the probability should be high that in reality it is below F_{lim}. This management framework is therefore risk avoiding and does not target reference points such as MSY. In practice, F_{pa} is often used as a *de facto* target F, because the TAC advice is usually the catch that is forecast under F_{pa}. Therefore, the advice is simulated corresponding to F_{pa} and the TACs and effort restrictions are set accordingly. Note that, whereas in reality the advice is not always followed, in our simulations it is.

The objective of our study was to monitor the two management scenarios over annual time-steps, and to evaluate their performance with respect to biological (e.g. sustainable stock development), economic (e.g. net revenues), and management criteria (stability of management measures).

Methods

The model was implemented in the FishLab simulation framework [the version used here is held by the first author; the FishLab framework is no longer maintained, because it has been transferred to FLR (Kell *et al.*, 2007)], representing a set of dynamic link libraries (DLLs) that can be called from within Excel (Kell *et al.*, 2002, 2005). The equations are listed in Appendix A. Simulations were run in a "Monte Carlo" set-up ($\times 100$) to evaluate the variability in the final outcomes. The structure of the model is illustrated in the flow diagram (Figure 1).

Basic model

The model consisted of two main parts: the operating model (OM) simulated the true system, and the management procedure (MP) simulated the perceived system and associated management decisions. The OM represented two age-structured populations that mimicked North Sea sole and plaice, respectively. These populations were developing in annual time-steps from a starting population in 1957, given annual recruitment and mortality (F and natural mortality M). The MP simulated (i) annual observations taken from the populations, such as commercial catch-at-age data and tuning series, (ii) annual stock assessments by XSA (extended survivors analysis; Darby and Flatman, 1994) and associated catch forecasts, and (iii) annual management decisions depending on the scenario. The XSA method, a calibrated variant of virtual population analysis (VPA), is currently used to assess flatfish (ICES, 2007).

In the MP, the annual "target F" (the fishing mortality averaged over ages 2–8, i.e. F_{2-8}) was set according to the F_{pa} determined by ICES: 0.4 for sole and 0.3 for plaice (ICES, 2004). In the TAC scenario, the catch forecast under F_{pa} of each species was set as the respective TAC. In the effort scenario, the catch forecasts under F_{pa} were translated into an allowable effort in the economic submodel (see below). All options in the simulated assessment and forecast procedure matched those used by ICES (2004), and settings were maintained for the entire simulation period. This is contrary to common practice of ICES Working Groups, which may make small changes to the settings each year.

The simulation consisted of a historic part spanning the years from 1957 to 2002 and a projected part from 2003 to 2015. Therefore, 2002 was the first year in which an assessment was carried out leading to a management decision for the next year. From 2003, F was affected by the management decision made the year before, through a feedback loop. The technical interaction between sole and plaice in the Dutch beam trawl fishery was assumed to reflect the linkage between the two species in all fisheries. Neither catching nor discarding of undersized fish was assumed to take place. Appendix B gives a detailed description of the input data in the basic model.

Economic submodel

The economic submodel was constructed to calculate the annual removals by the fishery, as well as costs and revenues. Economic parameters were obtained from an analysis of logbook data and data on costs and revenues by the LEI panel, a sample representing 25% of the Dutch fleet. The price-elasticity parameters of sole and plaice were based on Nielsen (1999). For the effort scenario, the submodel also functioned as a means to calculate the management decision, i.e. the allowable effort, from the catch forecasts under F_{pa}. The calculations were done through Cobb–Douglas production functions, the derivation of whose parameters is explained below. The spawning-stock biomasses of the two species and the Dutch portions of their respective catch forecasts in each year served as input for the economic submodel, which then calculated the associated annual true catches (see below). The calculated catches (after raising them back from the Dutch to the international level) were output and fed back into the basic model, as realized removals in the OM. Revenues were calculated from the landed catches and the prices per kilogramme, the latter being dependent on the landings, but subject to price elasticity. The price-elasticity model causes revenues to fluctuate less than landings. The costs were calculated from the effort deployed. Equations and parameters are listed in Appendix A.

The economic submodel was designed as a short-term model to predict adjustments within the existing fleet in response to management policies. For individual vessels, responses would consist of adjustments in seasonal or spatial effort allocation, in the use

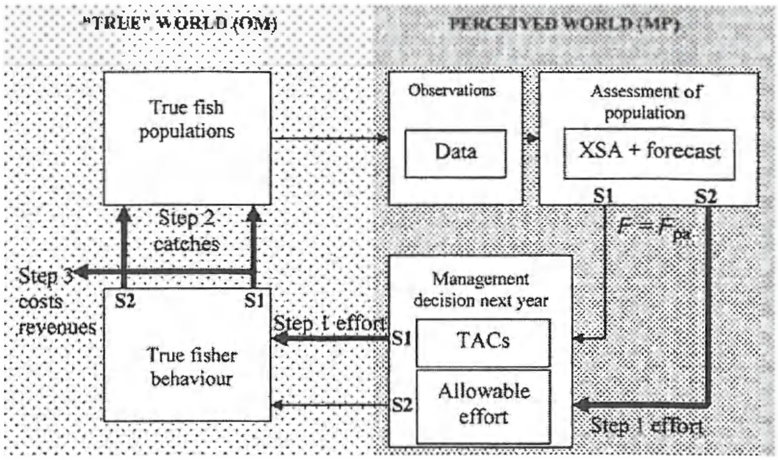

Figure 1. Flow diagram of the simulation. The thick arrows mark where calculations in the economic submodel are carried out, denoted by steps 1, 2, and 3, as explained in the main text. S1 and S2 denote scenarios 1 and 2, respectively.

of gears, or in the number of effective sea-days. Fishers were assumed to maximize their net revenues per unit of the restricting factor. In the TAC scenario, the restricting factor was taken to be the weighted value of the individual transferable quotas (ITQs) of sole and plaice, a value calculated by multiplying the sole and plaice ITQs with fixed average prices, then summing the products. Hence, fishers were assumed to maximize their net revenues, divided by the weighted value of the ITQs they used during that trip. In the effort-restriction scenario, fishers were assumed to maximize their net revenues per day at sea. The basic assumption was that with a management restriction on the fishery, the least efficient trips (those with lowest net revenues per unit of the restricting factor) would be dropped first.

To calculate the parameters for the submodel, we used the Dutch reported landings and effort data and corresponding economic data by fishing trip for 2002, but to reduce the number of records in the database, individual trips were grouped by vessel, month, and gear. These vessel–month–gear groups were sorted according to descending efficiency, i.e. descending net revenues per unit of the restricting factor. Subsequently, all records were cumulated, and landings were regressed against fishing effort. This relationship between effort and landings of each species led to a different production function for each management scenario, because fishers were expected to follow a different strategy according to the type of restriction.

The exponential regression of landings on fishing effort was designed as a classical Cobb–Douglas production function of the form

$$L_i = \alpha_i B_i E^{\beta}, \qquad (1)$$

where L denotes landings, E the fishing effort in horsepower-days (hp-days), B the spawning-stock biomass, i the species, and α and β are constants ($\beta < 1$). The actual equations and the estimated parameter values are listed in Appendix A. An important feature of this production function is that the rate of increase in landings decreases as effort increases. This reflects the assumption that if the fishery is restricted, the least efficient trips would be dropped first.

One assumption underlying the economic submodel is that the efficiency of month–vessel–gear groups is not distributed randomly, and that fishers may decide to cancel specific month–vessel–gear combinations to maximize net revenues. An analysis of variance of the net revenues per hp-day was performed to test this hypothesis. Selected explanatory variables were month, gear, interaction of month and gear, available hp-days per year, vessel, hp-group of vessel, interaction of gear and hp-group, and interaction of hp-group and available hp-days. For the entire fleet, 56% of the variance could be explained by these variables (Table 1), vessel, hp-group, and gear being the most important. For individual hp-groups (Table 2), the variance explained varied between 53% for Euro-cutters (engine power <300 hp) and 79% for vessels >2000 hp, and the most important explanatory variables were vessel and month. These results make it plausible that the industry should be able to cancel the least efficient trips when they faced a restriction. Therefore, net revenues would decrease less than proportionally with the restriction, by concentrating trips in the most efficient seasons, and by trading quota or hp-days from less-efficient vessels to more-efficient ones. In the long term, this may cause less-efficient vessels to be withdrawn from the fishery.

The variability in net revenues and catches of plaice and sole among groups of trips is probably less than among individual trips. This means that we may be underestimating the potential for fishers to optimize their behaviour, and hence the curvature of the production functions for the two species.

Scenarios

Scenario 1: TAC management

In this scenario, the two species are managed independently by single-species TACs. We make the assumption that the fisheries primarily target sole (the most valuable species by a factor of 4–5 in terms of price per kilogramme), and that fishing continues until the sole TAC has been fully taken, irrespective of the plaice TAC. In other words, exploitation of sole determines the behaviour of the fleets, resulting in under- or overexploitation of the plaice TAC. We assume that overquota catch of plaice is not landed, and therefore not accounted for in the assessment,

Table 1. Analysis of variance for the whole fleet: test of between-subjects effects with the dependent variable being net revenues per hp-day.

Source	Type 1 sum of squares	d.f.	Mean square	F	Significance
Corrected model	524 751[a]	60	8 702.3	9.95	<0.01
Intercept	209 706	1	209 706.0	239.80	<0.01
hp-group	235 545	5	47 109.1	53.87	<0.01
Gear	126 310	2	63 155.0	72.22	<0.01
Month	41 056 9	1	3732.4	4.26	<0.01
Gear * hp-group	106 866	6	17 811.0	20.36	<0.01
Gear * month	88 309.4	2	4014.0	4.59	<0.01
hp-group * hp- days	17 023.8	6	2837.3	3.24	<0.01
Vessel	463 240	55	8407.2	9.61	<0.01
Error	409 483	468	874.4	–	–
Total	955 206	528	–	–	–
Corrected total	934 235	528	–	–	–

[a]r^2 = 0.562 (adjusted r^2 = 0.505).

leading to a discrepancy between the true catch and the perceived catch (=landings). This scenario is thought to approximately reflect the current situation, although it clearly is a simplification: not all fleets target sole, and a small part of the Dutch beam trawl fleet even exploits the area north of 55°N, where sole are virtually absent and the target species is plaice.

The plaice catch taken is calculated from the effort required to deplete the sole TAC (see below), and may be below or exceed the plaice TAC. In both cases, the calculated plaice catch is fed back into the OM as the true catch. If the plaice catch exceeds the TAC, the landings (=perceived catch) in the MP are equal to the TAC, otherwise landings are equal to the true catch. The age distribution of the landings is assumed to be equal to the age distribution of the true catch.

Scenario 2: effort management

For this scenario, management is based on the allowable effort. We decided to simulate a management rule where the allowable effort is determined by the species for which the catch forecast imposes the most severe effort restriction. Hence, this effort is the lower of the two values estimated to yield the predicted catch under the target F for each species. Of course, managers may decide to apply a different type of effort-management regime, e.g. one where the allowable effort is determined by the catch forecast of the most valuable species, or of the species that needs least restriction. We did not run simulations of such management regimes.

The respective catches taken with the allowable effort serve directly as input for the true catch in the OM, as well as for the landings in the MP. In contrast to control by TACs, where discarding of overquota catch is implied under full compliance, if the quotas do not match, compliance under effort control implies that all fish caught may be landed.

Calculations

Because the parameters of the economic submodel are based on the Dutch fishery only, it was necessary to scale the forecast catches down to the Dutch portions of the total international catches. Similarly, the catches subsequently taken with Dutch effort had to be scaled up again to arrive at international catches. The proportions of the international sole and plaice catches taken by the Dutch fishery were assumed to be constant at 74% and 45%, respectively (averages over the period 1995–2002; ICES, 2004).

The calculation of annual "true" catches and economic results was performed in three steps.

(i) Using the inverse production functions (listed in Appendix A), effort was calculated from the spawning-stock biomass (B) and the Dutch part of the catch forecast under the target F. In scenario 1, the true effort equals the effort required to deplete the sole TAC, using the true B in the calculation. In scenario 2, the true effort equals the allowed effort (the lower estimate of the effort required to deplete the forecast catch of the two species), using the perceived B to mimic a management process in which only estimates are available, but not the true values.

(ii) True Dutch catches taken with the effort determined in step (i) were calculated, using the true B in the production functions listed in Appendix A. In scenario 1, only the plaice catch had to be calculated, because the sole catch was set equal to the TAC. The true Dutch catches were then raised to true international catches.

(iii) Prices, costs, revenues, and profits were calculated from Dutch landings and effort, using the costs and revenues functions listed in Appendix A. In scenario 2, landings were always equal to catches, whereas in scenario 1, potential overquota catches for plaice were not landed.

Uncertainty

Uncertainty was explicitly incorporated into the simulation framework, acknowledging the presence of a variety of sources (Rosenberg and Restrepo, 1994). In the simulation, this uncertainty caused the perception to deviate from the true world. Sources of error include: process error, attributable to natural variation in dynamic processes (e.g. recruitment); measurement error, generated when collecting observations from the populations; estimation error, arising from estimating parameters of the dynamic process during the assessment process; model error, because the true complexity of the dynamics can never be captured; and implementation error, attributable to imperfect implementation, e.g. when TACs are exceeded. In our case, model error was minimal because the equations used to construct the simulated populations were the same as those used to assess them. Variation among Monte Carlo runs came from two sources: random variability in recruitment around a Ricker relationship, and sampling error in the generation of the tuning series. Because the time-series of stock and recruitment data since 1957 do not provide evidence for one particular stock–recruitment relationship, the choice of a Ricker relationship was arbitrary, and we did not test robustness to alternative hypotheses (Machiels *et al.*, 2007; Pastoors *et al.*, 2007). Considering that recruitment variation is the major source of process error, we did not include additional forms, such as variability in weight-at-age or selectivity (see Pastoors *et al.*, 2007).

Table 2. Analyses of variance for hp-groups: test of between-subjects effects with the dependent variable being net revenues per hp-day.

hp-group	Source	Type 1 sum of squares	d.f.	Mean square	F	Significance
0–260	Corrected model	5 061 141.5[b]	263	19 243.884	5.098	<0.001
	Intercept	414 947.950	1	414 947.950	109.934	<0.001
	Gear	233 166.043	2	116 583.021	30.887	<0.001
	Month	124 604.387	11	11 327.672	3.001	0.001
	Gear * month	102 830.578	22	4810.481	1.274	0.178
	hp-days	14 199.940	1	14 199.940	3.762	0.053
	Vessel	4 583 340.5	227	20 190.927	5.349	<0.001
	Error	4 000 976.7	1060	3774.506	–	–
	Total	9 477 066.2	1324	–	–	–
	Corrected total	963.378	1323	–	–	–
261–300	Corrected model	21 890.617[d]	203	107.836	10.254	<0.001
	Intercept	27 472.774	1	27 472.774	2612.356	<0.001
	Gear	1279.602	2	639.801	60.838	<0.001
	Month	4493.482	11	408.498	38.844	<0.001
	Gear * month	2830.191	22	128.645	12.233	<0.001
	hp-days	546.835	1	546.835	51.998	<0.001
	Vessel	12 740.507	167	76.290	7.254	<0.001
	Error	19 749.935	1878	10.516	–	–
	Total	69 113.326	2082	–	–	–
	Corrected total	41 640.552	2081	–	–	–
301–800	Corrected model	574.644[e]	32	17.958	4.957	<0.001
	Intercept	145.440	1	145.440	40.150	<0.001
	Gear	16.036	1	16.036	4.427	0.037
	Month	170.649	11	15.514	4.238	<0.001
	Gear * month	37.386	7	5.341	1.474	0.181
	hp-days	27.998	1	27.998	7.729	0.006
	Vessel	322.575	12	26.881	7.421	<0.001
	Error	492.649	136	3.622	–	–
	Total	1212.733	169	–	–	–
	Corrected total	1067.293	168	–	–	–
801–1500	Corrected model	154.004[f]	35	4.400	5.228	<0.001
	Intercept	141.502	1	141.502	168.114	<0.001
	Gear	50.556	1	50.556	60.064	<0.001
	Month	27.557	11	2.505	2.976	0.002
	Gear * month	22.274	11	2.025	2.406	0.010
	hp-days	0.209	1	0.209	0.248	0.619
	Vessel	53.407	11	4.855	5.768	0.000
	Error	93.430	111	0.842	–	–
	Total	388.936	147	–	–	–
	Corrected total	247.433	146	–	–	–
1501–2000	Corrected model	738.795[c]	98	7.539	27.928	<0.001
	Intercept	1600.308	1	1600.308	5928.560	<0.001
	Gear	50.058	1	50.058	185.445	<0.001
	Month	369.071	11	33.552	124.298	<0.001
	Gear * month	10.119	5	2.024	7.497	<0.001
	hp-days	13.296	1	23.296	86.302	<0.001
	Vessel	286.251	80	3.578	13.256	<0.001
	Error	224.583	832	0.270	–	–
	Total	2563.686	931	–	–	–
	Corrected total	963.378	930	–	–	–

Continued

Table 2. *Continued*

hp-group	Source	Type 1 sum of squares	d.f.	Mean square	F	Significance
>2000	Corrected model	607.188[a]	68	8.929	30.274	<0.001
	Intercept	934.379	1	944.379	3 201.857	<0.001
	Gear	27.691	1	27.691	93.885	<0.001
	Month	210.317	11	19.120	64.824	<0.001
	Gear * month	0.168	2	9.309E−02	0.316	0.729
	hp-days	3.001	1	3.001	10.176	0.002
	Vessel	365.993	53	6.906	23.413	<0.001
	Error	166.645	565	0.295	–	–
	Total	1718.213	634	–	–	–
	Corrected total	773.834	633	–	–	–

[a] $r^2 = 0.785$ (adjusted $r^2 = 0.759$).
[b] $r^2 = 0\ 0.558$ (adjusted $r^2 = 0\ 0.449$).
[c] $r^2 = 0\ 0.767$ (adjusted $r^2 = 0\ 0.739$).
[d] $r^2 = 0\ 0.526$ (adjusted $r^2 = 0\ 0.474$).
[e] $r^2 = 0.538$ (adjusted $r^2 = 0\ 0.430$).
[f] $r^2 = 0\ 0.622$ (adjusted $r^2 = 0\ 0.503$).

Results

Scenarios

In both scenarios, the plaice stock recovered and generally stayed above $B_{\rm lim}$, whereas the sole stock generally remained above $B_{\rm lim}$ (Figure 2). The performance therefore conformed to the management goal of the precautionary approach, where the advice should ascertain that stocks would not drop below these limit reference points. In scenario 1, however, the initial rise of the spawning stock was followed by a decline, in contrast to a continuing rise in scenario 2. Moreover, F could rise steeply in scenario 1 when stock sizes were already decreasing, requiring drastic reductions subsequently. In contrast, F remained stable in scenario 2. This stability closer to the target F resulted in greater population growth than in scenario 1, and discrepancies between the perceived and true populations were smaller.

In scenario 1, there was overquota fishing of plaice throughout the simulation period in at least some Monte Carlo runs (Figure 3), resulting in the perceived catch (=landings) being lower than the true catch. This is the consequence of target Fs for sole and plaice not corresponding to similar effort levels (with the assumption that the sole TACs are adhered to), leading to conflicting management objectives. Dutch effort increased massively, but it declined towards the end of the simulation period (Figure 4).

In scenario 2, the effort restriction was more often (except the first year) determined by the plaice assessment than by that for sole (Figure 5). Despite the effort restriction alternating between being plaice- or sole-determined, the resulting effort level was fairly constant over the years (Figure 4). The conflicting level of effort required to meet the target Fs of the two species necessarily led to perceived "underexploitation" of one of the species (often alternating between the species among years). This term is used here to mean exploitation below the target catch forecast, whereas "overexploitation" means exploitation exceeding the target forecast. Formally, there should not be overexploitation in scenario 2, because the lower of two efforts was always used. However, there was sometimes overexploitation because the MP calculated allowable effort based on the predicted spawning-stock biomass, whereas the true spawning stock was sometimes larger, leading to higher catches than predicted.

In scenario 1, simulated prices first decreased and then rose, whereas they dropped first and then stabilized in scenario 2 (Figure 4). Net revenues followed the development of the landings (mirroring the development of the prices): in scenario 1, they increased but then declined, whereas they continued to rise in scenario 2 (Figure 4).

Comparative evaluation

We selected a range of evaluation criteria (Table 3) to review and compare the results of the two scenarios. In terms of the sustainable exploitation of the stocks, both management regimes generally led to avoidance of limit reference points. Scenario 2 performed slightly, but significantly, better than scenario 1: for sole, scenario 2 never violated $B_{\rm lim}$, whereas scenario 1 did so occasionally. For plaice, scenario 2 violated $B_{\rm lim}$ less often than scenario 1. Mean landings did not differ significantly, but the annual variation in sole landings was lower in scenario 1, and in plaice landings lower in scenario 2. From an economic perspective, scenario 2 appeared to perform better than scenario 1, with significantly higher net revenues in both the short and the long term. These higher net revenues seem to be due to lower variable costs at lower effort. Although net revenues over the whole period were similar, the net revenues in scenario 2 did not decline during the latter part of the simulation period, suggesting that if the simulation period had been prolonged, scenario 2 would have also yielded higher net revenues over the whole period. Economic stability between years was similar in the two scenarios. With respect to management, scenario 2 was more stable, requiring effort restrictions to vary annually by only 4% on average, whereas TACs varied annually by 14% and 12% for sole and plaice, respectively.

Shrinkage

In scenario 1, the perception of stock development lagged several years behind their true development (Figure 2). This can be seen more easily when the true population and the perceived population estimates from each Monte Carlo run are compared (Figure 6a). Because of this lag, there was a cyclical alternation of underestimation and overestimation.

We investigated whether this lag was caused by shrinkage, a technical setting in the XSA assessment model affecting the estimated F in the most recent years. The setting used traditionally in routine ICES assessments of North Sea plaice and sole is

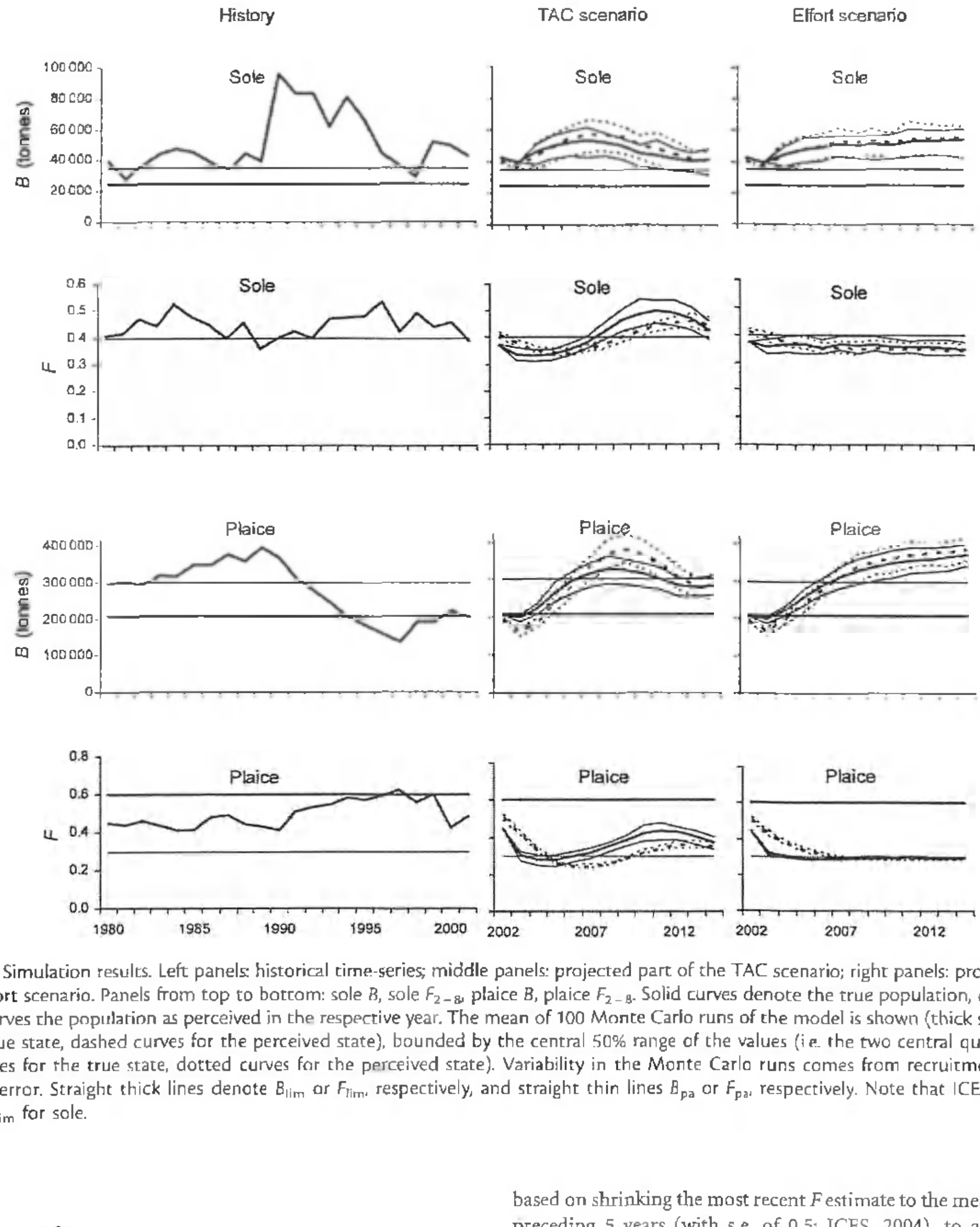

Figure 2. Simulation results. Left panels: historical time-series; middle panels: projected part of the TAC scenario; right panels: projected part of the effort scenario. Panels from top to bottom: sole B, sole F_{2-8}, plaice B, plaice F_{2-8}. Solid curves denote the true population, dashed and dotted curves the population as perceived in the respective year. The mean of 100 Monte Carlo runs of the model is shown (thick solid curves for the true state, dashed curves for the perceived state), bounded by the central 50% range of the values (i.e. the two central quartiles; thin solid curves for the true state, dotted curves for the perceived state). Variability in the Monte Carlo runs comes from recruitment and sampling error. Straight thick lines denote B_{lim} or F_{lim}, respectively, and straight thin lines B_{pa} or F_{pa}, respectively. Note that ICES has not defined F_{lim} for sole.

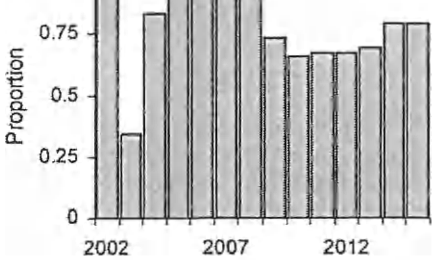

Figure 3. The proportion of Monte Carlo runs in which the plaice catch exceeded the plaice TAC in scenario 1, 2002–2015.

based on shrinking the most recent F estimate to the mean over the preceding 5 years (with s.e. of 0.5; ICES, 2004), to account for uncertainties in the non-converged part of the VPA. The choice of a specific degree of shrinkage reflects a trade-off between uncertainty and bias in the estimate of F. Therefore, the consequence of strong shrinkage is that actually realized changes in F in recent years are partly overruled by values of F in the more distant past. For comparison, we carried out 100 Monte Carlo runs with identical settings, except that F was shrunk less strongly towards the recent average (over 3 years with s.e. = 2.0). In this case (Figure 6b), the discrepancy between perceived and true population size was reduced markedly and the cycles disappeared. Therefore, the lag in perception of the true population development and the cyclical alternation of under- and overestimation appear to be caused by the strong shrinkage setting. Under low

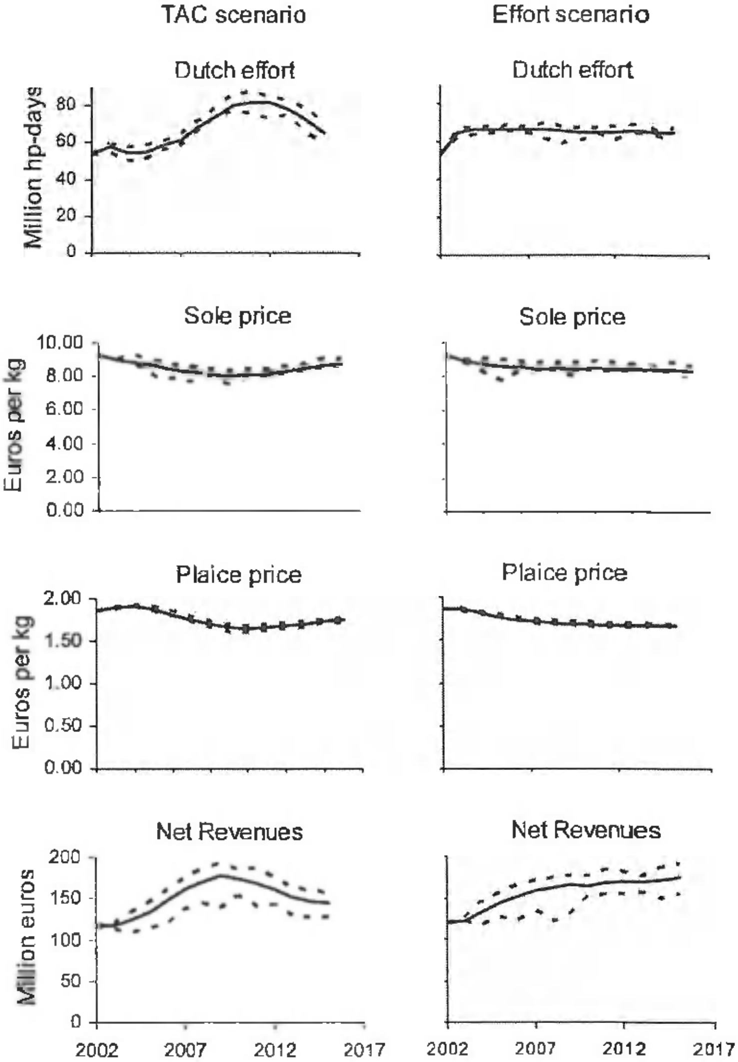

Figure 4. Simulation results. Left panels: projected part of the TAC scenario; right panels: projected part of the effort scenario. Panels from top to bottom: Dutch effort, Dutch sole price, Dutch plaice price, Dutch net revenues. Solid curves: mean; dotted curves: upper and lower quartiles.

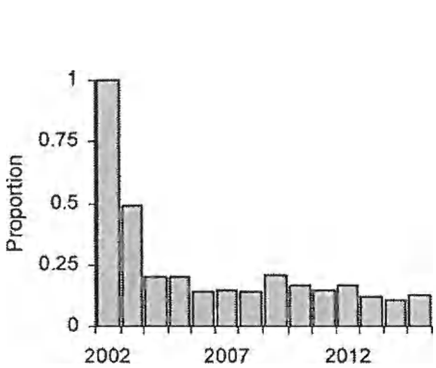

Figure 5. The proportion of Monte Carlo runs in which the sole catch forecast determined the effort restriction in scenario 2, 2002–2015.

shrinkage, spawning stock and F of the two species remain more stable (Figure 7). Although scenario 2 was run with the same strong shrinkage as scenario 1, there was no indication of cycles, because management stabilizes effort directly, and if F does not vary over time, shrinking to the mean has no additional effect.

Discussion

Assumptions

The construction of the true population was based on similar assumptions and the same equations as those used in the assessment of the perceived population. This may easily lead to overconfidence in the results, and must be acknowledged as a possible weakness in this type of analysis.

Moreover, the true natural mortality, maturity-at-age, and historical weights in the OM are assumed to be known in the MP, which is, of course, never the case. The same applies to the true

Table 3. Performance with respect to biological, economic, and management evaluation criteria for the two scenarios.

Parameter	Scenario 1 TAC	Scenario 2 effort restriction
Biological criteria for sole		
Frequency $F > F_{pa}$ (target F; %)	51.6	**15.2**
Frequency $B < B_{lim}$ (%)	1.9	**0.0**
Mean (true) landings ('000 t)	266	258
Mean absolute difference in landings between two consecutive years (%)	**14.3**	16.5
Biological criteria for plaice	–	–
Year of recovery (≥75% of runs with $B > B_{lim}$)	2004	2004
Frequency $F > F_{pa}$ (target F; %)	74.2	**36.0**
Frequency $B < B_{lim}$ from 2004 on (%)	4.5	**1.8**
Mean (true) landings ('000 t)	1130	1209
Mean absolute difference in landings between two consecutive years (%)	11.5	**7.2**
Economic criteria	–	–
Mean net short-term revenues (2003–2005; €million)	378	**400**
Mean net long-term revenues (2013–2015; €million)	444	**514**
Overall mean net revenues (2003–2015; €million)	1984	2049
Mean absolute difference in net revenues between two consecutive years (%)	9.3	9.2
Management criteria	–	–
Mean absolute difference in management measure between consecutive years (%)	Sole TAC: 14.3 Plaice TAC: 12.3	Allowable effort: **3.6**

Emboldened scores are significantly better than the score in the alternative scenario (t-test assuming equal variances, $p < 0.05$; proportions are arcsine-square-root transformed; no Bonferroni correction).

catch composition. The implicit but unrealistic assumption here is that the market-sampling programme gives precise estimates of the age composition of the landed catch. Another assumption is that the age composition is the same in the landed and the discarded overquota catches. This is an unrealistic assumption, because fishers may high-grade and selectively discard small fish of low value when approaching exhaustion of their quota. Consequently, the potential sources of variation and bias are underestimated in the approach, and the results may be overoptimistic with regards to how well the MP is able to monitor the true development of stocks. Moreover, there are other simplifications. For instance, historical recruitment series of sole and plaice appear to be correlated, whereas the projected recruitment of the two species in the simulations was uncorrelated.

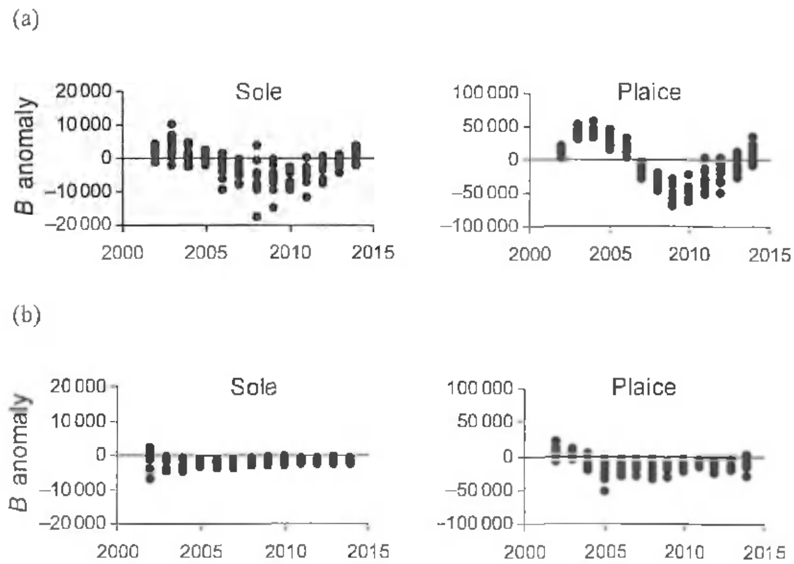

Figure 6. The difference between true and perceived estimates of B of 10 Monte Carlo runs of scenario 1. Left panels, sole; right panels, plaice. Positive values mean that the perceived population is underestimated compared with the true population. Variability in the Monte Carlo runs comes from recruitment and sampling error. (a) With strong shrinkage (over 5 years, s.e. = 0.5). (b) With weak shrinkage (over 3 years, s.e. = 2.0).

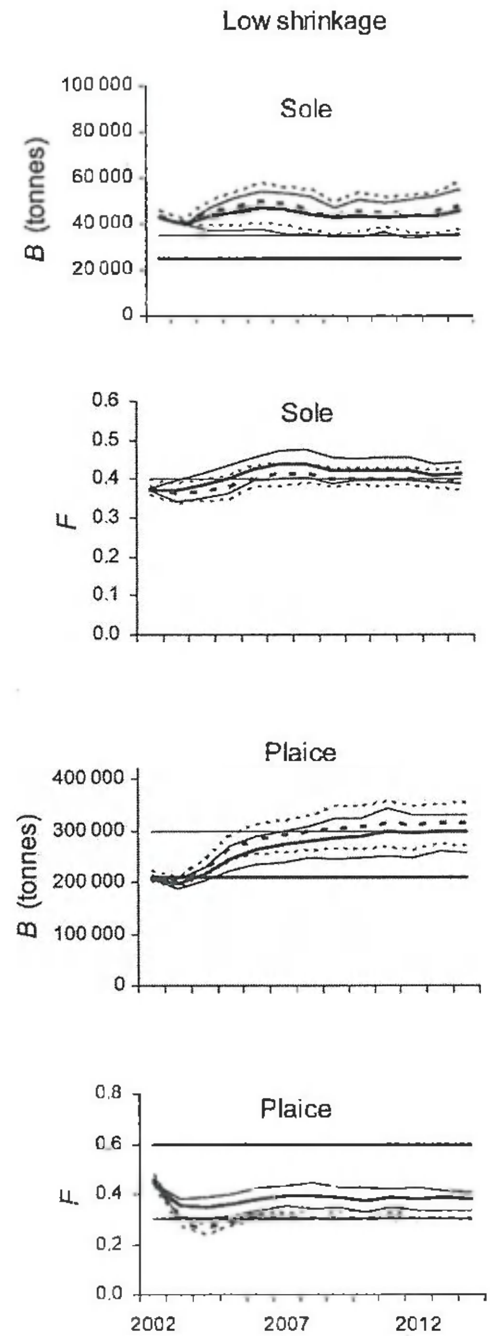

Figure 7. Scenario 1 with weak shrinkage (over 3 years, s.e. = 2.0). See legend of Figure 2 for full explanation.

An important restriction of this comparative analysis of two management scenarios is that many of the underlying data are based on the present behaviour of the fleets, which is a response to the present management system (single-species TACs). For instance, catch rates may change when single-species TACs are replaced by effort restrictions, because fishers will gain additional freedom to target individual species, and this may affect the direction in which a fishery develops. We have assumed that the fishery will not change from a sole-targeted one, but this cannot be guaranteed. In fact, the assumption that the sole TAC drives the plaice catch is a simplification, because a small part of the Dutch fleet fishes in an area where sole catches are negligible. Moreover, under effort restrictions, fishers may also maximize their profit by improving effort characteristics that are not restricted and increase the catchability (the so-called "technology creep" phenomenon; Pascoe *et al.*, 2001; Ulrich *et al.*, 2002; Rijnsdorp *et al.*, 2006). When this happens, parameters of the production functions change over time, and effective catches for a given amount of effort will differ from expected catches.

Another assumption relates to the behaviour of fishers with regard to high-grading. Under the existing single-species quota regime, catches may be higher than landings because of high-grading and discarding of overquota catches. Under effort restrictions, there is no incentive for high-grading and overquota discarding, so landings-per-unit-of-effort may increase. However, the production functions are based on (historical) landings data only, whereas they are used in the model for generating catch data. In reality, catches are likely to be bigger than the landings, which could cause underestimation of modelled catches.

The discarding of undersized plaice has been ignored in this analysis. Because of the different body shapes of sole and plaice, the selectivity of the prescribed 80 mm mesh size is very different for the two species. The 50% retention length in these nets is 27 cm for sole (minimum landing size 24 cm), but 18 cm for plaice (minimum landing size 27 cm). This implies that large quantities of undersized plaice are caught and discarded. Discard estimates were not available at the time of this study, but have recently been included in the plaice stock assessment even though their reliability is questionable (ICES, 2005). By neglecting the discarded fraction, the implicit assumption is that only fish above the respective minimum landing sizes are caught. Kell and Bromley (2004) pointed out that taking management decisions for North Sea plaice based on simulations without accounting for discarding leads to incorrect perceptions of stock development. However, the available estimates are based on crude extrapolations, and their incorporation might easily lead to unreliable results. Dickey-Collas *et al.* (2007) found that the inclusion of noisy discard estimates may make population trends hard to track with XSA. Although we acknowledge that assessments may be improved, this should not devalue the comparative value of the two scenarios assessed without these complications.

We also assumed full compliance to the management measures (allowing for legal discarding of overquota plaice catch in scenario 1). In reality, compliance with TACs may be a function of the discrepancies between the TACs set in consecutive years. Without directed research into fishers' responses to management measures, it is problematic to incorporate implementation bias into a model.

The technical interaction between sole and plaice in the Dutch beam trawl fishery is taken as a proxy for the linkage between the two species in all fisheries. This assumption imposes friction with another assumption that Dutch catches can be multiplied by fixed but different factors to arrive at the respective international catches (1.35 for sole, 2.25 for plaice). The friction arises because the different multipliers imply that, internationally, more plaice is caught than can be accounted for by the mixed-fishery interaction.

We further assumed that future management decisions were based on fishing at a target F that equals the respective F_{pa} (for both species in the TAC scenario, and for the species requiring stronger effort-restriction in the effort-regulation scenario). In reality, the EC management plan for North Sea plaice and sole adopted in 2007 and implemented for the first time in 2008 prescribes target *F*s that are reduced by 10% each year until the final objectives of $F = 0.3$ for plaice and $F = 0.2$ for sole will have been reached. With this plan becoming operational, the stocks should recover more quickly and lower *F*s would be realized. However, because our study was carried out before the new management plan existed, we did not incorporate any adjustments of target *F* in the simulations, which may have led to a more pessimistic development than the new plan would allow for. An earlier version of the current management plan has been evaluated by Machiels *et al.* (2007) and the Scientific, Technical, and Economic Committee for Fisheries (STECF, 2006), through simulation studies similar to this one.

Ulrich *et al.* (2002) undertook a comparable study on the mixed-flatfish fishery of the North Sea, also including economics. They modelled individual fleets, included discarding of undersized fish, and assumed that *F* is proportional to effort. Similarly, the simulation study that evaluated the draft EC management plan (Machiels *et al.*, 2007) assumed a proportional relation between effort and *F*. However, this assumption seemed problematic in the evaluation by STECF (2006). To our knowledge, a proportional relationship between effort and *F* has not been convincingly demonstrated. On the contrary, the relationship varies in response to technical, environmental, and behavioural factors (Rijnsdorp *et al.*, 2006). The last authors noted that through optimization behaviour of fishers, fishing mortality can be affected less than proportionally by effort reductions. We therefore believe that a production function based on the idea that a fishery faced with management restrictions will cancel the least profitable trips first is a more realistic approach. However, although we introduced economic data to mimic changes in fleet behaviour in response to management, the approach is still simplistic. In practice, species compositions of the catch may change outside the simulated range in response to the management strategy chosen.

Another limitation is the lack of sensitivity tests to evaluate the robustness of our conclusions by varying the various assumptions. It is clear that there is a tension between the level of detail that may seem desirable and the level that can be provided when specifying, conditioning, and validating OMs in management evaluation studies (Pastoors *et al.*, 2007).

To put all these assumptions into perspective, we stress that scenario 1 is supposed to reflect largely the management system as it has been in place until and including 2007. The outcomes of our simulations suggest that this system would guarantee that both stocks are managed effectively in avoiding the limit reference points, and that the TAC system is a safeguard against overexploitation. However, this statement contrasts sharply with the recent perceptions of the developments in both stocks, suggesting that scenario 1 is too optimistic as a representation of the recent regime.

Shrinkage

For many demersal stocks in northern Europe, the use of strong shrinkage of *F* towards the recent 5-year mean has been a common practice until recently. For North Sea plaice and sole, this practice has changed recently (ICES, 2007), when the results of the simulations reported here were made available to the assessment working group in an earlier draft (Kraak *et al.*, 2004). Our analysis shows that if *F* varies markedly, as in scenario 1, the use of strong shrinkage induces cyclical developments in the stock. This instability is caused by the perception of the stock being always out of phase with the true state. Shepherd (1999) commented that shrinking *F* to the recent mean might cause conflict with strong signals in the surveys, particularly if catch data are inaccurate. However, such a conclusion misses the point that even if the catch is well sampled and there are no strong signals in surveys, shrinkage may introduce a greater bias than the uncertainty originating from the unconverged part of the VPA that it is trying to avoid. Importantly, the strength of the shrinkage chosen reflects a trade-off between bias and uncertainty (Dickey-Collas *et al.*, 2007). In our simulations, many sources of uncertainty were absent (as explained above), but in ICES assessments, total uncertainty may be substantial.

The bias has marked effects on the advice and corresponding management measures. During a period with a strong trend in *F*, shrinkage results in a discrepancy in the perception of the true state of the stock (see also Dickey-Collas *et al.*, 2007). Therefore, the advice, and the measures taken, will be inappropriate relative to the reference points aimed for, and always lag behind, leading to overshoots and undershoots. These problems may explain the retrospective bias often observed in assessments. They may also lead to unwelcome and unnecessary variations in TACs, so undermining the credibility of the scientific advice (by causing large retrospective change) and prohibiting implementation of timely and stable management measures. As recovery plans or other conservative measures are often associated with strong reductions in *F*, as intended by the adopted EC management plan for North Sea plaice and sole, the perception of success of such plans would be jeopardized if shrinkage of *F* to the mean were applied.

Comparison of the two scenarios

This simulation exercise was undertaken to explore the methodology and to gain insight into the question of how alternative management scenarios may differentially affect stock development, the fishery, its economy, and its management. The projections are not meant to be viewed as quantitative stock forecasts or predictions to be used in North Sea flatfish management (Pastoors *et al.*, 2007). Rather, the two scenarios represent a qualitative comparison of two management systems for a particular set of simplifying assumptions. The conclusion as to which management scenario performs best should be viewed as conditional on many restrictive assumptions, because the conclusions may not hold if some of these were relaxed or changed.

Based on the evaluation criteria chosen, the effort-management scenario appears to be more suitable for managing the beam trawl fishery than the TAC scenario, because it results in a more positive biological development of both stocks, in higher short-term and long-term economic profits, and greater stability of annual management measures.

Although both scenarios generally led to sustainable exploitation (i.e. they kept *F* below F_{lim}, and the spawning stock above B_{lim}), effort management did so slightly better and led generally to larger stocks and lower effort. This is because the effort-restriction implemented was tuned to the species most in need of restriction, so acknowledging the mixed-species nature of the fishery. This often resulted in catches of the other species that were lower than might be taken with the single-species target F (F_{pa}). This, in turn, was favourable for the development of that stock, and was paid back later. However, other effort-management regimes can be envisaged and very different outcomes expected. For instance, a system aimed specifically at sustaining the fishery of the most valuable species could easily lead to overexploitation of the other.

Scenario 1, mimicking the current situation, performed worse because the sole catch determined effort, whereby the plaice stock was often overexploited and some plaice catches could not be landed and did not contribute to revenues. Still, this scenario is probably overoptimistic as a reflection of the current management regime, in view of the negative development of the stocks observed recently under the current management regime.

The development of the stocks was more stable under scenario 2, leading to lower assessment bias and error (lower discrepancy between perceived and true states of both stocks). Assessment bias caused by shrinkage reinforces instability. Conversely, the lack of assessment bias reinforces stability and makes the advice less sensitive to shrinkage. This should, in turn, lead to the setting of more appropriate measures (in this case, allowed effort).

To conclude, many of the differences between the scenario outcomes are attributable to the way the management regimes dealt with the conflicting target *F*s for sole and plaice and the inconsistent fleet effort required to execute these *F*s. Our study suggests that a management strategy that occasionally results in perceived underexploitation of the stocks may work best, given the existence of assessment error and bias. Our results also suggest that stability in fishing mortality reinforces itself, because assessment bias is lower under greater stability, and corrections become less necessary.

Acknowledgements

This study was financed by the Department of Fisheries of the Dutch Ministry of Agriculture, Nature and Food Quality (LNV) under contract Ond/2002-1/6c/1. We thank Laurie Kell for playing a large role in developing the simulation models and the concept of the paper, and Clara Ulrich, Trevor Hutton, and anonymous referees for their detailed comments on earlier versions of the manuscript.

References

Butterworth, D. S., and Punt, A. E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985–998.

Daan, N. 1997. TAC management in North Sea flatfish fisheries. Journal of Sea Research, 37: 321–341.

Darby, C. D., and Flatman, S. 1994. Virtual Population Analysis: Version 3.1 User Guide. MAFF Directorate of Fisheries Research, IT Report 1. 85 pp.

de Wilde, W. 1999. Measurement of economic impacts of fishery management decisions, The Dutch case (associated with the Klepper & Lasch work). EAFE Conference 1999, Dublin.

Dickey-Collas, M., Pastoors, M. A., and van Keeken, O. A. 2007. Precisely wrong or vaguely right: simulations of noisy discard data and trends in fishing effort being included in the stock assessment of North Sea plaice. ICES Journal of Marine Science, 64: 1641–1649.

EC. 2007. Council Regulation (EC) No 676/2007 of 11 June 2007 establishing a multiannual plan for fisheries exploiting stocks of plaice and sole in the North Sea. http://eur-lex.europa.eu/LexUriServ/site/en/oj/2007/l_157/l_15720070619en00010006.pdf.

Frost, H. 1997. Transition from quota to effort regulation. South Jutland University Press, DIFER Report Series 23/97. Report with Financial Assistance from the EU Commission DG XIV (DG14.A-94/03).

Harwood, J., and Stokes, K. 2003. Coping with uncertainty in ecological advice: lessons from fisheries. Trends in Ecology and Evolution, 18: 617–622.

ICES. 2001. Report of the ICES Advisory Committee on Fishery Management. ICES Cooperative Research Report, 242.

ICES. 2003. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) Copenhagen, Denmark, 11–20 June 2002. ICES Document CM 2003/ACFM: 02.

ICES. 2004. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) Boulogne sur Mer, France, 9–18 September 2003. ICES Document CM 2004/ACFM: 07.

ICES. 2005. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) Bergen, Norway, 7–16 September 2004. ICES Document CM 2005/ACFM: 07.

ICES. 2007. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK), ICES Headquarters, 5–14 September 2006. ACFM: 35. http://www.ices.dk/reports/ACFM/2006/WGNSSK/directory.asp.

Kell, L. T., and Bromley, P. J. 2004. Implications for current management advice for North Sea plaice (*Pleuronectes platessa* L.). 2. Increased biological realism in recruitment, growth, density-dependent sexual maturation and the impact of sexual dimorphism and fishery discards. Journal of Sea Research, 51: 301–312.

Kell, L. T., Kirkwood, G., Pilling, G., Abaunza, P., Aps, R., van Beek, F. A., Biseau, A. *et al.* 2002. Analysis of possibilities of limiting the annual fluctuations in TACs (MATES, Final Report FISH/2001/02/02). CEFAS, Lowestoft, UK.

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J-M., Garcia, D., Hillary, R., Jardim, E. *et al.* 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640–646.

Kell, L. T., Pilling, G. M., Kirkwood, G., Pastoors, M. A., Mesnil, B., Korsbrekke, K., Abaunza, P. *et al.* 2005. An evaluation of the implicit management procedure used for some ICES roundfish stocks. ICES Journal of Marine Science, 62: 750–759.

Kraak, S. B. M., Buisman, F. C., Dickey-Collas, M., Poos, J. J., Pastoors, M. A., Smit, J. P. G., Daan, N. *et al.* 2004. How can we manage mixed fisheries? A simulation study of the effect of management choices on the sustainability and economic performance of a mixed fishery. ICES Document CM 2004/FF: 11.

Machiels, M. A. M., Kraak, S. B. M., and Beek, F. A. van 2007. Evaluation of a management plan as proposed by the European Commission in 2006 for fisheries exploiting stocks of plaice and sole in the North Sea. Wageningen IMARES Report, C011/07. IJmuiden, the Netherlands. http://old-stecf.jrc.it/meetings/sgeca/0605/imares.pdf.

Nielsen, M. 1999. EU Seafood Markets—Integration and Demand, Working Paper No13/1999, SJFI, Copenhagen.

Pascoe, S., Andersen, J., and de Wilde, W. 2001. The impact of management regulation on the technical efficiency of vessels in the Dutch beam trawl fishery. European Review of Agricultural Economics, 28: 187–206.

Pastoors, M. A., Poos, J. J., Kraak, S. B. M., and Machiels, M. A. M. 2007. Validating management simulation models and implications for communicating results to stakeholders. ICES Journal of Marine Science, 64: 818–824.

Piet, G. J., and Rice, J. C. 2004. Performance of precautionary reference points in providing management advice on North Sea fish Stocks. ICES Journal of Marine Science, 61: 1305–1312.

Punt, A. E., Smith, A. D. M., and Cui, G. 2002. Evaluation of management tools for Australia's South East fishery. 1. Modelling the South East fishery taking into account technical interactions. Marine and Freshwater Research, 53: 615–629.

Rijnsdorp, A. D., Daan, N., and Dekker, W. 2006. Partial fishing mortality per fishing trip: a useful indicator of effective fishing effort in mixed demersal fisheries. ICES Journal of Marine Science, 63: 556–566.

Rosenberg, A. A., and Restrepo, V. R. 1994. Uncertainty and risk evaluation in stock assessment advice for U.S. marine fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 51: 2715–2720.

Sainsbury, K. J., Punt, A. E., and Smith, A. D. M. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES Journal of Marine Science, 57: 731–741.

Salz, P., and Frost, H. 2000. Model for economic interpretation of the ACFM advice (EIAA). *In* Proceedings of the XIIth Annual Conference of the European Association of Fisheries Economists (EAFE). Ed. by E. Lindebo, and N. Vestergaard. University of Southern Denmark.

Shepherd, J. G. 1999. Extended survivors analysis: an improved method for the analysis of catch at age data and abundance indices. ICES Journal of Marine Science, 56: 584–591.

STECF. 2006. Commission staff working paper, impact assessment: plaice and sole long-term management. STECF Sub-group SGECA-SGRST-06-05, Brussels, 26–29 September 2006. http://stecf.jrc.ec.europa.eu/20.

Sutinen, J. G. 1999. What works well and why: evidence from fishery-management experiences in OECD countries. ICES Journal of Marine Science, 56: 1051–1058.

Ulrich, C., Pascoe, S., Sparre, P. J., de Wilde, J-W., and Marchal, P. 2002. Influence of trends in fishing power on bioeconomics in the North Sea flatfish fishery regulated by catches or by effort quotas. Canadian Journal of Fisheries and Aquatic Sciences, 59: 829–843.

Appendix A: equations and symbols used in the model

Equations

Population dynamics

$$N_{i,a+1,y+1} = N_{i,a,y}e^{-Z_{i,a,y}} \tag{2}$$

$$N_{i,p,y} = N_{i,p-1,y-1}e^{-Z_{i,p-1,y-1}} + N_{i,p,y-1}e^{-Z_{i,p,y-1}} \tag{3}$$

Mortality rates

$$Z_{i,a,y} = F_{i,a,y} + M_{i,a,y} \tag{4}$$

$$F_{i,a,y} = S_{i,a,y}f_{i,y} \tag{5}$$

Catch equation

$$C_{i,a,y} = N_{i,a,y}\frac{F_{i,a,y}}{Z_{i,a,y}}(1 - e^{-Z_{i,a,y}}) \tag{6}$$

Stock–recruitment relationship (Ricker)

$$N_{i,r,y} = (\alpha_i B_{i,y-r}e^{-\beta_i B_{i,y-r}})e^{\varepsilon_{i,y}-\sigma^2/2} \tag{7}$$

$$\varepsilon_{i,y} \sim N(0, \sigma_i^2) \tag{8}$$

$$\sigma_i^2 = \ln(\mathrm{CV}_i^2 + 1) \tag{9}$$

Catch per unit effort models

$$U'_{i,f,a,y} = \frac{U_{i,f,a,y}}{A_{i,f,a,y}} \tag{10}$$

$$A_{i,f,a,y} = \frac{(e^{-\alpha_f Z_{i,a,y}} - e^{\beta_f Z_{i,a,y}})}{(\beta_f - \alpha_f)Z_{i,a,y}} \tag{11}$$

$$U'_{i,f,a,y} = q_{i,f,a}N_{i,a,y}e^{N(0,\varphi^2_{i,f,a})-\varphi^2_{i,f,a}/2} \tag{12}$$

Yield

$$Y_{i,y} = \sum_{a=r}^{p} C_{i,a,y}W_{C,i,a,y} \tag{13}$$

Spawning-stock biomass

$$B_{i,y} = \sum_{a=r}^{p} N_{i,a,y}W_{S,i,a,y}O_{i,a,y} \tag{14}$$

Single-species TAC scenario

$$C_{s,y} = Q_{s,y} \tag{15}$$

$$E_y = \alpha_0(M_{s,y})^{-\beta_0}Q_{s,y}^{\beta_0} \tag{16}$$

$$C_{p,y} = \alpha_{1,p}M_{p,y}E_y^{\beta_{1,p}} \tag{17}$$

Effort-restriction scenario

$$E_y = \mathrm{MIN}\left[\left(\alpha_{0,s}(M_{s,y})^{-\beta_0^s}Q_{s,y}^{\beta_0^s}\right), \left(\alpha_{0,p}(M_{p,y})^{-\beta_0^p}Q_{p,y}^{\beta_0^p}\right)\right] \tag{18}$$

$$C_{s,y} = \alpha_s M_{s,y}E_y^{\beta_s} \tag{19}$$

$$C_{p,y} = \alpha_p M_{p,y}E_y^{\beta_p} \tag{20}$$

Economic variables

$$P_{i,y} = P_i^0\left(\frac{L_{i,y}}{L_i^0}\right)^{e_i} \tag{21}$$

$$R_{O,y} = R_O^0\left(\frac{E_y}{E^0}\right) \tag{22}$$

$$R_{T,y} = \sum(L_{i,y}P_{i,y}) + R_{O,y} \tag{23}$$

$$C_{V,y} = \alpha_c + \beta_c E_y 1000 \tag{24}$$

$$R_{N,y} = R_{T,y} - C_{V,y} \tag{25}$$

Symbols used in equations

	Parameter	Definition
Population dynamics	$N_{i,a,y}$	Numbers of fish of species i of age a at the start of year y
	p	Age of the plus group; here, p = 15 years
	$Z_{i,a,y}$	Total mortality of species i at age a in year y
Mortality rates	$M_{i,a,y}$	Natural mortality of species i at age a in year y
	$F_{f,a,y}$	Fishing mortality of species i at age a in year y
	$f_{f,y}$	Year effect of fishing mortality of species i in year y
	$S_{i,a,y}$	Selection pattern of species i at age a in year y
Catch equation	$C_{i,a,y}$	Catch in numbers of species i at age a in year y
Stock–recruitment relationship	r	Age at first recruitment to the fishery; here, r = 1 year
	$B_{i,y}$	Spawning-stock biomass of species i in year y
	α_i, β_i	Constants; for i = sole $\alpha_i = 5.1055$, $\beta_i = 0.0000168$, and for i = plaice $\alpha_i = 3.81$, $\beta_i = 0.00000331$ (estimated as explained in Appendix B)
	$\varepsilon_{i,y}$	Recruitment residual of species i in year y
	σ_i	Standard error of recruitment residuals of species i
	CV_i	Coefficient of variation of recruitment; for i = sole $CV_i = 0.5$, and for i=plaice $CV_i = 0.35$ (estimated as explained in Appendix B)
Catch per unit effort models	$U_{i,f,a,y}$	Cpue of species i by fleet f at age a in year y
	$U'_{i,f,a,y}$	Cpue of species i by fleet f at age a adjusted to start of year y
	$A_{i,f,a,y}$	Averaging factor of species i for fleet f relating the population abundance-at-age a during the time at which the catch was taken to the population abundance-at-age a at the beginning of the year y
	$q_{i,f,a}$	Catchability, relationship between cpue and numbers of species i for fleet f at age a (source explained in Appendix B)

Continued

Continued

	Parameter	Definition
	α_f	Start of the period of fishing of fleet f, at 0.66 of the year
	β_f	End of the fishing period of fleet f, at 0.75 of the year
	$\varphi_{i,f,a}$	Standard error of cpue residuals of species i of fleet f at age a (source explained in Appendix B)
Yield	$W_{C,i,a,y}$	Body mass in the catch of species i at age a in year y
	$Y_{i,y}$	Total catch mass of species i of all ages in year y
Spawning-stock biomass	$W_{S,i,a,y}$	Body mass in the stock of species i at age a in year y
	$O_{i,a,y}$	Proportion mature of species i at age a in year y
Single-species TAC scenario and effort-restriction scenario	E_y	Fishing effort in million hp-days in year y
	$M_{i,y}$	B of species i in year y relative to the B of that species in the reference year (2002)
	$Q_{i,y}$	Catch forecast (tonnes) of species i for year y, using the equation for yield Y
	$C_{i,y}$	Catch (tonnes) of species i in year y
	s (subscript)	Sole
	p (subscript)	Plaice
	α_0, β_0, $\alpha_{0,s}$, $\alpha_{0,p}$, β_0^s, β_0^p, α_s, β_s, α_p, β_p	Constants; $\alpha_0 = 0.0009786$, $\beta_0 = 1.177$, $\alpha_{0,s} = 0.00011698$, $\alpha_{0,p} = 0.00001882$, $\beta_0^s = 1.4057$, $\beta_0^p = 1.4578$, $\alpha_s = 628.637$, $\beta_s = 0.7105$, $\alpha_p = 916.432$ and $\beta_p = 0.8496$ in scenario 1, $\alpha_p = 1747.43$ and $\beta_p = 0.6853$ in scenario 2 (estimated from regressions as explained in the main text)
Economic variables	$P_{i,y}$	Dutch price in euros per kg of species i in year y
	$C_{V,y}$	Dutch variable costs in 1000 euros in year y
	α_c, β_c	Constants; $\alpha_c = -1.629$ and $\beta_c = 1.4355$ in scenario 1, $\alpha_c = 0.4715$ and $\beta_c = 1.4021$ in scenario 2 (estimated from regressions, see main text)
	$R_{T,y}$	Dutch total revenues in thousand euros in year y

Continued

Continued

Parameter	Definition
$R_{N,y}$	Dutch net revenues in thousand euros in year y
$L_{i,y}$	Dutch landings of species i (tonnes) in year y
L_i^0	Dutch landings of species i (tonnes) in reference year (2002); $L_s^0 = 10\,611$ and L_p^0 26 977 in scenario 1, $L_s^0 = 10\,611$ and L_p^0 26 668 in scenario 2 (from LEI data)
P_i^0	Dutch average price in euros per kg of species i in reference year (2002); $P_s^0 = 9.26$ and $P_p^0 = 1.86$ (from LEI data)
e_i	Price elasticity of species i; $e_s = -0.3$ and $e_p = -0.2$ (after Nielsen, 1999)
$R_{O,y}$	Dutch revenues from other species in thousand euros in year y
R_O^0	Dutch revenues from other species in reference year (2002) in thousand euros; $R_O^0 = 46\,385$ (from LEI data)
E^0	Dutch fishing effort in reference year (2002) in million hp-days; $E^0 = 53.569$ in scenario 1, $E^0 = 53.346$ in scenario 2 (from LEI data)

Appendix B: input data

The input data required for the OM are:

(i) Initial (1957) population numbers-at-age ($N_{i,a,y}$); from ICES (2004);

(ii) Natural mortality-at-age M ($M_{i,a,y}$); assumed constant (ICES, 2004);

(iii) Maturity-at-age ($O_{i,a,y}$); assumed constant (ICES, 2004);

(iv) Mean weight-at-age of fish in the stock ($W_{S,i,a,y}$) and in the catch ($W_{C,i,a,y}$); for the period 1957–2001 from ICES (2004); for the years 2002–2015 the running average over the previous 3 years;

(v) Selection pattern (relative F-at-age, $S_{i,a,y}$); for 1957–2001 absolute F-at-age values from ICES (2003); for 2002–2015 assumed constant as the average over the period 1997–2001;

(vi) Recruits ($N_{i,1,y}$); for 1957–2001 from ICES (2004); for 2002–2015 according to a Ricker model with a lognormal error distribution (determined by regressing recruits on SSB based on 1957–2001 data; standard deviations chosen through visual inspection of the resulting variation). Recruitment in the period 2002–2015 is a source of variability between simulation runs.

(vii) Yield ($Y_{i,y}$); for 1957–2002 from ICES (2004); from 2003 on the yield is determined by the management decision set in the previous year.

From these input data, the following quantities are calculated for both stocks by year: catch-at-age ($C_{i,a,y}$), population numbers-at-age ($N_{i,a,y}$), F-at-age ($F_{i,a,y}$), F_{2-8} ($F_{i,2-8,y}$), and spawning-stock biomass ($B_{i,y}$).

The input required for the MP (as required by XSA) is as follows.

(i) Perceived catch-at-age; equal to true catch-at-age in the OM for sole, and for plaice if the true catch did not exceed the TAC; otherwise equal to the TAC with the same age composition;

(ii) Perceived M; equal to true M in the OM;

(iii) Perceived maturity-at-age; equal to true maturity in the OM;

(iv) Perceived weight-at-age of fish in the stock and in the catch; averages of true values of the previous 3 years in the OM.

In addition, for each species, two catch per unit effort (cpue) series ($U'_{i,f,a,y}$) are generated for tuning, from true N-at-age, true M-at-age, true F-at-age, and catchabilities ($q_{i,f,a}$, see below), without a power model. Both series are set to take place in late summer, taking into account the proportion of the year the fish have been exploited (survey starts at 0.66 and ends at 0.75). Cpue series 1 commences in 1984 and involves ages 1–9, and series 2 commences in 1982 and involves ages 1–3. Catchabilities $q_{i,f,a}$ and their standard errors $\varphi_{i,f,a}$ are given in Table B1. These values are taken from existing research vessel survey tuning series (ICES, 2003), except for ages 2–9 of sole series 1, which are from a commercial fleet (ICES, 2004). The generation of tuning series with random error contributes a second source of variability between simulation runs.

In the MP, the annual assessment is performed by XSA, using the perceived data, with the settings for both stocks as used in ICES (2004) for plaice. For runs with low shrinkage, the s.e. has been set at 2 (as opposed to 0.5 for default runs) over 3 years (as opposed to 5 years for the default runs). The assessment generates values for perceived F-at-age, perceived N-at-age, perceived F_{2-8}, and perceived B up to the last data year. Short-term forecasts were run, under the assumption that F_{2-8} in the intermediate year equalled F_{2-8} in the previous year, and with the respective target

Table B1. Catchability ($q_{i,f,a}$) and standard error ($\varphi_{i,f,a}$) of cpue series 1 and 2 for sole and plaice.

Age a	Sole as species i				Plaice as species i			
	$q_{i,1,a}$	$\varphi_{i,1,a}$	$q_{i,2,a}$	$\varphi_{i,2,a}$	$q_{i,1,a}$	$\varphi_{i,1,a}$	$q_{i,2,a}$	$\varphi_{i,2,a}$
1	8.89	0.29	3.94	0.18	7.29	0.54	2.47	0.37
2	6.22	0.50	4.92	0.31	7.70	0.31	3.62	0.43
3	5.27	0.24	5.57	0.50	8.62	0.24	4.93	0.49
4	5.11	0.22	–	–	9.48	0.21	–	–
5	5.09	0.22	–	–	10.13	0.22	–	–
6	5.27	0.19	–	–	10.45	0.28	–	–
7	5.31	0.26	–	–	10.73	0.29	–	–
8	5.31	0.27	–	–	10.76	0.34	–	–
9	5.31	0.20	–	–	11.08	0.38	–	–

*F*s (0.4 and 0.3 for sole and plaice, respectively) for the TAC year. The selection pattern in the TAC year and the intermediate year was assumed to be equal to the average over the previous 3 years. These forecasts generate a predicted *B* for the start of the TAC year and 1 year later and the forecast catch for the TAC year. Note that in reality in 2004 the age range for *F* as well as the pa reference points changed (ICES, 2005); we used the values before that change (ICES, 2004).

doi:10.1093/icesjms/fsn045