Aquat. Living Resour. 23, 77–94 (2010)

DOI: 10.1051/alr/2010005
www.alr-journal.org

Aquatic Living Resources

Evaluating the performance of survey-based operational management procedures

Marine Pomarede[1,a], Richard Hillary[1], Leire Ibaibarriaga[2], Johannes Bogaards[3,b] and Panayiota Apostolaki[4]

[1] Fisheries Group, Imperial College, Division of Biology, Silwood Park, Ascot SL5 7PY, UK
[2] AZTI, Herrera Kaia, Portualdea z/g, 20110 Pasajes, Spain
[3] IMARES, PO Box 68, Haringkade 1, 1970 AB, IJmuiden, The Netherlands
[4] CEFAS, Pakefield road, NR33 OHT, Lowestoft, UK

Received 13 March 2009; Accepted 15 October 2009

Abstract – The design and evaluation of survey-based management strategies is addressed in this article, using three case-study fisheries: North Sea herring, Bay of Biscay anchovy and North Sea cod, with a brief history and the main management issues with each fishery outlined. A range of operational management procedures for the case study stocks were designed and evaluated using trends that may be derived from survey indices (spawner biomass, year-class strength and total mortality) with an array of simple and more structured observation error regimes simulated. Model-free and model-based indicators of stock status were employed in the management procedures. On the basis of stochastic stock-specific simulations, we identified the following key determinants of successful management procedures: (i) adequate specification of the stock-recruit relationship (model structure, parameter estimates and variability), (ii) knowledge of the magnitude and structure of the variation in the survey indices, and (iii) explication of the particular management objectives, when assessing management performance. More conservative harvesting strategies are required to meet specified targets in the presence of increasing stochasticity, due to both process and observation error. It was seen that survey-based operational management procedures can perform well in the absence of commercial data, and can also inform aspects of survey design with respect to acceptable levels of error or bias in the surveys.

Key words: FLR fishery simulation system / Management strategy evaluation / Fish survey / Fishery-independent data / Herring / Anchovy / Cod / North Sea / Biscay

1 Introduction

A certain number of stocks in Europe are managed by fisheries quotas which are set to help achieve the objectives of the Common Fisheries Policy for the conservation and sustainable management of fish stocks. The goal of quotas is to control exploitation via Total Allowable Catch (TAC) limitations. Historically, and more particularly in recent years, important reductions of the biomass in several demersal fish stocks have been observed in European waters. Recent stock biomass declines induced reductions in TAC levels. As a consequence, the motivation for fishermen to discard, misreport or distort catch records increased (Cook 1997; Anonymous 2004) so that scientists do not know how many fish have been landed. This can lead to biases in the catch data, low stock abundance estimates by scientists, and even lower TACs, followed by even more misreporting.

Standard assessment methods are heavily reliant on fisheries dependent data, e.g. data on commercial landings and fishing effort (McAllister and Kirchner 2001; Nielsen and Lewy 2002; Lewy and Nielsen 2003; Chen et al. 2004; Cotter et al. 2004). Survey indices are also used in these methods but their role is to "tune" trends in commercial data (Beare et al. 2005). The concern with these methods is the predominance of reported catch-at-age data. In fact, as there is a risk of misreporting and as a consequence of bias in the catch data due to the reduction of TAC levels, the quality of the central data used in the assessments suggests that this type of assessment method might generate misleading scientific advice to then be provided for use in fisheries management, with potential implications for sustainability, conservation and recovery.

Time-series estimates from fish surveys and other fishery-independent sources may provide better information than those

[a] Corresponding author: marine.pomarede@imperial.ac.uk
[b] Present address: VU University Medical Centre, PO Box 7057, 1007 MB Amsterdam, The Netherlands

based on fishery-dependent sources for managing a fishery, especially when the available fishery-dependent data are of poor quality. Biological and spatial survey estimates that can be used for such purposes are discussed in Cotter et al. (2009); Woillez et al. (2009).

The concept of Management Strategy Evaluation (MSE) was developed as a tool in the 1990s by the Scientific Committee of International Whaling Commission (Kirkwood 1997; IWC 1999; Punt and Donavan 2007; Hammond and Donovan in press). It was further developed for fisheries management in South Africa (Butterworth and Bergh1993; Butterworth et al. 1997; Cochrane et al. 1998; Geromont et al. 1999; De Oliveira and Butterworth 2004; Johnston and Butterworth 2005) and Australia (Smith 1993; Punt and Smith 1999; Punt et al. 2001; Tuck et al. 2003; Campbell and Dowling 2005; Dichmont et al. 2005; Punt et al. 2005). It can be a powerful tool for assessing the potential performance of candidate management strategies (Butterworth et al. 1997; Butterworth and Punt 1999; Punt and Smith 1999; De Oliveira and Butterworth 2004; Campbell and Dowling 2005) and there are examples where fishery-independent, survey-based management procedures have been evaluated and are used in practice (De Oliveira and Butterworth 2004). The aim of MSE is to evaluate quantitatively by computer-based simulation the expected performance (relative to the uncertainties in the system and for a given set of objectives) of candidate management strategies prior to implementation.

The research described in this paper was undertaken as a contribution to the FISBOAT project, which is part of the sixth framework programme (policy oriented research) funded by the European Commission. The objectives of this project were to develop research survey-based tools allowing for fishery independent assessments and to evaluate the developed tools in their ability to provide quantitative advice on management options. With the aim of the FISBOAT project in mind, an evaluation framework is developed that includes the development of several modelling components (a population dynamics model (or "operating model"), an observation error model, an assessment model and a harvest control rules model) (Fig. 1) (Hillary 2009). The robustness of management procedures based on fisheries independent data in stock assessment and fisheries management will be evaluated using this particular evaluation framework (FISBOAT 2003).

The objective of this paper is to use MSE to evaluate fishery-independent management strategies. The implementation of the MSE in the FLR environment (Fisheries Library in R) as described in Hillary (2009) is applied to three different case studies: North Sea herring (*Clupea harengus*), Bay of Biscay anchovy (*Engraulis encrasicolus*) and North Sea cod (*Gadus morhua*). These case studies include species of different behaviours (demersal and pelagic) and life spans (long or short) that entail different survey types (bottom trawl, hydroacoustics and ichthyoplankton). In all three cases, traditional management has not shown convincing results, stressing the need for an alternative approach. In fact, the spawning stock biomass of the three case studies used in this work has been highly fluctuant in the last 25 years and all of them have shown a decline in recent years (ICES 2006a–c).

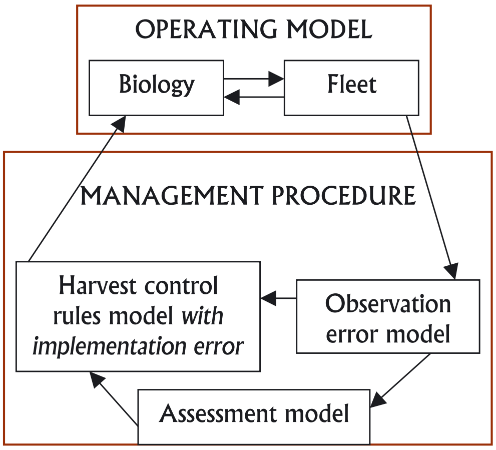

Fig. 1. The simulation evaluation platform.

First, we describe the three stocks and fisheries, and how the FLR operating models are conditioned for each stock. Then, we set out criteria to be used for assessing the performance of the candidate management strategies. Finally, we report results of the simulations, and summarise the general findings. A general discussion concludes the paper.

2 Summaries of fisheries for reference stocks

The three case studies presented here are important commercial species in Europe: North Sea herring, Bay of Biscay anchovy and North Sea cod. At the Rio Earth Summit in 1992 the principle of the precautionary approach was well supported. The sentence that might summarise it best is "the lack of scientific certainty is no reason to postpone action to avoid potentially serious or irreversible harm to the environment". This principle is internationally agreed and can be applied to any marine renewable resource. For the three case studies mentioned above precautionary spawning biomass levels have been defined (B_{PA} and B_{lim}) so as to provide managers with reference points with which to assess stock status and associated management actions.

2.1 North Sea herring

The herring fishery expanded in the 19th century but the rapid development of industrial fishing in the 20th century led to a collapse of the stock in the 1970s and subsequently, after a short recovery, in the middle of the 1990s. Since then, the 2001 cohort was the strongest seen but recruitment has declined since. In 2008, it corresponded to 10% of that estimated in 2001 (ICES 2008).

Between 1972 and 1995, assessments were performed using a Virtual Population Analysis (VPA) (ICES 1991) but because of the uncertainty in the assessment, the ICES Herring Assessment Working Group (HAWG) switched to an Integrated Catch Analysis (ICA) method (Patterson and Melvin 1996) in 1995. In 2007, ICES classified the stock as *"being at risk of having reduced reproductive capacity and at risk* *of being harvested unsustainably"* and estimated the spawning stock biomass at 1 230 000 t (ICES 2007a,b). $B_{\rm lim}$ is set at 800 000 t and $B_{\rm PA}$, at 1 300 000 t.

2.2 Bay of Biscay anchovy

Since 2002 the anchovy stock has been at very low levels - between $B_{\rm lim}$ and $B_{\rm PA}$ set at 21 000 t and 33 000 t, respectively – with 2005 the lowest point of the historical series. After the failure of the fishery in spring 2005 the fishery was closed allowing only a provisional quota in spring 2006 and some experimental fishing with spatio-temporal restrictions in spring 2007. Based on survey and commercial catch data, assessments of the stock using ICA have been conducted by ICES by the Working Group on the assessment of mackerel, horse mackerel, sardine and anchovy (WGMHSA) up until 2007 and by the Working Group on anchovy (WGANC) since then.

The stock has been traditionally managed by a fixed annual allowable catch (TAC). However, since this TAC is set regardless of the incoming recruitment, which forms the major part of the stock, it has little regulatory effect. Furthermore, in the case of a series of low recruitments, over-exploitation of the stock could occur rapidly. Currently, the European Commission is considering a long-term management plan for the anchovy.

2.3 North Sea cod

Cod in the North Sea is the main target species of the demersal fisheries of several European countries, e.g. Denmark, Germany and the United Kingdom. The stock decreased from the 1970s leading to a dramatic decline in catches from 1980. Since the late 1990s, several cod recovery plans have been adopted with the aim of increasing the spawning stock biomass (SSB) of North Sea cod above the precautionary limit ($B_{\rm PA}$) of 150 000 t. However, in 2007, ICES classified the North Sea cod stock as "*being at risk of being harvested unsustainably*" (ICES 2007a) and stock assessment models have estimated a continuing decline, SSB being well under the 70 000 t limit ($B_{\rm lim}$) below which the stock is expected to suffer reduced reproductive capacity. The extended survivor analysis (XSA) has been used as assessment method until 2004 and *B*-Adapt has been used since.

3 Methods

We describe the modelling implemented for all case studies within the FLR library and specificities for each case study are provided for each element of the simulation framework. The terminology used is that employed in the paper by (Rademeyer et al. 2007). The FLR library (Kell et al. 2007) was designed as an open-source framework, built within the R statistical environment, for the design and evaluation of harvesting strategies. We use the FLR library to design and test an array of survey-based operational management procedures for the given case-study stocks: North Sea herring, Bay of Biscay anchovy and North Sea cod.

Population dynamics models are based on expert knowledge, available data, assumptions and hypotheses. The "real" dynamics cannot be known with accuracy and the aim of MSE is to test the robustness of the strategies based on alternative hypotheses to these "true" dynamics and to meet requirements of the precautionary approach to fisheries management. The link between the "real" and the "observed" systems are observations that are based on outputs of the operating model and the implementation of management actions will act on the "real" system (FAO 1996; Kell et al. 2007). The "real" dynamics and the "observed" data on the system are developed in simulation models. The operating model corresponds to the "real" system. It simulates the biological dynamics and conditioned (parameters and structure estimated in some way) to be consistent with the current data and assumptions about the biology. The "observed" one contains three elements: the observation error model, the assessment model and the harvest control rules model, the last two being grouped in the management model. The observation error model simulates the observation of the stocks via surveys, where the level and structure of the error of observation is, if possible, consistent with available information. The assessment model (model-free or model-based) assesses the status of the stock and the harvest control rules model dictates (given any potential implementation error) the fishery actions for the next period (thus closing the loop).

The methods used in this work are the one described in Hillary (2009), adapted as required for each of the three case studies. Only a brief description of the methods is provided here. There are three major components to the operating model: the biological model, the observation error model and the management model. Each must be conditioned (Rademeyer et al. 2007) meaning that the parameters (and associated model outputs) should reasonably reflect the reality of the system they attempt to simulate (Table 1).

3.1 Operating model

The biological operating model is a yearly, seasonal (within-year periods permitted), and age-structured stochastic population and fishery model. The specifics of the model can be found in Hillary (2009) but at a base level two main processes must be parameterised to run the model:

1. The manner (magnitude, uncertainty and timing) in which recruitment (and its potential relationship with the spawning stock) occurs.
2. Future selectivity pattern(s) dictating the vulnerability/targeting of future age-classes by the fishery.

The biological model permits a wide-range of potential recruitment dynamics (stock-recruit relationship, median recruitment level with structured noise, bootstrap of historical recruitment series). For the three case studies considered in this paper an array of spawner-recruit relationships were fitted to the historical data, using the FLSR class and methods in the FLR (Kell et al. 2007) main package FLCore. Maximum likelihood methods are employed to estimate the relevant parameters of the spawner-recruit (S-R) relationship. While classical model

Table 1. List of components of the management strategy evaluation for each case study.

	North Sea herring	Bay of Biscay anchovy	North Sea cod
Initial conditions			
Age classes	10 (0 to 9+)	6 (0 to 5+)	7 (1 to 7+)
Historical data	1960-2006	1987-2005	1963-2005
Stock weight at age	Picked randomly from the historical data	Average of the historical series	Average of the historical series
Natural mortality	Constant over time	Constant over time	Constant over time
Maturity at age	Selected randomly from historical data and different over time	Constant over time	Constant over time
Starting conditions	Last historical year	– Start to apply the HCR immediately in 2006 – Start to apply the HCR after a fishery closure for the 1st 2 years	– Start to apply the HCR immediately in 2006 – Start to apply the HCR after a fishery closure for the 1st 3 years
Stock recruit model			
Ricker	Fitted for 1960-2006	Fitted for 1987-2005	Fitted for 1963-2005
Hockey stick	Fitted for 1960-2006	/	Fitted for: – 1963-2005 – 1998-2005 only (when recruitment was low)
Selectivity model	Logistic selectivity ogive (2006)	Double normal selectivity ogive (2005)	Logistic selectivity ogive (1996-2005)
Survey indices	SSB, Z	SSB, R, Alarm index	SSB, Z trend
Observation error			
Catchability	/	= 1	0.3, 0.9, 0.95, 1.2, 1.75, 1.75, 1.75
CV	Age specific	Ranging from 0 to 1	Ranging from 0.25 to 0.50
Correlation	/	Uncorrelated multiplicative errors, independent and log-normally distributed	Uncorrelated multiplicative errors, independent and log-normally distributed
Autocorrelation	Autocorrelation, ageing error	None	None
Bias	None	None	None
Assessment model	None	None	YCC
HCR (see Table 2)			
Misreporting	0 < random ⩽ 5%	None	0, 5, 10, 25%

selection methods may be used to select an optimal relationship the work in this paper considers the various candidate S-R relationships more as alternative population-dynamic hypotheses and, as such, are all included in the MSE process. Figure 2 shows the observed and predicted S-R relationships for each case study. The CVs on the SR residuals have been calculated and vary between 0.35 (North Sea cod with a Hockey stick SR function fitted between 1998 and 2005) and 0.99 (Bay of Biscay anchovy with a Ricker SR function). All values can be found in Figure 2.

The final base requirement is to parameterise the future selectivity pattern(s). These patterns are permitted to change over time and vary stochastically if required and are defined by the user so they are either assumed or estimated in some way. For all three cases the selectivity relationships were estimated using historical data. For a given subset of the historical fishing mortality (F) estimates, these mortalities were normalised by their maximum to obtain "observed" selectivity patterns:

$$s_{a,y}^{OBS} = \frac{F_{a,y}}{\max\left(F_{a,y}\right)} \quad (1)$$

This makes the assumption that, for each year, at least one age-class in the population was fully selected. From these "observed" selectivity patterns maximum likelihood methods were then employed using the FLOgive package in FLR to parameterise the relevant selectivity relationship. Figure 3 shows the fitted selectivity relationships used for each of the case studies.

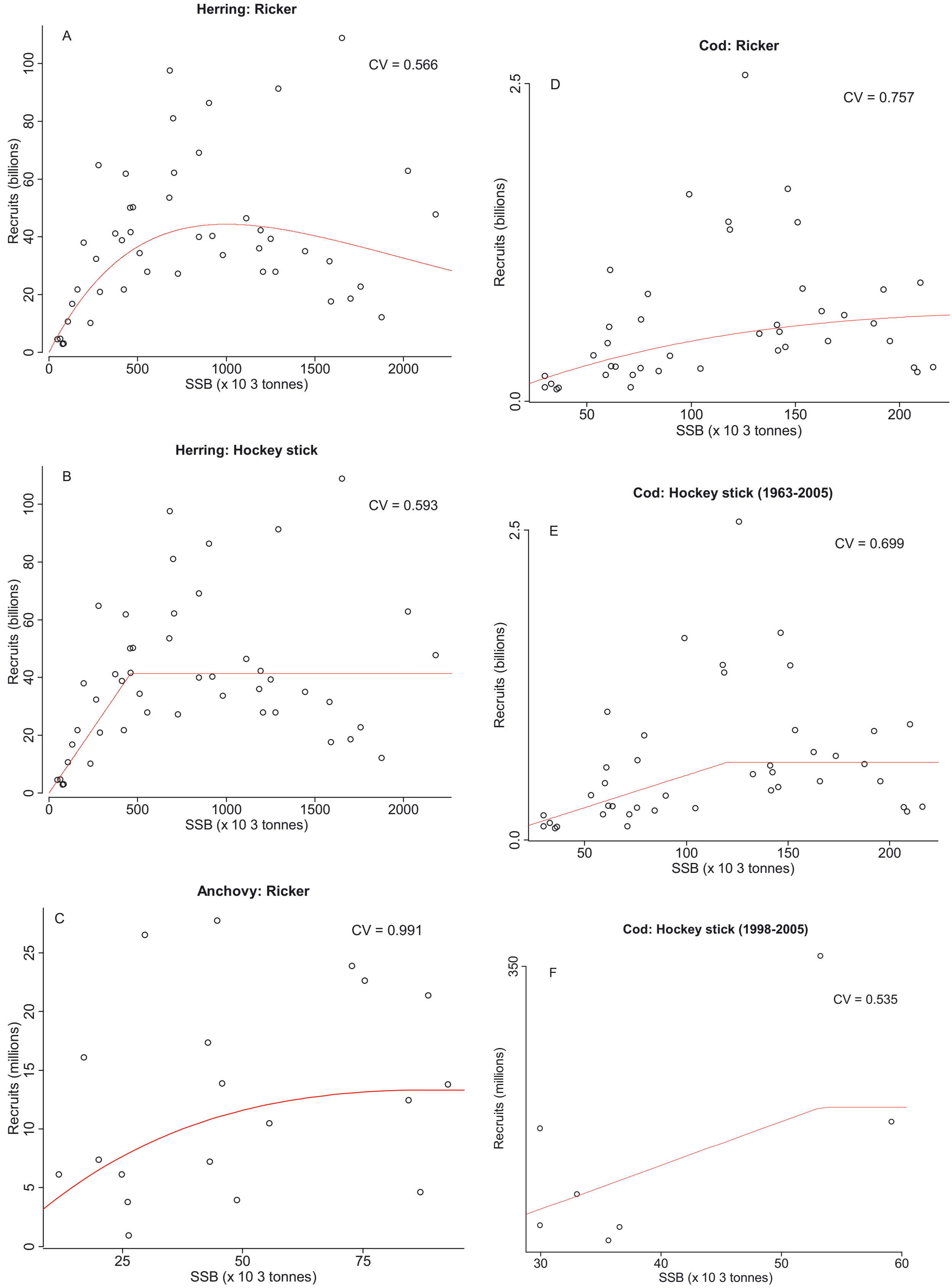

Fig. 2. Fitted and observed stock-recruit relationship for each case study. Dots: observed values, line: fitted curve. A) Herring with Ricker SR, B) Herring with hockey stick SR, C) Anchovy with Ricker SR, D) Cod with Ricker SR, E) Cod with hockey stick SR (1963-2005), F) Cod with hockey stick SR (1998-2005).

3.2 Observation error model

To evaluate the performance of survey-based harvest control rules one needs to be able to simulate the survey process which includes simulating the relevant quantity of the biological population observed by the survey and also the uncertainty (in terms of magnitude and structure) of those observations. Any observation model should at least look to parameterise these two base processes. As is the way with the biological operating model an observation error model should also be conditioned on data if available, although it seems that the level of information required to do this is not always readily available. What is required is the following:

- A clear understanding of what the survey is "observing" – is it SSB, recruitment, or perhaps a mixture of easily categorised population groupings?
- Some information on the uncertainty of the observations, preferably in terms of precision and potential covariance (if we have age/length/stage/spatially structured survey data)

For North Sea herring, the only survey simulated was the acoustic survey (the primary abundance and mortality data source) and a covariance matrix-at-age for the survey proportions was available (ICES 2001) meaning a full conditioning of the observation error model was feasible. Following the approach detailed in Hillary (2009) the delta method is used to translate the proportions-at-age covariance matrix into the logit domain, where observation error is then simulated as a multivariate normal process with the re-sampled survey data then transformed back through the inverse-logit function (and renormalised) to generate the survey data with the (approximately) appropriately covariant observation error. The catchability index (q) is used for relative SSB and for total mortality (Z) which (as long as q does not change with time) are unaffected by the value of q; $q = 1$ was chosen as in theory, the acoustic survey is a survey of the whole population covered (1–9).

For the Bay of Biscay anchovy case study there was no available information on the precision of the surveys. The approach adopted was to simulate unbiased observations of the spawning stock biomass (thus simulating the potential observations from the acoustic or egg production surveys currently undertaken) for a range of observation error CVs (assuming lognormal error in the observations). Secondly, a recruitment survey was also simulated to offer potential insights as to the utility of such a survey for survey-based management purposes if one were feasible/undertaken in the future. In the present study, as shown in Table 1, the catchability q is set equal to 1. In the assessment, q for the SSB from the recruitment survey is taken as 1 (and not estimated) and the q for the SSB from acoustics is estimated.

In the case of North Sea cod, no direct information on the precision/potential covariance-at-age of the observation error in the simulated International Bottom Trawl Survey (IBTS) was available. The variance in the residuals of the fits of the IBTS survey in the assessment for North Sea cod was used as an indicator of the potential level of observation error in the survey. It should be noted that this is always likely to be an over-estimator as the noise in these residuals may be considered as a sum of both observation and process error. Given these investigations a lognormal coefficient of variation (CV) of 0.25 (across all age classes) was used as the base-case. Various catchabilities were tested (Table 1).

3.3 Harvest control rules model

The management model consists of three processes: stock assessment, calculation of a suitable harvest level via the HCR, and implementation of the management recommendation(s).

Stock assessment, at least for the purposes of the work considered here, is anything that uses the data to both infer stock status and derive quantities for use in the harvest control rule. Model-based assessments are those which actively estimate quantities of interest in some statistical fashion with some underlying population/assessment model. Model-free assessments are where quantities are derived from the available data but not statistically estimate conditional on a population/assessment model. For example, the year-class curve (YCC) algorithm (Cotter et al. 2007) is used as a means of estimating cohort-specific total mortality, assuming a simple linear structure and represents a model-based assessment. Deriving total mortality-at-age directly from survey data assumes the same population model as YCC but is a model-free assessment as the quantities are derived directly from the data, not "filtered" in some statistical fashion via the population model. More detailed survey-based assessments do exist such as SURBA (Needle 2005) but were not available in the FLR suite of assessments in a stable enough format for use in the work presented, although their potential utility is discussed later on.

The key part of the management model is the form of the harvest control rules considered for the case-studies. Management was effected (for all case studies) via alterations in the TAC for each stock and in consideration of this various HCRs were considered based on the following general form:

$$\begin{cases} TAC_{y+1} = \min\left\{\max\left\{\exp\{u_y^1\}, \exp\{u_y^2\}\right\}, \exp\{u_y^3\}\right\} TAC_y \\ u_y = K_P\, e_y + K_I \sum\limits_{z=y-\delta}^{y} e_z + K_D\left(e_y - e_{y-1}\right). \end{cases} \tag{2}$$

Here, u_y denotes the control signal in year y that is used for TAC adjustment from year y to the next and in control theory this type of system is known as a PID (proportional, integral, derivative) controller. The control signal is calculated from e_y, giving the divergence of an index relative to a reference point. The survey estimate is calculated either directly from the latest survey or through an assessment conducted in year y. The desired closed-loop behaviour can be obtained by tuning the three parameters K_P, K_I and K_D with δ denoting the history that is considered in calculating the integrated/smoothed part of the control signal, e_y. The back-averaging control time-scale δ was set arbitrarily at five years in all simulations where it was used. It is important to note that only "moving targets" are considered, i.e. divergence is always relative to the index in the previous year (Apostolaki and Hillary 2009). This general formulation was then made specific for each case study-specific HCR as required and the details of these specifics are given later in the paper. The three control parameters of the HCR (K_P, K_I and K_D) can be tuned in various ways (O'Dwyer 2003). A common method, named after Ziegler and Nichols and often applied in industrial engineering, is to first set K_I and K_D to zero and increase K_P until the output of the control loop starts to oscillate with constant amplitude. The value of K_P at this point is called the critical gain and can be combined with the value of the oscillation period to arrive at standard tunings for K_P, K_I and/or K_D, depending on whether strictly proportional (P), proportional-integral (PI), or proportional-integral-derivative (PID) control is desirable. Here, K_P represents the *proportional* response given the control signal, K_I is the *integral* part of the response given the aggregation (sum) of recent control signals, and K_D is the *derivative* part of the response driven by the rate at which the control signal varies. This type of tuning was done only for the North Sea cod case (applied to deterministic model runs) to ensure that the number of simulation trials to be evaluated and summarized did not grow too large to detail in the paper.

The implementation model was decomposed into two processes: (1) random implementation error which assumes some unbiased variation around (2) a specified implementation bias multiplier i.e. a number set less than, equal to, or greater than one to simulate over-reporting, unbiased reporting, and under-reporting, respectively. For the North Sea herring case a level of mean under-reporting by 2.5% with a further 2.5% random error was set to simulate small amounts of average under-reporting but with reasonable uncertainty as to the specific value of the under-reporting. For the Bay of Biscay anchovy study no implementation error was considered and for the North Sea Cod case study levels of mean under-reporting of up to 25% were included in some simulation trials.

3.4 Application to the case studies

The various HCRs tested are listed in Table 2. A brief explanation of the HCRs is provided here and details can be found in the Appendix.

The first HCR applied to all stocks serves as a reference scenario and corresponds to the case in which no fishing is allowed (HCR_0). The second HCR, also applied to all case studies, consists on keeping the TAC at a constant level defined by the latest historical catch level (HCR_1). In case of North Sea herring and Bay of Biscay anchovy, HCRs considered are pure P-controllers which means that K_I and K_D are always equal to zero (Table 2). Also, if no min, max functions are used it is possible to simplify the Equation (2). By manipulating the values of K_P, K_I and K_D it is possible to recreate the HCRs implemented for herring and anchovy ($HCR_{2,3}$ and HCR_6), but also HCR_7 applied to North Sea cod. If min and max functions are used u_y^1, u_y^2 and u_y^3 have different values. This is the case for example in HCR_4 and HCR_6 applied to Bay of Biscay anchovy and North Sea herring respectively. As explained in the appendix, the aim of the HCR_4 is to constraint the variation of the TAC by a set percentage. The control signal is calculated and its value is compared to the lower limit of the allowed variation and the largest value is kept. The later is then compared to the higher limit of the allowed variation and the smallest is kept and used in the calculation of the TAC for the following year. In the case of HCR_6, two control signals are calculated and the smallest is used in the calculation of the TAC.

Contrary to herring and anchovy, in the case of the North Sea cod case study, various K_P, K_I and K_D tunings were tested (with $\delta = 5$) and u_y values were kept equal to each other, i.e. no use of min or max functions.

Following the two initial HCRs (HCR_0 and HCR_1), several SSB-based HCRs have been implemented for the three case studies (HCR_2 to HCR_5). In most cases, the index used is based directly on the SSB observations from the surveys ($S\hat{S}B_y$). To test the value of the indicator approach, a binary alarm ($\hat{A}_y$) was included in the simulations for the anchovy (HCR_3). When the true population biomass was below B_{lim}, the alarm was triggered with probability 0.9 ($P\left(\hat{A}_y = 1 | SSB_y < B_{\text{lim}}\right) = 0.9$), and when the true population biomass was above B_{lim}, it was triggered with probability 0.05 ($P\left(\hat{A}_y = 1 | SSB_y \geqslant B_{\text{lim}}\right) = 0.05$). The alarm operated independently of the abundance indices observed. Its purpose was to modify the HCR to a more restrictive one. This method modifies TAC according to the relative change in SSB, reducing the exploitation if an alarm is triggered. The TAC is reduced automatically by a fraction α in case the indicator triggers an alarm which depends on the true population biomass. Henceforth, the control parameter α can take three values: 0.25, 0.5 or 0.75.

The last three HCRs use indices based on total mortality rate, either coming directly from observation from the survey (HCR_6 and HCR_7) or from the output of the assessment (HCR_8). HCR_6 is based on age non-aggregated index. The index is estimated for 2 age groups of the stock and the minimum of the two ratios is used in the calculation of the TAC. The aim of this rule is to try and mirror aspects of the actual decision rule applied to North Sea herring, but based on total mortality levels on the juvenile and adult sections of the stock. Values of the total mortality rate at age at the precautionary level (Z_{PA}) for each age group derive from other parameter values defined by ICES ($Z_{[0,1]}^{PA} = 1.12$, $Z_{[2,6]}^{PA} = 0.4$; see Appendix for details).

Several starting conditions were assumed across the three case studies:

For the North Sea herring case study, only one starting condition has been implemented: the first year of projection is 2007 and starting conditions are those estimated by the Working Group. For the Bay of Biscay anchovy case study, two starting conditions have been tested, as shown in Table 1: (i) start to apply the HCR immediately in 2006 (condition A) and (ii) start to apply the HCR after a fishery closure for the first two years (2006 and 2007) to let the population recover slightly, followed by a TAC of 30 000 t for 2008 (condition B). Also, for the North Sea cod case study two initial starting conditions have been tested, as shown in Table 1: (i) start to apply the HCR immediately in 2006 and (ii) start to apply the HCR after a ban on cod fishery for the first three years (2006, 2007 and 2008) to let the population recover. After this, the HCR was applied with the TAC in 2009 being modified from the catch in 2005. A sensitivity analysis showed that when HCR_8 is used, the model does not provide the expected results. In fact, this assessment method may be incapable of fully detecting the signal coming from the survey data in the presence of observation error and given the restrictions in the model structure and estimate parameters. These reasons could explain the behaviour of the model simulations over time. As the SSB is

Table 2. Harvest control rules (HCR) for each case study. e_y: divergence of an index relative to a reference point; North Sea herring, North Sea cod, Bay of Biscay anchovy; K_P, K_I and K_D: three control parameters of the HCR.

		Case study	K_P	K_I	K_D
No exploitation					
HCR$_0$: $TAC_{y+1} = 0$		Herring	0	0	0
		Anchovy	0	0	0
		Cod	0	0	0
Constant *TAC*					
HCR$_1$: $TAC_{y+1} = TAC_y$		Herring	1	0	0
		Anchovy	1	0	0
		Cod	1	0	0
SSB-based					
HCR$_2$:	$e_y = \log\left(\frac{S\hat{S}B_y}{S\hat{S}B_{y-1}}\right)$	Herring	1	0	0
		Anchovy	1	0	0
		Cod	0.48	0.069	0.84
		Cod	0.48	0.027	0.84
		Cod	0.4	0	0
HCR$_3$:	$e_y = \log\left(\alpha(\hat{A}_y)\frac{S\hat{S}B_y}{S\hat{S}B_{y-1}}\right)$	Anchovy	1	0	0
HCR$_4$:	$e_y = \min\left\{\max\left\{\log\left(\frac{S\hat{S}B_y}{S\hat{S}B_{y-1}}\right), \log 0.8\right\} \log 1.2\right\}$	Anchovy	1	0	0
HCR$_5$:	$e_y = \log\left(\frac{\overline{SSB}_y}{\overline{SSB}_{y-1}} \frac{\hat{R}_y}{\hat{R}_{y-1}}\right)$	Anchovy	1	0	0
Z-based					
HCR$_6$:	$e_y = \min\left\{\log\left(\frac{Z^{PA}_{[0,1]}}{\hat{Z}_{y-1,[0,1]}}\right), \log\left(\frac{Z^{PA}_{[2,6]}}{\hat{Z}_{y-1,[2,6]}}\right)\right\}$	Herring	1	0	0
HCR$_7$:	$e_y = \log\left(\frac{\hat{Z}_y}{\hat{Z}_{y-1}}\right)$	Cod			
HCR$_8$:	$e_y = \log\left(\frac{\overline{Z}_y}{\overline{\overline{Z}}_{y-1}}\right)$	Cod	0.32	0.02	0
		Cod	0.35	0	0
		Cod	0.175	0	0

very high, in theory, catch should be possible but this model predicts a collapse of the fishery, i.e. catch almost null. In fact, with YCC, Z is constant over ages and the total mortality is cumulative. As YCC does not estimate an age-dependent Z, a high level of implementation error will have a larger effect on the older ages given differential selection. In fact, in theory, the total mortality-at-age for a given year is equal to the sum of the natural mortality-at-age and the product of the fishing mortality for this given year and the selectivity-at-age ($Z_{y,a} = M_a + F_y \times s_a$). When YCC is used as an assessment method and as the level of implementation increases the fishing mortality increases towards a value of 1. As the fishing mortality tends to 1, the total mortality-at-age will become almost independent of the age structured ($Z_{y,a} \approx Z_y$). As a consequence and to limit this behaviour of the model over time, it had been decided to impose the closure of the fishery over 3 years for all simulations using the HCR$_8$.

In terms of projections, the number of years is different for each case study. The generation time of North Sea herring is about 10 years, and the recent history of the stock has included two periods of major depletion over the last 40 years (a major crash occurred in the late 1970s and a minor one in the middle 1990s). Based on these observations, it is assumed that 10 years can be considered as a medium term simulation. In order to provide a strong basis for understanding long-term performance, all simulations were run for 20 years (2 generation times). In each of the runs for the Bay of Biscay anchovy stock, the population is projected forward for 10 years (3 generations). In the case of North Sea cod, the simulation period was fixed at 30 years, corresponding to approximately five generations. In all the case studies, performance statistics were calculated on the basis of 100 Monte Carlo iterations.

4 Performance statistics

Performance statistics were defined to summarise the results from the simulations and to evaluate the performance of the different HCRs. The statistics were related to the management objectives set for each of the stocks and fell into two groups:

a) Performance statistics related to the state of the stock

- Probability that spawning stock biomass (SSB) is below some biomass reference point, *Bref*, at least once in the time series;
- Probability that SSB is below some biomass reference point, *Bref* in the final year;
- Number of years necessary to get SSB above *Bref*.

b) Performance statistics related to yield

- Average catch over all years and iterations;
- Median percentage of inter-annual change in TAC.

5 Results

5.1 Basic validation of the management strategy

The first test of the management evaluation framework is to project the stock without catches and then with constant catches to obtain something with which one can compare the alternative dynamic management strategies. For the simplest comparative HCR (HCR_0) applied to all case studies, we see stock growth (as we would expect given the dynamics) for no catch taken – under the obvious assumption that the stock-recruit, growth, maturity and natural mortality dynamics are unchanged in the near future. In the case of anchovy, even if the simulations start from the lowest level of the population observed since 1987, the stock recovers rapidly, getting in four years to the highest SSB in the series.

The second HCR (HCR_1) evaluates the impact of a fixed annual TAC (set as the last historical catch) on the stock. In both the herring and cod cases, results show that, conditional on the parameterization of the model, the current official catch level (reported landings and estimated discards) is sustainable if compliance can be enforced. In fact, because the TAC is not adjusted in response to a change in the SSB and maintained at a level which does not compensate the growth of the stock, the harvest rate declines. This case is useful in the sense that the most recent catch may act as a benchmark given that catches of this level afford stock growth. With regard to anchovy, the traditional management procedure with a fixed annual TAC resulted in a high risk of stock collapse even if the stock was not over-exploited before the application of the TAC.

5.2 SSB-based harvest control rules

Several HCRs based on SSB survey indices have been implemented and tested. In the case of HCR_2, which is applied to the three case studies, various outcomes may be expected due to different initial stock levels for the three case studies, the lack of a specific reference point in the HCR, and the different life-histories involved.

In the case of North Sea herring, we expect to see a decrease in the SSB in the first years of the projection (given the weak year-classes of 2002 and 2003 moving through the mature population) followed by an increase and then a tendency to equilibrium. In fact, the SSB starts by decreasing in the first two years of the projection. It then increases for five years until it reaches quasi-equilibrium levels – values being well above B_{PA} – as catches have been reduced. In fact, the TAC follows the trend of the SSB with a delay induced by the way TAC is calculated. During the last 13 years of the projection, the SSB is not constant due to the uncertainty in the system. Over the 100 iterations, the probability of being below B_{lim} at least once in the projection is zero. The change of the stock-recruit function (Ricker instead of a segmented regression) resulted in very small changes in the final SSB – being slightly smaller with a Ricker stock recruit function – and a very small increase in the probability of being below B_{lim} at least once in the projection (0.01 instead of 0). The probability of being below B_{PA} in the final year increases from 0 to 0.03 with the change of stock recruit function from hockey-stick to Ricker (Fig. 4A). Catches decrease for the first two or three years and then increase and stabilise. Due to the behaviour of the different stock-recruit functions, average catch over all projected years is higher when a hockey-stick relationship is assumed than with a Ricker one (Fig. 4B, Table 4). The CV in the TAC in the final year is equal to 0.264. The use of a Ricker model implies smaller values of TAC in the final year as the TAC is related to the SSB and a slight increase in the value of the CV in the TAC (Table 3).

Table 3. Performance of the harvest control rules tested. K_P, K_I and K_D: three control parameters of the HCR: # $K_P = 1$, $K_I = 0$, $K_D = 0$ in all HCRs considered; * K_P, K_I, K_D: variable tunings considered. Performance of HCR in attempt to stabilise catches at an approximate historical level: (–), (+), or (++).

	North Sea herring #	Bay of Biscay anchovy #	North Sea cod
No exploitation			
HCR_0:	√ (NA)	√ (NA)	√ (NA)
Constant TAC			
HCR_1:	√ (NA)	√ (NA)	√ (NA)
SSB-based			
HCR_2:	√ (++)	√	√* (++)
HCR_3:		√ (+)	
HCR_4:		√ (+)	
HCR_5:		√ (++)	
Z-based			
HCR_6:	√ (++)		
HCR_7:			√* (-)
HCR_8:			√* (++)

When this HCR is applied to anchovy, the results obtained are noticeably different. When the observation error CV of the SSB index is 25%, the probability of SSB being below B_{lim} at least once in the 10 years of projection is 0.11 and takes less than 1 year on average to get the population back above B_{lim}. Based on the low starting level (*starting condition A*), this means that the population starts outside the biological limits but recovers rapidly at low exploitation levels (4000 t) minimizing the depletion probability. The CV of the SSB index has not an influence on the state of the stock, probably due to the low exploitation levels. However, the larger the CV, the larger the TAC and its variability. Similarly, when the rule is tested by starting the simulations after two years of fishery closure and an initial TAC of 30 000 t (*starting condition B*), the probability of SSB being below B_{lim} at least once in the 10 years of projection raises to 0.53, needing 2 years on average to get the population back above B_{lim}. In comparison with the other initial conditions, the average actual catch is larger (approximately 29 000 t) and the inter-annual change in TAC is centred on zero. In this case, the effect of increasing the CV of the SSB index leads to increasing number of years required to recover

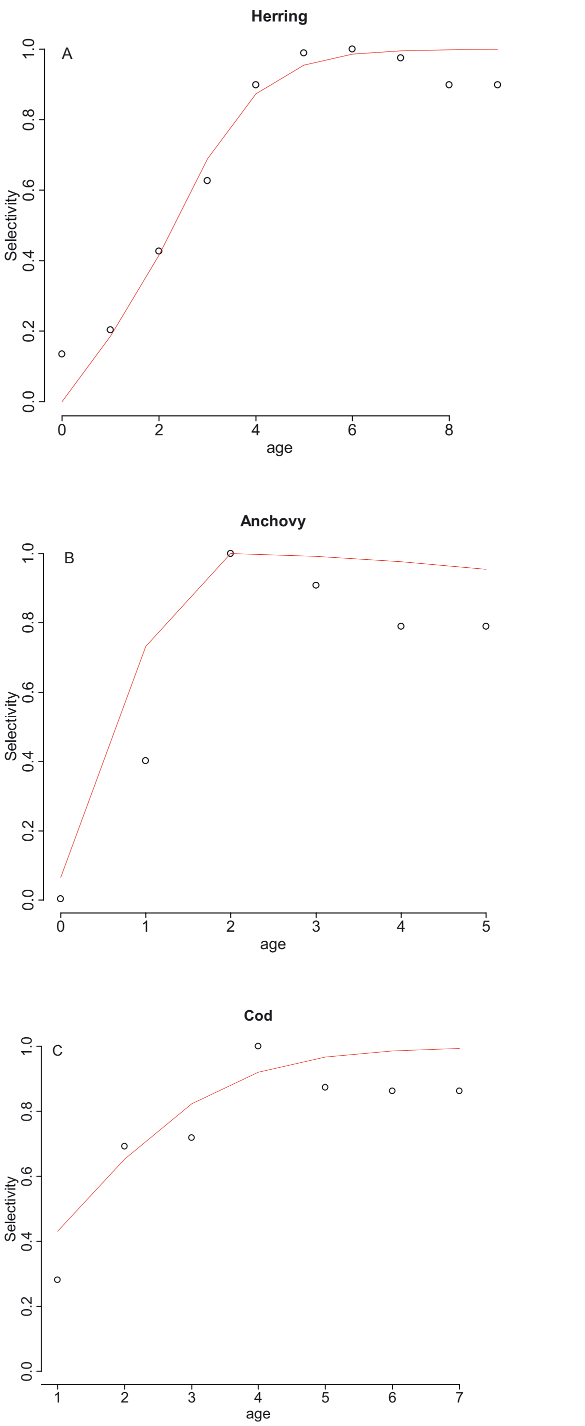

Fig. 3. Fitted (full line) versus stock assessment-derived "observed" selectivity (circles) for each case study. A) Herring, B) Anchovy, C) Cod.

from stock depletion and increasing probability of the TAC being larger than the exploitable stock (Table 3).

With regard to the other HCRs applied to anchovy, performance of all the HCRs except HCR_6 depended on the starting conditions, A or B, and especially on the starting TAC (Fig. 5, Table 5). Starting from the current stock situation with a very low TAC (*condition A*) resulted in stock rebuilding and a slow

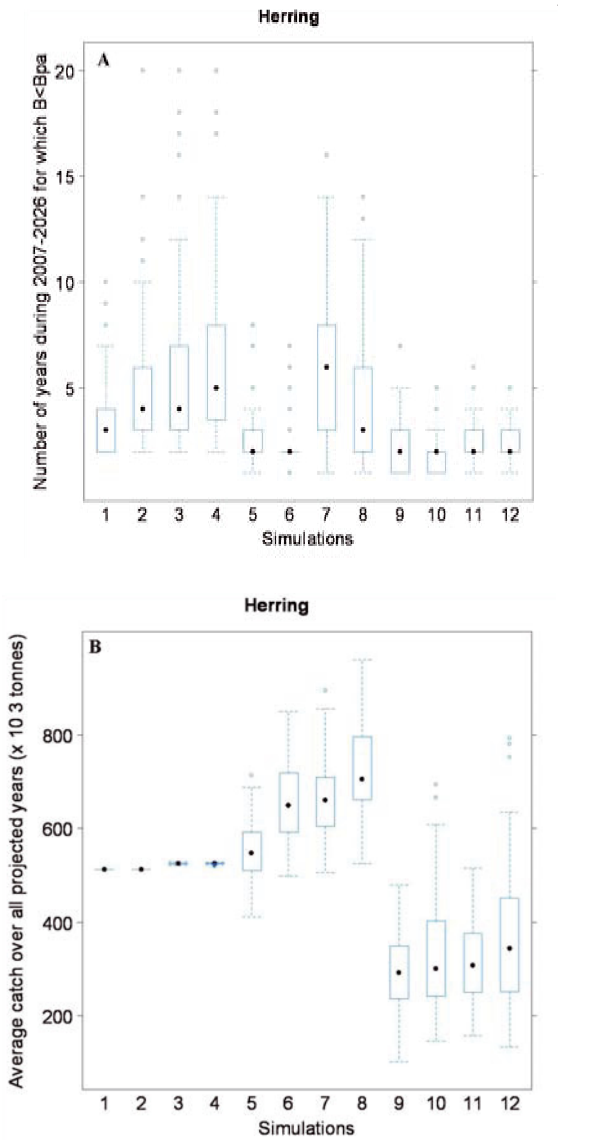

Fig. 4. Performance statistics for different stock recruitment functions, harvest control rules and levels of misreporting for North Sea herring. A) Summary of risk of biomass, $B < B_{PA}$, B) Summary of average catch ($\times 10^3$ tonnes) over all projected years.

Table 4. Simulations 1 to 12, North Sea herring.

Simulation	HCR	Stock-recruitment	Misreporting
1	1	Ricker	No
2	1	Hockey stick	No
3	1	Ricker	Yes
4	1	Hockey stick	Yes
5	2	Ricker	No
6	2	Hockey stick	No
7	2	Ricker	Yes
8	2	Hockey stick	Yes
9	6	Ricker	No
10	6	Hockey stick	No
11	6	Ricker	Yes
12	6	Hockey stick	Yes

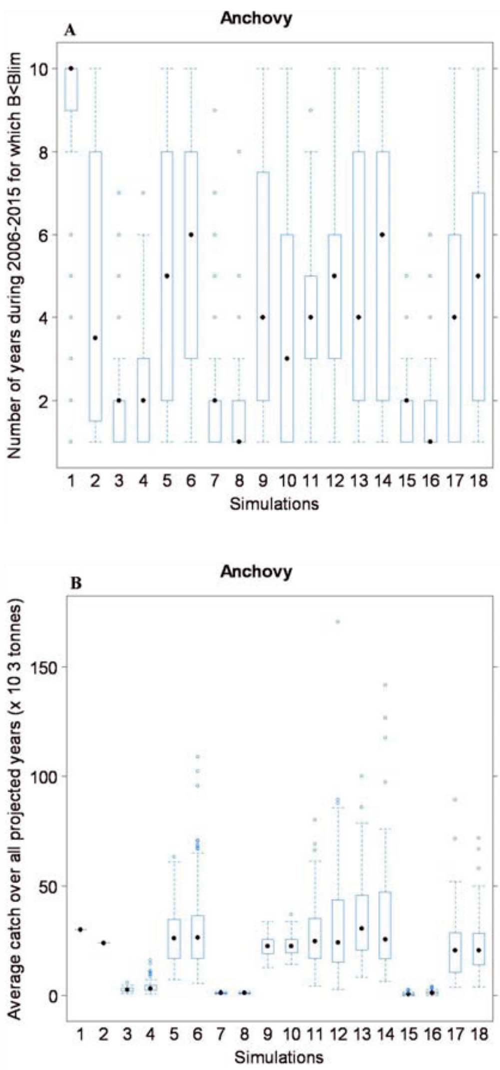

Fig. 5. Performance statistics for different harvest control rules, initial conditions, CVs for indices (SSB and recruitment) and levels of alarm trigger for Bay of Biscay anchovy. A) Summary of risk of $B < B_{\text{lim}}$, B) Summary of average catch ($\times 10^3$ tonnes) over all projected years.

increase in catches. However, closing the fishery for two years and then starting with the usual TAC of 30 000 t in 2008 (*condition B*), tended to favour high catches in the first years of the projection and resulted in a slightly decreasing rate in the projected SSB. The use of the juvenile abundance index, through the use of the simulated recruitment index, in HCR_6 allowed adjustment of the TAC with better knowledge of the state of the stock in the forthcoming years. This led to larger catches while keeping the probability of the stock falling below the relevant reference points relatively low (below 20%). The incorporation of an alarm triggering indicator with a highly risk-averse reduction factor provided better protection against stock depletion than HCR_2. As a consequence, the level of catch at the end of the projection was smaller when HCR_3 was considered (Table 3).

Table 5. Simulations 1 to 18, Bay of Biscay anchovy. Initial conditions: A) start to apply the HCR immediately in 2006 and B) start to apply the HCR after a ban on stock fishery for the first two years (2006 and 2007). α: control parameter of the HCR.

Simulation	HCR	Initial condition	CV SSB index	CV recruitment index	α
1	1	A	/	/	/
2	1	B	/	/	/
3	2	A	0.25	/	/
4	2	A	0.75	/	/
5	2	B	0.25	/	/
6	2	B	0.75	/	/
7	4	A	0.25	/	/
8	4	A	0.75	/	/
9	4	B	0.25	/	/
10	4	B	0.75	/	/
11	5	A	0.25	0.3	/
12	5	A	0.25	0.9	/
13	5	B	0.25	0.3	/
14	5	B	0.25	0.9	/
15	3	A	0.25	/	0.25
16	3	A	0.25	/	0.5
17	3	B	0.25	/	0.25
18	3	B	0.25	/	0.5

When considering HCR_2 for North Sea cod, using a "hockey-stick" stock recruitment function fitted to SSB and recruitment data for 1998-2005 only, or a Ricker-type stock recruitment function, both resulted in an increased risk of stock collapse (Fig. 6A). The low level "hockey-stick" stock recruitment function gave a probability of having a TAC above the Exploitable Stock Biomass (ESB) of 0.03, while the Ricker-type stock recruitment function gave a probability of 0.2. Both simulations had significantly reduced catches: 132 000 (standard error, SE 22) and 298 000 (SE 59) t, respectively (Fig. 6B) with corresponding inter-annual TAC CVs of 0.32 and 0.41. The lower level of recruitment in the alternative "hockey-stick" stock recruitment function significantly increased the time to stock recovery (SSB 95% of being above B_{PA} as of 2014).

In terms of recruitment, a slightly different shape is expected depending on the stock recruitment relationship assumed. If a hockey stick relationship is assumed, the recruitment stays constant all along the projection once the threshold SSB has been reached. In the case of a Ricker relationship, the recruitment begins to decreases due to the compensatory nature of the relationship. The harvest rate is kept relatively constant over time with a small decline in the first years of the projection due to the decline of catches, given the declining SSB trend in this period.

The consideration of the random misreporting varying between 0 and 5% has an important impact on the SSB in the projection. In fact contrary to previous thought the SSB increases between 2007 and 2015 and then decreases until the final year of the projection to reach values smaller than without misreporting. The final value is above the reference points but the probability of being below those values is not negligible and reaches 0.2 for the probability of being below B_{lim} at least once in the projection and 0.4 for the one of being below B_{PA} in the final year, when either stock-recruitment relationship is assumed.

5.3 Total mortality (*Z*-based) harvest control rules

For the herring and cod case studies age-structured indices are simulated and this makes it possible to (potentially) estimate total mortality (Z) for the relevant age classes observed in the surveys. The Z-based HCRs are detailed in Table 3 (HCR_6 to HCR_8) and the appendix. For North Sea herring HCR_6 was used to attempt to roughly approximate the current management procedure used for this stock (an F-based strategy concerned with juvenile (0–1) and adult (2–6) animals). The precautionary levels of Z were derived by simply adding the appropriate level of natural mortality (averaged over the relevant age classes) to the precautionary levels of F as defined by the Working Group. For North Sea cod a simpler yearly mean Z was used as the control variable of interest in the HCR ($HCR_{7,8}$) but this mean Z was both estimated from the raw survey data (model-free) and estimated using the YCC approach.

For the case of the herring the Z-based rule displayed a much more reactive and precautionary dynamic than the SSB-based HCRs. Figure 4A shows how the total mortality HCR manages to keep the stock well away from precautionary levels over the whole management simulation period. The reactive nature of this rule, relative to those based around SSB, is not surprising given that this rule will obtain information on changes in the stock much faster than an SSB-proxy signal would. Given changes in year-class strength and observation error one can expect relatively strong changes in the levels of Z observed on especially the juvenile portion of the stock, even though in this case juvenile mortality is effectively a mean of the total mortality felt by ages 0 and 1. While these variations may tend to be smoothed as the index observes the mature portion of the stock (ages 2–6) total mortality is still a more reactive and potentially informative index than simply the SSB in any given year. The apparently more precautionary action of the HCR also stems from its definition:

- The TAC will only be increased if both the juvenile and adult mortality levels are below their respective precautionary levels, and then only by the smallest ratio of Z_{PA}/Z.
- The TAC will be decreased if only *one* of the mortality levels is above their respective precautionary levels and if both are above then the decrease in TAC will be by a fraction equating to the smallest ratio of Z_{PA}/Z.

A major criticism of the use of reference points in HCRs is that they themselves are often derived from parameters such as the stock-recruit parameters, growth and in particular natural mortality, which can often be poorly estimated and/or highly confounded. The notion of a precautionary level of total mortality is used in the HCR_6 and these juvenile/adult Z_{PA} levels are themselves derived from F_{PA} and M. The reasons for doing this were two-fold:

- The use of these Z_{PA} levels and the particular form for the HCR were an attempt to mimic certain features of the current agreed HCR

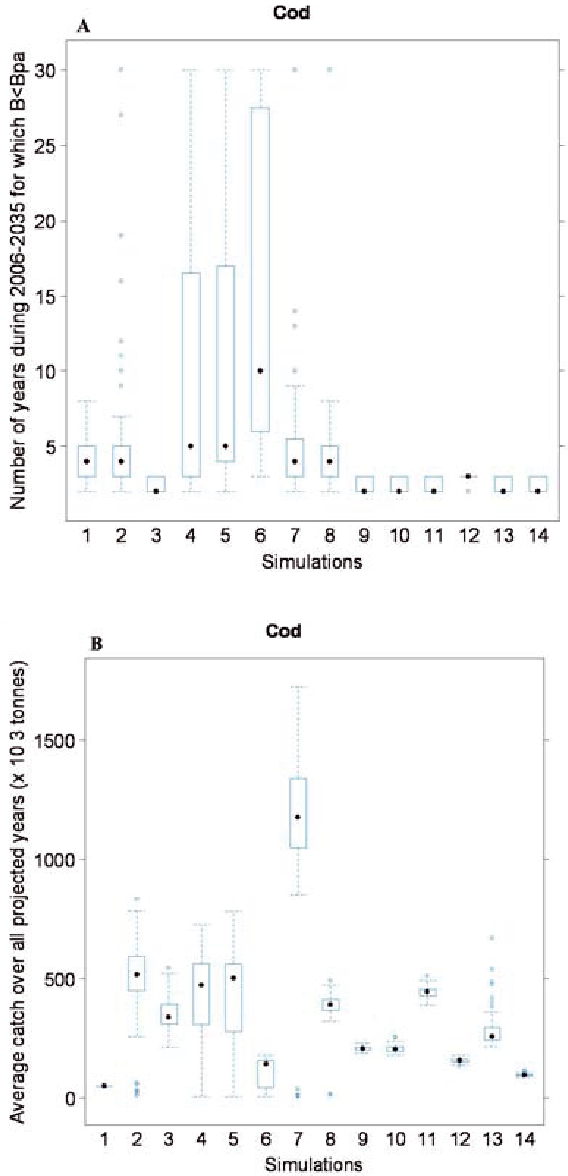

Fig. 6. Performance statistics for different stock recruitment functions, harvest control rules, initial conditions, survey noises, K parameters and levels of misreporting for North Sea cod. A) Summary of risk of $B < B_{PA}$, B) Summary of average catch ($\times 10^3$ tonnes) over all projected years.

- It is our contention that using Z and not F-based reference points are much less prone to the problems usually attributed to such reference points and are, as such, defensible for use within a HCR. The estimation of reference points such as F_{PA} (and not F_{MSY}) is usually the estimation of a level of total mortality that acts to reduce the stock (assuming equilibrium) below biologically safe levels - in essence it only becomes F_{PA} when one assumes a certain M vector for the age classes. Thus, when one reconverts F_{PA} back to Z_{PA} (using the M one assumed in the first

Table 6. Simulations 1 to 14, North Sea cod. Stock-recruit relationships: Ricker Hockey stick (HS) fitted between 1963 and 2005 then, between 1998 and 2005. Initial conditions: A) start to apply the HCR immediately in 2006 and B) start to apply the HCR after a ban on stock fishery for the first three years (2006, 2007 and 2008). K_P, K_I and K_D: three control parameters of the HCR.

Simulation	HCR	Stock-Recruitment	Initial condition	Survey noise	K_P	K_I	K_D	Misreporting
1	1	HS 1963-2005	A	Estimated	1	0	0	0
2	2	HS 1963-2005	A	Estimated	0.48	0.027	0.84	0
3	2	HS 1963-2005	B	Estimated	0.48	0.027	0.84	0
4	2	HS 1963-2005	A	Doubled	0.48	0.027	0.84	0
5	2	HS 1963-2005	A	Estimated	0.48	0.027	0.84	25%
6	2	HS 1998-2005	A	Estimated	0.48	0.027	0.84	0
7	2	Ricker	A	Estimated	0.48	0.027	0.84	0
8	2	Ricker	A	Estimated	0.4	0	0	0
9	8	HS 1963-2005	B	Estimated	0.315	0.02	0	0
10	8	HS 1963-2005	B	Doubled	0.315	0.02	0	0
11	8	HS 1963-2005	B	Estimated	0.315	0.02	0	25%
12	8	HS 1998-2005	B	Estimated	0.315	0.02	0	0
13	8	Ricker	B	Estimated	0.315	0.02	0	0
14	8	Ricker	B	Estimated	0.175	0	0	0

place) there is less dependence on the factor that was the most unknown: natural mortality.

Current catch levels result in levels of juvenile Z quite above their precautionary levels, given the estimated current population numbers and assumed selectivity function, and the decrease in the mature stock as a result of the low recruitments in the early 2000s also results in higher Z levels on the adult stock. Both these initial dynamics result in a rapid reduction in the TAC from current levels. This results in a strong increase in the SSB (after the initial dip given the low recruitments) and once the levels of Z (proxies for F in effect) have decreased below the trigger levels in about 2014 the TAC begins to gradually rise to just under recent levels after 20 years, all the while maintaining a low exploitation rate and lower inter-annual variation in the TAC than that observed using the SSB-based rules (Fig. 4B). One further observation based upon results not detailed in this paper is that the $Z-$based rule was more affected by alternate levels of observation error (in comparison with the SSB indices) becoming more reactive and resulting in more uncertain future stock and catch dynamics. This again stems from the fact that the SSB index is derived from age-structured acoustic data but the observation error is smoothed over the age-classes which does not happen to anywhere near the same degree when using the raw survey data to estimate Z in this manner.

For North Sea cod the PID tuning method was again employed to estimate an optimal HCR from the general HCR defined in Equation (2). When using a model-free estimate of annual mean Z (estimated directly from the survey), it was not possible to determine the value of K_P where the output of the control loop started to oscillate. Consequently, it was not possible to obtain a convergent set of tuning parameters and model-free Z-based HCRs were not considered viable within the PID paradigm. However, PID control was possible when using the YCC method, but only if a closure of the fishery is considered in the three first years of the projection, as explained earlier. Before addressing the reasons for this dynamic, the performance of the YCC-based HCR is detailed. In the absence of reporting bias (misreporting) and for the more "optimistic" stock-recruit regimes (based on the full data set, not the 1998-2005 subset) the Z-based rule performed well in terms of stock rebuilding, and low inter-annual variation in TAC – better than the SSB-based rules (Fig. 6A). Also the results were largely insensitive to increasing the observation error CV to 50% – mean levels of catch were the same with a higher variation in TAC and no added risk of stock collapse (Fig. 6B, Table 6). Using the truncated stock-recruitment data resulted in high risks of stock collapse (more so for the Ricker model) with catch becoming increasingly likely to be almost equal to the exploitable stock biomass in these cases.

6 Discussion

In the present work, several harvest control rules based on fishery independent data have been tested on three fish stocks with different biological characteristics (herring, anchovy and cod).

For North Sea herring, simple HCRs based upon changes in a single age-aggregated SSB index derived from the acoustic survey annually undertaken for this stock were reasonably capable of increasing the future SSB significantly above precautionary limits ($p < 0.05$) as well as increasing future TAC levels. The primary robustness test in these cases was the inclusion of covariant observation error in the acoustic survey (defined using actual survey variance information) and process error in the form of stochastic future recruitment. A Z-based HCR was also tested and proved to be a more conservative HCR – the size of the SSB increase was larger than that seen for the derived SSB trend HCRs and with an initial decrease in TAC followed by an eventual increase in TAC as the stock abundance increased. The Z-based HCR gave future median TAC levels below the current level. Not only does this HCR work with variables (age-specific mortalities) which experience stronger variation over time (in the presence of observation error and stock recruit variation) than SSB/age-aggregated indices but it also has both specific reference points and an implicit precautionary element. Once the SSB has increased, the TAC decreased and these weaker initial cohorts grow out of the age range under consideration (0–6) the incoming stronger cohorts cause a decrease in total mortality and a subsequent increase in the TAC although never back above current catch levels. Inter-annual variation in TAC was noticeably less in the case of the total mortality HCR and this particular effect seems at least in part driven by the presence of reference points in the total mortality HCR. In the SSB trend HCRs there are no target or limit reference levels, merely an adaption of the TAC as the mature biomass dynamics unfold. With the Z-based HCR there are effective target levels of total mortality for the juvenile and mature populations. This HCR acts to swiftly reduce the TAC to get to the population to these target levels then stabilises the TAC to maintain the stock around these target levels. This stability in the TAC results in the lower levels of inter-annual variation in catch levels.

Given the life-history of anchovy (short lived with fast maturation, high natural mortality and highly variable recruitment) the HCRs tested used SSB and recruitment survey indices only. For practical purposes the recruitment of one year forms the basis of the spawning stock in the following year making the idea of a total-mortality type rule practically redundant if one considers recruitment and SSB survey information. There were two clear conclusions with respect to the anchovy case study:

- Performance of all the HCRs bar one was strongly driven by the initialisation of the operating model.
- The HCR which utilised information on both recruitment and SSB appeared much less sensitive to the initialisation assumptions.

Given a depleted stock with anchovy-type life-history traits the results clearly outlined how more accurate surveys and a wider coverage of the population in terms of its life-cycle can improve the performance of a dynamic HCR, when one's key performance criterion is rebuilding.

For North Sea cod both the SSB and total mortality HCRs were evaluated over a range of robustness criteria such as stock-recruit model and data set, initial moratorium, fine-tuning the HCR using methods from control theory, and various misreporting regimes. In terms of initialisation influence, across all other robustness trials an initial catch moratorium greatly improved performance in terms of avoiding stock collapse. With respect to the particular stock-recruit function assumed the hockey-stick model yielded lower levels of both maximum recruitment and recruits-per-spawner at low stock levels, making scenarios assuming this model more prone to poor performance in terms of avoiding stock collapse. When restricting the data to the more recent period (1998-2005) both hockey-stick and Ricker models predicted even poorer recruits-per-spawner and as such meant that the application of the SSB-index HCRs resulted in a more marked risk of future stock collapse.

Of the two considered total mortality HCRs (one using age-averaged Z from raw survey data, the other using the YCC method to derive the yearly estimates of Z) only the YCC-based HCR was considered, given issues with the performance of the "raw" Z-based HCR when using the PID control scheme. The YCC estimated Z-based HCR performed well in terms of rebuilding the stock and not surprisingly the imposition of the initial moratorium accelerated the time at which rebuilding (expressed as the time taken to attain a probability of being above B_{PA} of 95%) could be expected to be achieved. This rebuilding performance was at the sacrifice of TAC levels: over almost all cases the average TAC levels were significantly lower than those seen when using the SSB-based HCR but, as was observed with the herring Z-based rules, the inter-annual variation in TAC was lower than that seen using the SSB-based HCRs.

One interesting issue explored in the cod case study was the idea of using ideas from control theory to develop more complicated and dynamic HCRs, even when using simple survey-derived indices such as those considered here. It was observed that, for the SSB-based HCRs, the fine-tuned (incorporating all three tuning parameters K_P, K_I and K_D) runs tended to increase the levels of TAC set but at the same time resulted in higher chances of stock collapse. The integral (back-average of the control variable) and derivative (rate of change of the control variable) parts of the scheme seemed to act so as to increase TAC. When reducing the reactivity of the HCR (K_P set to 0.4 and not 1 as in the herring and anchovy cases) and with the integral and derivative terms set to zero afforded the stock the shortest rebuilding time. For the SSB-based HCR with the Ricker model the fine-tuned case performed poorly given the potential over-compensation dynamics of the Ricker model as the SSB approaches the point at which recruitment then begins to decrease. This is perhaps not surprising given that the PID scheme assumes the linearity of the system being controlled and this type of Ricker dynamic is strongly non-linear in origin. For the fine-tuning of the total mortality case this performed very poorly with the Z estimates derived from the raw survey data, driven by the fact that the PID scheme also assumes a smooth system and the observation error driven noise in the Z estimates caused the fine-tuned scheme to perform very poorly, hence its exclusion. The YCC method did not suffer from this problem and it is because it is essentially a biological smoother, so it removed some of this stochastic variation in the Z signal in the survey data, affording the fine tuned PID scheme much improved performance.

The idea of smoothing the data, for whatever reason, leads to the potential utility of survey-based assessment methods within the framework of survey-based management. There are a number of methods that can use survey data to estimate trends in key population variables such as total mortality, SSB and recruitment (Needle 2005; Trenkel 2008; Bogaards et al. 2009). In this paper only the YCC method (used for estimating total mortality) was used largely on the basis of (a) simplicity in terms of integration within an MSE framework given its linear nature, and (b) availability as at least one of the more complex methods is being integrated into the FLR panoply of packages but was not at a point where its integration into the MSE framework was feasible. More complex models tend to be non-linear and require optimisation methods to estimate parameters and often human judgement to decide on data weighting, model choice and so on. This makes the automated running of such models in an MSE difficult in some cases but these may not be insurmountable problems. It is clear that having such assessment methods available in the survey-based management framework detailed in this paper would make the work more relevant for making actual management recommendations (these methods are used in the assessment process for many stocks), and would allow one to explore further the potential utility of survey-based management.

7 Conclusion

In this paper a variety of management procedures constructed using only survey-based data have been evaluated, using a selection of case studies (North Sea herring and cod, Bay of Biscay anchovy) chosen so as to provide a good coverage of potential life-histories, stock status dynamics and survey types. The aim was to both look at how certain management procedures might perform for the given case studies and also to see what can be inferred from these particular examples with respect to survey-based management of exploited fish stocks in general. A generic TAC adaption algorithm with a theoretical basis in classical control theory was defined which was then parameterised in a variety of different ways relevant to the particular case studies. Details of how to construct both the biological, observation and management sections of the overall operating model were given, with particular relevance to fisheries assessed using VPA-type algorithms and with age-structured surveys.

The scenarios considered in the three case studies showed that conservative HCRs are required to control a system with high variability regardless of whether that comes from external factors (e.g. high misreporting or imprecise survey data) or from the stock itself (e.g. yield is sustained by a couple of year classes, stock-recruit uncertainty etc.). Also the "coverage" of the surveys with respect to the life-cycle of the population was found to be important. For North Sea herring the total mortality HCR clearly utilised the information from the survey relating to the juvenile and adult portions of the stock; for anchovy it was observed that the potential inclusion of a recruitment survey (effectively closing the life-cycle loop with the SSB survey) made the associated management procedure more robust than all of those that lacked such recruitment information. Other important issues are the levels and structure of the observation errors; correlation among indicators thought to be independent; HCRs relying on models that in turn rely on assumptions that are not supportable; unaccounted-for lags between the application of the HCRs, management action and response by the stocks.

The simulations also clearly demonstrate the need for management procedures designed and tested on a case-specific basis which is neither a new nor surprising result, but does add further support to the utility of the management strategy evaluation process, with respect to designing and defining management regimes. Also, there is the potential for such work to inform on survey design. There were clear cases where survey precision was an important factor in determining performance (from both stock preservation and fishery perspectives) and these types of analyses could be the case for support for more survey effort/coverage. Also, the anchovy simulations showed the potential benefit of a recruitment survey, even though at present no such survey is undertaken. For North Sea herring the results are far more robust to our ideas about the stock-recruit dynamics than the North Sea cod case; for the anchovy case the importance of individual year-class strength was found to be important, whereas for the others the stochastic variation in recruitment was less important than the manner in which we define mean recruitment (i.e. stock-recruit curve), given the spawning stock. For North Sea herring we have a reasonable understanding of the error and also the error structure of the key surveys, but given the differing performance of all the rules (irrespective of the stock) for different levels of observation error, it is also clear that some rules may well be more robust (or acceptable to stakeholders) to higher levels of survey imprecision. Having multiple performance criteria (biological and economic for example) allows one to assess the potential trade-offs of one particular harvest strategy compared to another, depending on what weight is given to the various performance measures.

The results of the analyses show that survey-based management, in conjunction with the use of management strategy evaluation in the management design phase, could provide a workable alternative to catch-based methods for fisheries management. In terms of using the MSE process it is clear that using a survey-based approach does not remove the problems of a lack of understanding of key processes, in particular the spawner-recruitment dynamics, but at the very least the removal of the dependence of having to know natural mortality must be considered a key advantage. The case studies all showed a reasonable level of robustness to low levels of misreporting (specified as under-reporting) but that consistently high levels cannot be easily dealt with. While this is also true for most fisheries data-dependent approaches one key advantage of the survey-based approach is that there are no associated biases entrained in the data available for estimating stock status. The fishery models used in this analysis are quite simple and do not allow evaluation of the performance of the HCRs in terms of the fishery and economical and social issues. Although such aspects could be equally important, their consideration in the calculations was beyond the scope of this analysis. The use of assessment methods that use only survey data were very briefly explored using a simple model but as the various models available are further developed, used for assessment and hopefully incorporated into the software framework used in this work their inclusion in the survey-based MSE process is strongly encouraged.

Acknowledgements. This research was funded by the European Commission project FISBOAT (Fishery Independent Survey Based Operational Assessment Tools). The authors are grateful to John Cotter from CEFAS (UK) for his valuable observations and comments. John Simmonds from FRS Aberdeen (UK) is thanked for his help and comments on the North Sea herring case study. They also would like to thank Charlotte Deerenberg from IMARES (Netherlands) for her work in the initial steps of the FLR simulation evaluation work. Finally, Andres Uriarte from AZTI-Tecnalia (Spain) is thanked for his internal discussions on the Bay of Biscay anchovy case study. Thank you also to Coby Needle and Peter Shelton for their comments and suggestions which greatly improved the paper.

APPENDIX

Initial equation:

$$\begin{cases} TAC_{y+1} = \min\left\{\max\left\{\exp\{u_y^1\}, \exp\{u_y^2\}\right\}, \exp\{u_y^3\}\right\} TAC_y \\ u_y = K_P\, e_y + K_I \sum_{z=y-\delta}^{y} e_z + K_D\left(e_y - e_{y-1}\right) \end{cases} \tag{3}$$

- u_y: control signal in year y.
- e_y: divergence of an index relative to a reference point.
- K_P, K_I and K_D: three parameters.
- δ : history that is considered in calculating the control signal.

HCR$_0$:

- Serves as a reference scenario
- Corresponds to the case in which no fishing is allowed
- Conforms to $K_P = K_I = K_D = 0$

$$TAC_{y+1} = 0 \tag{4}$$

HCR$_1$:

- Consists on keeping the *TAC* at a constant level defined by the latest historical catch level
- $K_I = K_D = 0$, $K_P = 1$ and $e_y = \log(1)$

$$TAC_{y+1} = TAC_y \tag{5}$$

Pure P-controllers HCRs (North Sea herring and Bay of Biscay anchovy):

- $K_I = K_D = 0$

$$TAC_{y+1} = \min\left\{\max\left\{\exp\{u_y^1\}, \exp\{u_y^2\}\right\}, \exp\{u_y^3\}\right\} TAC_y \tag{6}$$

If no min, max functions are used:

- $u_y^1 = u_y^2 = u_y^3$

$$TAC_{y+1} = \exp\left\{K_p e_y\right\} TAC_y \tag{7}$$

By adjusting these factors, we can recreate HCRs implemented for herring and anchovy:

- $K_I = K_D = 0, \quad K_P = 1$
- $e_y = \log\left(I_y\right) - \log\left(I_{y-1}\right) = \log\left(I_y / I_{y-1}\right)$

$$TAC_{y+1} = TAC_y\left(\frac{I_y}{I_{y-1}}\right) \tag{8}$$

Examples of I_y based on *SSB*

- $I_y = S\hat{S}B_y$

$$TAC_{y+1} = TAC_y\left(\frac{S\hat{S}B_y}{S\hat{S}B_{y-1}}\right) \tag{9}$$

- $S\hat{S}B_y$: *SSB* observations from the surveys

Use of min and max functions:

- u_y^1, u_y^2 and u_y^3 have different values
- For example: $u_y^1 = \log\left(\frac{S\hat{S}B_y}{S\hat{S}B_{y-1}}\right)$, $u_y^2 = 0.8$, $u_y^3 = 1.2$ variation constrained by an upper and lower limit for the inter-annual variation in the *TAC* (20% in the example)

$$TAC_{y+1} = \min\left\{\max\left\{\frac{S\hat{S}B_y}{S\hat{S}B_{y-1}}, 0.8\right\}, 1.2\right\} TAC_y. \tag{10}$$

Examples of I_y based on Z (total mortality rate)

$$TAC_{y+1} = TAC_y * \min\left(\frac{Z^{PA}_{[0.1]}}{Z_{y-1,[0,1]}}, \frac{Z^{PA}_{[2.6]}}{Z_{y-1,[2,6]}}\right). \tag{11}$$

In the case of North Sea herring, the index is estimated for 2 age groups of the stock and the minimum of the two ratios is used in the calculation of the *TAC*. Two age groups (0 to 1 and 2 to 6) are distinguished in the calculation of the *TAC* so we are using an age-structured survey in this case. The *TAC* of following year is inversely proportional to the observed changes in total mortality estimated along cohorts from the survey. *TAC* is increased in proportion to the smallest relative increase in Z_{PA}/Z. Values of the total mortality rate at age at the precautionary level (Z_{PA}) for each age group derive from other parameter values defined by ICES. The total mortality rate, Z, is the sum of the natural mortality M and the fishing mortality F. For the group 0-1 year old, F_{PA}, defined by ICES, is equal to 0.12 (ICES 2007b) and M is estimated as 1 by ICES North Sea Multispecies Virtual Population Analysis, MSVPA (Pope 1991). The combination of both of them gives a value of 1.12 for $Z^{PA}_{[0,1]}$. For ages 2 to 6, F_{PA} has been fixed at 0.25 by ICES experts and the natural mortality rate is equal to 0.3, 0.2 and 0.1 for ages 2, 3 and 4+ respectively. The combination of the mean of the natural mortality rate at age and F_{PA} gives a value of 0.4 for $Z^{\mathrm{PA}}_{[2,6]}$. For each year, both ratios included in HCR are calculated and the minimum value is used to calculate the *TAC* for the following year.

Combination of indices:

- $I_y = S\hat{S}B_y\hat{R}_y$

$$TAC_{y+1} = TAC_y \frac{S\hat{S}B_y}{S\hat{S}B_{y-1}} \frac{\hat{R}_y}{\hat{R}_{y-1}} \tag{12}$$

- $S\hat{S}B_y$: SSB observations from the surveys
- $\hat{R}_y$: recruitment from the surveys

PID-controllers HCRs (North Sea cod)

- Various K_P, K_I and K_D values
- $u_y^1 = u_y^2 = u_y^3$
- $\delta = 5$.

References

Anonymous, 2004, Inquiry into the future of the Scottish fishing industry. The Royal Society of Edinburgh. Edinburgh, pp. 1–108.

Apostolaki P., Hillary R., 2009, Harvest control rules in the context of fishery-independent management of fish stocks. Aquat. Living Resour. 22, 217–224.

Beare D.J., Needle C.L., Burns F., Reid D.G., 2005, Using survey data independently from commercial data in stock assessment: an example using haddock in ICES Division VIa. ICES J. Mar. Sci. 62, 996–1005.

Bogaards J.A., Kraak S.B.M., Rijnsdorp A.D., 2009, Bayesian survey-based assessment of North Sea plaice (*Pleuronectes platessa*): extracting integrated signals from multiple surveys. ICES J. Mar. Sci. 66, 665–679.

Butterworth D.S., Bergh M.O., 1993, The development of a management procedure for the South African anchovy resource. In: Smith S.J., Hunt J.J., Rivard D. (Eds.), Risk evaluation and biological reference points for fisheries management, Can. Spec. Publ. Fish. Aquat. Sci., pp. 83–99.

Butterworth D.S., Cochrane K.L., De Oliveira J.A.A., 1997, Management procedures: a better way to manage fisheries? The South African experience. In: Pikitch E.K., Huppert D.D., Sissenwine M.P. (Eds.), Global trends: fisheries management. Bethesda, Maryland, Am. Fish. Soc. Symp. 20, 83–90.

Butterworth D.S., Punt A.E., 1999, Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56, 985–998.

Campbell R.A., Dowling N.A., 2005, Evaluating harvest strategies for a rapidly expanding fishery: the Australian broadbill swordfish fishery. In: Kruse G.H., Gallucci V.F., Hay D.E., Perry R.I., Peterman R.M., Shirley T.C., Spencer P.D. (Eds.), Fisheries assessment and management in data-limited situations. University of Alaska, Fairbanks, Alaska Sea Grant College Program Report 05–02, pp. 509–532.

Chen J.H., Thompson M.E., Wu C.B., 2004, Estimation of fish abundance indices based on scientific research trawl surveys. Biometrics 60, 116–123.

Cochrane K.L., Butterworth D.S., de Oliveira J.A.A., Roel B.A., 1998, Management procedures in a fishery based on highly variable stocks and with conflicting objectives: experiences in the South African pelagic fishery. Rev. Fish Biol. Fish. 8, 177–214.

Cook R.M., 1997, Stock trends in six North Sea stocks as revealed by an analysis of research vessel surveys. ICES J. Mar. Sci. 54, 924–933.

Cotter A.J.R., Burt L., Paxton C.G.M., Fernandez C., Buckland S.T., Pax J.X., 2004, Are stock assessment methods too complicated? Fish Fish. 5, 235–255.

Cotter A.J.R., Mesnil B., Piet G.J., 2007, Estimating stock parameters from trawl cpue-at-age series using year-class curves. ICES J. Mar. Sci. 64, 234–247.

Cotter A., Mesnil B., Witthames P., Parker-Humphreys M., 2009, Notes on nine biological indicators estimable from trawl surveys with an illustrative assessment for North Sea cod. Aquat. Living Resour. 22, 135–153.

De Oliveira J.A.A., Butterworth D.S., 2004, Developing and refining a joint management procedure for the multispecies South African pelagic fishery. ICES J. Mar. Sci. 61, 1432–1442.

Dichmont C.M., Deng A.R., Venables W.N., Punt A.E., Haddon M., Tattersall. K., 2005, A new approach to assessment in the NPF: spatial models in a management strategy environment that includes uncertainty. CSIRO Div. Marine Research, Hobart.

FISBOAT, 2003, Description of work Sixth framework programme priority [8.1] [Policy-oriented research]. Specific targeted research or innovation project. Annex I.

FAO, Food and Agriculture Organisation of the United Nations, 1996, Precautionary approach to fisheries. Part 2: Guidelines on the precautionary approach to capture fisheries and species introductions. In: FAO (Ed.) FAO Fish. Tech. Pap. 350, pp. 1–62.

Geromont H.F., De Oliveira J.A.A., Johnston S.J., Cunningham C.L., 1999, Development and application of management procedures for fisheries in southern Africa. ICES J. Mar. Sci. 56, 952–966.

Hammond P.S., Donovan G.P., The RMP: managing whales in an uncertain world. J. Cetacean Res. Manage. Spec. Issue 3, in press.

Hillary R., 2009, An introduction to FLR fisheries simulation tools. Aquat. Living Resour. 22, 225–232.

ICES, 1991, Herring assessment working group for the area south of 62^{o}N. ICES CM 1991/Assess.15.

ICES, 2001, Study group on evaluation of current assessment procedures for North Sea herring. ICES CM 2001/ACFM:22.

ICES 2006a, Herring assessment working group for the area south of 62^{o}N. ICES CM 2006/ACFM:20.

ICES, 2006b, Working group on the assessment of mackerel, horse mackerel, sardine and anchovy. ICES CM 2006/ACFM:08.

ICES, 2006c, Working group on the assessment of demersal stocks in the North Sea and Skagerrak. ICES CM 2006/ACFM:09.

ICES, 2007a, ICES Advisory Committee on Fishery Management, Advisory Committee on the Marine Environment and Advisory Committee on Ecosystems. ICES advice, 6.

ICES, 2007b, Herring assessment working group for the area south of 62°N. ICES CM 2007/ACFM:11.

ICES, 2008, Herring assessment working group for the area south of 62°N. ICES CM 2008/ACOM:02.

IWC, International Whaling Commission, 1999, The revised management procedure (RMP) for baleen whales. Annex N to the Report of Scientific Committee. J. Cetacean Res. Manage. 1 (Suppl.), 251–258.

Johnston S.J., Butterworth D.S., 2005, The evolution of operational management procedures for the South African west coast rock lobster fishery. N.Z. J. Mar. Freshw. Res. 39, 687–702.

Kell L.T., Mosqueira I., Grosjean P., Fromentin J.M., Garcia D., Hillary R., Jardim E., Mardle S., Pastoors M.A., Poos J.J., Scott F., Scott R.D., 2007, FLR: an open-source framework for the evaluation and development of management strategies. ICES J. Mar. Sci. 64, 640–646.

Kirkwood G.P., 1997, The revised management procedure of the International Whaling Commission. In: Pikitch E.K., Huppert D.D., Sissenwine M.P. (Eds.), Global trends: fisheries management. Bethesda, Maryland, Am. Fish. Soc. Symp. 20, 91–99.

Lewy P., Nielsen A., 2003, Modelling stochastic fish stock dynamics using Markov chain Monte Carlo. ICES J. Mar. Sci. 60, 743–752.

McAllister M.K., Kirchner C.H., 2001, Development of Bayesian stock assessment methods for Namibian orange roughy *Hoplostethus atlanticus*. S. Afr. J. Mar. Sci.- Suid-Afr. Tydsk. Seewetens. 23, 241–264.

Needle C.L., 2005, SURBA 3.0: Technical Manual (first draft). In: FRS Marine Laboratory Aberdeen (Ed.), pp. 10.

Nielsen A., Lewy P., 2002, Comparison of the frequentist properties of Bayes and the maximum likelihood estimators in an age-structured fish stock assessment model. Can. J. Fish. Aquat. Sci. 59, 136–143.

O'Dwyer A., 2003, Handbook of PI and PID controller tuning rules. Imperial College Press.

Patterson K., Melvin G.D., 1996, Integrated catch at age analysis. Version 1.2. Scottish Fish. Res. Rep. 58.

Pope J.G., 1991, The ICES multispecies assessment group: evolution, insights and future problems. ICES Mar. Sci. Symp. 23–33.

Punt A.E., Donavan G.P., 2007, Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. ICES J. Mar. Sci. 64, 603–612.

Punt A.E., Pribac F., Taylor, Bruce L., Walker, Terence I., 2005, Harvest strategy evaluation for school and gummy shark. J. Northw. Atl. Fish. Sci. 35, 387–406.

Punt A.E., Smith A.D.M., 1999, Harvest strategy evaluation for the eastern stock of gemfish (*Rexea solandri*). ICES J. Mar. Sci. 56, 860–875.

Punt A.E., Smith A.D.M., Cui G.R., 2001, Review of progress in the introduction of management strategy evaluation (MSE) approaches in Australia's South East fishery. Mar. Freshw. Res. 52, 719–726.

Rademeyer R.A., Plagányi E.E., Butterworth, D.S., 2007, Tips and tricks in designing management procedures. ICES J. Mar. Sci. 64, 618–625.

Smith A.D.M., 1993, Risk assessment or management strategy evaluation: what do managers need and want? ICES CM 1993/D:18.

Trenkel V.M., 2008, A two-stage biomass random effects model for stock assessment without catches: what can be estimated using only biomass survey indices? Can. J. Fish. Aquat. Sci. 65, 1024–1035.

Tuck G.N., Sainsbury K.J., Williams R., 2003, Abundance estimation and TAC setting for Patagonian toothfish (*Dissostichus eleginoides*) at Macquarie Island: a synopsis. Sub-Antarctic fisheries assessment group, SAFAG17, Agenda 5.

Woillez M., Rivoirard J., Petitgas P., 2009, Notes on survey-based spatial indicators for monitoring fish populations. Aquat. Living Resour. 22, 155–164.