

Development of a stock–recruit model for simulating stock dynamics for uncertain situations: the example of Northeast Atlantic mackerel (*Scomber scombrus*)

E. John Simmonds^{1*}, Andrew Campbell², Dankert Skagen³, Beatriz A. Roel⁴, and Ciaran Kelly²

¹Joint Research Centre IPSC, Maritime Affairs Unit, European Commission, TP 051, 21027 Ispra, Italy

²Marine Institute, Rinville, Oranmore, Co. Galway, Ireland

³Institute for Marine Research, PO Box 1870, 5817 Nordnes, Bergen, Norway

⁴Cefas, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK

*Corresponding Author: tel: +39 0332 785311; fax: +39 033278 9156; e-mail: john.simmonds@jrc.ec.europa.eu.

Simmonds, E. J., Campbell, A., Skagen, D., Roel, B. A., and Kelly, C. 2011. Development of a stock–recruit model for simulating stock dynamics for uncertain situations: the example of Northeast Atlantic mackerel (*Scomber scombrus*). – ICES Journal of Marine Science, 68: 848–859.

Received 23 April 2010; accepted 25 January 2011

The assumption of a relationship between recruitment and a spawning stock is the cornerstone of the precautionary approach and may constrain the use of a maximum sustainable yield (MSY) target for fisheries management, because the failure to include such a relationship suggests that providing a measure of stock protection is unnecessary. The implications of fitting different functional forms and stochastic distributions to stock-and-recruit data are investigated. The importance of these considerations is shown by taking a practical example from management: the management plan for Northeast Atlantic mackerel (*Scomber scombrus*), a fish stock with an average annual catch of 600 000 t. The historical range of spawning-stock biomass is narrow, and historical data from a stock assessment explain only a small proportion of the recruitment variability. We investigate how best to reflect the uncertainty in the stock–recruit relationship. Selecting a single model based on simple statistical criteria can have major consequences for advice and is problematic. Selecting a distribution of models with derived probabilities gives a more complete perception of uncertainty in dynamics. Differences in functional form, distribution of deviations, and variability of coefficients are allowed. The approach appropriately incorporates uncertainty in the stock–recruit relationship for F_{MSY} estimation.

Keywords: fisheries management, model probability, stock–recruit relationship.

Introduction

Developing and implementing management plans are requisites for ensuring that fish stocks, their ecosystems, and their environments are maintained in a productive state (FAO, 1997). Following the 2002 World Summit on Sustainable Development in Johannesburg (FAO, 2003), a number of nations agreed to implement fisheries management policy based on maximum sustainable yield (MSY). These requirements have led to the development or the adaptation of a number of scientific tools to support this policy. To provide advice on fish stock exploitation in this context, there is a requirement for management simulations to include the elements of stock dynamics (Stokes *et al.*, 1999; Cochrane, 2002; Butterworth, 2007; ICES, 2008a). In particular, this leads to the requirement to simulate a plausible caricature of the recruitment to allow simulation of one or multiple stocks. A critical aspect of this is the relationship between the recruitment (r) and the spawning stock (S) that produces it. Here, we investigate the characterization of the stock–recruit (S – r) relationship in terms of both functional form and uncertainty, and its influence on the estimation of exploitation targets such as MSY. We do not attempt to define S – r over all biomasses, e.g. the detail at the origin, but just those features required for MSY estimation. Nor do we report a full management strategy evaluation (MSE),

which should include *inter alia* additional sources of uncertainty, e.g. growth, maturation, spawning potential, natural mortality, fishery selection, and issues of observation and implementation uncertainty.

A universal truth is “that if you fish hard enough on any stock, you will reduce recruitment” (Hilborn and Walters, 1992). Myers and Barrowman (1996) state that there is a clear dependence of recruitment in some way or other on stock size. The precautionary approach (PA) to fisheries (Garcia, 1996; ICES, 2001) dealt with this issue by defining an objective: the avoidance of reduced recruitment. MSY can be considered complementary to the PA such that a necessary condition for MSY is compatibility with precautionary limits under the PA system (ICES, 2001). In the PA framework, managers were expected to select targets freely, which led often to short-term considerations and loss of catch. MSY adds to the PA concept of problem avoidance by providing a specific target and that of maximizing the sustainable (long-term) yield. For some stocks, yield-per-recruit, growth, and mortality considerations dominate the estimate of MSY, but for others, such as Northeast Atlantic (NEA) mackerel (*Scomber scombrus*), the precautionary limits attributable to declining recruitment define MSY. The assumption of some type of relationship between r and S becomes the cornerstone of both approaches,

because the failure to include such a relationship could suggest that providing a measure of stock protection was unnecessary. Such a view would be incompatible with the precautionary principle.

Although we consider that S – r relationships are crucial, they vary over time for some stocks (Brunel and Boucher, 2007; Simmonds and Keltz, 2007; Köster *et al.*, 2008; Drinkwater, 2009). In these examples, it is clear that recruitment depends on environmental influences, which dominate inter-year variability, but to assume no relationship between S and recruitment, because it is difficult to establish, could lead to quite unrealistic expectations when considering the tolerance to exploitation. In some cases, the clear dependence on a changing environment leads to the view that there are multiple states for a population, and both these and the potential for transition is acknowledged (King and McFarlane, 2006; Simmonds and Keltz, 2007; Yatsu *et al.*, 2008). On occasion, complexity through species linkage and mixed fishery exploitation may be added (Rätz *et al.*, 2007; Mackinson *et al.*, 2009). Nevertheless, in all these cases, there is still a need to provide a plausible functional relationship between stock and recruitment that can be considered sufficiently stable over time, although that period may be difficult to predict (Alheit and Niquen, 2004; Jacobson *et al.*, 2005; Bakun and Weeks, 2008).

An MSE is an evaluation of management under a range of plausible biological conditions, but also taking into account measurement uncertainty in the likely levels of compliance with regulation. Stock–recruit relationships are one of the most important aspects of such a simulation because they dominate the supply of fish available for exploitation. Although variability in recruitment is acknowledged in almost all cases, it is common for a single stock–recruit relationship to be selected explicitly for a particular stock study (ICES, 2005; Kell *et al.*, 2009), adding a stochastically varying component to describe inter-year variability. The inclusion of a stochastic component is a good solution to an inability to predict the necessary detailed environmental changes or how they will specifically influence recruitment, although some temporal autocorrelation between years may also be required. The functional form of the S – r model may be chosen for biological reasons, or selected on statistical fit criteria using, for example, the Akaike Information Criterion (AIC) to select among models (Shimoyama *et al.*, 2007). Sometimes, it is clear that a fixed S – r relationship is inadequate because the stock dynamics seem to have changed either with time (King and McFarlane, 2006; Simmonds and Keltz, 2007) or explicitly with climate (Stige *et al.*, 2006; Fogarty *et al.*, 2008). In some of these cases, several management options have been developed to account for different S – r regimes, although they are sometimes presented without any specific indication as to which scenario should be selected and how future uncertainty might be dealt with (Nash *et al.*, 2009). In others (King and McFarlane, 2006), detecting the regime shift is specifically addressed, although within a regime, a single functional form of the S – r model is chosen. The consideration of uncertainty in the model form was introduced by Michielsens and McAllister (2004) when they presented an approach using multiple models. In their case, model selection resulted in a very strong preference for one model form over another, and the results were insensitive to the complexity, suggesting that the approach might be unnecessary. Brodziak and Legault (2005) examined 12 models and made comparisons between informative and uninformative priors to evaluate

rebuilding targets for several groundfish stocks. They used Bayesian model averaging to evaluate the results. That approach recognizes the need to consider the different biological hypotheses implied by different relationships. They evaluated the precision of estimates for recovery under the different assumptions by averaging the models rather than allowing a range of potential models to exist simultaneously and considering the risks of depletion, as we do here. Rademeyer *et al.* (2007) postulated that, given a number of hypotheses related to population current status and productivity, the performance of a harvest control rule (HCR) should be integrated over all possible hypotheses/scenarios considered, with relative weights assigned to the output statistics, to account formally for the relative likelihoods of the hypotheses postulated. The approach we present here for hypotheses of the S – r formulation conforms directly to that philosophy. Our approach also fits with the suggestions of McAllister and Kirchner (2002) that “a set of alternative models [be selected] each of which specify a unique functional form of some biological or fishery process or density function where it is not convenient or possible to formulate a single generalized model from them”. In that paper, the authors examined some parametric modelling uncertainty in the stock assessment to provide advice on management.

For NEA mackerel, the modelling problem is made more difficult by the absence of informative data on the slope of the S – r function near the origin, both the biomass at which the decline in recruitment begins and also the Allee effect which may occur for pelagics (Quinn and Deriso, 1999). We consider that the former is important, but that the latter would be expected well outside the range of biomasses encountered under MSY exploitation and should not influence the analysis. One approach favoured by some (e.g. Myers and Barrowman, 1996; Brodziak and Legault, 2005) is to use a meta-analysis from other comparable populations to infer stock dynamics. However, the dynamics of a highly migratory pelagic stock are bound to be greatly influenced by the biotic and abiotic environment over which they range (if they were not, they would be ubiquitous). We consider that these aspects are going to have a greater influence over stock productivity and resilience than that controlled by their being the same species. This viewpoint has some support; there is only one other population of *S. scombrus*, which is found on the western side of the Atlantic. Moustahfid *et al.* (2009) present stock-and-recruit data for NWA mackerel under two differing modelling assumptions, and in both cases, a totally different type of recruitment dynamic from that of NEA mackerel is evident, with very much greater variability. The reported performance of the NWA mackerel assessment (Anon., 2010) is poor with major retrospective errors, which also raises questions about its usefulness. The report shows severe retrospective patterns in the estimates of S , F (fishing mortality), and recruitment, and for the early period (going back to the 1960s), landings data were questionable and the biological samples taken relatively poor. Nevertheless, the differences between NWA and NEA mackerel are stark. Whereas NEA mackerel has relatively consistent recruitment with a range of less than one order of magnitude, Moustahfid *et al.* (2009) show that, for the same period, NWA mackerel has a spread of recruitment of nearly two orders of magnitude. NEA mackerel spawning is spread smoothly over a wide range of latitudes and over a long season. Spawning in NWA mackerel is much less spatially contiguous and temporally more variable and discrete (Sette, 1950). This difference in the dynamics could be

due in part to the markedly different hydrography in the two areas. Recruitment dynamics are clearly different for the two stocks, so making inference from one to the other inappropriate.

This metadata approach may be applicable where the stock in question is one among several in the midrange of similar ecosystems. Here, we present an approach that relies explicitly on the data available, which we favour in this case as a way to account for our level of knowledge. However, for other situations where the biological arguments for inference among stocks are robust, the modelling approach described below could be applied equally well with informative priors. However, in such cases, the reliability of the results would be influenced by the increase in certainty resulting from the inclusion of informative priors, making the choice of priors critical to the analysis.

As discussed above and as seen in the many of the cases cited, the shape of the S - r functional form is basic to the stock dynamics implied in the simulations. There are two basic philosophical approaches to the problem of selection models to test. One may prefer to select or reject models based explicitly on knowledge of the dynamics and biological hypotheses of say the independence of recruitment at high biomass (hockey stick/Beverton–Holt), cannibalism (Ricker), or depensation at low biomass (Saila–Lorda/sigmoid Beverton–Holt). For NEA mackerel, all of these biological processes are potentially valid, and in cases such as cannibalism, there is some evidence (Olaso *et al.*, 2005). Throughout we have taken a statistical approach, allowing for a wide range of plausible biological hypotheses that can be formalized as S - r functions, taken from the review of Needle (2002).

In our example of NEA mackerel, we show the importance of S - r model uncertainty by taking a practical example from a specific management problem: the management plans for the stock. In 2007, ICES received a request for advice on long-term management of NEA mackerel (ICES, 2007), an important fish stock with an average annual catch of 600 000 t. The historical stock dynamics are available from a stock assessment performed annually (ICES, 2009). Based on those data, we investigate how best to reflect the uncertainty in the S - r relationship for the stock. Our results show that it is critical to use multiple functional relationships with distributions of coefficients, to account for uncertainty in S - r dynamics. It is also necessary to consider the form of the parametric distribution providing the stochastic variability. This distribution has often been assumed to be lognormal (ICES, 2005; Kell *et al.*, 2009), although other continuous distributions may provide more plausible stochastic variability for some stocks.

Methods

Stock and recruit data

The data used (Figure 1) are the time-series of S and r taken from the most recent stock assessment (ICES, 2009). S - r estimates from recent years (2007–2009) are excluded because of the lack of convergence of S for those years and the associated poor estimation of recent year classes in the assessment. Estimates of S in earlier years (1972–1979) are more uncertain (ICES, 2008b), because the data for separating the older ages in early catch information are not available (in the historical data, a plus group at age 4 in 1972 is incremented annually to age 12 in 1980). To account for the additional uncertainty in S in those years, extra variability is added in the modelling of these S data in the time-series (Supplementary material). An alternative solution to the

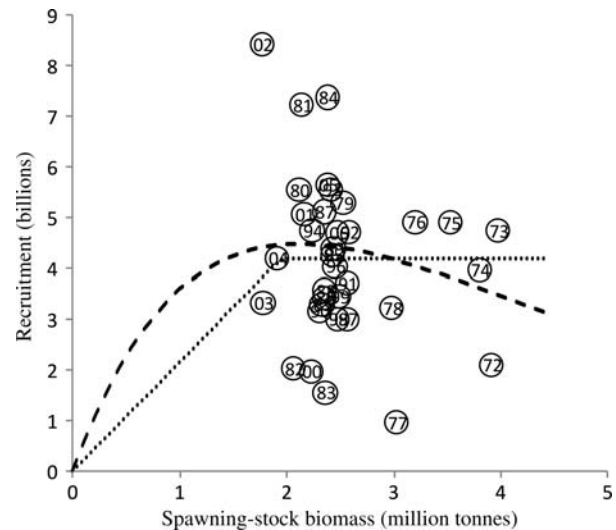


Figure 1. Point observations of spawning-stock biomass (S) and recruitment (r) from the ICES assessment (ICES, 2009), with numbers referring to year classes. Two fitted stock–recruit relationships with the lowest AICc are shown, Ricker (dashed) and hockey stick (dotted), both fitted with normal distribution of deviations.

problem would have been simply to exclude these data, but the recruit estimates for those years are as valid as the rest of the time-series, and the additional uncertainty in S is small relative to the information potentially available on stock dynamics. S in the earliest years (1972–1978 inclusive) is higher than that observed at any other time, so including the data in the analysis extends the information on stock dynamics.

Stock–recruit models

Needle (2002) provides a comprehensive review of stock models proposed since the S - r relationship was postulated mathematically in the 1950s. He presents seven functional relationships that have a continuous parametric form ideally suited to simulation (Table 1). He also presents other methods that are constrained to the observed dynamic range of the data, although these would not be suitable for a simulation of S that extends beyond the range of the observed values of S . The S - r model equations are fitted using the maximum log-likelihood method (Fisher, 1922) with normal, lognormal, or gamma distributions to characterize the stochastic component.

The likelihoods are calculated for each of the stochastic distributions as shown below. Normal distribution (N) with standard deviation σ proportional to the model mean (μ):

$$\Pi_{ji} = \frac{1}{\sigma_j \mu_{ji} \sqrt{2\pi}} \exp\left(-\frac{1}{2\mu_{ji}^2 \sigma_j^2} (r_i - \mu_{ji})^2\right). \quad (1)$$

Lognormal distribution (L) with standard deviation σ :

$$\Pi_{ji} = \frac{1}{r_i \sigma_j \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_j^2} (\ln(r_i) - \mu_{ji})^2\right). \quad (2)$$

Table 1. The functional form of the seven types of model used in the evaluations, with a , b , and g being fitted parameter values.

Model type	Reference	Formula
Models with asymptotic recruitment at high biomass		
BH (Beverton–Holt)	Beverton and Holt (1957)	$r = aS/(b + S)$
SB–H (sigmoidal Beverton–Holt)	Myers <i>et al.</i> (1995)	$r = aS^g/(b + S^g)$
HS (hockey stick)	Butterworth and Bergh (1993)	$r = aS/b : S < b = a$ $: S \geq b$
P (power)	Cushing (1973)	$R = aS^b : b \geq 0$
Models with peak in recruitment and decline with high biomass		
Rk (Ricker)	Ricker (1954)	$r = aS \exp(-bS)$
Sh (Shepherd)	Shepherd (1982)	$r = aS^g/[1 + (S/b)^g]$
SL (Saila–Lorda)	Iles (1994)	$r = aS^g \exp(-bS)$

Gamma distribution (G) with shape factor k :

$$\Pi_{ji} = r_i^{(k_j-1)} \exp(-r_i \theta_{ji}) \theta_{ji}^{k_j} / \Gamma(k_j) \quad \text{where } \theta_{ji} = k_j / \mu_{ji} \quad (3)$$

Here, μ_{ji} is the modelled mean recruitment given the S_i , the observed recruitment r_i , and the model parameters a , b , and g , using the equations for μ listed in Table 1.

The log-likelihood of model j , given the data S_i and r_i , is given by

$$\Pi_j = \sum_i \ln(\Pi_{ji}). \quad (4)$$

The fit of model j with k parameters based on n observations can be compared with other model fits using the AIC corrected for small sample size, AICc (Akaike, 1974), given by

$$\text{AICc}_j = 2k - 2\Pi_j + \frac{2k(k+1)}{n-k-1}. \quad (5)$$

Sensitivity to year ranges

To investigate the sensitivity of the analysis to the period, a limited number of models was fitted to subsets of the entire period (1972–2006). These periods were chosen as 1980–2006 (removing early data at high biomass) and 1980–2000 (removing early data and the most recent more-variable uncertain recruitment).

Distribution of coefficients

To establish the distributions of the model coefficients, while taking into account the correlation between parameters, the models were also fitted in a Bayesian Markov chain Monte Carlo (MCMC) framework (Gelman *et al.*, 2004) using WINBUGS software (Spiegelhalter *et al.*, 2003). The model parameters were loaded into the statistical package R using the R CODA package (Best *et al.*, 1997). The WINBUGS framework also provides a deviance information criterion metric, DIC (Spiegelhalter *et al.*, 2002; Gelman, 2003) for selecting models. DIC is a Bayesian equivalent to AIC, but incorporating the effect of correlation between variables through the model complexity parameter (pD). The DIC is used to compare the fit of the different functional models in WINBUGS. The priors were kept uninformative to give reliance

on the data, to reflect the uncertainty in the available information. Priors and initial values are listed in the Supplementary material.

Model selection and probability of different model types

The Bayes posterior probabilities of a model were calculated from the converged MCMC using the harmonic mean of the likelihoods of the l models taken from the chain (Kass and Raftery, 1995; Michielsens and McAllister, 2004):

$$\Pi_j = \frac{1}{\sum_l 1/\Pi_{j,l}}. \quad (6)$$

For each model, a set of 1000 coefficients was taken from the MCMC of 40 000 by thinning by 1/40.

In all, 21 models (seven functional forms and three stochastic distributions) are available. However, several of these can imply similar dynamic and management outcomes (see the “Results” section), and the method to obtain the probability of each is conditional on the number and types of models used in the analysis. If one evaluated ten models of one type and one of a different type, utilizing all the models would weight the results in favour of the larger group by a factor of 10, so the outcome would depend on the choice of model types tested. To avoid this methodological problem, an equal number of models of each type was selected in a two-stage process. During the first stage, models were selected based on their underlying dynamics, i.e. models having the same functional form through the data and the same probability distribution for the stochastic process. Where several models have the same functional form through the data, the model with the best fit based on DIC was selected.

The second stage of model selection assigned a relative probability to each model (given the others selected), which is calculated as the relative posterior probability of the model given the other models. For the six models selected, the probability of one (HS–N; for abbreviations, see Tables 1 and 2) is given by

$$P_{\text{HS–N}} = \frac{\Pi_{\text{HS–N}}}{\Pi_{\text{HS–N}} + \Pi_{\text{HS–G}} + \Pi_{\text{HS–L}} + \Pi_{\text{RN}} + \Pi_{\text{RG}} + \Pi_{\text{RL}}}. \quad (7)$$

The issue and influence of model selection is explored further in the discussion.

Population dynamics

The population dynamics implied by the different S – r relationships were simulated using a simple fisheries operating model. The software used was developed to evaluate HCRs for mackerel and was tailored to accommodate the recruitment scenarios derived here. The basic equations of the operating model can be expressed as a Leslie matrix (Leslie, 1945, 1948):

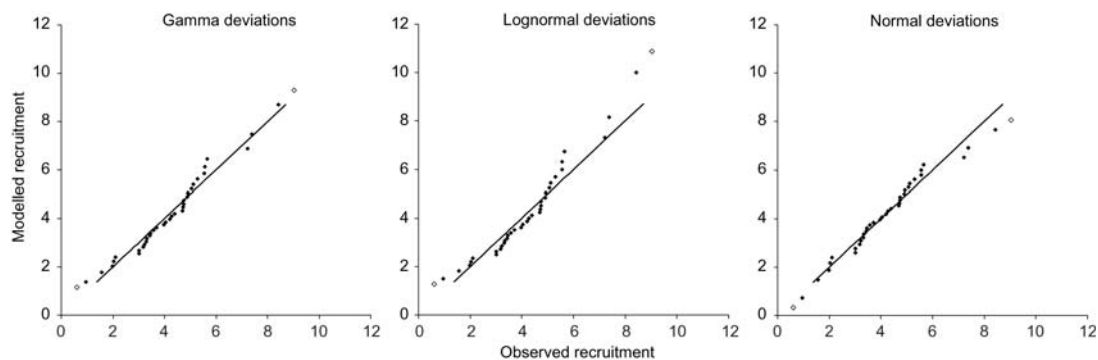
$$\begin{pmatrix} N_0 \\ N_1 \\ N_2 \\ \vdots \\ N_{11} \\ N_{12+} \end{pmatrix}_{t+1} = \begin{pmatrix} f(S)\varepsilon & \dots & \dots & \dots & \dots & \dots \\ \phi & \dots & \dots & \dots & \dots & \dots \\ \dots & \phi & \dots & \dots & \dots & \dots \\ \dots & \dots & \phi & \dots & \dots & \dots \\ \dots & \dots & \dots & \phi & \dots & \dots \\ \dots & \dots & \dots & \dots & \phi & \phi \end{pmatrix}_t \times \begin{pmatrix} 1 \\ N_0 \\ N_1 \\ \vdots \\ N_{10} \\ N_{11} \\ N_{12+} \end{pmatrix}_t, \quad (8)$$

where the survival (ϕ) is derived from a fixed age-based natural mortality (M) and fishing mortality ($F_{a,y}$) derived from a selection

Table 2. Estimates of fitted model parameters using the maximum log-likelihood method with L (lognormal), N (normal), and G (gamma) deviations around the model.

Deviation	Parameter	Model type (Table 1)						
		HS	Rk	P	B–H	S–L	Sh	SB–H
L (lognormal)	<i>a</i>	1.76	5.53	3.86	3.86	5.56	9.50	2 777
	<i>b</i>	3.86	0.50	0.00	0.00	0.14	0.93	718
	<i>g</i>	–	–	–	–	0.00	1.66	0.00
	σ	0.44	0.43	0.44	0.44	0.43	0.43	0.44
	AICc	74.66	74.43	74.66	74.66	76.65	76.64	77.22
	Rank	13	11	14	15	20	19	21
N (normal)	<i>a</i>	1.76	6.37	4.19	4.19	6.34	111.4	2 814
	<i>b</i>	4.19	0.52	0.00	0.00	0.16	0.16	669
	<i>g</i>	–	–	–	–	0.00	1.51	0.00
	σ	0.38	0.37	0.36	0.38	0.36	0.36	0.38
	AICc	72.49	71.86	72.57	72.50	73.59	73.31	75.05
	Rank	2	1	5	3	10	9	17
G (gamma)	<i>a</i>	1.76	6.04	4.19	4.19	6.01	62.77	20.05
	<i>b</i>	4.19	0.50	0.00	0.00	0.14	0.21	3.78
	<i>g</i>	–	–	–	–	0.00	1.46	0.00
	β	0.68	0.67	0.68	0.68	0.65	0.65	0.68
	AICc	72.77	72.53	72.78	72.78	74.70	74.59	75.34
	Rank	6	4	7	8	16	12	18

The estimated AICc is from maximum log-likelihood. The best fit under AICc criteria is the Rk (Ricker) model with N (normal) deviations.

**Figure 2.** Comparison between observed and fitted deviations from modelled *r* values, comparing gamma, lognormal, and normal deviations for the hockey stick model fit, showing the observed quantiles (points) and fitted quantiles (line) over the range of data. Estimated limits to *r* (diamonds) are based on a half increment further in each direction.

pattern taken from the final year of the ICES assessment (ICES, 2009) and a target *F*:

$$\phi_{a,y} = e^{-(M+F_{a,y})}, \quad (9)$$

with *S* derived from

$$S = \sum_{a=0}^{12+} N_a w_a m_a. \quad (10)$$

Here, w_a and m_a are the mean weight and mean proportion mature at age in the population taken from the ICES assessment (ICES, 2009). For these studies, no account is taken of measurement or implementation error that would normally be required for full management simulation.

Truncation of simulated recruitment values

The fitted models may simulate recruitment that can exceed the observed historical range considerably. Theoretical distributions

can yield a very small number of high deviation values that are correctly represented mathematically, but may be biologically unrealistic. This is particularly the case when the tails of the distributions fit poorly, as for example to the lognormal distribution illustrated in Figure 2. To select appropriate truncation limits, simulated recruitment was compared with observed recruitment using the range of *S*–*r* pairs from the assessment. Then, assuming linearly changing increments on the upper and the lower half of each distribution, the location of the end of the extent of the distribution can be estimated approximately by extending the lower and the upper ends of the distribution by one half increment (Figure 2). Although this defines the limits over the observed *S*, they still need to be allocated appropriately over all values of *S*. Three approaches were taken to determine the upper and the lower limits to the deviance for each distribution:

- limiting the proportion of simulated values to 1/70th of the total above and below the observed minimum and maximum recruitments;
- limiting σ to the same fixed limit in all models;

- (iii) limiting the σ on each model, to give an overall limit to simulated recruitment (for the range of S values observed).

When a number outside the determined range for the model was drawn in the simulation, it was rejected, and a new number drawn.

Results

Maximum likelihood fit

The model coefficients and AICc values for the seven potential S – r models and three types of distribution are listed in Table 2. The best fit using the AICc is a Ricker model with normal deviations (Rk–N), although the difference in AICc between this and the next best fitting model, the hockey stick with normal deviations (HS–N) is small. These two models are shown in Figure 1, along with the S – r dataset from the latest assessment. Fits to normal or gamma distributions generally have lower values of AICc than those fitted with lognormal distributions. Power (P), Beverton and Holt (B–H), and sigmoid Beverton and Holt (SB–H) models fit with flat recruitment throughout the observed range of S . The Shepherd (Sh) and Sella–Lorda (SL) four-parameter models have a peak in a similar place to the three parameter Ricker model, and in both these four-parameter cases, coefficients are poorly resolved. The hockey stick model fits to the mean over the observed range of biomass, with a straight line between the lowest observed biomass and the origin. In contrast, the Ricker model fits a falling slope through the observed biomass, and the slope towards the origin is controlled by the slope at the highest biomass, resulting in a much steeper slope at the origin.

The fit for the different distributions is similar for each functional form. Cumulative distributions and quantile–quantile (Q–Q) comparisons between data and model show the properties of different distributions for the example of the hockey stick form (Figure 2). The normal and the gamma distributions fit best with slight under- and overestimation of the skew, respectively. The log-normal distribution is more skewed than the gamma, fitting more poorly and with strong potential for more high recruitment than observed.

Sensitivity of the fitted models to the period of data

In these data, there was a trend in the residuals, with earlier years fitting with negative residuals to those models without an ability to yield reduced recruitment at high biomass. When such a signal is seen only once throughout an available time-series, it is not possible to determine whether it is a correlated environmental effect or biomass-dependent recruitment. To investigate the sensitivity of the models to the period of data, the hockey stick and Ricker models were fitted to different ranges of years. As both models have three parameters (a , b , σ), the differences between models can be considered as dependent only on the likelihood of the values of the residuals from the model fit, and not the number of parameters. The differences in the parameter values for the two ranges of years tested are small. Model values overlap completely in regions where the data are common to all periods. The only difference between the fits for the periods is in the estimated values of σ . Hence, the choice of period is informative only for variability, not for mean recruitment. The longest period with the greatest variability seems most appropriate to use in simulations. That shorter periods have less variability might infer autocorrelation,

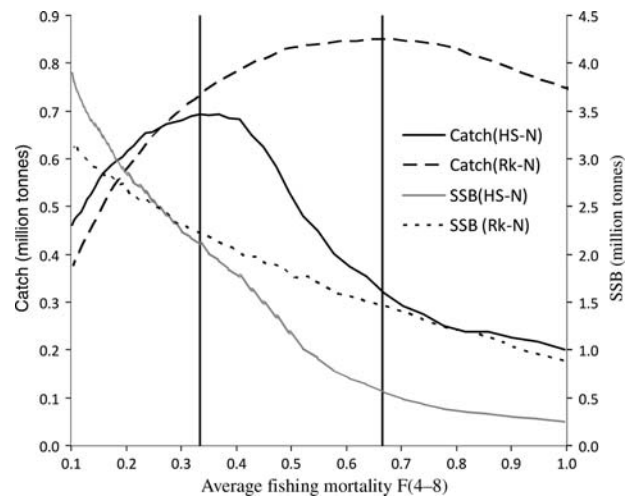


Figure 3. Population dynamics (mean S , yield vs. average F_{4-8}) for populations simulated with HS (hockey stick) and Rk (Ricker) models. F_{MSY} (black grid lines) are at $F_{4-8} = 0.66$ for Rk and $F_{4-8} = 0.33$ for the HS. The maximum yield for the Rk model is at a higher value of F than stock collapse with the HS ($F_{4-8} = 0.55$).

although there is little evidence that recruitment is autocorrelated because there is a non-significant negative term at lag 1.

Population dynamics

Using the AICc to select from the 21 models fitted, the population dynamics implied by the two top-ranked models (Rk–N and HS–N; see abbreviations in Tables 1 and 2) are contrasted in Figure 3. The dynamics implied and the resulting management advice differ markedly. The Rk model gives a value F_{MSY} of 0.7 and $F_{crash} > 1.2$. The HS gives $F_{MSY} = 0.3$ and $F_{crash} = 0.55$. This importantly different view of the stock dynamics derives mainly from the differing slope of the S – r relationship towards the origin (Figure 1).

Bayesian posterior probability

The SB–H, SL, and Sh models proved difficult to fit in WINBUGS. It was necessary effectively to reduce four-parameter models to three (two plus σ) by giving at least one parameter a very narrow prior distribution, effectively eliminating the independence of the third shape parameter. The DIC for each of these models is listed in Table 3. In that case, the best fit based on selection by DIC is Rk–N, with Rk–G and HS–N next. Again based on the functional form, the results fall into two main categories; those that are asymptotic to fixed recruitment at high S , and those that yield a peak in recruitment and a decline at high S .

Probability of the different models

Following the two-stage procedure described above, six models were selected from the potential 21, based on hockey stick and Ricker forms and the three distributions normal, gamma, and log-normal. The choice of hockey stick was based also on additional considerations. Whereas the hockey stick has the same asymptotic property as the Beverton–Holt and power forms, the behaviour of the Beverton–Holt and power forms near the origin has the potential to give infinite resilience to fishing, because they have unrealistic slope in the S – r function at the origin. By choosing the hockey stick constrained to the range of observed S , we maintained the independence of recruitment at high biomass,

Table 3. Model selection criteria for Bayesian models, model complexity (pD), and deviance information criteria DIC (with DIC analogous to AIC in the maximum likelihood method in Table 2).

S-r model	Distribution	pD	DIC	Rank
BH (Beverton–Holt)	N (normal)	2.068	135.6	5
	G (gamma)	2.010	136.1	8
	L (lognormal)	2.058	145.8	14
HS (hockey stick)	N (normal)	2.034	135.5	3
	G (gamma)	2.014	136.1	7
	L (lognormal)	2.052	139.9	11
P (power)	N (normal)	2.010	135.6	4
	G (gamma)	2.020	136.1	9
	L (lognormal)	2.037	139.9	12
Rk (Ricker)	N (normal)	2.376	134.8	1
	G (gamma)	0.867	135.2	2
	L (lognormal)	1.560	139.6	10
SB–H (sigmoidal Beverton–Holt)	N (normal)	2.115	135.7	6
	G (gamma)	2.093	155.7	15
	L (lognormal)	0.984	143.2	13

Table 4. Parameter estimates, the mean, and 0.025, 0.5, and 0.975 quantiles for the selected Bayesian models used in the recruitment simulation.

Model	Parameter	pdf	Mean	s.d.	2.5%	50.0%	97.5%
HS (hockey stick)	<i>a</i>	N	1.761	0.013	1.760	1.760	1.880
		G	1.761	0.019	1.760	1.760	1.940
		L	1.761	0.026	1.760	1.760	1.950
	<i>b</i>	N	4.195	0.278	3.653	4.189	4.754
		G	4.236	0.297	3.686	4.222	4.863
		L	3.895	0.304	3.330	3.881	4.533
	σ	N	0.393	0.058	0.300	0.386	0.527
		G	0.416	0.050	0.331	0.411	0.528
		L	0.455	0.058	0.358	0.449	0.586
Rk (Ricker)	<i>a</i>	N	0.628	0.126	0.385	0.627	0.877
		G	0.647	0.175	0.317	0.645	0.995
		L	0.401	0.285	0.002	0.488	0.843
	<i>b</i>	N	8.524	2.842	4.378	8.071	15.310
		G	9.470	4.453	3.726	8.567	20.540
		L	5.460	3.466	1.382	5.358	13.500
	σ	N	0.389	0.056	0.298	0.383	0.514
		G	0.420	0.052	0.331	0.415	0.538
		L	0.455	0.059	0.358	0.450	0.587

maintained the link to the data, and avoided unrealistic responses or the problem of finding a way to constrain the slope for the Beverton–Holt and power forms, but we miss the potential for steeper slopes at the origin. However, steeper slopes are included in the Ricker model set. The ranges of slopes and dynamics included in the analysis are dealt with in the “Discussion” section.

The parameter estimates of selected models are listed in Table 4. The use of a harmonic mean to evaluate the probability of model type is sensitive to the likelihoods of the least likely models because $1/\Pi$ increases rapidly as Π becomes small. The sensitivity of model probabilities to the Π of the least likely models is depicted in Figure 4. By removing the least likely values progressively, it is possible to see the influence of those in the extreme tails of the likelihood distributions. Once six models (0.6% of 1000) had been removed, the more-rapid changes in model probability ceased and the proportions were more stable, varying by $<0.5\%$ for each subsequent model. Although the changes over the next 2–3% of models are evident, these are considered to reflect the

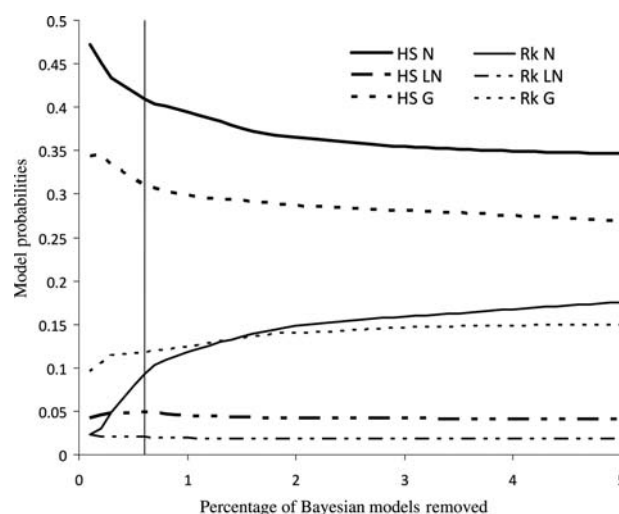


Figure 4. The proportion of models of each form relative to the percentage of least likely models removed from the set, showing how a very small proportion of models in the tail of the distribution can affect the relative probability of the different forms.

Table 5. Probability of different model types.

Model	Deviations	Probability
HS (hockey stick)	N (normal)	0.404
Rk (Ricker)		0.103
HS	L (lognormal)	0.047
Rk		0.020
HS	G (gamma)	0.306
Rk		0.120

appropriate influence of the tails in the distributions, not just noise from single models. The model probabilities are listed in Table 5. They imply 404, 103, 47, 20, 306, and 120 models for HS–N, Rk–N, HS–L, Rk–L, HS–G, and Rk–G, respectively, if 1000 populations are to be simulated.

The maximum likelihood fit, which considers only the most likely model, selects the Ricker model over the hockey stick. In contrast, the method of Kass and Raftery (1995) considers the co-varying distribution of coefficients, not just the most likely model. As the hockey stick coefficients are less correlated and more precisely estimated than those for Ricker, the hockey stick distribution fits to the S–r observations with more certainty, so receives a greater weighting, resulting in more populations with this characteristic.

The values of *S* before 1980 are uncertain owing to the changing plus group in the catch-at-age data from 1972 to 1979. The values of *r*, however, are as well established as elsewhere in the series. This uncertainty has been explicitly included in the modelling (Supplementary material), and the differences in the results are negligible.

Truncation of simulated values of recruitment

The lower and the upper limits to observations of *r* are 0.96 and 8.42, respectively. These limits are extended for simulation to limits of 0.61 and 9.03. Using a linear fit to the increments in the Q–Q plot, these overall limits are depicted in Figure 2 as the extreme non-filled points on the Q–Q plots.

Table 6. Truncation levels for different parametric distributions used to generate $S-r$ deviants.

Model	G (gamma)	L (lognormal)	N (normal)
Lower limit	0.00024	0.00001	0.0118
Upper limit	0.989	0.974	0.999

Using method 1 to constrain recruitment resulted in a small number of unrealistically high values of more than twice the highest previously observed recruitment. Method 2 gave a spread of limits that constrained some model maxima well below the observed maxima, while delivering others at $>2\times$ the observed maximum, suggesting that this was not an effective method. Method 3 seemed the most realistic, but required setting model-specific limits in each of the 1000 cases. The average probabilities at which the different distributions are clipped are listed in Table 6.

Very few draws from the gamma and lognormal distributions are discarded at the lower end, and similarly for high values from the normal distribution, but significant truncation is required at the high end for lognormal, because the distribution fits poorly (Figure 2). Including the highest 2.5% of values from a lognormal distribution would yield some very high values of recruitment if this was used without truncation. The gamma and normal distributions are less affected, losing just over 1% of high and low values, respectively. To deal with the truncation without creating bias, small adjustments were made to the model parameters to maintain the mean for each truncated model over the observed range of S .

Simulated values of recruitment

The probabilities of the different models are listed in Table 5 and were used to define a 1000 randomly selected model set with the parameters defined from the MCMC chains of the Bayesian analysis (Table 3), truncated as described above. The resulting distribution of simulated recruitment for different levels of S is illustrated in Figure 5a. A comparison of observed and simulated values of S is provided in Figure 5b. The match is a good compromise, better than any model shown above (Figure 2). The mean simulated recruitment is $<2.5\%$ greater than the observed value, and the distribution of deviations is a good match to the observed deviations as described, through either a comparison of cumulative distributions (Figure 5b) or a Q–Q plot (Figure 5c).

Population simulation with full $S-r$ variability

The results of simulating the 1000 selected populations, each parametrized with its own $S-r$ relationship and exploited at a range of constant F computed as the mean for ages 4–8 (F_{4-8}) are shown in Figure 6. For comparative purposes, included on the plots are the historical observed values of r , S , and catch against F_{4-8} from the 2009 assessment (ICES, 2009). These observed values lie within the range of simulated values, confirming that the simulations conform closely to the data. The plots show the relatively stable average recruitment for F_{4-8} of <0.35 . Above that level, there is a rapidly rising probability of declining r , catch, and S . The risk of S being below B_{lim} rises rapidly from $\sim 5\%$ for $F_{4-8} = 0.35$ to around 50% at $F_{4-8} = 0.4$. The wide intervals on S and catch at $F_{4-8} > 0.35$ illustrate the uncertainty in stock dynamics over that range of F . The values of F for MSY vary among the simulated populations, and the point values obtained above are now a pdf (Figure 6d), which indicates a peak at $F_{4-8} = 0.35$.

In this study, we were concerned with the $S-r$ relationship, but in practice, management will involve errors, which can be considered as probability distributions of realized values of F for a given point or variable target F_s . This uncertainty in real F will spread or smear the distributions of r , S , and catch horizontally, reducing the slope on the risk curves and increasing the risks at lower values of F . Therefore, for management purposes, given the uncertainty in realized F , the MSY target F_{4-8} resulting in a low probability of stock depletion will be significantly lower than $F_{4-8} = 0.35$ (ICES, 2007). The management plan derived from these simulations is given as a thick blue line in Figure 6b and c. The plan includes some protection for low biomass by reducing F as S declines below 2.2 million tonnes.

Discussion

Although this work was part of a full MSE for NEA mackerel, the sensitivity of the MSE to the $S-r$ relationship can be judged through its influence on the estimate of MSY for the population. Here, we identified one key area of uncertainty for managing the stock, the slope of the $S-r$ curve as it approaches the origin. It is clear from the maximum likelihood model selection that statistically barely indistinguishable models give totally different responses to exploitation (Figure 3). Had this fact been ignored and standard statistical fit criteria used, we would have given advice that did not recognize the real uncertainty in the situation. We chose to include model uncertainty by modelling multiple populations each with a different $S-r$ function drawn from parameter sets of statistically selected models with similar responses. There are other approaches. One would have been to select one or two models *a priori*, e.g. only the Beverton–Holt and Ricker models, and to evaluate them without consideration of other options. Another option would have been to choose the Shepherd model, which can take either form. By evaluating all three model options and more, though, we have determined whether there is sufficient signal in the data to resolve the issue of choosing the Shepherd formulation over the hockey stick and Ricker forms. Clearly, there are insufficient data to support the Shepherd formulation, the range of $S-r$ curves that result are similar in form to the Ricker form, and the addition of an extra parameter in the Shepherd function is not statistically justified.

The key issue for the estimation of MSY for this stock is the slope of $S-r$ at the origin, and the information available is poor. The methodology presented includes a range of slopes that are quite extreme, with the steepest $10\times$ the shallowest. The steepest slope comes from the Ricker type and implies much higher recruitment at biomasses of just over a million tonnes, well outside the range of the data. However, by including it, we acknowledge its possibility. The models with extreme slopes imply F_{4-8} for maximum yield of 1.5 and 0.29, although F_{crash} for the shallow slope is 0.31, just above the maximum making this level hard to sustain, implying a lower value of F_{MSY} . The extreme cases can be compared with the maximum likelihood models in Figures 1 and 3, which have a factor of 2 for $S-r$ slope at the origin and values of F_{4-8} for maximum yield of 0.55 and 0.33. The overall MSY is estimated at $F_{4-8} = 0.35$ (Figure 6). By explicitly allowing such extreme values to be included in an analysis, the uncertainty is acknowledged and accounted for by assigning probabilities derived directly from what can be shown as known.

This approach has its critics; Rochet and Rice (2009) argued that numerical approaches such as described here disguise a lack

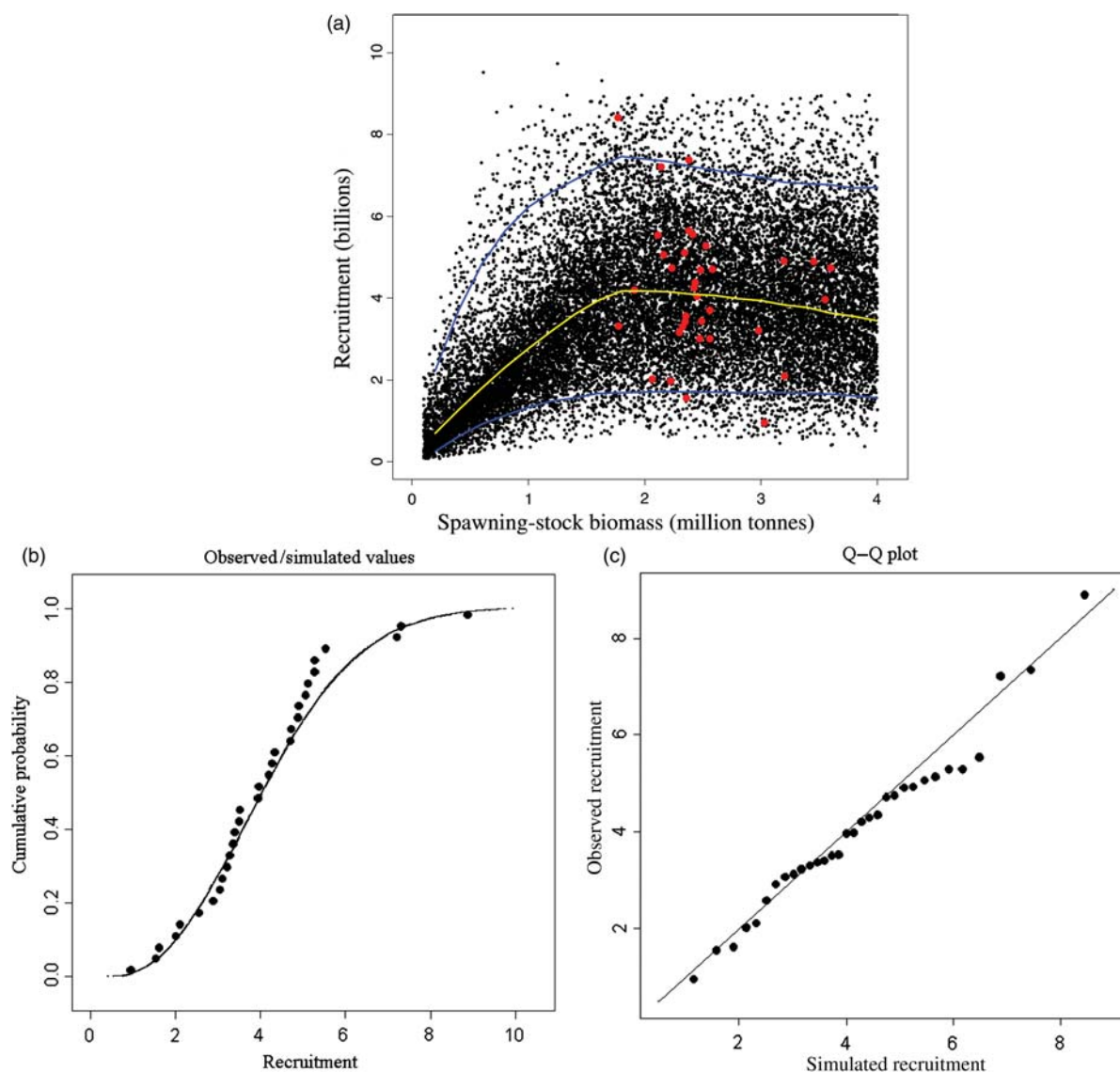


Figure 5. Comparison of observed (red dots) and simulated (small black dots) recruitment (r). (a) Spawning-stock biomass (S) from 0.1 to 4 million tonnes, with 5th (blue), 50th (yellow), and 95th (blue) percentiles, (b) cumulative probability distributions of observed and simulated r values over observed S ranges, and (c) Q–Q plot of observed and simulated r values over observed S ranges. Simulated values are derived from 1000 models with hockey stick and Ricker functional forms and truncated normal or lognormal stochastic deviation.

of knowledge, as in a cloak of scientific illusion. Indeed, if the approach here was described as representing the truth about NEA mackerel population dynamics, we would be guilty of exactly that. The use of an extensive numerical analysis of the type reported here is not an attempt to say that we have defined the recruitment response for the stock, but rather that we have attempted to characterize our ignorance of the dynamics in our evaluation by including a wide range of possible functions based explicitly on the data available. As the fishery is managed not with wise words, but with distinct quantities of catch, a requirement of any analysis is advice in this numerical form. The analysis here works because it is statistically based and results in estimates consistent with historical estimates of recruitment, S , and catch.

In giving advice, one needs to be conscious of the history of the stock. It is particularly important not to be overly optimistic about the reproductive potential for NEA mackerel at low stock size.

Iversen (1981) records the collapse of the North Sea mackerel stock at high, but uncertain, levels of fishing, and that component of the stock has still not recovered to previous levels some 30 years later. Acknowledging the uncertainty in the manner shown here adds the necessary caution.

This approach may appear superficial in its consideration of temporal stability and overelaborate in modelling the relationship with abundance. For some stocks, this approach would indeed be inappropriate. However, it is the uncertainty of the slope of the S – r relationship at the origin that is currently of most importance for this stock compared with concerns related to longer-term change. The apparent temporal stability in recruitment may be because the reproductive strategy for NEA mackerel involves extended spawning in both time and latitude, making recruitment success relatively robust to short-term variability and (relatively) small shifts in climate with respect to the changes through the

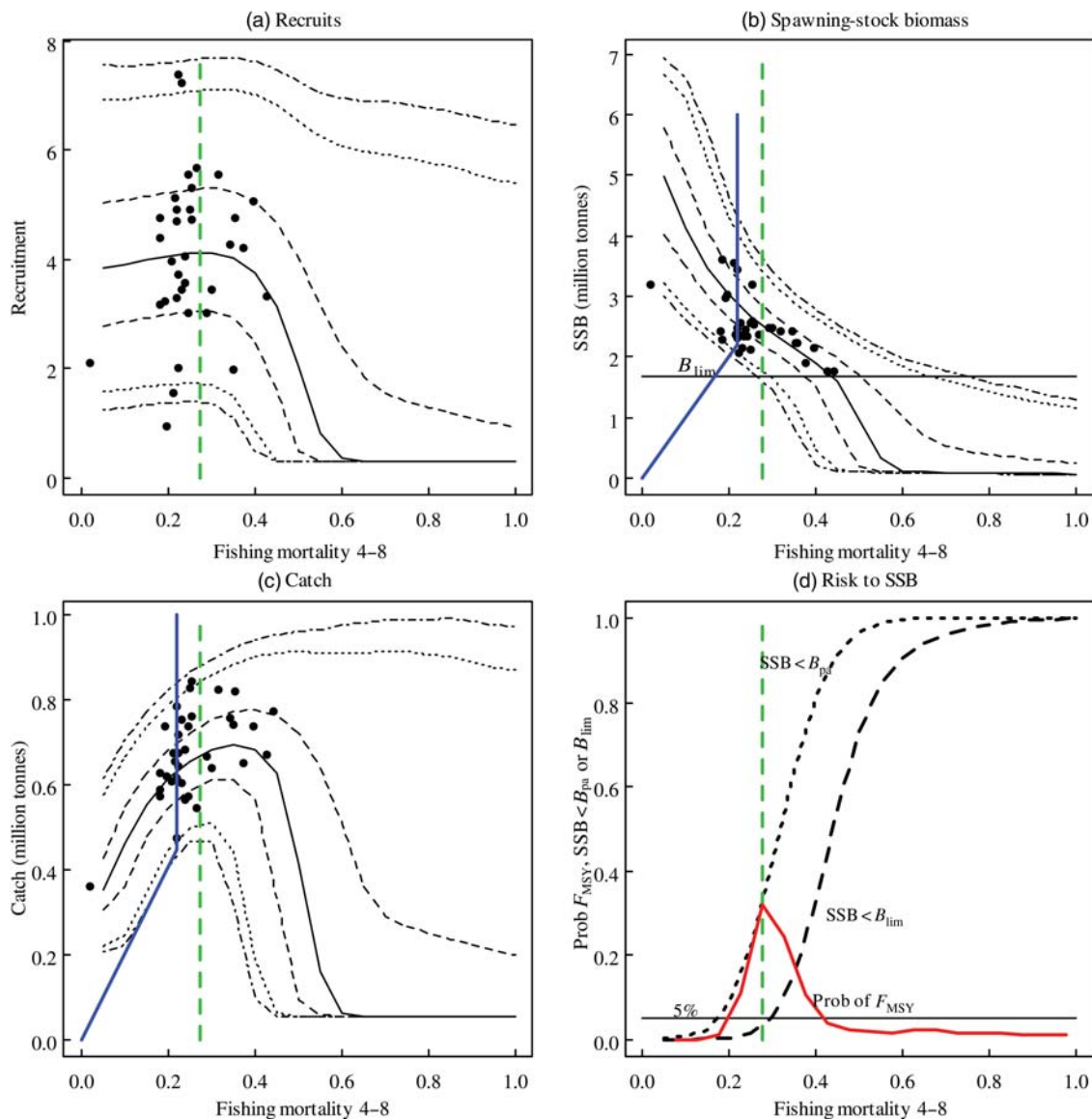


Figure 6. Results from the stochastic equilibrium evaluation of exploitation at constant F (F_{4-8}). Recruitment follows 1000 selected S – r relationships. Plots show 0.025, 0.05, 0.25, 0.5, 0.75, 0.95, and 0.975 quantiles of (a) simulated r , (b) S , and (c) catch. Points indicate historical values of r , S , and catch against F_{4-8} . Plot (d) shows the risk of S being below B_{lim} (thick dashed line), the risk of S being below B_{pa} (thick dotted line), and the distribution of $F = F_{MSY}$ (red solid line), and the F for exploitation at a 5% risk of $S < B_{lim}$ (vertical green dotted line on all panels). The uncertainty in dynamics at higher values of F is illustrated with the wide intervals on S and catch at $F_{4-8} > 0.35$. F_{MSY} and risks are estimated here without observation or implementation error. The current management plan including implementation and measurement error is given as the thick blue line on plots (b) and (c).

spawning season. Evaluations of spawning date (ICES, 2008b) and distribution show only small deviations over the past 20 years for the stock as a whole (Beare and Reid, 2002; Borja *et al.*, 2002).

Including uncertainty in the stock dynamics is an important aspect of many MSEs, and the methodology described here is important for stocks with poorly defined dynamics. Currently, it is rare for multiple biological hypotheses characterized by several S – r functional relationships to be included in evaluations of stock dynamics for setting MSY targets or carrying out an MSE. In 2008, ICES re-evaluated 16 management plans developed over the previous 6 years (ICES, 2008a). Several did include tests of multiple S – r functions, although those that did, such as North Sea

herring or North Sea cod, were limited to a single functional form with two or three values for a parameter to imply environmental change. Sometimes, model choice is important but not considered appropriately; for North Sea plaice, the choice of S – r model between hockey stick and Ricker implies a factor of 2 in F_{MSY} . Although the Ricker model is used in MSE, the implication for the estimate of F_{MSY} is not discussed in the report on MSE (STECF, 2007). For most of these ICES stocks, random noise is added to the values of a fixed function with fixed parameters. Effectively, that implies random variation in the parameter scaling the relationship, but without varying other parameters or the form of the function. Of the simulations reviewed by ICES,

it is only for NEA mackerel that the uncertainty is acknowledged and the type of approach described here used to give uncertainty in both functional form and parameter values.

In our opinion, the approach described here provides a workable methodology. Others have examined this issue in slightly different circumstances. For instance, Brodziak and Legault (2005) addressed the issue of functional form for rebuilding targets and used a similar approach for averaging the results, but their method is less useful for defining risk. The philosophy espoused by Rademeyer *et al.* (2007), in which different hypotheses are included with formal plausible weights, seems to be a good approach. The method provided here is a practical example of how to do this for a situation where the data are not particularly informative but the results depend directly on the choices. Where information on other closely related stocks is available, a move to informative priors may augment this approach.

Had we elected to use a single model here, the results would have been highly sensitive to choices driven by AIC or DIC. With scenario testing (ICES, 2008a), it would have been impossible to determine a composite risk and/or probability distribution for F_{MSY} . The weighting approach of Kass and Raftery (1995) used provides a means of assigning probabilities and providing a set of populations with appropriate weight for each type. This allows risks to be evaluated from the composite of populations, a useful approach and a step forward in providing robust advice better reflecting a lack of knowledge.

To conclude, when the information available does not explain a large proportion of the recruitment variability, care needs to be taken to ensure that the form of the functional relationship and the distribution of the variation are well founded. Under circumstances where the relationship is poorly supported, selecting a single model can have great influence on advised harvest strategies, so is not recommended. Selecting a distribution of models gives a more complete perception of the potential for different dynamics. In this case study, the hockey stick is favoured over the Ricker by 75–25%. The balance between different parametric deviation models is 50, 43, and 7% for normal, gamma, and lognormal, respectively. More complex models with four parameters are rejected by both Bayesian and maximum likelihood methods. For all models, the Bayesian method provides median values of σ slightly higher than that given by the maximum likelihood method. We consider that the multifunctional approach taken here to model recruitment is preferable to the use of a single model.

Supplementary material

Supplementary material is available at the ICES/JMS online version of this manuscript in the form of uncertainty in S before 1980, a table of prior distribution means (μ) and deviances ($\tau = 1/\sigma^2$), and truncation to exclude zero values for each parameter in each model fitted using WINBUGS. Initial values are also provided for three MCMC runs used to check for model convergence.

Acknowledgements

We thank all those in ICES who helped stimulate this work through their contributions to the discussion of the NEA mackerel MSE, and its poorly supported but critical S – r functionality, all those who provided feedback on the MSE for NEA mackerel that was developed through 2007 and 2008, and the referees for helpful criticism that improved the paper and contributed to better description of the work and of the issues included. We take full responsibility for the validity of the approach, however.

References

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19: 716–723.
- Alheit, J., and Niquen, M. 2004. Regime shifts in the Humboldt Current ecosystem. *Progress in Oceanography*, 60: 201–222.
- Anon. 2010. Transboundary Resources Assessment Committee Status Report, 2010. Atlantic mackerel in the Northwest Atlantic, NAFO Subareas 2–6. http://www2.mar.dfo-mpo.gc.ca/science/trac/TSRs/TSR_2010_01_E.pdf.
- Bakun, A., and Weeks, S. J. 2008. The marine ecosystem off Peru: what are the secrets of its fishery productivity and what might its future hold? *Progress in Oceanography*, 79: 290–299.
- Beare, D. J., and Reid, D. G. 2002. Investigating spatio-temporal change in spawning activity by Atlantic mackerel between 1977 and 1998 using generalized additive models. *ICES Journal of Marine Science*, 59: 711–724.
- Best, N. G., Cowles, M. K., and Vines, S. K. 1997. CODA: convergence diagnosis and output analysis software for Gibbs sampling output, version 0.4. MRC Biostatistics Unit, Cambridge. <http://www.mrc-bsu.cam.ac.uk/bugs/classic/coda04/readme.shtml>.
- Beverton, R. J. H., and Holt, S. J. 1957. On the dynamics of exploited fish populations. *Fishery Investigations*, London, Series II, 19. 533 pp.
- Borja, A., Uriarte, A., and Egaña, J. 2002. Environmental factors and recruitment of mackerel, *Scomber scombrus* L. 1758, along the north-east Atlantic coasts of Europe. *Fisheries Oceanography*, 11: 116–127.
- Brodziak, J., and Legault, C. M. 2005. Model averaging to estimate rebuilding targets for overfished stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, 62: 544–562.
- Brunel, T., and Boucher, J. 2007. Long-term trends in fish recruitment in the north-east Atlantic related to climate change. *Fisheries Oceanography*, 16: 336–349.
- Butterworth, D. S. 2007. Why a management procedure approach? Some positives and negatives. *ICES Journal of Marine Science*, 64: 613–617.
- Butterworth, D. S., and Bergh, M. O. 1993. The development of a management procedure for the South African anchovy resource. *Canadian Special Publications of Fisheries and Aquatic Sciences*, 120: 87–99.
- Cochrane, K. L. 2002. A Fishery Managers Guide Book: management measures and their application. *FAO Fisheries Technical Paper*, 424. 235 pp.
- Cushing, D. H. 1973. Dependence of recruitment on parent stock. *Journal of the Fisheries Research Board of Canada*, 30: 1965–1976.
- Drinkwater, K. 2009. Comparison of the response of Atlantic cod (*Gadus morhua*) in the high-latitude regions of the North Atlantic during the warm periods of the 1920s–1960s and the 1990s–2000s. *Deep Sea Research II: Topical Studies in Oceanography*, 56: 2087–2096.
- FAO. 1997. *FAO Technical Guidelines for Responsible Fisheries*, 4. 82 pp.
- FAO. 2003. *World Summit on Sustainable Development 2002 and its Implications for Fisheries*. <http://www.fao.org/docrep/meeting/005/y8294E.htm>.
- Fisher, R. A. 1922. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London, Series A*, 222: 309–368.
- Fogarty, M., Incze, L., Hayhoe, K., Mountain, D., and Manning, J. 2008. Potential climate change impacts on Atlantic cod (*Gadus morhua*) off the northeastern USA. *Mitigation and Adaptation Strategies for Global Change*, 13: 453–466.
- Garcia, S. M. 1996. The precautionary approach to fisheries and its implications for fishery research, technology and management: an updated review. *FAO Fisheries Technical Paper*, 350/2. 210 pp.

- Gelman, A. 2003. A Bayesian formulation of exploratory data analysis and goodness-of-fit testing. *International Statistical Review*, 71: 369–382.
- Gelman, A., Carlin, J. C., Stern, H., and Rubin, D. B. 2004. *Bayesian Data Analysis*, 2nd edn. Chapman and Hall, New York. 670 pp.
- Hilborn, R., and Walters, C. J. 1992. *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Chapman and Hall, New York. 570 pp.
- ICES. 2001. Report of the Study Group on the Further Development of the Precautionary Approach to Fishery Management. ICES Document CM 2001/ACFM: 11. 51 pp.
- ICES. 2005. Report of the *ad hoc* Group on Long Term Advice (AGLTA). ICES Document CM 2005/ACFM: 25. 126 pp.
- ICES. 2007. Report of the Working Group on Mackerel, Horse Mackerel, Sardine and Anchovy. ICES Document CM 2007/ACFM: 36. 725 pp.
- ICES. 2008a. Report of the Study Group on Management Strategies (SGMAS). ICES Document CM 2008/ACOM: 24. 74 pp.
- ICES. 2008b. Report of the Working Group on Widely Distributed Stocks (WGWISE). ICES Document CM 2008/ACOM: 13. 702 pp.
- ICES. 2009. Report of the Working Group on Widely Distributed Stocks (WGWISE). ICES Document CM 2009/ACOM: 12. 583 pp.
- Iles, T. C. 1994. A review of stock–recruitment relationships with reference to flatfish populations. *Netherlands Journal of Sea Research*, 32: 399–420.
- Iversen, S. A. 1981. Spawning and trends in spawning stock size of the North Sea mackerel during the period 1973–1980. ICES Document CM 1981/H: 16. 9 pp.
- Jacobson, L. D., Bograd, S. J., Parrish, R. H., Mendelssohn, R., and Schwing, F. B. 2005. An ecosystem-based hypothesis for climatic effects on surplus production in California sardine (*Sardinops sagax*) and environmentally dependent surplus production models. *Canadian Journal of Fisheries and Aquatic Sciences*, 62: 1782–1796.
- Kass, R. E., and Raftery, A. E. 1995. Bayes factors. *Journal of the American Statistical Association*, 90: 773–795.
- Kell, L. T., Dickey-Collas, M., Hintzen, N. T., Nash, R. D. M., Pilling, G. M., and Roel, B. A. 2009. Lumpers or splitters? Evaluating recovery and management plans for metapopulations of herring. *ICES Journal of Marine Science*, 66: 1776–1783.
- King, J. R., and McFarlane, G. A. 2006. A framework for incorporating climate regime shifts into the management of marine resources. *Fisheries Management and Ecology*, 13: 93–102.
- Köster, F. W., Vinther, M., MacKenzie, B. R., Eero, M., and Plikshs, M. 2008. Environmental effects on recruitment and implications for biological reference points of eastern Baltic cod (*Gadus morhua*). *Journal of Northwest Atlantic Fishery Science*, 41: 205–220.
- Leslie, P. H. 1945. On the use of matrices in certain population mathematics. *Biometrika*, 33: 183–212.
- Leslie, P. H. 1948. Some further notes on the use of matrices in population mathematics. *Biometrika*, 35: 213–245.
- Mackinson, S., Deas, B., Beveridge, D., and Casey, J. 2009. Mixed-fishery or ecosystem conundrum? Multispecies considerations inform thinking on long-term management of North Sea demersal stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, 66: 1107–1129.
- McAllister, M., and Kirchner, C. 2002. Accounting for structural uncertainty to facilitate precautionary fishery management: illustration with Namibian orange roughy. *Bulletin of Marine Science*, 70: 499–540.
- Michielsens, C. G. J., and McAllister, M. K. 2004. A Bayesian hierarchical analysis of stock–recruit data: quantifying structural and parameter uncertainties. *Canadian Journal of Fisheries and Aquatic Sciences*, 61: 1032–1047.
- Moustahfid, H., Link, J. S., Overholtz, W. J., and Tyrrell, M. C. 2009. The advantage of explicitly incorporating predation mortality into age-structured stock assessment models: an application for Atlantic mackerel. *ICES Journal of Marine Science*, 66: 445–454.
- Myers, R. A., and Barrowman, N. J. 1996. Is fish recruitment related to spawner abundance? *Fishery Bulletin US*, 94: 707–724.
- Myers, R. A., Barrowman, N. J., Hutchings, J. A., and Rosenberg, A. A. 1995. Population dynamics of exploited fish stocks at low population levels. *Science*, 269: 1106–1108.
- Nash, R. D. M., Dickey-Collas, M., and Kell, L. T. 2009. Stock and recruitment in North Sea herring (*Clupea harengus*); compensation and depensation in the population dynamics. *Fisheries Research*, 95: 88–97.
- Needle, C. L. 2002. Recruitment models: diagnosis and prognosis. *Reviews in Fish Biology and Fisheries*, 11: 95–111.
- Olaso, I., Gutierrez, J. L., Villamor, B., Carrera, P., Valdés, L., and Abuanza, P. 2005. Seasonal changes in the north-eastern Atlantic mackerel diet (*Scomber scombrus*) in the north of Spain. *Journal of the Marine Biological Association of the UK*, 85: 415–418.
- Quinn, T. J., and Deriso, R. B. 1999. *Quantitative Fish Dynamics*. Oxford University Press, New York. 560 pp.
- Rademeyer, R. A., Plagányi, É. E., and Butterworth, D. S. 2007. Tips and tricks in designing management procedures. *ICES Journal of Marine Science*, 64: 618–625.
- Rätz, H.-J., Bethke, E., Dörner, H., Beare, D., and Gröger, J. 2007. Sustainable management of mixed demersal fisheries in the North Sea through fleet-based management—a proposal from a biological perspective. *ICES Journal of Marine Science*, 64: 652–660.
- Ricker, W. 1954. Stock and recruitment. *Journal of the Fisheries Research Board of Canada*, 11: 559–623.
- Rochet, M.-J., and Rice, J. C. 2009. Simulation-based management strategy evaluation: ignorance disguised as mathematics? *ICES Journal of Marine Science*, 66: 754–762.
- Sette, O. E. 1950. Biology of the Atlantic mackerel (*Scomber scombrus*) of North America. 2. Migrations and habits. *Fishery Bulletin US*, 49: 251–358.
- Shepherd, J. G. 1982. A versatile new stock–recruitment relationship for fisheries and the construction of sustainable yield curves. *Journal du Conseil International pour l'Exploration de la Mer*, 40: 67–75.
- Shimoyama, S., Sakuramoto, K., and Suzuki, N. 2007. Proposal for stock–recruitment relationship for Japanese sardine *Sardinops melanostictus* in North-western Pacific. *Fisheries Science*, 73: 1035–1041.
- Simmonds, E. J., and Keltz, S. 2007. Management implications and options for a stock with unstable or uncertain dynamics: west of Scotland herring. *ICES Journal of Marine Science*, 64: 679–685.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van der Linde, A. 2002. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B*, 64: 583–616.
- Spiegelhalter, D. J., Thomas, A., Best, N., and Lunn, D. 2003. WinBUGS User Manual, version 1.4. <http://www.mrc-bsu.cam.ac.uk/bugs>.
- STECF. 2007. Report of the Working Group on Impact Assessment: Plaice and Sole Long Term Management, SGECA-SGRST 07-01. <https://stecf.jrc.ec.europa.eu/meetings/2007>.
- Stige, L. Chr., Ottersen, G., Brander, K., Chan, K.-S., and Stenseth, N. Ch. 2006. Cod and climate: effect of the North Atlantic Oscillation on recruitment in the North Atlantic. *Marine Ecology Progress Series*, 325: 227–241.
- Stokes, T. K., Butterworth, D. S., Stephenson, R. L., and Payne, A. I. L. 1999. Confronting uncertainty in the evaluation and implementation of fisheries-management systems: introduction. *ICES Journal of Marine Science*, 56: 795–796.
- Yatsu, A., Aydin, K. Y., King, J. R., McFarlane, G. A., Chiba, S., Tadokoro, K., Kaeriyama, M., et al. 2008. Elucidating dynamic responses of North Pacific fish populations to climatic forcing: influence of life-history strategy. *Progress in Oceanography*, 77: 252–268.