ICES Journal of Marine Science Advance Access published November 30, 2012

ICES Journal of Marine Science

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

ICES Journal of Marine Science; doi:10.1093/icesjms/fss183

Fitting a non-parametric stock–recruitment model in R that is useful for deriving MSY reference points and accounting for model uncertainty

Noel G. Cadigan*

Centre for Fisheries Ecosystems Research, Marine Institute of Memorial University of Newfoundland, PO Box 4920, St. John's, NL, Canada A1C 5R3

*Corresponding Author: tel: +1 709 778 0603; fax: 1 709 778 0669; e-mail: noel.cadigan@mi.mun.ca

Cadigan, N. G. Fitting a non-parametric stock–recruitment model in R that is useful for deriving MSY reference points and accounting for model uncertainty – ICES Journal of Marine Science, doi:10.1093/icesjms/fss183

Received 17 May 2012; accepted 6 November 2012.

Modelling the relationship between parental stock size and subsequent recruitment of fish to a fishery is often required when deriving reference points, which are a fundamental component of fishery management. A non-parametric approach to estimate stock–recruitment relationships is illustrated using a simulated example and nine case studies. The approach preserves compensatory density dependence in which the recruitment rate monotonically decreases as stock size increases, which is a basic assumption of commonly used parametric stock–recruitment models. The implications of the non-parametric estimates on maximum sustainable yield (MSY) reference points are illustrated. The approach is used to provide non-parametric bootstrapped confidence intervals for reference points. The efficacy of the approach is investigated using simulations. The results demonstrate that the non-parametric approach can provide a more realistic estimation of the stock–recruitment relationship when informative data are available compared with common parametric models. Also, bootstrap confidence intervals for MSY reference points based on different parametric stock–recruitment models often do not overlap. The confidence intervals based on the non-parametric approach tend to be much wider, and reflect better uncertainty due to stock–recruit model choice.

Keywords: Beverton–Holt model, confidence intervals, hockey-stick model, monotone spline, Ricker model.

Introduction

Modelling the relationship between parental stock size (S) and reproduction and subsequent recruitment (R) of juveniles to a fishery is widely recognized as a fundamental component of sustainable fishery management (Quinn and Deriso, 1999). For example, stock–recruit (SR) relationships are used to project future fish population dynamics in response to proposed management actions, and to determine management reference points (Needle, 2002). Many fisheries are managed using reference points (RPs), where prescribed actions should occur when stock size or fishing mortality rates exceed RP values. Some RPs are derived directly from the SR relationship. An example of this is the spawning stock size corresponding to 50% of maximum recruitment which may be taken as a biomass limit (Myers *et. al.*, 1994). This RP is usually estimated using an SR model. Other RPs take into account other aspects of stock productivity, and an important example is described in the next paragraph. RPs are widely considered an essential part of well-managed fisheries (e.g. Hilborn and Stokes, 2010; Hutchings and Rangeley, 2011). Reliable SR models are therefore important for successful fishery management.

Maximum sustainable yield (MSY) RPs have been adopted by many national (e.g. the USA and New Zealand) and international fishery management agencies (e.g. IWC, ICCAT, IATTC, EU, and NAFO). The fishing mortality rate that maximizes long-term yield (F_{MSY}) is often taken to be an upper limit for management purposes, while the resulting biomass at F_{MSY} (i.e. B_{MSY}) may be a default target. Although using MSY RPs is not without criticism (e.g. Hilborn, 2010; Legović *et al.*, 2010), if these RPs are to be used for management then it is important to have reliable estimates of them. MSY RPs are essentially derived from long-term stock projections over a range of fishing mortalities. If the projections are deterministic, then the calculation of MSY RPs is also deterministic. In this context, some theory has been developed related to MSY calculations (see below). MSY RPs are determined by the growth and mortality process of the stock, and by the age pattern in fishing mortality. The SR relationship is a fundamental component in the stock growth process. Consequently, the SR relationship has a major impact on MSY RPs. In reality, the growth and mortality processes will not be constant and will fluctuate depending on other changes in the stocks ecosystem; however, methods (e.g. Bousquet *et al.*, 2008; Cadigan, 2011) to deal with stochastic productivity are beyond the scope of this paper.

Parametric SR models are commonly used to compute MSY RPs, especially when MSY calculations include inferring R outside the range of observed stock sizes. A parametric SR model expresses R as an analytic function of S and a small number of unknown parameters θ that must be estimated. Two SR models commonly used are the Ricker (RK; Ricker, 1954) and the Beverton–Holt (BH; Beverton and Holt, 1957). There are other formulations (e.g. Needle, 2002), and most have a compensatory mortality property (CMP; e.g. Quinn and Deriso, 1999) in which the recruitment rate (R/S) declines monotonically as S increases; that is,

$$\frac{\mathrm{d}(R/S)}{\mathrm{d}S} < 0$$

Both the BH and RK have this property (see the Material and Methods). Some models, such as the RK, have overcompensation (Quinn and Deriso, 1999) in which $\mathrm{d}R/\mathrm{d}S < 0$ at larger stock sizes. The BH has pure compensation, in which compensation, but not overcompensation, occurs for all values of S.

The CMP ensures that for any level of fishing mortality (F) there is a unique equilibrium stock size and age distribution. This equilibrium is determined by both the production of R from S (i.e. the SR relationship) and the production of S from R (survival and growth). A unit of R survives fishing mortality F and natural mortality, and grows to produce some amount of S over the lifetime of the cohort. The relationship is assumed to be linear in R, with a slope determined by growth and survival parameters. Let $Z_a = F\,S_a + M_a$ be the total mortality rate which is the sum of age-specific fishing mortality (F times selectivity, S_a) and natural mortality (M_a) rates. Let $\Lambda_a = \sum_{i=0}^{a-1} \exp(-Z_a)$ be the cumulative per-recruit survival rate to age a at the beginning of the year, and let w_a and Ω_a be the stock weight and maturity at age a. For standard MSY calculations, M_a, S_a, Ω_a, and w_a are considered to be known and constant over time. Mid-year parental stock size is given by $S(R,F) = R\sum_{a=0}^{\infty} \Omega_a \Lambda_a w_a e^{-(F\,S_a+M_a)/2}$. The term $e^{-(F\,S_a+M_a)/2}$ is dropped if $S(R,F)$ is considered to be at the beginning of the year. Mortality and growth parameters may be affected by S, but this is beyond the scope of this paper. Using the Baranov catch equation, the equilibrium fishery yield from a cohort is

$$Y(R,F) = R\sum\nolimits_{a=0}^{\infty} \Lambda_a c_a \left\{1 - e^{-(F\,S_a+M_a)}\right\} \frac{F\,S_a}{Z_a},$$

where c_a is the catch weight-at-age. The intersection of the $S(F,R)$ line as a function of R with F fixed, and the SR curve, determines the equilibrium values of R and S at the fixed level of F [i.e. $R_e(F)$ and $S_e(F)$; e.g. see Sissenwine and Shepherd, 1987]; that is, if $SPR(F) = S(R,F)/R$, which is independent of R, and if $\mu(S)$ is the SR function, then $S_e(F)$ is the x solution to $\mu(x) = x/SPR(F)$ and $R_e(F) = \mu\{S_e(F)\}$. The equilibrium yield is $Y_e(F) = Y\{R_e(F),F\}$ and the value of F that maximizes $Y_e(F)$ is F_{MSY}. B_{MSY} is $S_e(F_{\mathrm{MSY}})$. If $\mu(S)/S$ declines strictly monotonically with S, then there is only one intersection between $S(R,F)$ and the SR curve. Otherwise, there may be multiple intersections and potentially multiple $S_e(F)$ values for some values of F. The CMP ensures that $Y_e(F)$ is continuous in F with a well-defined maximum at F_{MSY}.

The SR parameters θ are usually estimated by minimizing the log error sum of squares based on a time-series of SR observations, although other estimation procedures have been advocated (e.g. Walters, 1990; Jiao *et al.*, 2004; Michielsens and McAllister, 2004). SR data are usually noisy, and both the BH and RK models often fit the data almost equally as well. In this case, the choice of model is a partially subjective decision, although there may be *a priori* reasons based on the life-history characteristics of the species to favour one model over the other (e.g. Quinn and Deriso, 1999).This choice of SR model may have substantial impact on MSY RPs and other RPs (e.g. Bravington *et al.*, 2000; Brodziak and Legault, 2005). The fitting of the SR relationship at low S values will be influenced by values of R at high S values for parametric models, and some researchers find this objectionable. Characterizing this uncertainty and understanding its implications on management advice is important.

Some researchers have utilized non-parametric approaches as a way to avoid the sensitivity of advice to parametric SR model assumptions. Deriving RPs and associated measures of uncertainty using a non-parametric SR model may account more fully for model uncertainty. Evans and Rice (1988) estimated the probability density function of R at a given S using non-parametric algorithms based on the distribution of past R values at similar S values. Cook (1998) used the LOWESS smoother (Cleveland, 1981) to fit an SR relationship and derive a fishing mortality RP. Duplisea and Fréchet (2010, 2011) recently used a similar approach involving a cubic smoothing spline to derive a recruitment–overfishing stock size RP. Bravington *et al.* (2000) noted two problems with using the LOWESS smoother for fitting SR relationships: (i) the choice of the amount of smoothing and (ii) the possibility of biologically unreasonable fits. A LOWESS smoother will not in general have the CMP and, therefore, may not result in unique equilibrium results and MSY RPs. Cubic smoothing splines will suffer from the same problems. Bravington *et al.* (2000) proposed two non-parametric spline smoothers with the CMP property. Their DIMPOS smoother ensures CMP. Their CONCR smoother ensures smooth concavity, which is a strong form of CMP. For example, the Ricker model is CMP but not concave everywhere. Concavity implies CMP, but not the reverse. Bravington *et al.* (2000) noted that these shape restrictions on the smoothers greatly reduced the sensitivity of SR estimates to smoothing parameters, and this basically addressed their two concerns with the LOWESS approach.

Bayesian methods have also been used for non-parametric, or less parametric, SR model fitting and RP estimation. Munch *et al.* (2005) used a Gaussian process model approach for flexible modelling of SR relationships. Fronczyk *et al.* (2011) developed a more general and fully non-parametric Bayesian approach based on a bivariate normal mixture model for the joint distribution of S and $\log(R/S)$. Their goal was a flexible model for the joint density of S and $\log(R/S)$ that could accommodate skewness, excess variability, possible multimodalities, etc. The approaches of both Munch *et al.* (2005) and Fronczyk *et al.* (2011) do not ensure CMP. Brodziak and Legault (2005) used Bayesian model averaging to estimate rebuilding targets under alternative hypotheses about SR dynamics. Their objective was to account for some of the uncertainty in the SR model, including the form of the SR model (BH or RK), error structures, and assumptions about prior information. Simmonds *et al.* (2011) used a similar framework to Brodziak and Legault (2005), but they performed a stochastic MSY evaluation using random recruitment drawn from Bayesian model averaging of two parametric SR models and three error distributions. Neither Brodziak and Legault (2005) nor Simmonds *et al.* (2011) proposed Bayesian model averaging as a semi-parametric SR model. Their applications only partially accounted for SR model uncertainty because they considered only two model choices.

In this paper the *scam* (shape constrained additive models) package in R (Pya, 2012) is used to fit non-parametric SR curves with CMP, and to derive "SR-non-parametric" yield curves. This is an approach that should be easy to use for anyone familiar with the R programming language. The focus is on describing the range of SR relationships that are consistent with data, and the implications of this range on MSY RPs. Bootstrap methods are used to produce non-parametric confidence intervals (*CI*s) for the SR relationship, and for MSY RPs. Some simulation results are presented to examine the efficacy of estimates of F_{MSY} and B_{MSY} based on the non-parametric SR approach compared with estimates based on common parametric SR models.

Material and Methods

Let $\mu(s) = E(R|S = s)$ denote the SR model, that gives the expected value of R as a function of S. The BH model is $\mu(s) = \alpha s/(\beta + s)$, $\alpha,\beta > 0$. It is not difficult to show that $d\{\mu(s)/s\}/ds = -\alpha/(\beta + s)^2$ is always negative so this model has the CMP. The $\max_S \mu(S) = Rmax = \alpha$, and the value of S that corresponds to 50% of *Rmax* is $S50\% = \beta$. The slope at the origin is $Sao = \alpha/\beta$. The RK model is $\mu(s) = \alpha s\exp(-\beta s)$, $\alpha,\beta > 0$, and $d\{\mu(s)/s\}/ds = -\alpha\beta\exp(-\beta s)$ is always negative so this model also has the CMP. $Rmax = \alpha/\beta\exp(1)$ and $Sao = \alpha$. A closed form solution for *S50%* does not exist, but it can be found numerically. For this paper, both the BH and RK model parameters (α and β) were estimated using the *nls* function in the R software package (R Development Core Team, 2011), which is convenient for producing bootstrapped parameter estimates. A segmented regression SR model, often referred to as the hockey-stick (HS) model (Barrowman and Myers, 2000), was also investigated. To facilitate estimation using *nls*, a slightly smoothed version was used which is described in Mesnil and Rochet (2010),

$$\mu(s) = \alpha\left[s + \sqrt{\delta^2 + \gamma^2/4} - \sqrt{(s-\delta)^2 + \gamma^2/4}\right],$$

with $\gamma^2 = 0.1$. $Sao = 2\alpha$, $R\max \cong 2\alpha\delta$, and $S50\% \cong \frac{\delta}{2}$.

Parameters were estimated by minimizing the log error sum of squares. Occasionally, the *nls* minimization did not converge, particularly when bootstrapping parameter estimates. This was usually because the optimum solution was a straight line through the origin with a very large and unrealistic *Rmax*. This is a common problem when fitting and bootstrapping SR models (e.g. Overholtz, 1999). As a remedy, the BH and RK models were constrained to have *Rmax* less than the maximum value of R in the data. This was never a problem for the HS model, although for this approach it is necessary to constrain $S_{min} \le \delta \le S_{max}$. The objective function is flat for values of δ outside this range. These constraints are practical but they can result in inaccurate statistical inferences, especially when the constraints are not valid for the population model (e.g. *Rmax* is really greater than the maximum value of R in the data).

The *scam* function in R (ver. 1.1-2; Pya, 2012) was used to fit non-parametric SR curves with CMP. The *scam* function is part of the *scam* package, which is similar to the *mgcv*(*gam*) package (Wood, 2006, 2011) for fitting GAMs (generalized additive models), except that *scam* allows for a variety of shape constraints on the component functions of the linear predictor of the GAM. The constraints involve monotonicity (increasing or decreasing) with options to specify convex or concave shapes. Both of these packages use spline smoothers for non-parametric regression. Briefly, the data are assumed to be independent and from an exponential family distribution (e.g. Normal, Gamma, Binomial, etc.) with a mean μ that is a partially linear function of covariates $z_1, \ldots, z_d$, and $x_1, \ldots, x_p$; that is, $g(\mu) = \theta_o + \sum_{k=1}^{d}\theta_k z_k + \sum_{j=1}^{p} f_j(x_j)$, where g is a smooth monotone link function and the $f_j(x_j)$ values are unknown smooth functions of the x_j covariates. The θ_k and $f_j(x_j)$ values must be estimated. In the remainder of this section, the j index in f is dropped for simplicity, but basis functions and spline parameters are required for each smooth function.

The spline method involves approximating $f(x)$ as a linear combination of q known spline basis functions [i.e. $B_1^m(x), \ldots, B_q^m(x)$] and unknown spline parameters (i.e. $\gamma_1, \ldots, \gamma_q$) to estimate, $f(x) = \sum_{i=1}^{q} B_i^m(x)\gamma_i$, where m indicates the number of continuous derivatives. The *scam* package uses B-spline basis functions (e.g. Wood, 2006) which have some attractive properties for the purpose of *scam*. These basis functions are essentially a sequence of polynomials that are "centred" at different points across the range of x. The spline function is a weighted average of these polynomials. The basis functions require a sequence of "knots" to be specified, and the knots determine the location and shape of the polynomials. The *scam* package assumes the knots are evenly spaced and cover the range of x. The number of knots (q) should be large to avoid oversmoothing/underfitting; however, this means that the number of spline parameters will usually be large, which could result in overfitting of the data. To avoid the latter problem, a penalty function is used to control the variation in the γ values (see below). A smoothing parameter determines the contribution of the smoothing penalty function to the total fit function, and the *scam* package uses generalized cross-validation (GCV) or the Akaike information criterion to determine the value of the smoothing parameter.

The *scam* package uses various shape restrictions on the γ spline parameters to ensure the correct shape restrictions on f. For example, if $f(x)$ is strictly monotone increasing in x, then a sufficient condition for this shape constraint when using B-splines is $\gamma_i > \gamma_{i-1}$ for $i = 2, \ldots, q$. This is achieved by redefining the γ parameters as

$$\gamma_1 = \beta_1,\ \gamma_j = \beta_1 + \sum_{i=2}^{j}\exp(\beta_i),\ \text{for } j = 2, \ldots, q,$$

where the β_i values are unknown but unconstrained parameters to estimate. The penalty function is based on the squared differences $\beta_j - \beta_{j-1}$ for $j = 2, \ldots, q$. However, the shape restrictions themselves add much smoothing, and model results are usually not that sensitive to the choice of smoothing parameter. A more thorough description of the *scam* model approach is provided in Pya (2010). Details of how *scam* was used for estimating the SR relationship are given in the Supplementary material.

Although *scam* results were not highly sensitive to the value of the smoothing parameter, there were some differences especially when extrapolating *R* for *S* values less than the minimum observed or greater than the maximum observed, which could result in substantial differences in yield per recruit curves. To examine sensitivity to the smoothing parameter, *scam*s were also computed for degrees of freedom (df) fixed at 3 and 4. When the smoothing parameter is very large, then *scam* produces a straight-line fit to $\log(R/S)$ vs. *S*, which is the RK model. Hence, for a large smoothing parameter, the RK and *scam* model estimates are identical. When the SR data are very noisy, then the GCV statistic tends to be minimized by a large value of the smoothing parameter, and *scam* estimates will be very similar to RK estimates. Occasionally, *scam* SR model estimates were linear for all values of *S*, similar to the problems encountered with the BH and RK models. In this case and for MSY RP calculations, *Rmax* was set at the maximum of the observed recruitments. However, this was not necessary when calculating bootstrap *CI*s for the SR model, unlike the BH and RK models. For these latter models, the *nls* bootstrap function did not return a result when the SR parameters did not converge.

Scam uses a Bayesian approach to obtain a covariance matrix of the model coefficients and credible intervals for predictions. The resulting standard errors were really large for extrapolations and they did not seem to reflect the shape constraints. Also, propagating these standard errors into uncertainties about MSY RPs is not straightforward. A solution is to use bootstrap methods to derive standard errors and *CI*s (e.g. Overholtz, 1999). Residuals were resampled with replacement and added to model predictions to generate bootstrapped datasets. The *scam* model was fit to the bootstrapped datasets using the same smoothing parameter as determined for the original data, and this was used to obtain bootstrapped predictions of the SR relationship. These bootstrapped predictions were then used to compute *CI*s for MSY RPs, using the percentile *CI* method (e.g. Efron and Tibshirani, 1993). Two thousand bootstrap resamples were used. More accurate *CI* methods (e.g. bias-corrected, BC_a, bootstrap-t) are available but are beyond the scope of this paper. Bootstrapped distributions for MSY RPs were obtained using the intersections of the *S* per *R* line and the bootstrapped SR curves. This could be done analytically for the parametric SR models, and numerical methods were used for the non-parametric SR curves. The bootstrapped distributions for MSY RPs only reflect uncertainty about the SR relationship, and not uncertainties about future stock weights, maturities, natural mortality, and fishery selectivity.

The various estimators were illustrated using a simulated example. The SR simulation-generating model was a 50:50 mixture of BH and RK models, each with $Rmax = 600 \times 10^6$ and $S50\% = 250$ t. Also, some constant recruitment (i.e. 5% of *Rmax*) was added to represent a small amount of recruitment immigration. This is recruitment from another stock that is transported to the area of the stock in question where it merges with this stock. This is similar to the open-mixture model considered in Munch *et al.* (2005). One hundred SR observations were generated based on a broad range of *S* values; that is, lognormal with mean 1000 and a log s.d. of 1. Lognormal observations of *R* were generated using a log s.d. of 0.2. This is an SR-informative data generator, which was chosen to avoid problems with data deficiencies. Such problems will be highlighted later in real data examples. The various SR models were estimated and used to derive MSY RPs. Biological parameters (i.e. weights, maturities, etc.) for MSY RPs were taken as those for Barents Sea cod (see Table 1). The simulation true values for F_{MSY} and B_{MSY} were based on the open-mixture model without error.

Table 1. Stocks and sample sizes (*n*).

ID	Stock name, area	*n*
1	Cod (*Gadus morhua*), Subdivision 25–32	44
2	North-East Arctic cod (*Gadus morhua*), Subareas I and II	63
3	Icelandic cod (*Gadus morhua*), Division Va	54
4	Faroe Plateau cod (*Gadus morhua*), Subdivision Vb1	49
5	Cod in Division VIa (*Gadus morhua*), West of Scotland	33
6	Cod in Division VIIa (*Gadus morhua*), Irish Sea	43
7	Cod in Divisions VIIe-k (*Gadus morhua*), Celtic Sea	40
8	North-East Arctic saithe (*Pollachius virens*), Subareas I and II	49
9	Plaice in Division VIIe (*Pleuronectes platessa*), Western Channel	31

Data source: ICES Data Centre Stock Summary Database (http://www.ices.dk/datacentre/StdGraphDB.asp).

This basic procedure was also repeated 1000 times, and the results were used to approximate the bias and mean squared error (average squared difference in estimates and true value) of the various estimator of F_{MSY} and B_{MSY}. A subtle difference was the way the *S* values were generated. The procedure above will occassionaly produce an anomalous *S* value that can cause convergence problems for the parametric models. Such anomalies are not realistic because *S* values are usually highly autocorrelated in practice. As a remedy for this problem, the $\log(S)$ values were generated from an AR(1) process with a correlation of 0.5 and a stationary variance of 1, and these were used to generate *R* values. Lognormal bias correction was applied so that $E(S) = 1000$. These simulations were repeated using the BH and RK models as the simulation-generating models. The purpose of this was to examine the potential loss of efficiency in using a non-parametric approach compared with the correct parametric model.

The various estimators were also applied to real SR datasets (see Table 1) obtained from the ICES Data Centre Stock Summary Database (see http://www.ices.dk/datacentre/index.asp), which was accessed in January 2012. Stocks were chosen based on visual evidence of density dependence in *R* and also based on the availability of biological parameters for yield per recruit analyses. This information is provided in the *Yieldrecruit* tables of the database and includes the stock assessment range of ages, and age-specific values for the proportion mature, natural mortality, fishery selectivity, catch weights, and stock weights. Not all stocks in the database had this information. Data for some stocks will change in the future as new estimates become available. The data are only used to illustrate the methods, and no conclusions are otherwise drawn about these stocks. Some of the stocks were also considered by Fronczyk *et al.* (2011), and their results demonstrated that their non-parametric SR model does not ensure the CMP, and some of their model predictions had curious behaviours, such as transient increases in productivity as stock size increased, which do not seem realistic.

Results

The potential benefits of the non-parametric SR estimator are first illustrated using a simulated example involving a somewhat complicated mixture SR relationship. The BH, RK, and HS models did not fit the simulated data well (Figure 1; Table 2). Systematic discrepancies between model fits and the data were apparent. However, *scam* with the smoothing parameter selected to minimize the GCV statistic tracked the trend in the data well, and the true (i.e. simulation generator) SR relationship (grey curve) is mostly contained within the 95% bootstrapped *CI*s. The *scam* smoother had 6 df, whereas the parametric models only had 2 df. The equilibrium yield curve inferred from *scam* was very similar to the curve based on the true SR relationship (Figure 2), whereas this was not the case for the BH, RK, and HS models. Bootstrapped *CI*s for F_{MSY} based on the BH, RK, and HS models did not include the value based on the true SR curve (Table 2). Only the B_{MSY} *CI*s based on the *scam* and BH models contained the true B_{MSY}. MSY RPs based on the non-parametric smoother were very accurate.

The results from the 1000 simulations were consistent with the above example. The *scam* model had the lowest bias and the second lowest root mean squared error (RMSE) for F_{MSY} and the lowest RMSE for B_{MSY} (Table 3). The *scam* model also had low bias when data were generated from BH or RK models, although the approach was less efficient (i.e. higher RMSE) than the correct parametric model. The ratio of RMSE for the correct

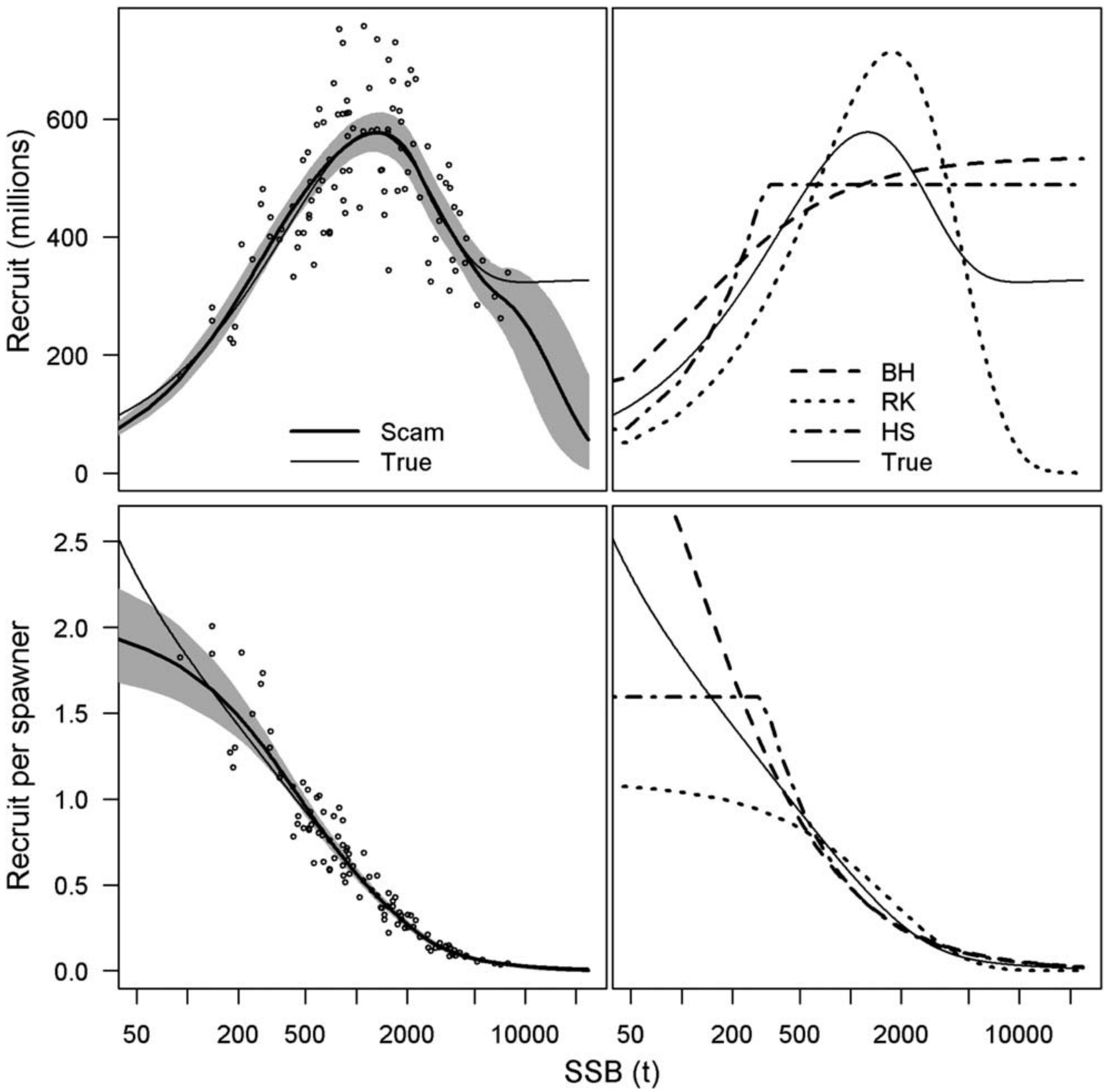

Figure 1. Model fits. See Table 2 for model names. Shaded regions indicate 95% bootstrapped confidence intervals based on the non-parametric *scam* model. The SSB axis is in the natural log scale.

Table 2. Stock–recruit (SR) model fits and associated MSY reference points for a simulated example.

					Percentiles			Percentiles	
SR model	MSE	df	GCV	F_{MSY} Est	LCI	UCI	B_{MSY} Est	LCI	UCI
Scam	0.030	6.0	0.034	0.44	0.39	0.50	1187	1028	1397
HS	0.056	2.0	0.058	0.54	0.54	0.54	773	737	808
BH	0.065	2.0	0.067	0.41	0.38	0.44	1110	969	1247
RK	0.102	2.0	0.106	0.41	0.39	0.43	1605	1526	1701

MSE is the mean squared lognormal residual, df is the model degrees of freedom, and GCV is the generalized cross-validation statistic.
HS, hockey-stick; BH, Beverton–Holt; RK, Ricker.
Estimates (Est) are followed by 2.5th (LCI) and 97.5th (UCI) bootstrap percentiles. $F_{MSY} = 0.45$ and $B_{MSY} = 1158$ based on the true SR relationship.

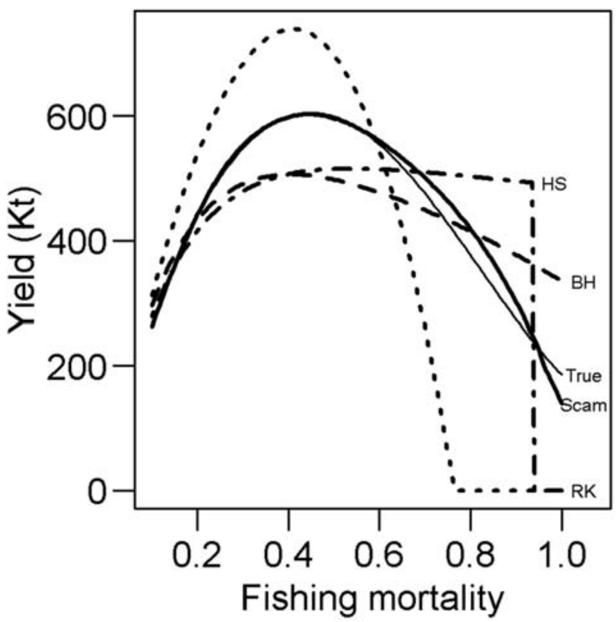

Figure 2. Equilibrium yield curves for different SR models. The true curve is based on the SR simulation generating model. See Table 2 for SR model names.

BH model was 7 for F_{MSY} and 4 for B_{MSY}. The efficiency of the *scam* model was much better when the true SR was RK. Estimates of F_{MSY} and B_{MSY} based on the HS model had substantial bias. The HS F_{MSY} estimates were always constant and equal to F_{max}—the value of F that maximizes yield per recruit. When averaged over the three SR simulation scenarios, the *scam* model had the lowest RMSE for F_{MSY} and B_{MSY}. Convergence in the simulations was very good, with only 1–2 cases not converging for either the BH or HS models. The RK and *scam* models always converged.

The *scam* fits for five of the nine real data examples were identical to the RK model fits when the *scam* smoothing parameter was estimated by minimizing the GCV statistic (i.e. *scam* GCV; Figures 3 and 4). This was the case for stocks 2, 4, 5, 7, and 9. For these stocks, the *scam* GCV and RK fits are coincident in Figures 3 and 4, and only the *scam* curves are visible. For all nine stocks, the fitted *scam* GCV had df < 4, which is why the *scam* GCV curves are smoother than the *scam* fits when df was constrained to be 4 (i.e. *scam* df = 4). The *scam* df = 3 results were smoother than the df = 4 results, as expected, but overall these results were similar. For simplicity, the *scam* df = 3 results are not shown in Figures 3–6. The overall pattern was that the BH and *scam* models had the same or higher *Sao* values (Figure 4) than the RK model, and the RK and *scam* models usually had lower *Rmax* values than the BH models (Figure 3). The HS fits are not shown in Figures 3 and 4 to simplify the presentation, but they resulted in the lowest *Sao* and *Rmax* values compared with the RK, BH, and *scam* models.

The GCV statistics were mostly similar for all six SR models (i.e. HS, RK, BH, *scam* df = 3, *scam* df = 4, *scam* GCV) for each stock, indicating similar fits to the data. Over all stocks, the maximum GCV (for the six SR models) was at most 18% greater than the minimum GCV, and on average it was 8% greater. The HS model did not fit as well for stocks 1 and 7, and, after discarding this model, the maximum GCV was 5% greater than the minimum. Patterns in the residuals were also similar (Figure 5). Plotting the raw SR residuals from each model on the same panel was not informative; therefore, model-specific trends in residuals were inferred using the *loess* smoother (R Development Core Team, 2011) with a span of 0.5. Smoothed trends were interpolated over a mesh of 1000 S values within the range of observed S values. The *scam* df = 4 model resulted in the lowest mean absolute smoothed error. This error was ranked across models for each stock. The average ranks across stocks were: 1.0 (*scam* df = 4), 2.7 (*scam* df = 3), 3.3 (*scam* GCV), 4.1 (RK), 4.9 (BH), and 5.0 (HS). Interestingly, the smooth pattern in residuals was similar for eight of nine stocks (Figure 5). This pattern was a decreasing trend (from positive to negative residuals) at low S values, followed by an increase and then a further decline or a levelling off. The exception is stock 2 (i.e. Barents Sea cod). Possible causes of this pattern are considered in the Discussion.

Bootstrap *CIs* for the two *scam* models in Figure 3 (i.e. *scam* GCV and *scam* df = 4) were very wide outside the range of the S data. Fits in this region are only determined by the monotonicity shape constraint on R/S and the second derivative penalty function. Bravington *et al.* (2000) reported a similar finding.

The equilibrium yield curves derived using the various SR models differed substantially for some stocks (Figure 6). The crash behaviour associated with the HS model is well known (see Mesnil and Rochet, 2010). An overall pattern was for higher yields implied by the BH model at low F values compared with the other models, and lower yields at low F values implied by the *scam* df = 4 model. The high BH yields at low F values usually involve extrapolations beyond the range of SR data. Conversely, at high F values, the *scam* df = 4 model indicated the highest yields, and the RK and HS models resulted in lower yields, including zero yields. Values for F_{MSY} (Figure 7) and B_{MSY} (Figure 8) differed substantially among methods for many of the stocks (i.e. 2, 3, 5, 6, 7, and perhaps 9). The F_{MSY} values

Table 3. Percentage bias and root mean squared error (RMSE) of estimators of F_{MSY} and B_{MSY} derived from four stock–recruit models (rows) based on 1000 simulations from three population models (columns).

SR model	F_{MSY}						B_{MSY}					
	Bias			RMSE			Bias			RMSE		
	BH	RK	MIX	BH	RK	MIX	BH	RK	MIX	BH	RK	MIX
Scam	2.0	0.2	−1.4	13.9	2.7	9.2	0.2	0.0	1.8	19.1	3.8	13.7
HS	55.1	9.9	7.3	55.1	9.9	7.3	−47.2	−23.2	−22.1	47.3	23.2	22.3
BH	0.1	−23.4	−16.5	2.0	24.7	16.7	0.3	6.3	17.7	4.8	28.0	19.3
RK	15.7	0.1	6.1	19.5	1.7	13.3	22.5	−0.1	28.1	34.2	2.0	42.2

The models are: BH, Beverton–Holt; RK, Ricker; HS, hockey-stick; MIX, mixture model.

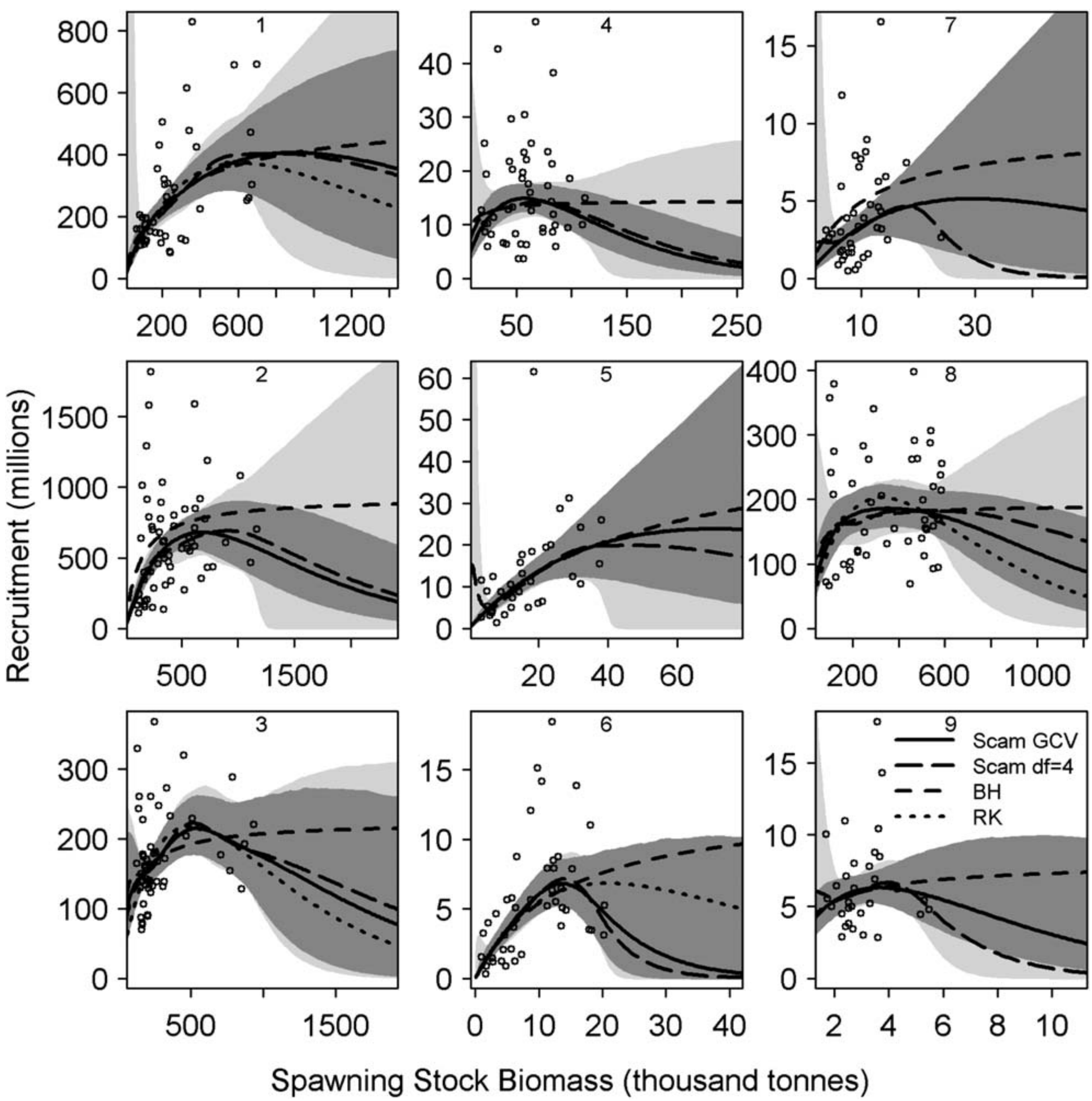

Figure 3. Model fits to SR data (points) for nine stocks, with stock IDs listed at the top of each panel. See Table 1 for stock names and Table 2 for model names. Shaded regions indicate 95% bootstrapped confidence intervals based on the non-parametric *scam* model with df = 4 (lighter shading) or smoothing parameter selected to minimize GCV (darker shading).

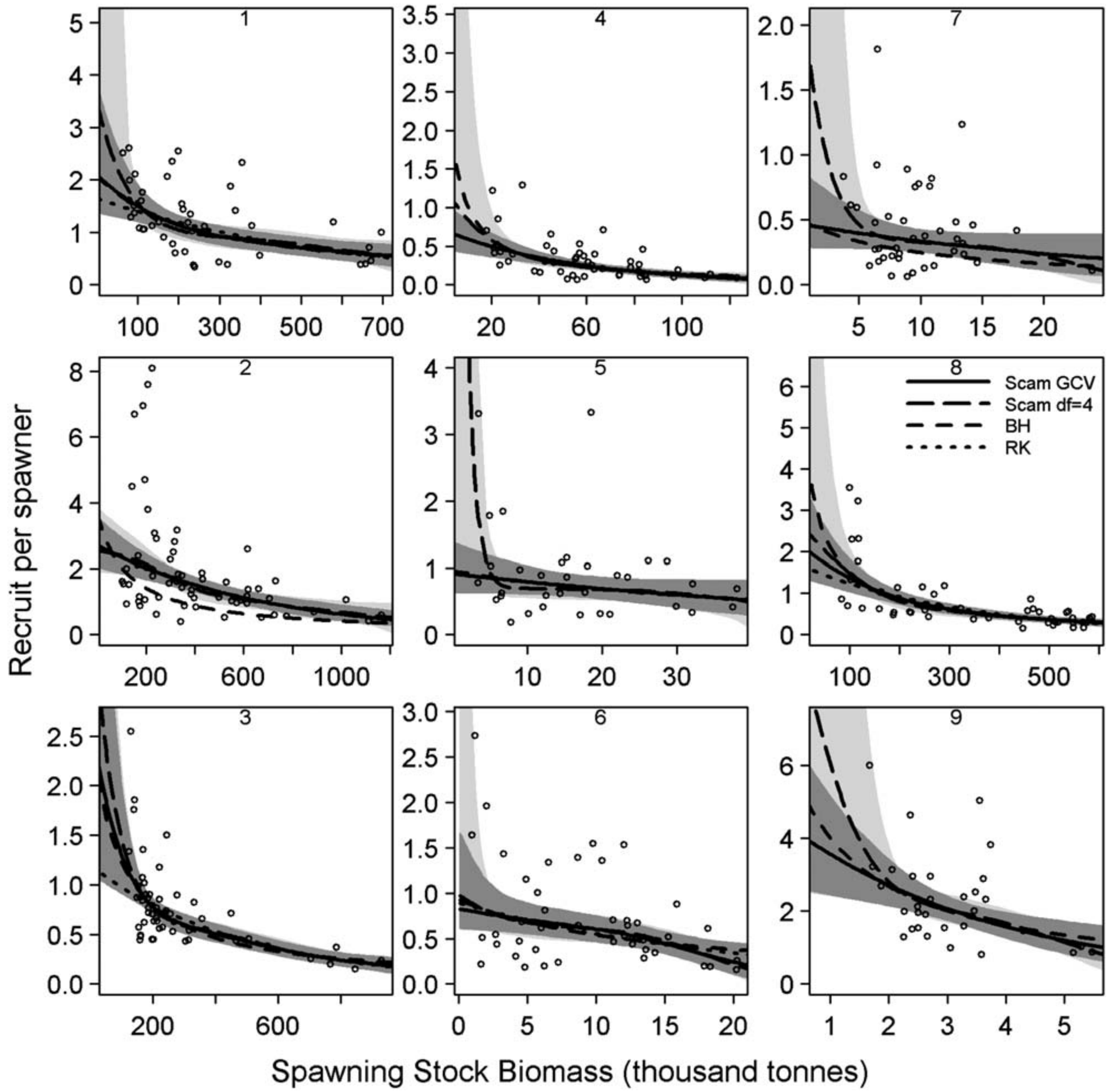

Figure 4. Model fits to recruit per spawner data (points) for nine stocks, with stock IDs listed at the top of each panel. See Table 1 for stock names and Table 2 for model names. Shaded regions indicate 95% bootstrapped confidence intervals based on the non-parametric *scam* model with df = 4 (lighter shading) or smoothing parameter selected to minimize GCV (darker shading). The biomass axis is in the natural log scale.

were ranked across models for each stock. F_{MSY} based on the BH model had the lowest average rank across stocks (1.8), followed by F_{MSY} values based on the *scam* df = 3 (3.1) *scam* GCV (3.4), *scam* df = 4 (3.4), RK (4.1), and HS (5.2) models. For six of nine stocks, the BH-based F_{MSY} was lowest among the five methods (Figure 7). The average ranks for B_{MSY} were: HS (1.6), RK (3.1), *scam* GCV (3.3), *scam* df = 3 (3.9), *scam* df = 4 (3.9), and BH (5.2). The HS-based B_{MSY} was the lowest for seven stocks and the BH-based B_{MSY} was highest for six stocks.

Bootstrapped *CI*s for B_{MSY} were usually right-skewed. *CI* widths were averaged across stocks for each method. F_{MSY} *CI*s based on the HS model were the most narrow (0.05), followed by those based on the BH (0.18), RK (0.23), *scam* GCV (0.33), *scam* df = 3 (0.97), and *scam* df = 4 (1.88). Similarly, B_{MSY} *CI* widths were: HS (89), RK (200), *scam* GCV (349), BH (495), *scam* df = 3 (1356), and *scam* df = 4 (1623). The *scam* df = 4 confidence limits for F_{MSY}, particularly the upper limit, were usually much greater than those of the other methods. Bravington *et al.* (2000) reported a similar result for the *F*crash RP, in which upper confidence limits were poorly defined using a non-parametric SR model. The range, over stocks, of the ratio of the largest to the smallest values of estimates of F_{MSY} for each of the 6 SR models was 1.4–2.3. This range was 1.6–4.6 for B_{MSY}.

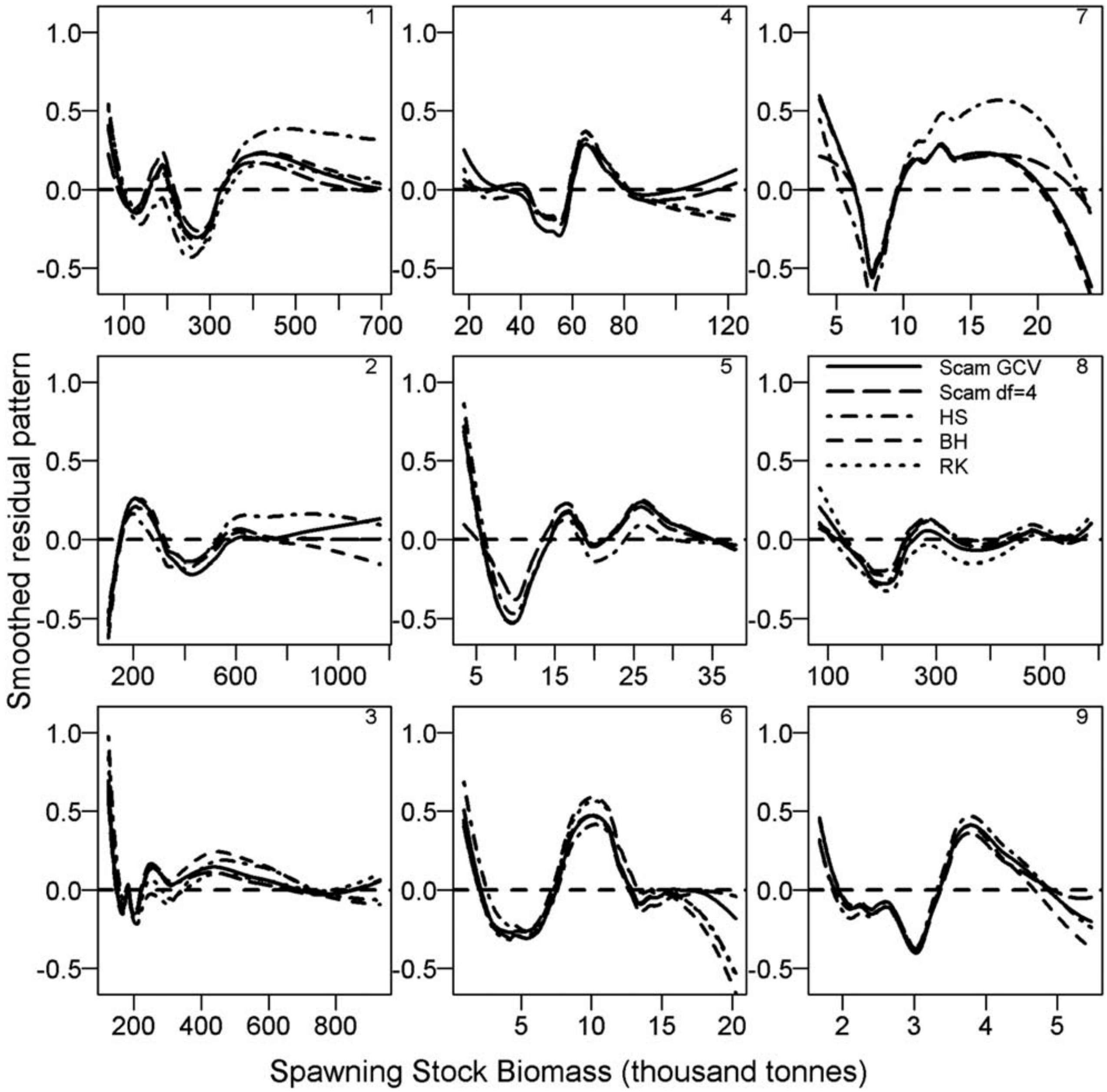

Figure 5. Smoothed residual patterns. Stock IDs are listed in the top corner of each panel. See Table 1 for stock names and Table 2 for model names.

The RK and *scam* GCV models were identical for Celtic Sea cod (stock 7), but the bootstrap results were not. This is because of the different constraints used to estimate these models. For the RK model, $Rmax <$ maximum observed R. This constraint was not used for fitting the *scam* SR curves, but it was used when deriving MSY RPs. If the *scam*-predicted recruitment exceeded the maximum observed R, then it was set equal to this maximum. Such constraints are fairly arbitrary and indicate the sensitivity of statistical inferences when the SR data are noisy.

Discussion

A non-parametric SR model that preserves the compensatory mortality property (e.g. Quinn and Deriso, 1999) was demonstrated in a simulated example to provide a more accurate estimation of the SR relationship and equilibrium yield curve when there are good SR data. The non-parametric approach performed fairly well in terms of simulation bias and mean squared error for RP estimates compared with the correct parametric model, and the approach performed better compared with RP estimates based on incorrect parametric models. The performance of the non-parametric model with nine real datasets was equivocal. It performed about the same in terms of prediction error (i.e. GCV statistic) compared with common SR models (i.e. BH, RK, and HS). Nonetheless, it is a useful method to evaluate the adequacy of proposed parametric models.

The non-parametric model was implemented using the *scam* function in R (Pya, 2012), and the main advantage of the *scam* package compared with Bravington *et al.* (2000) is ease of implementation. Anyone familiar with R should be able to implement

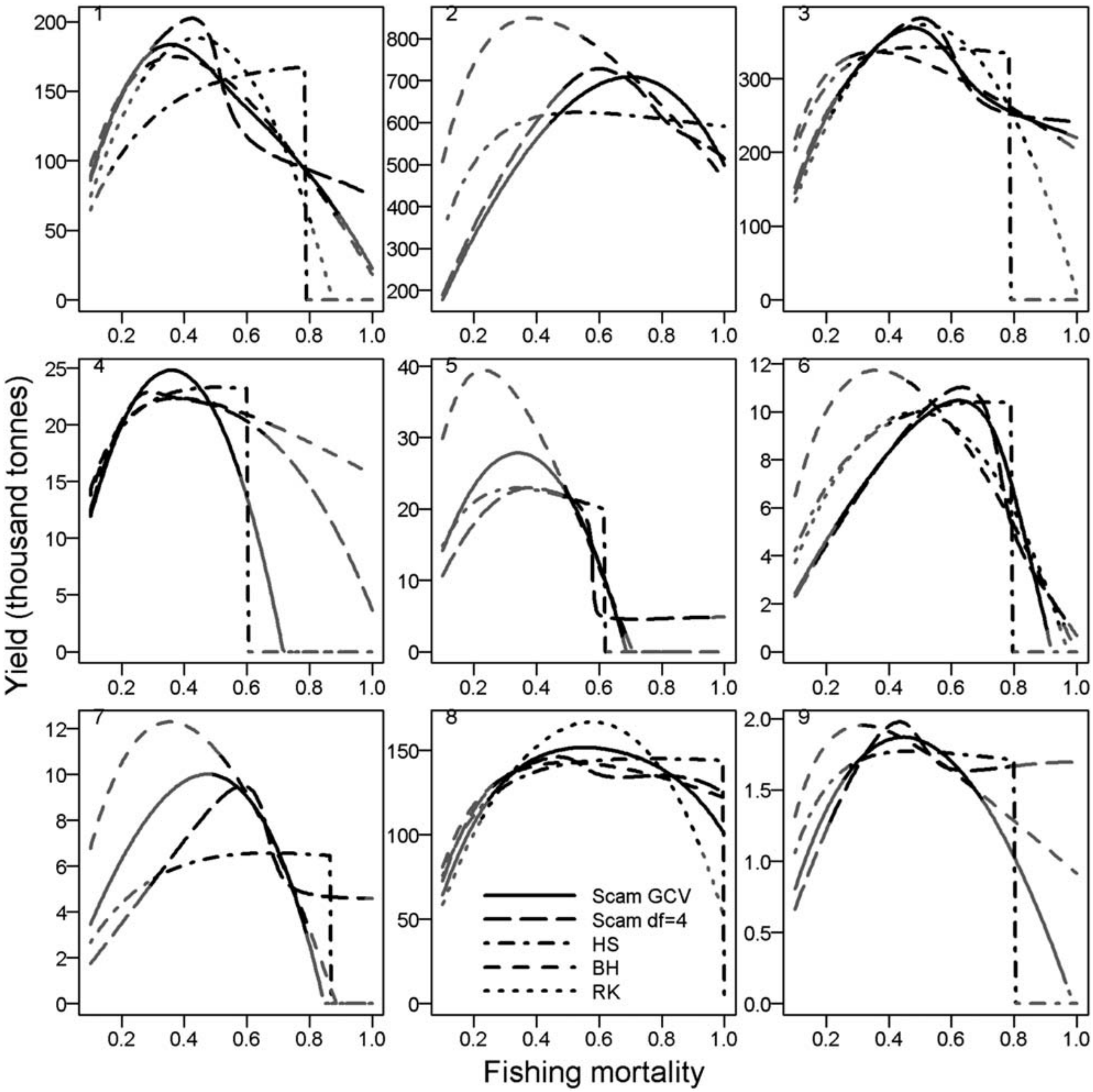

Figure 6. Equilibrium yield curves for nine stocks, with stock IDs listed in the top corner of each panel. See Table 1 for stock names and Table 2 for model names. Grey lines indicate yields inferred from extrapolations of recruitment outside of the range of the stock–recruit data.

the method in a typical stock assessment working group meeting. Smoothing approaches such as LOWESS are also useful for exploratory modelling of SR data. They can help orientate the eye to patterns in the data. However, such unrestricted smoothers are less useful for deriving RPs, and in particular MSY RPs. These smoothers can also result in unrealistic SR relationships caused by overfitting the substantial noise that is often present in SR data time-series. This is less of a problem with the shape-constrained non-parametric approach outlined herein.

Another potential advantage of the non-parametric SR model is more "honest" *CI*s that do not depend on parametric model assumptions that are very difficult to verify in practice. A bootstrapped uncertainty analysis was used to propagate the uncertainty about the SR relationship into *CI*s about MSY RPs. In the simulated example, the bootstrapped *CI*s based on the non-parametric SR model with GCV smoothing parameter covered the true values for F_{MSY}, whereas the *CI*s based on the BH, RK, and HS models did not. In the real data examples, the *CI*s for F_{MSY} based on the BH, RK, and HS SR models usually barely overlapped or did not overlap at all. This demonstrates that *CI*s for F_{MSY} derived from parametric SR models are sensitive to the parametric assumptions. The overlap for B_{MSY} was better. C*I*s based on the non-parametric model were usually wider than the parametric intervals, and much wider for some stocks, when the model was made flexible by fixing the df to be 4. Even when the df was fixed at 3, *CI*s for F_{MSY} were substantially wider for five of nine

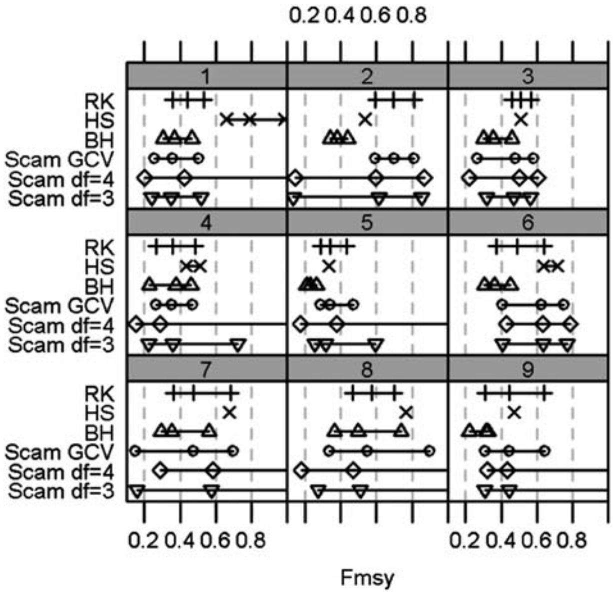

Figure 7. F_{MSY} with 95% bootstrap confidence intervals. Stock IDs are listed in the top strip of each panel. See Table 1 for stock names and Table 2 for model names. *Scam* upper limits were truncated at one in some panels.

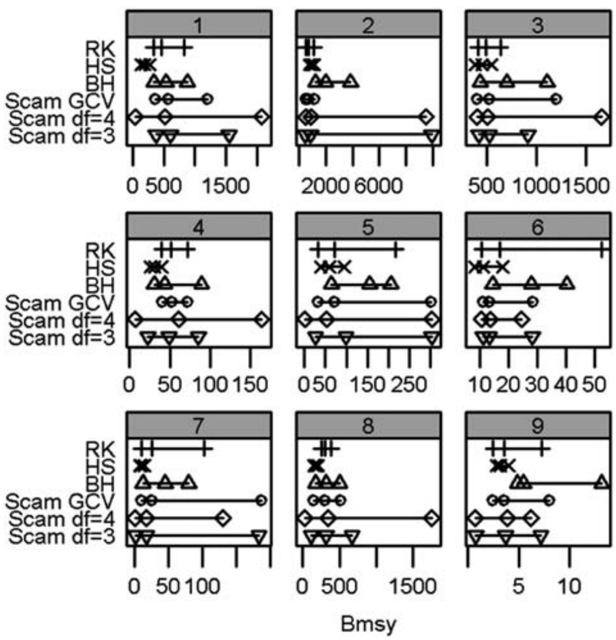

Figure 8. B_{MSY} with 95% bootstrap confidence intervals. Stock IDs are listed in the top strip of each panel. See Table 1 for stock names and Table 2 for model names.

stocks. As a result, these intervals usually overlapped the parametric SR model-based *CI*s. The non-parametric intervals reflected much of the uncertainty due to SR model choice, which is important information in practice. Wide *CI*s for MSY RPs reflect some of the substantial variability in stock productivity processes related to unaccounted multispecies interactions and impacts of environmental variability on recruitment (Overholtz, 1999) and others aspects of productivity, as well as stock assessment error.

A surprising result was the consistency in the smoothed pattern of the residuals obtained from fitting the various SR models. This pattern was similar in eight of nine real data examples. It was not present in the simulated example. The residual pattern suggests a steeper decline in recruits per spawner at low stocks sizes than the parametric or non-parametric SR models could accommodate. The residual pattern was worse for the RK and HS models, but still present for the BH and non-parametric models. It is possible to generate this basic pattern in simulated data by increasing the amount of recruitment immigration; however, this is not the only mechanism. For example, Cadigan (2009) found that measurement error in S tended to increase the *Sao*, and this is another possible mechanism for the residual pattern. Another possible cause is unaccounted ageing error which can smear out and reduce recruitment variability in the stock assessment, and make it seem possible that low stock sizes can produce large recruitments (i.e. higher slope at the origin).

The HS-based RPs were usually much more precise than the other methods, but they seem risky in that the HS F_{MSY} was usually greater than F_{MSY} values derived using other SR models, and B_{MSY} was usually much lower. This suggests the potential of incorrectly concluding the stock is not being overfished using HS-based RPs.

In the examples, it was necessary to impose constraints when estimating and bootstrapping the parametric SR curves. It was also necessary to impose constraints on the non-parametric estimates of SR curves, and in particular the extrapolations beyond the range of the data, when deriving MSY RPs. Such constraints complicate statistical inferences which will not be accurate if the constraints are wrong. MSY RPs will also contain additional uncertainty related to how other productivity processes will evolve in the future, but this was beyond the scope of this paper.

The variations in MSY RPs, as a function of the choice of SR model, estimated for most of the stocks examined in this paper would probably have substantial management implications if management was based only on stock status relative to point estimates of RPs. F_{MSY} was less sensitive overall to the choice of SR model; hence, evaluating stocks status in terms of F (i.e. $F_{current}/F_{MSY}$) seems to be less sensitive to the choice of SR model than evaluating status in terms of biomass (i.e. $B_{current}/B_{MSY}$). This suggests that F_{MSY} is a more reliable RP than B_{MSY}, which is a conclusion others have reached (e.g. ICES, 2011).

The lower confidence limits for F_{MSY} derived from the non-parametric SR models in some examples (i.e. stocks 1–3, 6–8) were much lower than the point estimates. Similarly, the upper limits for B_{MSY} were often much greater than the point estimates. The latter was also true for the RK and BH models in some examples (i.e. stocks 5–7, and 9). This suggests that there can be considerable uncertainty when evaluating the probability that current $F > F_{MSY}$, and current $B < B_{MSY}$. This will be caused by uncertainty in the values of the RPs and also uncertainty in the values of current F and B. Hence, a management framework that requires a low probability that current $F > F_{MSY}$ and current $B < B_{MSY}$ may provide for little fishery yield when there is substantial uncertainty about the values of F_{MSY} and B_{MSY}.

Supplementary material

Supplementary material is available at the *ICESJMS* online version of the paper and shows an example code for *scam*.

Acknowledgements

The work received financial support from an NSERC grant to N.C. I am grateful to the anonymous reviewers and the Editor for providing some stimulating comments which greatly improved the manuscript.

References

Beverton, R. J. H., and Holt, S. J. 1957. On the Dynamics of Exploited Fish Populations. Chapman and Hall, London. 533 pp.

Barrowman, N., and Myers, R. A. 2000. Still more spawner–recruit curves: the hockey stick and its generalizations. Canadian Journal of Fisheries and Aquatic Sciences, 57: 665–676.

Bousquet, N., Duchesne, T., and Rivest, L.-P. 2008. Redefining the maximum sustainable yield for the Schaefer population model including multiplicative environmental noise. Journal of Theoretical Biology, 254: 65–75.

Bravington, M. V., Stokes, T. K., and O'Brien, C. M. 2000. Sustainable fishing: the bottom line. Marine and Freshwater Research, 51: 465–475.

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.

Cadigan, N. G. 2011. Impact of stock–recruit and natural mortality process errors on MSY reference points. DFO Canadian Science Advisary Secretariat Research Document 2012/075. iii+20 pp.

Cook, R. M. 1998. A sustainability criterion for the exploitation of North Sea cod. ICES Journal of Marine Science, 55: 1061–1070.

Cleveland, W. S. 1981. LOWESS: a program for smoothing scatterplots by robust locally weighted regression. The American Statistician, 35: 54.

Duplisea, D., and Fréchet, A. 2010. Precautionary reference point estimates for northern Gulf of St. Lawrence (3Pn4RS) cod (*Gadus morhua*) and methods for their calculation. DFO Canadian Science Advisary Secretariat Research Document 2009/097. iv+24 pp.

Duplisea, D., and Fréchet, A. 2011. Updated reference point estimates for northern Gulf of St. Lawrence (3Pn4RS) cod (*Gadus morhua*) based on revised beginning of year weights at age. DFO Canadian Science Advisary Secretariat Research Document 2011/003 iv+8 pp.

Efron, B., and Tibshirani, R. J. 1993. An Introduction to the Bootstrap. Chapman-Hall, New York. 456 pp.

Evans, G. T., and Rice, J. C. 1988. Predicting recruitment from stock size without the mediation of a functional relation. Journal du Conseil International pour l'Exploration de la Mer, 44: 111–122.

Fronczyk, K., Kottas, A., and Munch, S. 2011. Flexible modeling for stock–recruitment relationships using Bayesian nonparametric mixtures. Environmental and Ecological Statistics, 19: 183–204.

Hilborn, R., and Stokes, K. 2010. Defining overfished stocks: have we lost the plot? Fisheries, 35: 113–120.

Hutchings, J. A., and Rangeley, R. W. 2011. Correlates of recovery for Canadian Atlantic cod (*Gadus morhua*). Canadian Journal of Zoology, 89: 386–400.

ICES. 2011. Report of the ICES Advisory Committee, 2011. ICES Advice, 2011. Book 1, 226 pp.

Jiao, Y., Schneider, D., Chen, Y., and Wroblewski, J. 2004. An analysis of error structure in modeling the stock–recruitment data of gadoid stocks using generalized linear models. Canadian Journal of Fisheries and Aquatic Sciences, 61: 134–146.

Legović, T., Klanjčšek, J., and Geček, S. 2010. Maximum sustainable yield and species extinction in ecosystems. Ecological Modeling, 221: 1569–1574.

Mesnil, B., and Rochet, M.-J. 2010. A continuous hockey stick stock–recruit model for estimating MSY reference points. ICES Journal of Marine Science, 67: 1780–1784.

Michielsens, C. G. J., and McAllister, M. K. 2004. A Bayesian hierarchical analysis of stock–recruitment data: quantifying structural and parameter uncertainties. Canadian Journal of Fisheries and Aquatic Sciences, 61: 1032–1047.

Munch, S. B., Kottas, A., and Mangel, M. 2005. Bayesian nonparametric analysis of stock–recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences, 62: 1808–1821.

Myers, R., Rosenberg, A., Mace, P., Barrowman, N., and Restrepo, V. 1994. In search of thresholds for recruitment overfishing. ICES Journal of Marine Science, 51: 191–205.

Needle, C. L. 2002. Recruitment models: diagnosis and prognosis. Reviews in Fish Biology and Fisheries, 11: 95–111.

Overholtz, W. J. 1999. Precision and uses of biological reference points calculated from stock recruitment data. North American Journal of Fisheries Management, 19: 643–657.

Pya, N. 2010. Additive models with shape constraints. PhD thesis. University of Bath, Department of Mathematical Sciences. 195 pp.

Pya, N. 2012. scam: shape constrained additive models. R package version 1.1-1. http://CRAN.R-project.org/package=scam.

Quinn, T. J., and Deriso, R. B., 1999. Quantitative Fish Dynamics. Oxford University Press, New York. 560 pp.

R Development Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Ricker, W. E. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada, 11: 559–623.

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.

Sissenwine, M. P., and Shepherd, J. G. 1987. An alternative perspective on recruitment overfishing and biological reference points. Canadian Journal of Fisheries and Aquatic Sciences, 44: 913–918.

Walters, C. J. 1990. A partial bias correction factor for stock–recruitment parameter estimation in the presence of auto-correlated environmental effects. Canadian Journal of Fisheries and Aquatic Sciences, 47: 516–519.

Wood, S. N. 2006. Generalized Additive Models. An Introduction with R. CRC Press, New York, 416 pp.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society B 73: 3–36.

Handling editor: Verena Trenkel