Model-based estimates of reference points in an age-based state-space stock assessment model

Albertsen, Christoffer Moesgaard; Trijoulet, Vanessa

Published in:
Fisheries Research

Link to article, DOI:
10.1016/j.fishres.2020.105618

Publication date:
2020

Document Version
Peer reviewed version

Citation (APA):
Albertsen, C. M., & Trijoulet, V. (2020). Model-based estimates of reference points in an age-based state-space stock assessment model. *Fisheries Research*, *230*, Article 105618. https://doi.org/10.1016/j.fishres.2020.105618

Model-based estimates of reference points in an age-based state-space stock assessment model

Christoffer Moesgaard Albertsen[a,*], Vanessa Trijoulet[a]

[a]*National Institute of Aquatic Resources, Technical University of Denmark, Kemitorvet 201, DK-2800 Kgs. Lyngby, Denmark*

Abstract

Reference points are central in the current management of marine living resources. However, reference points are estimated from data and model estimates. Therefore, they are inherently uncertain. We present two objective methods for estimating reference points and quantifying their uncertainty. The first method uses per-recruit calculations, while the second method relies on a long-term forecast of the managed system. Confidence intervals are calculated through a combination of the implicit function theorem and the delta method. Both methods are illustrated for 12 recruitment models using data from the Northeast Arctic cod assessment. Finally, the methods are validated in a simulation study.

Keywords: MSY, per-recruit, recruitment models, reference points, stock assessment

---

*Corresponding author

Email addresses: cmoe@aqua.dtu.dk (Christoffer Moesgaard Albertsen), vtri@aqua.dtu.dk (Vanessa Trijoulet)

Postprint accepted for publication in Fisheries Research. DOI: 10.1016/j.fishres.2020.105618

1. Introduction

Fisheries management aims to find an optimal trade-off between ecological conservation, economic yield, and social considerations for marine living resources. To this end, biological and economic reference points provide invaluable guidance for management in evaluating the current status of a stock and the possibility of exploiting it. However, since reference points are derived from data and assessment model estimates, they are subject to variability in data that for instance may be updated every year, and uncertainty in model estimates and observations. This must be kept in mind when using reference points for management and conservation.

Reference points can be based on fishing mortality and/or biomass. Some are calculated as optimum targets. For instance, in the US, biomass corresponding to maximum sustainable yield B_{MSY} is used as a biomass target (National Research Council, 2014). In Australia, biomass and fishing mortality corresponding to maximum economic yield (B_{MEY} and F_{MEY} respectively) are used as target reference points (Australian Government, 2018). Given the uncertainty in reference point estimates and the adoption of the precautionary approach to fisheries management (Garcia, 1994), reference points are also used in scientific advice as limits not to exceed (e.g., F_{MSY} in the US and in the International Council for the Exploration of the Sea, ICES) or not to fall below (e.g., the biomass limit reference point, B_{lim}) (Caddy and Mahon, 1995). For data rich stocks, reference points are often based on stock-recruitment considerations where the spawning component of the stock is conserved to allow for the stock to retain its reproductive capacity. These reference points necessitate estimating a stock-recruitment relationship. However, this is not always possible, notably for data poor stocks, and in these cases, reference points based on simple levels of fishing mortality or biomass (e.g., percent of unfished spawning biomass-per-recruit, $F_{x\%}$) and on per-recruit (PR) analyses (e.g., fishing mortality at maximum yield-per-recruit, F_{max}) are often used. These reference points do not account for the possible correlation between recruitment and spawning stock size, notably the fact that recruitment decreases at low stock size. This can sometimes result in fishing mortality reference points that are unsustainable because they are above F_{Crash} or not providing an optimal economic or biomass yield for the fishery because they differ from F_{MEY} or F_{MSY}, respectively.

In some assessment models, such as surplus production models, certain reference points can be calculated directly from, or correspond to, model parameters (eg., Pedersen and Berg, 2017). In these cases, evaluating their uncertainty follows directly from the maximum likelihood estimator of model parameters. However, in more complex age- or length-based statistical assessment models, this is not the case. In these models, reference points are calculated through optimization of derived quantities such as catch, not as an explicit expression from model parameters. Therefore, additional steps are needed to obtain reference point estimates and their uncertainty.

The SAM assessment model is an age-based state-space single stock assessment model (Nielsen and Berg, 2014). The model was used as the basis of the assessment for more than 20 stocks in 2019 assessed by ICES (ICES, 2020). The model assumes an exponential decay of cohorts and links the population process to catch observations via the Baranov catch equation and to survey indices via an assumption of proportionality. The SAM model currently allows for three types of recruitment: random walk on log-scale, Ricker, and Beverton-Holt stock-recruitment functions. Recently, SAM has been extended to consider several observational models (Berg and Nielsen, 2016; Albertsen et al., 2017) and to model several stocks (Albertsen et al., 2018). In the SAM model, reference points cannot be expressed explicitly from model parameters.

Advice given by ICES for data rich stocks (category 1 and 2 stocks) using for instance, the SAM model, involves a post hoc analysis based on visual inspection of the spawner-recruit relationship and on a long-term forecast simulation is used to obtain reference points (ICES, 2017). First, the biomass reference point, B_{lim}, is either obtained by fitting a segmented regression (hockey stick) to the SSB and recruitment pairs or by a subjective choice following ICES guidelines for estimation of B_{lim} (ICES, 2017). Then, a simulation-based approach is used to estimate F_{MSY} based on a weighted average of Ricker, Beverton-Holt and hockey stick stock-recruitment curves. In addition to estimates of spawning stock biomass, number of recruits, fishing selectivity, weight at age and natural mortality used as input, this approach is independent of the assessment model and may involve different assumptions about the system being managed.

In contrast, several North American assessment models calculate reference points as part of the model output. The Woods Hole Assessment Model (WHAM, version 0.0.0.9000; Miller and Stock, 2019) calculates F_{MSY} (Miller et al., 2016a) and $F_{x\%}$ (Miller et al., 2018) reference points based on PR calculations. Using 10 Newton steps, the yield and biomass-per-recruit are optimized as part of the model implementation. Through the TMB R-package (Kristensen et al., 2016), derivatives of the 10 step Newton optimization are obtained via automatic differentiation and, in turn, the variance of the estimates are provided. Likewise, Stock Synthesis 3 (SS3, version 3.30.10; Methot and Wetzel, 2013) implemented in AD Model Builder (Fournier et al., 2012) calculates $F_{0.1}$, $F_{x\%}$, F_{MSY}, and F corresponding to a biomass target as part of the model implementation. For the optimization, a grid search is used followed by a fixed number of Newton iterations. Again, automatic differentiation provides derivatives of the Newton optimization along with variance estimates.

While the Newton optimization with a fixed number of steps in, for instance, TMB provides the correct result, the number of steps must be chosen carefully. A sufficiently high number of iterations must be chosen to ensure convergence. However, more Newton steps will increase the computational complexity. The optimal choice will depend not only on the model, but also on the input data and model parameters.

We introduce a framework for estimating reference points and their uncertainty that allows for the use of an adaptive optimization algorithm, both in the number of steps and the step sizes, by adding the optimization criteria to the likelihood function. Reference point uncertainties are calculated through a combination of the implicit function theorem and the delta method. Two methods were considered for calculating reference point optimization criteria. The first method uses PR calculations to find equilibrium yield, biomass, and recruitment, while the second method uses a long-term forecast of the system for the same purpose.

For this study, the framework was implemented in the SAM model along with nine new recruitment models. The methods for estimating reference points are illustrated using data from the 2019 assessment of Northeast Arctic cod (ICES, 2019). To exemplify general applicability, reference points are estimated for each of the 12 recruitment models and validated through a simulation study. The simulation study investigates the properties of the reference point estimators by generating new data from the model. Thereby, variability in reference points arises from variability in observations and, in turn, variability in parameter estimates. Through the simulations, we show that the method provides accurate, unbiased estimates of reference points with correct confidence intervals under the true model. The method is not limited to the SAM nor state-space models. Further, the method is not limited to the reference points considered here, but can be applied for any derived quantity that requires an optimization.

2. Methods

Reference point estimation was implemented in the age-based state-space stock assessment model SAM (we refer to Nielsen and Berg, 2014; Berg and Nielsen, 2016; Albertsen et al., 2017; and Albertsen et al., 2018, for full details) using the R-package Template Model Builder (Kristensen et al., 2016). The source codes can be found at `https://github.com/fishfollower/SAM/tree/reference_points` (commit b56f2ec). The original SAM model was extended to include nine new recruitment options. Further, two methods used to calculate reference points were implemented. The first method was based on a PR analysis, while the second was based on a long-term forecast of the assessment model. The two methods were compared in a case study using data from the 2019 assessment of Northeast Arctic cod and validated through simulations.

2.1. Per-recruit and related concepts

To introduce the notation used below, we briefly reiterate the PR calculations needed for the reference point estimation (see e.g., Laurec and Le Guen, 1981; Sissenwine and Shepherd, 1987; Mace, 1994; Mesnil and Rochet, 2010, for further details).

Starting from a cohort containing a single recruit, $N_0 = 1$, the cohort size is projected by the population model used for the assessment: $N_{a+1} = \exp\left(-\tilde{F} \cdot s_a - M_a\right) N_a$, where $\tilde{F}$ is the fully-selected fishing mortality, s_a and M_a are the fishing selectivity and natural mortality at age a, respectively. For simplicity, we let $a = 0$ denote the age of recruitment to the fishery; since a is only used as an index, this does not influence the calculations. Now, the spawning biomass-per-recruit (SPR) is calculated by

$$S_R(\tilde{F}) = \sum_{a=0}^{\infty} w_a^{(s)} \cdot p_a \cdot N_a,$$

where p_a is the proportion mature at cohort age a after recruitment and $w_a^{(s)}$ is the average weight at age in the stock. While we consider spawning biomass, spawners-per-recruit could be used by letting $w_a^{(s)} = 1$. Likewise, yield-per-recruit (YPR) can be calculated by

$$Y_R(\tilde{F}) = \sum_{a=0}^{\infty} w_a^{(c)} \cdot \frac{\tilde{F} \cdot s_a}{\tilde{F} \cdot s_a + M_a} \left(1 - \exp\left(-\tilde{F} \cdot s_a - M_a\right)\right) N_a,$$

where $w_a^{(c)}$ is the weight at age a in the catch. In the analyses below, weights were used to scale the abundance at age in the two equations above. However, in general, other scalings such as relative fecundity can be used if relevant. For simplicity, we do not distinguish between catch and landings. However, if catch and landings differ, a landing fraction may be multiplied to the summands and $w_a^{(c)}$ can be changed to landing weights. Further, we only consider a single fishing fleet. However, the equation for Y_R may be generalized to account for several fleets.

Based on the SPR, equilibrium spawning biomass, S_e, is obtained by solving the equation:

$$S_e(\tilde{F}) = S_R(\tilde{F}) \cdot R(S_e(\tilde{F})); \tag{1}$$

that is, the spawning biomass that will produce the same biomass. Finally, equilibrium recruitment,

$$R_e(\tilde{F}) = R(S_e(\tilde{F})) = \frac{S_e(\tilde{F})}{S_R(\tilde{F})},$$

and yield,

$$Y_e(\tilde{F}) = Y_R(\tilde{F}) \cdot R_e(\tilde{F}) = \frac{Y_R(\tilde{F}) \cdot S_e(\tilde{F})}{S_R(\tilde{F})},$$

can be calculated. While SPR and YPR can be calculated without assuming a stock-recruitment relationship, a functional relationship is needed to calculate S_e, R_e, and Y_e.

2.2. Recruitment models

In this study, 12 recruitment models were considered to cover a broad range of stock-recruitment shapes (Table 1). By default, SAM was built with three possible configurations for the recruitment model: (i) a random walk time series model on log-scale where recruitment in one year only depends on recruitment the year before, (ii) the recruitment model by Ricker (1954), and (iii) the recruitment model by Beverton and Holt (1957). In addition, to these three options, nine recruitment models were implemented. A power function (e.g., Cushing, 1971) and a bent hyperbola (Mesnil and Rochet, 2010) of the spawning-stock biomass were considered. The bent hyperbola relationship is a differentiable version of the hockey stick, or segmented regression, recruitment used for ICES fisheries management reference points (ICES, 2017). Moreover, the recruitment model proposed by Shepherd (1982) was implemented. Further, the generalization of the Ricker model proposed by Saila-Lorda (e.g., Iles, 1994; Needle, 2001) along with two generalizations of the Beverton-Holt were also added to the package: a sigmoidal version (Myers et al., 1995) and the version proposed by both Hassell (1975) and Deriso (1978). Most of these three parameter models can mimic both the Ricker and the Beverton-Holt recruitment functions. The latter is parameterized in a way similar to Schnute (1985). However, in common with Deriso (1978), we enforce stronger constraints on the parameters to ensure positive recruitment for positive spawning biomass. While Schnute (1985) allowed the shape parameter γ to be negative, we only consider positive values. Finally, an AR(1) time series model on log-scale and two spline models were implemented. Both spline models were implemented on the logarithm of the recruits-per-spawner (see Supplementary Material for details). Similar to Cadigan (2012), one spline was implemented to have a non-increasing recruitment rate (R/S) as S increases to ensure compensatory mortality, while the second spline had no restriction on the parameters.

For the 11 functional models (all except the random walk), either a closed form or numerical solution can be found for $S_e(\tilde{F})$ (Table 1). For the random walk, recruitment is not stationary. Therefore, an equilibrium spawning biomass, recruitment, or yield cannot be found. However, in a forecast the mean recruitment is constant; that is, $E(\log R_{T+n} \mid \log R_T, Y_{1:T}) = \log R_t$ for any $n > 0$ where $Y_{1:T}$ denotes the data available from time 1 to time T. Therefore, projected values will tend to a stable mean value with increasing variance. Assuming constant recruitment, these long-term projected values can be used as a proxy for the equilibrium quantities when calculating reference points.

For three of the models, equilibrium biomass is not always uniquely defined. For certain parameter values, the Saila-Lorda, sigmoidal Beverton-Holt, and the general spline model do not have compensatory mortality, where recruits-per-spawning-biomass is decreasing as a function of spawning biomass. As a result, the replacement line may cross the stock-recruitment curve twice. Consequently, there are two solutions to equation (1); a stable solution at high biomass and an unstable solution at low biomass. The numerical methods will aim to find the stable solution.

Table 1: Stock-recruitment models and the corresponding equilibrium biomass. In the formulas, S denotes the spawning stock biomass, $S_R(\tilde{F})$ is the spawners-per-recruit for fishing mortality $\tilde{F}$, and $\alpha, \beta, \gamma > 0$, $\delta \in (0,1)$, and $\theta \in \mathbb{R}$ are parameters. Boldface indicates vectors. For the random walk, recruitment in year y follows a Gaussian distribution where the mean is the recruitment from the year before. $W(x)$ denotes the Lambert W function which solves $x = W\exp(W)$ (e.g., Fritsch et al., 1973). In the spline models, I_Φ is an integrated spline basis of Gaussian functions, $\boldsymbol{\kappa}$ is a vector of knots, while CMP is short for "compensatory mortality property". Further details on the numerical solvers used can be found in the supplementary material.

Model	$R(S)$	$S_e(\tilde{F})$
Random walk	$R_y \sim N(R_{y-1}, \sigma_R^2)$	Not applicable
Ricker	$\alpha S \exp(-\beta S)$	$\frac{\log\left(\alpha S_R(\tilde{F})\right)}{\beta}$
Beverton-Holt	$\frac{\alpha S}{1+\beta S}$	$\frac{\alpha S_R(\tilde{F})-1}{\beta}$
AR(1)	$R_y \sim N\left(\alpha + \left(\frac{2}{1+\exp(-\theta)} - 1\right) \cdot (R_{y-1} - \alpha), \sigma_R^2\right)$	$S_R(\tilde{F})\alpha$
Bent hyperbola	$\beta\left(S + \sqrt{\alpha^2+\gamma^2/4} - \sqrt{(S-\alpha)^2+\gamma^2/4}\right)$	$\frac{2\sqrt{\alpha^2+\gamma^2/4}-2\alpha-2\sqrt{\alpha^2+\gamma^2/4}}{(\beta S_R(\tilde{F}))^{-2}-2(\beta S_R(\tilde{F}))^{-1}}$
Power-law	αS^δ	$\left(\alpha S_R(\tilde{F})\right)^{\frac{1}{1-\delta}}$
Shepherd	$\frac{\alpha S}{\left(\frac{S}{\beta}\right)^\gamma+1}$	$\beta\left(\alpha S_R(\tilde{F}) - 1\right)^{\frac{1}{\gamma}}$
Hassell/Deriso	$\frac{\alpha S}{(1+\gamma\beta S)^\gamma}$	$\frac{\left(\alpha S_R(\tilde{F})\right)^{\frac{1}{\gamma}}-1}{\gamma\beta}$
Saila-Lorda	$\alpha S^\gamma \exp(-\beta S)$	$\begin{cases} \frac{1-\gamma}{\beta} W\left(\frac{\beta}{1-\gamma}\left(\alpha S_R(\tilde{F})\right)^{\frac{1}{1-\gamma}}\right), & \gamma < 1 \\ \text{Solved numerically}, & \gamma \geq 1 \end{cases}$
Sigmoidal Beverton-Holt	$\frac{\alpha S^\gamma}{1+\beta S^\gamma}$	Solved numerically
CMP Spline	$S \cdot I_\Phi(S; \boldsymbol{\kappa}, \boldsymbol{\alpha}, \theta)$	Solved numerically
General Spline	$S \cdot I_\Phi(S; \boldsymbol{\kappa}, \boldsymbol{\theta})$	Solved numerically

2.3. Reference points

To illustrate the two estimation methods, six fishing mortality reference points (Table 2) were considered and compared to the fishing mortality in the last year of the assessment ($F_{\text{Status quo}}$). Of the six reference points, three depend on the recruitment function (F_{MSY}, F_{Crash}, F_{lim}) while three can be derived from PR calculations without assuming a specific stock-recruitment relationship ($F_{0.1}$, $F_{\text{x\%}}$, F_{max}).

For the PR method, fishing mortality reference points at which (i) the equilibrium yield is maximized (F_{MSY}), (ii) the YPR curve is maximized (F_{max}) (iii) the slope of the YPR curve is 10% of its slope at the origin ($F_{0.1}$), (iv) S_R is x% of the unfished spawning biomass-per-recruit ($F_{\text{x\%}}$), (v) the slope of the replacement line equals the slope of the recruitment function (F_{Crash}), and (vi) the reference point at which the spawning stock biomass equals the limit biomass (B_{lim}) where the stock is considered to have reduced reproductive capacity (F_{lim}). The F_{lim} reference point is only considered for the bent hyperbola, where B_{lim} is included as a parameter.

For the PR method, each of the six reference points can be defined mathematically through a criterion, $\kappa(\theta, \lambda, \tilde{F})$, to optimize (Table 2). In general, the criterion will be a function of model parameters, θ, latent variables, λ, and the fishing mortality to optimize over, $\tilde{F}$. For each of the fishing mortality reference points, corresponding equilibrium SPR, YPR, biomass, and recruitment can be obtained through the PR calculations outlined above. Here, we focus on comparing estimated fishing mortality reference points and the corresponding equilibrium biomass.

For the forecast-based method, only F_{MSY} was considered. However, the method can be applied to any of the reference points considered for the PR method (see subsection 2.3.2). For this method, F_{MSY} is the fishing mortality that optimizes the long-term yield. Likewise, F_{max}, for instance, could be obtained as the fishing mortality that maximizes the long-term yield divided by the long-term recruitment. This would, however, require an additional forecast of the system for each additional reference point considered. Again, the optimization criterion, $\kappa(\theta, \lambda, \tilde{F})$, is a function of parameters, latent variables, and a fishing mortality. However, the value is calculated from the long-term projections. Likewise, proxies for equilibrium SPR, YPR, biomass, and recruitment can be obtained from the state of the system in the last year of the long-term forecast.

Table 2: Minimization criteria for reference points estimated through a per-recruit calculation. $\mathrm{F}_{\mathrm{lim}}$ is only estimated for the Bent Hyperbola model where B_{lim} corresponds to the model parameter α.

Reference point	Criterion $\kappa(\theta, \lambda, \tilde{F})$
$\mathrm{F}_{\text{Status quo}}$	Not applicable
$\mathrm{F}_{\mathrm{MSY}}$	$-\log Y_e(\theta, \lambda, \tilde{F})$
$\mathrm{F}_{\max}$	$-\log Y_R(\theta, \lambda, \tilde{F})$
$\mathrm{F}_{0.1}$	$\left(0.1 \frac{\partial}{\partial f} Y_R(\theta, \lambda, f)\Big\vert_{f=0} - \frac{\partial}{\partial f} Y_R(\theta, \lambda, f)\Big\vert_{f=\tilde{F}}\right)^2$
$\mathrm{F}_{\mathrm{x\%}}$	$\left(\frac{x}{100} S_R(\theta, \lambda, 0) - S_R(\theta, \lambda, \tilde{F})\right)^2$
$\mathrm{F}_{\mathrm{lim}}$	$\left(\log B_{\mathrm{lim}} - \log S_e(\theta, \lambda, \tilde{F})\right)^2$
$\mathrm{F}_{\mathrm{Crash}}$	$\left(\frac{\partial}{\partial f} R(\theta, \lambda, f)\Big\vert_{f=0} - \frac{1}{S_R(\theta,\lambda,\tilde{F})}\right)^2$

2.3.1. Estimation through PR analysis

To estimate reference points using the PR method, the joint negative log-likelihood of data and latent variables, $\ell(\theta, \lambda)$, is augmented by the relevant criterion:

$$\ell(\theta, \lambda, \tilde{F}) = \ell(\theta, \lambda) + \kappa(\theta, \lambda, \tilde{F}).$$

Then, reference point estimates are obtained as

$$\hat{F}(\hat{\theta}, \hat{\lambda}_{\hat{\theta}}) = \mathrm{argmin}_{\tilde{F}}\, \ell(\hat{\theta}, \hat{\lambda}_{\hat{\theta}}, \tilde{F}) = \mathrm{argmin}_{\tilde{F}}\, \kappa(\hat{\theta}, \hat{\lambda}_{\hat{\theta}}, \tilde{F}),$$

where $\hat{\theta}$ is a vector of maximum likelihood estimates of model parameters, while $\hat{\lambda}_{\hat{\theta}}$ are the predicted latent variables given data and the estimates. Note that multiple criteria can be added to the augmented likelihood to estimate multiple reference points at once.

2.3.2. Estimation through forecast

In the forecast method, a reference point is estimated using an iterative procedure. For a given $\tilde{F}$, the stock process is forecasted for n years to reach the equilibrium state of the system, thereby obtaining catch, recruitment, and spawning-stock-biomass as proxies for $Y_e(\tilde{F})$, $R_e(\tilde{F})$, and $S_e(\tilde{F})$, respectively. Technically, the random processes, $\log F$ and $\log N$ are elonged by n time steps such that $E(\log F_t) = \log \tilde{F}$ for all t after the data period. To propagate uncertainties from the $\log F$ process in the assessment period, $\tilde{F}$ is calculated as a parameter multiplied by F in the final assessment year. Given $\log F$, the $\log N$ process, with recruitment as the first age, is forecasted using the estimated model, including the process covariance. Then, the joint log-likelihood is optimized to predict the random processes $\log F$ and $\log N$, thereby giving $\hat{\lambda}_{\theta}$. Subsequently, $\kappa(\theta, \hat{\lambda}_{\theta}, \tilde{F})$ is calculated in the final year and optimized over $\tilde{F}$ to obtain the reference point estimate. Note that the method requires a forecast per reference point estimate. In the case study and simulations below, the processes were forecasted for $n = 100$ years.

To calculate the equilibrium catch proxy, an average over the final 20 years was used. For recruitment models where recruitment decreases with SSB after a certain point can exhibit oscillatory behaviour around the equilibrium values (e.g., Ricker, 1954). Forecasting the system until the oscillations are damped can be impractical. Therefore, averages were used to limit the necessary number of years in the forecast.

To re-use computations in TMB, in particular the automatic differentiation, the criterion can be added by augmenting the joint log-likelihood:

$$\ell(\theta, \lambda, \delta, \tilde{F}) = \ell(\theta, \lambda) + \delta \cdot \kappa(\theta, \lambda, \tilde{F})$$

Then,

$$\ell(\theta, \lambda, 0, \tilde{F}) = \ell(\theta, \lambda)$$

and

$$\frac{\partial}{\partial \delta} \ell(\theta, \lambda, \delta, \tilde{F}) \bigg|_{\delta=0} = \kappa(\theta, \lambda, \tilde{F}).$$

Hence, the program and automatic differentiation only needs to run with $\delta = 0$. Including the criterion in the augmented joint log-likelihood does not influence the result of parameter estimation or latent process prediction. However, the criterion is easily obtained by evaluating the gradient at the maximum likelihood estimates and predicted latent processes.

Similar to the PR method, the reference point estimate is obtained as

$$\begin{aligned} \hat{F}(\hat{\theta}, \hat{\lambda}_{\hat{\theta}}) &= \operatorname{argmin}_{\tilde{F}} \frac{\partial}{\partial \delta} \ell(\hat{\theta}, \hat{\lambda}_{\hat{\theta}}, \delta, \tilde{F}) \bigg|_{\delta=0} \\ &= \operatorname{argmin}_{\tilde{F}} \kappa(\hat{\theta}, \hat{\lambda}_{\hat{\theta}}, \tilde{F}). \end{aligned}$$

2.3.3. Calculating uncertainties

In both estimation methods, reference points are estimated on log-scale, since they must be positive. The reference point estimate is a deterministic function of the model parameters and processes for a given data set. To calculate the variance of a reference point estimator on log-scale, the following function could be defined:

$$\begin{aligned} g(\theta, \lambda, \log \hat{F}(\theta, \lambda)) &= \frac{\partial}{\partial \log \tilde{F}} \left(\ell(\theta, \lambda, \log \hat{\lambda}) + \kappa(\theta, \lambda, \exp(\log \tilde{F})) \right) \bigg|_{\log \tilde{F} = \log \hat{F}(\theta, \lambda)} \\ &= \frac{\partial}{\partial \log \tilde{F}} \kappa(\theta, \lambda, \exp(\log \tilde{F})) \bigg|_{\log \tilde{F} = \log \hat{F}(\theta, \lambda)}. \end{aligned}$$

Now, reference point estimates can be defined implicitly by

$$g(\theta, \lambda, \log \hat{F}(\theta, \lambda)) = 0.$$

Therefore, the implicit function theorem gives the derivative of the reference point estimate with respect to the model parameters and processes,

$$\frac{\partial}{\partial(\theta, \lambda)} \log \hat{F}(\theta, \lambda) = -J^{-1}_{g, \log \hat{F}(\theta, \lambda)} J_{g, (\theta, \lambda)},$$

where $J_{g,\log \hat{F}(\theta,\lambda)}$ is the Jacobian matrix of g with respect to $\log \hat{F}(\theta, \lambda)$ and, likewise, $J_{g,(\theta,\lambda)}$ is the Jacobian of g with respect to (θ, λ). Consequently, the variance of the reference point estimator can be approximated asymptotically by the delta method,

$$\mathrm{Var}\left(\log \hat{F}(\hat{\theta}, \hat{\lambda}_{\hat{\theta}})\right) = \left(\log \tilde{F}'(\hat{\theta}, \hat{\lambda}_{\hat{\theta}})\right)^T \Sigma_{\hat{\theta},\hat{\lambda}_{\hat{\theta}}} \log \tilde{F}'(\hat{\theta}, \hat{\lambda}_{\hat{\theta}}), \tag{2}$$

where $\log \tilde{F}'(\hat{\theta}, \hat{\lambda}_{\hat{\theta}}) = \frac{\partial}{\partial(\theta,\lambda)} \log \hat{F}(\theta, \lambda)\Big|_{\theta=\hat{\theta},\lambda=\hat{\lambda}_{\hat{\theta}}}$ and $\Sigma_{\hat{\theta},\hat{\lambda}_{\hat{\theta}}}$ denotes the joint covariance of parameters and latent processes. Finally, confidence intervals can be calculated on log-scale and transformed to the natural scale to ensure that fishing mortality reference points are positive.

When calculating derived quantities, such as the current stock status relative to a reference point, the correlation between the estimates must be accounted for. To this end, Equation (2) is modified to calculate the joint covariance of parameters, latent processes and reference points. In turn, the delta method can be used to calculate the variance of the derived quantity as a function of, for instance, the reference point and latent processes.

2.4. Case study: Northeast Arctic cod

To illustrate the two methods, we calculated reference points for Northeast Arctic cod (*Gadus morhua*). The Northeast Arctic cod assessment in 2019 used the SAM model (ICES, 2019). Data used in the assessment includes 73 years of commercial catches and four scientific surveys (for further details, see supplementary Figure S1 and ICES, 2019). We used the same data and model configuration as the assessment. However, the model was fitted using each of the 12 recruitment models outlined above: random walk, Ricker, Beverton-Holt, AR(1), bent hyperbola, power function, Shepherd, Hassell/Deriso, Saila-Lorda, sigmoidal Beverton-Holt, CMP spline, and general spline.

For each of the 12 recruitment models, reference points and corresponding 95% confidence intervals were estimated using the PR-based method and the forecast-based method. Further, $\mathrm{F}_{\text{status quo}}$ was compared between the methods. For the PR method, $\mathrm{F}_{\mathrm{MSY}}$, $\mathrm{F}_{\max}$, $\mathrm{F}_{0.1}$, $\mathrm{F}_{35\%}$, $\mathrm{F}_{\mathrm{lim}}$, and $\mathrm{F}_{\mathrm{Crash}}$ were estimated. In the forecast method, only one reference point can be estimated for each long-term forecast. Therefore, only $\mathrm{F}_{\mathrm{MSY}}$ was estimated for this method. For the random walk recruitment, $\mathrm{F}_{\mathrm{MSY}}$ can not be estimated by the PR method while $\mathrm{F}_{\mathrm{Crash}}$ is not defined for the random walk, AR(1), and power-law recruitment models. Likewise, $\mathrm{F}_{\mathrm{Crash}}$ is not defined for the Saila-Lorda and sigmoidal Beverton-Holt recruitment functions in the presence of depensatory recruitment; that is, when $\gamma > 1$. Similarly, $\mathrm{F}_{\mathrm{Crash}}$ was not estimated for the general spline.

In both methods, selectivity in the last year of the assessment was used for calculating the reference points, while input data such as stock weights and natural mortalities were averaged over the last 15 years. Assumptions about the long-term selectivity, stock weights and natural mortalities must be made for both methods. However, other years to average over could be chosen.

2.5. Simulation study

To validate the methods, a simulation study was conducted. Based on each of the 12 case study model fits, 500 data sets were generated. Given the estimated parameter values from the case study, new $\log F$ and $\log N$ (including recruitment) processes were simulated.

For the $\log F$ process, selectivity was simulated using the estimated selectivity patterns from the case study while average fishing mortality was simulated from the auto-regressive process:

$$\log \bar{F}_y \sim N\left(\log \mu_y + 0.7\left(\log \bar{F}_{y-1} - \log \mu_{y-1}\right), 0.05^2\right)$$

with

$$\mu_y = \frac{0.6}{1 + \exp\left(-2\cos\left(\frac{y}{4} + \pi\right)\right)}.$$

This process was used instead of the random walk used for estimation to produce more realistic fishing mortality patterns and to reduce the probability of stock collapse in the simulations.

Given the fishing mortality, the $\log N$ process was simulated from the process model. Recruitment was simulated using the respective stock recruitment functions. All remaining parameters, and input data were taken from the case study. Based on the new $\log F$ and $\log N$ processes, data were simulated from the observation model.

For each simulated data set, the true model was fitted and all reference points were estimated by the PR method. In addition, $\mathrm{F_{MSY}}$ was estimated by the forecast method. Estimated reference points were compared to the true values calculated through a PR analysis using the true parameter values and processes. Bias in reference points was estimated using relative errors. Likewise, estimated recruitment model parameters were compared to the true values.

3. Results

3.1. Case study: Northeast Arctic cod

In the case study, 12 recruitment models were fitted to data from the 2019 assessment of Northeast Arctic cod (Figure 1). While the 12 models reflect very different assumptions about the stock-recruitment relationship, all models provided similar estimates of spawning-stock biomass (SSB), and average fishing mortality ($\bar{F}$). Compared to the average over all models, SSB estimates were within 4.83% while $\bar{F}$ estimates were within 3.51% for all models. In contrast, estimated recruitment values were more variable. While 75% were within 1.56% of the average and 95% were within 5.07%, the highest difference was 31.9%.

Of the four models generalizing the Ricker and Beverton-Holt recruitment, three were estimated to resemble the Ricker model; namely, the Shepherd, Hassel/Deriso, and the Saila-Lorda models. In contrast, the sigmoidal Beverton-Holt model was estimated to have depensatory mortality with an estimated γ of 4.268 (95% confidence interval: 2.195; 8.301).

Similar to SSB and $\bar{F}$, $\mathrm{F_{Status\ quo}}$ was similar for all 12 recruitment models (Figure 2). Estimated values ranged from 0.397 $\mathrm{y^{-1}}$ to 0.418 $\mathrm{y^{-1}}$, while the length of corresponding 95% confidence intervals ranged from 0.505 $\mathrm{y^{-1}}$ to 0.529 $\mathrm{y^{-1}}$. Likewise, as expected, the reference

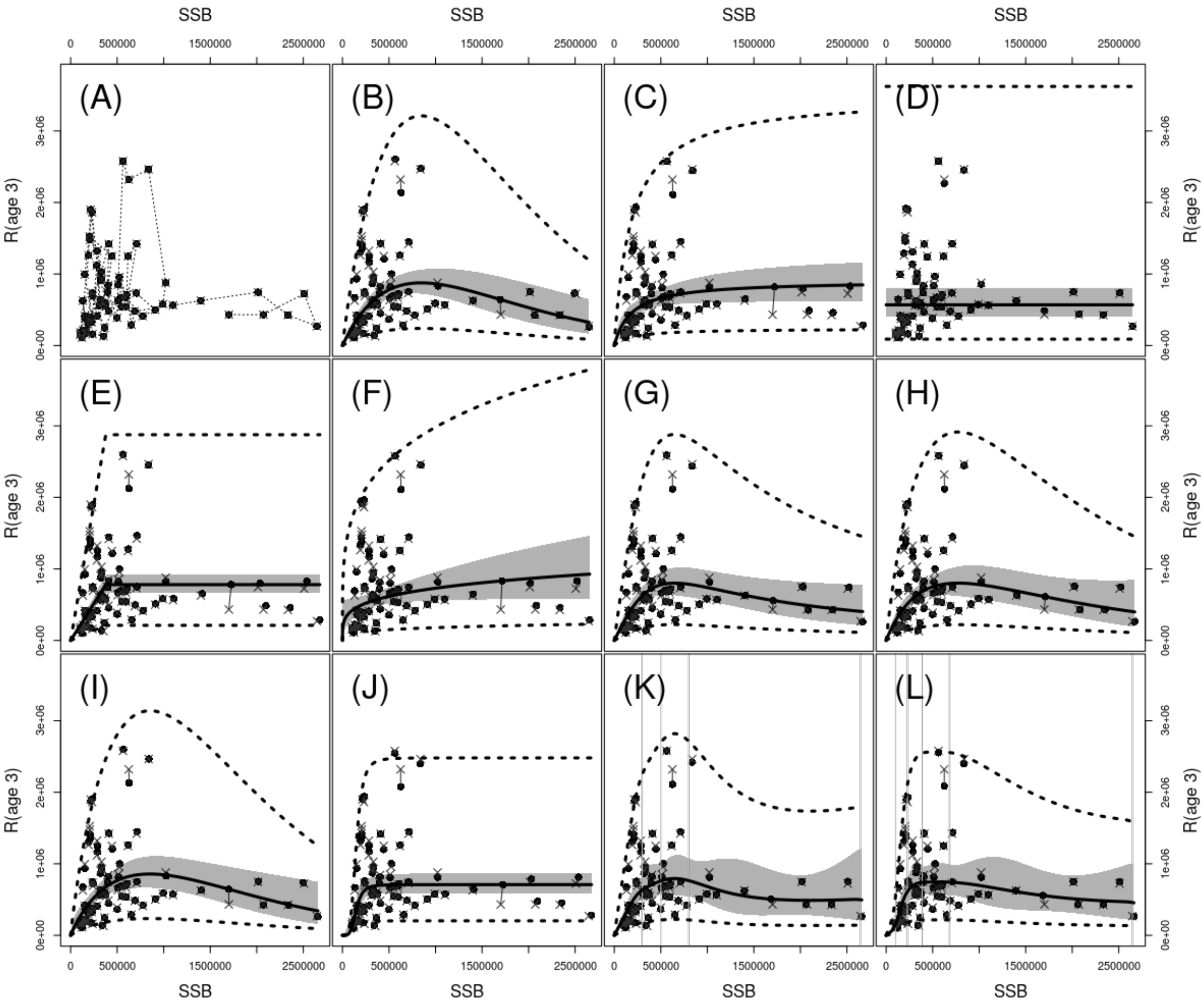

Figure 1: Estimated spawning-stock biomass (SSB), recruitment (R), and stock-recruitment relationship (full black line) with 95% confidence intervals (grey area) and 95% prediction intervals conditional on the model (dashed black lines) in the Northeast Arctic cod case study for the stock-recruitment functions: Random Walk (A), Ricker (B), Beverton-Holt (C), AR(1) (D), Bent Hyperbola (E), Power (F), Shepherd (G), Hassell/Deriso (H), Saila-Lorda (I), Sigmoidal Beverton-Holt (J), CMP Spline (K), and General Spline (L). Estimated (SSB,R) points from the 2019 assessment is indicated by a grey cross and connected to the corresponding new estimate by a thin grey line. In panel (A), the temporal order is indicated by a dotted line. For the spline models, knot positions are indicated by a vertical grey line.

points that did not depend on the stock recruitment relationship were similar between the fitted models. For $F_{0.1}$, estimated values ranged from 0.176 y^{-1} to 0.179 y^{-1} with 95% confidence intervals ranging in length from 0.313 y^{-1} to 0.318 y^{-1}. Estimated values of F_{Max} ranged from 0.313 y^{-1} to 0.316 y^{-1}, while estimated $F_{35\%}$ ranged from 0.144 y^{-1} to 0.145 y^{-1}. Again, 95% confidence intervals had similar lengths for both of these reference points. For F_{Max}, the lengths ranged from 0.549 y^{-1} to 0.557 y^{-1} while they ranged from 0.255 y^{-1} to 0.258 y^{-1} for $F_{35\%}$.

In contrast to the recruitment independent reference points, F_{MSY} and F_{Crash} differed substantially between the models. In the models where it was defined, F_{Crash} ranged from 0.992 y^{-1} for the bent hyperbola model to 2.3 y^{-1} for the Saila-Lorda recruitment model. For the bent hyperbola, F_{lim} was estimated along with the other reference points. For this model, F_{lim} was estimated to 0.989 y^{-1} (0.418; 2.349), which is almost identical to F_{Crash}, estimated to be 0.989 y^{-1} (0.417; 2.349). This is expected since F_{Crash} and F_{lim} are two ways of defining a limit reference point such that higher fishing mortality rates will, eventually, lead to stock extinction.

F_{MSY} was estimated using both the forecast and the PR method. The two methods provided identical reference point estimates. However, 95% confidence intervals were consistently wider for the PR method. The smallest difference was found for the AR(1) model, where the PR-based confidence interval was 37.5% wider than the forecast-based confidence interval. The largest difference was found for the general spline recruitment model where the PR-based interval was 228.4% wider. For the random walk, AR(1), bent hyperbola, and sigmoidal Beverton-Holt model, recruitment was constant above a certain SSB. Therefore, F_{Max} and F_{MSY} were identical for these four models.

3.2. Simulation study

3.2.1. Model parameters

In general, recruitment model parameters were estimated accurately (Figure 3). For the five two parameter models, median relative bias was between -0.037 and 0.012. In these models, the parameter furthest from the truth was the β parameter of the bent hyperbola; that is, half the slope of the recruitment curve at the origin. This was also the parameter with the highest uncertainty Likewise, the four models with three recruitment parameters generally provided accurate estimates. However, the β parameter of the sigmoidal Beverton-Holt model was problematic for some simulations. While the median was close to the true value, some of the simulations had estimates with a relative bias of 0.35. A similar pattern was seen for the CMP spline where the estimates of the first two knot parameters fell in two groups. Again, the median estimates for one of the groups were close to the true values. For both models, we suspect that some of the simulations did not cover a sufficient range of spawning biomasses to accurately identify the parameters.

Consequently, confidence interval coverages were generally close to the nominal 95% for most parameters (Supplementary Material Figures S3-S13). For the bent hyperbola, coverage of the α parameter was 77.0%, while the coverage was 78.0% and 87.0% for the second and fourth knot parameter of the CMP spline, respectively. Likewise, the coverage was 82.2% for

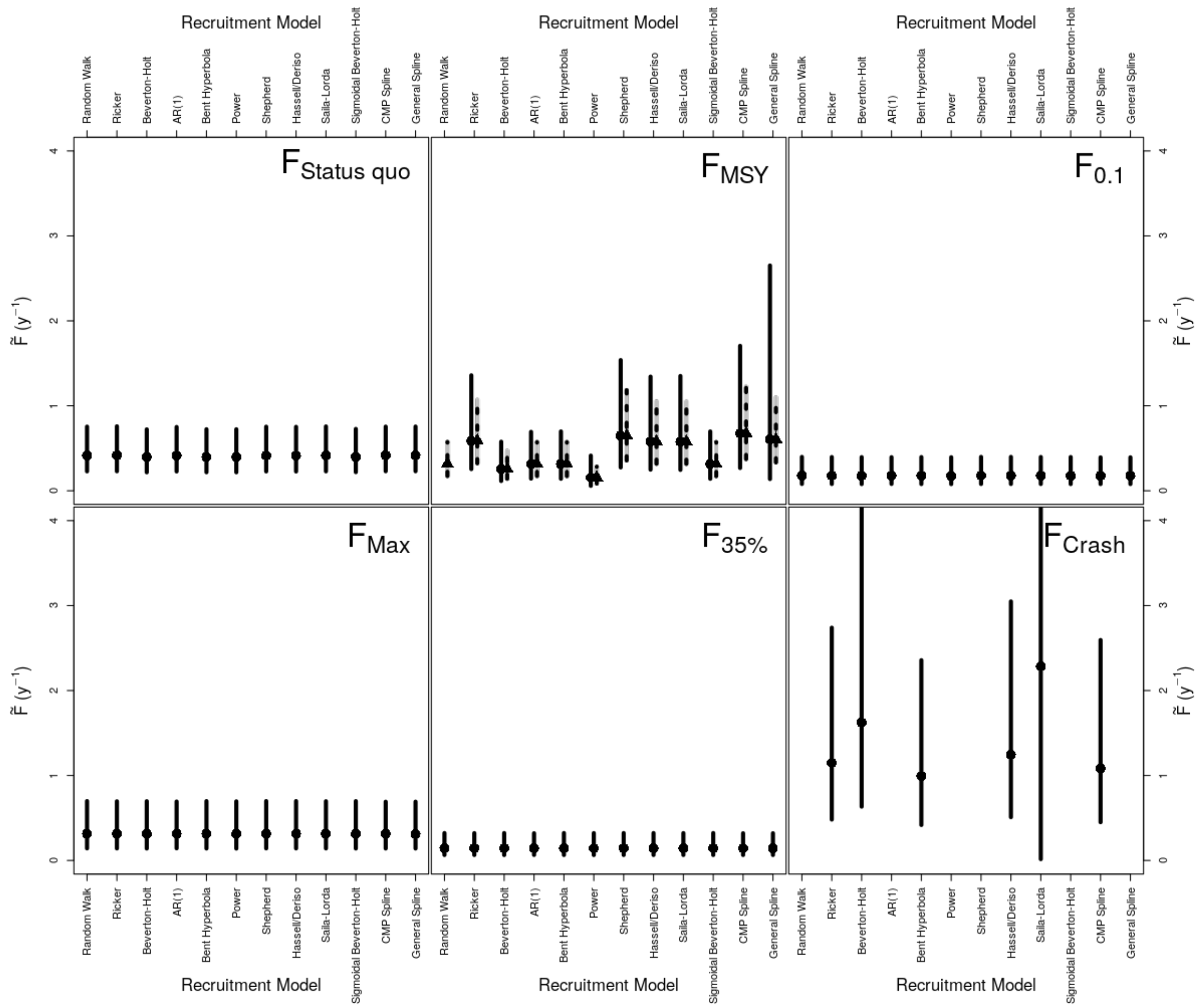

Figure 2: Estimated fishing mortality reference points with 95% confidence intervals in the Northeast Arctic cod case study. For $\mathrm{F_{MSY}}$, the dots are the PR-based estimates and the triangles the forecast-based estimates.

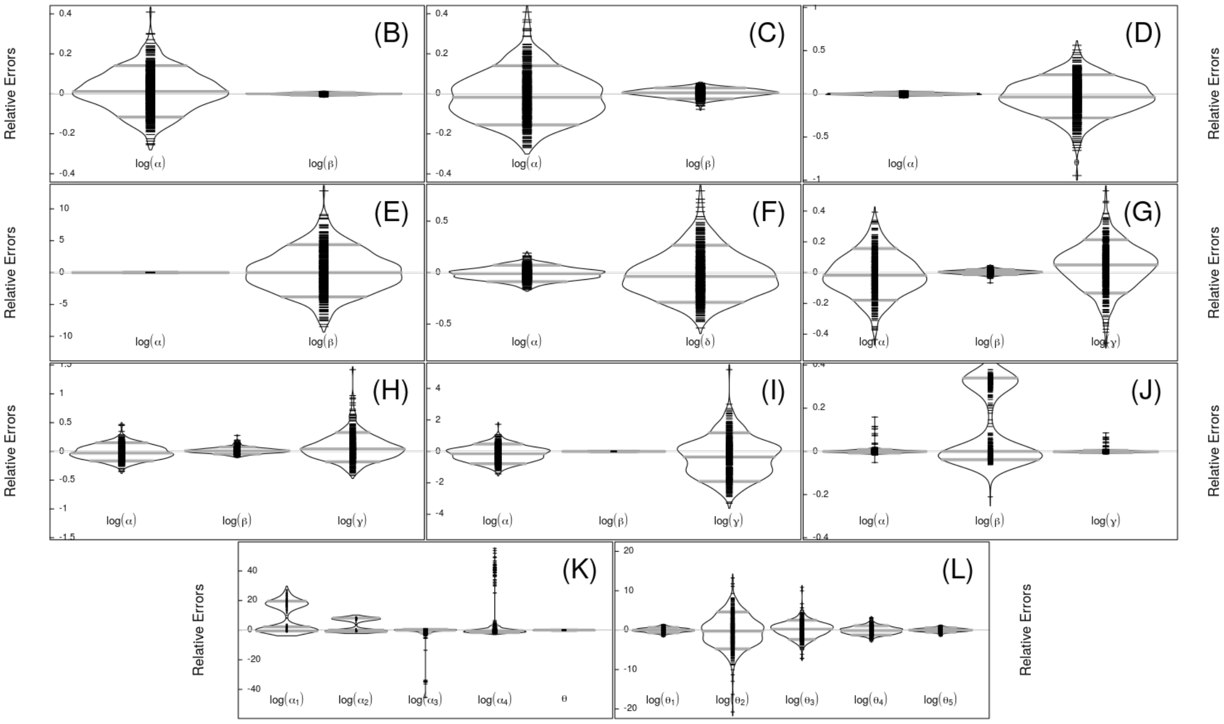

Figure 3: Relative errors in recruitment parameter estimates in the simulation study, for the different recruitment assumptions: Ricker (B), Beverton-Holt (C), AR(1) (D), Bent Hyperbola (E), Power (F), Shepherd (G), Hassell/Deriso (H), Saila-Lorda (I), Sigmoidal Beverton-Holt (J), CMP Spline (K), and General Spline (L). The horizontale grey lines show the 10%, 50% and 90% quantiles. Note the change in y-axis.

the β parameter of the sigmoidal Beverton-Holt. Remaining parameters had coverages above 92.6%.

3.2.2. Fishing mortality reference points

Despite the estimation issues for a few parameters, reference points were well estimated by the PR method for all 12 models (Figure 4). For the ten parametric models and the CMP spline, median relative bias ranged from -0.036 to 0.053 across all PR-based reference points. For the general spline, the relative bias of F_{MSY} was -0.125 while it ranged from -0.027 to -0.005 for the remaining reference points.

Overall, the PR method provided conservative confidence intervals. For almost half of the estimated reference points, the confidence interval always contained the true value, while 65% had a coverage above the nominal 95%. For the sigmoidal Beverton-Holt model, however, the coverage was 63% for each of the reference points, while the AR(1) model had coverages around 71%. Further, the bent hyperbola had coverages around 90%, while the coverages for the CMP spline was 92%, except for the second knot parameter, which had a coverage of

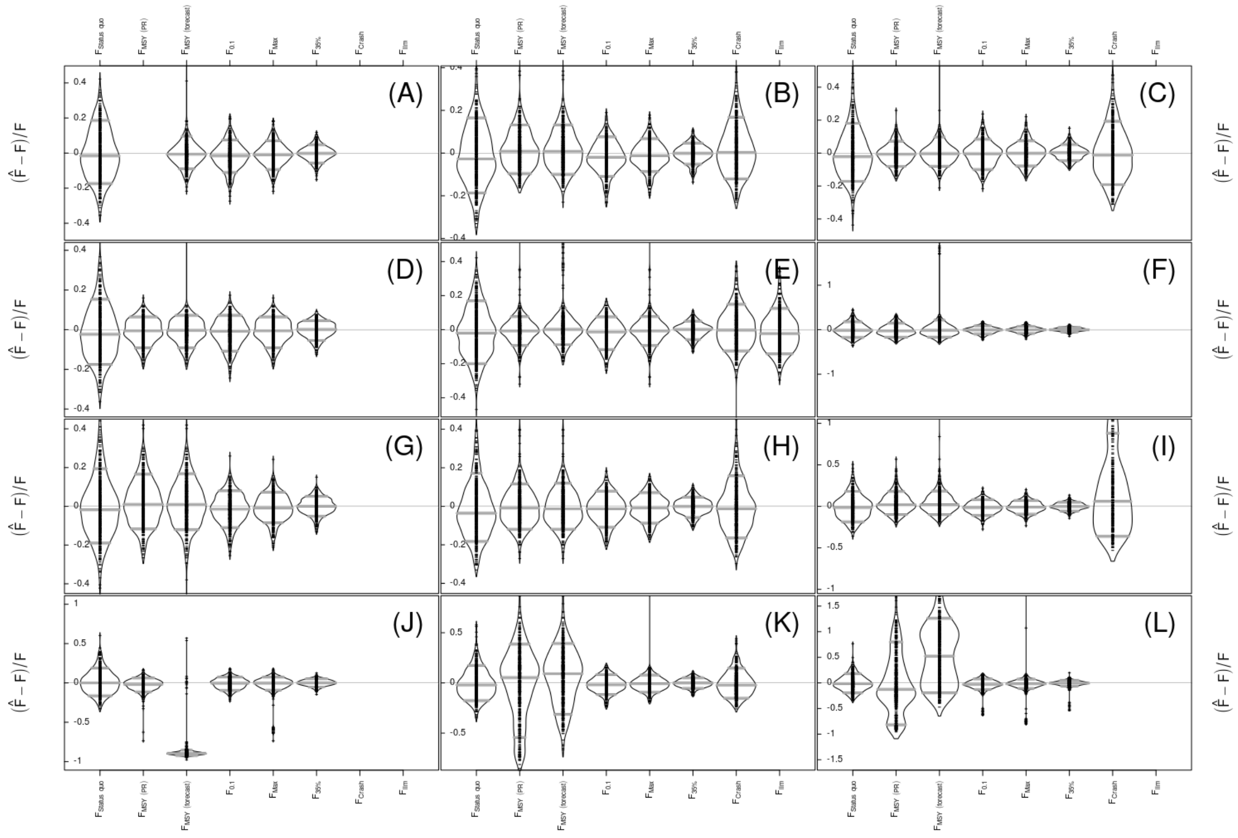

Figure 4: Relative errors in fishing mortality reference point estimates in the simulation study, for the different recruitment assumptions: Random Walk (A), Ricker (B), Beverton-Holt (C), AR(1) (D), Bent Hyperbola (E), Power (F), Shepherd (G), Hassell/Deriso (H), Saila-Lorda (I), Sigmoidal Beverton-Holt (J), CMP Spline (K), and General Spline (L). The horizontale grey lines show the 10%, 50% and 90% quantiles. Note the change in y-axis.

87.6%.

For most of the models, the forecast-based method also provided accurate estimates of F_{MSY}. The mean relative bias ranged from -0.012 to 0.008 for the five recruitment models with two parameters. Likewise, the median relative bias was -0.006 for the random walk, 0.009 for the Shepherd model, -0.013 for the Hassel/Deriso model, 0.020 for the Saila-Lorda model, and 0.091 for the CMP spline. For the sigmoidal Beverton-Holt and general spline, however, the method had substantial relative bias of -0.896 and 0.523, respectively. Consequently, confidence interval coverages were 2.8% for the sigmoidal Beverton-Holt model and 60% for the general spline (Supplementary Material Figures S3-S13). In the remaining models, coverages ranged from 93.6 to 99.6% (Supplementary Material Figures S14-S25).

3.2.3. Equilibrium biomass

Besides estimated fishing mortality reference points, corresponding equilibrium biomasses were compared to the true values (Figure 5). Again, both methods generally provided accurate results. Apart from the spline models and the sigmoidal Beverton-Holt model,

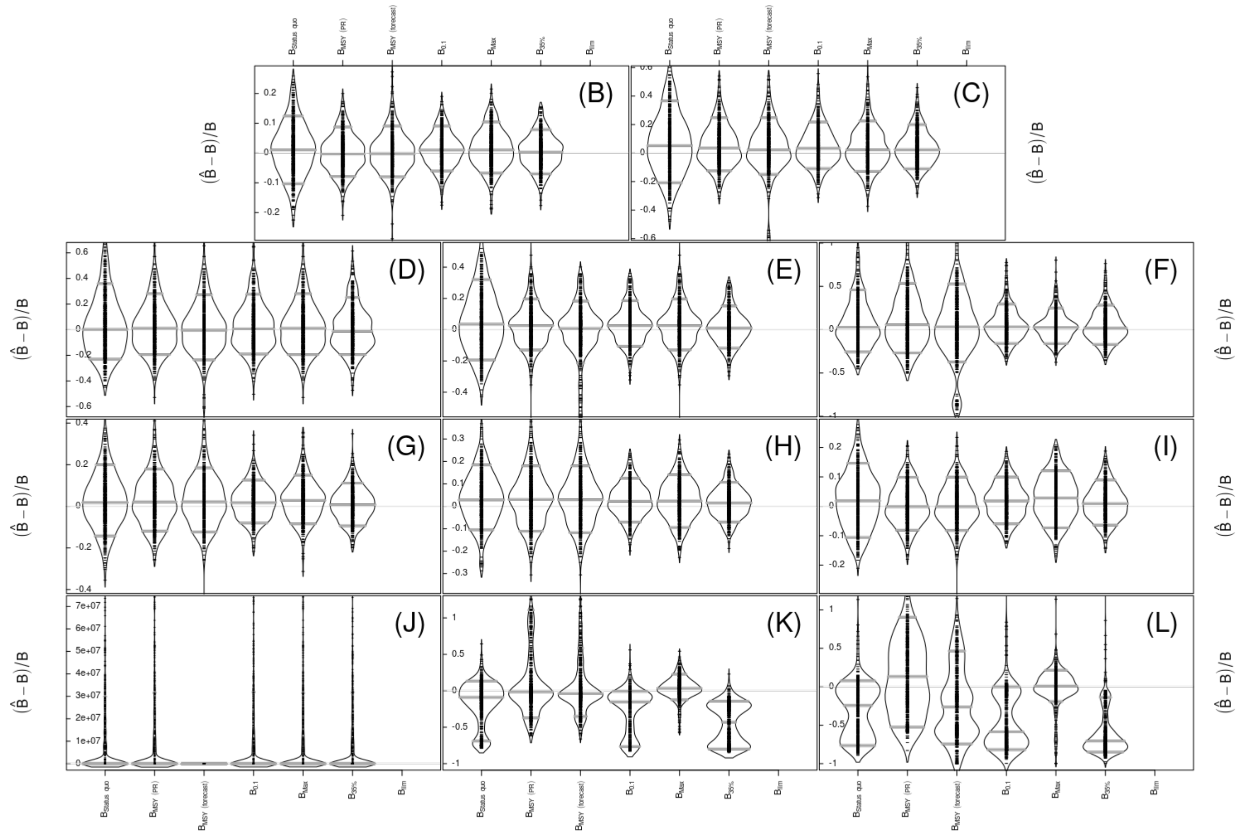

Figure 5: Relative errors in biomass reference point estimates in the simulation study, for the different recruitment assumptions: Ricker (B), Beverton-Holt (C), AR(1) (D), Bent Hyperbola (E), Power (F), Shepherd (G), Hassell/Deriso (H), Saila-Lorda (I), Sigmoidal Beverton-Holt (J), CMP Spline (K), and General Spline (L). The horizontale grey lines show the 10%, 50% and 90% quantiles. Note the change in y-axis.

median relative bias in the PR method ranged from -0.013 to 0.056. Relative bias was larger for the three models using numerical methods to find equilibrium biomass. For the sigmoidal Beverton-Holt, median relative bias ranged from -0.102 to -0.052 while it was between -0.705 and 0.133 for the spline models. For the PR method, equilibrium biomass could not be calculated for the random walk recruitment model. Confidence interval coverages were similar to the corresponding fishing mortality reference points.

Likewise, the forecast-based method generally provided accurate estimates of B_{MSY}. Except from the sigmoidal Beverton-Holt and general spline models, median relative bias ranged from -0.038 to 0.043 while confidence interval coverages were above 93.8%. For the general spline model, median relative bias was -0.264 with a coverage of 97.0%. For the sigmoidal Beverton-Holt model, median relative bias was -1, corresponding to zero biomass, with a coverage of 2.6%, indicating that care is needed with the method in combination with depensatory mortality and multiple optima for equilibrium biomass.

4. Discussion

Reference points are essential for management and conservation of marine living resources. However, reference point estimates are subject to uncertainty that should be quantified and accounted for. To this end, we provided a method for estimating reference points and their uncertainty in general fisheries stock assessment models. Two methods were used to estimate reference points: a method based on per-recruit calculations (PR) and a method based on a long-term forecast. For both methods, the same procedure was used to obtain confidence intervals. By thinking of reference point estimates as implicit functions of the model parameters, a combination of the implicit function theorem and the delta method was used to provide confidence intervals for the estimates. Through the implicit function theorem, Zheng et al. (2019) recently used similar ideas to quantify local sensitivity of MSY reference points to perturbations in, amongst others, natural mortality, selectivity, growth, and maturity. Here, the focus was instead on quantifying uncertainty in estimates given the model assumptions and data at hand.

To illustrate the general applicability of the methods, reference points were calculated for data from the 2019 Northeast Arctic cod assessment (ICES, 2019) using 12 recruitment functions. The recruitment models included two time-series models, four two-parameter models, four three-parameter models, and two spline models; thereby, representing a wide range of model classes (see, e.g., Punt and Cope, 2019, for a list of other models not considered here). Estimates for reference points that do not directly depend on recruitment (i.e, $F_{\text{Status Quo}}$, $F_{0.1}$, F_{Max}, and $F_{35\%}$) were practically identical for all 12 models. In contrast, F_{MSY} and F_{Crash} were highly dependent, as expected, on the imposed stock-recruitment relationship. Comparing, or selecting between, the models was outside the scope of this paper. However, techniques such as AIC (e.g., Albertsen et al., 2017, 2018) and prediction residuals (e.g., Thygesen et al., 2017) have been used for state-space stock assessment models. For any validation method, the model should be adequately validated before reference points are considered. Likewise, knot selection for the spline models was not considered here. Selecting the position of knots for regression splines is difficult. For fully observed statistical models, advice can be found in textbooks (e.g., Harrell, 2013). However, in the present model, the regressor (i.e., SSB) is unobserved. Therefore, using, for example, quantiles of the regressor is not possible without first fitting another stock-recruitment model. For simplicity, SSB estimates from the original assessment model was used to place knots in the case study. However, this is generally not recommended as it requires a double use of the data to fit the model. Such a procedure may lead to overfitting. Alternatively, subject knowledge may be used to place the knots a priori, numerous knots can be used while penalizing the parameters (e.g., Eilers and Marx, 1996), or several knots combinations could be tested until the best convergence is obtained.

For F_{MSY}, both a PR and a forecast-based method were used for estimation. For all 12 recruitment models in the case study, both methods provided identical reference point estimates but confidence intervals differed. In agreement with the case study, the PR method provided wider confidence intervals, reflected by a larger coverage, in the simulation study. The method for calculating uncertainties relies on differentiability of the optimization criterion. Therefore, confidence intervals may be affected when reference point estimates are close to F_{Crash}. Particularly, when the optimization criterion depends on equilibrium biomass. Further, Hessian- and delta method-based confidence intervals rely on asymptotic normality, which is not parameterization invariant. Therefore, different ways of obtaining the reference points may lead to different confidence intervals for (statistically) small sample sizes. This applies to considerations between the PR and forecast method as well as different optimization criteria giving the same optimum. When sample sizes are small, confidence intervals based on asymptotic normality may be inappropriate. In this case, other methods can be used to quantify uncertainty in fisheries stock assessments (e.g., Magnusson et al., 2013). In particular, simulation-based methods such as bootstrapping (e.g., Punt and Butterworth, 1993; Overholtz, 1999; Cadigan, 2012), Monte Carlo simulations (e.g., Grabowski and Chen, 2004; Hart, 2013; Braccini et al., 2015), or simulation studies (e.g., García-Carreras et al., 2015) have been used for reference points. In state-space models, however, special attention must be given for resampling-based methods to account for the temporal dependence of data (e.g., Stoffer and Wall, 1991).

Despite this difference in confidence intervals, we show that both the PR and forecast method can be robustly used to estimate MSY reference points. In the simulation study, both methods generally provided accurate reference points estimates with similar relative errors. However, both methods have limitations. The number of years for the projection in the forecast method can be limiting as the projection should be long enough to reach the equilibrium but long projections would also be computationally demanding. The number of years necessary to get to the equilibrium will vary depending on the fish stock biology, notably its reproduction and growth, but also dependng on the fishing mortality. For instance, long-lived species may necessitate a longer forecast to reach the steady state. Similarly, low fishing mortality may induce slow stock development as the exponential decay over time will converge less quickly to the equilibrium. This method requires therefore extra checks by the user to verify that the equilibrium is reached at the end of the forecast. It should be straightforward when estimating reference points for a stock to start with a low number of years (e.g., 20-30 years) and gradually increase the years if needed until the stock stabilize. The necessary verification can be made by, for instance, visual inspection of the forecasted time series. The PR method does not present this problem when equilibrium biomass is known analytically. However, when numerical methods are needed, extra checks are needed to verify the results. In the simulations, equilibrium biomass was difficult to estimate in the sigmoidal Beverton-Holt model. When calculating equilibrium values, the PR method does not account for stochasticity in processes such as survival and fishing mortality compared to the forecast method, but similarly accounts for uncertainty in the stock-recruitment relationship and the parameter estimates. Ignoring the errors in the hypothetical future abundance and fishing mortality processes did not seem to affect the median reference points estimates.

Besides MSY reference points, the PR method was used to calculate $F_{0.1}$, F_{Max}, $F_{35\%}$, and F_{Crash}. Further, F_{lim} was calculated for the bent hyperbola (or smooth hockey-stick) recruitment model. For the bent hyperbola model, F_{lim} and F_{Crash} were identical. This was expected, since both are limit reference points where the stock can no longer sustain itself. The ICES definition of F_{lim} does, however, not easily transfer to other recruitment models than hockey-sticks. B_{lim} is defined by ICES to be "a deterministic biomass limit below which a stock is considered to have reduced reproductive capacity" (ICES, 2017). Then, F_{lim} is the fishing mortality rate leading to B_{lim} in the equilibrium. However, for the power-law recruitment, reproductive capacity is reduced by any reduction in spawning biomass. Likewise, for the Ricker recruitment model, reproductive capacity is both reduced for small and large spawning biomasses. In contrast to F_{lim}, F_{Crash} can be used for any recruitment model with the compensatory mortality property.

Similar to F_{MSY}, the PR method was able to accurately estimate the five other reference points for each of the 12 recruitment models. Both the fishing mortality reference points and corresponding equilibrium biomasses had median relative biases close to zero. Overall, the method performed better for recruitment models where a closed form solution was known for equilibrium biomass. For two recruitment models where numerical methods were needed, the PR method had difficulties for some equilibrium biomasses. Likewise, the forecast method had issues with these two models: both with depensatory mortality.

In the presence of depensatory mortality, there are multiple possible positive equilibrium biomasses. This poses a potential problem when estimating reference points. However, in applications, the estimated values can easily be validated graphically through the stock-recruitment curve and replacement line Even for recruitment models with a unique positive equilibrium biomass, estimation problems for reference points can occur. In the model used here, reference points depend on the stock recruitment function as well as the estimated selectivity determined by the fishing mortality process. Therefore, the latent abundance and fishing mortality processes must be estimated well, which can be difficult in state-space models (Auger-Méthé et al., 2016).

The methods presented here only propagates uncertainty in parameter estimates and latent processes to the reference point confidence intervals. Therefore, uncertainty in quantities considered known in the model, as well as the model itself, is not incorporated. This relates to, for instance, natural mortality, weight-at-age, and mortality. Instead, uncertainty related to these can be evaluated directly within the assessment model (Methot and Wetzel, 2013) or by, for example, local sensitivity (e.g., Cadigan and Wang, 2016; Zheng et al., 2019) or sensitivity analyses (e.g., Zhu et al., 2012). Moreover, the reference point estimates ignore external factors that are not explicitly considered in the models such as environmental effects (e.g., Miller et al., 2016b), spatial variation (Punt et al., 2020), mixed fisheries considerations (e.g., Albertsen et al., 2018; Rindorf et al., 2016) or multispecies interactions (e.g., Trijoulet et al., 2018, 2019, 2020).

For assessments by ICES, reference point estimates are often calculated through a post hoc analysis (ICES, 2017). The assumptions of the post hoc analysis will often be distinct from the assumptions of the assessment model. Further, using model estimated stock-recruitment pairs as data in a post hoc analysis may lead to bias in the results (Brooks and Deroba, 2015). In contrast, the methods presented here ensures that assumptions are consistent between the assessment model and subsequent reference point estimates and automatically treats stock-recruitment pairs as uncertain model estimates. For the case study considered here, stock-recruitment pairs only differed slightly between the recruitment models. This indicates that the stock-recruitment pair estimates are stable. Therefore, reference points are likely influenced more by different assumptions about recruitment than by a post hoc analysis. This should be investigated in the future.

Reference points for both methods were estimated assuming selectivity in the last year of the assessment, last 15 years average for input data to the assessment (e.g., weight at age, natural mortality, proportion of mature fish at age), and these were kept constant in the forecast method. The recruitment was also estimated assuming one stock-recruitment relationship for the entire time-series. The average years chosen or considering dynamic values for these variables may affect the reference point estimates (Berger, 2019). Different methods could be chosen but the same assumptions should be used in the estimation model and in calculating reference points. Therefore, dynamic values may necessitate calculating different reference points for each year corresponding to a hypothetical equilibrium. However, if the dynamic values are modelled by a process, the forecast method can be used to forecast the system to an equilibrium, assuming it exists.

Although the reference point estimates are ensured to be consistent with model assumptions, they are estimated in a separate step after model parameters are estimated. This allows the use of any general adaptive optimizer. Further, starting values for the optimizer can easily be altered and tested. Finally, this allows the method to be used with derivative information obtained by any method: analytical, numerical, symbolic, or through automatic differentiation. Instead, the SS3 (Methot and Wetzel, 2013) and WHAM (Miller et al., 2016a, 2018) models calculate reference points by applying a fixed number of newton iterations as part of the model implementation in AD Model Builder and TMB, respectively. While this method will give correct results, it requires pre-specifying a sufficient number of steps and does not easily apply without automatic differentiation.

Not only are the methods presented here applicable with any differentiation method, they are also directly applicable for several other model classes. In particular, the method is directly applicable for state-space assessment models for multiple stocks (e.g., Albertsen et al., 2018) or species (e.g., Trijoulet et al., 2020). Further, the methods are not limited to state-space models, but can be applied for general statistical assessment models. Likewise, the method is neither constrained to the recruitment models nor the reference points considered here. For reference points, it is, however, a limitation that it must be possible to express the reference point as a function of model parameters, latent variables, or derived quantities.

When a precautionary approach to management and conservation of marine living resources is warranted, the inherent uncertainty in reference point estimates should be accounted for. Therefore, accurate evaluations of reference point uncertainties are necessary for managers, and policy-makers to make informed decisions about the trade-offs between conservation and exploitation of fish stocks. We presented a method for estimating model consistent reference points and their uncertainty. Reference points based on the same assumptions as the assessment model make the assessment process more transparent and allow accurate estimates of reference point uncertainty needed for precautionary management.

Acknowledgements

This work was funded by a grant from the European Fisheries and Maritime Fund (33113-B-17-106, Ministry of Environment and Food in Denmark). We thank Andre Punt and an anonymous reviewer for providing useful feedback on this manuscript.

References

Albertsen, C. M., Nielsen, A., Thygesen, U. H., 2017. Choosing the observational likelihood in state-space stock assessment models. Can. J. Fish. Aquat. Sci. 74 (5), 779–789.
`https://doi.org/10.1139/cjfas-2015-0532`

Albertsen, C. M., Nielsen, A., Thygesen, U. H., 2018. Connecting single-stock assessment models through correlated survival. ICES J. Mar. Sci. 75 (1), 235–244.
`https://doi.org/10.1093/icesjms/fsx114`

Auger-Méthé, M., Field, C., Albertsen, C. M., Derocher, A., Lewis, M. A., Jonsen, I. D., Flemming, J. M., 2016. State-space modelsø dirty little secrets: even simple linear gaussian models can have parameter and state estimation problems. Sci. Rep. 6, 26677.
`https://doi.org/10.1038/srep26677`

Australian Government, 2018. Guidelines for the Implementation of the Commonwealth Fisheries Harvest Strategy Policy. Tech. rep., Department of Agriculture and Water Resources,.

Berg, C. W., Nielsen, A., 2016. Accounting for correlated observations in an age-based state-space stock assessment model. ICES J. Mar. Sci. 73 (7), 1788–1797.
`https://doi.org/10.1093/icesjms/fsw046`

Berger, A. M., 2019. Character of temporal variability in stock productivity influences the utility of dynamic reference points. Fish. Res. 217, 185 – 197, recruitment: Theory, Estimation, and Application in Fishery Stock Assessment Models.
`http://www.sciencedirect.com/science/article/pii/S0165783618303424`

Beverton, R. J. H., Holt, S. J., 1957. On the dynamics of exploited fish populations. Fishery Investigations Series 2, volume 19, H.M.S.O, London, U.K.

Braccini, M., Brooks, E. N., Wise, B., McAuley, R., 2015. Displaying uncertainty in the biological reference points of sharks. Ocean Coast. Manag. 116, 143 – 149.
`https://doi.org/10.1016/j.ocecoaman.2015.07.014`

Brooks, E. N., Deroba, J. J., 2015. When “data” are not data: the pitfalls of post hoc analyses that use stock assessment model output. Can. J. Fish. Aquat. Sci. 72 (4), 634–641.
`https://doi.org/10.1139/cjfas-2014-0231`

Caddy, J. F., Mahon, R., 1995. Reference points for fisheries management. Vol. 374. Food and Agriculture Organization of the United Nations Rome.

Cadigan, N. G., 2012. Fitting a non-parametric stock-recruitment model in R that is useful for deriving MSY reference points and accounting for model uncertainty. ICES J. Mar. Sci. 70 (1), 56–67.
`https://doi.org/10.1093/icesjms/fss183`

Cadigan, N. G., Wang, S., jan 2016. Local sensitivity of per-recruit fishing mortality reference points. J. Biological Dyn. 10 (1), 525–545.
`https://doi.org/10.1080/17513758.2016.1225831`

Cushing, D. H., 1971. The Dependence of Recruitment on Parent Stock in Different Groups of Fishes. ICES J. Mar. Sci. 33 (3), 340–362.
`https://doi.org/10.1093/icesjms/33.3.340`

Deriso, R. B., 1978. Non-linear age-structured models for seasonally breeding populations. Ph.D. thesis, University of Washington.

Eilers, P. H. C., Marx, B. D., 1996. Flexible smoothing with *b*-splines and penalties. Stat. Sci. 11 (2), 89–102.
`http://www.jstor.org/stable/2246049`

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., Sibert, J., 2012. Ad model builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. 27 (2), 233–249.
https://doi.org/10.1080/10556788.2011.597854

Fritsch, F. N., Shafer, R. E., Crowley, W. P., 1973. Solution of the transcendental equation wew = x. Commun. ACM 16 (2), 123–124.
https://doi.org/10.1145/361952.361970

Garcia, S. M., 1994. The Precautionary Principle: its implications in capture fisheries management. Ocean Coast. Manag. 22 (2), 99–125.
https://doi.org/10.1016/0964-5691(94)90014-0

García-Carreras, B., Jennings, S., Le Quesne, W. J. F., 11 2015. Predicting reference points and associated uncertainty from life histories for risk and status assessment. ICES J. Mar. Sci. 73 (2), 483–493.
https://doi.org/10.1093/icesjms/fsv195

Grabowski, R., Chen, Y., 2004. Incorporating uncertainty into the estimation of the biological reference points f0.1 and fmax for the maine green sea urchin (strongylocentrotus droebachiensis) fishery. Fish. Res. 68 (1), 367 – 371.
http://www.sciencedirect.com/science/article/pii/S0165783604000293

Harrell, F., 2013. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer Series in Statistics. Springer New York.
https://doi.org/10.1007/978-1-4757-3462-1

Hart, D. R., 01 2013. Quantifying the tradeoff between precaution and yield in fishery reference points. ICES J. Mar. Sci. 70 (3), 591–603.
https://doi.org/10.1093/icesjms/fss204

Hassell, M. P., 1975. Density-dependence in single-species populations. J. Anim. Ecol. 44 (1), 283–295.

ICES, 2017. ICES fisheries management reference points for category 1 and 2 stocks. ICES Advice Technical Guidelines 12.4.3.1, International Council for the Exploration of the Sea.
https://doi.org/10.17895/ices.pub.3036

ICES, 2019. Arctic Fisheries Working Group (AFWG). ICES Scientific Reports 1:30, International Council for the Exploration of the Sea.
https://doi.org/10.17895/ICES.PUB.5292

ICES, 2020. Library. Accessed: 2020-01-28.
http://www.ices.dk/publications/library/Pages/default.aspx?k=ALL(SAM)

Iles, T., 1994. A review of stock-recruitment relationships with reference to flatfish populations. -l. J. Sea Res. 32 (3), 399 – 420.
https://doi.org/10.1016/0077-7579(94)90017-5

Kristensen, K., Nielsen, A., Berg, C., Skaug, H., Bell, B., 2016. Tmb: Automatic differentiation and laplace approximation. J. Stat. Softw. 70 (5), 1–21.
https://doi.org/10.18637/jss.v070.i05

Laurec, A., Le Guen, J.-C., 1981. Dynamique des populations marines exploitées. tome 1. concepts et modèles. Rapports scientifiques et techniques N°45, Publications du Centre National Pour l'Exploration Des Océans, Brest, France.

Mace, P. M., 1994. Relationships between common biological reference points used as thresholds and targets of fisheries management strategies. Can. J. Fish. Aquat. Sci. 51 (1), 110–122.
https://doi.org/10.1139/f94-013

Magnusson, A., Punt, A. E., Hilborn, R., 2013. Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and mcmc. Fish Fish. 14 (3), 325–342.
https://doi.org/10.1111/j.1467-2979.2012.00473.x

Mesnil, B., Rochet, M.-J., 2010. A continuous hockey stick stock-recruit model for estimating MSY reference points. ICES J. Mar. Sci. 67 (8), 1780–1784.
https://doi.org/10.1093/icesjms/fsq055

Methot, R. D., Wetzel, C. R., 2013. Stock synthesis: A biological and statistical framework for fish stock

assessment and fishery management. Fish. Res. 142, 86 − 99.
https://doi.org/10.1016/j.fishres.2012.10.012

Miller, T., Stock, B., 2019. wham: Woods Hole Assessment Model (WHAM). R package version 0.0.0.9000.
https://github.com/timjmiller/wham

Miller, T. J., Hare, J. A., Alade, L. A., 2016a. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to southern new england yellowtail flounder. Can. J. Fish. Aquat. Sci. 73 (8), 1261–1270.
https://doi.org/10.1139/cjfas-2015-0339

Miller, T. J., Hare, J. A., Alade, L. A., 2016b. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to southern new england yellowtail flounder. Can. J. Fish. Aquat. Sci. 73 (8), 1261–1270.
https://doi.org/10.1139/cjfas-2015-0339

Miller, T. J., O'Brien, L., Fratantoni, P. S., 2018. Temporal and environmental variation in growth and maturity and effects on management reference points of georges bank atlantic cod. Can. J. Fish. Aquat. Sci. 75 (12), 2159–2171.
https://doi.org/10.1139/cjfas-2017-0124

Myers, R. A., Barrowman, N. J., Hutchings, J. A., Rosenberg, A. A., 1995. Population dynamics of exploited fish stocks at low population levels. Sci. 269 (5227), 1106–1108.
https://doi.org/10.1126/science.269.5227.1106

National Research Council, 2014. Evaluating the Effectiveness of Fish Stock Rebuilding Plans in the United States. The National Academies Press, Washington, DC.
https://www.nap.edu/catalog/18488/evaluating-the-effectiveness-of-fish-stock-rebuilding-plans-in-th

Needle, C. L., 2001. Recruitment models: diagnosis and prognosis. Rev. In Fish Biology Fish. 11 (2), 95–111.
https://doi.org/10.1023/A:1015208017674

Nielsen, A., Berg, C. W., 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fish. Res. 158, 96–101.
https://doi.org/10.1016/j.fishres.2014.01.014

Overholtz, W. J., 1999. Precision and uses of biological reference points calculated from stock recruitment data. North Am. J. Fish. Manag. 19 (3), 643–657.
https://doi.org/10.1577/1548-8675(1999)019<0643:PAUOBR>2.0.CO;2

Pedersen, M. W., Berg, C. W., 2017. A stochastic surplus production model in continuous time. Fish Fish. 18 (2), 226–243.
https://doi.org/10.1111/faf.12174

Punt, A. E., Butterworth, D. S., 1993. Variance estimates for fisheries assessment: their importance and how best to evaluate them. In: Smith, S. J., Hunt, J. J., Rivard, D. (Eds.), Risk evaluation and biological reference points for fisheries management. Canadian Special Publication of Fisheries and Aquatic Sciences 120, pp. 145–162.

Punt, A. E., Cope, J. M., 2019. Extending integrated stock assessment models to use non-depensatory three-parameter stock-recruitment relationships. Fish. Res. 217, 46 − 57, recruitment: Theory, Estimation, and Application in Fishery Stock Assessment Models.
https://doi.org/10.1016/j.fishres.2017.07.007

Punt, A. E., Dalton, M. G., Foy, R. J., 2020. Multispecies yield and profit when exploitation rates vary spatially including the impact on mortality of ocean acidification on north pacific crab stocks. Fish. Res. 225, 105481.
https://doi.org/10.1016/j.fishres.2019.105481

Ricker, W. E., 1954. Stock and recruitment. J. Fish. Res. Board Can. 11 (5), 559–623.
https://doi.org/10.1139/f54-039

Rindorf, A., Cardinale, M., Shephard, S., De Oliveira, J. A. A., Hjorleifsson, E., Kempf, A., Luzenczyk, A., Millar, C., Miller, D. C. M., Needle, C. L., Simmonds, J., Vinther, M., 06 2016. Fishing for MSY: using "pretty good yield" ranges without impairing recruitment. ICES J. Mar. Sci. 74 (2), 525–534.
https://doi.org/10.1093/icesjms/fsw111

Schnute, J., 1985. A general theory for analysis of catch and effort data. Can. J. Fish. Aquat. Sci. 42 (3), 414–429.
https://doi.org/10.1139/f85-057

Shepherd, J. G., 1982. A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. ICES J. Mar. Sci. 40 (1), 67–75.
https://doi.org/10.1093/icesjms/40.1.67

Sissenwine, M. P., Shepherd, J. G., 1987. An alternative perspective on recruitment overfishing and biological reference points. Can. J. Fish. Aquat. Sci. 44 (4), 913–918.
https://doi.org/10.1139/f87-110

Stoffer, D. S., Wall, K. D., 1991. Bootstrapping state-space models: Gaussian maximum likelihood estimation and the kalman filter. J. Am. Stat. Assoc. 86 (416), 1024–1033.
https://doi.org/10.2307/2290521

Thygesen, U. H., Albertsen, C. M., Berg, C. W., Kristensen, K., Nielsen, A., 2017. Validation of ecological state space models using the laplace approximation. Environ. Ecol. Stat. 24 (2), 317–339.
https://doi.org/10.1007/s10651-017-0372-4

Trijoulet, V., Fay, G., Curti, K. L., Smith, B., Miller, T. J., 2019. Performance of multispecies assessment models: insights on the influence of diet data. ICES J. Mar. Sci. 76 (6), 1464–1476.
https://doi.org/10.1093/icesjms/fsz053

Trijoulet, V., Fay, G., Miller, T. J., 2020. Performance of a state-space multispecies model: What are the consequences of ignoring predation and process errors in stock assessments? J. Appl. Ecol. 57 (1), 121–135.
https://doi.org/10.1111/1365-2664.13515

Trijoulet, V., Holmes, S. J., Cook, R. M., 2018. Grey seal predation mortality on three depleted stocks in the West of Scotland: What are the implications for stock assessments? Can. J. Fish. Aquat. Sci. 75 (5), 723–732.
https://doi.org/10.1139/cjfas-2016-0521

Zheng, N., Wang, S., Cadigan, N., 2019. Local sensitivity equations for maximum sustainable yield reference points. Theor. Popul. Biology 130, 143–159.
https://doi.org/10.1016/j.tpb.2019.07.010

Zhu, J., Chen, Y., Dai, X., Harley, S. J., Hoyle, S. D., Maunder, M. N., da Silva, A. M. A., 2012. Implications of uncertainty in the spawner–recruitment relationship for fisheries management: An illustration using bigeye tuna (thunnus obesus) in the eastern pacific ocean. Fish. Res. 119-120, 89 – 93.
https://doi.org/10.1016/j.fishres.2011.12.008