Detecting and correcting underreported catches in fish stock assessment: trial of a new method

Nicolas Bousquet, Noel Cadigan, Thierry Duchesne, and Louis-Paul Rivest

Abstract: Landings from fisheries are often underreported, that is, the true landings are greater than those reported. Despite this bias, reported landings are widely used in fish stock assessments, and this might lead to overoptimistic exploitation strategies. We construct a statistical stock assessment model that accounts for underreported landings using the theory of censoring with sequential population analysis (SPA). The new model is developed and implemented specifically for the cod stock (*Gadus morhua*) from the southern Gulf of St. Lawrence (Canada). This stock is known to have unreported overfishing during 1985–1992. We show in simulations that for this stock, the new censored model can correctly detect the problematic landings. These corrections are nearly insensitive to subjective boundaries placed on real catches and robust to modifications imposed in the simulation of landings. However, when surveys are too noisy, the new SPA for censored catches can result in increased uncertainty in parameters used for management such as spawning stock biomass and age-structured stock size.

Résumé : Les relevés de pêche effectués lors des débarquements sous-estiment souvent la pêche réelle. Malgré ce biais, ces relevés sont largement employés pour l'évaluation des points de repère biologiques, ce qui peut mener à une exploitation trop optimiste des ressources. Nous construisons un modèle statistique pour l'estimation des stocks de poissons qui permet l'utilisation de relevés de pêches sous-estimés, en incorporant de la censure au modèle d'analyse de population séquentielle. Ce nouveau modèle est ensuite développé et mis en oeuvre pour le stock de morue (*Gadus morhua*) dans le sud du Golfe du Saint-Laurent (Canada) qui a subi une forte surexploitation non relevée pendant les années 1985–1992. Nous montrons à l'aide de simulations que le modèle avec censure détecte correctement les relevés de pêche problématiques. Les corrections sont peu sensibles à des bornes arbitraires imposées aux vraies prises et robustes à des modifications du modèle pour la simulation des relevés. Cependant, lorsque les enquêtes sont bruitées, le nouveau modèle avec prises censurées peut donner des résultats peu précis pour des paramètres utilisés pour la gestion, tels que le niveau du stock ou des mesures de stock par âge.

Introduction

The management of commercial fisheries and especially the determination of the total allowable catch (TAC) usually rely on information from population dynamics models. Cohort models are often used to assess stocks when an age structure can be established. Typically, observed data are total catches per year, catch proportions at age, and annual age-based abundance indices (Pope 1988) and other biological data on growth and maturation rates.

Common analytic assessment methods rely on the assumption that the catches are known without error. To be more precise, the catches are assumed to be measured without observation errors by harvesters and reported accurately to control authorities. This is the basis of the ADAPT method introduced by Gavaris (1988), which is commonly used when modelling Northwest Atlantic groundfish stocks and for extended survivor analysis (Shepherd 1999). Both of these approaches are versions of sequential population analysis (SPA; cf. Lassen and Medley 2000). However, in many cases, the reported landings are suspected to be lower than the true catches because of discards of less valuable fish, errors in reports, illegal fishing, control difficulties, etc. (Pitcher et al. 2002). For example, Punt et al. (2006) indicated that discard proportions in trawl fisheries could be as large as 50%. Note, however, that overreporting can also occur because of misidentifications (Walsh et al. 2007) or when fishermen want to secure track records or induce market distortion (Pearson 2001). Such errors will introduce bias in estimated population sizes derived from SPA. Mohn (1999), Cadigan and Farrell (2005), and Radomski et al. (2005), among others, have highlighted, through simulation and sensitivity studies, the limitations of ADAPT-type methods assuming that catches are measured without error.

Models with observation errors have been applied to catch-at-age data by Fournier and Archibald (1982), Gudmundsson (1986), and Aanes et al. (2007) and to total catch data by Kimura (1990), Patterson (1998), and Nielsen and Lewy (2001). Though the normal distribution has sometimes

Received 26 February 2009. Accepted 30 April 2010. Published on the NRC Research Press Web site at cjfas.nrc.ca on 23 July 2010.
J21081

Paper handled by Associate Editor Terrance Quinn II.

N. Bousquet,[1] T. Duchesne, and L.-P. Rivest. Department of Mathematics and Statistics, Pavillon Vachon, Université Laval, 1045 avenue de la Médicine, Québec, QC G1V 0A6, Canada.
N. Cadigan. Fisheries and Oceans Canada, P.O. Box 5667, St. John's, NL A1C 5X1, Canada.

[1]Corresponding author (e-mail: nicolas.bousquet@edf.fr).

been used (Lewy 1988), the lognormal distribution has been the most frequent choice for the catch observation error and survey data error. The models with observation errors assume that the reported landings, while uncertain, are unbiased with respect to the real catches.

The quality of stock assessment is directly linked to the quality of catch data; therefore, correcting underreported catches is important (Pitcher et al. 2002). Numerous correction approaches have been considered. For instance, misreported catches have been estimated by comparison of landings with fishmeal factory outputs (Castillo and Mendo 1987; Patterson et al. 1990; Pitcher and Stokes 1990). Another approach is to place observers onboard the fishing vessels to evaluate the discards and correct the official landings (Knuckey et al. 1999). Among others, Fox and Starr (1996) and Walsh et al. (2002, 2007) have compared reports from logbooks and observers to correct catch histories. Knowing the particular problems associated with a specific fishery such as its history and economic stakes, Pitcher et al. (2002) proposed corrections for potentially misreported catches. Punt et al. (2006) included data on discards in Bayesian population assessments, building hypotheses on the discarding pattern of fishing fleets. These approaches require additional data or information that will usually not be available for many stock assessments or for the entire time frame on which the stock is based.

In this article, we develop a methodology, based only on typically available data, to estimate the magnitude of underreporting when catches are measured with error. It thus can help to make up for the lack of data needed for direct corrections (usually obtained from observers). The methodology relies on two steps: first, a statistical detection of underreported catches, and then a correction of the catch history. Our approach is motivated by Hammond and Trenkel (2005), who considered catches as censored observations. "Censoring" refers to a technique commonly used in survival analysis and industrial reliability (Lawless 2003), which, however, seems not to have been considered much in fish stock assessment. A censored observation is one for which we do not know the exact value but about which we have partial information such as a lower bound (right censoring) or an interval known to cover the exact value (interval censoring).

Hammond and Trenkel (2005) proposed a censoring framework to model total landings (L_y^*) that are lower bounds for the total catches (C_y^*) that should have been observed with error and to constrain corrections ($C_y^* - L_y^*$) to reasonably low values. They assumed that the unknown value of C_y^* was in the interval $[L_y^*, 2L_y^*]$. These constraints were included in a Schaefer surplus production model, and parameters were estimated using Bayesian methods. They concluded that their censored model improved estimates of maximum sustainable yields and depletion indices, and that handling misreporting with censored catches can outperform some traditional assessment methods.

This paper extends the censoring framework to age-structured models with a particular focus on catch estimation and detection of misreporting. The model is developed specifically for the southern Gulf of St. Lawrence (SGSL; 4T in Fig. 1) cod stock. The assessment information is provided by Chouinard et al. (2006), and commercial catches are known to be underreported in some years. We consider censored total catches (i.e., summed over ages) because the age-composition estimates are thought to be sufficiently accurate for this stock. For example, cohorts can be clearly tracked over time. Pope (1988), Patterson (1998), and Cotter et al. (2004) also favored adding observation error to the total catch only because it is often reported independently of the landings age structure, which can be estimated more reliably.

The paper is organized as follows. First, we present the features of the models used for the assessment of the SGSL cod stock and provide the likelihood terms associated with censored catches. Focusing on underreporting, we assume that a censored observation C_y^* lays in an interval $[L_y^*, \beta_y L_y^*]$. It is assumed that stock experts can propose a reasonable value for β_y in most cases, but a less restrictive framework with $\beta_y = \infty$ is also considered. We use maximum likelihood estimation of the model parameters, with a simulated annealing algorithm to numerically maximize the likelihood and get standard errors. The effectiveness of incorporating censoring is highlighted in simulations using several censoring strategies. A heuristic method for the detection of underreporting is presented. The impact of catch corrections on SPA estimates of stock size is also discussed. Finally, extensions of the methodology to other stocks are considered in the Discussion.

Materials and methods

We consider the evolution of an exploited marine population, such as the SGSL cod stock, through years $y = 1, \ldots, Y$ and A age classes indexed by $a = 1, \ldots, A$. Without loss of generality, A can be an A+ class, aggregating the population for ages $\geq A$.

Three notations are used to differentiate the true catches, the observed catches, and the landings. First, denote the true total catches by $\mathbf{C} = (C_1, \ldots, C_Y)$, which are the exact numbers of fish taken out of the water. Next, denote the measured catches as $\mathbf{C}^* = (C_1^*, \ldots, C_Y^*)$, which are, in mean, equal to the true catches $\mathbf{C}$ but incorporate some observational errors. Finally, denote $\mathbf{L}^* = (L_1^*, \ldots, L_Y^*)$ as the total landings reported by the harvesters.

The framework

Observed catches are assumed to be linked to the real catches C_y through the typical observation equation

$$(1) \qquad C_y^* = C_y \exp(\nu_y), \quad y = 1, \ldots, Y$$

where $\nu_y \overset{\text{iid}}{\sim} \mathcal{N}(-\sigma^2/2, \sigma^2)$. Note that the mean $-\sigma^2/2$ is required so that $E[C_y^*] = C_y$. Let $\theta_1 = (C_1, \ldots, C_Y, \sigma^2)$ denote the set of $Y + 1$ parameters of primary interest. If the C_y^* are known, then the log-likelihood for θ_1 associated with eq. 1 is

$$(2) \qquad \ell_C(\theta_1; \mathbf{C}^*) = -Y \log \sigma - \frac{1}{2\sigma^2} \sum_{y=1}^{Y} \left(\log C_y^* - \log C_y + \frac{1}{2}\sigma^2 \right)^2$$

Because there are only Y independent observations, more data are required for a nontrivial and statistically valid assessment of θ_1, or the dimension of θ_1 must be reduced. Because it seems quite difficult and arbitrary to connect C_y to C_{y-1} through a deterministic relationship with a small number of parameters, finding additional data is more desirable. We use survey indices to define a second term in the log-likelihood.

We assume that annual age-based survey indices $\{I^*_{a,y}\}$ of stock size are related to stock abundance $N_{a,y}$ through the observation equation

$$(3) \qquad I^*_{a,y} = Q_a N_{a,y} \exp(\varepsilon_{a,y}), \quad y = 1, \ldots, Y,\ a = 1, \ldots, A$$

where $Q_1, \ldots, Q_A$ are catchability coefficients and $\varepsilon_{a,y} \overset{\text{iid}}{\sim} \mathcal{N}(0,\ \tau^2)$. From eq. 3, $E(I^*_{a,y}) = Q_a N_{a,y} \exp(\tau^2/2)$. We do not require $E(\varepsilon) = 1$ because the bias term $\exp(\tau^2/2)$ only affects the value of Q_a and not the estimation of the $N_{a,y}$ (Aanes et al. 2007). Let $\mathbf{I}^*$ denote the matrix of survey indices. The associated log-likelihood, ignoring the bias term, is

$$(4) \qquad \ell_I(\theta;\ \mathbf{I}^*) = -AY \log \tau - \frac{1}{2\tau^2} \sum_{A,Y} (\log I^*_{a,y} - \log Q_{\mathrm{a}} - \log N_{a,y})^2$$

where θ denotes the vector of all the parameters, including τ^2, the catchabilities, and all $N_{a,y}$. When the $N_{a,y}$ are modeled as functions of C_y and of some baseline abundances, the survey indices supplement the information in eq. 2 to estimate the C_y.

We assume that the abundance $N_{a,y}$ follows an exponential decline at a rate $Z_{a,y} = F_{a,y} + M_{a,y}$, where $F_{a,y}$ is the fishing mortality rate and $M_{a,y}$ is the natural mortality rate:

$$(5) \qquad N_{a+1,y+1} = N_{a,y} \exp(-Z_{a,y})$$

We make the common assumption that $M_{a,y}$ is known for all ages and years. We consider uncertainty about $M_{a,y}$ in the Discussion. The real catch-at-age $C_{a,y}$ is related to $N_{a,y}$ and $F_{a,y}$ through the Baranov equation:

$$(6) \qquad C_{a,y} = N_{a,y} \frac{F_{a,y}}{Z_{a,y}} \{1 - \exp(-Z_{a,y})\}, \qquad a = 1,\ \ldots,\ A-1,\ y = 1,\ \ldots,\ Y$$

An F constraint (Gavaris and Ianelli 2002) is used to reduce the number of parameters:

$$(7) \qquad F_{A,y} = \frac{1}{2}(F_{A-1,y} + F_{A-2,y}), \quad y = 1,\ \ldots,\ Y$$

The rationale behind this choice is the similarities in size between the fish of the oldest ages, which should imply similar fishing mortality. Pope's approximation (Pope 1972) of eq. 6 leads to

$$(8) \qquad N_{A,y} = \frac{C_{A,y} \exp(-M_{A,y}/2)}{1 - \exp(-F_{A,y})}, \quad y = 1,\ \ldots,\ Y$$

Hence, using eqs. 5, 6, and 8, all pairs ($N_{a,y}$, $F_{a,y}$) can be computed knowing the catches at age and the survivors in the last year, $N_{a,Y}$, $a = 1, \ldots,\ A - 1$.

Finally, the $C_{a,y}$ are obtained as the product of C_y and the catch proportions at age, $\rho_{a,y}$. As in Patterson (1998), we assume that the $\rho_{a,y}$ are estimated with negligible errors by survey and biological sampling, independently from the landings L^*_y. This hypothesis is reasonable for stocks with industrial fisheries such as the SGSL cod stock that have good port sampling of landed fish, which results in accurate estimation of the age composition of the catch. In the Discussion, we consider how this assumption can be relaxed. Hence, we reduce the parameter vector to $\theta = (\theta_1,\ \theta_2)$, where

$$(9) \qquad \theta_1 = (C_1,\ \ldots,\ C_Y,\ \sigma^2)$$

and

$$(10) \qquad \theta_2 = (N_{1,Y},\ \ldots,\ N_{A-1,Y},\ Q_1,\ \ldots,\ Q_A,\ \tau^2)$$

Correctly reported catches

Two notions of "correct" reporting can be considered, leading to two population models.

ADAPT model

Our first population model assumes that there is no measurement error, namely $C_y = C^*_y = L^*_y$. Thus $\sigma = 0$, θ is reduced to θ_2, and the assessment of biological quantities is made by maximizing $\ell_I(\theta;\ \mathbf{I}^*)$ only. Because the likelihood is derived from the lognormal, its maximization is similar to the least squares estimation promoted by Gavaris (1988) and Rivard and Gavaris (2003) under the name "ADAPT". Despite the increasing number of research articles that take into account catch measurement error, ADAPT remains one of the most widely used stock-assessment models (Lassen and Medley 2000; Patterson et al. 2001).

FA model

In the FA model, from Fournier and Archibald (1982), it is assumed that observed catches C^*_y are noisy but unbiased measures of the C_y and that there is no misreporting, i.e., $L^*_y = C^*_y$. The model is also defined by the parameter θ and the main term of the log-likelihood $\ell(\theta) = \ell_C(\theta_1;\ \mathbf{C}^* = \mathbf{L}^*) + \ell_I(\theta;\ \mathbf{I}^*)$.

Misreported catches

In the censored FA model, some L^*_y are lower bounds for C^*_y. Each reported landing L^*_y is associated with a doublet $(\alpha_y,\ \beta_y) \in IR_+ \times [\alpha_y,\ \infty)$ such that $\alpha_y L^*_y \le C^*_y \le \beta_y L^*_y$, where C^*_y is an unknown realization of the observed catches described in eq. 1. The variable C^*_y is said to be interval censored. If the catch is not misreported, then $\alpha_y = \beta_y = 1$. We focus here on underreporting and use $\alpha_y = 1$ and $\beta_y \ge 1$ throughout the paper.

In survival analysis, the contribution of an interval censored observation $(\alpha_y L^*_y, \beta_y L^*_y)$ to the likelihood is usually expressed as the probability that C^*_y belongs to $(\alpha_y L^*_y, \beta_y L^*_y)$, i.e., $F_{C^*_y}(\beta_y L^*_y) - F_{C^*_y}(\alpha_y L^*_y)$, where $F_{C^*_y}$ is the lognormal cumulative distribution function (cdf) of C^*_y emanating from eq. 1; see chapter 2 of Lawless (2003). To show that this is appropriate for underreported catches, assume that $C^*_y = u_y L^*_y$, where u_y follows the uniform distribution on $[\alpha_y,\ \beta_y]$. In other words, once α_y and β_y are specified, all values in the interval $(\alpha_y L^*_y, \beta_y L^*_y)$ are equally likely to be the true value of C_y^*. The likelihood term for a censored C_y^* is then

$$\frac{1}{\beta_y - \alpha_y}\int_{\alpha_y}^{\beta_y} f_{C_y^*}(u_y L_y^* | \alpha_y, \beta_y, L_y^*, u_y)\,\mathrm{d}u_y$$

where $f_{C_y^*}$ is the density of $F_{C_y^*}$. This is proportional to $F_{C_y^*}(\beta_y L_y^*) - F_{C_y^*}(\alpha_y L_y^*)$, the standard likelihood contribution for interval-censored variables presented above. Thus, the likelihood for censored catches can be constructed as a standard likelihood for interval-censored lifetime data. If $\beta_y = \infty$, then C_y^* is right-censored, as only a lower bound for its unknown value, $\alpha_y L_y^*$, is available. Its likelihood contribution then reduces to the survival function $1 - F_{C_y^*}(\alpha_y L_y^*)$.

Finally, denote $\mathbf{Y}_{[\mathrm{U}]}$ the subset of $\{1, \ldots, Y\}$ such that the $(C_y^*)_{y\in\mathbf{Y}_{[\mathrm{U}]}}$ are reported correctly, $\mathbf{Y}_{[\mathrm{U}]} = \{k \in \{1, \ldots, Y\}: \alpha_k = \beta_k = 1\}$, and denote Y_U the cardinal of $\mathbf{Y}_{[\mathrm{U}]}$. Similarly, denote $\mathbf{Y}_{[\mathrm{C}]} = \{1, \ldots, Y\}/\mathbf{Y}_{[\mathrm{U}]}$ the subset of years with misreported catches, and let $Y_C = Y - Y_U$ denote the number of censored catches. Then the log-likelihood of observed catches $\ell_C(\theta_1; \mathbf{C}^*)$, given in eq. 1, is replaced by

$$(11)\quad \ell_C(\theta; \mathbf{L}^* | \alpha_y, \beta_y) = \sum_{y\in\mathbf{Y}_{[\mathrm{C}]}} \log\left\{ \Phi\left(\frac{\log\beta_y L_y^* - \log C_y + \frac{1}{2}\sigma^2}{\sigma}\right) - \Phi\left(\frac{\log\alpha_y L_y^* - \log C_y + \frac{1}{2}\sigma^2}{\sigma}\right)\right\} - Y_U \log\sigma - \frac{1}{2\sigma^2}\sum_{y\in\mathbf{Y}_{[\mathrm{U}]}}\left(\log L_y^* - \log C_y + \frac{1}{2}\sigma^2\right)^2$$

with $\Phi(\cdot)$ being the cdf of the standard normal distribution.

Maximum likelihood estimation

Profile likelihoods

Consider the parameter vector θ whose component subvectors (θ_1, θ_2) are detailed in eqs. 9–10. We focus on obtaining the maximum likelihood estimator (MLE) of catches

$$\widehat{\theta}_1 = \arg\max_{\theta_1} \ell\left(\theta_1, \widehat{\theta}_2(\theta_1)\right)$$

where $\widehat{\theta}_2(\theta_1)$ is the profiled MLE obtained by maximizing eq. 4 for given θ_1. In other words, $\widehat{\theta}_2(\theta_1)$ is the ADAPT estimate of stock size obtained with the landings set equal to the catches $C = (C_1, \ldots, C_Y)$ in θ_1.

Denote by $\mathbf{S} = (N_{1,Y}, \ldots, N_{A-1,Y})$ the set of survivors. Recall that all the $N_{a,y}$ are functions of $(\mathbf{C}, \mathbf{S})$. Catchabilities and variances can be replaced in eq. 4 by their conditional MLE, leading to a profile loglikelihood $\ell_I^p(\mathbf{C}, \mathbf{S}; \mathbf{I}^*)$ based on the observed survey indices. The MLE of τ is indeed

$$(12)\quad \widehat{\tau}^2(\mathbf{C}, \mathbf{S}) = \frac{1}{AY}\sum_{a=1}^{A}\sum_{y=1}^{Y}\left(\log I_{a,y}^* - \log\widehat{Q}_a(\mathbf{C}, \mathbf{S}) - \log N_{a,y}(\mathbf{C}, \mathbf{S})\right)^2$$

As suggested by Patterson (1998), we add an estimation constraint on the series of catchabilities at age $(Q_1, \ldots, Q_A)$. To be biologically plausible, real catchability should be a smooth function of age, say $Q(a|\delta)$, depending on a parameter vector δ. Following Patterson (1998), we choose $Q(\cdot|\delta)$ as a natural cubic spline function. The rationale for this choice and additional details about the estimation of δ are given in Appendix A. Replacing such a nuisance parameter vector by ad hoc estimators in the likelihood is considered by Gong and Samaniego (1981). They called the resulting estimates pseudo maximum likelihood. They show that under suitable regularity conditions, pseudo and true maximum likelihood estimates have similar sampling properties. Thus we overlook this slight deviation from maximum likelihood estimation and consider

$$\ell_I^p(\mathbf{C}, \mathbf{S}; \mathbf{I}^*) = -\mathrm{AY}\log\widehat{\tau}^2(\mathbf{C}, \mathbf{S}) + \text{constant}$$

as the profile log-likelihood for the estimation for $(\mathbf{C}, \mathbf{S})$.

In a similar way, σ^2 can be profiled out of the catch log-likelihoods eqs. 2 and 11. In the uncensored framework $(C_y^* = L_y^*)$, a closed form expression for $\widehat{\sigma}^2(\mathbf{C})$ maximizing eq. 11 is given by

$$(13)\quad \widehat{\sigma}^2(\mathbf{C}) = 2\left\{\sqrt{1 + \frac{1}{Y}\sum_{y=1}^{Y}(\log L_y^* - \log C_y)^2} - 1\right\}$$

In a censored framework, a simple Newton–Raphson method starting from eq. 13 can be used to compute $\widehat{\sigma}^2(\mathbf{C})$. The profile catch log-likelihood is then

$$\ell_C^p(\mathbf{C}; \mathbf{L}^*) = \ell_\mathrm{C}(\mathbf{C}, \widehat{\sigma}^2(\mathbf{C}); \mathbf{L}^*)$$

Estimation of the FA parameters and their uncertainties

Our objective is to maximize the following profile log-likelihood:

$$(14)\quad \ell^p(\mathbf{C}, \mathbf{S}) = \ell_C^p(\mathbf{C}; \mathbf{L}^*) + \ell_\mathrm{I}^\mathrm{p}(\mathbf{C}, \mathbf{S}; \mathbf{I}^*)$$

Nelder–Mead simplex methods (Lange 2001) have been successfully used to fit ADAPT-type models, namely to estimate $\widehat{\mathbf{S}}(\mathbf{C})$ for a given vector $\boldsymbol{C}$. Starting points for the survivors may simply be chosen as multiples of the landings at age reported in year Y. In FA models, however, our experience has shown that a two-step procedure that first maximizes eq. 14 with respect to $\mathbf{S}$ for fixed $\mathbf{C}$ with the ADAPT algorithm and then maximizes $\ell^P(\mathbf{C}, \widehat{\mathbf{S}}(\mathbf{C}))$ with respect to $\mathbf{C}$ is numerically prohibitive. For this reason, total catches and survivors have been estimated simultaneously. Thus we redefine θ as $(\mathbf{C}, \mathbf{S})$ to simplify the notation, as the variance components σ^2 and τ^2 have been profiled out of eq. 14.

From a theoretical viewpoint, classic frequentist tools such as Hessian inversion (Patterson et al. 2001) for the estimation of parameter uncertainty are not relevant for a FA model. Indeed, the justification of Hessian inversion is based on the asymptotic normality of $\widehat{\theta}$; see Koenker and Machado (1999) for additional details. In the FA model, the dimension of θ increases linearly with the size Y of the data, which invalidates the direct application of classical asymptotic results to this model. Bootstrap methods have also been considered. However, these methods also require an asymptotic justification and were found to be computationally prohibitive for the SPA models that we consider.

We therefore adopted the following scheme. (*i*) Estimate $\widehat{\theta}$ using a Nelder–Mead simplex method. (*ii*) Attempt to improve $\widehat{\theta}$ with a sampling method that avoids local maxima and provides Monte Carlo estimates of uncertainty for calculating confidence intervals. We used simulated annealing (SA) for this purpose. Theoretical justifications can be found in Van Laarhoven and Arts (1987), and a didactic presentation with an overview of the main results is provided in Robert and Casella (2004, § 5.2.3). This sampling method is often used to find the mode of a posterior distribution in a Bayesian framework, for instance by Andrieu and Doucet (2000) and Doucet et al. (2002). Because the MLE is the posterior mode when all priors are uniform, the SA algorithm can be used to estimate θ and build confidence regions; this is described in Appendix A.

Illustration

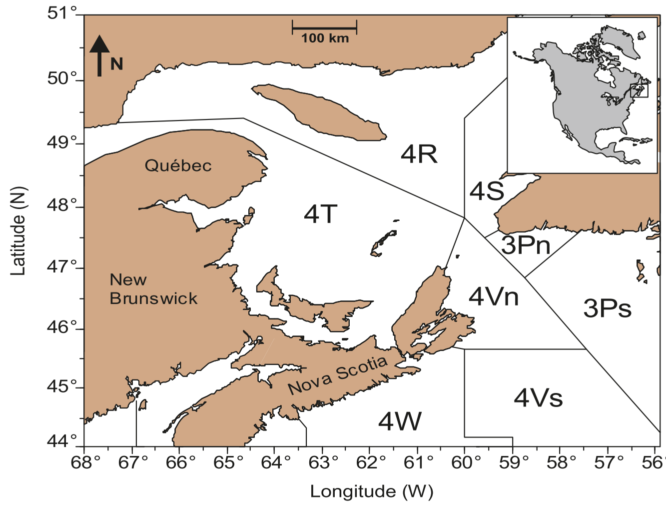
We fit the models to data from Chouinard et al. (2006) on the cod stock in the southern Gulf of St. Lawrence (SGSL, Canada, NAFO Subdivisions 4T and 4Vn; cf. Fig. 1) between 1971 and 2005 (Y = 35 years). This stock is recognized as having severely suffered from overfishing (Swain and Chouinard 2008). A simulation study is also carried out to investigate whether various censored FA models can identify underreporting periods and provide good estimates of real catches.

An assessment of SGSL cod catches

The SGSL cod stock has remained depleted since a fishing moratorium was imposed in 1993. The fishery was closed between September 1993 to May 1998. It was reopened with a TAC of 3000 tonnes (t) for an index fishery in 1998. The TACs were set at 6000 t from 1999 to 2002. The directed fishery was closed again in 2003 but was reopened with a TAC of 3000 t in 2004. In 2005, the TAC was increased to 4000 t. When the directed fishery was closed, removals were only due to scientific fishery. The collapse of this stock and others in the Northwest Atlantic in the early 1990s has been attributed to overfishing (Sinclair and Murawski 1997), in the sense that the TACs were exceeded and underreporting occurred. High overfishing for SGSL cod has been suspected during the years 1985–1992, just before the alarming decline of the resource. In addition, note that the cod population has suffered from an important increase of the natural mortality in the mid-1980s (Chouinard et al. 2006). In the model, the known rates of mortality are $M_{a,y}$ = 0.2 from 1971 to 1985 and $M_{a,y}$ = 0.4 from 1986 to 2005 for all ages, as suggested by Chouinard et al. (2006). We used the data reported in their paper for ages 2 to 15+. Thus A = 14 age classes were used in the model. The known catch proportions $\rho_{a,y}$ were calculated using the catch-at-age data of Chouinard et al. (2006).

We first fitted the ADAPT model with $L_y^* = C_y^* = C_y$ and the uncensored FA model with $L_y^* = C_y^*$, $y = 1, \ldots, Y$. The estimated catch history $\{C_y\}$ for the FA model is displayed (Fig. 2); for all practical purposes, it is equal to the reported landings $\mathbf{L}^*$, with a negligible standard deviation, $\widehat{\sigma} = 0.0043$. Because $\widehat{\tau}_{\mathrm{FA}} \simeq \widehat{\tau}_{\mathrm{ADAPT}} \simeq 0.68$, the fits of the uncensored FA model and of the ADAPT model are very similar. This high survey CV leads to a relatively imprecise

Fig. 1. Northwest Atlantic Fisheries Organization (NAFO) areas in the Gulf of St. Lawrence. Inset shows the study location off the eastern coast of Canada.



fit of the ADAPT-predicted indexes during the years 1980–1990 for which the observed indices are highly variable (Fig. 3). However, the magnitude of estimates in this period reflects an ADAPT overestimation of fish later caught between 1985 and 1989, resulting from catch underreporting. A continuous underreporting during the predecline period (1989–1991), associated with a pessimistic abundance underestimation, could explain the persistence of this underestimation after the decline.

Next we implemented the censored FA model with three censoring schemes, where a "censoring scheme" refers to the choice of censored years and corresponding β_y. We assumed α_y = 1 for all years. Scheme 1 (broken lines in Fig. 2) included three time periods of censoring: 1985–1992, 1993–1998, and 2003 for which β_y is set to 2, 1.25, and 1.25, respectively. Scheme 2 (dotted lines in Fig. 2) assumed censoring only for the years 1971–1992 with β_y = 2. Scheme 3 (dashed–dotted lines in Fig. 2) assumed censoring for all years with β_y = 2.

Three time periods are always distinguishable in the estimated catch histories. The censored estimated catches C_y fit the landings L_y^* for the years 1971–1984 (period P_1) and 1993–2005 (period P_3). However, during the period 1985–1992 (period P_2), we estimate $C_y \geq L_y^*$, and this difference appears more significant for the period 1988–1992. This is especially true under scheme 3. Furthermore, the estimates of σ and τ are similar for all schemes ($\widehat{\sigma} \simeq 0.072$ and $\widehat{\tau} \simeq 0.58$). The similarity of these results suggests that a crude censoring strategy has the ability to detect the years with underreported catches. The decrease of $\widehat{\tau}$ from 0.68 for the FA model to 0.58 for the censored FA model shows that adding censoring to the observation model for catches decreases the residual variance of the model for the indices by 27% (= 1 – $0.58^2/0.68^2$). Thus censored catches give the assumed SPA model an additional flexibility that permits the indices data $I_{a,y}^*$ to be fitted more closely. As will be shown in the simulations below, this makes the censored FA model heavily reliant on the accuracy of the survey indices.

Fig. 2. Reported landings and total catch estimates according to the uncensored and three censored FA models.

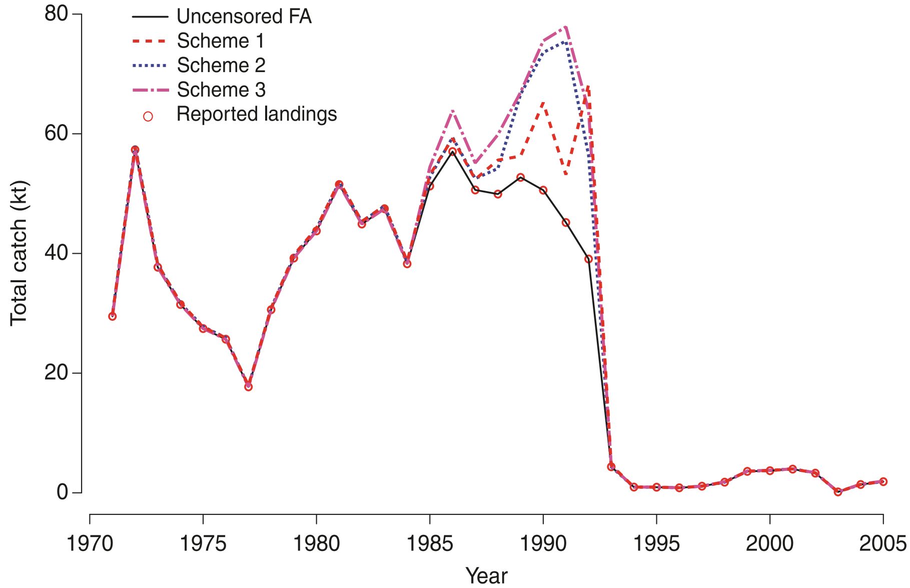
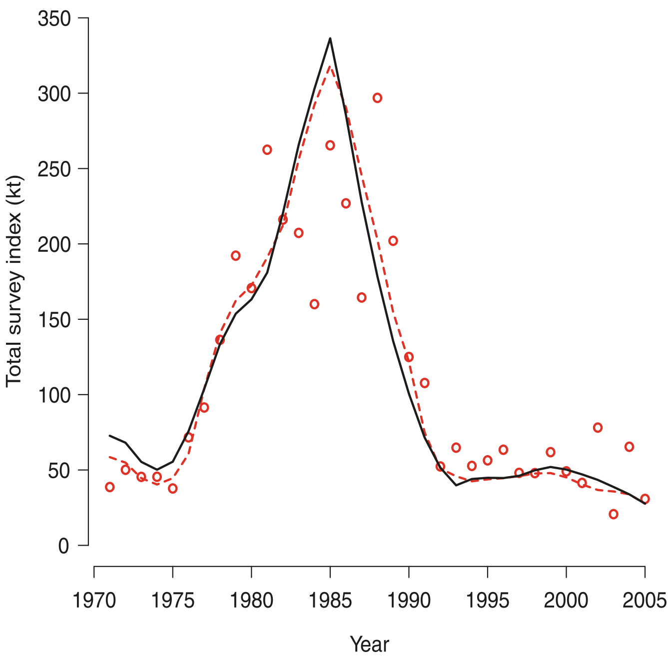


Fig. 3. Observed and predicted total survey indexes for the southern Gulf of St. Lawrence cod: ADAPT, continuous line; censored FA model, broken line.

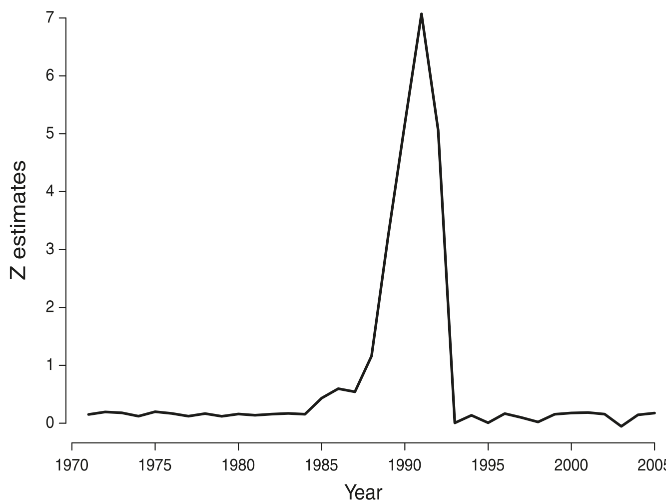


A heuristic detection of the underreported catches can been obtained with a time plot of

$$(15) \qquad Z_y = \frac{\log \widehat{C}_y - \log L_y^* + \frac{\widehat{\sigma}^2}{2}}{\widehat{\sigma}}$$

where $\widehat{C}_y$ and $\widehat{\sigma}$ are the MLEs of C_y and σ, respectively. Under the hypothesis that the L_y^* are not misreported, Z_y should approximatively follow a distribution with mean zero and unit variance. If the L_y^* are underreported during a certain period, the Z_y for this period should be significantly higher than during periods of correct reporting. The values of Z_y for the full-censored scheme 3 have been plotted (Fig. 4). The

Fig. 4. Evolution of values of Z_y estimated in the real case study by the full-censored FA model with $\beta_y = 2$ (scheme 3 in Fig. 2).



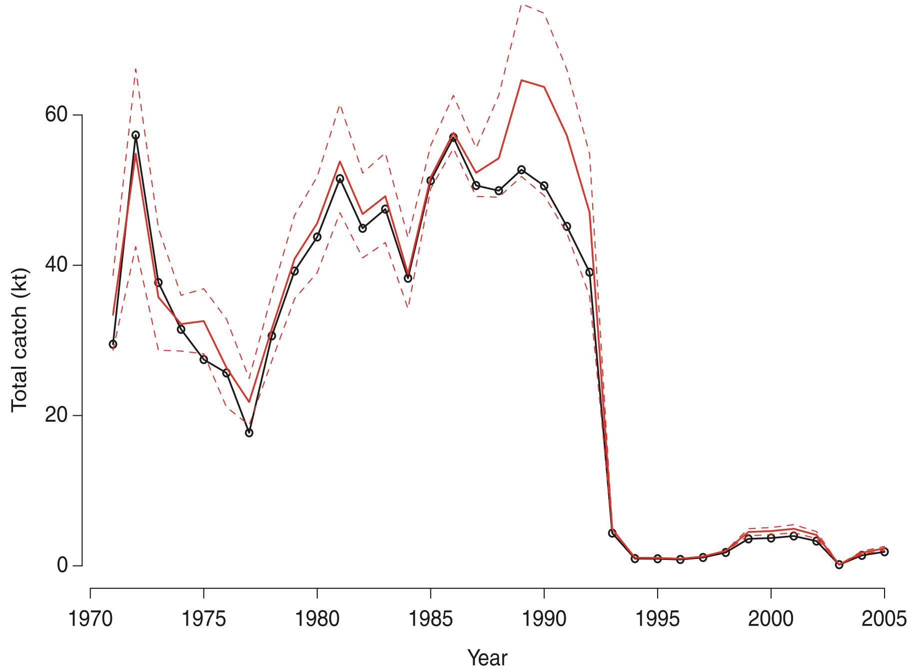
break point in the estimated catch history between years 1985 and 1992 appears quite clearly, in accordance with what was suspected to be the main period of underreporting.

For a reduced censoring scheme for 1985–1992 (period P_2) and assuming $\beta_y = 2$, the MLE of the total catches and an 80% confidence interval (estimated using the SA algorithm) are displayed (Fig. 5). Our motivation for using 80% was to show that the probability that true catches are equal or less than reported is only 10% during the period P_2.

Simulation study

The SGSL cod example suggests that incorporating censoring can lead to quantitative results that are in agreement with the experts' qualitative opinion about the real catches. We conducted a simulation study to investigate the accuracy of the catch estimators and of the biological indices used in fishing management. The Monte Carlo (MC) results are based on 80 data sets $\mathbf{X}^* = (\mathbf{I}^*, \mathbf{L}^*)$ independently simulated from the FA model with real catches, abundances, and

Fig. 5. Reported landings (line with circles) and MLE (shaded line) of total catches for the FA censored model (with confidence limits (broken lines)). Observed total catches for the years 1985–1992 are considered censored with $\beta_y = 2$.



catchabilities as estimated under the reduced censoring scheme at the end of the previous section (summarized in Fig. 5 and Appendix A, Table A1). Though 80 MC simulations may appear small, estimates of relative bias and root mean squared error (RMSE) stabilized after as few as 40 or 50 simulations. Four values (0.05 0.2, 0.4, 0.8) for the coefficient of variation (CV) of survey indices, defined here by $\sqrt{\exp(\tau^2) - 1}$, and two values (0.072, 0.2) for the CV of total catches ($\sqrt{\exp(\sigma^2) - 1}$) were used. The 0.05 value is realistic; a similar CV was observed by Chassot et al. (2009) for the cod stock in the northern Gulf of St. Lawrence (NGSL). Two schemes of censoring were used to simulate the observed landings. Denote by C_y^* a realization of the lognormal distribution (eq. 1).

In scheme A, the simulation model follows the assumptions made in the Material and methods section: a landing L_y^* for $y = 1985, \ldots, 1992$ is sampled from an inverse uniform distribution, i.e., $L_y^* = C_y^*/U_y$, where $U_y \sim \mathcal{U}[1, 2]$. In scheme B, the simulation model is different from the model in the Material and methods section and rather uses landings L_y^* for $y = 1985, \ldots, 1992$ given by $L_y^* = C_y^*/(1 + \kappa)$, for a fixed $\kappa \in [0, 1]$. Three values of κ (0.25, 0.5, and 0.75) are considered.

Comparing the model performance

Besides catches, the models are compared in terms of characteristics useful to fishery managers such as spawning stock biomass (SSB), recruitment (R), and fishing mortality rates (F). The SSB for year y (SSB_y) is defined as $\sum_{a=1}^{A} \omega_{a,y} p_{a,y} N_{a,y}$, where ω_a and $p_{a,y}$ are the mean weights and proportions mature at age a in year y, respectively, taken from tables 9 and 23 of Chouinard et al. (2006). The recruitment in year y is defined as the average size of the populations of biological ages 3 and 4, whereas the fishing rate in year y is taken as the average over all ages.

Relative biases and RMSEs are reported for these four sets of parameters. For instance, denote $\mathbf{C_{o,p}} = (C_{o,1}, \ldots, C_{o,p})$ a p vector of true catches for p of the Y years under study. Denote $\widehat{\mathbf{C}}_{\mathbf{p}}^{(i)} = (\widehat{C}_1^{(i)}, \ldots, \widehat{C}_p^{(i)})$ the MLE of these catches for the ith simulated data set. The performance of the estimator of $\mathbf{C_{o,p}}$ is measured using the (empirical) relative RMSE:

$$(16) \qquad \text{RMSE}(\mathbf{C_p}) = \frac{\sqrt{\sum_{j=1}^{p} M^{-1} \sum_{i=1}^{M} \left(\widehat{C}_j^{(i)} - C_{o,j}\right)^2}}{\sqrt{\sum_{j=1}^{p} C_{o,j}^2}}$$

and the mean relative bias

$$(17) \qquad B(\mathbf{C_p}) = \frac{1}{pM} \sum_{j=1}^{p} \sum_{i=1}^{M} \frac{\left(\widehat{C}_j^{(i)} - C_{o,j}\right)}{C_{o,j}}$$

where $M = 80$ is the number of simulated data sets. Thus when $B(\mathbf{C_p}) > 0$ [respectively, $B(\mathbf{C_p}) < 0$], the catches are overestimated [respectively, underestimated].

In the next section, the relative biases and RMSEs are aggregated for three periods: P_1 = 1971–1984, P_2 = 1985–1992, and P_3 = 1993–2005. Four models are fitted: ADAPT, CS_1 (the uncensored FA model), CS_2 (the censored FA model with known years of censoring), and CS_3 (the censored FA model that assumes censoring every year).

Simulation results

RMSE and relative biases obtained under scheme A are reported for four sets of parameters (Tables 1 and 2), as

Table 1. RMSE of total catches (C), annual fishing rates (F), spawning stock biomass (SSB), and recruitment (R), based on 80 simulated data sets for each combination of survey and catch CVs.

CV		ADAPT			CS_1			CS_2						CS_3					
								$\beta_y = 2$			$\beta_y = \infty$			$\beta_y = 2$			$\beta_y = \infty$		
Survey	Catches	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3
C																			
0.8	0.2	**21**	**37**	**20**	100	98	99	94	80	99	115	91	131	230	113	427	246	124	430
0.8	0.072	**13**	**34**	**8**	99	98	97	78	73	97	95	82	101	155	114	185	155	135	194
0.4	0.2	**21**	**37**	**20**	101	100	100	93	70	101	93	78	102	169	75	346	172	76	370
0.4	0.072	**12**	**34**	**8**	99	99	97	77	65	96	79	67	96	122	67	152	123	71	177
0.2	0.2	**21**	**37**	**21**	100	99	101	92	48	101	92	48	102	137	62	274	172	76	370
0.2	0.072	**12**	**34**	**7**	99	100	101	82	53	99	83	57	103	106	62	132	110	63	163
F																			
0.4	0.2	**14**	**23**	**16**	99	98	99	89	93	101	101	100	105	136	105	134	139	120	143
0.4	0.072	**11**	**16**	**12**	100	102	101	86	83	99	94	92	99	131	101	136	137	117	137
0.2	0.2	**12**	**23**	**15**	97	96	99	91	63	96	97	69	101	101	65	102	108	72	108
0.2	0.072	**10**	**21**	**13**	100	104	100	90	60	98	92	64	99	95	61	108	99	71	107
0.05	0.2	**10**	**22**	**14**	96	100	97	82	57	91	82	61	92	81	59	101	85	62	103
0.05	0.072	**9**	**21**	**13**	95	105	96	78	55	91	79	56	91	82	57	100	83	59	104
SSB																			
0.4	0.2	**14**	**22**	**11**	103	101	115	158	165	123	160	171	132	232	227	247	240	221	248
0.4	0.072	**11**	**14**	**7**	102	98	105	150	140	115	153	139	121	225	209	226	231	213	245
0.2	0.2	**11**	**17**	**7**	100	98	111	131	92	104	138	99	110	224	163	324	235	168	332
0.2	0.072	**10**	**15**	**5**	102	99	109	127	84	98	132	86	102	211	156	328	221	160	334
0.05	0.2	**8**	**17**	**3**	100	99	105	99	62	96	102	64	95	155	81	232	153	83	232
0.05	0.072	**8**	**15**	**3**	101	99	103	101	57	91	99	58	94	149	73	221	151	72	224
R																			
0.4	0.2	**14**	**19**	**12**	99	102	128	158	124	141	159	127	150	227	187	218	234	187	224
0.4	0.072	**13**	**15**	**10**	99	100	131	151	112	130	155	122	143	230	183	202	228	185	211
0.2	0.2	**13**	**17**	**6**	99	98	109	118	83	122	111	89	116	189	131	240	195	136	254
0.2	0.072	**12**	**17**	**6**	100	99	115	102	68	98	106	72	101	181	125	227	187	133	235
0.05	0.2	**10**	**17**	**5**	96	97	107	101	59	114	103	64	115	157	88	189	162	85	194
0.05	0.072	**10**	**17**	**4**	98	99	105	92	51	109	93	54	104	144	68	177	141	71	181

Note: RMSEs are reported separately for the periods P_1 = 1971–1984, P_2 = 1985–1992, and P_3 = 1993–2005. ADAPT RMSEs are in bold type. Relative RMSEs with respect to ADAPT RMSEs, in percentage, are given in Roman type. Results are presented for the uncensored ADAPT and CS_1 models, for models with a known (CS_2) and an unknown (CS_3) underreporting period, and for interval-censored ($\beta_y = 2$) and right-censored ($\beta = \infty$) catches. Underreported landings are simulated according to scheme A.

Table 2. Relative bias estimates in percentage for total catches (C), annual fishing rates (F), spawning stock biomass (SSB), and recruitment (R), based on 80 data sets for each combination of survey and catch CVs.

								CS_2						CS_3					
CV		ADAPT			CS_1			$\beta_y = 2$			$\beta_y = \infty$			$\beta_y = 2$			$\beta_y = \infty$		
Survey	Catches	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3
C																			
0.8	0.2	–2	–29	1	–2	–28	1	1	4	1	3	7	5	15	8	21	15	10	25
0.8	0.072	–2	–27	1	–2	–25	–1	0	–7	–1	1	7	1	14	9	18	18	9	23
0.4	0.2	–2	–30	1	–2	–30	1	0	–3	2	0	–1	2	12	–8	16	16	–7	21
0.4	0.072	–2	–27	0	–2	–25	0	0	–6	0	0	–	1	13	–3	16	13	–4	20
0.2	0.2	–2	–29	1	–2	–29	–1	0	–4	1	0	–4	2	10	–9	13	13	–9	17
0.2	0.072	–2	–30	1	–2	–30	1	0	–7	1	0	–7	1	10	–4	14	11	–7	17
F																			
0.4	0.2	2	0	18	1	–1	16	–2	–1	12	–3	–1	13	0	–2	31	0	–2	33
0.4	0.072	2	0	17	1	0	15	–2	–2	13	–2	–1	14	0	–2	31	1	–2	32
0.2	0.2	1	0	16	1	–1	14	–2	–1	15	–1	–1	11	–1	–2	29	0	–1	32
0.2	0.072	1	4	18	0	3	16	–1	–2	12	0	–1	13	0	–1	27	1	0	29
0.05	0.2	1	–1	11	1	–1	10	–1	0	7	–1	–1	5	–2	–2	14	0	–1	15
0.05	0.072	1	–2	9	0	–1	9	0	0	6	0	–1	6	–1	–1	12	–1	–2	12
SSB																			
0.4	0.2	6	–12	–2	5	–11	–2	5	9	6	6	9	6	12	16	13	13	16	21
0.4	0.072	5	–12	–1	5	–11	0	5	8	6	5	7	5	12	15	12	13	15	18
0.2	0.2	4	–12	–1	3	–11	–1	4	4	3	4	4	4	6	5	11	7	5	13
0.2	0.072	4	–15	–4	4	–14	–4	5	4	2	5	5	3	5	5	8	6	6	12
0.05	0.2	2	–12	–2	2	–11	–1	3	4	2	3	4	2	4	5	8	5	5	10
0.05	0.072	2	–11	–3	2	–10	–2	1	3	2	2	3	2	3	4	7	6	4	9
R																			
0.4	0.2	4	–6	–3	4	–5	–3	7	6	9	7	6	10	10	9	19	12	11	22
0.4	0.072	3	–5	–2	3	–5	–1	7	5	8	7	6	9	9	8	16	11	10	19
0.2	0.2	3	–6	–1	2	–5	–1	4	4	5	4	5	5	6	6	7	7	6	7
0.2	0.072	4	–8	–2	4	–6	–2	5	4	4	5	5	4	6	4	6	6	5	6
0.05	0.2	4	–7	–1	4	–6	–1	4	4	4	5	4	4	5	5	6	5	6	7
0.05	0.072	2	–8	–2	3	–8	–1	3	4	4	4	5	3	5	4	6	5	4	7

Note: Findings are reported separately for the periods P_1 (1971–1984), P_2 (1985–1992), and P_3 (1993–2005). Results are presented for the uncensored ADAPT and CS_1 models, for models with a known (CS_2) and an unknown (CS_3) underreporting period, and for interval-censored ($\beta_y = 2$) and right-censored ($\beta = \infty$) catches. Underreported landings are simulated according to scheme A.

Table 3. RMSE (%) and relative bias (in parentheses) of total catch estimation over the complete period 1971–2005.

CV					CS_2		CS_3	
Survey	Catches	κ	ADAPT	CS_1	$\beta_y = 2$	$\beta_y = \infty$	$\beta_y = 2$	$\beta_y = \infty$
0.8	0.072	0.25	**16.8** (–4.7)	99 (–4.3)	154 (2.3)	204 (3.1)	255 (25.4)	258% (25.5)
		0.5	**26.6** (–7.7)	99 (–7.4)	82 (1.1)	141 (2.1)	162 (21.8)	165 (22.3)
		0.75	**33.8** (–9.9)	99 (–9.6)	51 (0.2)	122 (1.7)	135 (19.1)	138 (19.2)
0.4	0.072	0.25	**17.0** (–4.7)	98 (–4.4)	154 (–1.6)	159 (–1.5)	130 (18.5)	130 (18.8)
		0.5	**26.7** (–7.7)	99 (–7.4)	75 (–2.5)	80 (–2.3)	91 (16.8)	91 (16.9)
		0.75	**33.7** (–9.9)	99 (–9.6)	60 (–3.0)	72 (–2.7)	69 (15.2)	69 (15.5)
0.2	0.072	0.25	**16.9** (–4.8)	99 (–4.6)	80 (–2.6)	82 (–2.4)	92 (14.6)	92 (14.3)
		0.5	**26.6** (–7.8)	97 (–7.6)	65 (–3.4)	65 (–3.3)	66 (16.5)	67 (16.7)
		0.75	**33.7** (–10.0)	100 (–9.8)	58 (–3.8)	59 (–3.7)	53 (16.0)	54 (16.2)
0.8	0.2	0.25	**23.5** (–4.8)	100 (–4.5)	128 (2.6)	163 (3.4)	184 (25.5)	184 (25.6)
		0.5	**30.8** (–7.8)	98 (–7.6)	85 (1.0)	141 (2.7)	143 (21.9)	145 (21.8)
		0.75	**36.8** (–9.9)	99 (–9.7)	63 (–0.3)	127 (1.9)	126 (20.6)	129 (21.2)
0.4	0.2	0.25	**23.8** (–5.0)	99 (–4.6)	89 (–1.4)	91 (–1.2)	108 (19.7)	109 (19.9)
		0.5	**30.8** (–8.0)	100 (–7.6)	76 (–2.6)	78 (–2.1)	83 (17.1)	83 (18.0)
		0.75	**36.9** (–10.0)	100 (–9.7)	65 (–3.0)	66 (–2.5)	67 (16.2)	69 (16.4)
0.2	0.2	0.25	**23.7** (–5.0)	100 (–4.8)	83 (–2.2)	84 (–2.0)	84 (19.6)	84 (19.6)
		0.5	**30.8** (–8.1)	98 (–7.9)	69 (–3.4)	71 (–3.1)	68 (17.9)	72 (18.4)
		0.75	**36.9** (–10.2)	100 (–10.0)	62 (–3.8)	63 (–3.6)	58 (18.1)	61 (18.2)

Note: ADAPT RMSEs are in bold type. Relative RMSEs with respect to ADAPT RMSEs, in percentage, are given in Roman type. Results are presented for the uncensored ADAPT and CS_1 models, for models with a known (CS_2) and an unknown (CS_3) underreporting period, and for interval-censored ($\beta_y = 2$) and right-censored ($\beta = \infty$) catches. Underreported landings are simulated according to scheme B, indexed by κ.

well as the estimation of the catches under scheme B (Table 3). The censored FA models did not give reasonable estimates of SSB, R, and F for a survey CV of 0.8; these results are not presented. In a similar way, the censored FA models estimated the catches very well, with a survey CV of 0.05; these are omitted. Thus, three survey CVs are considered for each set of parameters.

ADAPT and CS_1 estimates always have similar properties. They perform reasonably well when there is little or no underreporting. Even if CS_1 allows observation errors, it does not do better than ADAPT. The CS_2 model, with β_y = 2, gives the best results for both bias and RMSE among all censored models. For estimating catches (C) and fishing mortality rates (F), it also outperforms ADAPT and CS_1. For both CS_2 and CS_3, the estimates obtained with $\beta_y = \infty$ have larger RMSEs, especially in the most extreme case where the survey CV = 0.8 and the catch CV = 0.072. However, in general, going from $\beta_y = 2$ to $\beta_y = \infty$ when estimating the parameters does not change the results much.

Model CS_3 performs better than both ADAPT and CS_1 for catch estimation in the period of underreporting, except when the survey CV is 0.8. However, the CS_3 model is sensitive but not specific: if there is underreporting, it will flag it, but it can show underreporting in years when there is none. Note also that CS_2 outperforms ADAPT and CS_1 in bias and RMSE for estimating SSB_y in the period of underreporting if the survey CV is at most 0.2.

Large biases have been highlighted (Table 2). In all of the models, the absolute abundance estimates SSB and R are driven by the catches. The uncensored model underestimates both in the years of underreporting. Allowing censored catches gives larger SSB and R values; thus, all the censored models overestimate SSB and R. All of the models overestimate the fishing mortality rates in the third period. This might be an artefact of the small size of the stock in that period. The SGSL cod stock has an estimated survey CV of 0.63; thus, the censored FA model is unlikely to yield better estimates of the age-specific abundances $N_{a,y}$ and of SSB and R than the ADAPT or the uncensored FA model.

To check the ability of Z_y given by eq. 15 to detect underreported catches, we displayed yearly box plots of the 80 replicates of this statistic obtained with the CS_3 model with $\beta_y = 2$, for data sets generated with a survey CV of 0.4 and a catch CV of 0.072 (Appendix A, Fig. A2). The break points between 1985 and 1992 are clearly visible. This agrees with the results obtained in the SGSL case study. The lack of specificity noticed earlier shows up in this plot: all the Z_y values are positive and suggest some small underreporting at the beginning and end of the observation period.

Results on the estimation of the catches when the censoring scheme used for the simulation differs from the scheme used for the estimation are given (Table 3). Model CS_2 provides nearly unbiased catch estimates, even if the survey CV is large. The censored model appears to be robust to the specification of the censoring mechanism for estimating catches. Indeed, the RMSEs and the biases of ADAPT and the CS_1 model are higher than those of the censored models, except when κ = 0.25 (Table 3). Note that in the latter case, the landings are close to the true catches. Even if it overestimates the catches by 10% to 20%, model CS_3 outperforms both ADAPT and CS_1 in most cases, provided that the survey CV is equal to or lower than 0.4. In addition, when the survey CV is under 0.8, taking $\beta_y = 2$ and $\beta_y = \infty$ yields similar results. Thus, the subjective selection of a β_y value can be avoided by systematically taking $\beta_y = \infty$.

Discussion

In this article, we extended the censoring of catches introduced by Hammond and Trenkel (2005) to age-structured models. We considered the true catches as unknown model parameters and landings as lower bounds for noisy catch measurements. We built a model for the SGSL cod stock and showed, through a simulation study, that it can reconstruct the true catch history when the landings are underreported.

This approach involves two steps. From simulations based on a real case study, we showed that crudely censored likelihoods are able to detect underreported catches when the natural mortality is assumed known. Full-censored strategies are diagnostic tools. Once problematic catches are identified, additional information can be sought to determine a more precise censoring scheme. This can lead to improved estimation of the true catches. We also succeeded in getting catch estimates that are better, in mean squared error, than simulated landings and uncensored estimates. Even when the model for the censoring of the landings is incorrect (scheme B), censored methods still provided reliable estimates.

Extensions of the model

We developed a model suitable for the SGSL cod stock assessment. In the rest of this section, we outline some generalizations that could make the model useful for other stocks.

First, for long-lived and not intensively fished species, approximation (eq. 8) is inappropriate and can be replaced by the standard plus-group equation:

$$N_{A,y+1} = N_{A,y} \exp(-Z_{A,y}) + N_{A-1,y} \exp(-Z_{A-1,y})$$

Second, the $I^*_{a,y}$ likelihood assumed that the estimated indices were all independent. A more realistic model might have correlated survey indices within years, as discussed in Myers and Cadigan (1995). This only impacts the survey likelihood. Following Chassot et al. (2009), one then has to define one likelihood term for the total survey indexes (summed over ages) and another term for the proportions at age in the population. In addition, the behavior of the proposed censored FA model when $I^*_{a,y}$ are missing for some values of a or y should be investigated.

Another important issue is age-specific underreporting, which can be due to highgrading. This can be accommodated by partitioning the age classes into two sets — one with misreporting (set 1) and one without (set 2). The likelihood would then be written in terms of the yearly landings in the two sets, $L_y^{*(1)}$ and $L_y^{*(2)}$. Censoring is included only for $L_y^{*(1)}$.

The assumption of known values of $\rho_{a,y}$ could be relaxed by adding a term to the likelihood for the estimation of these parameters using either the landings-at-age data or some biological sampling data. Likelihoods for complex sampling schemes used for size and age information (i.e., length-stratified age samples, cluster samples) are complex, and in practise, simple approximations are used (e.g., Quinn and Deriso 1999).

The treatment of the natural mortality rate M could also be refined. If the real natural mortality rates are suspected to exceed the fixed values chosen in this paper, a part of the estimated catches may actually represent natural deaths, which should be removed before correcting the landings. Therefore, preliminary studies must be made to provide reliable estimates of M using survey data, as in Sinclair (2001). Otherwise, a separability assumption may probably be necessary to estimate M.

Finally, terms can be added to the likelihood to account for other sources of uncertainty in the assessment method. Each additional term makes the likelihood maximization more complex. More complicated models could be implemented in a Bayesian framework, where computer-intensive estimation methods can be used for parameter estimation. One interest of the relatively simple model that we proposed for the SGSL cod stock is the ability to carry out inference in a standard frequentist framework, which might be preferable when sources of reliable prior information are unavailable, as might be the case in stock assessment.

Comparison with other methods and future work

The censored approach may provide advantages when compared with alternatives for population modelling and catch correction such as the BADAPT model (Darby 2005). The latter model requires a time period of reliable catch data and survey data that overlap the period of reliable catch data and the period in which misreporting occurs. The censored FA model only requires crude bounds on the true catches. This is especially interesting for modelling many species with large historic catches for which important misreporting was unlikely for many years, because it was not possible to blackmarket or discard large amounts of catch. These years can be given β values only slightly larger than 1.

The relatively poor performance of the censored FA model when estimating management parameters needs further investigation. The solution might be to use simultaneously several abundance indices and also to constrain interannual variations in fishing mortality within realistic values, perhaps using a random walk approach.

Changes to the estimation procedure itself could also be investigated. Alternatives to the SA algorithm could be considered. The threshold acceptance strategies of Dueck and Scheuer (1990) could improve its convergence. In addition, adapting the SA algorithm to assessment models that feature a random process error is a challenging problem. The advances proposed by Andrieu and Doucet (2000) and de Valpine and Hilborn (2005) seem promising in this regard.

Acknowledgements

We thank Douglas P. Swain, Ghislain A. Chouinard, Roderick Morin, and Hugues Benoît from Fisheries and Oceans Canada (Moncton, New Brunswick) for providing helpful comments and references about the SGSL cod case study. We also thank A.E. Punt, an anonymous reviewer, and the Associate Editor for their comments that greatly helped to improve the paper. This work was supported by the ARC Program between the Fisheries and Oceans Canada, Université Laval, and the Canada Research Chair in Statistical Sampling and Data Analysis.

References

Aanes, S., Engen, S., Saether, B.-E., and Aanes, R. 2007. Estimation of the parameters of fish stock dynamics from catch-at-age data and indices of abundance: can natural and fishing mortality be separated? Can. J. Fish. Aquat. Sci. **64**(8): 1130–1142. doi:10.1139/F07-074.

Andrieu, C., and Doucet, A. 2000. Simulated annealing for maximum *a posteriori* parameter estimation of hidden Markov models. IEEE Trans. Inf. Theory, **46**(3): 994–1004. doi:10.1109/18.841176.

Cadigan, N., and Farrell, P.J. 2005. Local influence diagnostics for the retrospective problem in sequential population analysis. ICES J. Mar. Sci. **62**(2): 256–265. doi:10.1016/j.icesjms.2004.11.015.

Castillo, S., and Mendo, J. 1987. Estimation of unregistered Peruvian anchoveta (*Engraulis ringens*) in official catch statistics, 1951–1982. *In* The Peruvian anchoveta and its upwelling ecosystem: three decades of change. *Edited by* D. Pauly and I. Tsukayama. ICLARM Studies and Reviews No. 15. pp. 109–116.

Chassot, E., Duplisea, D., Hammill, M., Caskenette, A., Bousquet, N., Lambert, Y., and Stenson, G. 2009. Role of predation by harp seals *Pagophilus groenlandicus* in the collapse and non-recovery of northern Gulf of St. Lawrence cod *Gadus morhua*. Mar. Ecol. Prog. Ser. **379**: 279–297. doi:10.3354/meps07897.

Chouinard, G.A., Currie, L., Poirier, G.A., Hurlbut, T., Daigle, D., and Savoie, L. 2006. Assessment of the southern Gulf of St. Lawrence cod stock. DFO Can. Sci. Advis. Sec. Res. Doc. No. 2006/023.

Cotter, A.J.R., Burt, L., Paxton, C.G.M., Fernandez, C., Buckland, S.T., and Pan, J.-X. 2004. Are stock assessment methods too complicated? Fish Fish. **5**: 235–254. doi:10.1111/j.1467-2679.2004.00157.

Darby, C.D. 2005. Appendix 4: estimating systematic bias in the North Sea cod landings data. *In* Report on the assessment of demersal stocks in the North Sea and Skaggerrak, 7–16 September 2004, Bergen, Norway. ICES CM 2005/ACFM:07. pp. 693–713.

de Valpine, P., and Hilborn, R. 2005. State–space likelihoods for nonlinear fisheries time-series. Can. J. Fish. Aquat. Sci. **62**(9): 1937–1952. doi:10.1139/f05-116.

Doucet, A., Godsill, S.J., and Robert, C.P. 2002. Marginal maximum *a posteriori* estimation using Markov chain Monte Carlo. Stat. Comput. **12**(1): 77–84. doi:10.1023/A:1013172322619.

Dueck, G., and Scheuer, T. 1990. Threshold accepting: a general purpose optimization algorithm appearing superior to simulated annealing. J. Comp. Physiol. **90**: 161–175. doi:10.1016/0021-9991(90)90201-B.

Fournier, D., and Archibald, C.P. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. **39**(8): 1195–1207. doi:10.1139/f82-157.

Fox, D.S., and Starr, R.M. 1996. Comparison of commercial fishery and research catch data. Can. J. Fish. Aquat. Sci. **53**: 2681–2694. doi:10.1139/cjfas-53-12-2681.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. Res. Doc. No. 88/29.

Gavaris, S., and Ianelli, J.N. 2002. Statistical issues in fisheries' stock assessment. Scand. J. Stat. **29**(2): 245–267. doi:10.1111/1467-9469.00282.

Gong, G., and Samaniego, F.J. 1981. Pseudo maximum likelihood estimation: theory and applications. Ann. Stat. **9**(4): 861–869. doi:10.1214/aos/1176345526.

Gudmundsson, G. 1986. Statistical considerations in the analysis of catch-at-age observations. ICES J. Mar. Sci. **43**(1): 83–90. doi:10.1093/icesjms/43.1.83.

Hammond, T.R., and Trenkel, V.M. 2005. Censored catch data in fisheries stock assessment. ICES J. Mar. Sci. **62**: 1118–1130. doi:10.1016/j.icesjms.2005.04.015.

Kimura, D. 1990. Approaches to age-structured separable sequential population analysis. Can. J. Fish. Aquat. Sci. **47**(12): 2364–2374. doi:10.1139/f90-263.

Knuckey, I.A., Grieve, C., and Smith, D.C. 1999. Evolution of the integrated scientific monitoring programme in Australia's South East Fishery. *In* Proceedings of the International Conference on Integrated Fisheries Monitoring, Sydney, Australia, 1–5 February 1999. FAO, Rome. pp. 231–248.

Koenker, R., and Machado, J.A.F. 1999. GMM inference when the number of moment conditions is large. J. Econom. **93**(2): 327–344. doi:10.1016/S0304-4076(99)00014-7.

Lange, K. 2001. Numerical analysis for statisticians. Springer, New York.

Lassen, H., and Medley, P. 2000. Virtual population analysis: a practical manual for stock assessment. FAO Fish. Tech. Pap. No. 400, Rome, Italy.

Lawless, J.F. 2003. Statistical models and methods for lifetime data. 2nd ed. John Wiley and Sons, Hoboken, New Jersey.

Lewy, P. 1988. Integrated stochastic virtual population analysis: estimates and their precision of fishing mortalities and stock sizes for the North Sea whiting stock. ICES J. Mar. Sci. **44**(3): 217–228. doi:10.1093/icesjms/44.3.217.

Mohn, R. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES J. Mar. Sci. **56**(4): 473–488. doi:10.1006/jmsc.1999.0481.

Myers, R.A., and Cadigan, N.G. 1995. Statistical analysis of catch-at-age data with correlated errors. Can. J. Fish. Aquat. Sci. **52**: 1265–1273. doi:10.1139/f95-123.

Nielsen, N.A., and Lewy, P. 2001. Comparison of the frequentist properties of Bayes and the maximum likelihood estimators in an age-structured fish stock assessment model. Can. J. Fish. Aquat. Sci. **59**(1): 136–143. doi:10.1139/f01-204.

Patterson, K.R. 1998. Assessing fish stocks when catches are misreported: model, simulation tests, and application to cod, haddock, and whiting in the ICES area. ICES J. Mar. Sci. **55**(5): 878–891. doi:10.1006/jmsc.1998.0351.

Patterson, K.R., Castello, M., Santos, M., and Barcia, M. 1990. Validation and recomputation of catch and effort data in the near-shore fleet of Ecuador for the period 1981–1989. Report to INP (El Instituto Nacional de Pesca), Guayaquil, Ecuador. Available from www.inp.gov.ec [in Spanish].

Patterson, K.R., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., Punt, A., Restrepo, V., Skagen, D.W., and Stefánsson, G. 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish Fish. **2**: 125–157. doi:10.1046/j.1467-2960.2001.00042.x.

Pearson, H. 2001. China caught out as model shows net fall in fish. Nature (London), **414**(6863): 477. doi:10.1038/35107216. PMID:11734811.

Pitcher, T.J., and Stokes, T.K. 1990. Assessment of Ecuadorean commercial pelagic fisheries. Contract Report to the Overseas Development Administration, United Kingdom. Marine Resources Assessment Group, Imperial College, London.

Pitcher, T.J., Watson, R., Forrest, R., Valtýsson, H.P., and Guénette, S. 2002. Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish Fish. **3**: 317–339. doi:10.1046/j.1467-2979.2002.00093.x.

Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. Res. Bull. Int. Comm. Northw. Atl. Fish. **9**: 65–74.

Pope, J.G. 1988. Collecting fisheries assessment data. *In* Fish population dynamics. 2nd ed. *Edited by* J.A. Gulland. John Wiley, Chichester, UK. pp. 63–82.

Punt, A.E., Smith, D.C., Tuck, G.N., and Methot, R.D. 2006. Including discard data in fisheries stock assessments: two case studies from south-eastern Australia. Fish. Res. **79**(3): 239–250. doi:10.1016/j.fishres.2006.04.007.

Quinn, T.J., II, and Deriso, R.B. 1999. Quantitative fishery dynamics. Oxford University Press, New York.

Radomski, P., Bence, J.R., and Quinn, T.J., II. 2005. Comparison of virtual population analysis and statistical kill-at-age analysis for a recreational, kill-dominated fishery. Can. J. Fish. Aquat. Sci. **62**(2): 436–452. doi:10.1139/f04-228.

Rivard, D., and Gavaris, S. 2003. Appendix 3: St. Andrews (S. Gavaris) version of ADAPT. Estimation of population abundance. NAFO (Northwest Atl. Fish. Organ.) Sci. Counc. Stud. **36**: 201–249.

Robert, C.P., and Casella, G. 2004. Monte Carlo statistical methods. 2nd ed. Springer, New York.

Shepherd, J.G. 1999. Extended survivors analysis: an improved method for the analysis of catch-at-age data and abundance indices. ICES J. Mar. Sci. **59**: 592–605. doi:10.1016/j.fishres.2008.12.021.

Sinclair, A.F. 2001. Natural mortality of cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. ICES J. Mar. Sci. **58**(1): 1–10. doi:10.1006/jmsc.1999.0490.

Sinclair, A.F., and Murawski, S.A. 1997. Why have groundfish stocks declined? *In* Northwest Atlantic groundfish: perspectives on a fishery collapse. *Edited by* J. Boreman, B.S. Nakashima, J.A. Wilson, and R.L. Kendall. American Fisheries Society, Bethesda, Maryland. pp. 71–93.

Swain, D.P., and Chouinard, G.A. 2008. Predicted extirpation of the dominant demersal fish in a large marine ecosystem: Atlantic cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. Can. J. Fish. Aquat. Sci. **65**(11): 2315–2319. doi:10.1139/F08-175.

Van Laarhoven, P.J., and Arts, E.H.L. 1987. Simulated annealing: theory and applications. Reidel, Amsterdam.

Walsh, W.A., Kleiber, P., and McCracken, M. 2002. Comparison of logbook reports of incidental blue shark catch rates by Hawaii-based longline vessels to fishery observer data by application of a generalized additive model. Fish. Res. **58**(1): 79–94. doi:10.1016/S0165-7836(01)00361-7.

Walsh, W.A., Bigelow, K.A., and Ito, R.Y. 2007. Corrected catch histories and logbook accuracy for billfishes (*Istiophoridae*) in the Hawaii-based longline fishery. NOAA Tech. Memo. No. NMFS-PIFSC-13.

Appendix A

Cubic spline catchability function

The cubic spline function $Q(\cdot|\delta)$, chosen to structure the catchability at age, can be described as follows: between the knots defined as ages (0, 1, …, A), A piecewise three-degree polynomials $Q(\cdot|\delta) = Q_0(\cdot|\delta), \ldots, Q_{A-1}(\cdot|\delta))$, of total dimension dim(δ) = 4A, are fitted under the following constraints:

1. interpolating constraints: $Q(0|\delta) = 0$, and for $a = 1,\ldots,A$,

$$\text{(A1)} \qquad Q(a|\delta) = \frac{\sum_{y=1}^{Y} I^{*}_{a,y}}{\sum_{y=1}^{Y} N_{a,y}(\mathbf{C},\mathbf{S})}$$

2. twice continuity and differentiability constraints: for $a = 1,\ldots,A-1$,

$$Q_{a-1}(a \mid \delta) = Q_a(a \mid \delta)$$

$$Q^{(i)}_{a-1}(a|\delta) = Q^{(i)}_{a}(a|\delta) \quad \text{for } i \in \{1,2\}$$

Table A1. Estimated cod survivors (N_a, tonnes) in 2005 and catchabilities (Q_a).

Age a	$N_{a,Y}$	Q_a
1	9 023	7.543×10^{-5}
2	17 266	1.833×10^{-4}
3	15 344	2.712×10^{-4}
4	10 355	3.440×10^{-4}
5	7 211	3.949×10^{-4}
6	5 922	3.807×10^{-4}
7	5 479	3.983×10^{-4}
8	4 719	3.783×10^{-4}
9	3 091	4.165×10^{-4}
10	1 256	4.648×10^{-4}
11	316	4.514×10^{-4}
12	515	7.152×10^{-4}
13	325	9.703×10^{-4}
14		1.246×10^{-3}

Note: Numerical age a represents biological age a + 1.

The interpolating points (eq. A1) have been chosen to reduce the influence of yearly trends on the age structure of the catchability. The choice of a cubic spline ensures smoothness and minimizes the interpolating oscillations under the supplementary constraint $Q''(1|\delta) = Q''(A|\delta) = 0$, which is required by a classic logistic form of the selectivity at age (Harley and Myers 2001) and appears necessary to ensure the determination of polynomial parameter vector $\boldsymbol{\delta}$ (Hastie 1992).

The SA algorithm

The principle of the SA algorithm is to maximize the function $f_T(\theta) = \exp(\ell^p(\theta)/T)$, where $\ell^p(\theta)$ is the profile log-likelihood (eq. 14), and for which $T > 0$ decreases towards 0 in successive steps. Indeed, $f_T(\theta)$ becomes concentrated around the MLE $\widehat{\theta}$. Denote by θ_i and T_i the estimate of $\widehat{\theta}$ and the value of T at step i (θ_0 being the value obtained with Nelder–Mead's method). Following Pibouleau et al. (2005), a practical value for T_0 is $|\ell^p(\theta_0)|$. To obtain a fast method, a geometric rate of $T_i = 0.99^i T_0$ is often chosen for practical implementation. To explore the MLE region and avoid possible local maxima, the renewal of θ_i at step i is made following a Metropolis procedure (Robert and Casella 2004): a new value of θ_i is sampled from an instrumental distribution $\pi(\theta|\theta_{i-1})$. It is accepted or refused according to its probability of increasing $f_T(\theta)$ with respect to θ_{i-1}: the larger the probability, the closer we are from the MLE. In our case, $\theta = (\mathbf{C},\mathbf{S})$ and $\pi(\theta|\theta_i)$ is chosen as a truncated multivariate normal distribution $\mathcal{N}(\theta_i, \Sigma_i)$, discretized on $\mathbb{N}^{A-1+Y}$. Lower bounds are 0, and upper bounds have been set to five times the reported landings and the ADAPT estimates of survivors. Indeed, instrumental sampling on finite state–space ensures the convergence of the SA chains to $\widehat{\theta}$ (Hàjek 1988; Winkler 1995). The instrumental variance Σ_i has been chosen as $\lambda_i\theta_0\mathbf{I}$, where $\mathbf{I}$ is the identity matrix on $\mathcal{R}^{A-1+Y}$ and where the one-dimensional coefficient λ_i is empirically calibrated at each step to ensure an acceptance rate of 25% during the first 2000 iterations (λ_i is increased or de-

Fig. A1. Reconstitution of catch histories using 11 various censoring schemes placed on real SGSL cod landings (circles). Two schemes consider all reported landings as censored random values, with $\beta_y = 2$ and $\beta_y = 1.5$. Two schemes consider only the period 1971–1992 as censored, with $\beta_y = 2$ or $\beta_y = 1.5$. Two schemes consider the years 1985–1992 and 2003–2005 as censored, with $\beta_y = 2$ or $\beta_y = 1.5$ for the first period and $\beta_y = 1.25$ for the second period. Two schemes consider the years 1971–1992 and 2003–2005 as censored, with $\beta_y = 2$ or $\beta_y = 1.5$. Two schemes consider the period 1985–1992 and the year 2003 as censored, with $\beta_y \in \{2, 1.5, \infty\}$ for 1985–1992 and $\beta_y = 1.25$ for 2003. Finally, one scheme considers three time periods of censoring: years 1985–1992, 1993–1998, and 2003, with $\beta_y = 2$, 1.25, and 1.25, respectively.

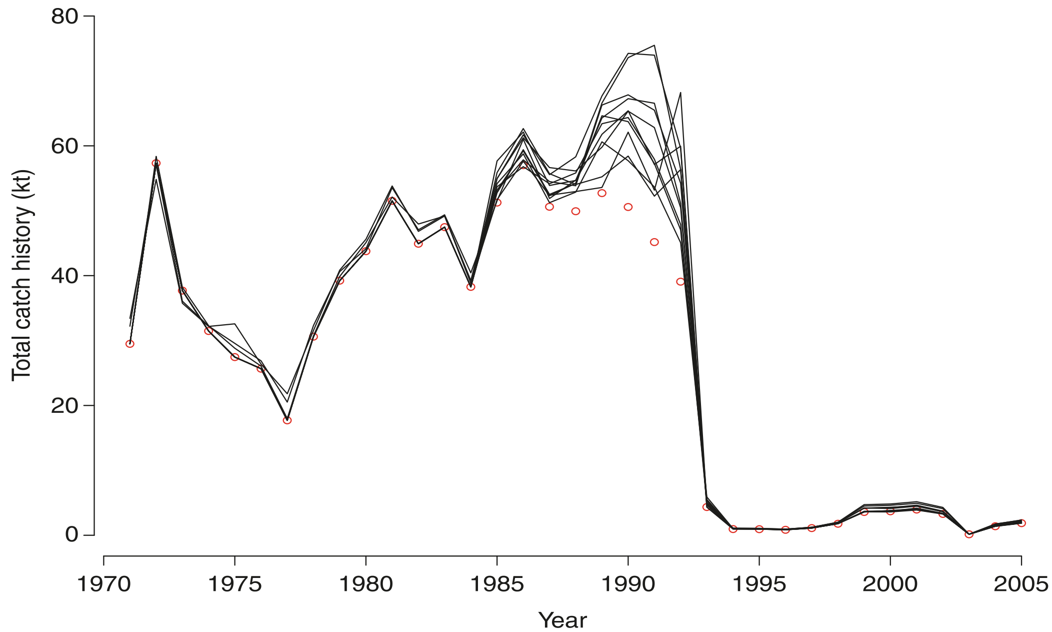
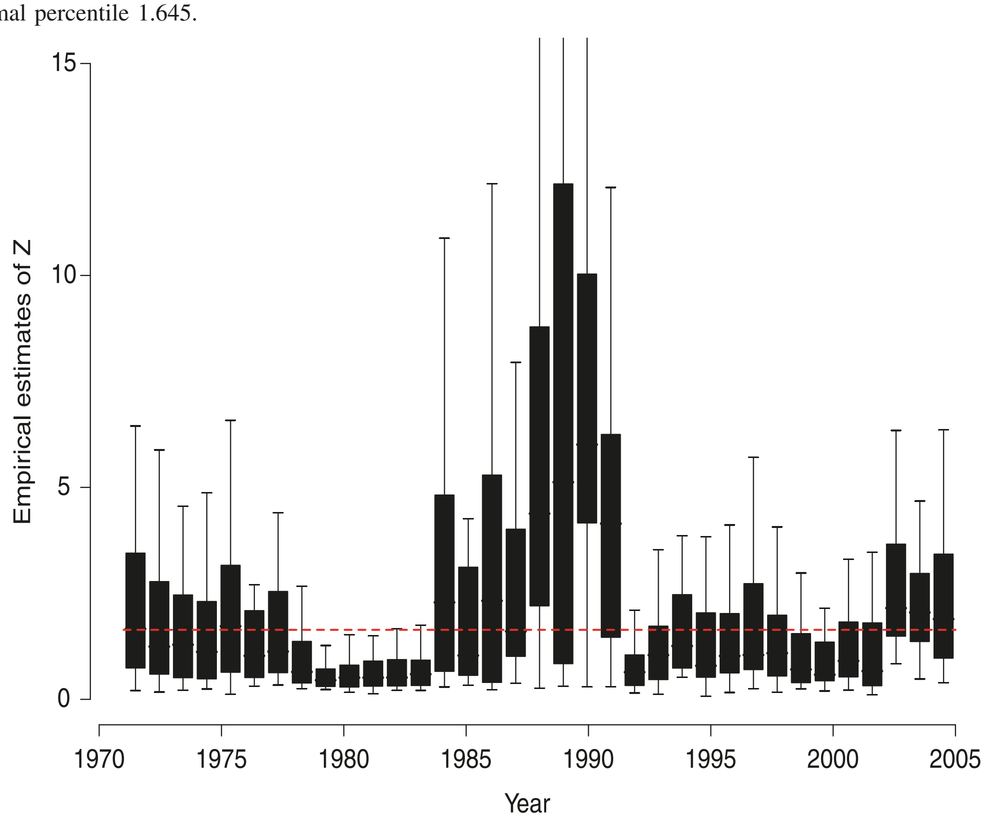


Fig. A2. Box plots of values Z_y estimated by the full-censored FA model with $\beta_y = 2$, based on 80 independent data sets with underreported catches simulated according to scheme A. The survey CV is 0.4 and the catch CV is 0.072. The horizontal broken line indicates the 95%-order standard normal percentile 1.645.



creased such that four steps are necessary, on average, to select a new value of θ). This value has been recommended by Roberts et al. (1997) for high-dimensional models. The choice of a diagonal Σ_i does not take into account the possible correlation between the parameters. Finally, the SA scheme can be summarized as follows.

Simulated annealing algorithm

Step 0: let θ_0 be a Nelder–Mead estimate of $\widehat{\theta}$ and select $T_0 = |\ell^p(\theta_0)|$.

Step i:

1. sample $\tilde{\theta}$ from $\pi(\theta|\theta_{i-1})$;
2. accept $\theta_i = \tilde{\theta}$ with probability $\min(1, \exp(\{\ell^p(\tilde{\theta}) - \ell^p(\theta_{i-1})\}/T_{i-1}))$;
3. update $T_i = 0.99^i T_0$.

Stop when θ_i remains nearly unchanged after several iterations of T_i.

The approximate SA distribution was produced using the principle of importance resampling, which became very popular in Bayesian analysis (Rubin 1988). This was carried out from exploration data generated by the addition of a random local exploration step to the classic SA algorithm. On average, we experimented with 1500 iterations before reaching stability of the estimation. By switching the SA algorithm to a random walk with a small coefficient of variation (10%), we got 500 different values close to the stable estimation. We then considered the cloud of the last 1000 points as a good set of values with which to explore the area of the MLE. Then we (i) computed the values of the corresponding likelihood and renormalized them between 0 and 1 to get probabilities and (ii) resampled the catch estimates, with replacement, according to the multinomial distribution defined by these probabilities. This step is necessary to zoom in on the MLE area.

There are, of course, identical and correlated values in the final set of estimates, but this does not differ from any posterior Bayesian approximation obtained by common MCMC or importance sampling approaches. Variances and confidence intervals were finally estimated by selecting randomly 250 vectors from among the 1000 resampled ones. This crude heuristic (following Rubin 1988) allowed the influence of correlations to be diminished. The wider the range of estimates, yielding close likelihoods, the larger is the variance of the MLE.

References

Hàjek, B. 1988. Cooling schedules for optimal annealing. Math. Oper. Res. **13**(2): 311–329. doi:10.1287/moor.13.2.311.

Harley, S.J., and Myers, R.A. 2001. Hierarchical Bayesian models of length-specific catchability of research trawl surveys. Can. J. Fish. Aquat. Sci. **58**(8): 1569–1584. doi:10.1139/cjfas-58-8-1569.

Hastie, T.J. 1992. Generalized additive models. *In* Statistical models in S. Chap. 7. *Edited by* J.M. Chambers and T.J. Hastie. Chapman & Hall/CRC, New York. pp. 249–308.

Pibouleau, L., Domenech, S., Davin, A., and Azzaro-Pantel, C. 2005. Expérimentations numériques sur les variantes et paramètres de la méthode du recuit simulé. Chem. Eng. J. **105**(3): 117–130 [in French]. doi:10.1016/j.cej.2004.06.013.

Robert, C.P., and Casella, G. 2004. Monte Carlo statistical methods. 2nd ed. Springer, New York.

Roberts, G., Gelman, A., and Gilks, W. 1997. Weak convergence and optimal scaling of random walk Metropolis algorithms. Ann. Appl. Probab. **7**(1): 110–120. doi:10.1214/aoap/1034625254.

Rubin, D. 1988. Using the SIR algorithm to simulate posterior distributions. *In* Bayesian statistics 3. *Edited by* J. Bernardo, M. DeGroot, D. Lindley, and A. Smith. Oxford University Press, New York. pp. 396–402.

Winkler, G. 1995. Image analysis, random fields and Markov chain dynamic Monte Carlo methods: a mathematical introduction. Springer-Verlag, New York.